



Model Reduction and Translation for Coordinated Expansion Planning Studies

Final Project Report

S-98G

Power Systems Engineering Research Center
*Empowering Minds to Engineer
the Future Electric Energy System*



Model Reduction and Translation for Coordinated Expansion Planning Studies

Final Project Report

Project Team

James McCalley, Professor and Project Leader

Ali Jahanbani-Ardakani

Iowa State University

Nicolas Lhuillier

Oliver Despouys

RTE-France

Graduate Students

Yanda Jiang

Zohreh Parvini

Iowa State University

PSERC Publication 24-05

December 2024

For information about this project, contact:

James McCalley
Iowa State University
Department of Electrical and Computer Engineering
Ames, IA 50011
Phone: 515-460-5244
Email: jdm@iastate.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: <http://www.pserc.org>.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

PSERC members are given permission to copy without fee all or part of this publication for internal use if appropriate attribution is given to this document as the source material. This report is available for downloading from the PSERC website.

© 2024 Iowa State University. All rights reserved.

Acknowledgments

We would like to express our deepest gratitude to PSERC for their invaluable support and funding throughout this project. Their commitment to advancing research in power systems has been instrumental in the success of our work.

We also extend our heartfelt thanks to the teams at RTE and MISO for their collaboration, insights, and the data they provided. Their practical expertise and real-world perspectives have greatly enriched the quality and relevance of this research.

The collective efforts and contributions from these organizations have made this project possible, and we are truly grateful for their partnership.

Executive Summary

This report presents the development and application of a comprehensive network reduction procedure that is specifically designed for use by Coordinated Expansion Planning (CEP) tools. CEP tools are optimizers designed to identify generation and transmission investments that minimize the long run (e.g., 20 years) investment and operational costs of running a power system. Although state-of-the-art hardware, parallelization, and advanced optimization algorithms offer ways to provide high-speed computing, the desire for increased modeling fidelity motivates deployment of these methods *together with* network reduction, and not *instead of*. The primary objective of this research was to identify network reduction methods to minimize the computational time required for CEP while maintaining high modeling fidelity. This is important due to the increasing complexity and size of power systems, which makes full-scale CEP computationally intensive and often impractical.

The motivation for this research arises from the need to manage the computational burden of CEP, which is exacerbated by the large size and complexity of modern power systems. Traditional network reduction methods, though well-established, are not directly applicable to CEP due to the specific requirements of CEP, such as the need to handle multiple operating conditions and the necessity to model equivalent branches with both flow capacity and expansion cost attributes.

The research methodology involved developing an eight-step network reduction procedure tailored to CEP. This procedure starts by identifying “key branches,” i.e., those likely to receive investment in the multi-year CEP; and then “trimming” the network by eliminating certain buses. The system is then divided into zones, followed by the selective retention and elimination of buses using a guided Ward elimination approach. For each zone, the remaining network, which is dominated by generator buses, is aggregated; capacities and expansion costs are then estimated for equivalent branches. Finally, CEP is applied to the reduced network, and the results are translated back to the full network.

The key achievement of this research lies in the development of a comprehensive network reduction procedure that addresses different aspects of model reduction and effectively balances the trade-off between computational efficiency and model fidelity. The procedure was tested on several networks, including the IEEE 118-bus system and large-scale networks of the French grid and of the MISO grid. The results demonstrated significant reductions in network size, with minimal impact on the fidelity of the CEP outcomes. Specifically, the developed procedure maintained a high degree of accuracy in representing the power flows and investment decisions of the full network while achieving substantial reductions in computational time.

There were seven specific innovations in this work: (1) Efficient identification of key branches using a high-speed rolling simulation method; (2) Application of trimming methods to significantly reduce network size; (3) Identification of zones within the network using a minimum spanning tree algorithm; (4) Development of a guided Ward elimination procedure using a genetic algorithm to strategically eliminate buses and produce equivalent branches to minimize CEP compute time; (5) Aggregating generator buses within zones while maintaining the impedances seen from one boundary bus to another; (6) Estimating capacities and assigning expansion costs

to equivalent branches using specialized optimization methods; (7) Performing translation from reduced model results to the full model using an iterative linear program.

The implications of this research are significant for the field of power systems engineering, particularly in the context of expansion planning. The developed network reduction procedure enables more efficient CEP analyses, making it feasible to apply CEP to larger and more complex networks. This contributes to the broader goal of ensuring reliable and cost-effective expansion of power systems in the face of increasing renewable integration and the corresponding need for significant transmission investments.

Future research will focus on further refining the network reduction procedure, particularly in the context of handling even larger networks and more complex operating conditions. Additionally, there is potential for developing software tools based on this procedure, which could be made available to the power systems research community and industry practitioners for more efficient CEP studies.

Project Publications:

- [1] Y. Jiang, Z. Parvini, J. McCalley, N. Lhuillier, O. Despouys, A. Figueroa-Acevedo, and J. Okullo. "Network Reduction for Power System Planning: Zone Identification." In *2023 North American Power Symposium (NAPS)*, pp. 1-6. IEEE, 2023.

Student Theses:

- [1] Yanda Jiang. "Power system reliability evaluation and long-term planning: integrating network reduction and adaptive co-optimization expansion planning," Ph. D. dissertation, Iowa State University, Nov 2024.
- [2] Zohreh Parvini. Power system coordinated expansion planning: model reduction. Ph. D. dissertation, Iowa State University, May 2026

Table of Contents

1. Introduction.....	1
1.1 Background, Motivation, and Objective	1
1.2 Network Reduction for CEP.....	2
1.3 A New Network Reduction Procedure	3
1.4 Run-time and Fidelity Assessment.....	4
1.4.1 Measuring Run-time.....	4
1.4.2 Measuring Modeling Fidelity	4
1.4.2.1 Based on OPF Flows	5
1.4.2.2 Based on CEP Objective Function Values	5
1.4.2.3 Based on Investment Comparison	6
1.5 Report Organization	6
2. Key Branch Identification and Related Preprocessing	8
2.1 Key Branch Identification	8
Rolling simulation	9
2.2 Trim and Map.....	9
2.2.1 1 st and 2 nd Degree Trimming	9
2.2.2 Binding Buses.....	11
2.2.3 Load and Generation Mapping.....	11
2.2.4 Trimming Results	12
2.3 Identify Study System	13
2.4 Reduce External System.....	14
3. Zone Identification.....	15
3.1 Weighted LP Multi-cut Method	15
3.2 MST method.....	17
4. Guided Ward elimination.....	20
4.1 Introduction	20
4.2 CEP Fidelity and Network Structure	21
4.2.1 Network Structure Similarity Measures	22
4.2.2 Similarity of Structures.....	23
4.3 CEP Complexity and Problem Size.....	24
4.4 Problem Formulation.....	26

4.5	Bus Selection: Solving Problem 1	27
4.5.1	Lowest Connection Degree (LCD).....	27
4.5.2	Approximate Minimum Degree (AMD)	27
4.5.3	Nested Dissection (ND).....	28
4.5.4	Genetic Algorithm (GA).....	28
4.5.5	Exhaustive Search (ES)	28
4.6	Results and Discussion.....	28
4.6.1	Application to Entire 118-Bus System.....	28
4.6.2	Application to Individual 118-Bus System Zones.....	30
4.6.3	Assessing Run-time/Fidelity Tradeoff	31
4.7	Conclusion.....	32
5.	Bus Aggregation	33
5.1	Topology-based Aggregation.....	33
5.2	Bus Aggregation Method – Quotient Graph (QG).....	36
5.3	Checking mechanism	42
5.4	Case Study.....	43
6.	Treatment of Equivalent Branches.....	46
6.1	Capacity Estimation.....	46
6.1.1	Basic Formulation for Capacity Estimation	46
6.1.2	Extended Formulation for Capacity Estimation.....	49
6.2	Expansion Cost Assignment.....	50
6.2.1	Formulation 1 for cost assignment	50
6.2.2	Formulation 2 for cost assignment	54
6.2.3	Dual variable method for cost assignment	55
6.3	Conclusion.....	56
7.	Expansion Planning with Translation.....	57
7.1	Expansion Planning Formulations.....	57
7.1.1	Coordinated Expansion Planning (CEP)	57
7.1.2	Adaptive Coordinated Expansion Planning (ACEP).....	58
7.2	Solution Features	59
7.2.1	Modeling Environment, Solver, and Optimization Method.....	59
7.2.2	Scaling	60
7.2.3	Hardware	60

7.2.4 Results before translation	60
7.3 Translation	61
7.4 Conclusion	63
8. Summary and Conclusions	64
8.1 Summary of Key Findings.....	64
8.2 Conclusions	64
8.3 Future Work.....	64
References.....	66

List of Figures

Figure 1-1: Influence on computational intensity for expansion planning problems	2
Figure 1-2: Network reduction procedure developed in this project	4
Figure 2-1: Illustration of 1 st -degree trimming	10
Figure 2-2: Illustration of 2 nd -degree trimming	10
Figure 2-3: Equivalent branches in 2 nd -degree trimming	11
Figure 2-4: Binding buses associated with a key branch	11
Figure 3-1: Zone identification for IEEE 118-bus system using relaxed multi-cut method.....	16
Figure 3-2: An example of minimum spanning tree	18
Figure 3-3: Zone identification for IEEE 118-bus system using MST method.....	19
Figure 4-1: Eliminating two different sets of bus ($\Delta N_{bus}=1$)	20
Figure 4-2: Eliminating different numbers of buses	21
Figure 4-3: System used for showing communicative property of bus elimination	27
Figure 4-4: ΔN_{br} for different levels of bus reduction, ΔN_{bus} , by various methods	29
Figure 4-5: Optimal bus sets for each ΔN_{bus} for zone 2 by various methods.....	30
Figure 4-6: ΔN_{br} of each bus set elimination for the 118-bus system.....	31
Figure 4-7: Run time and CEP error for each bus set for the 118-bus system	32
Figure 5-1: Generator Aggregation Effect of One Zone in IEEE 118	33
Figure 5-2: Checking mechanism based on OPM	43
Figure 5-3: Key branch flow and zonal generation comparison between original network and the Ward + topology-based reduced network	43
Figure 5-4: Key branch flow and zonal generation comparison between original network and the Ward + topology-based + QG reduced network	44
Figure 5-5: Key branch flow and zonal generation comparison between original network and the Ward + topology-based RTE reduced network	45
Figure 5-6: Key branch flow and zonal generation comparison between original network and the Ward + topology-based + QG RTE reduced network	45
Figure 6-1: Sample four-bus system	47
Figure 6-2: Reduced four-bus system	48
Figure 6-3: Operational conditions for maximizing angular difference of buses 1 and 4	48
Figure 6-4: Depicting transfer capacity in a small test system	51
Figure 6-5: Expansion cost estimation process.....	53
Figure 6-6: Five-bus test system, full model (left) and reduced model (right).....	53

Figure 7-1: Illustration of the effect of β on the investment solution	59
Figure 7-2: Step 2 of the translation process	62

List of Tables

Table 2-1: Key branch identification results comparison	9
Table 2-2: Trimming results for IEEE 118-bus network	12
Table 2-3: Trimming results for RTE 617-bus network	13
Table 2-4: Trimming results for RTE 6515-bus network	13
Table 3-1: Identified zones by relaxed multi-cut method	17
Table 4-1: Comparing the similarity of two graphs and their reduced versions.....	23
Table 4-2: Network reduction results by different methods for $\Delta N_{bus}=20$	29
Table 4-3: Comparing ΔN_{br} , for different methods applied to zones.....	31
Table 6-1: Case studies showing effects of capacity estimation on OPM after Ward reduction..	49
Table 6-2: Case studies showing effects of capacity estimation on OPM after Ward reduction and aggregation.....	50
Table 6-3: Transfer capacity in the full 5-bus and reduced network	54
Table 6-4: Expansion cost of equivalent branches (\$/MW)	54
Table 7-1: CPM results before translation for IEEE 118-bus network.....	61
Table 7-2: CPM results before translation for RTE 672-bus network.....	61
Table 7-3: CPM results after translation for IEEE 118-bus network.....	63
Table 7-4: CPM results after translation for RTE 672-bus network.....	63

1. Introduction

1.1 Background, Motivation, and Objective

Generation and transmission expansion planning software tools are useful in guiding decisions related to investment in electric power systems. When they focus on generation alone, they are called generation expansion planning (GEP) or even just capacity expansion planning tools; when they focus on transmission alone, they are called transmission expansion planning (TEP) tools, and when they do both, they are called cooptimized expansion planning or coordinated expansion planning tools (CEP). There are several expansion planning tools available today, including, for example, Plexos (Energy Exemplar), Encompass (Anchor Power Systems), Resolve (E3), ReEDS (National Renewable Energy Laboratory. NREL), and REGEN (Electric Power Research Institute, EPRI), and CGT-Plan (Iowa State University). A summary of these tools is available [1].

Expansion planning tools have been of interest for many decades, as indicated by the excellent 1985 monograph done on this subject [2], but they are receiving increased attention today as climate change, government subsidies, and the economics of wind and solar motivate a transformation from fossil-fueled electricity generation to renewable electricity generation. Because the best wind and solar locations differ from existing fossil fuel power plants, the transformation requires significant transmission investment. As a result, decisions on which type of renewables to build and where to build them involves assessing the tradeoffs between resource quality (as measured by capacity factors) with transmission investment cost. This drives the need for, and interest in, CEP.

Expansion planning tools are optimizers, identifying the investments over multiple years that optimize an objective. Most commonly, expansion planning tools minimize the net present value (NPV) of those investments plus the NPV of the operational costs incurred over that time frame, subject to the fundamental constraint that the loads are supplied such that the power grid operates reliably. Environmental (e.g., emissions) and investment constraints (maximum capacity) are often added. While the objective is usually expressed in monetary terms, expansion planning formulations can also target other objectives, such as maximizing resilience, reducing CO₂ emissions, or minimizing the use of critical materials (e.g., copper). Using the most common approach, a high-level expression of this problem, in economic terms, follows:

Minimize:

$$\sum_{t=1,N} \text{NPV} \{ \text{InvCosts}(\underline{x}(t)) + \{ \text{OpCost}(\underline{x}(t)) \} \quad (1-1)$$

Subject to

constraints on network, operations, environment, investments

Here, t is the time period, N is the number of time periods, and $\underline{x}(t)$ is the vector of generation and transmission investments made in time period t . This problem may be formulated as, in order of increasing computational intensity and solution difficulty, a linear program (LP), a mixed integer linear program (MILP), a nonlinear program (NLP), or a mixed integer nonlinear program (MINLP). But even when using the most tractable of these, the LP, the problem can still be extremely intensive, with the major influence on solve times being model granularity (spatially and temporally) and model fidelity (investment options, transmission representation). The problem

can expand even further if uncertainty is represented (requiring the modeling of multiple scenarios or futures) or if reliability and/or resilience is assessed.

Solve times for realistic-sized grid models using high-end servers can range upwards from hours to days or even weeks. The various modeling dimensions that contribute to long solve times are illustrated on the left-hand-side of Figure 1-1 [3]. In addition to problem structure and modeling granularity, there are approaches one can take to reduce solve times; as illustrated on the right-hand-side of Figure , these may be categorized into hardware solutions, optimization method, algorithm design, use of a particular solver, solver parameters, and modeling system, and/or numerical scaling. Each of these represents useful approaches that should be considered in CEP design.

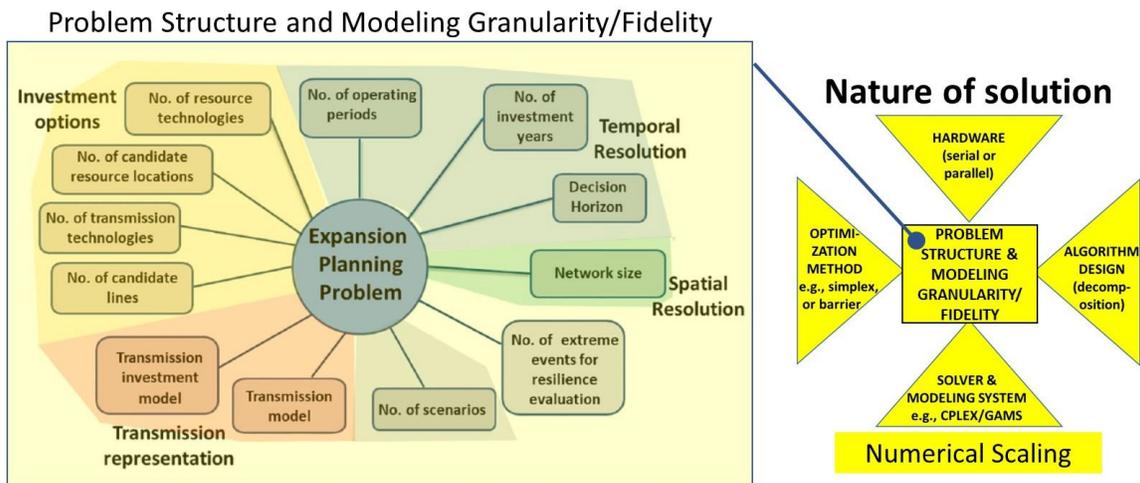


Figure 1-1: Influence on computational intensity for expansion planning problems

Another way to reduce solve times is to address the spatial resolution by reducing network size, and to do so with minimal effect on modeling fidelity. This leads to the objective of this project, which is to minimize CEP compute time but retain good CEP model fidelity; we will accomplish this via development of an overall network reduction approach for CEP.

1.2 Network Reduction for CEP

Network reduction is a relatively old topic that has been well-researched, dating back to the 1939 book by Gabriel Kron, where the so-called Kron reduction procedure was introduced [4]. Kron reduction is a method of performing Gaussian elimination on the network admittance matrix to develop a network having fewer buses, yet the same impedance is seen looking into the reduced network from remaining buses, i.e., from those not eliminated. Ten years later, J.B. Ward extended this method to what we call Ward elimination [5]. Ward elimination utilizes Kron’s Gaussian elimination method to reduce the admittance matrix, but operations are also applied to the network load vector to distribute load to the remaining buses so that (assuming all generation buses are retained), retained buses and lines see the same voltages and flows, respectively, as they did in the original system. Following Ward’s work, there were various contributions to the network reduction literature, but the most important of those were reported by Podmore and Germond in a 1977 report of the Electric Power Research Institute (EPRI) [6]. Although the focus of this project was to

produce an equivalent reduced order representation of the network and the generator dynamics for transient stability analysis, they made significant contributions towards reducing the static network as well. Of most importance, they developed a method of aggregating generator buses to reduce the network size while retaining the identity of each individual generator. This method of generator aggregation was further developed in 1992 by McCalley and colleagues [7].

A great deal of work on network reduction was performed in the ensuing years, the more significant of which includes that by Singh and Srivastava [8], Cheng and Overbye [9], Oh [10], and Shi and Tylavsky [11]. However, there has been relatively little work focusing on network reduction methods for general expansion planning analysis, exceptions include [12, 13, 14, 15, 16], and almost none for CEP analysis. This is important because network reduction for CEP poses new problems that do not arise when developing reduced networks for power flow or transient stability analysis. There are three main reasons for this.

- *Multiple operating conditions*¹⁷: The reduced network must perform well for not just a single operating condition but rather for as many operating conditions as are modeled in the CEP, typically hundreds to thousands (it is typical that CEP models a representative set of operating conditions in each year of the decision horizon to maintain a tractable optimization model). This is unlike the case of power flow or transient stability analysis where the reduced network must perform well for one or just a few operating conditions.
- *Equivalent branch modeling*: In CEP, equivalent branches must have two additional attributes beyond what is typical for power flow analysis: a flow capacity, and an expansion cost. These attributes are not typically provided by most bus elimination or bus aggregation methods.
- *Ward elimination effect on CEP branch-related decision variables*: When Ward elimination is applied to the network, run-time improvement from bus elimination may be offset by the tendency of Ward elimination to create “fills” in the admittance matrix, thus causing the number of branches to decrease relatively little, or even to increase. Because CEP run time is sensitive to the number of decision variables, and because each investible branch contributes decision variables, network reduction due to an “unguided” Ward elimination can have little effect on CEP run time.

1.3 A New Network Reduction Procedure

Motivated by the need for decreased CEP compute-time and the relatively lightly addressed problem in the literature, we have developed a new network reduction procedure for CEP application. This network reduction procedure is illustrated in Figure 1-2.

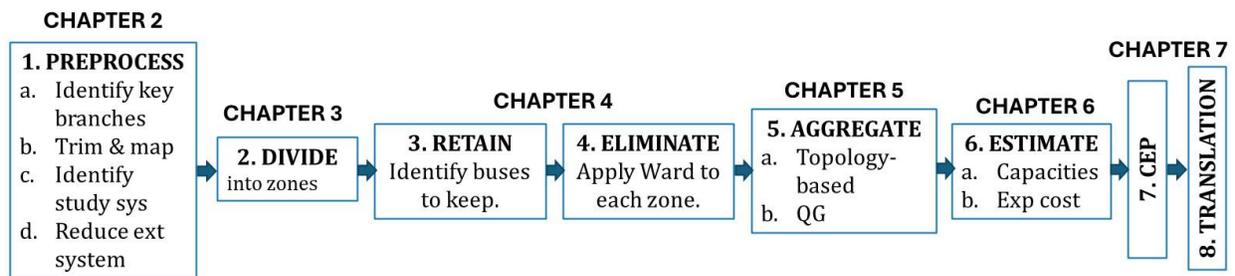


Figure 1-2: Network reduction procedure developed in this project

Below, we provide a brief description of each of the eight steps. A complete description of each of the steps, together with illustrative results, is provided in the subsequent chapters of this report.

1. Preprocess: The preprocessing step performs three functions. First, key branches, defined as branches for which investment is likely, are identified. Second, buses connected to just one or two other buses are “trimmed” and their generation and load “mapped” to adjacent buses. Third, the portion of the system to which CEP will be applied, i.e., the portion to be expanded, is identified. Finally, the rest of the system, called the “external area,” is reduced.
2. Divide: Based on key branch identification, we identify zones within the study area. Network reduction will be applied independently to each zone, with interzonal boundary buses retained.
3. Retain: In addition to the interzonal boundary buses, we identify buses to be retained to minimize CEP run time (by minimizing CEP decision variables) and maximize CEP fidelity.
4. Eliminate: We use Ward elimination to eliminate all load buses that have not been retained. All generator buses are retained.
5. Aggregate: In each zone, generator buses are aggregated to one or more new buses, and an equivalent network between aggregated buses and boundary buses is constructed.
6. Estimate: Branch capacities are estimated, and expansion costs are assigned, for each equivalent branch.
7. CEP: The CEP is run on the reduced equivalent network. Generation and branch investments are identified; some of these investments are at equivalent buses or on equivalent branches.
8. Translate: Investments made on the equivalent buses or on equivalent branches are “translated” to the full system.

1.4 Run-time and Fidelity Assessment

A key feature of our work is the ability to quantify the extent to which we achieve our objective, which is to “minimize CEP compute time but retain good CEP model fidelity.” And so we need the ability to measure these two attributes of our model.

1.4.1 Measuring Run-time

In this work, CEP is an LP; its run-time increases with the number of decision variables, and so part of our work focuses on reducing the network in such a way that the number of decision variables in the CEP is minimized (see Chapter 4). However, in actually quantifying run-time, we use measures provided by our LP solver that depend only on the platform (combination of hardware and software). To this end, for the solver CPLEX, run-time is measured in ticks. For the solver Gurobi, run-time is measured in work-units.

1.4.2 Measuring Modeling Fidelity

We use three ways to measure modeling fidelity: based on flows, based on CEP objective function value, and based on investment comparison.

1.4.2.1 Based on OPF Flows

This approach provides modeling fidelity assessment of the reduced model based on the solution to a DC optimal power flow (OPF) calculation and therefore utilizes a single operating condition to make the assessment. Although DC-OPF is simplified relative to CEP in that it addresses only a single loading condition, it is a step beyond a DC power flow calculation in that it optimizes dispatch. As such, we view DC-OPF as a good intermediate means to evaluate the fidelity of a CEP model. Satisfying criteria using this approach is a necessary condition (but not sufficient) for testing the model for CEP performance. It is useful in measuring modeling fidelity for certain intermediate steps in the reduction process.

We refer to the metric used in this approach as the OPF-based performance metric (*OPM*). The *OPM* is the root-mean-square flow deviation between reduced and full networks over all retained branches as a percentage of the average full network flow of those retained branches, given by eq. (1-2).

$$OPM = \frac{\sqrt{\frac{\sum_{b=1}^N (\Delta f_b)^2}{N}}}{\frac{\sum_{b=1}^N |f_b|}{N}} \quad (1-2)$$

Here, N is the number of branches, f_b is the flow on branch b in the full model, and Δf_b is the difference between branch b flow in full and reduced model. An *OPM* value of 0% indicates perfect model fidelity in retaining branch flows. We have established that *OPM* values below 20% to be very good and *PM* values between 20% and 40% to be acceptable. Where the flow of retained branches in the reduced network is the same as the full network and the *OPM* value is zero, we can compare the OPF objective value to evaluate the detailed impact of the reduction process on the operation modeling.

The fact that *OPM* assesses only a single operating condition and not all operating conditions within the CEP decision horizon is both a negative and a positive. The negative is that it does not reflect the model's fidelity for the various other operating conditions. The positive is that it depends on only OPF and not CEP. Because OPF can be run efficiently on large-models, it is possible to obtain the full-model result. This is in contrast to measures which capture the model's performance based on a CEP result on both full and reduced model. Such measures are not obtainable when the network is very large because it is not possible to run CEP on very large networks.

1.4.2.2 Based on CEP Objective Function Values

This approach provides modeling fidelity assessment of the reduced model based on the solution to a CEP calculation. Because the CEP calculation must be done on both reduced and full networks, it is not possible to use this measure for assessing model fidelity when the original network is very large. That is, this measure can only be used when the network is relatively small so that CEP can be performed on the full network. Although we view this approach as a sufficient condition for assessing model fidelity, it cannot be used on industry-sized networks because CEP on unreduced industry-sized networks is too computationally intensive. Thus, we use it to establish model fidelity for small-size networks; satisfactory values of this measure on small-size networks suggests that the reduction process will produce models of satisfactory fidelity on large-sized networks as well.

We refer to the metric used in this approach as the CEP-based performance metric (*CPM*). The *CPM* is the ratio of the

- CEP solution's objective function value for the full network to the
 - CEP solution's objective function value for the reduced network
- as given by eq. (1-3).

$$CPM = \frac{CEPObjFuncValue(reducedmodel)}{CEPObjFuncValue(fullmodel)} \quad (1-3)$$

A *CPM* value of 1.0 indicates perfect model fidelity. We have established that values of *CPM* between 0.97 (CEP objective function value for reduced model is 97% of CEP objective function value for the full model) and 1.03 (CEP objective function value for reduced model is 103% of the CEP objective function value for the full model) to be very good, and values outside of this range but between 0.95 and 1.05 to be acceptable.

1.4.2.3 Based on Investment Comparison

Similar to the measure based on CEP objective function values, this approach provides fidelity assessment of the reduced model based on the solution to a CEP calculation. And for the same reason (because the CEP calculation must be done on both reduced and full networks), it is not possible to use this measure for assessing model fidelity when the original network is very large. Thus, it can only be applied when assessing model fidelity for small-size networks.

Whereas both *OPM* and *CPM* provide high-level quantitative measures of model fidelity, this assessment does not; rather, it provides a method to study the *causes of infidelity*. That is, when *OPM* and/or *CPM* indicate the reduced model is performing poorly, then this approach enables identification of the investments (or lack thereof) that are causing the poor performance. This is accomplished via bar charts showing pairs of bars corresponding to a particular generation or transmission investment, where one bar is the investment for the full model, and the other bar is the investment for the reduced model. A visual scan of the various bar pairs enables quick assessment of where investments are very similar and where they are significantly different.

This approach is effective for comparing investments made in the full model and in the reduced model on retained buses and branches (since both models will have these retained buses and branches), and indication of good agreement for these investments is a strong indication that the reduced model has high fidelity. However, this method cannot compare investments made in the full model on non-retained buses or branches (since those buses and branches do not exist in the reduced model), and this method cannot compare investments made in the reduced model on aggregated buses or equivalent branches (since those buses and branches do not exist in the full model). One approach to addressing this issue is to compare full model investments with reduced model investments that have been translated (see step 8 in Figure 1-2) back to the full model.

1.5 Report Organization

The remainder of this report is organized as follows. The “steps” mentioned below refer to those portrayed in Figure 1-2. Chapter 2 describes the preprocessing step 1, which includes key branch

identification. Chapter 3 identifies the methods used to identify network zones in step 2. Chapter 4 describes the “guided Ward” method of bus elimination deployed in steps 3 and 4. Chapter 5 identifies two methods of bus aggregation performed in step 5. Chapter 6 describes the step 6 methods used to perform capacity estimation and expansion cost assignment for equivalent branches. Chapter 7 provides the step 7 CEP formulation and the ACEP formulations used in this project. Conclusions are drawn in Chapter 8.

2. Key Branch Identification and Related Preprocessing

This chapter describes four main preprocessing tasks. The first task is to identify the key branches. The second step is to trim the network, eliminating buses only connected to one or two other buses and then mapping their generation and load to the remaining buses. The third step is to identify the section of the system designated for CEP application and expansion. The final step is that the remainder of the system, referred to as the "external area," is reduced to boundary buses. Each of these steps is explained in this chapter.

2.1 Key Branch Identification

A central step in the entire reduction procedure is to identify *key branches*. Key branches are those that are most likely to receive investment during the planning horizon. If the branches receiving investment in the original full network are preserved and the power transfer patterns between the reduced and full networks are similar, then the expansion of the reduced network should reflect the corresponding situation in the full network. However, running expansion planning for the full system is computationally intractable. To address this, we explore alternative methods for forecasting which branches will be key.

One approach is to use the congested branches from the first year, based on the assumption that these branches will be the ones receiving investment over the next twenty years. However, as generation grows and load increases, some branches that were initially congested may no longer be, while others that were initially not congested might become congested in the future. This method may result in a poor key branch forecast.

Another approach is to conduct expansion planning for the entire system on a rolling basis such that a single simulation is performed every Δ years up to the last year of the decision horizon T , according to the below algorithm:

1. $k=1$
2. Run CEP for year k to year $k+\Delta-1$
3. Update model for identified investments from step 2 CEP.
4. If $k=T-\Delta$ stop; else $k=k+\Delta$ and go to Step 2.

We refer to this as a rolling simulation. In using it for key branch selection, we make an initial selection of key branches as those with the highest expansion cost. Unselected branches, being less costly to expand (and relatively speaking, free to expand), appear as infinite capacity branches, as their investment cost is negligible. After completing one round of rolling simulation over 20 years, if certain branches are never selected for investment, we can replace them with cheaper alternatives. This process can be repeated over several iterations until all branches have been tested.

To evaluate the effectiveness of different methods for accurately identifying key branches, we conducted a case study. Using the IEEE 118-bus system, we assign a branch capacity of 200 MW and simulate a 7.5% annual load increase over twenty years. We then perform a CEP analysis to determine optimal investments in generation and transmission needed to meet future demand. The

investment costs for various generator types and transmission lines are derived from data provided by MISO.

According to the expansion planning results, 12 branches received investment throughout the entire planning horizon, with a total expansion amounting to 1,522.2 MW. The effectiveness of the two methods mentioned earlier has been evaluated, and the comparative results are presented in Table 2-1. We observe in Table 2-1 that for key branch identification, rolling simulation outperforms use of first year congested branches.

Table 2-1: Key branch identification results comparison

Key branch identification methods	Accurately captured	Not captured
Rolling simulation	11 branches (1511.4 MW)	1 branch (10.78 MW)
Using first year congested branches	1 branch (36.3879 MW)	11 branch (1485.8 MW)

The key branches are integral to the reduction process in that they define the potential for transmission investment. As a final comment in this section, there are two ways in the reduction process where error in this step (i.e., missing a key branch), can be compensated. The first is in the capacity estimation and cost assignment of equivalent branches, as described in Chapter 6. The second is in the last step, “translation,” where expansion results on the reduced model are “translated” to the full model.

2.2 Trim and Map

The purpose of this step is to reduce the network size without actually making significant changes to the structure of the network. This is achieved by iteratively eliminating buses connected to one or two other buses, which targets buses radially connected and buses connected in series in long lines. We call this process “*trimming*” and do it in two levels: 1st-degree trimming and 2nd-degree trimming. The generation buses can also be eliminated in this process; however, the generation on the eliminated buses is moved to the closest remaining bus. We refer to this step as mapping.

2.2.1 1st and 2nd Degree Trimming

The term *1st-degree* refers to buses connected to only one bus, i.e., buses with degree one. The 1st-degree process is iterative, therefore, when a single-degree bus is eliminated and the bus connected to it becomes of degree one, it will also be eliminated in the iterative process. The process is continued until no bus of degree one is left. This way, all the radially connected buses will be removed. Figure 2-1Figure shows an example of applying iterative 1st-degree trimming on a sample graph. In the first iteration, buses 1, 9, and 13 are eliminated as the buses with degree one. Then, bus 8 becomes degree one, so it is eliminated in the second iteration. The 1st-degree trimming process is complete when no degree one bus is left.

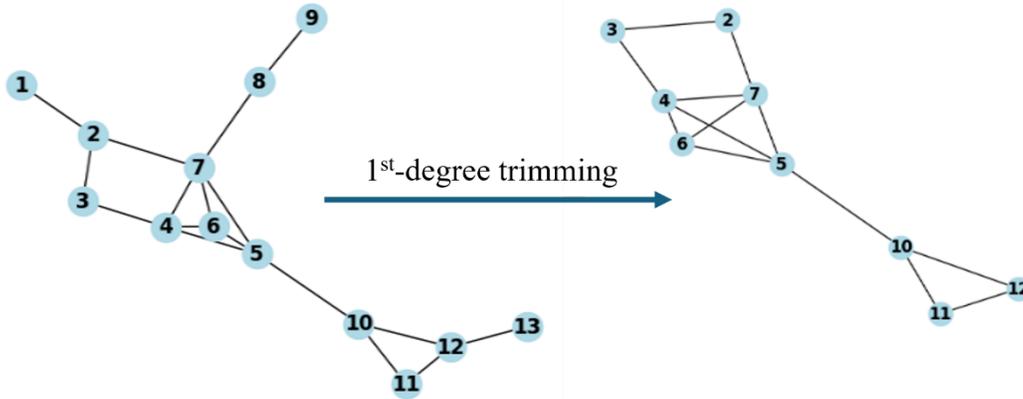


Figure 2-1: Illustration of 1st-degree trimming

In the 1st-degree trimming, no new branches are created in the network, and only radial branches are removed. The next level is 2nd-degree trimming, where buses with a degree of two are eliminated iteratively. This means that all 2nd-degree buses are eliminated first. Then, any buses of degree two or lower are also eliminated. This process continues until no bus of degree two or lower is left. Figure 2-2 shows the process in the example network. In the first iteration, buses 2, 3, 11, and 12 are eliminated. The degree of bus 10, then, changes from three to one and thus is eliminated in the next iteration. All the other remaining buses have a degree of higher than two, so the process stops.

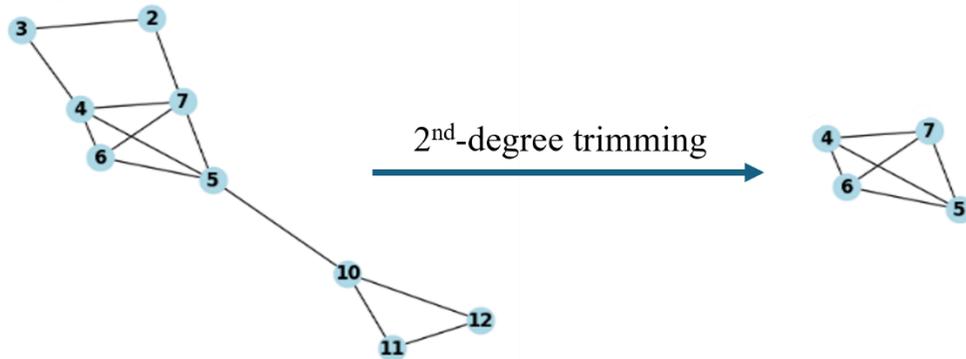


Figure 2-2: Illustration of 2nd-degree trimming

In the 2nd-degree trimming, the network structure does not change much either because the main connections are maintained, and the equivalent branches would be straightforward connections which means it is not complicated to define their capacity and cost. For example, if a 2nd-degree bus is eliminated, the neighboring buses are connected by a new equivalent branch, as illustrated in Figure 2-3.

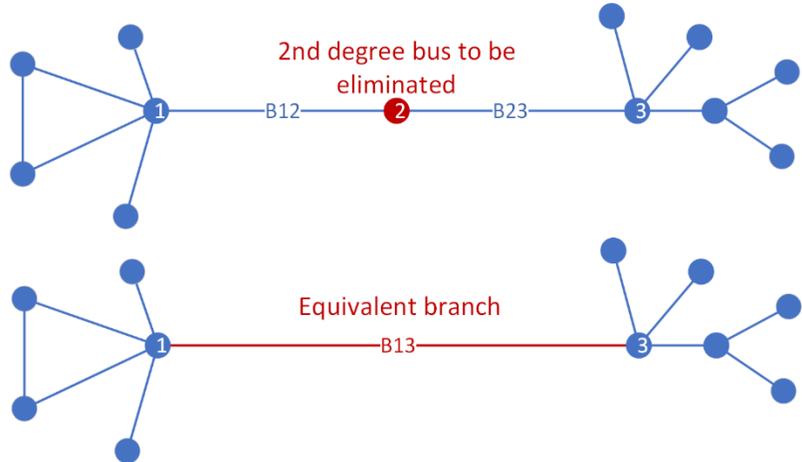


Figure 2-3: Equivalent branches in 2nd-degree trimming

2.2.2 Binding Buses

To ensure that the key branches in the reduced model maintain their original attributes, we not only retain their terminating buses but also, for one terminating bus, all buses connected to it. We call these additional retained buses binding buses and exclude them from the trimming process even if they are degree one or two. They are necessary to prevent reduced paths parallel to the key branch from being “folded in” to the (retained) key branch, increasing its admittance by the admittance of the reduced parallel path.

Figure 2-4 illustrates the concept of binding buses connected to a key branch. We choose the terminating bus with the least number of connections to serve as the one for which we identify the binding buses, as doing so reduces the number of retained buses and allows for higher reduction levels. Binding buses are excluded from the trimming process regardless of their degree. The list of binding buses contains the buses connected to each key branch plus all the neighboring buses connected to one side of it.

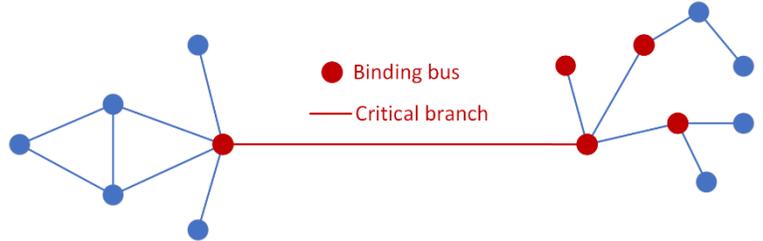


Figure 2-4: Binding buses associated with a key branch

2.2.3 Load and Generation Mapping

In the trimming process, we allow generation buses to be eliminated if they are not included in the binding buses. Therefore, we need to relocate the generators connected to them to the nearest

retained buses. We call this process mapping, where each generator is mapped to a new bus that is retained.

For the 1st-degree trimming, the mapping is straightforward because the eliminated bus is only connected to one bus, so there is only one bus to move the generator to. Even in a radial subnetwork, the last eliminated bus is only connected to one bus. So, identifying the new bus to map the generators to is simple in 1st-degree trimming.

For the 2nd-degree trimming, since each bus is connected to two other buses, there are two options to choose from. In this case, we choose the bus that already has a generator connected to it, if available, or the one that is connected to other eliminated generator buses. This helps to reduce the number of generator buses in the network, which will be important in the next reduction step, where we want to retain the generator buses. The lower the number of generator buses, the higher reduction levels we can achieve. Loads on eliminated buses are distributed based on the load factor matrix in the Ward reduction approach.

2.2.4 Trimming Results

To test the effectiveness of the trimming step in reducing network size, we apply it to different networks and assess the reduction level and accuracy. The measures that we consider are the number of eliminated buses and branches, the number of generator buses, the number of equivalent branches, the flow of key branches, and the OPF results. In these cases, the congested lines in the full network and all the buses connected to one end are considered binding buses and are retained to maintain the OPF similar to the full network.

For the IEEE 118-bus test network, which has 54 generator buses and 179 branches, the trimming results are provided in Table 2-2. In this case, there are 10 key branches leading to 22 binding buses that are retained. The number of buses is reduced to 54 after the trimming, and the branches are reduced to almost half. Moreover, the number of generator buses is reduced from 54 in the full network to 37 after 2nd deg trimming, while the OPM results are identical in all cases because the key branches that are retained have the same flow as in the full network (so it is not included in the table), and DC OPF results are very close. The minor differences in OPF results are due to the changes in power flows in the rest of the network caused by load and generation mapping. This reduction in the number of generator buses is of interest because, in the next step, we need to retain generator buses. Thus, the lower number of generator buses means that we have more flexibility in reduction.

Table 2-2: Trimming results for IEEE 118-bus network

IEEE 118	Full Network	1 st -degree	2 nd -degree
N_{Bus}	118	112	54
N_{GBus}	54	51	37
N_{Br}	179	173	95
OPF (\$/hr)	128946	128946	128913

This step is also applied to two RTE test networks of different sizes, RTE 617 and RTE 6515 [18,19]. Trimming the RTE617 network with 922 branches reduces the network size to 193 buses and 325 branches while maintaining the flows of the key branches close to their original flow in the full network. The difference between OPF objective functions in the full and reduced network is less than 1%. The results for the RTE 617-bus network are provided in Table 2-3.

Table 2-3: Trimming results for RTE 617-bus network

RTE 617	Full Network	1 st -degree	2 nd -degree
N_{Bus}	617	388	193
N_{GBus}	163	92	76
N_{Br}	922	566	325
OPF (\$/hr)	347575	347756	349651

Trimming RTE 6515 with 9037 branches leads to a network with 1210 buses and 2290 branches. This reduces the network size by 80%, while the difference between OPF objective functions in the full and reduced network is around 8%. After mapping the trimmed generator buses to the retained buses in the reduced network, the number of generator buses is reduced from 1389 to 448. Table 2-4 presents the results for this case.

Table 2-4: Trimming results for RTE 6515-bus network

RTE 6515	Full Network	1 st -degree	2 nd -degree
N_{Bus}	6515	3780	1210
N_{GBus}	1389	739	448
N_{Br}	8104	5369	2290
OPF (\$/hr)	2788806	2553350	2553707

Since we use DCOPF in the CEP, we expect that similarity in OPF results would also apply if we used CEP.

2.3 Identify Study System

We identify the portion of the network to be studied and potentially expanded as the internal system, and the remaining part of the network is referred to as the external system. The entire network thus comprises the internal system, the external system, and one or more boundary buses

that connect these two parts of the overall system. We apply reduction methods to the internal system, while the external network is represented only by the boundary buses.

2.4 Reduce External System

By simplifying the external network representation, we can more effectively manage the computational burden of CEP and concentrate our efforts on the internal system. This targeted approach allows for more precise modeling and analysis of the power system under study.

The external system is reduced to only the boundary buses that connect to the internal system. By applying reduction techniques such as Ward elimination and retaining only these boundary buses, all the generation and load are aggregated into these representative buses. This significantly reduces the network size, allowing for a more manageable and focused analysis of the internal system.

3. Zone Identification

The main objective of this chapter is to partition the entire system into several zones. Our goal is to ensure that the power transfer patterns between these zones in the reduced model closely resemble those in the original full model. This way, any power expansion implemented in the reduced model will accurately reflect the results in the full model.

We consider zones to be subnetworks of the overall network, partitioned by selected key branches each of which have terminals that are members of the set of interzonal boundary buses. Each zone's buses can be replaced by a subnetwork of one or just a few buses, connected to the boundary buses, so that the power flow pattern of the entire network, as indicated by flows on key branches, remains very nearly the same. We desire to identify zones to guide our reduction efforts so that we may achieve a desired level of network reduction while meeting the targeted performance.

Zone identification is not a new problem for network reduction. Indeed, a relatively recent effort [16] proposed a weighted linear program (LP) multi-cut formulation to partition the system. In this section, we propose a new zone identification method based on minimum spanning tree (MST). The weighted LP multi-cut method and the MST method are described in the following subsections, respectively.

The core idea of zonal division is to partition the system network into several zones based on key branches, with each key branch having its terminals in different zones. The subsequent reduction process occurs within each zone, while key branches are preserved to maintain the network's backbone.

We described identification of key branches in Section 2.1. Once the key branches are identified, we use them to divide the system into zones. This section describes two different algorithms that can be used for this zonal division.

It is important to note that while identifying more key branches typically leads to a greater number of zones and a larger reduced network, this process can also, in some instances, divide larger zones into smaller ones, thereby simplifying the aggregation process (the aggregation process is described in Chapter 5). This aspect of the analysis is nuanced and can be quite complex.

3.1 Weighted LP Multi-cut Method

The weighted LP multi-cut method uses the notion of a key branch (i,j) , which as indicated in Sections 1.3 and 2.1, is an existing branch, projected to be congested. A key bus is a bus terminating a key branch. The method selects key branches in advance and then partitions the two terminals of these key branches into different zones [16]. It accomplishes this by (1) minimizing the sum of the weights of the edges cut by the zonal boundaries subject to: (2) terminals s_k and t_k of each of the K key branches are in different zones; (3) if bus u belongs to the same zone as bus v , and bus v belongs to the same zone as bus w , then buses u and w belong to the same zone; (4) all branches have non-negative distance. This problem is described analytically as follows:

$$\min \sum_{(i,j) \in E} d(i,j) * W_{i,j} \quad (3.1)$$

subject to:

$$d(s_k, t_k) \geq 1, k \in \{1, \dots, K\}, \quad (3.2)$$

$$d(u, v) + d(v, w) \geq d(u, w), u, v, w \in \{1, \dots, N\}, \quad (3.3)$$

$$d(u, v) \geq 0, \quad (3.4)$$

where:

- $d(i,j)$ are formally binary decision variables, indicating for any pair of buses (i,j) whether the two buses are within a single zone $d(i,j)=0$ or in different zones $d(i,j)=1$; for key branches, $d=1$. However, similar to what was done in [16], we solve a relaxed LP version of this problem where bus pair (i,j) is within one zone if $d(i, j) \leq 0.5$ and between two zones if $d(i, j) > 0.5$.
- W_{ij} , real valued weights on bus pair (i,j) , where $W_{ij} = 0$ if bus pair (i,j) represents a key branch; $W_{ij} = 1$ if either bus i or j are key buses (but not both); $W_{ij} = 2$ if neither i or j are key buses;
- N is the number of buses;
- K is the number of key branches;
- E is the set of lines in the initial network.

We illustrate this method on the IEEE 118 bus system Pena et al. (2017) [20] by selecting all non-radial branches having flow exceeding 200 MW as key branches. Using the above-described relaxed multi-cut method, we identify zones as shown in Figure 3-1, where key branches include those highlighted in red, which are branches 5-8, 17-30, 26-30, 37-38, 49-66, and 89-92. Figure 3-1 shows the identification of four distinct zones, summarized in Table 3-1.

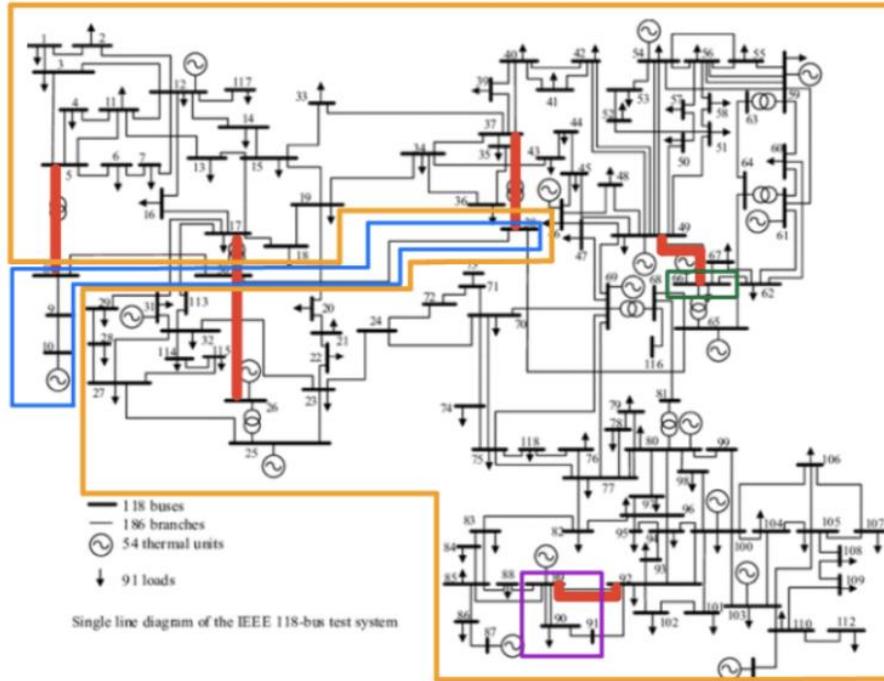


Figure 3-1: Zone identification for IEEE 118-bus system using relaxed multi-cut method

Table 3-1: Identified zones by relaxed multi-cut method

Zone #	Zone size
1	3 buses (89, 90, 91)
2	1 bus (66)
3	5 buses (8, 9, 10, 30, 38)
4	109 buses (all remaining ones)

A typical industry model has over 10,000 buses, e.g., models of the US Eastern Interconnection often have 90,000 buses or more. Assuming a 10,000-bus model, then the number of variables d , (there is a value of d for each pair of buses), is given by $N*(N-1)/2=10,000*9999/2 \approx 5 \times 10^7$. Consequently, the optimization problem becomes intractable for large systems. Our experience in testing the multi-cut method indicates it often results in what we call heterogeneous (or nonuniform) partitions, which are partitions of several “small” zones, i.e., zones containing just one or just a few buses, and a single “large” zone containing all remaining buses. Such solutions minimize the objective function (1) because they tend to minimize the number of additional boundary buses between zones (i.e., boundary buses not terminating key branches (s_k, t_k)) and thus the total number of branches (i,j) in the partition for which buses i and j are in different zones (for such branches, $d(i,j)$ are decision variables in (1) and take on the value of 1 in the optimal solution).

These solutions, as heterogeneous partitions, are problematic to reduce, for two reasons. First, for “small zones,” the extent to which the number of buses and/or circuits may be reduced is extremely limited; indeed, if the zone contains just one bus, no reduction is possible. Second, although “large” zones offer significant reduction potential, the zone size may constrain that potential since excessive reduction in a particular area can eliminate important structural features of that area and consequently degrade the fidelity of the reduced model. Overall, having a single zone that contains most of what comprises the original system means that the partitioning effect, in the end, offers little benefit. To address this, to identify more uniform partitions, we have developed a new MST-based method, described in Section 3.2.

3.2 MST method

In graph theory, an MST, also known as a minimum (weighted) spanning tree, is a subset of edges in a graph that connects all nodes while minimizing the total sum of edge weights Sedgewick and Wayne (2011) [21]. In other words, it is the minimum weight set of edges that links all nodes. Figure 3-2, adapted from Wikipedia contributors (2023) [22], illustrates the MST as the dark thick lines for a simple network.

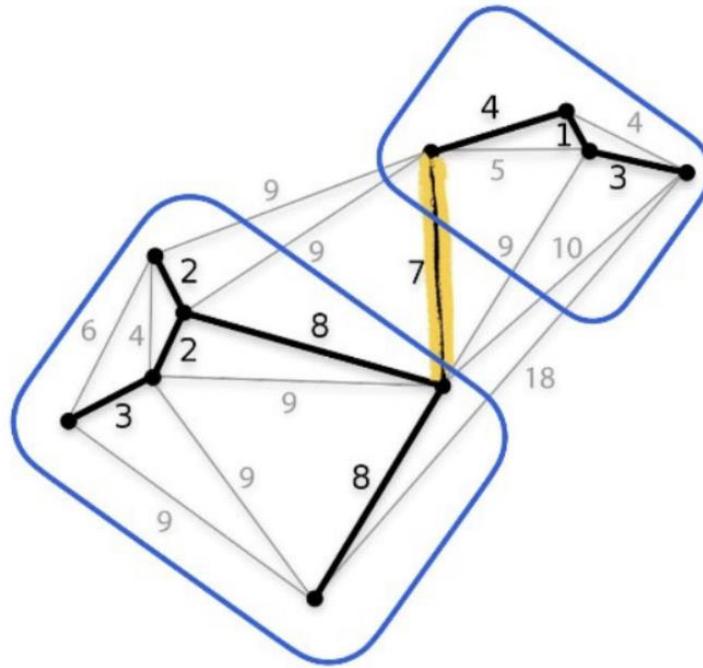


Figure 3-2: An example of minimum spanning tree

We have developed an MST-zone identification method to identify zones in a network [23]. This method is based on identification of key branches. A key branch, similar to the key branch employed in the multi-cut method, is used to partition the network into two zones. Key branches must have low weights so that they are selected into the MST and preserved in the reduced model. Figure 3-2 highlights in yellow one key branch and the resulting partitioned zones. The key branches are likely to have a significant influence on the ultimate expansion planning result. Given that our approach, as illustrated in Chapter 1, performs both generation and transmission expansion on the reduced model but concludes with a second transmission expansion step (via “translation”) on the generation-expanded full model, we observe that capturing the network’s power flow pattern is important to retaining congested branches since the power flow pattern significantly influences the amount and location of invested generation. Thus, we select key branches as those having flow exceeding a specified threshold in the network.

Although key branches are selected before branch weighting is performed, branch weighting should ensure that key branches are selected for the MST. If we were to follow the multi-cut thinking, we would use the reciprocal of percent loading where the normalization is given with respect to the flow limit of each branch. This approach yields a proxy for congestion weighting, where lines operating closer to their capacity are assigned lower weights and are more likely to be selected into the MST. However, this approach can select low-capacity branches that, although congested, do not have much influence on the overall network power flow pattern. To solve this problem, we select large-flow branches and medium-flow congested branches. If there are more than one load condition (scenario), we can use the average of branch flows over all conditions as representative flows for branches. This approach is consistent with our key branch selection criteria described in the previous paragraph which prioritizes preservation of the overall network

power flow pattern in the reduced model and incorporates advantages of the multi-cut paper’s idea. This prioritization is important because, although we perform both generation and transmission expansion on the reduced model, we retain from this step only the generation expansion (the final transmission expansion is performed following translation as illustrated in Chapter 1).

Once the MST is identified, we remove key branches on the MST one by one, where each key branch removal results in partitioning the MST into two subnetworks containing one or more buses, one of which is a terminal of the removed key branch. We identify each subnetwork as a zone. With N key branches, the MST results in $N+1$ zones which can however be recombined to a lesser number (if there is at least one branch and no key branch between 2 zones).

The MST method is applied to divide the system into zones, considering the same key branches used in the multi-cut method described in the last section. The calculation was completed in less than one second using MATLAB, demonstrating the speed of the MST method, making it scalable to large industry-sized systems. In addition, there are no single-bus zones in the result unless all branches connecting to one bus in MST are selected as key branches, and most zones contain over 10 buses. Although there is one larger zone, it is much smaller than the largest zone identified in the multi-cut method; in addition, this can be easily addressed by identifying one or more additional key branches within this zone. These results indicate that the MST method shows promise in effectively identifying power system zones when compared to the multi-cut method.

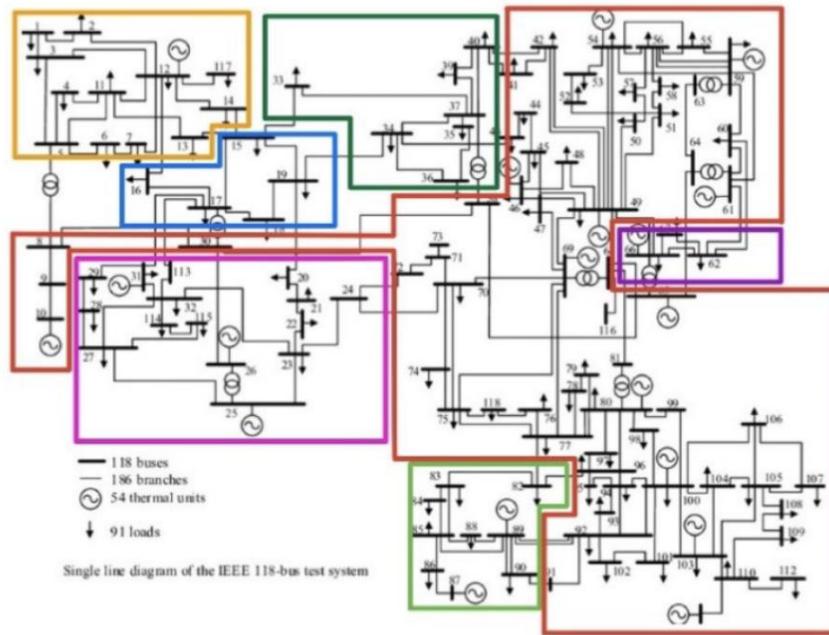


Figure 3-3: Zone identification for IEEE 118-bus system using MST method

4. Guided Ward elimination

4.1 Introduction

This chapter focuses on step 4 of Figure 1-2, the bus selection and elimination via the Ward method [5], within the internal system, while neighboring networks are considered the external system. This focus is motivated by the observation that CEP run-time is influenced by both the number of buses and branches. While network reduction using Ward's approach is effective in reducing the number of buses, it often does so at the expense of a relatively small decrease, and possibly even increase, in the number of branches; this affects both run-time and accuracy of the reduced equivalent when used for CEP. This chapter addresses this issue; specifically, we address the following problem:

In the system under study, generator buses are retained for CEP purposes, and non-generator buses are candidates for elimination. We desire to identify a set of buses to eliminate, S_e , that achieves a good tradeoff between reducing CEP run-time and maintaining modeling fidelity.

For a specified bus-reduction target ΔN_{bus} , CEP run time reduces with more reduction in the number of branches, ΔN_{br} . Moreover, having fewer equivalent branches in the network improves the CEP fidelity. Therefore, finding the best buses to eliminate that produce less dense (more sparse) admittance matrices is essential, resulting in fewer branches in the reduced network.

To illustrate, Figure 4-1 is an example that compares two different selections of eliminated buses, showing the influence of eliminating certain key buses.

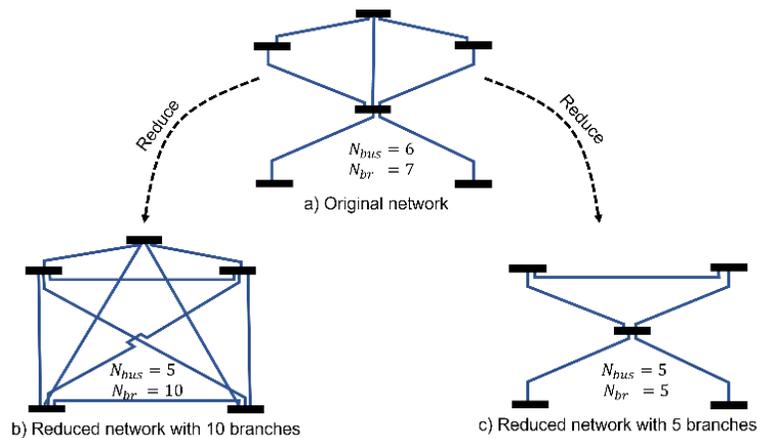


Figure 4-1: Eliminating two different sets of bus ($\Delta N_{bus}=1$)

Figure 4-2 demonstrates how retaining an extra bus results in a less complicated network with fewer buses and branches. These examples highlight the significance of carefully selecting the number of buses to eliminate.

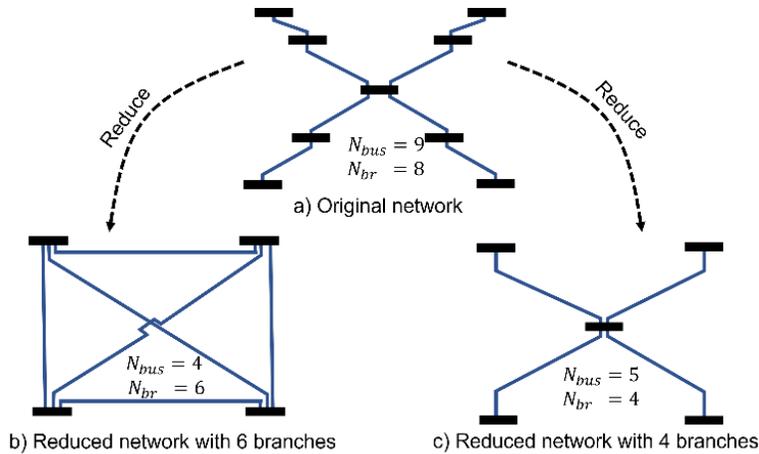


Figure 4-2: Eliminating different numbers of buses

The above examples motivate an additional observation. In both cases, the network of (c) is topologically more similar to the original network than that of (b). We hypothesize that topological similarity lends itself to better modeling fidelity of the CEP problem because topological similarity facilitates the calculation of capacities and expansion cost assignments of the equivalent branches and improves the accuracy of the translation step. Chapter 6 addresses the calculation of capacities and expansion cost assignments of the equivalent branches and the translation step is presented in Section 7.3.

Bus eliminations leading to sparser admittance matrices may serve as promising candidates for achieving the run-time/modeling fidelity tradeoff we desire. Therefore, the questions we are trying to answer are: *Which buses should we eliminate? And how many buses to eliminate?*

In the following, Section 4.3 provides CEP complexity analysis, and we establish a rationale for minimizing the number of variables. Section 4.4 provides a formulation for identifying the bus set S_e that reduces the number of branches for a given target bus reduction ΔN_{bus} , and Section 4.5 introduces methods of solving them. Section 4.6 provides results and discussion, and Section 4.7 concludes.

4.2 CEP Fidelity and Network Structure

The fidelity of a CEP applied to a reduced network is highly influenced by how well key features of the full network are preserved. Key factors include the number of retained buses and branches, which are essential for maintaining the network's integrity. Additionally, retaining key branches and buses, especially those associated with congested or constraining lines, is crucial. Binding constraints and feasible regions also play a significant role in CEP accuracy, making it vital to preserve the same solution space in the reduced network. This can be addressed during the key branch identification step, where congested branches are identified, and their corresponding buses are retained from elimination. Maintaining the overall structure of the network is desirable, to ensure that the reduced model accurately reflects the original system. Minimizing the introduction of equivalent lines helps to avoid unnecessary complexities, while accurately estimating the capacity and cost of these equivalent branches is vital for maintaining the precision of the reduced

network model. Together, these factors contribute to the high fidelity of the CEP, ensuring that the reduced network effectively replicates the behavior of the full network.

Among the aforementioned factors, this section focuses on the optimal selection of buses to reduce, which is crucial for preserving the network structure and, consequently, results in higher CEP fidelity on the reduced network.

4.2.1 Network Structure Similarity Measures

The goal is to understand how a reduced network compares to the original network. The reduction process may involve the removal of nodes, edges, load and generation displacement and mapping or other simplifications of the network structure. To ensure that the essential properties of the original network are preserved, we evaluate the similarity between the original and reduced networks.

Various measures can be employed to compare the topological structure similarity between graphs, which can be categorized into direct structural, centrality, and spectral measures [24].

- **Direct Structural Measures:** These include the clustering coefficient, average path length, and diameter. They offer direct insights into the structural properties of the network.
- **Centrality Measures:** This category includes betweenness centrality and degree distribution. These measures facilitate understanding the roles of specific nodes and the overall distribution of connections within the network.
- **Spectral Measures:** Examples include the spectral radius, algebraic connectivity, and the full spectrum of the Laplacian matrix. Spectral measures provide a global perspective on the network's structure, capturing complex properties such as robustness, connectivity, and potential community structures

While all the metrics provide valuable insights into different aspects of network similarity, spectral measures and clustering coefficient stand out as the most comprehensive indicators of whether the reduced network retains the essential characteristics of the original network. Spectral measures capture the global structure and robustness, while the clustering coefficient and average path length ensure that local and overall connectivity patterns are preserved.

- **Clustering Coefficient:** Clustering coefficient is a measure of the degree to which nodes in a network tend to cluster together. It is defined for a single node as the ratio of the number of actual connections between its neighbors to the number of possible connections between them. For the entire network, the average clustering coefficient is often used, representing the tendency of the whole network to form tightly knit groups. A high clustering coefficient indicates a network with a high degree of local interconnectedness, where many of the nodes' neighbors are also connected to each other.
- **Average Degree:** The average degree is a measure of the typical number of connections each node in a network has. It is calculated as the total number of edges in the network divided by the total number of nodes, giving the mean degree across all nodes. In network

analysis, the average degree provides insight into the overall connectivity of the network. A higher average degree suggests a more densely connected network, where nodes generally have more connections, while a lower average degree indicates a sparser network with fewer connections per node.

- **Spectral Radius:** The spectral radius is the largest absolute value of the eigenvalues of a matrix associated with the network, often the adjacency matrix. In the context of network analysis, it provides information about the network's connectivity properties. A higher spectral radius indicates that the network may have nodes or regions with a high degree of connectivity that are key points in the system.
- **Algebraic Connectivity:** Algebraic connectivity, also known as the second-smallest eigenvalue of the Laplacian matrix, is a measure of how well-connected a network is. A higher algebraic connectivity implies that the network is more robust to node or edge removals, indicating strong overall connectivity and resilience. It also reflects the network's ability to facilitate synchronization and flow.
- **Full Spectrum of the Laplacian Matrix:** The full spectrum of the Laplacian matrix consists of all its eigenvalues, which encode various structural properties of the network. The smallest eigenvalue is always zero and corresponds to the network's connected components, while the other eigenvalues provide insights into the network's connectivity, robustness, and other dynamic behaviors. Analyzing the full spectrum allows for a comprehensive understanding of the network's global structure.

It is useful to assess multiple structural and spectral measures for a holistic assessment of the two subnetworks' similarity. We focus on “subnetworks” rather than the entire network to facilitate network reduction that is performed one zone (or subnetwork) at a time. Each measure captures different aspects of the network's topology. This targeted approach ensures that the similarity measures accurately represent the degree of similarity between the original and reduced networks, focusing on the critical areas of their topology and function.

4.2.2 Similarity of Structures

Using the discussed measures, we can evaluate the topological similarity between the original and reduced networks. For the examples in Figure 4-1 and Figure 4-2, these measures are calculated and presented in Table 4-1: Comparing the similarity of two graphs and their reduced versions Table 4-1.

Table 4-1: Comparing the similarity of two graphs and their reduced versions

Graph	Ave Clustering Coefficient	Average Degree	Spectral Radius	Algebraic Connectivity
G _{1a}	0.48	2.33	2.81	1
G _{1b}	1	4	3	4
G _{1c}	0.43	2	2.34	1
G _{2a}	0	2	2.36	0.38
G _{2b}	1	3	3	4
G _{2c}	0	1.6	2	1

G_1 refers to the network (graph) in Figure 4-1, and G_2 refers to the network (graph) in Figure 4-2, while the letters a, b, and c refer to the three networks in each of Figure 4-1 and Figure 4-2. As different measures show, the reduced graph G_{1c} for Figure 4-1 and G_{2c} for Figure 4-2 is more similar to the original ones. This shows that eliminating the buses that lead to more sparse admittance matrices also helps preserve the topological structure of graphs, as illustrated in Figure 4-1 and Figure 4-2.

4.3 CEP Complexity and Problem Size

Several factors contribute to the compute time of the CEP problem, including network size and structure, problem structure, binding constraints, optimization model and solver, and the range of input data. The network size is determined by the number of buses and branches, as well as generators and loads, and the network structure is how different buses are interconnected by branches. The CEP structure includes the model itself, the number of decision variables and constraints, and the continuous or integer nature of these variables. Binding constraints, such as branch constraints, planning reserve constraints, and CO₂ emission reduction constraints, are critical in shaping the solution space. The form of the optimization model and the solver further influence complexity, involving aspects like solver choice, pre-solving techniques, and the matrix range. Additional factors like coefficient magnitudes, degeneracy, feasibility and boundedness, data precision, matrix conditioning, implementation quality, hardware, and parallelization can significantly impact the computational effort required to solve the CEP problem.

Among these factors, we can only control some to reduce the CEP complexity. This work focuses on efficiently reducing the network size to improve CEP run time and tractability. The variables we control in the reduction process are the number of buses (N_{bus}) and the number of branches (N_{br}). The key question is, what is the relation between these variables and their contribution to the CEP complexity and fidelity?

To answer this question, we first define network size as a function of changes in these variables our reduction process:

$$\Delta N_{net} = \alpha * \Delta N_{bus} + \beta * \Delta N_{br} \quad (4-1)$$

The network size is a direct function of ΔN_{bus} and ΔN_{br} , but depending on the CEP formulation, the number of buses and branches can affect the problem size and its complexity differently. There are various CEP problems with different objective functions and constraints addressing different concerns in planning. Therefore, based on the formulation used for the CEP, we can identify the weights for ΔN_{bus} and ΔN_{br} to minimize the problem size.

Here, we focus on quantifying problem size and the CEP run-time based on the number of buses and branches. We propose an analytical method to identify the optimal bus elimination set specific to the CEP application to reduce its runtime. By analyzing the CEP problem formulation, we determine the impact of buses and branches on problem size and, thus, its runtime.

CEP is often formulated as a linear programming (LP) problem in which the run-time depends on the optimizer used and the algorithm implemented. For interior point methods, which we use in solving CEP, it is known that time complexity is polynomial in the number of variables [25, 26, 27]. The number of variables in the CEP model depends on the problem structure. A scenario-based CEP process is modeled in [28], which identifies core and adaptive investments for a flexible investment portfolio over the planning horizon based on all the scenarios under study. A simplified version of this model with only one scenario without considering the reliability and policy constraints is used here as a basic CEP formulation. For this basic CEP, which is carried out over N_y periods using N_{oc} operating conditions (blocks) per period, the decision variables include:

- capacities of expandable generators: $N_y \times N_g$
- capacities of retired generators: $N_y \times N_g$
- power generation of generators (one for every bus in every period and block): $N_y \times N_{oc} \times N_g$
- bus angles: $N_y \times N_{bus} \times N_{oc}$
- capacities of expandable branches: $N_y \times N_{br}$

Therefore, the number of CEP decision variables is given by

$$N_{var} = N_y \{2N_g + N_{oc}N_g + N_{oc}N_{bus} + N_{br}\} \quad (4-2)$$

Constraints in the CEP problem also impact its run time. For the basic CEP considered here, the constraints are:

- generator limits: $2 \times N_y \times N_{oc} \times N_g$
- power balance equations: $N_y \times N_{oc} \times N_{bus}$
- branch flow limits: $2 \times N_y \times N_{oc} \times N_{br}$

The number of equations that show up in the constraints is:

$$N_{con} = N_y N_{oc} \{2N_g + N_{bus} + 2N_{br}\} \quad (4-3)$$

Considering all occurrences of variables in the problem and constraints, we have a total number of variables given by:

$$N_{prob} = N_y \left\{ \underbrace{(3N_{oc} + 2)N_g}_{Fixed} + \underbrace{2N_{oc}N_{bus} + (2N_{oc} + 1)N_{br}}_{Variable} \right\} \quad (4-4)$$

A Ward-based network reduction that eliminates load buses and retains generator buses results in a reduction of ΔN_{bus} buses and changes the number of branches ΔN_{br} ; so, the first part is fixed, and we only have control over the second part in two ways to reduce the size of CEP problem: (i) reduce the number of buses N_{bus} , and (ii) reduce the number of branches N_{br} . Based on the above analysis of the basic CEP formulation, N_{bus} appears $2N_{oc}$ times, and N_{br} shows up $2N_{oc} + 1$ times in the problem formulation for each year. Since the number of operating conditions, N_{oc} , is typically much more significant than 1, the coefficients of N_{bus} and N_{br} are almost the same, so we can assign the same weights to them and have $\alpha = \beta = 1$.

Consider a given subnetwork in a power system for which the number of load buses that can be eliminated is $\Delta N_{bus,tot}$. It may be possible to eliminate a lesser number, i.e., $\Delta N_{bus} < \Delta N_{bus,tot}$, without

significantly increasing the network size, N_{net} , if those buses are chosen so that the number of eliminated branches, ΔN_{br} , is increased. This is attractive because having more buses (fewer reduced buses) generally translates to improved modeling fidelity. Moreover, a smaller number of branches results in fewer variables and constraints in the CEP problem, improving its run time. Therefore, we aim to minimize the problem size N_{prob} by maximizing $\Delta N_{net} = \Delta N_{bus} + \Delta N_{br}$ in our reduction process.

The approach taken in this chapter is to identify, for each possible value of bus eliminations $\Delta N_{bus} = 1, \dots, \Delta N_{bus,tot}$, the bus set $S_e(\Delta N_{bus})$, which maximizes ΔN_{br} . For each of these bus sets, we develop the reduced model and assess its accuracy by comparing it to the results of the original network. We then create a plot of accuracy vs run-time, and from this plot, we choose the bus elimination value that provides good modeling fidelity and significant run-time reduction.

4.4 Problem Formulation

To reduce the network size, we first choose the part of the network to reduce. For small systems, we may operate on the entire network. However, for large industry-sized systems, we typically divide the network into several zones and then reduce the size of a subnetwork (a zone) within the larger network. We aim to reduce zones by eliminating buses using Ward reduction. In the following, we define Problem 1, where, for a specified number of buses to eliminate (ΔN_{bus}), the buses are chosen to maximize the reduction in the number of branches (ΔN_{br}). We also define Problem 2, where the number of buses is selected to maximize ΔN_{net} . Problem 1 is a subproblem to Problem 2, meaning Problem 2 can be solved by repeatedly solving Problem 1 for different values of ΔN_{bus} .

Problem 1:

Find the best set of a specified number of buses to eliminate:

$$\begin{aligned} & \max \Delta N_{br}(\underline{x}) \\ & \text{s. t.} \\ & \sum_j x_j = J \\ & x_j = 0 \forall j \in \text{retained bus list} \end{aligned}$$

Problem 2:

Find the best set of any number of buses to eliminate:

$$\begin{aligned} & \max \Delta N_{net}(\underline{x}) \\ & \text{s. t.} \\ & \sum_j x_j \leq J_{max} \\ & x_j = 0 \forall j \in \text{retained bus list} \end{aligned}$$

Denoting the total number of buses in the zone as J_{tot} , \underline{x} is a $J_{tot} \times 1$ vector with binary elements 0 or 1 ($x_j \in \{0,1\}$). Each x_j indicates whether its bus b_j in the zone should be retained (if $x_j=0$) or

eliminated (if $x_j=1$). J_{max} is the maximum number of buses allowed to be eliminated in the zone, and J is the designated number of buses to be eliminated, satisfying $J \leq J_{max}$. As indicated in Section II, ΔN_{br} is the number of branch reductions, and ΔN_{bus} is the number of bus reductions. For a given ΔN_{bus} , ΔN_{br} is obtained by performing Ward reduction on the zone.

The evaluation of ΔN_{br} and ΔN_{bus} using Ward reduction is independent of the order in which buses, or zones, are eliminated, i.e., bus elimination is commutative. This is supported by the fact that for a system containing a retained subsystem "I", and two subsystems to be eliminated "E₁" and "E₂", as shown in Figure 4-3, the solution is the same, given by (4-5), independent of which subsystem is eliminated first.

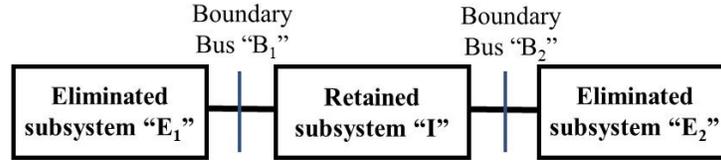


Figure 4-3: System used for showing commutative property of bus elimination

$$Y_{reduced} = \begin{bmatrix} Y_{II} & Y_{IB_1} & Y_{IB_2} \\ Y_{B_1I} & Y_{B_1B_1} - Y_{B_1E_1} \times Y_{E_1E_1}^{-1} \times Y_{E_1B_1} & Y_{B_1B_2} \\ Y_{B_2I} & Y_{B_2B_1} & Y_{B_2B_2} - Y_{B_2E_2} \times Y_{E_2E_2}^{-1} \times Y_{E_2B_2} \end{bmatrix} \quad (4-5)$$

4.5 Bus Selection: Solving Problem 1

Various approaches are considered to provide reasonable solutions to Problem 1, i.e., for a specified ΔN_{bus} , to identify sets of eliminated buses that yield a sparse Y_{bus} matrix and, therefore smaller number of branches. These approaches are described in the following five subsections.

4.5.1 Lowest Connection Degree (LCD)

This approach eliminates the first k buses with the lowest connection degree, where the connection degree is determined by the number of branches connected to each bus. It is a straightforward and fast method to choose buses and reduce the network. However, this approach does not consider the subsequent effects of the elimination process. While this method may result in lesser-quality solutions, it can be used to reduce the search space when initiating other methods.

4.5.2 Approximate Minimum Degree (AMD)

AMD selects buses that introduce the fewest new branches. It involves trial simulations of every feasible alternative at each step and provides near-optimal results. We eliminate buses in the obtained order and skip buses that should be retained. The difference between this method and the exhaustive search is that instead of trying all the possible combinations, AMD uses techniques based on the quotient graph for matrix factorization to obtain computationally cheap bounds for the minimum degree [27]. This method can also be modified to better fit the problem by using it iteratively after eliminating each suggested bus until the desired number of buses is eliminated [29], which we have implemented here.

4.5.3 Nested Dissection (ND)

The ND method is a multilevel graph partitioning algorithm employed to generate orderings of sparse matrices that minimize fill-ins. The algorithm operates by iteratively simplifying the adjacency matrix of the given graph. This simplification is achieved by the strategic aggregation of nodes and edges. The initial phase of the algorithm merges nodes and edges and then reorders the reduced graph. Subsequently, it employs refinement steps to reverse the simplification and produce a reordering of the original graph [30].

4.5.4 Genetic Algorithm (GA)

Problem 1 may be solved using a genetic algorithm (GA), where the GA is employed to identify the best set of buses to eliminate. The algorithm starts with an initial population of possible solutions (sets of buses), with each individual denoted by x , which is, as indicated in Section III, a vector containing binary elements 0 or 1. Individuals with higher fitness (better reduction quality) are selected through successive generations, undergo genetic operations (crossover and mutation), and produce offspring. This process promotes the evolution of better solutions. By iteratively refining the population, the GA efficiently searches the solution space to identify the optimal set of buses to eliminate. The parameters used for the GA in this work are 100 generations, 50 individuals per population, 20 parents, and a 0.1 mutation rate. Then, the GA is run ten times, and the best outcome is saved as the optimal solution.

We may use GA with a randomly chosen initial population. Still, the GA finds reasonable solutions with fewer iterations if the initial population is seeded with solutions from one or more of the methods described above; we refer to this approach as Seeded GA.

4.5.5 Exhaustive Search (ES)

This method involves considering all possible combinations of buses as a set to be reduced, providing the absolute optimal solution. However, given this method's computational intensity, it is unsuitable for large subnetworks. Nevertheless, this method is feasible in the case of a small network or in applying the reduction to small zones. Applying ES involves two main steps: (1) enumerate all possible sets of buses, and (2) test each set to evaluate its impact on the network reduction. By employing this approach, we are guaranteed to find the optimal bus set; the result can then be used to evaluate other methods' effectiveness.

4.6 Results and Discussion

We use the IEEE 118-bus test system to compare the performance of the selected methods. Initially, the full 118-bus system is analyzed, followed by a zonal analysis to confirm Problem 1 solutions. This zonal approach aligns with the practical applications of bus elimination. Lastly, we demonstrate the run-time/modeling fidelity trade-off for a subsystem of the 118-bus network.

4.6.1 Application to Entire 118-Bus System

Our investigation starts by analyzing the performance of various methods when eliminating a fixed number of buses. We examine the full 118-bus network, retaining 40 buses and targeting 20 for

elimination from the remaining 78 to maximize ΔN_{net} . Methods described in Section 4.5 (except for the exhaustive search) are applied to identify the "best" set of 20 buses to be eliminated. Then, Ward reduction is used on the IEEE 118-bus system with 179 branches to obtain the reduced equivalent network. Table 4-2 presents the new number of branches and total variable reduction for each method. Here, the obtained ΔN_{net} depends on which method is used. The Seeded GA method provides the best solution, outperforming ND, AMD, and LCD methods.

Table 4-2: Network reduction results by different methods for $\Delta N_{bus}=20$

Method	N_{br}	ΔN_{net}
Random Selection	169	30
Lowest Connection Degree (LCD)	157	42
Approximate Minimum Degree (AMD)	156	43
Nested-Dissection (ND)	154	45
Genetic Algorithm (GA)	145	54
Seeded GA	143	56

To address problem 2, we solve problem 1, using each of the methods above, for a range of different bus elimination values. Results are shown in Figure 4-4; the following observations are made:

- Each method results in a different ΔN_{br} for a fixed ΔN_{bus} and for various values of ΔN_{bus} .
- While the GA with random seed initialization proves effective for a moderate number of buses marked for elimination, its effectiveness wanes when dealing with larger numbers of buses.
- The Seeded GA, which is initialized with solutions from the other methods, has the best performance, always finding bus elimination sets having the largest value of ΔN_{br}^{max} .
- It is observed in GA results that eliminating more buses does not necessarily lead to more reduction in the number of branches. Therefore, it is important to find the optimum ΔN_{bus} with the maximum of ΔN_{br} . This motivates the approach described in Subsection 4.5.

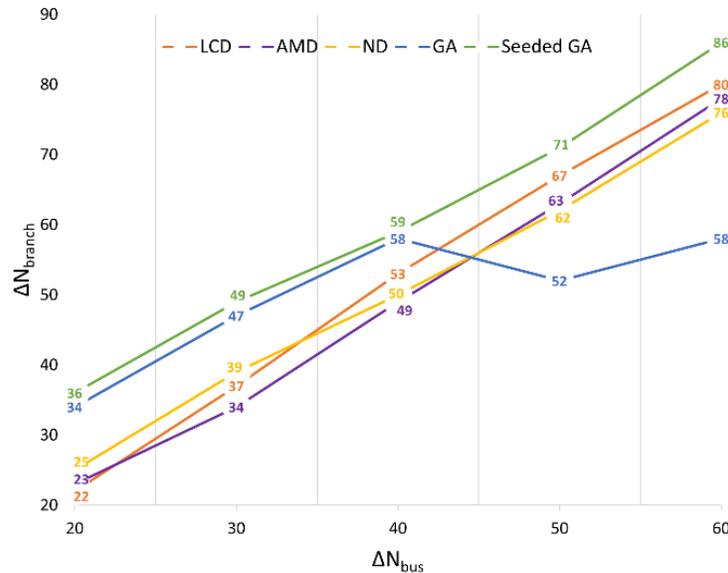


Figure 4-4: ΔN_{br} for different levels of bus reduction, ΔN_{bus} , by various methods

4.6.2 Application to Individual 118-Bus System Zones

The proposed approach is then applied to several zones within the 118-bus system. The smaller size of the network makes it possible to run an exhaustive search and compare its results with our proposed seeded GA method. Here, we use the zones identified by the minimum spanning tree (MST) method, as described and illustrated in Section 3.2. This method divides the network into 8 zones. Our goal here is to choose the best buses to eliminate within each zone. For the zones with more than three buses to select from for elimination, we apply the methods above and find $\Delta N'_{br,max}$, and the associated number of buses to eliminate, $\Delta N'_{bus}$. Because the buses to select from are relatively few, we are able to apply the exhaustive search method here, which enables assessment of our various search methods.

Results for zone 2 which is the largest zone in the network, are provided in Figure 4-5 for each number of ΔN_{bus} . We see that the reduction in the number of branches is not a linear function of number of eliminated buses, i.e. eliminating more buses does not always result in more reduction in the number of branches. Inspection of these results also indicates that the three methods LCD, AMD, and ND do not always obtain optimal set of buses that result in the maximum reduction in the number of branches. In this case, both methods GA, Seeded GA obtain the optimal results for all the ΔN_{bus} as the exhaustive search.

Among different ΔN_{bus} and their optimal set of buses, we choose the one that leads to the maximum ΔN_{net} , which is $\Delta N_{bus}=14$ for zone 2 with $\Delta N_{bus}=16$ and $\Delta N_{net}=30$.

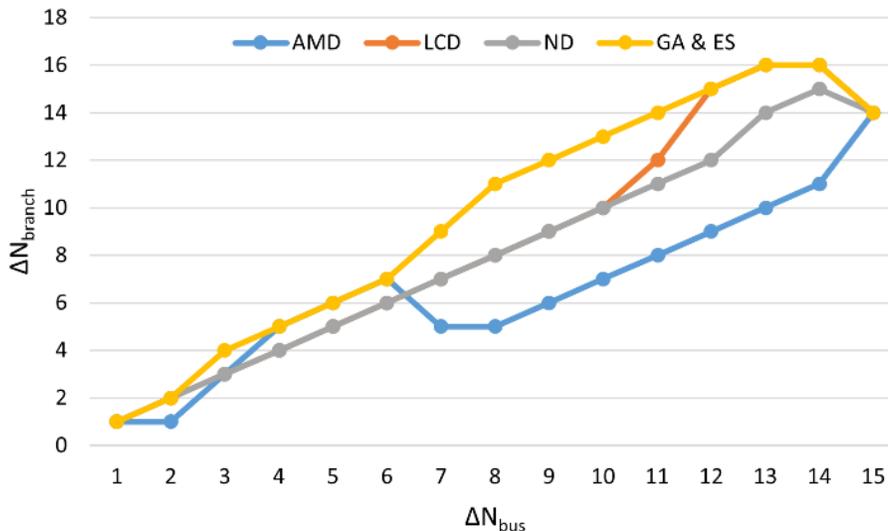


Figure 4-5: Optimal bus sets for each ΔN_{bus} for zone 2 by various methods

Final results of each method for the other zones are provided in Table 4-3. As expected, the graph theory methods do not always obtain optimal results. However, the GA and Seeded GA methods obtain correct results for all zones. These results, combined with the result from Subsection 4.6.1, confirm that the Seeded GA is an effective approach, as it takes advantage of the results from the LCD, AMD, and ND methods and improves on them via heuristic search.

Table 4-3: Comparing ΔN_{br} , for different methods applied to zones

Zone	Buses to select from	LCD	AMD	ND	GA	Seeded GA	ES
1	5	4	4	4	4	4	4
2	15	16	14	15	16	16	16
4	7	7	5	5	7	7	7
7	4	6	6	6	6	6	6
8	12	17	17	17	17	17	17

4.6.3 Assessing Run-time/Fidelity Tradeoff

The selection of the appropriate bus elimination set involves a three-step process, which we illustrate using the IEEE 118-bus system, consisting of 186 branches and 54 generators. In this test, we retain the generator buses and those connected to congested branches.

First, we illustrate the resulting ΔN_{br} values for each bus set in Figure 4-6, showing that reducing the number of buses does not necessarily decrease the number of branches. Thus, it is important to select a bus set that avoids introducing additional branches to the network.

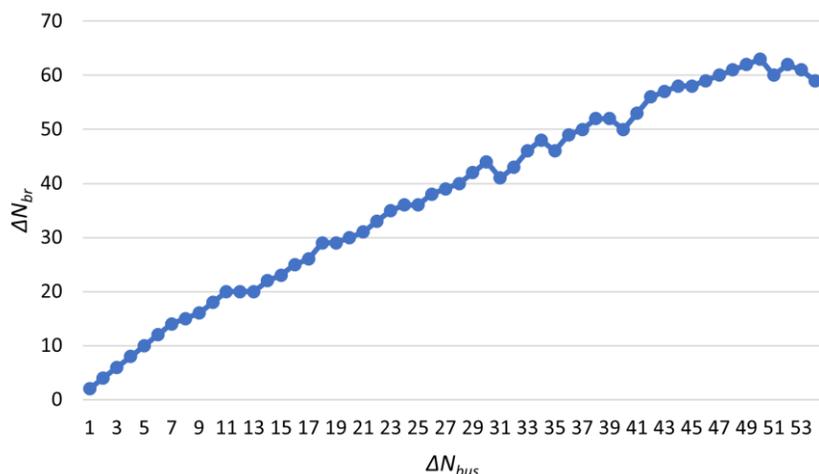


Figure 4-6: ΔN_{br} of each bus set elimination for the 118-bus system

Second, we develop the reduced equivalent for each bus set, including equivalent circuit capacity estimates. We are developing and using an alternative capacity estimation approach, which is presented in Section 6.1. The CEP is run on the equivalent network of the last 15 sets of selected buses, and their results are compared with the full network, demonstrating CEP fidelity in terms of relative error.

Third, we develop visuals to observe tradeoffs between model fidelity and run-time (or a run-time proxy, "ticks" which indicate run-time but without the influence of machine loading or memory usage). Figure 4-7 plots run time vs. CEP error, with the number of variables labeled beside each point. Desirable bus elimination sets should be computationally fast (low in ticks) and highly accurate (low in error), such as points with ΔN_{net} of 94 ($\Delta N_{bus}=41$, $\Delta N_{br}=53$) and 111 ($\Delta N_{bus}=49$, $\Delta N_{br}=62$). These points have a smaller ΔN_{bus} and higher ΔN_{br} , indicating that eliminating fewer

buses leads to higher fidelity in the CEP model, consistent with the principle that more buses lead to improved modeling fidelity.

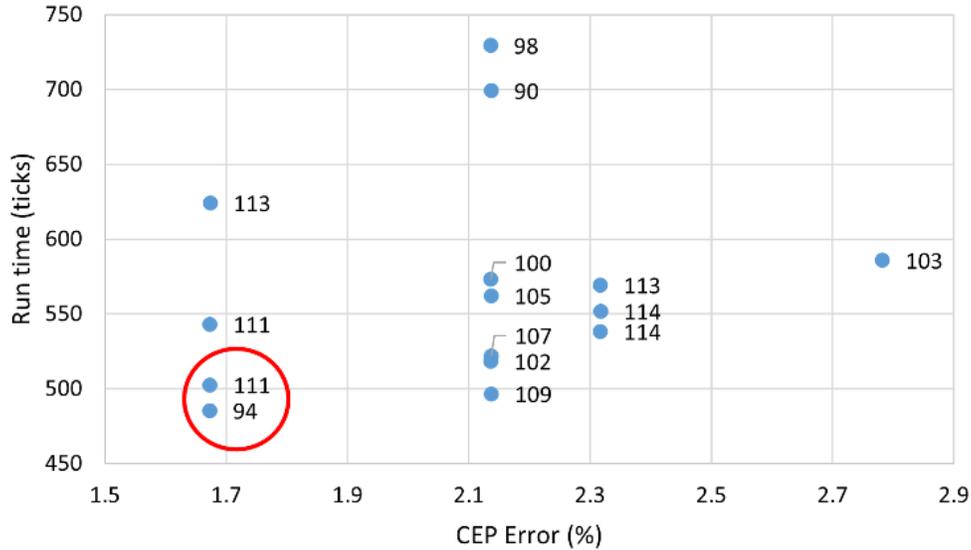


Figure 4-7: Run time and CEP error for each bus set for the 118-bus system

This analysis highlights the importance of selecting the right combination of ΔN_{bus} and ΔN_{br} to achieve the highest fidelity and fastest computation. However, determining the appropriate weights for ΔN_{bus} and ΔN_{br} depends on the network’s topological structure, k . These weights, α_k and β_k , can be determined by analyzing the fidelity and run-time of each specific structure. This analysis can then be applied to similar topological structures in real-world power networks. To accomplish this, we first need to identify frequent topological structures in the power networks by applying graph mining approaches to a significant database of networks. Then, the proper weights for each structure can be obtained by using the same experiments proposed in this project. The final step is to detect similar structures in the network under study and apply the corresponding coefficients to achieve the best tradeoff between fidelity and complexity.

4.7 Conclusion

The focus of this chapter is on improving network reduction in power system planning by identifying desirable sets of buses for elimination. We find that considering the reduction in the number of branches is helpful in this process. By comparing methods, the study determines a suitable approach for selecting a bus set that maximizes ΔN_{br} given a targeted number of buses to eliminate. Repeating this approach for all possible bus elimination values identifies candidate bus sets. A bus set is chosen that provides a good tradeoff between the number of variables (a proxy for CEP run-time) and modeling fidelity. Network topological structures are another important aspect that plays a role in the CEP run time and fidelity, which we aim to address in our future work.

5. Bus Aggregation

Once the internal load buses have been eliminated, our subsequent objective is to reduce the number of internal generator buses. If we choose bus elimination to address the generator buses, it will result in the introduction of numerous partial generators, thereby making the evaluation of their parameters complex. Hence, we resort to bus aggregation.

One type of aggregation method is based on similarities between vectors of power transfer distribution factors (PTDFs), where each vector corresponds to a bus in the network and each element (row) of the vector corresponds to a branch in the network, so that an element of a vector indicates the effect of an injection at the vector's bus on the flow of the element's branch, see [11]. Therefore, similarity of two vectors, as detected by a clustering algorithm such as K-means, indicates that injections at the two vectors have similar effects on flows in the network. A key problem here is the similarity detection and the identification of suitable hyperparameters controlling the algorithm as their selection is influential, and it becomes challenging when utilizing the clustering algorithms for large-scale systems. As a result, we have devised an alternative bus aggregation method based on the system's topology. This idea was originally proposed in [7]. In that work, only a few specific types of network topologies were considered. In this endeavor, we aim to develop a more comprehensive method that can be applied to networks of any topology.

5.1 Topology-based Aggregation

5.1.1 Topology Building for Aggregated Generator Buses

This step involves determining the topology of the resulting aggregated subnetwork. When a boundary bus can be reached by one internal generator bus through another internal generator buses, those internal generator buses are directly connected to that boundary bus. In cases where multiple internal generator buses share the same set of directly connected boundary buses, they can be merged into a single generator bus. For instance, in Figure 5.1, buses 1 and 6 of the left-hand diagram share an identical set of boundary buses (5 and 12) to which they are directly connected. Consequently, they can be consolidated into a new bus denoted as Bus 119. However, an additional internal generator bus, Bus 4, has direct access to boundary buses 5, 12, and 13, which differs from generator buses 1 and 6. Hence, it cannot be combined with these two buses.

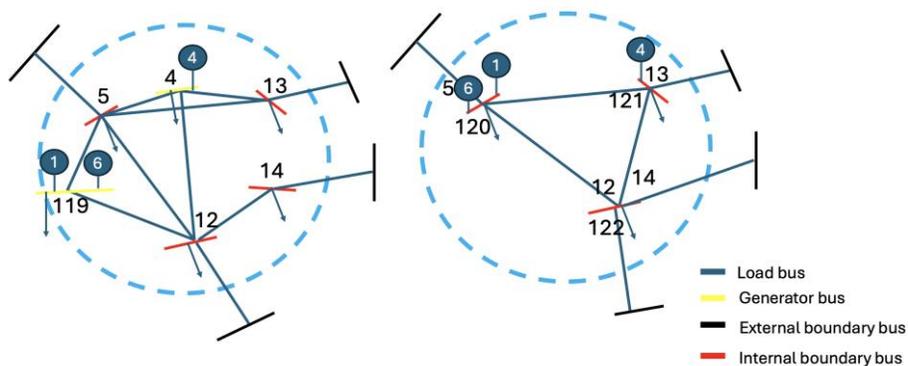


Figure 5-1: Generator Aggregation Effect of One Zone in IEEE 118

5.1.2 Impedance Calculation

After determining a subnetwork's topology, the next step is to compute the impedance of newly formed branches. From Ohm's law, we have

$$Y_{Bus} \cdot V_{Bus} = I_{inj} \quad (5-1)$$

If we set $I_{inj}^{1,2}$ as $[1, -1, 0, \dots, 0]^T$, then using $Y_{Bus}^{-1} \cdot I_{inj}^{1,2}$ can give us bus voltage vector $V_{Bus}^{1,2}$ when one unit of current is injected at bus 1 and withdrawn at bus 2. $V_{Bus-j}^{1,2}$ is the element j of $V_{Bus}^{1,2}$. If we go one step further, $\frac{V_{Bus-1}^{1,2} - V_{Bus-2}^{1,2}}{1}$ can calculate the Thevenin impedance $Z_{Thevenin}^{1,2}$ seen from buses 1 and 2. Here, a 1 in the denominator means one unit of injected current. From this perspective, we can put boundary buses ahead of internal buses in the sorting process and then define a transform matrix as

$$T_r = \begin{bmatrix} 1 & 1 & 1 & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ \vdots & \vdots & \vdots & \vdots & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [I_{inj}^{1,2}, \dots, I_{inj}^{N_{boundary}-1, N_{boundary}}] = I_{inj} \quad (5-2)$$

The size of T_r is $N_{bus} \cdot \frac{N_{boundary} \cdot (N_{boundary}-1)}{2}$, where N_{bus} is the number of buses within this zone and $N_{boundary}$ is the number of boundary buses in this zone. Using T_r , we can define the Thevenin impedance matrix as:

$$Z_{thevenin} = T_r^T \cdot V_{Bus} = T_r^T \cdot Y_{Bus}^{-1} \cdot I_{inj} = T_r^T \cdot Y_{Bus}^{-1} \cdot T_r \quad (5-3)$$

where $V_{Bus} = [V_{Bus}^{1,2}, \dots, V_{Bus}^{N_{boundary}-1, N_{boundary}}]$, and $V_{Bus}^{i,j}$ is the bus voltage vector when there is one unit of current being injected at boundary bus i and withdrawn at boundary bus j.

Diagonal elements of $Z_{thevenin}$ are the Thevenin impedances seen from each two boundary buses. The meaning of its off-diagonal elements is voltage differences measured at two boundary buses when one unit of current is injected at another two boundary buses. For example, $Z_{thevenin}^{3,3}$ means the Thevenin impedance of the third pair boundary buses, i.e bus 1 and bus 3, and $Z_{thevenin}^{3,4}$ means the voltage differences measured at the third pair boundary buses 1 and 3 when one unit of current is injected at the fourth pair, buses 1 and 4.

We want the aggregated subnetwork to have similar external characteristics as the original full network; therefore we can assume they have the same $Z_{thevenin}$ matrix. Then we have

$$T_r^T \cdot Y_{Bus}^{-1} \cdot T_r = Z_{thevenin} = T_{r\ reduced}^T \cdot Y_{Bus\ reduced}^{-1} \cdot T_{r\ reduced} \quad (5-4)$$

The $T_{r_reduced}^T$ is

$$T_{r_reduced} = \begin{bmatrix} 1 & 1 & 1 & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ \vdots & \vdots & \vdots & \vdots & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The $T_{r_reduced}^T$ appears similar to T_r , except its number of rows is N_{BUS} reduced, which is determined by aggregated topology in the last section. In addition, the Y_{BUS} reduced can be expressed as:

$$Y_{BUS\ reduced} = A_R^T \cdot D_Y^R \cdot A_R \quad (5-5)$$

where A_R is the incidence matrix of the reduced internal network and D_Y^R is $diag(\frac{1}{X_R})$, formed by the objective branch admittances of the reduced network that has similar external characteristics with original network. We define $K = pinv(A_R)$ and $B = pinv(T_{r_reduced})$ where function $pinv$ indicates the pseudo-inverse. Combining equation (5.4), (5.5), we obtain a least-square estimation for $Y_{BUS\ reduced}^{-1}$ according to

$$B^T \cdot Z_{thevenin} B = Y_{BUS\ reduced}^{-1} = \left[A_R^T \cdot D_Y^R \cdot A_R \right]^{-1} = K \cdot D_X^R \cdot K^T \quad (5-6)$$

where $Y_{BUS\ reduced}^{-1}$ is the estimator of $Y_{BUS\ reduced}^{-1}$, and D_Y^R is the estimator of D_Y^R , and D_X^R is the inverse of D_Y^R , representing the estimated diagonal matrix formed by branch impedances in the reduced network. And thus, X_R is estimated via

$$D_X^R = A_R \cdot Y_{BUS\ reduced}^{-1} \cdot A_R^T = A_R \cdot B^T \cdot Z_{thevenin} \cdot B \cdot A_R^T \quad (5-7)$$

whose diagonal elements are estimated branch impedances for the reduced network, i.e., X_R .

If the reduced network does not have the same list of Thevenin impedances as the original subnetwork, then the result of this approach results is the best (in the sense of least square error) approximation.

After that, we can use equation (5-7) to estimate X_R from the previous $Y_{BUS\ reduced}^{-1}$. Despite the expectation that the estimated diagonal matrix D_X^R solved in equation (5-7) would be diagonal, it is consistently observed to be non-diagonal. Using the equation

$Z_{thevenin} = T_{r_reduced}^T \cdot K \cdot D_X^R \cdot K^T \cdot T_{r_reduced} = T_{r_reduced}^T \cdot K \cdot (D_X^R + N_{TK}) \cdot K^T \cdot T_{r_reduced}$, where N_{TK} is in the null space of $T_{r_reduced}^T \cdot K$, we can take advantage of N_{TK} to diagonalize

matrix D_X^R via the least squares method. If we define N_{TK}^i as matrix N_{TK} without row i and d_X^i as i^{th} column of D_X^R without i^{th} row, then we can use $\left[(N_{TK}^i)^T \cdot N_{TK}^i \right]^{-1} N_{TK}^i \cdot d_X^i$ to obtain the least-square solution of s^i , which forces all elements, except row i , in column i of D_X^R to be zero. By applying analogous procedures to the remaining columns, we can ultimately make D_X^R matrix with a more diagonal shape using the equation below

$$D_X^{R\text{new}} = D_X^R - N_{TK} \cdot [s_1, s_2, \dots, s_n].$$

The newly determined impedance values on the diagonal parts of the D_X^R contribute to ensuring that the reduced network exhibits Thevenin impedances that are more closely aligned with those of the original network.

As this method serves as an approximate solution to achieve a reduced subnetwork with external characteristics closely resembling the original subnetwork, it is essential to ensure the effectiveness of the reduction. To assess the reduction's impact, we employ a method that involves testing the performance metric of flow (identified as the OPM in Section 1.4.2.1) on selected key branches after constructing a reduced subnetwork. If the similarity between the reduced subnetwork and the original network exceeds a predetermined threshold deemed acceptable, we refrain from further aggregation within that zone.

Branches with impedance exceeding 20 pu are eliminated as they are considered open circuits (such branches can appear as a result of Ward reduction). Due to the removal of these branches, buses (along with their associated generators and loads) which become isolated from the main system are eliminated. This is due to the significant impedances that separate the two sides, making it difficult for the power from the rest of the system to interact with them.

5.2 Bus Aggregation Method – Quotient Graph (QG)

A significant reduction in system size may be achieved via the implementation of topology-based aggregation on the internal generator buses of each zone, as described in Section 5.1.1. However, considering our primary objective of minimizing system size, we utilize another form of bus aggregation method as a second-round aggregation following the topology-based aggregation.

This second aggregation method is based on the influence of buses on system branch flows. The QG aggregation is positioned after the topology-based generator aggregation due to greater confidence in the latter approach. The primary challenge with QG aggregation lies in selecting appropriate hyperparameters. To mitigate this, hierarchical clustering was employed to group buses by mutual distances, offering a more precise alternative to cluster iteration, though it remains somewhat approximate. The inclusion of the QG step aims to further reduce system size, but with some inaccuracy. We have found that performing the QG step first reduces the effectiveness of the topology-based aggregation, which has proven to be more accurate in various test cases.

The bus aggregation method, based on similar bus features, was initially proposed by [8] and [9]. The authors of [8] used locational marginal prices (LMP) for aggregation, while those of [9] employed Power Transfer Distribution Factors (PTDFs) to reduce the PTDF matrix while retaining the system's structural characteristics. However, this method's dependence on the operational state

can cause variations in the interzonal impedances of the reduced system, making it challenging to achieve a single, consistent network under varying load conditions or scenarios. Reference [10] addressed this by using the pseudo-inverse for impedance calculations, and [11] later refined the approach using PTDF columns for clustering and the least squares method for computing interzonal impedances. Despite these advancements, the method's dependency on operational conditions remains a limitation, preventing the formation of a unique network when multiple scenarios exist. The next section addresses this issue and proposes a solution.

5.2.1 Quotient Graph Theory

In the DC formulation, the branch flow and bus power injections are linearly related to bus voltage angles as:

$$P_{inj} = B_{Bus} \cdot \theta \quad (5-8)$$

and

$$P_{flow} = D \cdot A \cdot \theta = B_{branch} \cdot \theta \quad (5-9)$$

where

- P_{inj} represents the bus net power injection vector
- P_{flow} represents the branch power flow vector
- B_{bus} is the bus susceptance matrix
- D is the primitive susceptance matrix
- A is the node-arc incidence matrix.
- B_{branch} is equal to $D \cdot A$.

If we combine (5-8) and (5-9), we then obtain

$$P_{flow} = B_{branch} \cdot B_{Bus}^{-1} \cdot P_{inj} \quad (5-10)$$

From (5-10) we can also obtain the system PTDF matrix [31] as:

$$\phi = B_{branch} \cdot B_{Bus}^{-1} \quad (5-11)$$

Assuming the system has N buses and L branches, then the size of the PTDF matrix is $L \times N$. Each row of the PTDF matrix corresponds to one branch and each column represents one bus. In an ideal scenario, if two PTDF columns are exactly identical, the net power injection at one bus will have an equivalent impact on the system power flow as the net power injection at another bus. This implies that the combined net power injections of these buses collectively influence the system power flow, rather than individual contributions. As a result, we can group these buses into a single zone and utilize a single aggregated bus to represent them. Furthermore, the loads and generation associated with the buses in the original system will be aggregated onto these newly designated representative buses.

5.2.2 Clustering Algorithm

Clustering, as an unsupervised learning algorithm, focuses on grouping objects with similar characteristics together while separating distinct objects into different clusters. In our case, each bus represents an object, and the elements of the corresponding PTDF column serve as its defining properties. Consequently, we assign buses to different clusters based on the similarity of their PTDF columns. Methods used to detect similarity among PTDF columns are as follows:

- **Centroid-based method (K-Means):** K-means clustering, which was proposed in [32], is widely recognized as the predominant centroid-based clustering method. It involves representing each cluster by a central vector, which may not necessarily be a member of the

dataset. The centroid-based approach relies on two hyperparameters: the specified number of clusters (k) and the initial center points. Points close to a center point are assigned as members of that cluster, and their mean value becomes the new center point. Through multiple iterations, the algorithm strives to reach a local optimum where the squared distances from the cluster centroids are minimized. There exist several notable variations of K-means clustering, including K-medoids (which employs actual data points as center points instead of mean values), K-medians (which uses the median value instead of the mean), and K-means++ (which employs a less random approach for selecting initial center points). While K-medoids and K-medians are more robust against outliers due to their non-reliance on mean values as center points, they require additional iterations, resulting in longer convergence times.

- **Connectivity-based method (Hierarchical clustering):** Connectivity-based clustering relies on the fundamental concept that objects in close proximity exhibit a stronger relationship. As a result, the two closest objects are initially merged into a single object, followed by the merging of the subsequent closest pair of objects. This merging process continues until all objects have been merged into a cohesive structure. Hierarchical clustering, which was proposed in [33] initially, encompasses a diverse range of methods, distinguished by their approaches to distance computation (such as Euclidean or Mahalanobis distances), linkage types (such as average or complete linkage), and the direction of clustering (either agglomerative or divisive).
- **Distribution-based method (GMM):** The distribution-based method operates under the assumption that data points are generated according to specific statistical distribution models. For instance, GMM (Gaussian Mixture Model) clustering, proposed in [34], relies on the assumption of Gaussian distribution. However, this assumption can be overly restrictive in many cases, and it may not be possible to find a suitable distribution type for many real-world datasets.
- **Other methods:** In addition to the methods mentioned earlier, two other types of clustering methods are commonly used: the Density-based method (DBSCAN) and the graph-based method (spectral clustering). However, applying DBSCAN may lead to incorrect results because it clusters buses into two layers, potentially merging buses with significantly different PTDF columns. On the other hand, the graph-based method requires the construction of a connected network between buses, which is unnecessary and troublesome for our purpose. Therefore, these two methods, along with the Gaussian Mixture Model (GMM), are deemed unsuitable for our current scenario. Hierarchical clustering and K-means are two suitable types of clustering algorithms, each with its own advantages. Hierarchical clustering is generally faster compared to other algorithms, but it may result in a majority of buses being grouped into a small number of clusters, while the remaining buses are divided into numerous other clusters. On the other hand, K-means and K-medoids clustering tend to produce more evenly distributed results, but they are slower due to their iterative nature.

5.2.3 Calculation of Parameters

Once the system buses have been grouped into different clusters, the next step is to construct the reduced system network, specifically the inter-group branches, based on the original system. We use the term “inter-group” here to differentiate it from the previously used term “interzonal,” as QG aggregation is applied upon each zone individually and the groups are considered as subzones in this context. If there are branches directly connecting two clusters, a single branch

should be present in the reduced network. Since the subsequent expansion planning study will utilize DC power flow to assess operational costs, our focus here is to determine the impedance of these inter-group branches, disregarding any resistance considerations.

Our primary objective is to ensure that the inter-group power flows in the reduced network closely match the corresponding values in the original full network. This alignment provides that expansion planning conducted on the reduced network will generate similar inter-group transmission investments as those obtained from the full network.

We desire to express the inter-group power flows as a function of the PTDF matrix Φ and the power injections P_{inj} . To do so, we recall from (5-10) and (5-11) that

$$P_{flow} = \Phi P_{inj} \quad (5-12)$$

We now define a transform matrix Π which is used to convert branch flows to inter-group flows, i.e.,

$$P_{flow}^{inter-group} = \Pi P_{flow} \quad (5-13)$$

Substitution of (5-12) into (5-13) results in

$$P_{flow}^{inter-group} = \Pi \Phi P_{inj} \quad (5-14)$$

Transform matrix Π is used to convert branch flows to inter-group flows. If the number of inter-group branch is L_R and the number of branches in the original network is L , then Π is an $L_R * L$ matrix. Assuming it is the branches i_1 and i_2 that connect group 1 and group 2, and inter-group 1-2 is regarded as the first inter-group branch in the reduced network, then the elements corresponding to i_1 and i_2 in first row of the Π matrix are non-zeros. If these two branches are in the same direction as the inter-group branch, then they will be 1 in Π , or they will be -1. The remaining elements in first row of Φ will be zero.

The inter-group power flows in the reduced network can be represented as:

$$\left(P_{flow}^{inter-group}\right)_R = \Phi_R \left(P_{inj}\right)_R = (B_{branch})_R \cdot (B_{Bus}^{-1})_R \cdot \left(P_{inj}\right)_R \quad (5-15)$$

Here, Φ_R , $(B_{branch})_R$, $(B_{bus})_R$ and $\left(P_{inj}\right)_R$ are the PTDF matrix, branch susceptance matrix, susceptance matrix, and power injection vector, respectively, of the reduced network. Then we have

$$(B_{branch})_R = \text{diag}\left(\frac{1}{X_R}\right) * C_R \quad (5-16)$$

$$(B_{bus})_R = C_R^T * (B_{branch})_R \quad (5-17)$$

where C_R^T is the incidence matrix of reduced network. Combining equations (5-15)-(5-17), we have

$$\left(P_{flow}^{inter-group}\right)_R = \text{diag}\left(\frac{1}{X_R}\right) * C_R * \left(C_R^T * \text{diag}\left(\frac{1}{X_R}\right) * C_R\right)^{-1} * \left(P_{inj}\right)_R \quad (5-18)$$

The ideal condition is we can find a perfect solution of X_R , which makes equation (5-14) exactly the same as equation (5-18). Under that condition, we can have

$$\Pi \Phi P_{inj} = \text{diag}\left(\frac{1}{X_R}\right) * C_R * \left(C_R^T * \text{diag}\left(\frac{1}{X_R}\right) * C_R\right)^{-1} * \left(P_{inj}\right)_R \quad (5-18)$$

However, we cannot solve X_R from equation (5-18). To address this, we define matrix Γ to transform $\left(P_{inj}\right)_R$ into P_{inj} according to

$$P_{inj} = \Gamma * \left(P_{inj}\right)_R \quad (5-19)$$

Division matrix Γ is used to transform between P_{inj} and $(P_{inj})_R$. It is a $N * N_R$ matrix where each column corresponds to a bus in the reduced system and each row corresponds to a bus in the original system. If bus p belongs to group k , then (p,k) of Γ will be nonzero and its value is the ratio of net power injection at bus p to total power injection in group k . And thus, equation (5.18) becomes

$$\pi\phi\Gamma * (P_{inj})_R = \phi_R * (P_{inj})_R = \text{diag}\left(\frac{1}{X_R}\right) * C_R * \left(C_R^T * \text{diag}\left(\frac{1}{X_R}\right) * C_R\right)^{-1} * (P_{inj})_R \quad (5-20)$$

To have equation (5-20) hold for all operation states, it must be the case that

$$\phi_R = \pi\phi\Gamma = \text{diag}\left(\frac{1}{X_R}\right) * C_R * \left(C_R^T * \text{diag}\left(\frac{1}{X_R}\right) * C_R\right)^{-1} \quad (5-21)$$

which can be further converted to

$$\phi_R * \left(C_R^T * \text{diag}\left(\frac{1}{X_R}\right) * C_R\right) = \text{diag}\left(\frac{1}{X_R}\right) * C_R \quad (5-22)$$

And then

$$(\phi_R * C_R^T - I) * \text{diag}\left(\frac{1}{X_R}\right) * C_R = 0 \quad (5-23)$$

We define c_r^i as the i^{th} column of C_R and x_r^i as branch i 's impedance, then equation (5.23) can be rewritten as:

$$(\phi_R * C_R^T - I) * \text{diag}(c_r^i) * \frac{1}{x_r^i} = 0, i=1, \dots, N_R \quad (5-24)$$

Equation (5-24), which contains a vector of equations, can be written in matrix multiplication format as below.

$$\Lambda * \frac{1}{X_R} = 0 \quad (5-25)$$

where

$$\Lambda = \begin{bmatrix} (\phi_R * C_R^T - I) * \text{diag}(c_r^1) \\ (\phi_R * C_R^T - I) * \text{diag}(c_r^2) \\ \vdots \\ (\phi_R * C_R^T - I) * \text{diag}(c_r^{N_R}) \end{bmatrix} \quad (5-26)$$

In addition, one more constraint can be added, which is

$$\frac{1}{x_r^i} = \left(\sum_{k=1}^{K_i} \frac{1}{x^k}\right)^{-1} \quad (5-27)$$

x^k is the impedance of k^{th} branch in original full network. K_i is the total number of branches in the original network which corresponds to inter-group branch i in the reduced network. The underlying meaning of this equation is that it represents the impedance of parallel circuit. We use it to define approximate value of inter-group branches' impedance.

Combining the above two equations, we obtain

$$\begin{bmatrix} I \\ \Lambda \end{bmatrix} * \frac{1}{X_R} = \begin{bmatrix} \left(\sum_{k=1}^{K_i} \frac{1}{x^k}\right)^{-1} \\ 0 \end{bmatrix}, i=1, \dots, N_R \quad (5-28)$$

If we define

$$\Lambda^* = \begin{bmatrix} I \\ \Lambda \end{bmatrix} \quad (5-29)$$

then, we can solve X_R using the least squares method as below

$$\frac{1}{X_R} = [(\Lambda^*)^T * \Lambda^*]^{-1} * (\Lambda^*)^T * \begin{bmatrix} \left(\sum_{k=1}^{K_i} \frac{1}{x^k}\right)^{-1} \\ 0 \end{bmatrix} \quad (5-30)$$

However, there is a challenge in this regard. In the expansion planning study, numerous scenarios and load conditions are considered throughout the planning horizon. As a result, the net power injection at each bus fluctuates over time, which implies that the division matrix used in the reduction process would also vary, and thus Φ_R and Λ will change as well, and therefore does the solution X_R . This is undesirable, since we do not want network parameters to vary with the varying operating conditions imposed by an expansion planning study. To solve this problem, we update the PTDF matrix Φ using the following approach. First, we assume we can represent all PTDF columns within a cluster of buses as the mean of the individual PTDF columns within that cluster; this assumption is reasonable if the similarity of PTDF columns within a cluster is very high (i.e., the variability of PTDF columns within a cluster is very low).³⁵ Then the equation (5-21) becomes

$$\phi_R^{updated} = \pi \phi^{updated} \Gamma \quad (5-31)$$

where $\phi^{updated}$ is the PTDF matrix updated using the above method and $\phi_R^{updated}$ is the updated group level PTDF matrix. Let us use group k as an example to explain why this method works.

$$\pi \phi^{updated} \Gamma_k = \pi [\phi_{k_1}^{updated} \quad \dots \quad \phi_{k_n}^{updated}] \Gamma_k \quad (5-32)$$

where Γ_k is column k of Γ matrix and $\phi_{k_i}^{updated}$ are group k 's update PTDF columns. Γ_k has N elements, and each element represents one bus. If there are k_n buses in group k , then Γ_k elements corresponding to those buses will be nonzero, and the exact value is the ratio of that bus's power injection to the total of the group's power injection. Therefore, the sum of Γ_k is 1. Since $\phi_{k_i}^{updated}$ ($i = 1, \dots, n$) are always the average of group k 's PTDF columns, we can have

$$\sum_{i=1}^{n_k} \phi_{k_i}^{updated} * \Gamma_k^i = \phi_{k_i}^{updated} * \sum_{i=1}^{n_k} \Gamma_k^i = \phi_{k_i}^{updated} * 1 \quad (5-33)$$

where n_k is the number of buses in the cluster k and $\phi_{k_i}^{updated}$ is the mean PTDF columns corresponding to buses of the cluster k . The equation (5-33) proves equation (5-32) is independent from operation state³⁶. Similarly for other groups, hence (5-31) is also independent of the operation state. Then Λ in (5-28) and Λ^* in (5-29) are both independent of the operation state, and thus X_R , which is solved from (5-30), is also independent of the operation state.

We use an example to clarify this concept. Consider a small zone consisting of five buses, where Bus 1 and Bus 2 form Group 1, and Bus 3, Bus 4, and Bus 5 make up Group 2. Assume the net power injections for these buses are [1, 2, 3, -2, 4], respectively. Additionally, there are four external buses [6, 7, 8, 9] with net power injections of [1, 5, -7, -7]. These external buses will be grouped into a single cluster. Then the

$$\phi_R = \pi \phi \Gamma = \pi \phi * \begin{bmatrix} 1/3 & 0 & 0 \\ 2/3 & 0 & 0 \\ 0 & 3/5 & 0 \\ 0 & -2/5 & 0 \\ 0 & 4/5 & 0 \\ 0 & 0 & -1/8 \\ 0 & 0 & -5/8 \\ 0 & 0 & 7/8 \\ 0 & 0 & 7/8 \end{bmatrix}.$$

When operation condition changes, division matrix Γ will change, causing ϕ_R and Λ^* to vary, ultimately leading the X_R to change. By applying the updated method, then ϕ will be changed to $\phi^{updated}$, which only has three types of columns, i.e. $\phi_{k_1}^{updated}$, $\phi_{k_2}^{updated}$, $\phi_{k_3}^{updated}$, corresponding to three clusters. Each column represents the average of all buses' PTDF columns within the same cluster. Given that,

$$\phi_R = \pi * \begin{bmatrix} \phi_{k_1}^{updated} & \phi_{k_1}^{updated} & \phi_{k_2}^{updated} & \phi_{k_2}^{updated} & \phi_{k_2}^{updated} & \phi_{k_3}^{updated} & \phi_{k_3}^{updated} & \phi_{k_3}^{updated} & \phi_{k_3}^{updated} \end{bmatrix}$$

$$* \begin{bmatrix} 1/3 & 0 & 0 \\ 2/3 & 0 & 0 \\ 0 & 3/5 & 0 \\ 0 & -2/5 & 0 \\ 0 & 4/5 & 0 \\ 0 & 0 & -1/8 \\ 0 & 0 & -5/8 \\ 0 & 0 & 7/8 \\ 0 & 0 & 7/8 \end{bmatrix} = \pi * \begin{bmatrix} \phi_{k_1}^{updated} & \phi_{k_2}^{updated} & \phi_{k_3}^{updated} \end{bmatrix},$$

which is independent from division matrix Γ , and therefore, it is independent of the operation state.

5.3 Checking mechanism

The two aggregation methods described above may not necessarily produce optimal results. For topology-based aggregation, it relies on the least squares method, which is a form of approximation. If there are too many internal generator buses to represent as a single bus, the aggregation may lose accuracy. Similarly, quotient graph aggregation may also underperform, as its success depends on clustering hyperparameters, such as the number of clusters or the distance threshold. These parameters must be determined by the user, often based on experience or intuition.

To address this issue, we developed a checking mechanism based on the performance measure OPM. In simple terms, we evaluate the OPM after each additional bus aggregation. If the OPM meets the desired criteria, we accept the aggregation. Otherwise, we either adjust the hyperparameters or abandon the aggregation for that specific zone. The process is illustrated in Figure 5-2. The first half is topology-based aggregation; the second half is quotient graph aggregation.

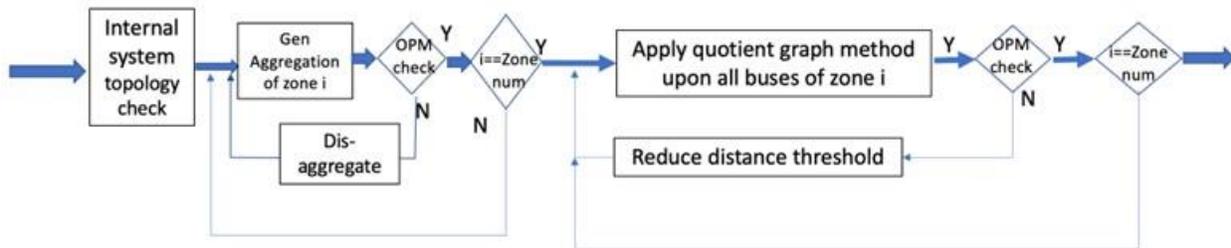


Figure 5-2: Checking mechanism based on OPM

5.4 Case Study

In this section, we present two examples to illustrate the effectiveness of the proposed reduction process.

5.4.1 IEEE 118

The first example is based on the IEEE 118-bus system. This case was employed to compare the effectiveness of rolling simulation versus using the congested branches from the first year in identifying key branches. Once key branches are identified, they are used to reduce the network.

The first diagram below illustrates the comparison between the original system and the reduced system after applying Ward reduction and topology-based aggregation. The top diagram compares key branch flows, while the bottom diagram compares zonal generation. The original case contains 118 buses. After merging the radial paths, 109 buses remain. Following the Ward reduction, the number of buses is reduced to 70. After applying topology-based aggregation, 54 buses remain, with an OPM at this stage of 0.1127.

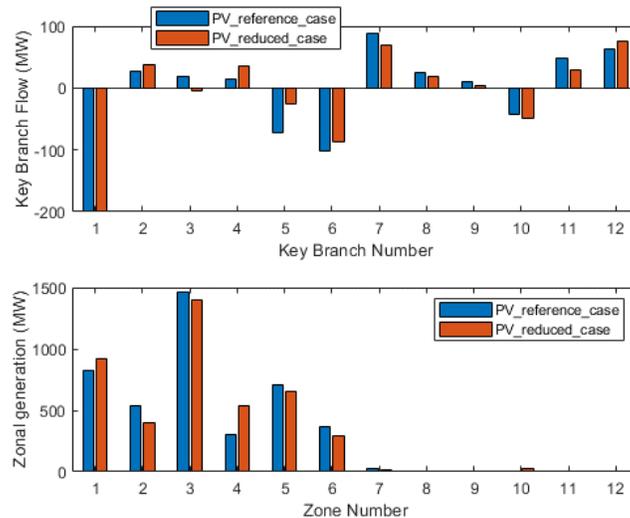


Figure 5-3: Key branch flow and zonal generation comparison between original network and the Ward + topology-based reduced network

The second diagram is similar to the previous one, but it includes the comparison after applying Ward reduction, topology-based aggregation, and quotient graph aggregation. After quotient graph aggregation, there are 18 buses and 54 branches at the end. The OPM at this step is 0.328.

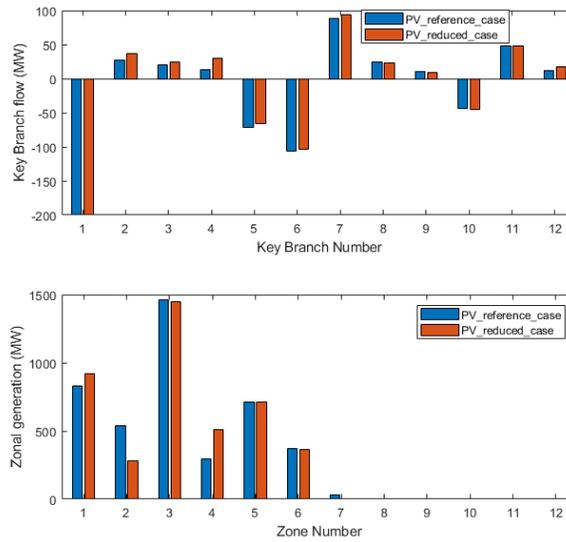


Figure 5-4: Key branch flow and zonal generation comparison between original network and the Ward + topology-based + QG reduced network

5.4.2 RTE Case

The second example involves a case from RTE, starting with 617 buses and 922 branches. Among these, 51 branches were identified as key branches. After merging radial paths, the network was reduced to 376 buses. Ward reduction further decreased this to 195 buses, with an OPM of 0.0004. Following Ward reduction and topology-based aggregation, 145 buses remained, with the OPM increasing to 0.2446. Finally, after quotient graph aggregation, the network was reduced to 76 buses, and the OPM rose to 0.3984. The detailed comparison is included in the below diagrams.

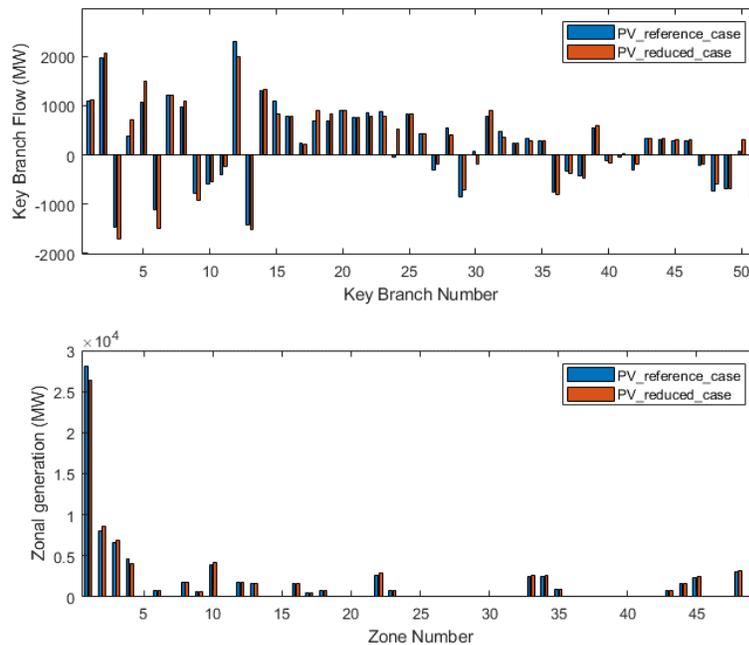


Figure 5-5: Key branch flow and zonal generation comparison between original network and the Ward + topology-based RTE reduced network

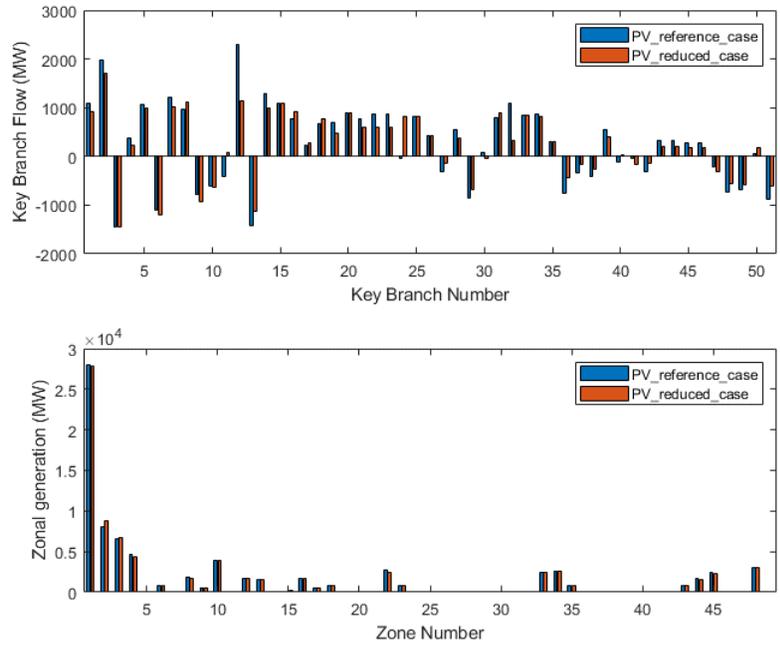


Figure 5-6: Key branch flow and zonal generation comparison between original network and the Ward + topology-based + QG RTE reduced network

6. Treatment of Equivalent Branches

In the process of power system reduction, using Ward’s approach, a set of buses are eliminated. When a bus is eliminated, the connected branches to that bus are also eliminated. New branches can be generated in the process of reduction, representing the eliminated branches. These branches are called “equivalent branches.” From the perspective of a CEP, in its basic form, a branch is defined by its terminal buses, its electrical characteristics such as resistance, reactance, and susceptance, its capacity, and its expansion cost.

The branches in a reduced network are either retained or equivalent. The retained branches are the branches that exist in reduced and full networks. These branches retain all their characteristics from the full network. For a line to be retained, its terminal buses must be retained, and also, for one terminating bus, all buses connected to it (“binding” buses) must be retained. The need to retain binding buses is explained in Section 2.2.2.

For the equivalent branches, the electric characteristics can be obtained from the reduction results. However, the capacity and expansion cost of the equivalent branches are unknown, as the equivalent branches represent more than one branch with potentially different electrical and economic characteristics. To ensure that the reduced model behaves similarly to the full network in the CEP calculations, the capacity and expansion cost of the equivalent branches should represent those of the branches in the full model represented by the equivalent branch.

In this section, new approaches are proposed to calculate the capacity and expansion cost of equivalent branches in the reduced networks. Section 6.1 proposes an approach that calculates the capacities of equivalent branches, and Section 6.2 proposes an approach for calculating expansion cost of equivalent branches.

6.1 Capacity Estimation

A basic formulation for capacity estimation of equivalent branches is provided in Section 6.1.1, and an extended formulation is provided in Section 6.1.2.

6.1.1 Basic Formulation for Capacity Estimation

Internal to CEP, a DCOPF is solved for each operational block to identify the optimal output of generators subject to operational and investment constraints. In the first step toward ensuring the CEP results of the full and reduced network are the same, the DCOPF results of the full and reduced network must match. In this section, we develop a capacity estimation approach based on the objective that DCOPF results in both the full and reduced models should be identical.

A DCOPF calculates the set of bus voltage angles (θ). The difference between the angles of two buses terminating a branch, multiplied by the inverse of the branch’s reactance, gives the flow on the connecting branch between the buses, as in (6-1):

$$f_\ell = \frac{\theta_i - \theta_j}{x_\ell} \quad (6-1)$$

The angle difference between two buses is a function of the operational conditions. Therefore, it is possible to calculate the maximum angular difference between two buses (MAD(i,j)). This value for each bus pair must be the same in the reduced and full model. MAD represents the maximum angular difference, and therefore maximum flow, between two buses. An optimization formulation is devised to find MAD for every bus pair which are terminal buses for equivalent lines (i,j). Once MAD is calculated for each bus pair, the capacity of the equivalent branch between the bus pair is calculated as the maximum flow possible on that equivalent branch. To limit the number of calculations, MAD is only calculated for the bus pairs that an equivalent line exists between them in the reduced network.

The devised optimization problem that calculates the capacity of equivalent branches is shown below. In this formulation, the MAD for each bus pair with equivalent branches are calculated. In the objective function, the MAD is divided by the reactance of the equivalent branch and therefore maximum possible flow, i.e., capacity, of the equivalent branch is calculated.

The following comments are applicable to the proposed optimization problem, given by equations (6-2)-(6-5):

- 1) Calculation of capacity for equivalent branches are independent of each other, and therefore the below formulation can be solved independently for all equivalent branches; this enables these calculations, one for each equivalent branch, to be performed in parallel.
- 2) MAD is calculated on the full network, i.e., for each bus pair that have an equivalent branch between them in the reduced network, MAD is calculated in the full network. The implication is that this approach cannot estimate capacity of a branch terminated by one or two buses not existing in the original network; we address this situation in Section 6.1.2.
- 3) The calculation of MAD is independent of operational condition of the full network. MAD for each bus pair represents maximum angular difference under any operational conditions. Therefore, the goal of the optimization problem is to identify an operational condition under which the angular difference between two bus pairs is maximized.

$$\max_{\theta} \quad S_{base} \cdot \frac{\theta_k - \theta_p}{x_{kp}} \quad (6-2)$$

s.t.

$$f_l = S_{base} \cdot \frac{\theta_i - \theta_j}{x_l}, \quad \forall l(i,j) \in L \quad (6-3)$$

$$f_l^{min} \leq f_l \leq f_l^{max} \quad \forall l(i,j) \in L \quad (6-4)$$

$$P_i - \sum_j f_{i,j} + \sum_j f_{j,i} = L_i \quad \forall i \in B \quad (6-5)$$

We illustrate the method using a 4-bus model, shown in Figure 6-1.

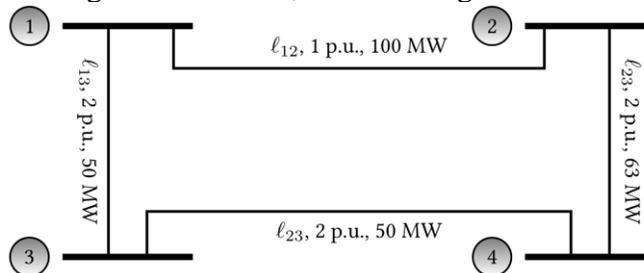


Figure 6-1: Sample four-bus system

By eliminating bus 3, the reduced network is as shown in Figure 6-2. The dashed line indicates the resulting equivalent branch; it has a reactance of 4 p.u., calculated using Ward reduction, but its capacity is unknown (indicated by a question mark in Figure 6-2).

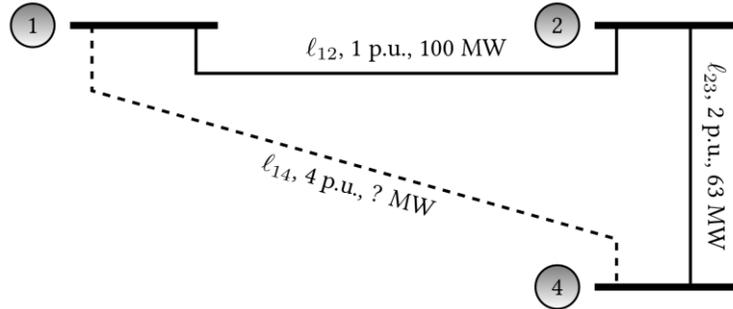


Figure 6-2: Reduced four-bus system

The lines between buses 1 and 2 and between 2 and 4 are retained and therefore, their capacity stays the same as in the full network. However, the capacity of the newly formed equivalent branch between buses 1 and 4 is unknown. Intuitively, based on the original system shown in Figure 6-2, one could conclude that the capacity of the equivalent branch is 50 MW.

We now compute the MAD for bus pair (1,4) by applying the optimization problem, resulting in $MAD(1,4)=2.0$ radians. This angular difference is realized under the operational conditions depicted in Figure 6-3.

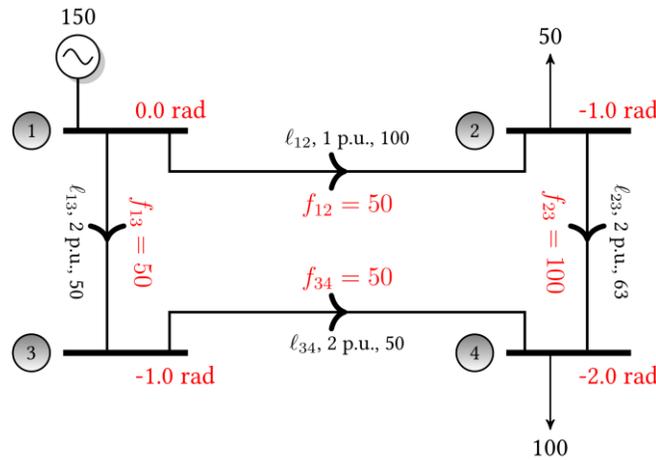


Figure 6-3: Operational conditions for maximizing angular difference of buses 1 and 4

Dividing the $MAD(1,4)$ by the reactance of the equivalent branch between buses 1 and 4 in the reduced network (4 p.u.), the maximum flow (capacity) of the equivalent branch is calculated:

$$f_{1,4}^{max} = \frac{2.0}{4.0} \times 100 = 50 \text{ MW} \quad (6-6)$$

In the above calculation, 100 is the S base of the p.u. calculations. Using the calculated capacity of the equivalent line, the OPF results of the full and reduced network match.

Applying the proposed approach to the 118-bus IEEE system and to the 617-bus RTE system shows that the OPF results of the reduced model match OPF results of the full model, as indicated

in Table 6-1 (recall from Section 1.4.2.1 that the smaller the OPM value, the better the agreement between full and reduced models).

The calculation of MAD was applied to the Eastern Interconnection (EI) network to test the computational burden of the formulation. The results show that the optimization problem can be solved within a few minutes for each bus pair. Moreover, by parallelizing the calculation of MAD for each bus pair, the calculation of thousands of branch capacities for the EI can be completed in less than an hour.

Table 6-1: Case studies showing effects of capacity estimation on OPM after Ward reduction

	IEEE 118	RTE 617 buses	MISO [first operating condition, 2024]
Without capacity evaluation	2.2378e-14	0.0132	0.0062
With capacity evaluation	1.8864e-11	0.0004	0.0062

6.1.2 Extended Formulation for Capacity Estimation

The basic formulation of Section 6.1.1 is effective for models reduced using only the Ward elimination method, but it does not apply to reduced systems with aggregated buses. The reason why it does not apply to reduced systems with aggregated buses is because the formulation of Section 6.1.1 requires that both terminal buses exist in the full network (as indicated by comment #2 in Section 6.1.1). For reduced systems having aggregated buses, those aggregated buses are new buses that did not exist in the original network. To address this issue, we developed an extended formulation for capacity estimation.

Following the approach outlined in the previous section, we consider a new branch with terminals k and p , for which we desire to estimate the capacity. However, here we assume that one or both of these terminals do not exist in the full model of the network. Instead, here, for purposes of illustration, we assume bus k is an aggregation of buses 1, 2, and 3, while bus p is an aggregation of buses 4, 5, and 6. The key principle in the basic capacity evaluation is to ensure that the maximum angle difference limit between bus k and bus p remains consistent between the original network and the reduced case. Given that the new buses are aggregated from several existing ones, the goal is to maintain the angle difference condition between the aggregated groups of buses 1, 2, 3, and buses 4, 5, 6. To achieve this, we use the average angles of the involved buses to represent the newly aggregated bus. Consequently, the extended formulation for capacity estimation can be constructed as follows:

$$\max_{\theta} \quad S_{base} \cdot \frac{\frac{\sum_{i=1}^{n_k} \theta_{k_i}}{n_k} - \frac{\sum_{i=1}^{n_p} \theta_{p_i}}{n_p}}{x_{kp}} \quad (6-7)$$

s.t.

$$f_l = S_{base} \cdot \frac{\theta_i - \theta_j}{x_l}, \quad \forall l(i, j) \in L \quad (6-8)$$

$$f_l^{min} \leq f_l \leq f_l^{max} \quad \forall l(i, j) \in L \quad (6-9)$$

$$P_i - \sum_j f_{i,j} + \sum_j f_{j,i} = L_i \quad \forall i \in B \quad (6-10)$$

In the objective function of (6-7), the term $\frac{\sum_{i=1}^{n_k} \theta_{k_i}}{n_k}$ represents the average angle of buses aggregated into bus k , and the term $\frac{\sum_{i=1}^{n_p} \theta_{p_i}}{n_p}$ represents the average angle of buses aggregated into bus p . Applying this extended approach to the 118-bus IEEE system and to the 617-bus RTE system shows that ...???... as indicated in Table 6-1 (recall from Section 1.4.2.1 that the smaller the OPM value, the better the agreement between full and reduced models).

Table 6-2: Case studies showing effects of capacity estimation on OPM after Ward reduction and aggregation

	IEEE 118	RTE (617 bus)
Without capacity evaluation	0.337	0.3381
With capacity evaluation	0.216	0.3381

6.2 Expansion Cost Assignment

The expansion cost of branches affects the CEP in that expansion of low-cost branches is favored over expansion of high-cost branches, ultimately influencing the flows used to transfer power from generation to load. Considering our goal of minimizing differences between CEP results of the full and reduced network, the CEP should perceive an equal cost for the same path in the full and reduced network. However, in the reduced network, some buses are eliminated, new paths are formed, and some paths are removed. In this section, we propose two formulations to assign expansion costs to the equivalent branches such that the CEP results are the same or very close for the full/reduced network. Section 6.2.1 identifies Formulation 1 for cost assignment. Section 6.2.2 identifies Formulation 2 for cost assignment.

6.2.1 Formulation 1 for cost assignment

In this subsection, we provide a “basic formulation” of one approach used for cost assignment. This formulation for expansion cost assignment involves calculating several characteristics in both full and reduced networks and solving sets of linear equations repetitively. One of these characteristics is transfer capacity.

Transfer Capacity: Transfer capacity between buses p and k is defined as the maximum power that can be injected in bus p and withdrawn in bus k without violating any branch capacity.

The following formulation calculates the transfer capacity of the bus pair (k, p) .

$$\max \quad P_k \quad (6-11)$$

s.t.

$$f_l = S_{base} \cdot \frac{\theta_i - \theta_j}{x_l}, \quad \forall (i, j) \in L \quad (6-12)$$

$$f_l^{min} \leq f_l \leq f_l^{max} \quad \forall (i, j) \in L \quad (6-13)$$

$$P_i - \sum_j f_{i,j} + \sum_j f_{j,i} = L_i \quad \forall i \in B \quad (6-14)$$

$$P_i = 0 \quad \forall i \in B \quad i \neq k \quad (6-15)$$

$$L_j = 0 \quad \forall j \in B \quad j \neq p \quad (6-16)$$

In this formulation, the injection and withdrawal (generation and load) are limited to the terminal buses of equivalent branches. Figure 6-4 depicts the concept of transfer capacity.

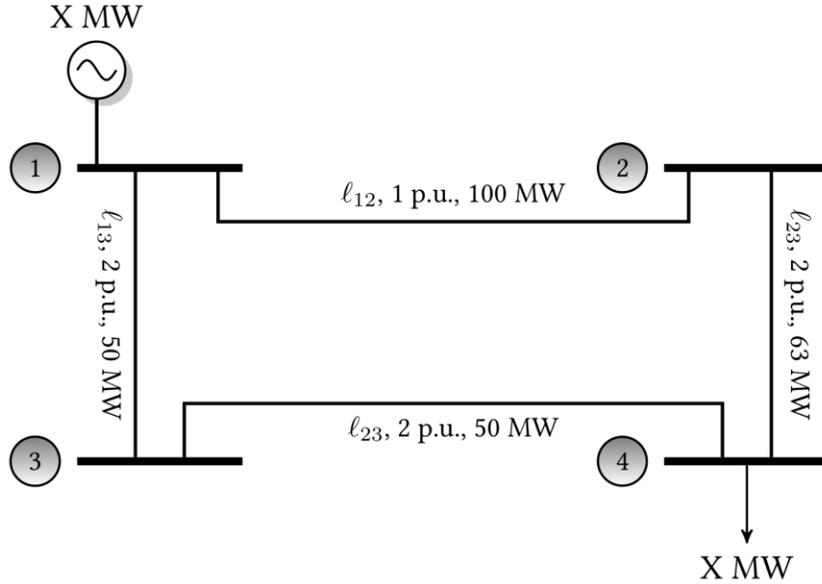


Figure 6-4: Depicting transfer capacity in a small test system

The previously developed capacity estimation approach focuses on ensuring the results of OPF for the full and reduced network are the same. However, this approach does not retain the transfer capacities between the bus pairs, resulting in inaccurate cost assignment. Therefore, we propose another capacity calculation approach to preserve bus pairs' transfer capacity. This approach is exclusively applied to the cost assignment and is not used to estimate the capacity of equivalent lines in the ACEP.

$$\max_{\theta} \quad S_{base} \cdot \frac{\theta_k - \theta_p}{x_{kp}} \quad (6-25)$$

s.t.

$$f_l = S_{base} \cdot \frac{\theta_i - \theta_j}{x_l}, \quad \forall (i, j) \in L \quad (6-26)$$

$$f_l^{min} \leq f_l \leq f_l^{max} \quad \forall (i, j) \in L \quad (6-27)$$

$$P_i - \sum_j f_{i,j} + \sum_j f_{j,i} = L_i \quad \forall i \in B \quad (6-28)$$

$$L_j = 0 \quad \forall j \in B \quad j \neq p \quad (6-29)$$

$$L_p = P_k \quad (6-30)$$

This differentiating aspect of this formulation is that the injection and withdrawals are limited to the bus pairs under the study.

Cost Assignment Philosophy:

The philosophy behind the cost assignment approach is to ensure that the cost for increasing the transfer capacity between buses p and k in the full and reduced network is the same.

We assume the transfer capacity of the bus pair (p,k) in the full network is TC_{pk}^{full} MW. We recall that the transfer capacity was calculated by maximizing the injection and withdrawal levels at buses p and k , respectively. Now, the injection amount is increased by 1 MW, that is $TC_{pk}^{full} + 1$ MW. This new amount of power is injected and withdrawn in buses p and k , respectively. Since this amount exceeds the transfer capacity, at least one branch must be expanded to accommodate the extra flow. A transmission expansion planning (TEP) problem can be solved to minimize the cost of transmission expansion while increasing the transfer capacity by 1 MW. The formulation for the TEP is as follows:

$$\min \sum_l f_l^{inv} C_l \quad (6-17)$$

s.t.

$$f_l = S_{base} \cdot \frac{\theta_i - \theta_j}{x_l}, \quad \forall l(i,j) \in L \quad (6-18)$$

$$-f_l^{inv} - f_l^{max} \leq f_l \leq f_l^{max} + f_l^{inv} \quad \forall l(i,j) \in L \quad (6-19)$$

$$P_i - \sum_j f_{i,j} + \sum_j f_{j,i} = L_i \quad \forall i \in B \quad (6-20)$$

$$P_i = 0 \quad \forall i \in B \quad i \neq k \quad (6-21)$$

$$P_k = TC + 1 \quad (6-22)$$

$$L_j = 0 \quad \forall j \in B \quad j \neq p \quad (6-23)$$

$$L_p = P_k \quad (6-24)$$

In the objective function, f_l^{inv} and C_l are the investment capacity and expansion cost of line l , respectively. In the constraints, f_l is the flow of line l . The angles of buses i and j are represented by θ_i and θ_j . S_{base} is the MVA base for conversion to per unit, which is assumed to be 100 MW. The branch ratings are indicated by f_l^{max} . Variables P and L represent the power injections and withdrawals at each bus indicated by their subscript.

To relate the cost of transmission expansion for the full and reduced network, the same injection amounts are applied to buses p and k in the reduced network. Since the injection level is more than the transfer capacity by 1 MW, at least one branch and possibly more are carrying flows over their capacities. Knowing the injection and withdrawal levels and the line reactances, we can calculate the flow of each branch using DC power flow equations. Moreover, considering that we have the total cost of expansion of branches from the TEP of the full network for this level of transfer capacity, we can form a set of linear equations by multiplying the excess flow of each branch and their expansion cost.

If the branch with the excess power is a retained branch, its expansion cost is known. However, if the branch is an equivalent branch, the expansion cost is unknown. We form a linear equation representing the expansion cost of the equivalent branch as an unknown variable. The right-hand side of the equation is the total expansion cost obtained from the full network TEP. The equation can be solved if only one equivalent line is overloaded in this iteration. The process is iterated by increasing the transfer capacity more and forming linear equations until all equivalent branches are represented in the set of linear equations. If there are N equivalent branches, we form N equations. Figure 6-5 shows a flowchart of the cost assignment process.

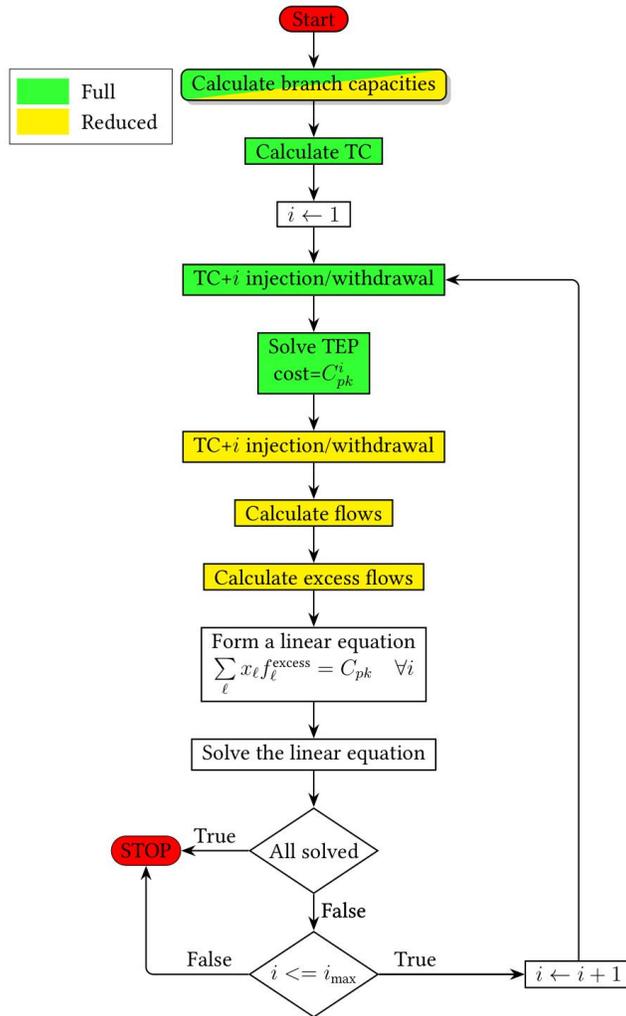


Figure 6-5: Expansion cost estimation process

To demonstrate the effectiveness of this approach, it is applied to the small test system shown in at the left of Figure 6-6; the reduced network is shown to the right. The branch data given is reactance (p.u.), capacity (MW), and expansion cost (\$/MW).

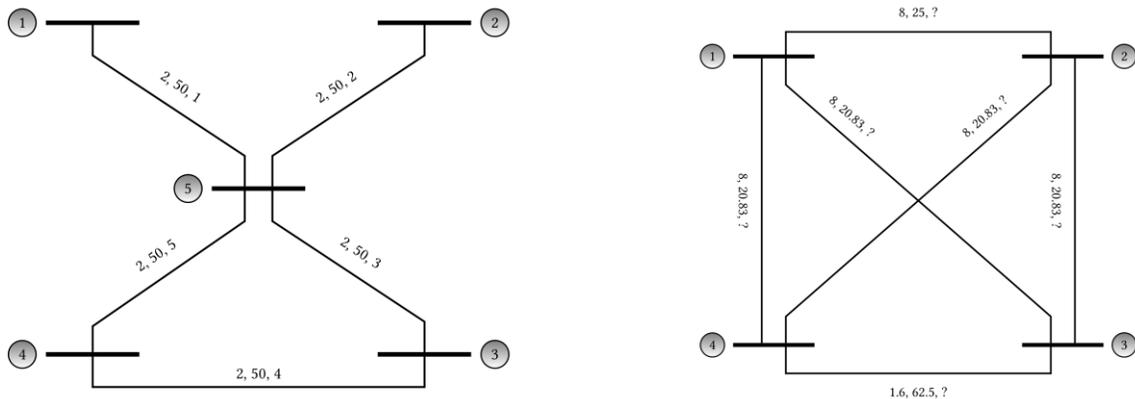


Figure 6-6: Five-bus test system, full model (left) and reduced model (right)

The reduced network does not have any retained branches, so all branches in the reduced model are equivalent. Therefore, we need to calculate capacity and expansion cost for all reduced model branches. We have calculated the capacities of the equivalent branches as shown in the reduced network in Figure 6-6. The transfer capacities between pairs of buses in the full network are shown in Table 6-3.

Table 6-3: Transfer capacity in the full 5-bus and reduced network

From	To	TC (MW)
1	2	50
1	3	50
1	4	50
2	3	50
2	4	50
3	4	75

The expansion costs of the equivalent branches are calculated using the proposed approach. Table 6-4 shows the results of cost assignment for this reduced network.

Table 6-4: Expansion cost of equivalent branches (\$/MW)

From	To	Cost (\$/MW)
1	2	3
1	3	1
1	4	1
2	3	2
2	4	2
3	4	2

6.2.2 Formulation 2 for cost assignment

Formulation 1 is effective when applied to a Ward-reduced network, where nodes are eliminated, and all remaining nodes in the reduced network exist in the full network. However, when performing aggregation (as described in Chapter 5), new nodes are *created* from the aggregation process (these nodes do not exist in the full network), and so the cost assignment approach of Formulation 1, described in Section 6.2.1, is not applicable. We describe a different approach in this section, developed explicitly for the situation where one or both terminating buses are aggregations of a group of constituent buses.

Consider a new branch with terminals k and p for which we need to assign an expansion cost. However, these terminals do not exist in the original network; instead, bus k is an aggregation of buses 1, 2, and 3, while bus p is an aggregation of buses 4, 5, and 6. The key principle we use here is that the cost of one additional MW of injection and withdrawal between buses k and bus p , respectively, remains consistent between the original network and the reduced case.

We assume that the power injection is evenly distributed among all the aggregated buses. In the example above, this means buses 1, 2, and 3 each inject 1/3 MW, and buses 4, 5, and 6 each receive 1/3 MW. The investment cost for equivalent branch $k-p$ is expressed using the following equation:

$$P_{kp} = P^{ori} * PTDF * Inj_{kp} \quad (6-25)$$

where P^{ori} is a row vector corresponding to investment cost of branches in the original network, $PTDF$ is original network's PTDF matrix, and Inj_{kp} is the injection column corresponding to power injection at buses k and p . For example, if bus k is an aggregation of buses 1, 2, and 3, while

bus p is aggregated by buses 4, 5, and 6, then the $Inj_{kp} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \\ -1/3 \\ -1/3 \\ -1/3 \\ \vdots \end{bmatrix}$. Assume the original network

has N buses and N_r branches, then P^{ori} is a row vector with N_r elements, $PTDF$ is an $N_r \times N$ matrix, Inj_{kp} is a column vector with N elements, and P_{kp} is the desired (scalar) expansion cost of branch $k-p$. This calculation is made for each equivalent branch in the network. It results in a single expansion cost value for each of these equivalent branches.

Results from using this method is provided in Chapter 7.

6.2.3 Dual variable method for cost assignment

The previous two sections focus on evaluating the investment cost of equivalent branches within the reduced network. These equivalent branches are obviously not key branches (key branches are retained and so the expansion cost for key branches is known from the full model). Our selection of key branches is made with the intent that the selection includes all branches in which investment is made. And so it is expected that equivalent branches should not see much investment. Nonetheless, it is typical that we find some investment in equivalent branches, and the methods described in the previous two sections provide ways to account for this, albeit in an approximate fashion.

In this section, we provide an alternative approach that, although it focuses on key branch investment, also addresses the potential for investment in other branches. To illustrate, consider the basic capacity evaluation formulation: if branch flow limits are relaxed as one key branch is expanded, it is likely that other branches' capacities will need to increase as well. Since this increase is not intentional but results from expanding key branches, we refer to it as supplementary expansion. Dual variables in the basic formulation of capacity evaluation, i.e., equations (6-2) to (6-5), can be used to quantify the extent of this supplementary expansion.

We formulate this approach as follows. Assume there are N_{key} key branches, and the expansion on key branch j is ΔF_j , while the supplementary expansion on another branch n caused by ΔF_j is Δb_n^j . The dual variable used to quantify the impact of ΔF_j on Δb_n is:

$$\lambda_n^j = \frac{\Delta b_n^j}{\Delta F_j} \quad (6-26)$$

Then the total amount of supplementary expansion on branch n can be described as

$$\Delta b_n = \sum_{j=1}^{N_{key}} \Delta b_n^j = \sum_{j=1}^{N_{key}} \lambda_n^j * \Delta F_j \quad (6-27)$$

where expansion on key branches, i.e ΔF_j , are independent variables in expansion planning, and expansion on other branches, i.e., Δb_n , depend on them.

After aggregation, the newly aggregated branches will inherit the total supplementary expansion from the branches that were combined to form them. For instance, if branch $k-p$ is newly aggregated, where bus k is formed by aggregating buses 1, 2, and 3, and bus p is formed by aggregating buses 4, 5, and 6, then the supplementary expansion of branch $k-p$ will be the sum of the supplementary expansions of all branches that connect any bus from group 1, 2, 3 to any bus from group 4, 5, 6. The dual variable method differs from the previous two methods, since it is not a cost assignment method. Instead, it is used to address certain side effects on equivalent branches that arise after investing in key branches. The reason we use this approach is due to the application of capacity evaluation for equivalent branches. Both the second and the third method are employed to deal with equivalent branches during the CEP step.

As a final note in this section, we observe that the dual variable method for cost assignment does not require that equivalent branches have estimated capacities, i.e., it can be applied without performing either of the capacity estimation methods of Sections 6.1.1 or 6.1.2. We have found, however, that performing equivalent branch capacity estimation before applying this method typically results in a higher fidelity reduced model, a feature attributed to the tendency of capacitated equivalent branches to yield more accurate flows than they would if they were uncapacitated. Results from using this method is provided in Chapter 7.

6.3 Conclusion

In this section, we proposed three approaches for estimating capacity of equivalent lines and for assigning expansion cost to equivalent lines. Results for the first approach were provided in this section. Results for the second two approaches are provided in Chapter 7.

7. Expansion Planning with Translation

Expansion planning for electric power system infrastructure is the problem of identifying investments over a decision horizon, typically between 5 and 25 years, to minimize the net present value of the total investment and operational costs over that decision horizon, subject to network, operational, environmental, and investment constraints. As indicated in Section 1.1, expansion planning problems may be formulated to identify generation investments only, transmission investments only, or both. There have been formulations to also include investments at the distribution level, motivated by the growth of distributed energy resources, but doing so typically requires highly simplified distribution system representation to avoid computational intractability [37]. In this work, we have focused on expansion planning applications that identify generation and transmission investments³⁸ to which we refer as coordinated expansion planning (CEP) if it is deterministic and adaptive coordinated expansion planning (ACEP) if it treats uncertainty in future conditions.

In this chapter, Section 7.1 provides high-level formulations of our expansion planning problems, including one formulation for CEP and one formulation for ACEP. Section 7.2 summarizes practical features of the solution we apply to these problems. Section 7.3 describes the need to translate expansion planning results from the reduced model on which they are obtained to the full model to correspond to the investments that will actually occur in the field. Section 7.4 concludes this chapter.

7.1 Expansion Planning Formulations

The expansion planning formulations provided in this section are communicated at a level to facilitate conceptual understanding. Rigorous analytic formulations may be found for CEP in [37] and for ACEP in [39, 40]. The essential difference in these two formulations lies in the treatment of the future conditions, specified as a set of parameters. Such parameters are typically load growth, technology build costs, fuel prices, environmental constraints or costs (e.g., on carbon emissions), operational performance (e.g., wind and solar capacity factors), and policies (e.g., presence and level of renewable portfolio standards). The (deterministic) CEP assumes only a single value for each parameter and therefore considers only one future. The (stochastic) ACEP assumes two or more possible values for each of these parameters. If there are, for example, 10 uncertain parameters each of which may take either of two values, then there are $2^{10}=1024$ possible futures; if the 10 parameters may each take one of three possible values, then there are $3^{10}=59,049$ possible futures.

7.1.1 Coordinated Expansion Planning (CEP)

Using “NPV” to denote “net present value,” the CEP problem is formulated as follows:

Minimize:

$$\sum_{t=1,N} \text{NPV} \{ \text{InvCosts}(\Delta \underline{x}(t)) + \{ \text{OpCost}(\underline{x}(t)) \} \}$$

(7-1)

Subject to
constraints on network, operations, environment, investments

Here, t is the time period, N is the number of time periods, $\Delta\mathbf{x}(t)$ is the vector of generation and transmission investments made during year t (all of which is assumed to occur at the beginning of year t), and $\mathbf{x}(t)$ is the vector of generation and transmission available in year t , which is related to the year t investments according to

$$\mathbf{x}(t) = \mathbf{x}_0(t) + \Delta\mathbf{x}(t) \quad (7-2)$$

where $\mathbf{x}_0(t)$ represents the generation and transmission infrastructure existing at the beginning of year t but before any year t investments are made.

7.1.2 Adaptive Coordinated Expansion Planning (ACEP)

Again, using “NPV” to denote “net present value,” the ACEP problem is formulated as follows:

Minimize:

$$\text{NPV} \left\{ \sum_{t=1, N} \left\{ \text{InvCosts}(\Delta\mathbf{x}_c(t)) + \beta \times \sum_{k=1, K} \left\{ \text{Pr}_k \times \text{AdaptationCost}(\Delta\mathbf{x}_k(t)) + \text{OpCost}(\mathbf{x}(t)) \right\} \right\} \right\} \quad (7-3)$$

Subject to, for each future k
constraints on network, operations, environment, investments

Here, once again, t is the time period and N is the number of time periods. However, in (7-3), we represent K futures, and investments have been split into two types: the core investments $\Delta\mathbf{x}_c(t)$ are made for all futures, but the adaptations $\Delta\mathbf{x}_k(t)$ are only made for future k . Thus, the investments made during year t may be core or adaptations, so that $\mathbf{x}(t)$, the vector of generation and transmission available in year t , is related to the year t investments according to

$$\mathbf{x}(t) = \mathbf{x}_0(t) + \Delta\mathbf{x}_c(t) + \Delta\mathbf{x}_k(t) \quad (7-4)$$

where $\mathbf{x}_0(t)$ once again represents the generation and transmission infrastructure existing at the beginning of year t but before any year t investments are made. We provide four comments pertaining to (7-3):

- *Number of futures:* Problem (7-3) becomes intractable if too many futures are used. It is typical to choose between 5 and 15 futures, depending on the network size, the number of time periods represented, and the hardware deployed. The choice of futures is important; there are various ways to identify a small set of futures that most effectively represents the entire space.
- *Probabilities:* The Pr_k represent the probability of occurrence for future k .
- *Relation to traditional stochastic programming:* Problem (7-3) differs from the traditional expansion planning formulation of a stochastic program (TSP) such as those developed in [41, 42], in that the “here and now” investments are a trajectory through time – the core. The “wait and see” investments are the adaptations made to the core at each time period.
- *Significance of β :* The parameter β in (7-3) is called the robustness parameter. Its value is user-specified according to the desired level of robustness. If β is small, then adaptation costs appear inexpensive to the optimizer, and it invests very little in the core; in this case, the core is cheap but not very robust. If β is large (e.g., 2.0) then adaptation costs appear very expensive to the

optimizer, and it invests mostly in the core and depends very little on the adaptations; in this case, the core is expensive and highly robust, i.e., it will accommodate almost all the futures without adaptation. It is useful for a planner to inspect different investment solutions for a wide range of robustness levels.

Figure 7-1 illustrates two different ACEP solutions, one for a low value of β (left) and one for a large value of β (right).

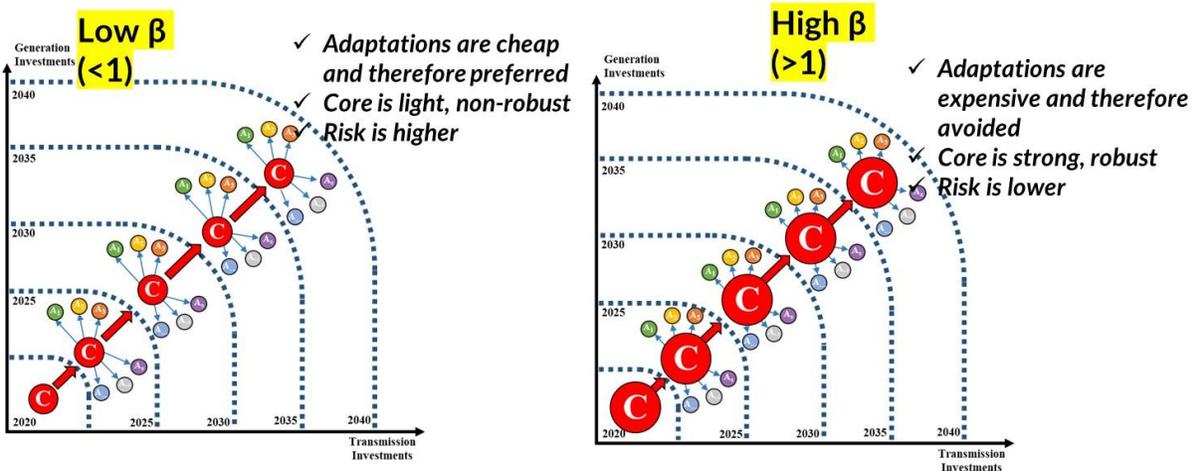


Figure 7-1: Illustration of the effect of β on the investment solution

7.2 Solution Features

For industry-sized network models, the optimization problems described in Sections 7.1.1 and 7.1.2 are very high-dimensional and typically intractable. This is the motivation for this project, to obtain a reduced network for which the CEP and ACEP problems are tractable. However, solving these optimization problems even for reduced models is challenging, and significant user-expertise must be employed. We describe the most important of these decisions in this section.

7.2.1 Modeling Environment, Solver, and Optimization Method

There are several modeling environments for solving developing optimization models, but the most common are the *Advanced Interactive Multidimensional Modeling System* (AIMMS), *A Mathematical Programming Language* (AMPL), and the *General Algebraic Modeling System* (GAMS). We have used GAMS in the work described in this report.

Like the other modeling systems, GAMS supports the development of the model one wishes to solve, and it does so for various types of optimization problems, e.g., linear, mixed-integer linear, nonlinear, mixed-integer nonlinear and others, but it does not itself perform the solution. For that, it calls any one of solvers to which it interfaces and that is capable of handling the specific type of optimization problem modeled. A list of 38 solvers to which GAMS interfaces is available [43] and includes, for example, CPLEX, Gurobi, LINDO, MOSEK, and XPRESS. We generally use Gurobi.

Most solvers offer various optimization methods for a given problem type. In this project, all of our problem types were linear programs (LPs). Gurobi, like many other solvers, offers primal simplex, dual simplex, and parallel barrier. Our problems are generally solved using the parallel barrier method followed by simplex.

7.2.2 Scaling

ACEP models are at crossroad of electrical and economic data. Therefore, the scale of the input data to the ACEP models can vary significantly depending on the units that are used. For example, the reactances of reduced network can vary from 10 p.u. to 0.001 p.u. Moreover, the generation and transmission investment cost and generators variable and fixed operation and maintenance costs can vary via several orders of magnitude. These differences make solving the resulting model more challenging [44]. Therefore, one of the techniques is to scale the variables and equations such that the difference between the values in the model are not significant [45]. Each model has a coefficient matrix which shows the range of the multipliers of variables in the constraints, right-hand side values, and variables bounds. As a general guideline, the ratio of the maximum to minimum values of the coefficient matrix should not exceed 10^6 .

To ensure the solver does not face any numerical issues, we scale constraints and variables to reduce the differences in the coefficient matrix. The process is that the corresponding constraints to the small coefficients of the coefficient matrix are identified and scaled. If the scaling improves the coefficient matrix, the scaling is imposed on the model.

Solvers such as Gurobi will auto-scale the model to improve the coefficient matrix; however, it is always preferred that the user scales the constraints and variables given the intimate knowledge of the model by the user. Moreover, solvers offer parameters to be tuned to tell the solver that this model may have numerical issues. These parameters tell the solver to be cautious of numerical issues when it comes to rounding, multiplying, and dividing. We have set these parameters to ensure the model is cognizant of potential numerical issues.

7.2.3 Hardware

The models are run on servers located at the Department of Electrical and Computer Engineering at Iowa State University. These servers are running Linux RedHat as their operating system with 40 CPUs (Intel(R) Xeon(R) Gold 6354 CPU @ 3.00GHz) with 256 GB of memory.

7.2.4 Results before translation

We present results comparing CEP performance for full and corresponding reduced cases. The full cases are the IEEE 118-bus system and the 617-bus RTE system. CEP performances for their reduced versions are compared with the full cases using CEP performance metric (CPM - as described in Section 1.4.2.2). Results for the PTDF method (described in Section 6.2.2) are provided in Table 7-1; results for the dual variable method (described in Section 6.2.3) are provided in Table 7-2. In these two tables, the CPM is shown for the reduced case after Ward reduction, after Ward and topology-based (TB) aggregation, and after Ward, TB, and quotient-graph (QG) aggregation. Recalling that the CPM is the ratio of CEP reduced model objective

function value to CEP full model objective function value (and therefore CPM=1 indicates best performance), we observe that the performance of the PTDF and the DV methods are reasonably close; we consider both to be acceptable.

Table 7-1: CPM results before translation for IEEE 118-bus network

	PTDF	DV
Ward (70 buses)	0.9781	1.2287
Ward+TB (54 buses)	0.9419	0.9423
Ward+TB+QG (18 buses)	0.8640	0.8640

Table 7-2: CPM results before translation for RTE 672-bus network

	PTDF	DV
Ward (195 buses)	1.0005	1.0207
Ward+TB (145 buses)	1.0479	1.2487
Ward+TB+QG (76 buses)	0.9156	0.9630

7.3 Translation

The previous two sections identified expansion planning formulations and its solution features. These formulations are intended for application on a reduced network. Once expansion planning is completed on the reduced network, the results need to be transferred back to the original full network. This process is referred to as *translation*, which is described in this section.

The translation process requires two steps.

Step 1, generation translation: This step involves transferring expanded generation from the reduced network to the full network. There are two cases involving the bus representation in the reduced network and the bus representation in the full network. In the one-to-one case, the reduced model generation expansion is performed on a bus that exists in the full network; in this case, the expansion is identical in both networks. However, in the one-to-many case, the reduced model generation expansion is performed on a bus that represents several buses in the full model. This case arises because of the aggregation step (described in Chapter 5). If generation is expanded at aggregated buses, we must then decide how to allocate the expanded generation at the one bus in the reduced model to the many buses in the full model. This is done using heuristic rules; typical such rules include:

1. Allocate renewable units to buses in proportion to the buses locational capacity factor (CF), a step which serves as a proxy to generation siting based on economics.
2. Ensure expanded capacity does not exceed the bus $N-1$ contingency limit (if there are N branches connected to a bus, then bus generation capacity should not exceed the sum of capacity on the $N-1$ least capacity branches)
3. Individual generating units are retired in the full model based on heat rate until the reduced model retirement level is reached.
4. Reduced model thermal expansions are added to full model buses where full model thermal plants were retired.

Step 2, transmission translation: A fast transmission expansion planning (TEP) optimization is applied to the full model with the generation expansion represented as described in Step 1. This

optimization is non-linear, given each transmission investment changes the circuit capacity and the circuit reactance. To address this, the TEP performs a sequence of linear programs (LPs), where each LP minimizes the total transmission investment cost (subject to DC power flow equations), and only circuit capacity is treated as a decision variable, while circuit reactance is held constant. In spite of the large network, each LP is fast because it is for transmission only (generation expansion has already been performed), and because only a single investment period is allowed (the final year of the decision horizon). Following the LP solution, the reactance of each invested circuit is updated to reflect the change in capacity, after which the LP is rerun. The iterations are terminated when the circuit with the largest change in capacity relative to the previous iteration is within a specified tolerance. This step 2 of the translation process is illustrated in Figure 7-2.

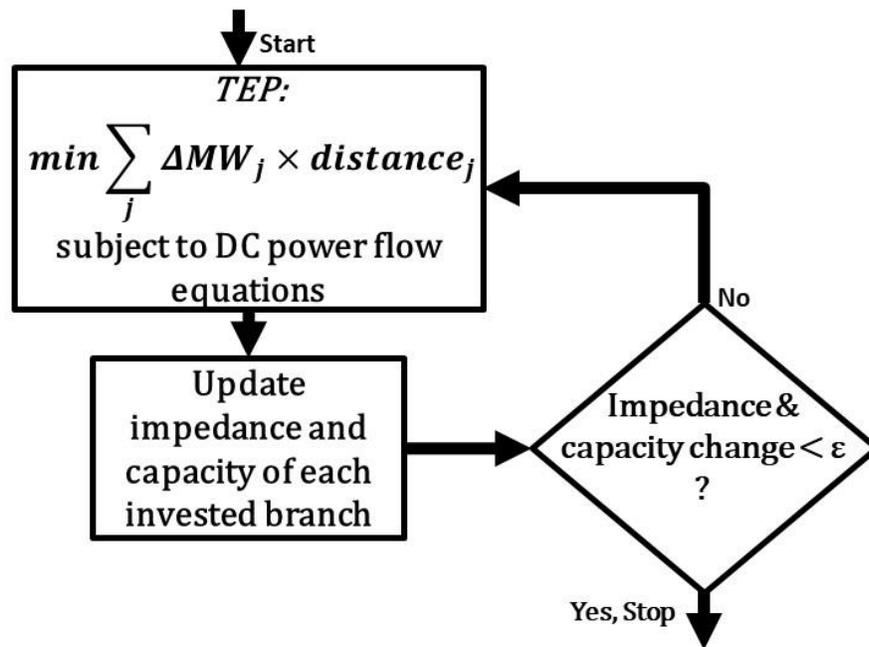


Figure 7-2: Step 2 of the translation process

In what follows, we repeat the analyses of Table 7-1 and Table 7-2 in Section 7.2.4, comparing CEP performance for full and corresponding reduced cases, except here, the analyses is performed on the translated model. As before, the full cases are the IEEE 118-bus system and the 617-bus RTE system. CEP performances for their *translated* versions are compared with the full cases using CEP performance metric (CPM - as described in Section 1.4.2.2). Results for the PTDF method (described in Section 6.2.2) are provided in Table 7-3; results for the dual variable method (described in Section 6.2.3) are provided in Table 7-4. In these two tables, the CPM is shown for the translated models after Ward reduction, after Ward and topology-based (TB) aggregation, and after Ward, TB, and quotient-graph (QG) aggregation. Recalling that the CPM is the ratio of CEP reduced model objective function value to CEP full model objective function value (and therefore CPM=1 indicates best performance), we observe that the performance of the PTDF and the DV methods are not as good as the corresponding results before translation, an observation that results

from comparing Table 7-1 to Table 7-3 and Table 7-2 to Table 7-4. This is expected since translation involves heuristics and an approximate TEP.

Table 7-3: CPM results after translation for IEEE 118-bus network

	PTDF, translation	DV, translation
Ward (70 bus)	1.0176	1.0398
Ward+TB (54 bus)	1.2655	1.2771
Ward+TB+QG (18 bus)	1.3946	1.3946

Table 7-4: CPM results after translation for RTE 672-bus network

	PTDF, translation	DV, translation
Ward (195 bus)	1.0002	1.0157
Ward+TB (145 bus)	1.0410	1.1068
Ward+TB+QG (76 bus)	1.2787	1.2832

7.4 Conclusion

In this section, we introduced expansion planning formulations, including CEP and ACEP, and we discussed key solution features. Additionally, we described the translation procedure, which involves transferring the expansion results from the reduced network back to the full network.

8. Summary and Conclusions

This project focused on developing a new network reduction procedure tailored for CEP in power systems. The increasing complexity and size of modern power networks necessitate methods that can efficiently reduce the network size without compromising the fidelity of the expansion planning results. The research presented in this report provides a comprehensive framework for achieving this balance.

8.1 Summary of Key Findings

1. **Network Reduction Procedure:** The research introduced an eight-step network reduction procedure that addresses the unique challenges of CEP. This procedure is designed to reduce computational time while maintaining high model fidelity. The steps include key branch identification, trimming and mapping of buses, zonal division, guided Ward elimination, aggregation, capacity and cost estimation of equivalent branches, and investment translation.
2. **Application and Validation:** The procedure was applied to several test networks, including the IEEE 118-bus system and large-scale networks provided by RTE and MISO. The results demonstrated that the procedure effectively reduced the network size by up to 80%, while preserving the key characteristics of the full network, such as power flows and investment patterns.
3. **Fidelity Assessment:** The research introduced new metrics for assessing the fidelity of the reduced network compared to the full network. These metrics include flow deviations, CEP objective function values, and investment comparison. The results showed that the reduced networks retained high fidelity, with minimal deviations in the critical aspects of the network's performance.
4. **Trade-offs:** The research also explored the trade-offs between network size reduction and model fidelity. It was found that while significant reductions in network size are achievable, careful consideration must be given to the selection of buses for elimination and the estimation of equivalent branch capacities to avoid compromising the accuracy of CEP results.

8.2 Conclusions

The network reduction procedure developed in this project offers a practical solution for applying CEP to large and complex power networks. By focusing on both computational efficiency and model fidelity, the procedure allows for more scalable and accurate expansion planning analyses. This is particularly important as power systems continue to grow in size and complexity, driven by the integration of renewable energy sources and the need for significant transmission investments.

8.3 Future Work

The work presented in this report lays the foundation for several avenues of future research and development. The following areas are identified as key opportunities for further exploration:

Temporal Resolution: One of the key areas not fully explored in this project is temporal resolution. Future work should consider approaches to "downsample" the original set of time points, capturing representative hours within representative days. This would help manage the large temporal datasets often involved in CEP. A recent study presented at PSCC 2024, [46], suggests techniques to isolate time points while capturing key temporal patterns, making it suitable for energy storage applications like batteries. While battery storage is not the core focus of our work, these down-sampling techniques could offer insights into preference sampling methods that improve computational efficiency without sacrificing accuracy.

Key Branch Identification via Umbrella Constraint Identification: Another area for future exploration is the identification of key branches using the "umbrella constraint" identification method [47]. This approach could provide a more rigorous way to pinpoint critical transmission lines for retention during network reduction. By focusing on umbrella constraints, it may be possible to ensure that the reduced network retains branches that have the greatest impact on operational and investment decisions. Exploring this method could lead to improvements in both accuracy and computational performance.

Structural Considerations: The structural integrity of the reduced network is critical to maintaining the fidelity of CEP results. In section 4.2, we discussed the structural issues that arise during network reduction, particularly regarding the elimination of certain buses and branches. Future work should aim to quantify the extent to which structural modifications affect the overall performance of the reduced model. By developing reduction techniques that preserve the network structure, we can achieve reduced networks with higher fidelity. This would involve focusing on methods that maintain key branches and the surrounding structure, while still reducing the model's size efficiently.

Development of Software Tools: To facilitate broader adoption of network reduction techniques, the development of automated software tools is a logical next step. These tools could integrate seamlessly with existing CEP platforms and allow practitioners to apply the reduction methods developed in this project more easily.

Multi-Objective Optimization and Expansion to Other Planning Areas: Lastly, future work could explore multi-objective optimization within the network reduction procedure. This could include balancing different objectives, such as cost minimization, emission reduction, and reliability enhancement. Additionally, the network reduction techniques could be adapted to other planning areas like stochastic planning and reliability analysis, expanding the utility of the work presented in this report.

By pursuing these future research directions, the network reduction procedure developed in this project can be further improved to handle more complex, larger-scale power systems and adapt to a wider range of applications.

References

- [1] Energy Systems Integration Group (ESIG), Capacity Expansion Model summary table, [Online.] Available: <https://www.esig.energy/wp-content/uploads/2022/10/ESIG-CEM-Developer-Survey-Summary-Table.pdf>.
- [2] International Atomic Energy Agency, "Expansion Planning for Electric Generating Systems: A Guidebook," Technical Reports Series No. 241, 1984. [Online.] Available: <https://www.iaea.org/publications/1343/expansion-planning-for-electrical-generating-systems>.
- [3] P. Maloney, P. Chitkara, J. McCalley, B. Hobbs, C. Clack, M. Ortega-Vazquez, A. Tuohy, A. Gaikwad, and J. Roark, "Research to Develop the Next Generation of Electric Power Capacity Expansion Tools: What Would Address the Needs of Planners?" International Journal of Electrical Power & Energy Systems, for the special issue on, Recent Advancements in Electric Power System Development Planning with High-Penetration of Renewable Energy Resources and Dynamic Loads, Volume 121, 2020, ISSN 0142-0615, <https://doi.org/10.1016/j.ijepes.2020.106089>.
- [4] G. Kron, "Tensor analysis of networks," 1939, the General Electric Company.
- [5] J. B. Ward, "Equivalent circuits for power flow studies," AIEE Transactions, Vol. 68, 1949.
- [6] R. Podmore and A. Germond, "Development of dynamic equivalents for transient stability studies," EPRI EL-456, Research Project 763 of the Electric Power Research Institute, April, 1977.
- [7] J. McCalley, J. Dorsey, J. Luini, R. Mackin, and G. Molina, "Subtransmission reduction for voltage instability analysis," IEEE Transactions on Power Systems, March, 1993.
- [8] H. Singh and S. Srivastava, "A reduced network representation suitable for fast nodal price calculations in electricity market," Proc. Of the 2005 IEEE Power Engineering Society General Meeting, pages 2070–2077, 2005.
- [9] X. Cheng and T. Overbye, "Ptdf-based power system equivalents," IEEE Transactions on Power Systems, 20(4):1868–1876, 2005.
- [10] H. Oh, "A new network reduction methodology for power system planning studies." IEEE Transactions on Power Systems, 25(2):677–684, 2009.
- [11] D. Shi and D. Tylavsky, "A novel bus-aggregation-based structure-preserving power system equivalent," IEEE Transactions on Power Systems, 30(4):1977–1986, 2014.
- [12] E. Shayesteh, B. Hobbs, L. Söder and M. Amelin, "ATC-based system reduction for planning power systems with correlated wind and loads", IEEE Trans. Power Syst., vol. 30, no. 1, pp. 429-438, Jan. 2015.
- [13] S. Lumbreras et al., "Network partition based on critical branches for large-scale transmission expansion planning", Proc. 2015 IEEE Eindhoven PowerTech, pp. 1-6, 2015.
- [14] C. Fezeu, K. Bell, J. Ding, P. Panciatici and M. Debry, "Simplified representation of a large transmission network for use in long-term expansion planning", Proc. 2014 Power Syst. Comput. Conf., pp. 1-7, 2014.
- [15] J. Jia, "Transmission Expansion Planning in Large Power Systems Using Power System Equivalencing Techniques", 2014.

-
- [16] Q. Ploussard, L. Olmos and A. Ramos, "An Efficient Network Reduction Method for Transmission Expansion Planning Using Multicut Problem and Kron Reduction," in *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6120-6130, Nov. 2018.
- [17] Here, and in the remainder of this report, the term "operating conditions" refer to loading, wind and solar output, thermal generation dispatch, the contingency state, and the influence of network controllers. Of these, the inclusion of control influence is atypical, and so we further comment. Control influence originates from HVDC converter stations, phase shifting transformers, various flexible AC transmission system (FACTS) controllers, and transmission topology optimization. Ultimately, it is of interest to model thousands of operating conditions (indeed, RTE considers 8760 hourly load/generation conditions for each year over 10 years, combined with 100-800 contingency conditions, resulting in at least 87600 x 100 operating conditions).
- [18] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
- [19] C. Jozs, S. Fliscounakis, J. Maeght, and P. Panciatici. "AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE." arXiv preprint arXiv:1603.01533 (2016).
- [20] I. Pena, C. Martinez-Anido, and B. Hodge. "An extended IEEE 118-bus test system with high renewable penetration." *IEEE Transactions on Power Systems* 33.1 (2017): 281-289.
- [21] R. Sedgewick and K. Wayne. *Algorithms*. Addison-Wesley Professional, 2011.
- [22] Wikipedia contributors (2023). Minimum spanning tree — Wikipedia, the free encyclopedia. [Online; accessed 29-May-2023].
- [23] Y. Jiang, et al. "Network Reduction for Power System Planning: Zone Identification." 2023 North American Power Symposium (NAPS). IEEE, 2023.
- [24] M. E. J. Newman, *Networks: An Introduction*, Oxford University Press, 2010.
- [25] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 60, No. 1, pp. 150-163, 2012.
- [26] N. Karmarkar, "A New Polynomial-Time Algorithm for Linear Programming." In *Proceedings of the 16th Annual ACM Symposium on Theory of Computing (1984)*, pp. 302-11. New York: ACM. Revised version in *Combinatorica* 4: 373-95.
- [27] N. Megiddo, "On the complexity of linear programming," in: *Advances in economic theory: Fifth world congress*, T. Bewley, ed., Cambridge University Press, Cambridge, 1987, pp. 225-268.
- [28] C. Newlun, J. McCalley, R. Amitava, A. Ardakani, A. Venkatraman and A. Figueroa - Acevedo, "Adaptive Expansion Planning Framework for MISO Transmission Planning Process," 2021 IEEE Kansas Power and Energy Conference (KPEC), Manhattan, KS, USA, 2021, pp. 1-6.
- [29] P. Amestoy, T. Davis and I. Duff, "An approximate minimum degree ordering algorithm," *SIAM Journal on Matrix Analysis and Applications*, Vol. 17, No. 4, pp.886-905, 1996.
- [30] J. Rommes and W. Schilders "Efficient methods for large resistor networks," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 29 No.1 pp28-39 2009.

-
- [31] C. Murillo-Sanchez, F. Wu, and J. Fuller, (2008). Power transfer distribution factors in power flow studies. *IEEE Transactions on Power Systems*, 23(4):1716–1724.
- [32] J. Hartigan, and M. Wong. "A k-means clustering algorithm." *Applied statistics* 28.1 (1979): 100-108.
- [33] S. Johnson, (1967). Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254.
- [34] S. Geman, and D. Geman, (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741.
- [35] Another reason we chose to use the mean here (instead of median) is that by choosing it, it mitigates the impact of variation caused by the division matrix (resulting from the variation of operational states).
- [36] In the original quotient-graph method, the division matrix is determined by each bus's power injection/cluster power injection. And thus, when operation states change, the final inter-group impedance will also change. Here, we ensured the division matrix is independent of operating states so that the solved impedance values are constant.
- [37] S. Sharma, "Planning electric power grid with centralized and distributed technologies," Ph. D. Dissertation, 2021, Iowa State University.
- [38] The value of expansion planning for organizations that own only generation or only transmission or neither (e.g., RTOs) is that it provides the solution that is most economically attractive and therefore a baseline for what may be considered to be the best investment plan. Where it is known that other investments (outside of the most economic solution) may or will occur, such investments can be "forced" using appropriate constraints.
- [39] P. Maloney, "Methods for cooptimization planning and plan validation under uncertainty," Ph. D. Dissertation, 2019, Iowa State University.
- [40] C. Newlun, "Cooptimized expansion planning for power system resilience and adaptation," Ph. D. Dissertation, 2021, Iowa State University.
- [41] A. van der Weijde and B. Hobbs, "The economics of planning electricity transmission to accommodate renewables: Using two-stage optimization to evaluate flexibility and the cost of disregarding uncertainty," *Energy Economics*, vol. 34, no. 6, pp. 2089 – 2101, 2012.
- [42] Y. Liu, R. Sioshansi, and A. Conejo, "Multistage stochastic investment planning with multiscale representation of uncertainties and decisions," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 781 – 791, 2017.
- [43] GAMS Solver Manuals, [Online.] Available:https://www.gams.com/latest/docs/S_MAIN.html.
- [44] Gurobi Solver, <http://www.gurobi.cn/download/GuNum.pdf>, Accessed: 08/20/2024.
- [45] GAMS Solver Manuals, [Online] Available: <https://www.gams.com/blog/2017/08/scaling/> Accessed: 08/20/2024
- [46] M. Moradi-Sepahvand and S. Tindemans, "Representative days and hours with piecewise linear transitions for power system planning," *Proc. of the Power Systems Computational Conference, 2024, Paris*.
- [47] A. J. Ardakani and F. Bouffard, "Identification of Umbrella Constraints in DC-Based Security-Constrained Optimal Power Flow," in *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3924-3934, Nov. 2013