

STRUCTURED ANALYSIS AND MODELING OF POWER SYSTEMS

by

Lisa Ming Fern Tang

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Industrial Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2021

Date of final oral examination: 08/17/2021

The dissertation is approved by the following members of the Final Oral Committee:

Michael C. Ferris, Professor, Computer Sciences (Chair)

Vicki M. Bier, Professor, Industrial and Systems Engineering

Christopher L. DeMarco, Professor Emeritus, Electrical & Computer Engr.

Bernard C. Lesieutre, Professor, Electrical and Computer Engineering

Jeffrey T. Linderth, Professor, Industrial and Systems Engineering

© Copyright by Lisa Ming Fern Tang 2021
All Rights Reserved

Wholeheartedly dedicated to:

My amazing parents, George & Audrey, for inspiring me to greater heights.

My dear husband, Chris, whose unwavering support kept this dream alive.

My sweet Arthur & Cinnabon, who bring such joy and richness to to my life.

and Michael Ferris, for setting this journey in motion.

ACKNOWLEDGMENTS

Throughout the long duration of my research, I have been most fortunate to have Professor Michael Ferris as my advisor, who has gone above and beyond the call of any mentor. He began my passion for Optimization and subsequent journey into the PhD program, and it is not hyperbole to state that this, or any thesis would be nonexistent without his dedication. I have gained so much from his invaluable guidance, insights and high expectation of my research efforts, while his patience, confidence and steadfastness went a long way in sustaining my interest and drive. For these reasons and much more, I am forever grateful.

I have also received much encouragement, valuable input and perceptive feedback from the members of my doctoral committee, Professors Vicki Bier, Chris DeMarco, Bernard Lesieutre, and Jeff Linderoth. By taking the time to evaluate my body of work, their individually unique insights have improved and uplifted the content of this thesis document. Professors DeMarco and Lesieutre have additionally been instrumental in crafting and validating my research efforts, in particular Chapters 1 and 2. Their inside perspective from the viewpoint of Electrical and Power Systems Engineering is evident throughout. To all above, I wish to record my deep appreciation.

There are many excellent folks affiliated with the Optimization and Power Systems groups, many of whom the line between friends and collaborators has been blurred. A great many thanks to Jesse Holzer, Yanchao Liu, Youngdae Kim, Byungkwon Park, Martha Loewe, Zev Friedman, Steve Wangen, Scott Greene, Dan Molzahn and Adam Christenson. Your time, contributions to this thesis, support and company throughout the years has been invaluable.

This dissertation would also not be possible without the generosity of many institutions. The departments of Computer Science and Industrial

Systems and Engineering at the University of Wisconsin-Madison have supported me generously through teaching and research positions. Research assistantships have been made possible through the grants funded by the Federal Energy Regulatory Commission, Department of Energy and the Advanced Research Projects Agency - Energy. I also gratefully acknowledge the contributions, time and insights of my collaborators at the Federal Energy Regulatory Commission, National Oceanic and Atmospheric Administration, and the department of Atmospheric and Oceanic Science at UW-Madison. In addition, I would also like to thank Eugene Litvinov, Feng Zhao and Tongxin Zheng for the rewarding and fruitful internship experience at ISO New England. Special thanks to Mandy Humbert, my manager at Schneider National, for her support and encouragement in completing my doctoral degree.

And finally, I am greatly indebted to my family. Parents, George and Audrey, have served as inspirations and encouraged me to keep aiming higher. Their selfless sacrifices, enduring love and faith are the reasons I am here today. To my sisters Grace and Doreen, and in-laws Siong, Fred, John and Pam, thank you so very much for the constant encouragement and wonderful times spent together. It would be remiss of me to not thank Mark and Mary for their loving welcome into the family and helping care for us these past few years. Last but not least, a very special acknowledgement to my husband, Chris, who has been a pillar of strength, especially on the last lap of my dissertation. Without his constant encouragement, sacrifices, patience, and unwavering support, the completion of this thesis would have been that much harder. Thank you for seeing this through with me.

CONTENTS

Contents iv

List of Tables ix

List of Figures xi

Abstract xii

1	A toolbox for exploring optimal power flow models	1
1.1	<i>Nomenclature</i>	3
1.2	<i>Model archive</i>	5
1.2.1	DC approximations	6
1.2.2	AC formulations	9
1.3	<i>Data Archive</i>	17
1.3.1	D-curves	18
1.3.2	Demand bidding	20
1.4	<i>Numerical experiments and observations</i>	21
1.4.1	Exploring different AC formulations	21
1.4.2	Initial starting conditions	22
1.4.3	Solver performances	25
1.5	<i>Sparse Tableau Formulation</i>	26
1.5.1	Overview on Constructing the STF Circuit	27
1.5.2	STF Network Elements Modeling	30
1.5.3	Construction of the incidence matrix A	32
1.5.4	Optimal Power Flow Model with STF	35
1.5.5	Sparse Tableau Formulation in the OPF Toolbox	39
1.5.6	Empirically Observed Computation Results	39
1.6	<i>Conclusions</i>	40

2	Large-scale datasets in power systems	41
2.1	<i>Introduction</i>	42
2.2	<i>Missing reactive demand</i>	43
2.2.1	Power factors	44
2.2.2	Number of buses with non-zero reactive demand .	45
2.2.3	Ratio of withdrawals to injections	45
2.3	<i>Solution methodology for large-scale datasets</i>	46
2.3.1	An AC feasible unit commitment profile	47
2.3.2	Initial starting point for ACOPF	48
2.4	<i>Numerical integrity and quality assurance</i>	50
2.4.1	Solution quality of UCDC models	50
2.4.2	Solution quality of reactive power values	52
2.4.3	Solution quality of AC OPF models	53
2.5	<i>Additional Issues in Network Feasibility</i>	55
2.5.1	Tap transformers	55
2.5.2	Line limit relaxation	56
2.6	<i>Epigrids Data</i>	57
2.6.1	Data Utility Tool Updates	58
2.6.2	PowerFlow Feasibility of the datasets	59
2.7	<i>Conclusions</i>	60
3	Hierarchical Framework for Long Term Planning Models	62
3.1	<i>Hierarchical Framework</i>	64
3.1.1	Transmission Planning Model	65
3.1.2	Price Response Function	66
3.1.3	Model Assumptions	69
3.1.4	Generator upgrades	70
3.2	<i>Solution Method</i>	72
3.2.1	Optimization Model	72
3.2.2	MCP	73
3.2.3	Restricted MCP	74

3.2.4	Comparison and Stability Analysis	76
3.3	<i>Numerical Examples</i> 78	
3.3.1	A function evaluation of $p_j^\omega(x)$ for fixed x	80
3.3.2	Transmission Line Expansion	81
3.3.3	Benchmark Comparison	84
3.4	<i>Conclusion</i> 86	
4	Real-time Co-optimization of Energy and Reserve Markets with Security Constraints 88	
4.1	<i>Introduction</i> 88	
4.2	<i>Model Formulation</i> 90	
4.2.1	Background	90
4.2.2	Iterative Modified Benders	96
4.2.3	Differentiating contingency events	97
4.3	<i>Solution method</i> 101	
4.3.1	Algorithm Overview	101
4.3.2	Contingency Handling	102
4.3.3	Algorithmic enhancements	105
4.4	<i>Numerical Results</i> 109	
4.4.1	14-bus IEEE Testcase	109
4.4.2	IEEE Polish Cases	111
4.5	<i>Contingency Modeling Extensions in the OPF Toolbox</i> 113	
4.6	<i>Conclusion</i> 113	
5	Stochastic Modeling with Renewable Energy115	
5.1	<i>Day-Ahead Operational Model with Reserve Co-optimization</i> 117	
5.2	<i>Data sources for modeling exercise</i> 119	
5.2.1	New England Transmission System	119
5.2.2	CAPS Ensemble Surface Insolation Forecast	120
5.2.3	Hourly Demand Data	122
5.3	<i>Solution Heuristic</i> 122	

5.4	<i>Scenario extension to GDX datafile definition</i>	124
5.5	<i>Observed Numerical Results and Conclusions</i>	125
6	Future Work	127
A	Model options for Toolbox Model Archive	128
A.1	<i>Model descriptions</i>	128
A.1.1	Power Flow Models	128
A.2	<i>Power Flow Model usage</i>	129
A.2.1	OPF model options	130
A.2.2	Unit commitment model options	131
A.2.3	DC specific model options	133
A.2.4	AC specific model options	134
A.3	<i>Additional Notes</i>	136
B	Collection of data utilities	137
B.1	<i>Model descriptions</i>	137
B.1.1	Format conversion tools	137
B.1.2	Other tools	138
B.2	<i>Model Options</i>	138
B.2.1	Format conversion model options	138
B.2.2	Model to_gdx.gms options	138
B.2.3	Model calc_S_matrix.gms options	140
B.2.4	Model make_demand_curves.gms options	140
B.3	<i>Prime Mover Abbreviations</i>	141
B.4	<i>Fuel Info Abbreviations</i>	142
C	Decoupled AC Model	145
C.1	<i>The Power Flow Equations</i>	145
C.2	<i>$P - \theta$ subproblem</i>	147
C.3	<i>$Q - V$ subproblem</i>	148

D Generator Capability Curve Approximation150*D.1 Armature Current Limit*150*D.2 Field Current Limit*151*D.3 End Region Heating Limit*154*D.4 Prime Mover Limits*156*D.5 Limitation of the D-curve*158*D.6 Nonphysical Data*158**E Unit Commitment Approximation**160**References**162

LIST OF TABLES

1.1	Description of Sets	3
1.2	Description of Parameters	4
1.3	Description of Variables	5
1.4	Rectangular vs. D-curve constraints	19
1.5	Comparison of OPF formulations	21
1.6	Comparison of initial conditions, Polar ACOPF	24
1.7	Comparison of solver performance	25
1.8	Description of Sets	36
1.9	Description of Functions and Parameters	37
1.10	Description of Decision Variables	37
1.11	Comparison of STF OPF to standard ACOPF solve times	40
2.1	Large-scale Data: Statistics	42
2.2	Line losses: Testcases	47
2.3	Initial conditions: large-scale datasets	48
2.4	Comparing Initial conditions: Winter & Summer	50
2.5	Examiner results: IPOPTH	54
2.6	Examiner results: CONOPT	54
2.7	Examiner results: Winter/Summer	55
2.8	Summary of line infeasibility for Summer dataset	57
3.1	Description of Sets	64
3.2	Description of Variables	65
3.3	Description of Parameters	65
3.5	Equilibrium Model Parameters	79
3.6	Example 1: Scenario information	79
3.7	Optimization Model Parameters	83
4.1	Description of Sets	90

4.2	Description of Parameters and Variables	91
4.3	Differences in post-contingency corrective actions	105
4.4	Comparing runtimes of JAVA and GUSS implementations, based on size of contingency list and number of threads	105
4.5	Comparison of Numerical Results using 24 threads	107
4.6	Iteration analysis of 14-bus testcase	110
4.7	Reserve allocation results	112
4.8	Data Definitions for Contingency Analysis	113
5.1	Data Definitions for Stochastic Modeling	124
5.2	Solution Comparison with different ramping capabilities	126
B.1	Prime Mover Abbreviations	142
B.2	Fuel Abbreviations	144
D.1	Typical P^{min} for Steam Turbines and Combined Cycle Prime Movers (Figures 15 and 23 in [40])	157
D.2	Typical P^{min} for Combustion Turbine Prime Movers (Figure 25 in [40])	157
D.3	Typical P^{min} Without Prime Mover Data (Figure 10 in [40]) . .	157

LIST OF FIGURES

1.1	Typical Generator Capability Curve (Figure 1 from [45])	18
1.2	Example of D-curve using the toolbox	19
1.3	Example of Demandbid curve using the toolbox	21
1.4	Two-port representation for network element k	28
1.5	π -equivalent circuit for transmission line	30
1.6	Nonlinear current source element as one port element	34
3.1	Representative decision-making timescales in electric power systems, courtesy of C. DeMarco	62
3.2	Illustration of investment variable's effect on cost curve	71
3.3	Graphical depiction of 14 bus example	78
4.1	Example of post-contingency line limit requirements	91
4.2	(a) Contingency 2 is intrinsically infeasible as there is no overlap in the feasible region of the base case with its corresponding subproblem. (b) Both contingencies 1 and 2 are not simultane- ously feasible with the base case and create conflicting cuts.	95
4.3	Flowchart showing an overview of the Parallel Benders algorithm	101
4.4	Graphical depiction of 14-bus testcase	109
5.1	Example of Solar Ensemble Forecast	121
D.1	Generator Capability Curve ($Q^{max} \geq P^{max}$)	152
D.2	Generator Capability Curve ($Q^{max} \leq 0.6P^{max}$)	153
D.3	Generator Capability Curve ($P^{max} \geq Q^{max} > 0.6P^{max}$)	153
D.4	Generator Capability Curve ($ Q^{min} \leq 0.31R^{max}$)	155
D.5	Generator Capability Curve ($R^{max} > Q^{min} > 0.31R^{max}$)	155
D.6	Example of when to impose the D-curve constraint	158
E.1	Approximation of Unit Commitment	161

ABSTRACT

Technological and economic trends point towards a continuing and significant increase in our nation's reliance on the power grid. Meanwhile, the landscape of today's power industry is rapidly changing, especially with the recent evolution of renewable generation, electricity storage and electric vehicles. These issues not only challenge our ability to provide accurate demand and supply forecasts, but they also raise questions on how best to operate the power grid in a reliable, fair and cost-efficient manner. Arising from the need to address these changes, interest in using mathematical models to answer these increasingly complex questions has grown. This thesis discusses the analytical tools we have developed, and our contribution of solution methods, modeling frameworks and approaches in power systems research.

Chapter 1 provides an overview of some of the challenges faced by the power industry and presents a collaborative and unified toolbox for power systems researchers. The usefulness of this toolbox is underscored by the examples of structured analysis comparing different models, solvers, starting points and formulations. Chapter 2 expands on the curation process of large datasets within the Toolbox's data archive and continues to explore solution methodologies for solving large and realistic datasets. We also discuss the need for quality assurance in solutions involving realistic datasets and provide suggestions on how to ensure solution integrity within our GAMS models.

In Chapter 3, we propose a hierarchical framework to investigate the complexities of long-term power planning decisions which is designed to capture, in a flexible and tractable manner, the richness of the model using transparent interactions between multiple agents making decisions at different timescales. We also investigate an iterative process to solve the shorter-scale Nash equilibrium problem.

Chapter 4 presents a fast and effective iterative Benders algorithm to solve an “N-1” Security Constrained Economic Dispatch (SCED) Model in realistic settings. In addition to solution speed, this section contributes to existing literature by directly applying security constraints to a co-optimized dispatch model. In Chapter 5, we present a stochastic model and solution method for solving a day-ahead planning model with renewable energy scenarios. By using a moving-window 2-stage stochastic process with look-ahead information, we can maintain a reasonable model size for tractability purposes. Furthermore, we provide versions of the SCED and stochastic models in the Chapter 1 Toolbox, along with data definition enhancements for contingency and scenario planning. These additions are made in order to adapt the application of the toolbox to a wider and richer set of research topics, thus improving the collaborative process within Power Systems planning and operations.

1 A TOOLBOX FOR EXPLORING OPTIMAL POWER FLOW MODELS

Well-accepted estimates cite a 35% growth in electricity demand over the next 20 years, based on technological and economic trends which point towards a continuing and significant increase in our nation's reliance on the power grid [27]. The landscape of today's power industry is rapidly changing, especially with the recent evolution of renewable generation, storage and electric vehicles [44] which has had a detrimental effect on our ability to provide accurate demand and supply forecasts. As a result of the sector addressing these changes, interest in using mathematical models to answer these increasingly complex questions has grown. The Federal Energy Regulatory Commission (FERC) in their 2009-2014 strategic plan, states their objective to "promote operational efficiency in wholesale market through the exploration and encouragement of the use of software and hardware that will optimize market operations" [31, p. 12]. At the heart of power systems operations lies the AC Optimal Power Flow (ACOPF) model which describes the physical properties of the network and thus helps guarantee network stability during operation.

System reliability and generator commitment is typically overseen by regional Independent System Operators (ISO) that operate as a go-between and system regulator for generators and wholesale consumers. In its most general form, the optimal power flow (OPF) problem is a cost minimization problem with equality constraints enforcing Kirchhoff's current law (i.e. power balance at each bus in the network) and inequality constraints enforcing physical and stability limits on power generation and flow. While the constraints describe the physical relationships involved in equipment and electricity flow, the objective minimizes cost of dispatch to fulfill the consumer demand. The OPF problem is well-researched and the standard models have gone through little change in the last 50 years [17]. While

significant progress has been made in algorithms, decomposition and stochastic methods, there lacks a cohesiveness in comparative literature due to the different models, file formats, software and solvers. Additionally, software to solve complex simulations of power system models are typically either highly specialized to answer a specific type of problem or proprietary and lack transparency. Two of the most commonly used softwares for solving OPF models is Matpower which is an open-source Matlab power system simulation package provided by Zimmerman et.al. [79], and PSS/E which is a widely used power system dynamics simulation software package that is sold by the Siemens Corporation. The OPF Toolbox introduced in this chapter is an an open source toolbox implemented in GAMS, and is comprised of the following three components:

1. A model archive with multiple OPF formulations.
2. A data library that includes well-researched IEEE datasets as commonly used in literature.
3. A data utility archive for compiling, enhancing and converting datasets between 3 popular formats used in power systems analysis (GAMS, Matpower, PSS/E-31).

While the transparent objective of this toolbox is to provide a standardized environment in which to test and compare different OPF formulations and algorithms, the more subtle aim is to facilitate the modeling of complex power systems models, many of which rely on the OPF model. By providing core models, established datasets and import/export tools to interface with popular formats, we encourage the user not just to use the toolbox as provided, but also to take advantage of the robust modeling capabilities in the GAMS environment, such as the ability to easily switch between multiple powerful solvers, as well as the ease of testing different model reformulations with the convenience of its high level modeling language. The models presented in following sections serve a dual purpose.

In addition to introducing the core of toolbox, they also provide the reader an overview of many OPF formulations that exist in literature.

1.1 Nomenclature

The following notation will be used in describing the different OPF formulations. Note that not all parameters and variables will appear in every model, and set $\mathcal{T}$ is typically only used in unit commitment models. Thus, index $t \in \mathcal{T}$ which is used to describe time periods, is not used in all the formulations although it is described and indexed within these tables. Another point of note is that every line $ijc \in \mathcal{E}$, occurs at most once in this set, that is to say that if $ijc \in \mathcal{E}$, then $jic \notin \mathcal{E}$. The convention as used in our models is that a positive flow on a line represents a withdrawal at its source end and an injection at its terminating end. This usage is further clarified in the introductory passage of Section 1.2.2.

Set	Description
$\mathcal{N}$	Set of buses in the transmission network
$\mathcal{G}$	Set of generators in the transmission network
$\mathcal{I}$	Set of interfaces in the transmission network
$c \in \mathcal{C}$	Set of transmission connections in network
$t \in \mathcal{T}$	Set of time periods
$\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N} \times \mathcal{C}$	Set of lines in the transmission network
$\mathcal{E}_i \subseteq \mathcal{E}$	Subset of lines $\mathcal{E}$ belonging to interface $i \in \mathcal{I}$
$\mathcal{G}_i \in \mathcal{G}$	Subset of generators $\mathcal{G}$ at bus $i \in \mathcal{N}$

Table 1.1: Description of Sets

Parameters	Description
$\tilde{c}_i(\cdot)$	Cost function for generation at bus $i \in \mathcal{N}$
d_{it}^P, d_{it}^Q	Real and reactive power demand at bus i in time $t \in \mathcal{T}$

r_{ijc}, x_{ijc}	Resistance and reactance on line $ijc \in \mathcal{E}$
g_i^s, b_i^s	Shunt conductance and susceptance at bus $i \in \mathcal{N}$
$g_{ijc}^{\mathcal{E}}, b_{ijc}^{\mathcal{E}}$	Conductance and susceptance on line $ijc \in \mathcal{E}$
b_{ijc}^C	Branch charging susceptance of line $ijc \in \mathcal{E}$
τ_{ijc}, ϕ_{ijc}	Transformer tap ratio and angle on line $ijc \in \mathcal{E}$
Y_{ij}^P, Y_{ij}^Q	Real & imaginary component of Y-bus admittance matrix for line $ij \subseteq \mathcal{N} \times \mathcal{N}$
u_{it}	Generator i 's on/off status at time $t \in \mathcal{T}$
$\bar{F}_{ijc}^P$	Real power limit on line $ijc \in \mathcal{E}$
$\bar{F}_i^{\mathcal{I}}$	Real power limit on interface $i \in \mathcal{I}$
$\bar{I}_{ijc}$	Current magnitude limit on line $ijc \in \mathcal{E}$
$\underline{P}_i, \bar{P}_i$	Lower and upper active power injection limits of generator $i \in \mathcal{G}$
$\underline{Q}_i, \bar{Q}_i$	Lower and upper reactive power injection limits of generator $i \in \mathcal{G}$
$\underline{U}_i^{\text{run}}, \bar{U}_i^{\text{run}}$	Minimum up and downtime of generator $i \in \mathcal{G}$
$\underline{U}_i^{\text{ramp}}, \bar{U}_i^{\text{ramp}}$	Rampup and rampdown rate of generator $i \in \mathcal{G}$
$\underline{V}_i, \bar{V}_i$	Voltage magnitude lower and upper limits at bus $i \in \mathcal{N}$

Table 1.2: Description of Parameters

Variables	Description
F_{ijct}^P, F_{ijct}^Q	Real and reactive power flowing on line $ijc \in \mathcal{E}$ in time $t \in \mathcal{T}$
I_{ijct}^P, I_{ijct}^Q	Real and reactive current flowing on line $ijc \in \mathcal{E}$ in time $t \in \mathcal{T}$
P_{it}, Q_{it}	Real and reactive power generated at bus $i \in \mathcal{N}$ in time $t \in \mathcal{T}$
V_{it}, θ_{it}	Voltage magnitude and angle at bus $i \in \mathcal{N}$ in time $t \in \mathcal{T}$

V_{it}^P, V_{it}^Q	Real & imaginary rectangular voltage components at bus $i \in \mathcal{N}$ in time $t \in \mathcal{T}$
$U_{it}^{\text{on}}, U_{it}^{\text{off}}, U_{it}^{\text{run}}$	Start, stop and run status of generator $i \in \mathcal{G}$ at time $t \in \mathcal{T}$

Table 1.3: Description of Variables

1.2 Model archive

The model archive currently consists of a collection of 12 OPF models, and various model options (called using the command line) allow the user to experiment with different features and relaxations. The user may, for example, specify specific time intervals to solve, choose between 4 objective functions (feasibility, linear, quadratic, piecewise linear), select from multiple starting point options, relax generator or line operating limits, and enforce D-curve or demand bidding constraints. A complete list of GAMS models and corresponding model options can be found in Appendix A.

Several different formulations exist for the power flow equations, the most popular of which are the polar power-voltage formulation (P), the rectangular power-voltage formulation (R), the rectangular current-voltage (IV), and the DC approximation (DC) which is a popular linear approximation of the problem. Cost functions (1.1) can be given as quadratic coefficients or as a list of points specifying a piecewise linear function. Piecewise linear functions must be convex at this time.

$$\min \quad \sum_i \tilde{c}_i(P_i) \quad (1.1)$$

The following sections provide detail on the mathematical formulations of the models provided in the archive.

1.2.1 DC approximations

The DCOPF approximation stems from a simplification of the physics of the transmission network. It linearizes the problem by treating resistance and line losses as negligible, assuming that the per unit voltage magnitude at each bus is 1, and uses small angle approximation for sine and cosine. Small angle approximation assumes that the phase angle difference between the buses, $(\theta_i - \theta_j) \forall i \in \mathcal{N}$, is small enough and approximates the cosine and sine functions with $\cos(i) = 1$ and $\sin(i) = i$ respectively.

With these assumptions, reactive power in the network becomes zero and can be ignored. Since losses are also zero in this model, the power for each ijc line, F_{ijc} , is only computed once (1.2) as the flow in the opposite direction F_{jic} is the same. The resulting formulation is a linear model that solves for variables P, θ, F^P subject to node balance constraints (1.3), interface flow limits (1.4), limits on the angle difference between connected buses (1.5), and bound constraints on P and F^P (1.6). Note that generator statuses are provided and this is a single time-period model, therefore the index $t \in \mathcal{T}$ is not used.

$$\begin{aligned} \min_{P, F^P, \theta} \quad & \sum_i \tilde{c}_i(P_i) \\ \text{s.t.} \quad & F_{ijc}^P = \frac{-1}{\tau_{ijc} x_{ijc}} (\theta_j - \theta_i + \phi_{ijc}) \quad \forall ijc \in \mathcal{E} \end{aligned} \quad (1.2)$$

$$\begin{aligned} \sum_{k \in G_i} P_k - \sum_{(jc): ijc \in \mathcal{E}} F_{ijc}^P + \sum_{(jc): jic \in \mathcal{E}} F_{jic}^P \\ - d_i^P - g_i^s = 0 \quad \forall i \in \mathcal{N} \end{aligned} \quad (1.3)$$

$$\sum_{ijc, jic \in \mathcal{E}_k} F_{ijc}^P \leq \bar{F}_k^{\mathcal{I}} \quad \forall k \in \mathcal{I} \quad (1.4)$$

$$\frac{-\pi}{3} \leq \theta_i - \theta_j \leq \frac{\pi}{3} \quad \forall (ij) : ij \in \mathcal{E} \quad (1.5)$$

$$\begin{aligned} u_i \underline{P}_i &\leq P_i \leq u_i \bar{P}_i & \forall i \in \mathcal{G} \\ -\bar{F}_{ijc}^P &\leq F_{ijc}^P \leq \bar{F}_{ijc}^P & \forall ij \in \mathcal{E} \end{aligned} \quad (1.6)$$

DC using Shift Factor Matrices

The DC power flow model can also be used to compute the sensitivities of branch flows to changes in nodal real power injections, sometimes called injection shift factors [79]. These matrices are typically large and dense, $|\mathcal{E}| \times |\mathcal{N}|$, and each element describes the expected change in real power flow on a line, $\mathcal{E}$, as it reacts to a 1 unit increase in the power injection at bus $\mathcal{N}$, under the strict assumption that the additional unit of power is based on some slack distribution. Additional detail about shift matrices can be found in [79].

Through use of the shift matrix, we can eliminate bus voltage angles, θ , as an intermediate variable, and the shift matrix model is reduced to (1.7-1.10), where H is the shift matrix. Due to the matrix density however, this formulation is not recommended over the standard DCOPF formulation, especially in large models.

$$\begin{aligned} \min_{P, F^P} \quad & \sum_i \tilde{c}_i(P_i) \\ \text{s.t.} \quad & F_{ijc}^P = \sum_{k \in \mathcal{N}} H_{ijck} \sum_{l \in \mathcal{G}_k} (P_l - d_k^P - g_i^s) \quad \forall ij \in \mathcal{E} \end{aligned} \quad (1.7)$$

$$\sum_{k \in \mathcal{G}_i} P_k - d_i^P - g_i^s = 0 \quad \forall i \in \mathcal{N} \quad (1.8)$$

$$\sum_{ijc, jic \in \mathcal{E}_k} F_{ijc}^P \leq \bar{F}_k^{\mathcal{I}} \quad \forall k \in \mathcal{I} \quad (1.9)$$

$$\begin{aligned} u_i \underline{P}_i &\leq P_i \leq u_i \bar{P}_i & \forall i \in \mathcal{G} \\ -\bar{F}_{ijc}^P &\leq F_{ijc}^P \leq \bar{F}_{ijc}^P & \forall ij \in \mathcal{E} \end{aligned} \quad (1.10)$$

Unit commitment DC model

Unit commitment models allow the modeler to choose the set of dispatched generators for each time period, and is typically useful when the OPF model needs to be solved over a continuous block of time. When solved over multiple time periods, $t \in \mathcal{T}$, the model includes additional generator operational constraints such as generator ramping, minimum up-time and minimum down-time. Equations (1.11 - 1.14) are equivalent to (1.2 - 1.5) with index $t \in \mathcal{T}$, and unit commitment variables $U = \{U^{\text{on}}, U^{\text{off}}, U^{\text{run}}\}$ are factored into generator power limits in (1.15). Binary variable relationships are modeled in (1.16), minimum up and down time in (1.17-1.18), and generator ramping conditions in (1.19-1.20), while (1.21) defines the remaining variable bounds. The constraints in (1.16-1.18) for the generator unit commitment problem is based on the work done by Hedman, O'Neill and Oren in [39].

$$\begin{aligned} \min_{P, F^P, \theta, U} \quad & \sum_t \tilde{c}_i(P_{it}) \\ \text{s.t.} \quad & F_{ijct}^P = \frac{-1}{\tau_{ijc} x_{ijc}} (\theta_{jt} - \theta_{it} + \phi_{ijc}) \quad \forall ij c \in \mathcal{E}, t \in \mathcal{T} \quad (1.11) \end{aligned}$$

$$\begin{aligned} \sum_{k \in \mathcal{G}_i} P_{kt} - \sum_{(jc): ij c \in \mathcal{E}} F_{ijct}^P \\ + \sum_{(jc): jic \in \mathcal{E}} F_{jict}^P - d_{it}^P - g_{it}^s = 0 \quad \forall i \in \mathcal{N}, t \in \mathcal{T} \quad (1.12) \end{aligned}$$

$$\sum_{ijc, jic \in \mathcal{E}_k} F_{ijct}^P \leq \bar{F}_k^{\mathcal{I}} \quad \forall k \in \mathcal{I}, t \in \mathcal{T} \quad (1.13)$$

$$\frac{-\pi}{3} \leq \theta_{it} - \theta_{jt} \leq \frac{\pi}{3} \quad \forall (ij) : ij c \in \mathcal{E}, t \in \mathcal{T} \quad (1.14)$$

$$U_{it}^{\text{run}} * \underline{P}_i \leq P_{it} \leq U_{it}^{\text{run}} * \bar{P}_i \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.15)$$

$$U_{it}^{\text{on}} - U_{it}^{\text{off}} = U_{it}^{\text{run}} - U_{i,t-1}^{\text{run}} \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.16)$$

$$\sum_{t_0=t-\bar{U}_i^{\text{run}}+1}^t U_{i,t_0}^{\text{on}} \leq U_{it}^{\text{run}} \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.17)$$

$$U_{it}^{\text{run}} \leq 1 - \sum_{t_0=t-\bar{U}_i^{\text{run}}+1}^t U_{i,t_0}^{\text{off}} \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.18)$$

$$P_{it} \leq P_{i,t-1} + U_{it}^{\text{run}} \bar{U}_i^{\text{ramp}} + U_{it}^{\text{on}} (\bar{P}_i - \bar{U}_i^{\text{ramp}}) \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.19)$$

$$P_{i,t-1} \leq P_{i,t} + U_{it}^{\text{run}} \underline{U}_i^{\text{ramp}} + U_{it}^{\text{off}} (\bar{P}_i - \underline{U}_i^{\text{ramp}}) \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.20)$$

$$-\bar{F}_{ijc}^P \leq F_{ijc}^P \leq \bar{F}_{ijc}^P \quad \forall ijc \in \mathcal{E}, t \in \mathcal{T}$$

$$U_{it}^{\text{on}}, U_{it}^{\text{off}}, U_{it}^{\text{run}} \in \{0, 1\} \quad \forall i \in \mathcal{G}, t \in \mathcal{T} \quad (1.21)$$

1.2.2 AC formulations

In the true physics of an AC power flow model, power transmitted through lines $ijc \in \mathcal{E}$ may experience loss, and this is reflected in the calculations of active and reactive power. That is to say that the power leaving bus $i \in \mathcal{N}$ on line $ijc \in \mathcal{E}$ may not necessarily equal the power entering the bus $j \in \mathcal{N}$ on the other end, i.e. F_{ijc} may not equal $-F_{jic}$. Recall that the convention used in these models is that a positive flow on a line, F_{ijc} represents a withdrawal at the source $i \in \mathcal{N}$ and an injection at its terminating end $j \in \mathcal{N}$. In this section, we provide three formulations of the AC power flow model, namely polar power-voltage in Section 1.2.2, rectangular power-voltage in 1.2.2 and rectangular current-voltage in 1.2.2.

Polar Power-Voltage Formulation (P)

The polar power-voltage formulation uses the polar form of complex quantities and explicitly uses sines and cosines in the power flow constraints. Variables Q_i and V_i model the reactive power support provided by genera-

tors and the bus voltage at bus $i \in \mathcal{N}$ respectively, while reactive power on a line is modeled by F_{ijc}^Q .

Due to possible line losses, real power flow on lines $ijc \in \mathcal{E}$, which are approximated in the DCOPF by (1.2), is modeled here by (1.22 - 1.23), and reactive power flow on lines is computed through (1.24 - 1.25). Node balance equations are updated in (1.26 - 1.27), limits on line interface flows and bus angle differences are provided in (1.28 - 1.29), and (1.30) defines bounds on the remaining variables.

$$\begin{aligned} & \min_{P, Q, F^P, F^Q, \theta, V} \quad \sum_i \tilde{c}_i(P_i) \\ \text{s.t.} \quad & F_{ijc}^P = \frac{1}{\tau_{ijc}^2} g_{ijc}^{\mathcal{E}} V_i^2 \\ & - \frac{1}{\tau_{ijc}} V_i V_j \left(g_{ijc}^{\mathcal{E}} \cos(\theta_i - \theta_j - \phi_{ijc}) \right. \\ & \left. + b_{ijc}^{\mathcal{E}} \sin(\theta_i - \theta_j - \phi_{ijc}) \right) \quad \forall ijc \in \mathcal{E} \quad (1.22) \end{aligned}$$

$$\begin{aligned} & F_{jic}^P = g_{ijc}^{\mathcal{E}} V_j^2 \\ & - \frac{1}{\tau_{ijc}} V_i V_j \left(g_{ijc}^{\mathcal{E}} \cos(\theta_j - \theta_i + \phi_{ijc}) \right. \\ & \left. + b_{ijc}^{\mathcal{E}} \sin(\theta_j - \theta_i + \phi_{ijc}) \right) \quad \forall ijc \in \mathcal{E} \quad (1.23) \end{aligned}$$

$$\begin{aligned} & F_{ijc}^Q = -\frac{1}{\tau_{ijc}^2} \left(b_{ijc}^{\mathcal{E}} + \frac{b_{ijc}^C}{2} \right) V_i^2 \\ & - \frac{1}{\tau_{ijc}} V_i V_j \left(g_{ijc}^{\mathcal{E}} \cos(\theta_i - \theta_j - \phi_{ijc}) \right. \\ & \left. - b_{ijc}^{\mathcal{E}} \sin(\theta_i - \theta_j - \phi_{ijc}) \right) \quad \forall ijc \in \mathcal{E} \quad (1.24) \end{aligned}$$

$$\begin{aligned} & F_{jic}^Q = -\left(b_{ijc} + \frac{b_{ijc}^C}{2} \right) V_j^2 \\ & - \frac{1}{\tau_{ijc}} V_i V_j \left(g_{ijc}^{\mathcal{E}} \cos(\theta_j - \theta_i + \phi_{ijc}) \right. \\ & \left. - b_{ijc}^{\mathcal{E}} \sin(\theta_j - \theta_i + \phi_{ijc}) \right) \quad \forall ijc \in \mathcal{E} \quad (1.25) \end{aligned}$$

$$\begin{aligned} \sum_{k \in \mathcal{G}_i} P_k - \sum_{(jc):ijc \in \mathcal{E}} F_{ijc}^P - \sum_{(jc):jic \in \mathcal{E}} F_{ijc}^P \\ - d_i^P - (V_i)^2 g_i^s = 0 \end{aligned} \quad \forall i \in \mathcal{N} \quad (1.26)$$

$$\begin{aligned} \sum_{k \in \mathcal{G}_i} Q_k - \sum_{(jc):ijc \in \mathcal{N}} F_{ijc}^Q - \sum_{(jc):jic \in \mathcal{E}} F_{ijc}^Q \\ - d_i^Q + (V_i)^2 b_i^s = 0 \end{aligned} \quad \forall i \in \mathcal{N} \quad (1.27)$$

$$\sum_{ijc, jic \in \mathcal{E}_k} F_{ijc}^P \leq \bar{F}_k^{\mathcal{I}} \quad \forall k \in \mathcal{I} \quad (1.28)$$

$$\frac{-\pi}{3} \leq \theta_i - \theta_j \leq \frac{\pi}{3} \quad \forall (ij) : ijc \in \mathcal{E} \quad (1.29)$$

$$\begin{aligned} u_i \underline{P}_i \leq P_i \leq u_i \bar{P}_i \quad , \quad u_i \underline{Q}_i \leq Q_i \leq u_i \bar{Q}_i \quad \forall i \in \mathcal{G} \\ \underline{V}_i \leq V_i \leq \bar{V}_i \quad \forall i \in \mathcal{E} \\ -\bar{F}_{ijc}^P \leq F_{ijc}^P \leq \bar{F}_{ijc}^P \quad \forall ijc \in \mathcal{E} \end{aligned} \quad (1.30)$$

Rectangular Power-Voltage Formulation (R)

The second AC formulation we provide uses the rectangular form of complex quantities, resulting in quadratic power flow constraints with respect to these quantities. Unlike the polar formulation, the sines and cosines are of constant parameters and the bus voltage is separated into real and imaginary parts, that is $V_i = ((V_i^P)^2 + (V_i^Q)^2)^{1/2} \quad \forall i \in \mathcal{N}$. Therefore, equations (1.22 - 1.27) from the polar model which define real and reactive power on lines and node balance equations are rewritten in the rectangular fomulation as (1.31-1.36). Additionally, the voltage magnitude limit is no longer a simple bound constraint but is enforced by the quadratic inequality in (1.37). Similar to (1.28-1.30) in the polar formulation, interface limits are imposed in (1.38) and variable bounds are defined in (1.39), while bus angle limits are not explicitly imposed in this formulation.

$$\begin{aligned}
& \min_{P,Q,F^P,F^Q,V} \sum_i \tilde{c}_i(P_i) \\
\text{s.t. } & F_{ijc}^P = \frac{1}{\tau_{ijc}^2} g_{ijc}^\mathcal{E} \left((V_i^P)^2 + (V_i^Q)^2 \right) \\
& - \frac{1}{\tau_{ijc}} \left(g_{ijc} \cos(\phi_{ijc}) - b_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_i^P V_j^P + V_i^Q V_j^Q) \\
& - \frac{1}{\tau_{ijc}} \left(b_{ijc}^\mathcal{E} \cos(\phi_{ijc}) + g_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_j^P V_i^Q - V_i^P V_j^Q) \quad \forall ijc \in \mathcal{E}
\end{aligned} \tag{1.31}$$

$$\begin{aligned}
& F_{jic}^P = g_{ijc} \left((V_j^P)^2 + (V_j^Q)^2 \right) \\
& - \frac{1}{\tau_{ijc}} \left(g_{ijc} \cos(\phi_{ijc}) + b_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_j^P V_i^P + V_j^Q V_i^Q) \\
& - \frac{1}{\tau_{ijc}} \left(b_{ijc}^\mathcal{E} \cos(\phi_{ijc}) - g_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_i^P V_j^Q - V_j^P V_i^Q) \quad \forall ijc \in \mathcal{E}
\end{aligned} \tag{1.32}$$

$$\begin{aligned}
& F_{ijc}^Q = -\frac{1}{\tau_{ijc}^2} \left(b_{ijc}^\mathcal{E} + \frac{b_{ijc}^C}{2} \right) \left((V_i^P)^2 + (V_i^Q)^2 \right) \\
& - \frac{1}{\tau_{ijc}} \left(g_{ijc}^\mathcal{E} \cos(\phi_{ijc}) - b_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_j^P V_i^Q - V_i^P V_j^Q) \\
& - \frac{1}{\tau_{ijc}} \left(b_{ijc}^\mathcal{E} \cos(\phi_{ijc}) + g_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_i^P V_j^P + V_i^Q V_j^Q) \quad \forall ijc \in \mathcal{E}
\end{aligned} \tag{1.33}$$

$$\begin{aligned}
& F_{jic}^Q = -\left(b_{ijc}^\mathcal{E} + \frac{b_{ijc}^C}{2} \right) \left((V_i^P)^2 + (V_i^Q)^2 \right) \\
& - \frac{1}{\tau_{ijc}} \left(g_{ijc}^\mathcal{E} \cos(\phi_{ijc}) + b_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_i^P V_j^Q - V_j^P V_i^Q) \\
& - \frac{1}{\tau_{ijc}} \left(b_{ijc}^\mathcal{E} \cos(\phi_{ijc}) - g_{ijc}^\mathcal{E} \sin(\phi_{ijc}) \right) (V_j^P V_i^P + V_j^Q V_i^Q) \quad \forall ijc \in \mathcal{E}
\end{aligned} \tag{1.34}$$

$$\begin{aligned}
\sum_{k \in \mathcal{G}_i} P_k - \sum_{(jc): ijc \in \mathcal{E}} F_{ijc}^P - \sum_{(jc): jic \in \mathcal{E}} F_{ijc}^P - d_i^P \\
- \left((V_i^P)^2 + (V_i^Q)^2 \right) g_i^s = 0 \quad \forall i \in \mathcal{N}
\end{aligned} \tag{1.35}$$

$$\begin{aligned}
\sum_{k \in \mathcal{G}_i} Q_k - \sum_{(jc): ijc \in \mathcal{E}} F_{ijc}^Q - \sum_{(jc): jic \in \mathcal{E}} F_{ijc}^Q - d_i^Q \\
+ \left((V_i^P)^2 + (V_i^Q)^2 \right) b_i^s = 0 \quad \forall i \in \mathcal{N}
\end{aligned} \tag{1.36}$$

$$\underline{V}_i^2 \leq (V_i^P)^2 + (V_i^Q)^2 \leq \overline{V}_i^2 \quad \forall i \in \mathcal{N} \tag{1.37}$$

$$\sum_{ijc, jic \in \mathcal{E}_k} F_{ijc}^P \leq \overline{F}_k^{\mathcal{I}} \quad \forall k \in \mathcal{I} \tag{1.38}$$

$$\begin{aligned}
u_i \underline{P}_i \leq P_i \leq u_i \overline{P}_i, \quad u_i \underline{Q}_i \leq Q_i \leq u_i \overline{Q}_i \quad \forall i \in \mathcal{G} \\
- \overline{F}_{ijc}^P \leq F_{ijc}^P \leq \overline{F}_{ijc}^P \quad ijc \in \mathcal{E}
\end{aligned} \tag{1.39}$$

Rectangular Current-Voltage Formulation (IV)

The third AC model presented here is the rectangular current-voltage formulation which considers the flow of current instead of power on a line. Therefore, the model computes real and reactive current on a line, $\{I_{ijc}^P, I_{ijc}^Q\} \forall ijc \in \mathcal{E}$, instead of $\{F_{ijc}^P, F_{ijc}^Q\} \forall ijc \in \mathcal{E}$ which is the real and reactive power on a line. Similar to the rectangular power-voltage model in Section 1.2.2, the IV formulation uses the rectangular form of complex quantities, $V_i = \left((V_i^P)^2 + (V_i^Q)^2 \right)^{1/2} \forall i \in \mathcal{N}$. Therefore, line flow constraints are once again quadratic in nature with constant sine and cosine quantities.

Equations (1.40-1.43) define real and reactive current flow on a line, and (1.44-1.45) define the node balance constraints. Equations (1.46-1.47) impose bounds on the voltage magnitude and current magnitude respectively, and (1.48) defines other variable bound constraints. Note that the current

formulation does not contain interface flow constraints because converting current to power is counter-intuitive in a current-based formulation. In addition, interface constraints were historically used to control LMP prices while current based formulations were used to consider network stability. The IV formulation also imposes limits on current magnitude in (1.47) instead of using simple variable bounds on the real such as in the rectangular model ((1.39) as we find this data is more easily available.

$$\begin{aligned}
& \min_{P,Q,I^P,I^Q,V} \quad \sum_i \tilde{c}_i(P_i) \\
\text{s.t.} \quad I_{ijc}^P &= \frac{1}{\tau_{ijc}^2} \left(g_{ijc}^{\mathcal{E}} V_i^P - \left(b_{ijc}^{\mathcal{E}} + \frac{b_{ijc}^C}{2} \right) V_i^Q \right) \\
& \quad - \frac{1}{\tau_{ijc}} \left(g_{ijc}^{\mathcal{E}} V_j^P - b_{ijc}^{\mathcal{E}} V_j^Q \right) \cos(\phi_{ijc}) \\
& \quad + \frac{1}{\tau_{ijc}} \left(g_{ijc}^{\mathcal{E}} V_j^Q + b_{ijc}^{\mathcal{E}} V_j^P \right) \sin(\phi_{ijc}) \quad \forall ij \in \mathcal{E} \quad (1.40)
\end{aligned}$$

$$\begin{aligned}
I_{jic}^P &= \left(g_{ijc}^{\mathcal{E}} V_j^P - \left(b_{ijc}^{\mathcal{E}} + \frac{b_{ijc}^C}{2} \right) V_j^Q \right) \\
& \quad - \frac{1}{\tau_{ijc}} \left(g_{ijc}^{\mathcal{E}} V_i^P - b_{ijc}^{\mathcal{E}} V_i^Q \right) \cos(-\phi_{ijc}) \\
& \quad + \frac{1}{\tau_{ijc}} \left(g_{ijc}^{\mathcal{E}} V_i^Q + b_{ijc}^{\mathcal{E}} V_i^P \right) \sin(-\phi_{ijc}) \quad \forall ij \in \mathcal{E} \quad (1.41)
\end{aligned}$$

$$\begin{aligned}
I_{ijc}^Q &= \frac{1}{\tau_{ijc}^2} \left(g_{ijc}^{\mathcal{E}} V_i^Q + \left(b_{ijc}^{\mathcal{E}} + \frac{b_{ijc}^C}{2} \right) V_i^P \right) \\
& \quad - \frac{1}{\tau_{ijc}} \left(g_{ijc}^{\mathcal{E}} V_j^Q + b_{ijc}^{\mathcal{E}} V_j^P \right) \cos(\phi_{ijc}) \\
& \quad - \frac{1}{\tau_{ijc}} \left(g_{ijc}^{\mathcal{E}} V_j^P - b_{ijc}^{\mathcal{E}} V_j^Q \right) \sin(\phi_{ijc}) \quad \forall ij \in \mathcal{E} \quad (1.42)
\end{aligned}$$

$$\begin{aligned}
I_{jic}^Q &= \frac{1}{\tau_{ijc}^2} \left(g_{ijc}^\mathcal{E} V_i^Q + \left(b_{ijc}^\mathcal{E} + \frac{b_{ijc}^C}{2} \right) V_i^P \right) \\
&\quad - \frac{1}{\tau_{ijc}} \left(g_{ijc}^\mathcal{E} V_i^Q + b_{ijc}^\mathcal{E} V_i^P \right) \cos(-\phi_{ijc}) \\
&\quad - \frac{1}{\tau_{ijc}} \left(g_{ijc}^\mathcal{E} V_i^P - b_{ijc}^\mathcal{E} V_i^Q \right) \sin(-\phi_{ijc}) \quad \forall ijc \in \mathcal{E} \quad (1.43)
\end{aligned}$$

$$\begin{aligned}
\sum_{k \in \mathcal{G}_i} P_k - d_i^P - V_i^P &\left(\sum_{(jc):ijc \in \mathcal{E}} I_{ijc}^P + \sum_{(jc):jic \in \mathcal{E}} I_{ijc}^P \right) \\
&- V_i^Q \left(\sum_{(jc):ijc \in \mathcal{E}} I_{ijc}^Q + \sum_{(jc):jic \in \mathcal{E}} I_{ijc}^Q \right) \\
&- \left((V_i^P)^2 + (V_i^Q)^2 \right) g_i^s = 0 \quad \forall i \in \mathcal{N} \quad (1.44)
\end{aligned}$$

$$\begin{aligned}
\sum_{k \in \mathcal{G}_i} Q_k - d_i^Q + V_i^P &\left(\sum_{(jc):ijc \in \mathcal{E}} I_{ijc}^Q + \sum_{(jc):jic \in \mathcal{E}} I_{ijc}^Q \right) \\
&- V_i^Q \left(\sum_{(jc):ijc \in \mathcal{E}} I_{ijc}^P + \sum_{(jc):jic \in \mathcal{E}} I_{ijc}^P \right) \\
&+ \left((V_i^P)^2 + (V_i^Q)^2 \right) b_i^s = 0 \quad \forall i \in \mathcal{N} \quad (1.45)
\end{aligned}$$

$$\underline{V}_i^2 \leq (V_i^P)^2 + (V_i^Q)^2 \leq \overline{V}_i^2 \quad \forall i \in \mathcal{N} \quad (1.46)$$

$$\left((I_{ijc}^P)^2 + (I_{ijc}^Q)^2 \right) \leq \overline{I}_{ijc}^2 \quad \forall k \in \mathcal{I} \quad (1.47)$$

$$u_i \underline{P}_i \leq P_i \leq u_i \overline{P}_i \quad , \quad u_i \underline{Q}_i \leq Q_i \leq u_i \overline{Q}_i \quad \forall i \in \mathcal{G} \quad (1.48)$$

Note that the current flow equations (1.40-1.43) are linear and that (1.44-1.46) involve quadratic and linear terms. This would lead us to hope that since the Hessian of the current-voltage constraints is constant, we might see some benefit in solution time, though in practice, this has not been found to be the case in general. Additionally, because this model limits apparent current rather than apparent power on lines, the solutions tend to be slightly different than the other ACOPF models.

Y-bus formulations

The Y-bus formulations of the above AC models use the Y-bus admittance matrix to calculate in the node balance equations, instead of explicit line flow variables $F_i^P jc, F_i^Q jc$ for example and further discussion on computing the Y-bus admittance matrix can be found in [10, p. 325-6].

Although the Y-bus matrix formulation benefits by eliminating line flow parameters in node balance constraints, it loses this benefit in the case where line flow variables still need to be defined in order to maintain line flow limits. For this reason, we find that the standard line power (SLP) models (polar, rectangular, IV) typically outperform the Y-bus models due to the additional computation required. This document only provides the mathematical formulation for the Y-bus polar model as the other two formulations are very similar in implementation and can be inferred from here.

$$\begin{aligned} \min_{P, Q, F^P, F^Q, \theta, V} \quad & \sum_i \tilde{c}_i(P_i) \\ \text{s.t.} \quad & (1.22) - (1.25) \end{aligned} \tag{1.49}$$

$$\begin{aligned} \sum_{k \in \mathcal{G}_i} P_k - d_i^P - (V_i)^2 g_i^s \\ - V_i \sum_{j \in \mathcal{N}} \left(V_j Y_{ij}^P \cos(\theta_i - \theta_j) \right. \\ \left. - Y_{ij}^Q \sin(\theta_i - \theta_j) \right) = 0 \end{aligned} \quad \forall i \in \mathcal{N} \tag{1.50}$$

$$\begin{aligned} \sum_{k \in \mathcal{G}_i} Q_k - d_i^Q + (V_i)^2 b_i^s \\ - V_i \sum_{j \in \mathcal{N}} \left(V_j Y_{ij}^P \sin(\theta_i - \theta_j) \right. \\ \left. - Y_{ij}^Q \cos(\theta_i - \theta_j) \right) = 0 \end{aligned} \quad \forall i \in \mathcal{N} \tag{1.51}$$

$$\sum_{ijc, jic \in \mathcal{E}_k} F_{ijc}^P \leq \overline{F}_k^T \quad \forall k \in \mathcal{I} \tag{1.52}$$

$$\frac{-\pi}{3} \leq \theta_i - \theta_j \leq \frac{\pi}{3} \quad \forall (ij) : ij \in \mathcal{E} \quad (1.53)$$

$$\begin{aligned} u_i \underline{P}_i &\leq P_i \leq u_i \overline{P}_i, & u_i \underline{Q}_i &\leq Q_i \leq u_i \overline{Q}_i & \forall i \in \mathcal{G} \\ \underline{V}_i &\leq V_i \leq \overline{V}_i & & & \forall i \in \mathcal{E} \\ -\overline{F}_{ijc}^P &\leq F_{ijc}^P \leq \overline{F}_{ijc}^P & & & \forall ij \in \mathcal{E} \end{aligned} \quad (1.54)$$

1.3 Data Archive

The data archive currently consists of many IEEE power systems datasets which are commonly used in many publications, and range from 14 buses in the simple cases to 3375 buses in the Polish cases. The majority of this data is extracted from Matpower's data library, and were converted from Matpower formatted .m files to .gdx files using conversion utilities provided in the toolbox.¹ In addition to the single period IEEE cases, we also compiled six 24-hour datasets based on the the 1996 Reliability Test System [76]. These 6 datasets represent weekday and weekend peak load for the Winter, Summer and Spring/Fall seasons and their inclusion in the data library provides the modeler with a set of well-researched datasets that can be used to test multi-period models.

To allow for a more comprehensive modeling environment, these traditional datasets have been augmented with information that enable the modeler to consider more complex features of power systems modeling. The following sections discuss two key data augmentations that are available with the toolbox. Section 1.3.1 discusses generator capability curves (D-curves) which provide a more accurate representation of a generator's production capabilities and section 1.3.2 discusses our implementation of demand bidding which considers price elastic response from consumers.

¹Model descriptions and usage details of the data utility archive are provided in Appendix B

1.3.1 D-curves

Simplistic generator models often use “rectangular constraints” for active and reactive output limits, which imposes independent limits on active and reactive power output. Although the rectangle constraints model provides a good first approximation, more detailed modeling is necessary to accurately characterize generator capability curves, also called “D-curves”. The reactive power output of a synchronous generator is constrained by several factors: armature current limit, field current limit, and end region heating limit [49]. Each of these limits is modeled as circles in the active-reactive power output plane and the machine must operate within the intersection of these circles in addition to their regular operating bounds. Fig. 1 from reference [45] (reproduced as Fig. 1.1 in this document) shows the capability curve for a typical synchronous generator. The upper portion of the curve is the circle from the field current limit, the right portion of the curve is the circle from the armature current limit, and the lower portion of the curve is the circle from the end region heating limit.

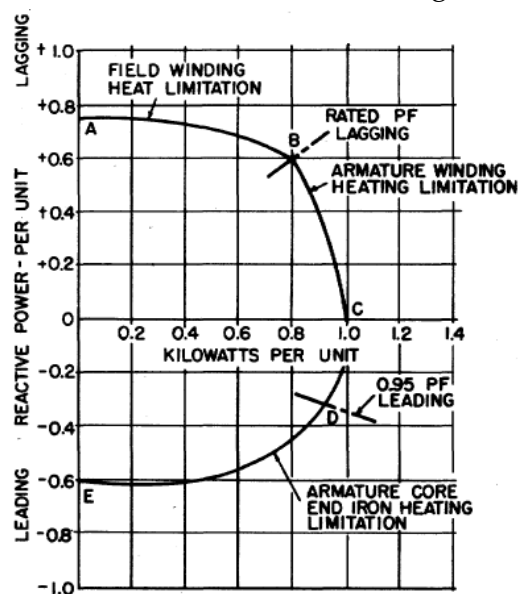


Figure 1.1: Typical Generator Capability Curve (Figure 1 from [45])

Dataset	Time		Objective	
	Standard	D-curves	Standard	D-curves
case14	0.3548s	0.424s	8.08153e+03	8.09162e+03
case118	0.734s	0.704s	1.29661e+05	1.29913e+05
case300	1.2314s	0.871s	7.19725e+05	7.20176e+05
case2737sop	7.9692s	8.689	7.77629e+05	7.77649e+05
case3120sp	11.9338s	11.279s	2.14270e+06	2.15042e+06
case3375wp	16.9038s	23.046s	7.41203e+06	7.43363e+06

Table 1.4: Rectangular vs. D-curve constraints

By using the approximation method outlined in Appendix D, the toolbox provides a D-curve such as that of Figure 1.2 whenever possible. As the image clearly shows, the application of D-curve constraints reduces the feasible region and will therefore affect the solution point. Table 1.4 provides initial numerical results using the D-curve approximations provided in this section, and the higher objective values echo the results above showing a reduced feasibility region. As expected, the addition of these constraints do impact the solution time and more research is needed to understand the importance (or not) of considering these details within an OPF model.

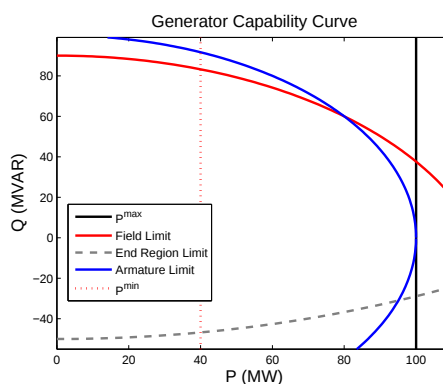


Figure 1.2: Example of D-curve using the toolbox

1.3.2 Demand bidding

Demand bidding is a demand-side decision that models a demand node's price-responsive behavior where consumers reduce their planned consumption in exchange for a monetary compensation from the ISO. While a common practice, this data is not generally available in IEEE datasets. According to [51], time-varying prices obtained from Electric Reliability Council of Texas(ERCOT) resulted in a peak price elasticity of -0.02 for 6% of the ERCOT demand. That is, for the 6% of ERCOT's load is considered price responsive, doubling the price results in a short-term demand reduction of 2% of peak load. These results indicate that electric demand is very price inelastic. The experiments surveyed in [28] show a substantially larger elasticity than that estimated for ERCOT in the PSERC study. This may be due to the fact that [28] only considers residential demand, while the PSERC study is likely to include a large proportion of industrial and commercial demand. It therefore may be appropriate to scale the y-axis of the curve on [28, p.45] to obtain price-responsive demand curves for other customer types if corresponding curves for industrial and commercial customers are unavailable.

We have therefore chosen to approximate the price-responsive curve using [51] as a baseline, with a little additional elasticity built in to account for the results discussed in [28]. By applying procedure 1, this produces the a 9 piece (10 breakpoints) piece-wise linear demandbid curve as shown in 1.3.

Procedure 1: Demandbid curve approximation

// Solve ACOPF to obtain $LMP_{it} \forall i \in N, t \in T$*

Initialization: $qlen_{it} \leftarrow d_{it}^P, \hat{p}_{it} = 0, p_{it}^b = q_{it}^b = 0$ for $b = 1$

1 **Loop for** $b = 2..10$

2 $q_{it}^b \leftarrow qlen_{it} * (b - 1)$

3 $p_{it}^b \leftarrow \hat{p}_{it} + qlen_{it} * (LMP_{it} * (11 - b) / 9)$

4 $\hat{p}_{it} \leftarrow p_{it}^b$

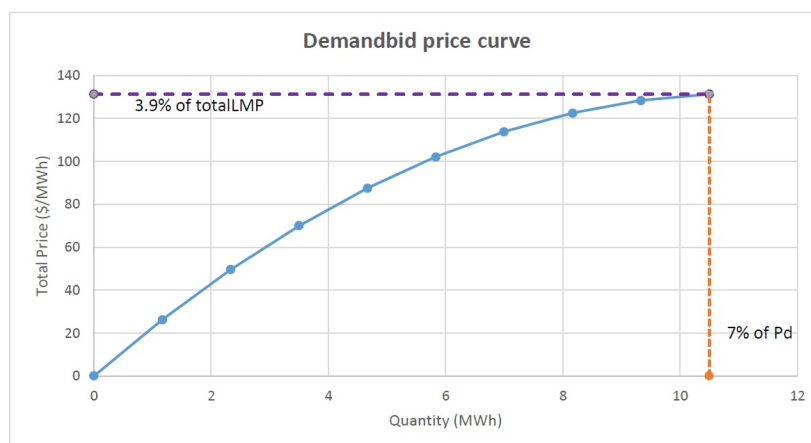


Figure 1.3: Example of Demandbid curve using the toolbox

1.4 Numerical experiments and observations

1.4.1 Exploring different AC formulations

As discussed in 1.2.2, the toolbox focuses on 3 main OPF formulations: polar (P), rectangular power-voltage (R) and rectangular current-voltage (IV). The rectangular formulation is often favored due to the quadratic structure of its constraints which can be exploited more easily over the polar form of quantities that explicitly uses sines and cosines.

Dataset	Polar	Rect-PV	Rect-IV
case118	0.702s	0.757s	0.843s
case300	1.2314s	1.339s	1.369s
case2737sop	7.9692s	9.483s	9.357s
case3120sp	11.9338s	14.411s	12.269s
case3375wp	16.9038s	15.553s	36.412s
rts96_winter_wend (UCAC)	15.931s	33.981s	infeas

Table 1.5: Comparison of OPF formulations

Based on the results in Table 1.5, the polar formulation consistently outperforms both the other two and the variance in solution time becomes clearer as the size of dataset becomes larger. This result was consistent

even in the unit commitment case of `rts96_winter_wend` which is models the winter weekend peak-load over 24 hours. While these results are promising, it should not discount the possibility of exploiting model structure in both rectangular and IV formulations especially when considering decomposition models. What the results do show is that current solvers favor the polar formulation when solving a standard AC model.

Note that these results were consistent over initial conditions and solver types which will be discussed in Sections 1.4.2 and 1.4.3. For the remainder of the results, we will be focusing on the polar formulation for brevity.

1.4.2 Initial starting conditions

Due to the difficulty of solving AC models, we have provided multiple starting point options for the SLP models where the variables are initialized using different methods. This is determined by the GAMS model option `ic=#`, and takes inputs 1, 2... 9. We explain some of these options below.

1. `ic=0`: [Midpoint]

$$V_i = \frac{\bar{V} + \underline{V}}{2}, \theta_i = 0 \quad \forall i \in \mathcal{N}$$

$$P_k = \frac{\bar{P}_k + \underline{P}_k}{2}, Q_k = \frac{\bar{Q}_k + \underline{Q}_k}{2} \quad \forall k \in \mathcal{G}$$

2. `ic=1`: [Random]

$$V_i = \text{Uniform}(\bar{V}, \underline{V}), \theta_i = \text{Uniform}(-\pi, \pi) \quad \forall i \in \mathcal{N}$$

$$P_k = \text{Uniform}(\bar{P}_k, \underline{P}_k), Q_k = \text{Uniform}(\bar{Q}_k, \underline{Q}_k) \quad \forall k \in \mathcal{G}$$

3. `ic=2`: [Flat]

$$V_i = 1, \theta_i = 0 \quad \forall i \in \mathcal{N}$$

$$P_k = 0, Q_k = 0 \quad \forall k \in \mathcal{G}$$

4. `ic=3`: [Random/AC]

$$V_i = \text{Uniform}(\bar{V}, \underline{V}), \theta_i = \text{Uniform}(-\pi, \pi) \quad \forall i \in \mathcal{N}$$

P_k, Q_k derived from node balance constraints (model dependent)
with V_i, θ_i defined $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

5. ic=4: [DC/AC]

$$V_i = 1 \quad \forall i \in \mathcal{N}$$

θ_i, P_k initialized solving DCOPF (1.2-1.6) $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

Q_k derived from node balance constraints (model dependent) with

V_i, θ_i defined $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

6. ic=5: [DC-/AC]

$V_i = 1, \theta_i$ initialized solving DCOPF (1.2-1.6) $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

P_k, Q_k derived from node balance constraints (model dependent)

with V_i, θ_i defined $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

7. ic=6: [Decoupled]

V_i, θ_i, P_k, Q_k initialized by solving the Decoupled AC model (refer to Appendix C).

8. ic=7: [DCLoss]

$$V_i = 1 \quad \forall i \in \mathcal{N}$$

θ_i, P_k initialized solving DCOPF (1.2-1.6) with option `--lineloss=1.055`
 $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

Q_k derived from node balance constraints (model dependent) with

V_i, θ_i defined $\forall i \in \mathcal{N}, k \in \mathcal{G}$.

Table 1.6 compares the impact that the different initial conditions have on the solution time of the polar ACOPF model on different datasets. Here, INFES means that the solver terminated at a locally infeasible point, while X indicates that the solver failed to reach its termination point or return a solution within the time frame allotted (1000 seconds), and F indicates that a feasible solution was returned but the solver terminated at the 1000 second limit. The expectation here is that more intuitive starting points such as ic=4-7 would outperform the simpler starting point options ic=0-3, especially in the larger datasets where finding a feasible point can be alot more challenging. While the results do show that the efficiency of DC

starts (ic=4,5,7) improves as the size of network increases, these methods still fail to outperform the default “midpoint” option ic=0. Another point of note is the badly performing ic=3 in comparison to ic=1. Initial condition ic=3 is supposed to improve upon the random draw option (ic=1) by only using random draws to initialize the minimal subset of variables and deriving starting points of the remaining variables from from node balance constraints, such as equations (1.26 - 1.27) in the case of the polar formulation. This is not shown to be the case in table 1.6 as ic=3 consistently terminates at infeasible solutions which the random starting point does not seem to suffer from.

Testcase	case14	case118	case300	2383	3012	3120	3375
ic=0	0.06	0.44	0.81	10.46	11.63	12.26	14.43
ic=1	0.05	0.34	0.88	33.98	46.06	103.80	183.78
ic=2	0.05	0.56	1.72	18.82	33.72	34.28	INFES
ic=3	INFES	X	INFES	X	X	X	X
ic=4	0.05	0.27	0.76	7.69	11.41	9.75	15.71
ic=5	0.03	0.27	0.62	8.47	12.62	11.58	16.69
ic=6	0.05	0.23	X	X	X	X	X
ic=7	0.04	0.22	0.62	8.12	12.09	10.50	18.43

Table 1.6: Comparison of initial conditions, Polar ACOPF

Initial condition 6 (Decoupled ACOPF) does not seem to work very well even in small cases, and more research needs to be carried out to look into this issue. Notice also that in the case of flat start (ic=2), it works very well in small sized cases, but is unable to manage the largest IEEE dataset which is case3375wp. This merely illustrates a known fact, that algorithms heuristics tested in small scale environments are not good indicators in realistic sized models. Chapter 2 focuses on this further by providing large-scale datasets with realistic information that can be used with the toolbox.

1.4.3 Solver performances

Numerical results from the previous sections were obtained using IPOPTH but we also tested the performance of different solvers AC models such as KNITRO, CONOPT, SNOPT and MINOS. A condensed version of these results are shown in Table 1.7.

KNITRO							
Testcase	case14	case118	case300	2383	3012	3120	3375
ic=1	0.12	0.36	1.03	INFES	INFES	X	INFES
ic=2	0.1	0.50	2.60	190.77	957.89	558.73	455.41
ic=4	0.10	0.33	0.86	11.51	X	482.68	13.2
ic=5	0.10	0.35	0.83	621.94	20.15	61.37	19.95
ic=7	0.08	0.35	0.943	10.37	15.65	15.34	20.56
CONOPT							
Testcase	case14	case118	case300	2383	3012	3120	3375
ic=1	0.21	1.58	9.40	961.45	921.17	X	924.90
ic=2	0.04	2.55	23.16	145.95	400.43	435.92	X
ic=4	0.01	0.83	4.45	135.02	107.61	260.76	581.83
ic=5	0.01	1.02	14.00	156.45	200.29	140.28	781.27
ic=7	0.01	0.85	6.61	161.48	223.12	111.01	326.69
SNOPT							
Testcase	case14	case118	case300	2383	3012	3120	3375
ic=1	0.06	1.89	6.85	891.97	X	X	X
ic=2	0.05	2.28	6.05	439.59	469.18	782.32	X
ic=4	0.04	1.78	6.36	516.83	441.945	780.64	X
ic=5	0.04	1.76	5.75	200.50	81.11	536.38	X
ic=7	0.04	1.92	6.31	541.88	378.65	732.57	F

Table 1.7: Comparison of solver performance

The results clearly show that in addition to being fast, IPOPTH is also

the most reliable in solving these types of problems. Consider for example, the numerical results with KNITRO in $ic=5$ for testcase 2383. It takes KNITRO over 5 minutes to solve that problem, but only takes a minute to solve the next 3 larger datasets. CONOPT, while fairly reliable, takes too long to solve these datasets, and was unable in fact to solve the large datasets in Chapter 2 without extremely good or lucky starting points.

1.5 Sparse Tableau Formulation

The vast majority of texts formulate AC network constraints based on the bus admittance matrix Y_{bus} , which represents a choice of “strict” nodal analysis (NA) from a circuit-theoretic perspective [60]. The Y_{bus} matrix encapsulates network information in a compact form by eliminating intermediate variables of branch element currents and voltages. In many classes of circuit analysis, elimination of variables can have drawbacks[56]:

1. The Y_{bus} lumps multiple network elements together, sometimes masking the individual role of network element parameters in defining Y_{bus} entry values.
2. Changes in circuit breaker status may require rebuilding the whole Y_{bus} , possibly with different dimension.
3. Standard Y_{bus} construction algorithms often require that models of individual grid components be represented as interconnections of ideal two-terminal circuit elements, sacrificing the flexibility of multi-port models
4. Ill-conditioning may be introduced via roundoff errors due to the elimination of intermediate variables

To overcome the limitations of the Y_{bus} , recent work [60] has proposed a very general modeling approach for electric grid representation using Sparse Tableau Formulation (STF) [58].

Sparse Tableau Formulation is a very general node-breaker formulation that provides the flexibility to capture atypical Network elements which cannot be represented by the Y_{bus} construction. Additionally, we will see that this formulation, when used within Optimal Power Flow (OPF) models, does not sacrifice computational efficiency and is comparable in solution time to the well-established Y_{bus} formulations provided in the toolbox. Given the growing need to capture intricacies in grid optimization problems, we believe that including this formulation in the OPF Toolbox can further advance future power systems research. As such, we have added an extension to the toolbox that can compute and add the necessary STF data components to the .gdx datafiles, and provided two STF based Optimal Power Flow (OPF) models in the model archive.

The Sparse Tableau Formulation and models discussed in this Chapter are a result of the work performed in [58, 60, 59], and heavily source their material. Their descriptions of formulation and models are included for reference purposes, and serve to clarify the STF formulation and its impact on Power Systems research.

1.5.1 Overview on Constructing the STF Circuit

Here, we review the key steps in constructing STF circuit constraint equations [50]. As case of interest in power systems, we give special attention to two-port circuit elements, and assume that the circuit analysis is conducted with respect to complex phasor branch or port currents, denoted i , complex phasor branch or port voltages, denoted v , and complex node voltages, denoted V . With this complete set of circuit quantities maintained as explicit variables, the first key advantage of STF follows: it provides a very easy and clear separation of the three classes of equality constraints

governing any circuit: (i) KCL, (ii) KVL, (iii) element constitutive equations.

Step 1. Write a complete set of linearly independent Kirchhoff's current law (KCL) equations at every bus, employing the node-to-element incidence matrix A :

$$Ai = I \quad (1.55)$$

Step 2. Write a complete set of linearly independent Kirchhoff's voltage law (KVL) equations:

$$v - A^T V = 0 \quad (1.56)$$

Step 3. Write the element constitutive equations. Consider the representative example of transmission lines and transformers in balanced operation, in which case the resulting behavior of each element is well-captured in a linear two-port model [9]. Presentation to follow will generalize to a wider class for elements, but for now restricting to these linear transformer and transmission line models, element constitutive relations may be written in the simple form:

$$F_v v + F_i i = 0 \quad (1.57)$$

While additional "driving" sources must be subsequently be added as non-zero terms on the right hand side of (1.55), (1.56), and (1.57), these equations form the core of the tableau equations. For an element represented as two-port, v and i quantities appear in "port-pairs," as illustrated in Fig. 1.4.

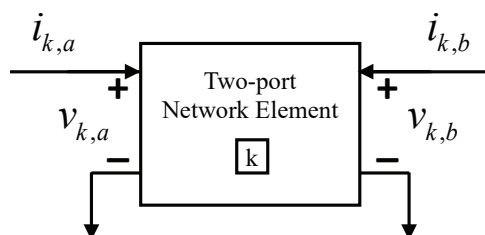


Figure 1.4: Two-port representation for network element k

To further examine element constitutive relations, consider the transmission line, perhaps the most common of power system circuit components. In this context, port a quantities are typically “sending end” positive sequence voltage and current, and port b quantities are “receiving end.” The two-port element’s constitutive relations place two independent algebraic constraints on the four variables (v_a, i_a, v_b, i_b) to specify the element’s behavior. If one wishes to allow for a general, nonlinear two-port element k , in the case of complex, phasor-based analysis, these constraints take the general implicit form

$$f_k : \mathbb{C}^4 \longrightarrow \mathbb{C}^2 \quad (1.58)$$

$$f_k(v_{k,a}, i_{k,a}, v_{k,b}, i_{k,b}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Typical power system elements within the transmission network are linear, passive, and time-invariant; however, the formulation retains its structure even if one allows for more general nonlinear or active transmission components.

DEFINITION. (The Tableau Matrix) Consider (1.55), (1.56) and (1.57) that constitute the tableau equations of a power system transmission network; these may be conveniently summarized in the single matrix equation

$$\underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & A \\ -A^T & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & F_v & F_i \end{bmatrix}}_{\mathbf{T}} \underbrace{\begin{bmatrix} V \\ v \\ i \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{u}} \quad (1.59)$$

Consistent with the method’s name, the tableau matrix T in (1.59) is extremely sparse (notably more sparse than Y_{bus}).

1.5.2 STF Network Elements Modeling

A key step to construct (1.59) for a power network requires consideration of the standard power system network circuit elements. Again, we note that many such elements are naturally considered as two-ports, as depicted in Fig. 1.4. Yet in power flow modelling, transmission lines are instead specified in terms of their π -equivalent circuit, with their data is provided as the three real-value parameters R , X and B , with associated complex series impedance $Z = R + jX$, and shunt $Y = jB$. Standard textbook derivations (e.g. [9]) demonstrate that at the sending end and receiving end ports, terminal behavior matching that of the [ABCD]-transmission matrix is indeed captured by this π -equivalent, as shown in Fig. 1.5. However, in the pedagogical approach being advocated here, one may simply ask: “why bother with this conversion to the π -equivalent - why not just use the two-port [ABCD] matrix directly?”

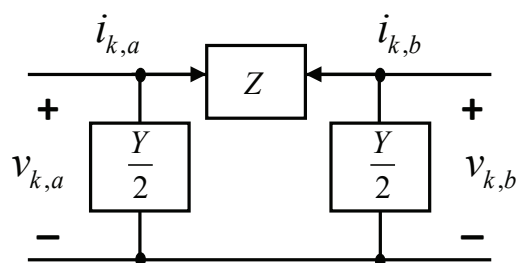


Figure 1.5: π -equivalent circuit for transmission line

Accepting that the long-established conventions for specifying transmission line data are unlikely to change, one may use the parameters R , X and B to recover the two-port constitutive relation consistent with the

[ABCD]-transmission matrix, and with (1.58), as follows:

$$\begin{aligned} \begin{bmatrix} v_{k,a} \\ i_{k,a} \end{bmatrix} &= \begin{bmatrix} 1 + \frac{ZY}{2} & Z \\ Y(1 + \frac{ZY}{4}) & 1 + \frac{ZY}{2} \end{bmatrix} \begin{bmatrix} v_{k,b} \\ -i_{k,b} \end{bmatrix} \\ \Rightarrow \underbrace{\begin{bmatrix} [cc|cc]1 & 0 & -(1 + \frac{ZY}{2}) & Z \\ 0 & 1 & -Y(1 + \frac{ZY}{4}) & (1 + \frac{ZY}{2}) \end{bmatrix}}_{\text{Representative circuit element } F_k \in \mathbb{C}^{2 \times 4}} \begin{bmatrix} v_{k,a} \\ i_{k,a} \\ v_{k,b} \\ i_{k,b} \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned} \quad (1.60)$$

To formulate equation (1.60) in a manner consistent with STF (1.59), one simply re-writes as

Linear Element Equation for transmission line:

$$\underbrace{\begin{bmatrix} 1 & -(1 + \frac{ZY}{2}) \\ 0 & -Y(1 + \frac{ZY}{4}) \end{bmatrix}}_{\mathbf{F}_{k,v}} \begin{bmatrix} v_{k,a} \\ v_{k,b} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & Z \\ 1 & (1 + \frac{ZY}{2}) \end{bmatrix}}_{\mathbf{F}_{k,i}} \begin{bmatrix} i_{k,a} \\ i_{k,b} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (1.61)$$

This yields the corresponding F_v and F_i matrices for the transmission line. The other typical network element is a transformer. Here, voltage gain of transformer is expressed as complex scalar n , allowing for the possibility of phase shifting transformers. Then, the corresponding transmission matrix representation is

$$\begin{bmatrix} v_{k,a} \\ i_{k,a} \end{bmatrix} = \begin{bmatrix} n & 0 \\ 0 & \frac{1}{n^*} \end{bmatrix} \begin{bmatrix} v_{k,b} \\ -i_{k,b} \end{bmatrix} \quad (1.62)$$

This can be equivalently re-written as

Linear Element Equation for transformer:

$$\underbrace{\begin{bmatrix} 1 & -n \\ 0 & 0 \end{bmatrix}}_{F_{k,v}} \begin{bmatrix} v_{k,a} \\ v_{k,b} \end{bmatrix} + \underbrace{\begin{bmatrix} [c|c]0 & 0 \\ 1 & \frac{1}{n^*} \end{bmatrix}}_{F_{k,i}} \begin{bmatrix} i_{k,a} \\ i_{k,b} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (1.63)$$

For each of these cases, it is straightforward to identify the corresponding constant F_v and F_i matrices. Note that for these very typical network elements, (1.61) and (1.63) contain only constant coefficients and have no independent sources.

1.5.3 Construction of the incidence matrix A

Remaining constraints are simple linear expressions imposing KVL and KCL interconnection constraints. Since a node-to-element incident matrix $A \in \mathbb{R}^{N \times 2l}$ is defined over all network elements where N and l are number of network buses and elements, we need to organize all network element variables (port voltages and port currents):

$$v \triangleq \begin{bmatrix} [c]v_{1,a} \\ v_{1,b} \\ \vdots \\ v_{l,a} \\ v_{l,b} \end{bmatrix}, \quad i \triangleq \begin{bmatrix} [c]i_{1,a} \\ i_{1,b} \\ \vdots \\ i_{l,a} \\ i_{l,b} \end{bmatrix}$$

Thus, $v, i \in \mathbb{C}^{2l}$. Goal of KCL is to efficiently assemble the right-hand side of the general current balance equation. To this end, the incidence matrix is composed entirely of values of 1 or -1 or 0. In keeping with

standard textbook presentations, we define:

$$A(j, r) \triangleq \begin{cases} 1/ -1, & \text{if } r\text{th component of } i \text{ corresponds to an} \\ & \text{elements' sending or receiving terminal} \\ & \text{leaving/entering node } j \\ \hline 0, & \text{otherwise} \end{cases} \quad (1.64)$$

Therefore, the current conservation law of KCL is written simply as

$$Ai = I \quad (1.65)$$

where $I \in \mathbb{C}^N$ is the nodal complex current injection from generators or loads; $i \in \mathbb{C}^{2l}$ is the complex branch current carried away from node by network elements. We can also use A to relate port voltages to bus voltages in a manner that guarantees KVL is automatically satisfied. So the linear constraints of KVL are written as

$$v - A^T V = 0 \quad (1.66)$$

where $V \in \mathbb{C}^N$ is bus voltages. The equation (1.66) serves to assign the correct bus voltage to any port voltage of a port connected to that bus. Now, to construct the sparse tableau matrix (1.59), $F_v \in \mathbb{C}^{2l \times 2l}$ and $F_i \in \mathbb{C}^{2l \times 2l}$ can be constructed in block diagonal form as

$$F_v = \begin{bmatrix} F_{1,v} & 0 & \cdots & 0 \\ 0 & F_{2,v} & 0 & : \\ : & : & \ddots & 0 \\ 0 & \cdots & \cdots & F_{l,v} \end{bmatrix}, \quad F_i = \begin{bmatrix} F_{1,i} & 0 & \cdots & 0 \\ 0 & F_{2,i} & 0 & : \\ : & : & \ddots & 0 \\ 0 & \cdots & \cdots & F_{l,i} \end{bmatrix} \quad (1.67)$$

where F_v and F_i are block diagonal matrices composed of previously described $F_{k,v}$, $F_{k,i}$. Finally, with matrices F_v , F_i , A and variables v , i , V , I

as defined above, we have STF description of the power network consistent with (1.59).

The remaining step to capture the power flow equations requires specification of the current source elements I . The approach here introduces a special class of nonlinear one-port element, as shown in Fig. 1.6, to accommodate standard generator and load models as represented in the power flow.

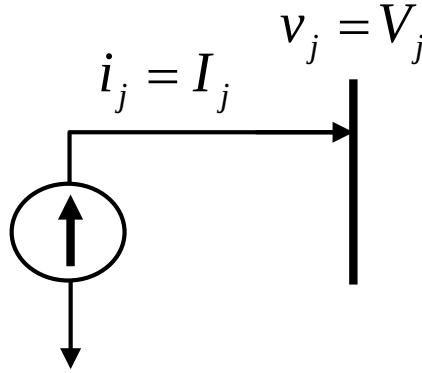


Figure 1.6: Nonlinear current source element as one port element

Then, nonlinear element equation as an equation (1.58) for current source I_j for bus j can be defined by

Nonlinear Element Equations:

$$f_j(v_j, i_j) = 0 \triangleq i_j = \frac{(S_{g,j} - S_{d,j})^*}{v_j^*} \quad (1.68)$$

Notice that $S_j = S_{g,j} - S_{d,j}$, $i_j = I_j$, and $v_j = V_j$ implying

$$\implies I_j - \frac{S_j^*}{V_j^*} = 0 \quad (1.69)$$

where $S_{g,j}$ and $S_{d,j}$ are specified apparent power generation and load at bus j .

1.5.4 Optimal Power Flow Model with STF

Given a connected power system with $\mathbf{N}$ the set of all buses, $\mathbf{G}$ the set of all generators and $\mathbf{L}$ set of all transmission line, a compact form of the OPF Formulation is as following:

$$\min_{P,Q,v,i,V,I} \sum_{j \in \mathbf{G}} \tilde{c}_j(P_{g,j}) \quad \text{subject to} \quad (1.70a)$$

$$\text{Linear Element: } F_v v + F_i i = 0 \quad (1.70b)$$

$$\text{KCL: } I - Ai = 0 \quad (1.70c)$$

$$\text{KVL: } v - A^T V = 0 \quad (1.70d)$$

$$\text{Nonlinear Element: } S - V \odot (I)^* = 0 \quad (1.70e)$$

$$\text{Gen. Limit: } P_j^{min} \leq P_{g,j} \leq P_j^{max} \quad (1.70f)$$

$$Q_j^{min} \leq Q_{g,j} \leq Q_j^{max}, \forall j \in \mathbf{G} \quad (1.70g)$$

$$\text{Vol. Limit: } V_j^{min} \leq |V_j| \leq V_j^{max}, \forall j \in \mathbf{N} \quad (1.70h)$$

$$\text{Line Limit: } |i_{k,a/b}| \leq i_k^{max}, \forall k \in \mathbf{L} \quad (1.70i)$$

The GAMS model uses the rectangular form of complex variables for bus voltage and current value as $V_c = V_c^d + \mathbf{j}V_c^q$ and $I_c = I_c^d + \mathbf{j}I_c^q$. The nonlinear element equations for power balance constraints are $S_j = V_j(I_j)^*$. In rectangular coordinate, it is equivalent to $S_j = V_j(I_j)^* = V_j^d I_j^d - \mathbf{j}V_j^d I_j^q + \mathbf{j}V_j^q I_j^d + V_j^q I_j^q$.

The modeling of a transmission line with a transformer can be done by combining the data from both elements. Doing so reduces the problem size and can lead to faster solution times. Using transmission matrix representation (TR), parameters for the line that has a transformer (TR_{total})

can be expressed as

$$TR_{total} = \underbrace{\begin{bmatrix} 1 + \frac{ZY}{2} & Z \\ Y(1 + \frac{ZY}{4}) & 1 + \frac{ZY}{2} \end{bmatrix}}_{TR_{line}} \underbrace{\begin{bmatrix} n & 0 \\ 0 & \frac{1}{n^*} \end{bmatrix}}_{TR_{transformer}} = \begin{bmatrix} \alpha & \beta \\ \zeta & \tau \end{bmatrix}$$

Through here, we can generate the necessary computational elements $\alpha, \beta, \tau, \zeta$ using matrix algebra and thus, the GAMS formulation is as follows.

Nomenclature

Set	Description
B	Set of circuit breakers between buses.
C	Set of contingency cases
G	Set of generators.
$G_j \in G$	Subset of generators G at bus j .
K	Set of network elements.
L	Set of lines in the transmission network.
N	Set of buses.
$N_g \in N$	Subset of buses having generation attached.

Table 1.8: Description of Sets

Parameters	Description
$\tilde{c}_j(\cdot)$	Cost function for generator $j \in G$.
g_j^s, b_j^s	Shunt conductance and susceptance at bus $j \in N$.
A	Node-to-element incidence matrix.
F_v, F_i	Matrices for elements' $v - i$ behavior.
$I_k^{\max}$	Current flow limit on line $k \in L$.
$P_{d,j}, Q_{d,j}$	Real, reactive power demand at bus $j \in N$.
$P_j^{\min}, P_j^{\max}$	Lower/upper active power generation limits for generator $j \in G$.
$Q_j^{\min}, Q_j^{\max}$	Lower/upper reactive power generation limits for generator $j \in G$.
$V_j^{\min}, V_j^{\max}$	Lower/upper voltage magnitude limits at bus $j \in N$.
η_j	Droop coefficient for generator $j \in G$.

Table 1.9: Description of Functions and Parameters

Decision Variables	Description
P_g	Active power generation dispatch.
Q_g	Reactive power generation.

Table 1.10: Description of Decision Variables

Rectangular Sparse Tableau Formulation Model

$$\min_{P_g, Q_g, v, i, V, I, \Delta\omega} \sum_{j \in G} \tilde{c}_j(P_{g,j}) \quad \text{subject to} \quad (1.71a)$$

Transmission line status (TLS):

$$\begin{bmatrix} 1 & -\alpha_k \\ 0 & -\zeta_k \end{bmatrix} \begin{bmatrix} v_{k,a} \\ v_{k,b} \end{bmatrix} + \begin{bmatrix} 0 & \beta_k \\ 1 & \tau_k \end{bmatrix} \begin{bmatrix} i_{k,a} \\ i_{k,b} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, k \in L \quad (1.71b)$$

Kirchhoff's voltage law (KVL):

$$\text{real}(v) = A^T V^d \quad (1.71c)$$

$$\text{imag}(v) = A^T V^q \quad (1.71d)$$

Active power balance:

$$\begin{aligned} V_j^d I_j^d + V_j^q I_j^q + g_j^s \left((V_j^d)^2 + (V_j^q)^2 \right) \\ = \left(\sum_{r \in \mathbf{G}_j} P_{g,r} - \eta_r \Delta \omega \right) - P_{d,j} \end{aligned} \quad j \in N \quad (1.71e)$$

Reactive power balance:

$$\begin{aligned} -V_j^d I_j^q + V_j^q I_j^d - b_j^s \left((V_j^d)^2 + (V_j^q)^2 \right) \\ = \sum_{r \in \mathbf{G}_j} Q_{g,r} - Q_{d,j} \end{aligned} \quad j \in N \quad (1.71f)$$

Generator output limit:

$$P_j^{\min} \leq P_{g,j} - \eta_j \Delta \omega \leq P_j^{\max} \quad \forall j \in G \quad (1.71g)$$

$$Q_j^{\min} \leq Q_{g,j} \leq Q_j^{\max} \quad \forall j \in G \quad (1.71h)$$

Voltage magnitude limit:

$$V_j^{\min} \leq \sqrt{(V_j^d)^2 + (V_j^q)^2} \leq V_j^{\max} \quad \forall j \in N \quad (1.71i)$$

Line thermal limit:

$$(i_{k,a/b}^d)^2 + (i_{k,a/b}^q)^2 \leq (I_k^{\max})^2 \quad \forall k \in L \quad (1.71j)$$

1.5.5 Sparse Tableau Formulation in the OPF Toolbox

In the data utility section of the OPF toolbox, we have provided a script that takes a .gdx case file, generates the F-matrices (F_v and F_i 1.67) and incidence matrix A 1.64. These, along with other necessary modeling components, are then added into the .gdx case file, and can be extracted for use with STF models. The F-matrix generated and stored in the casefiles is the F-maxtrix for each individual element, as opposed to a combined matrix as shown in 1.71a. The reason for this is to allow for flexibility in how a user would like to utilize the STF data.

To highlight this, the OPF toolbox includes two STF models. First is what we refer to as the "Separate STF model", where each element (line and transformers) is modeled separately in its own constraint. The benefit of this is a cleanly formulated model and data source (with a single F-matrix to represent all elements). However, pre-computing a combined line + transformer F-matrix in 1.71a reduces the dimensionality of the problem and does lead to faster solve times as a result. Therefore, we refer to the model which uses the combined line + transformer F-matrix as simply the "Rectangular STF model" as we expect that to be the standard STF model used within the toolbox. To go from the "general F-matrix" .gdx data to the combined line + transformer data, the STF model simply has a translation script to generate the combined F-matrix from the general F-matrix.

1.5.6 Empirically Observed Computation Results

This section compares the solution time between the "STF model" to the polar power-voltage formulation (P) and Rectangular Power-Voltage Formulation (R) discussed in Section 1.2.2 of Chapter 1. Computational time is evaluated on standard IEEE instances available from MATPOWER including one synthetic system produced under the ARPA-E GRID DATA

project, the EPIGRIDS 1600-bus system [36]. Note that the Rectangular STF formulation solve times is comparable in performance.

Table 1.11: Comparison of STF OPF to standard ACOPF solve times

	POLAR- Y_{bus}		REC-IV- Y_{bus}		REC-STF	
	Obj (\$/h)	Time (sec)	Obj (\$/h)	Time (sec)	Obj (\$/h)	Time (sec)
case118	129660.68	0.3	129660.68	0.3	129660.68	0.3
case300	719725.07	0.6	719725.07	2	719725.07	0.6
case2383wp	1868511.82	5.6	1862367.02	5.2	1862367.02	6.3
case3012wp	2591706.57	5.8	2582670.47	5.9	2582670.47	6
case3120sp	2142703.76	5.8	2141532.10	6.3	2141532.10	6.5
case3375wp	7412030.67	54	7404635.99	11.7	7404635.99	11.4
EPIGRIDS 1600	212840.95	3.5	212882.35	2.4	212882.35	2.5

1.6 Conclusions

In conclusion, this toolbox is useful for carrying out rigorous testing in a standardized fashion, and by providing interfaces with Matlab and PSS/E, this allows the user to import existing research and work into this environment. In addition, we believe this toolbox encourages power systems researchers to take advantage of the rich features within the GAMS environment, which also provides the flexibility to experiment with pre-existing and well-established solvers. What we have not illustrated here but plan to add, are examples that extend the traditional OPF formulation, such as stochastic unit commitment (SUC) or security constrained unit commitment (SCED) models, which is easily implemented using the EMP tool in GAMS. Additionally, including examples of planning models such as the one discussed in Chapter 3 will provide the user with good illustrations of this toolbox's utility as an asset in complex power systems modeling.

2 LARGE-SCALE DATASETS IN POWER SYSTEMS

While the primary structure of the ACOPF model has undergone little change since its formulation in 1962, it remains a current research topic as these models are difficult to solve in large-scale. In fact, these models are solved in industry every few minutes using approximations or judgment as there is not yet a commercially viable full ACOPF model for practical instances [17]. Despite the spotlight on solving these types of models, realistic large-scale datasets are difficult to come by. In a survey of ACOPF solution approaches [21, Table 2], test problems ranged from 14-bus systems to 42000 buses. Of the 60 test problems recorded, the largest publicly available data source was the single period 3120 IEEE Polish system. The largest 3, and arguably, the only realistic-sized test problems ($> 10,000$ buses), do not have publicly sourced data and only appear in 2 of the publications surveyed. This reflects a lack of testing on realistic sized datasets which can be attributed to lack of availability. Recently, the EPIGRIDS project has been working on constructing synthetic power grid datasets that mimic the characteristics of the actual grid, but do not disclose sensitive information. The data utility scripts have been updated to incorporate the content coming out of the EPIGRIDS project, and this is further discussed in Section 2.6 of Chapter 5.

In this chapter, I discuss the compilation of 4 large-scale datasets that represent realistic network and operational conditions, and that are compatible with the OPF toolbox from Chapter 1, difficulties in dealing with missing and questionable data, and discuss numerical results and quality assurance of solving ACOPF models in large datasets. The goal here is to test the various model formulations and commercial solvers with truly large-scale datasets and consider the complexities in finding solutions for real life applications.

As an extension after the initial Toolbox publication, we have added

new synthetic but realistic datasets from the EPIGRIDS [1] project into the data archive, which we discuss in Section 2.6. We detail the specific data utility tool updates that were made to accommodate this extension in 2.6.1, and briefly discuss the AC powerflow feasibility issues we encounter with this collection of datasets in 2.6.2.

2.1 Introduction

Four datasets were compiled using a mix of public and confidential data. The underlying network information (i.e. transmission lines and connections) is provided under Critical Energy Information Infrastructure (CEII) usage agreements as detailed by FERC. The remainder of the data was based on publicly available information provided in the FERC e-Library, but requires a bus-mapping to the network which was also covered under the CEII agreement. Table 2.1 lists some basic statistics of the 4 datasets. The regular and RTO versions, while relying on the same network and generator information, differ only in the number of time periods provided and their load-side demand data. More precisely, the regular versions provide 24-hour demand data while the RTO versions provide a different set of demand data but only for 6 discontinuous periods. One can view the regular Winter and Summer Peak datasets (henceforth referred to as Winter and Summer) as being day-ahead planning information, while the RTO datasets (henceforth referred to as Winter/Summer RTO) provide a snapshot of the actual operating constraints realized on a particular day.

Testcase	Network	Num. bus/gen/lines	Time periods
Winter Peak	Winter	13867/1043/18790	1 - 24
Winter Peak RTO	Winter	13867/1043/18790	1,5,9,13,17,22
Summer Peak	Summer	13981/1011/18626	1 - 24
Summer Peak RTO	Summer	13981/1011/18626	1,5,9,13,17,22

Table 2.1: Large-scale Data: Statistics

Similar to the IEEE data files referenced Chapter 1, these data files are also augmented with information to implement D-curve constraints and approximate demand side price elasticity. In the case of demand bidding, the FERC e-Library does provide a single (p,q) point which is then used in place of $(0.07 * LMP, qlen)$ in Procedure 1.

2.2 Missing reactive demand

Reactive power supports the voltages that must be controlled for system reliability in order to prevent damage to equipment (e.g. overheating) and voltage collapse, as well as maintain the ability of the system to withstand disturbances [30]. Our data sources for these large-scale datasets did not include reactive demand values (d^Q) at load buses, and this data is necessary to solve the ACOPF model. A key characteristic of reactive power demand is the magnitude and speed at which it changes over times [30], and for this reason we have found that even among industry experts, it is unclear what a network's reactive demand profile should look like. The task of finding reasonable and feasible reactive demand values is further complicated by the sheer size of the dataset and the difficulty of solving the AC model at this scale, since the data would need to be AC feasible for it to be any useful. Based on industry experts knowledge of general stability and demand with regards to reactive power, we propose the following necessary conditions for defining a realistic demand profile across a network.

1. Good power factor values at each load bus.
2. "Reasonable" number of buses with reactive demand.
3. Larger ratio of withdrawals to injections in the overall network.
4. Feasibility of the ACOPF model.

Conditions 1, 2 and 3 are further explained in Sections 2.2.1, 2.2.2 and 2.2.3 respectively while condition 4 is self-explanatory. For details on the process of finding feasible solutions to satisfy condition 4, please refer to Section 2.3.

2.2.1 Power factors

In an AC electrical power system, power factors (PF) are defined as the ratio of real power to apparent power (2.2) where apparent power is defined by (2.1). This leads to a dimensionless number between $[-1, 1]$.

$$S = \sqrt{P^2 + Q^2} \quad (2.1)$$

$$PF = \frac{P}{S} = \frac{P}{\sqrt{P^2 + Q^2}} \quad (2.2)$$

It is generally desirable to maintain a power factor value of close to ± 1.0 , but power factors as low as ± 0.9 are generally considered to be good practice. Provided that we have a non-zero active power generation at a load bus, it is possible to compute a bound for a bus' reactive demand if we limit the power factor to be no worse than 0.9, as in (2.3).

$$PF^2 = \frac{P^2}{P^2 + Q^2}$$

$$\text{For } 0.9 \leq |PF|,$$

$$0.81(P^2 + Q^2) \leq P^2$$

$$|Q| \leq 0.484P$$

$$|Q| \leq 0.484P \quad (2.3)$$

For load buses with no active power, the power factor calculation does not apply as it results in either a 0 or an undefined number. In these cases,

the reactive demand at these buses is arbitrarily restricted to half a unit, or ± 0.5 , as no additionally defining information is available.

2.2.2 Number of buses with non-zero reactive demand

In order to obtain a “reasonable” number of buses with non-zero reactive demand, we found in practice that an weighted objective function combining the norm-1 and norm-2 of reactive demand d_i^Q worked best. Individually, a norm-1 objective yielded too few values (less than 100 buses), while the second norm provided too many small values (over 13000 buses) which likely stems from numerical accuracy problems caused by the dataset and model complexity. To balance both these extremes, we found a weighted combination to be helpful. The third component of the objective works to minimize the difference between the final generator’s active power output (P) to the solution of the DCOPF approximation ($\hat{P}$). And finally, for model feasibility, the underlying ACOPF system constraints are modeled using the polar power-voltage formulation(2.4 - 2.7). Note that d_i^Q in (2.4 - 2.7) is a variable instead of a parameter which is the convention for ACOPF models.

$$\min_{d^Q, d^{Q+}, d^{Q-}} \sum_i w_i^1 (d_i^{Q+} + d_i^{Q-}) + w_i^2 (d_i^Q)^2 + w_i^3 (P_i - \hat{P}_i) \quad (2.4)$$

$$\text{s.t. } d_i^Q = (d_i^{Q+} + d_i^{Q-}) \quad (2.5)$$

$$(1.22) - (1.30) \quad (2.6)$$

$$d_i^{Q+}, d_i^{Q-} \geq 0 \quad \forall i \in \mathcal{N} \quad (2.7)$$

2.2.3 Ratio of withdrawals to injections

Reactive power can be positive or negative, and using similar convention to d^P in Chapter 1, a positive number represents consumption, meaning a

negative number represents an injection back into the system. Load buses are typically not expected to have equipment to provide reactive power support which is why we believe it reasonable to require condition 3. Note that condition 3 is not strictly imposed as a constraint in (2.4 - 2.7). Rather, we require it to be a necessary byproduct of the model solution and will reject any solution that does not satisfy these requirements. In general, we found this to not be an issue, but believe that is important to maintain as one of the conditions.

2.3 Solution methodology for large-scale datasets

Our experiences when finding feasible solutions in these large-scale problems echoed a couple of our earlier results from Chapter 1, namely that the solvers preferred the polar formulation (1.22-1.30), and the most robust solver for these types of models is IPOPTH. Because the data sources do not provide a feasible unit commitment profile across all time periods, what is really needed is a 24 hour unit commitment solution that is AC feasible for each $t \in \mathcal{T}$. Solving a UC_AC at this magnitude is an intractable problem, and we therefore propose Procedure 2 as a tractable solution. Note that while this procedure provides AC feasible solutions for each time period and adheres to ramping and minimum up/down time constraints needed for unit commitment in the DC sense, it falls just short of a complete UC_AC solution which will be explored in future work.

Procedure 2: Finding AC solution for large datasets

```

// Find a commitment profile across 24 hours
1  $(\tilde{P}, \tilde{\theta}, U) \leftarrow \text{Solve UC\_DC } \text{--lineloss}=1.055$ 
// Solve ACOPF for each  $t$  using commitment from previous step
2  $(P_t, Q_t, \theta_t, V_t) \leftarrow \text{Solve polar\_acopf}(\tilde{P}_t, \tilde{\theta}_t, U_t) \text{ --ic}=5 \ \forall t \in \mathcal{T}$ 

```

Step 1 attempts to determine a unit commitment profile that is both AC feasible and adheres to time sensitive constraints such as generator ramping and minimum up and down time, while step 2 solves the ACOPF model for each individual time period.

2.3.1 An AC feasible unit commitment profile

The U.S. Energy Information Administration estimates that electricity transmission losses average about 6% [71]. Table 2.2 tests this theory by comparing the amount of total active power generation in an AC OPF model to the total demand for multiple IEEE test cases. Total line loss is computed dividing the amount of extra generation by the total demand. The resulting line losses for the IEEE testcases however are much lower than expected due to the lack of binding line limits of these datasets. To arrive at a better estimate, we also run the comparison with the first time period from the Winter dataset (for which a feasible unit commitment was provided), and arrive at a line loss rate of 5.22%.

Testcase	Total demand (MWh)	Total generation(MWh)	Total line loss
case118	4242	4319.4	1.28%
case300	23525.4	25118.1	1.29%
case2383wp	24558.4	25118.1	2.28%
case3120sp	21181.5	21604.4	2.00%
case3375wp	48363	49189	1.71%
Winter, t=1	78595	82699	5.22 %

Table 2.2: Line losses: Testcases

By utilizing the `lineloss` option in the GAMS toolbox and solving a 24-hour UCDC model for the Winter dataset with a line loss rate of 5.5%, we find that the unit commitment profile is AC feasible for all timeperiods, $t \in \mathcal{T}$. The line losses range between [5.06, 5.48]% and the average line loss is computed to be 5.25%. We also find that while the line loss rate of 5.5%

also works to provide a AC feasible commitment profile in the Summer dataset, the eventual line losses in this case range between [5.94%, 8.39%] with the average line loss coming up as 7%. Further research needs to be conducted to determine if solving the 24-hour UCDC model with a higher line loss approximation rate would result in changes to these numbers, and if so, it would be interesting to experiment with this from a cost perspective.

2.3.2 Initial starting point for ACOPF

While we have been able to determine that the datasets are AC feasible for each time period through ad-hoc methods, tailored model options and unconventional starting points, we would like to determine the usefulness of the toolbox in solving realistic sized models such as this. Based on the results of Section 1.4.2, we find that initial conditions 0,4,7 are the best options moving forward with a large scale dataset. Initial testing as shown in table 2.3 clearly shows that the midpoint initial condition ($ic=0$) is not effective for large-scale datasets, despite being a solid choice in the IEEE testcases.

Testcase	ic=0	ic=4	ic=7
Winter, t=20	14m 28.69s	4m 27.31s	4m 35.52s
Summer, t=18	infeas	4m 34.74s	3m 50.66s

Table 2.3: Initial conditions: large-scale datasets

Table 2.4 provides a more in-depth comparison between the two selections and by running the solvers with a time limit of 3 hours. The ideal situation here is to find an ACOPF solution in under 20 minutes/time period. As the results indicate, there is no one-size-fits-all solution and solution times vary significantly. Initial condition 4 seems to outperform initial condition 7 which is surprising given that regular DC models do not account for line losses. The Summer dataset seems to be easier, with

20 models solving under 20 minutes compared to 14 in the Winter dataset. There are 5 periods in the Winter dataset that did not find a solution, whereas only 1 in the Summer. The asterisks in the Winter dataset indicate that solutions could be obtained using $ic=5$, but performance results of that initial condition are not provided here.

These results highlight the difficulty of solving real-life models and point to the reason most or all of industry currently choose to solve approximations instead.

Time(t)	Winter (solu time (s))		Summer (solu time (s))	
	ic=4	ic=7	ic=4	ic=7
1	1036	803.9	6112.3	682.7
2	*		344.2	938.2
3	1138.1	8932.8	409.8	416.2
4	1410.8	4852.5	600.4	630.8
5	781.5	3638.6	433.5	429.8
6	697.0	2940.4	459.5	327.0
7			360.8	517.4
8	544.7	407.8	589.7	539.4
9	441.1	687.8	442.5	1639.7
10	2443	973.4	287.2	609.9
11	447.7	380.2	348.8	602.9
12	523.7	889.2	543.1	659.4
13	301.8	548.4	288.3	7307.4
14			656.2	4077.7
15	*	7833.1	341.2	634.9
16	1301.4			
17	*		800.1	733.2
18	680.5	620.1	562.5	680.2
19	1303.4	679.9	660.8	506.1
20	739.6	726.8	568	652.8

21	495.9	496.8	442.6	601.7
22	1021.7	394.9	1570.7	1099.1
23	746	638.5	1439.9	450.3
24	5712.2	1186.8	298.2	878.9
STATISTICS				
AVG time (s)	1145.6	1980.1	807.0	1111.4
Solves <20min	14	14	20	20

Table 2.4: Comparing Initial conditions: Winter & Summer

2.4 Numerical integrity and quality assurance

When it comes to the process of finding feasible solutions in large-scale models, it is natural to accept a solver's claim of solution feasibility and this is especially true when dealing with commercial solvers. It is uncommon to question solution integrity in large datasets when manual checks of the returned data is cumbersome and not necessarily intuitive on such a scale, and especially so when the solution is provided by commercial and well-established solvers. Paradoxically, solution checking is probably more important in situations with large data as numerical issues may be magnified at this scale leading solvers to provide false claims of feasibility and/or optimality. The two examples discussed in Sections 2.4.1 and 2.4.2 highlight these problems and Section 2.4.3 discusses a tool that can be used to check feasibility.

2.4.1 Solution quality of UCDC models

In this section, we discuss the results of solving the UCDC model from the GAMS toolbox in Chapter 1 using the Winter and Summer datasets. When using the CPLEX solver to find commitment solutions for the UCDC model (1.11-1.21) with 24 consecutive time periods, the returned solutions

were not within the expected tolerance. Consider the example CPLEX output log displayed below, and note that a gap of almost 200 between the MIP solution and final solve. This gap points hints towards numerical inaccuracies due to the data size and model complexity.

Fixed MIP status(1): optimal

Solution satisfies tolerances.

MIP Solution: 35535818.059876

Final Solve: 35536002.311811

To deal with this issue, we experimented with multiple built-in CPLEX options and manual scaling of the data, and eventually found that a combination of CPLEX options `scaind` and `numeralemphasis` which focus on numerical precision and aggressive scaling, adequately dealt with minimizing the gap. In order to compare the different techniques, CPLEX options `MIPKappaStats` and `quality` provide detailed printouts regarding the conditioning of matrices and numerical integrity at the solution. An example of what the MIP Kappa distribution report looks like is as follows:

MIP Kappa distribution Report:

Percentage of stable bases ($\kappa < 1e+7$):	0.00%
Percentage of suspicious bases ($1e+7 < \kappa < 1e+10$):	35.08%
Percentage of unstable bases ($1e+10 < \kappa < 1e+14$):	64.92%
Percentage of illposed bases ($1e+14 < \kappa$):	0.00%
Max condition number:	5.0089e+11
Attention index (if >0.03 caution is advised)	0.20

In this example, the high percentage of unstable bases and large attention index indicate problems with the numerical integrity of the final solution. Our aim is to provide an attention index of ≤ 0.03 , which means that most of the bases fall within the stable or suspicious range, with lit-

tle (<2%) to none falling in the unstable range onwards. The following printout is example output from using the afore-mentioned CPLEX options, and shows decent solution quality, especially considering the model complexity.

MIP Kappa distribution Report:

Percentage of stable bases ($\kappa < 1e+7$):	19.23%
Percentage of suspicious bases ($1e+7 < \kappa < 1e+10$):	80.77%
Percentage of unstable bases ($1e+10 < \kappa < 1e+14$):	0%
Percentage of illposed bases ($1e+14 < \kappa$):	0.00%
Max condition number:	1.9341e+08
Attention index (if >0.03 caution is advised)	0.01

2.4.2 Solution quality of reactive power values

Returning to the problem of finding AC feasible solutions in the reactive demand model (2.4-2.7), a closer look at the reactive demand values, d^Q , revealed that a very large majority of the load buses had a very small solution value associated with it, typically in the range of $[1e^{-8}, 1e^{-16}]$. More specifically, this problem appeared to be linked to our use of interior point solvers such as IPOPTH and Knitro. In cases when CONOPT was able to find a solution to the model, it resulted in a “cleaner” solution in which many of these tiny extraneous values eventually converged to 0. While the latter is considered to be a more realistic reactive demand profile, CONOPT typically has immense difficulty with these models. In this case, we propose Procedure 3 as an extension of Procedure 2 to find

reactive demand values.

Procedure 3: Finding reactive values for large datasets

```

// Find a commitment profile across 24 hours
1  $(\tilde{P}, \tilde{\theta}, U) \leftarrow \text{Solve UC\_DC } \text{--line loss}=1.055$ 
// Find a good starting point
2  $(\tilde{P}^2, \tilde{Q}^2, \tilde{\theta}^2, \tilde{V}^2, \tilde{d}^Q) \leftarrow \text{Solve polar\_acopf}(\tilde{P}, \tilde{\theta}, U) \text{ --ic}=7$ 
NLP=IPOPTH
// Solution polishing based on IPOPTH starting point
3  $(P, Q, \theta, V, d^Q) \leftarrow \text{Solve polar\_acopf}(\tilde{P}^2, \tilde{Q}^2, \tilde{\theta}^2, \tilde{V}^2, \tilde{d}^Q, U)$ 
NLP=CONOPT

```

While this procedure worked reasonably well, there was no single process that could reliably solve each time period, and a mixture of CONOPT options that make specific algorithmic had to be tried in the more difficult situations. For example, option `lsanrm` uses a steepest edge method instead of steepest descent and `LMMXST` uses a method based on bending method to determine step length.

By combining the conditions laid out in Section 2.2 and the solution finding process discussed here, we were able to obtain reactive demand values for each of the timeperiods in Winter, Summer and their RTO cases, which are also AC feasible. Note that in the RTO cases, the step of finding a unit commitment profile using UCDC is done individually for each time period as they are not consecutive periods and as such, do not need to adhere to time specific constraints as a 24 hour planning model would.

2.4.3 Solution quality of AC OPF models

Based on our experiences in finding reactive demand values, it is prudent to perform a feasibility check on solutions returned by IPOPTH and we do so using Examiner, which is a GAMS tool for solution verification [22].

In Table 2.5, we run diagnostics on some solutions provided by IPOPTH and compare them to Table 2.6 which are results from using CONOPT.

Testcase	Var bound(P)	Var bound(D)	Constraints(P)	Constraints(D)
case14	OK	OK	OK	OK
case118	OK	OK	OK	OK
case300	OK	OK	OK	OK
case2383wp	OK	OK	OK	Max viol: $2.0e^{-6}$
case3120sp	OK	OK	OK	Max viol: $-1.7e^{-6}$
case3375wp	OK	OK	OK	Max viol: $1.5e^{-5}$

Table 2.5: Examiner results: IPOPTH

The first two columns check if variables are within their bounds for both primal and dual problems, and the next two columns check for constraint feasibility again both in primal and dual. Default tolerances for all four checks are set at $1e^{-6}$, which are tight tolerances, meaning that the test favors false negatives over false positives [25].

Testcase	Var bound(P)	Var bound(D)	Constraints(P)	Constraints(D)
case14	OK	OK	OK	OK
case118	OK	OK	OK	OK
case300	OK	OK	OK	OK
case2383wp	OK	OK	OK	Max viol: $1.1e^{-4}$
case3120sp	OK	OK	OK	Max viol: $7.1e^{-3}$
case3375wp	OK	OK	OK	Max viol: $7.8e^{-3}$

Table 2.6: Examiner results: CONOPT

While IPOPTH performed badly in Section 2.4.2 when searching for feasible reactive demand values, it seems to outperform CONOPT not only in solver running time, but also in solution quality. A point of note is that both solvers return the same objective value and therefore these results regarding solution quality may not be explicitly obvious without using the Examiner tool. We therefore test the IPOPTH solution of the Winter and Summer datasets in Table 2.7 and see that it produces similar results to the IPOPTH tests on the larger IEEE testcases. Our conclusion therefore is that the IPOPTH solution for the large-scale datasets is acceptable and feasible,

specifically when considering the size and complexity of the model. It is not unreasonable to relax certain tolerances in these conditions, and if we relax the dual constraint tolerance to $1e-5$, we see that they will easily pass all 4 feasibility tests.

Testcase	Var bound(P)	Var bound(D)	Constraints(P)	DConstraints(D)
W, t=1	OK	OK	OK	OK
W, t=5	OK	OK	OK	Max viol: $-3.0e^{-6}$
W, t=20	OK	OK	OK	Max viol: $2.4e^{-6}$
S, t=1	OK	OK	OK	OK
S, t=18	OK	OK	OK	Max viol: $-6.70e^{-6}$

Table 2.7: Examiner results: Winter/Summer

2.5 Additional Issues in Network Feasibility

While going through the exercise of finding AC feasible solutions for the reactive demand model (2.4-2.7) using the Procedure 3, we came across additional feasibility issues that either a) originated from erroneous data, or b) stemmed from actual network infeasibility. In both cases, feasible AC solutions were difficult to come by and we address those issues below.

2.5.1 Tap transformers

Tap transformers, when installed, may change the voltage ratio on a line which in turn enables stepped voltage regulation on that line. With relation to the ACOPE, this shows up in two components, τ_{ijc} , θ_{ijc} which are the transformer tap ration and angle on line ijc respectively, and affects active and reactive power. Build and maintenance costs of these devices limit the number of them in a network.

Based on network information, both Winter and Summer datasets contain roughly 30 active tap transformers. Analyzing initial solutions found that a number of lines at their active power limit were associated

with unusually high tap angle values, θ_{ijc} , at $\pm 84.99^\circ$. In fact, these values seemed to be outliers, while the rest of the data varied between $[-30^\circ, 30^\circ]$ or 0.524 radians. Note also that in DC model, $\sin(i) = i$. Therefore, $\sin(\theta_i - \theta_j - \phi_{ijc})$ in (1.2) may be misrepresented and provide bad initial starting points when used in the ACOPF model. A simple test of replacing them with 0 resulted in a marked improvement in our ability to find feasible solutions. Based on likely physical limitations of these transformers, we proceeded to solve the ACOPF feasibility model below (2.8-2.10), where θ_{ijc}^* refers to the outlier tap angle values.

$$\min \sum_{ijc \in \mathcal{E}} \theta_{ijc}^* \quad (2.8)$$

$$s.t. \quad (1.22) - (1.30) \quad (2.9)$$

$$\theta_{ijc}^* \in [-0.524, 0.524] \quad (2.10)$$

Further testing and analysis of the feasibility model revealed very small solution values for θ_{ijc} , and tests indicated that replacing them with 0 did not significantly impact the solution. Therefore, the erroneous tap angles were replaced with 0 in all model runs.

2.5.2 Line limit relaxation

Given that the demand data is based on peak consumption for both Winter and Summer seasons, it is not unusual to encounter an infeasible model especially when demand bidding is turned off. What was unusual in the Summer network however was that the models were DC infeasible for 12 hours throughout the day. Our goal was to make the models AC feasible, even with demand bidding turned off, which allows the user more flexibility in running these datasets. With that in mind, we solved a minimum AC feasibility model by adding slack variables to equations (1.22

- 1.27) which relaxes line limits in (1.22-1.25) and generator operating limits (generator buses) or consumer demand (load buses) in (1.26-1.27). The resulting solution pointed to 3 lines as infeasibility sources, summarized in Table 2.8.

Line	Infeasible periods(t)	Max violation(%)
A	11-23	32%
B	13-22	14%
C	15-19	4%

Table 2.8: Summary of line infeasibility for Summer dataset

To maintain the integrity of the network information, we also experimented with only adding slack variables to (1.22-1.25) but found that the solutions to those problems were unwieldy (required many changes) and distinctly different from period to period, such that no general insights could be drawn. We also experimented with allowing slacks in the $n + 3$ subgraph (starting from the infeasible lines listed in Table 2.8), but this did not yield any actionable results. Therefore, we uniformly applied a 40% line limit increase to the 3 lines above, which allows the modeler some flexibility when experimenting with different demand inputs. Finally, timeperiod $t = 17$ in Summer RTO required small changes to the active demand of 2 buses in order to retain feasibility. The alternative was to increase the line limits of 3 additional lines but we felt this to be excessive.

2.6 Epigrids Data

The EPIGRIDS project [1] aims to construct realistic transmission system models and scenarios that will serve as test cases to reduce barriers to the development and adoption of new technologies in grid optimization and control. The motivation for creating these datasets is because the research community lacks high-fidelity and public power systems data that realistically represent current and evolving grid characteristics. Typically, real

world data needed to validate research advancements are restricted due to grid security concerns. The CEII datasets discussed in 2.1 perfectly exemplify this problem as they fall under strict non-disclosure agreements and can only be provided under the Critical Energy Information Infrastructure (CEII) usage agreements as detailed by FERC. To help drive practical and useful innovation in the electric power industry, there is a need for datasets that more closely mimic the characteristics of the actual grid without disclosing sensitive information. The Grid Optimization (GO) Competition conducted by the Advanced Research Projects Agency-Energy (ARPA-E) (include citation and link to the GO website) already utilizes some of the datasets generated by the EPIGRIDS project, to accelerate the development of transformational and disruptive methods for solving problems related to the electric power grid and to provide a transparent, fair, and comprehensive evaluation of new solution methods.

2.6.1 Data Utility Tool Updates

To accommodate the data arising from the EPIGRIDS project, the data conversion tools in the OPF Toolbox have been updated and tested to ensure that the data utility models are able to incorporate the EPIGRIDS .raw electricity grid files into our testing environment. Due to differing assumptions about the contents of input files, the following updates had to be made to the raw2gdx data utility tool that converts a .raw file to a .gdx file, that is consistent with the Toolbox's GAMS models.

1. Non-integer valued circuits are changed to integers before being added to line data information. Integer valued circuits are required in the data description.
2. Zone information from the LOAD DATA portion in the .raw file are ignored and not checked against the AREA information from

the BUS DATA. This is because the EPIGRIDS data does not always maintain this fidelity.

3. Details of load buses with a non-active (0) status are ignored and not included in the .gdx file because the current data descriptions does not allow for this added dimension. Since the load buses are inactive, they should not impact the demand profile description of the network. This could be a possible extension in the future.
4. The control mode of the switched shunt data is no longer validated, because the epigrids datasets do not use this field reliably. Previously, only fixed (0) and continuous adjustment (2) modes were encoded into the .gdx file. This issue will likely need to be revisited in the future.

These updates allowed us to reliably convert the EPIGRIDS datasets to the Toolbox's .gdx formats, and thus utilize the provided models and tools. The datasets published in the GO Competition were used for testing.

2.6.2 PowerFlow Feasibility of the datasets

Using the models in the OPF Toolbox, all the datasets were proven to be DCOPF feasible, even when used with the `--lineLoss=1.055` option. This option crudely simulates line loss in an ACOPF model by increasing the demand profile across all loads in the network. Chapters 1 and 2 provide example usage cases of this.

Unfortunately, in the majority of the datasets, we were not able to obtain feasible AC power flow solutions, regardless of starting points or OPF formulation. In some of the cases, the datasets were PowerWorld feasible in the AC model, but the PowerWorld solution was either infeasible in our models or did not provide a good starting point. We believe that this is due to the slacks in the PowerWorld solutions that the GAMS models do not represent or account for. Thus, the addition of slacks to the ACOPF

models will likely be needed to solve many of these datasets within the framework of the Toolbox. In the testcases where we were able to solve the ACOPF and find a feasible AC operational point, the starting condition `ic=7` where the real power and voltage angle values are initialized using a DCOPF model with line loss approximation (`--lineloss=1.055`) was instrumental in finding these solutions. These findings add another data reference point to the starting point analysis conducted in Sections 2.3.2 and 1.4.2, showing that realistic datasets can be challenging to solve reliably, and further highlighting the importance of the work coming out of the EPIGRIDS project.

2.7 Conclusions

In conclusion, we produced 4 large-scale datasets with 13900 buses that is sourced using realistic data, has AC feasible data solutions for each time period, and that adhere to the unit commitment constraints of a UCDC model that incorporates line loss by a simple increase in demand of 5.5%. We are also able to provide exact ACOPF solutions (instead of approximations) for these datasets using default options in 5 minutes, and with further research, believe we can extend these results to all 24 periods.

These are difficult problems to solve, and even more so in “real-time” for the purposes of day-ahead planning. The compilation of these large-scale and realistic datasets however allow us to push the modeling boundaries for operation and planning models, for future implementation in

industry settings.

Procedure 4: Finding AC solution for large datasets

```

// Find a commitment profile across 24 hours
1  $(\tilde{P}, \tilde{\theta}, \tilde{U}) \leftarrow \text{Solve UC\_DC } \text{--line loss}=1.055$ 
// Solve ACOPF using commitment profile from previous step
2  $(P_t, Q_t, \theta_t, V_t) \leftarrow \text{Solve polar\_acopf}(\tilde{P}_t, \tilde{\theta}_t, \tilde{U}_t) \text{ --ic}=4 \ \forall t \in \mathcal{T}$ 
// Find UCAC solution using provided starting points
3  $(P, Q, \theta, V, U) \leftarrow \text{Solve UC\_AC}(\tilde{P}, \tilde{Q}, \tilde{V}, \tilde{\theta}, \tilde{U})$ 

```

As final note regarding the reactive demand profile as described in Section 2.2, one limitation of the conditions is that they do not allow for significant reactive support equipment such as synchronous condensers to be installed in load buses with 0 demand. This may be a limiting factor of AC feasibility, and should be looked into.

Lastly, in an effort to increase the accessibility to the high-fidelity power systems datasets generated by the EPIGRIDS project, we updated to our utilities to be able to bring those datasets in to the OPF Toolbox. We believe this will greatly enrich the testing and model validation capabilities of the research community.

3 HIERARCHICAL FRAMEWORK FOR LONG TERM PLANNING MODELS

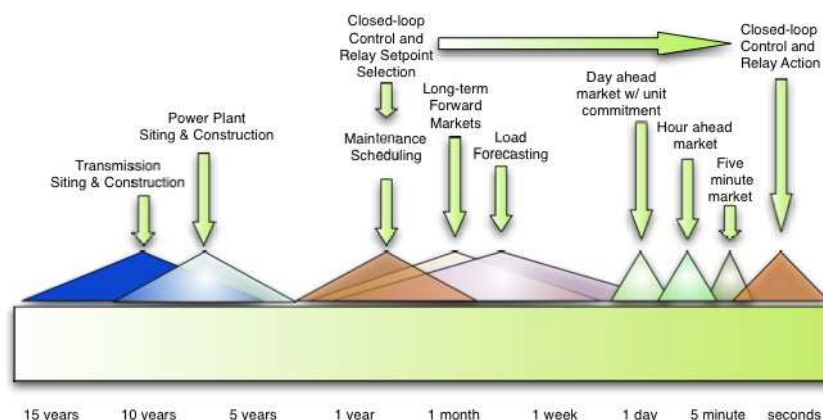


Figure 3.1: Representative decision-making timescales in electric power systems, courtesy of C. DeMarco

Planning and operating the Next Generation Electricity Grid is a complex matter that involves decisions across varying timescales; from reacting to changes within fractions of a second, to 5-minute and day ahead market clearing models, as well as decisions like major grid enhancements which are typically made on the scale of 10-15 years in advance. Figure 3.1 serves to illustrate the complexity of the power system by highlighting some of the many operations and decisions as controlled by multiple agents that take place at hugely differing timescales. While it is clear that information flow from decisions made in the long time scale models may affect outcomes and behaviors in the shorter time scales, what makes this setting particularly interesting is the less obvious constraints that operations at very fast time scales (e.g. requirements for grid resilience against cascading failures) potentially impose on longer term decisions like maintenance scheduling and electricity grid expansion. Due to the rapidly changing landscape through the introduction of new technologies such as electric

cars, renewable energy and storage, it is clear that better modeling and solution techniques for power systems are required for continued reliable operation.

The creation of regional transmission organizations (RTO) by the Federal Energy Regulatory Commission (FERC) in the last decade was one of the responses to electricity market deregularization which opened the doors to competition among generators. By FERC Order 2000A, one of the organization's responsibilities is planning necessary transmission line enhancements to provide efficient, reliable and non-discriminatory service [29]. Various approaches to this decision making process have been explored, including a model used to calculate a network's congestion status [52], the usage of a probabilistic reliability criteria [46], a MINLP based on the minimization of construction cost and transmission congestion in [77], as well a bilevel model in [47]. As a long time scale decision problem, transmission line expansion is clearly confounded by uncertainties in estimates of electricity demand, fuel cost, and even of visible technologies such as plug-in hybrid electric vehicles and renewable energy sources like wind and solar power. Bottlenecks within the network further prevent perfect competition among market participants because firms may use their market power to maintain prices above marginal cost [67, 14]. Adding to the complexity is the structural makeup of the system whereby interaction, and sometimes interdependence between issues like transmission line planning, generator building and upgrading, as well as the operational concerns of independent service operators (ISO) is difficult to capture; a result of non-cooperative agents and often competing principles such as economic revenue, policy and environmental influences and system stability. These issues not only stress the importance of a framework that is able to represent the different agents, timelines and uncertainties in a cohesive manner, but also highlight a need for model flexibility which enables easy adaptation to evolving information flow in the absence of

validation data.

To that end, this paper presents a hierarchical framework designed to capture the interactions between a long timescale model such as transmission line planning with shorter timescale decisions such as grid operation and generator behavior, all the while maintaining a clear partition of timelines and agents. Beyond the advantage of transparent information flow, the framework presented in Section 3.1 benefits from its ability to be broken down into smaller and disparate pieces to facilitate the solution method proposed in Section 3.2. Section 3.3 then provides illustrative numerical examples while Section 3.4 concludes with some future points of debate.

3.1 Hierarchical Framework

By using a hierarchical framework, it is intuitive to maintain the integrity of different timelines by keeping their respective decision models within separate hierarchies. For this example, the problem is simply split into two timelines, with the long timescale transmission expansion model occupying the upper hierarchy while the shorter timescale decision models reside in the hierarchy below. Tables 3.1, 3.2 and 3.3 provide easily referenced descriptions of the sets, variables and parameters used in the decision models described in the rest of this section.

Table 3.1: Description of Sets

Set	Description	Indices	Model
Ω	Demand scenarios	ω	All
N	Location nodes	i, j	All
X	Transmission expansion set	-	RTO
F	Firms	f	Firm
$J^f \subset N$	Location of firm f 's generators	j	Firm
$A \subset (N, N)$	Transmission Lines	(ij)	ISO

Table 3.2: Description of Variables

Decision Variable		Description
RTO	x	Investment in transmission line expansion
Firm	y_j	Investment in generator at j
ISO	z_{ij}^ω	Real power flowing along ij in scenario ω
	θ_j^ω	Voltage phase angle at j in scenario ω
	q_j^ω	Real power at j in scenario ω
	o_j^ω	Dispatch of generator at j in scenario ω

Table 3.3: Description of Parameters

Parameter	Description	Model
S_{ij}	Susceptance of line ij	ISO
$[\underline{u}_j, \bar{u}_j]$	Operating limits of generator at j	ISO
b_{ij}	Capacity of transmission line ij	ISO
d_j^ω	Demand at j in scenario ω	All
$\mu c(x)$	Cost of line investment	RTO
$g(x)$	Budgetary/engineering limitations	RTO
π_ω	Probability of scenario ω	RTO, Firm
$C_j(q_j^\omega, y_j)$	Cost function of generator at j	Firm, ISO
$\eta r(y)$	Cost of generator investment	Firm
$h_f(y)$	Budgetary constraints of firm f	Firm

3.1.1 Transmission Planning Model

For the uppermost hierarchy, we start with a basic formulation which balances investment cost with consumer welfare. The latter is defined as the total expected payment collected by the ISO from its wholesale customers and is calculated using locational marginal prices (LMP). The node-to-node disparity of these prices point to the existence of congestion within the network that restricts the ability of generation at low-priced

locations to supply power to customers at high priced locations [67], thus making it a useful social welfare measurement. It is possible and probably desirable to consider a more inclusive or robust objective function but for expository purposes, we use the formulation below.

Expansion Model (RTO)

$$\min_x \quad \overbrace{\mu c(x)}^{\text{investment cost}} + \overbrace{\sum_{\omega} \pi_{\omega} \sum_i d_i^{\omega} p_i^{\omega}(x)}^{\text{consumer welfare}} \quad (3.1)$$

$$\text{s.t. } g_k(x) \leq 0 \quad \forall k \quad (3.2)$$

Variable x represents expansion on specific lines and the objective balances expansion cost $\mu c(x)$ with expected cost to the wholesale customer. Demand scenarios d_i^{ω} and their corresponding probabilities π_{ω} represent aggregations of both forecast uncertainties and variations in consumption patterns, such as seasonal or daily fluctuations and $p_i^{\omega}(x)$ is the resulting LMP which is a response variable from the lower hierarchy. This decision may be subject to budgetary, physical and engineering constraints (3.2), and the dimension of x is assumed to be small since we expect physical and infrastructural limitations as well as political and geographical complexities to severely limit possible layouts. This formulation is very general and can be expanded to include other constraints, or manipulated to consider different objectives as necessary for accurate representation of the system, without upsetting the lower hierarchy.

3.1.2 Price Response Function

The bulk of the detail in this model actually lies in the process of obtaining the response variable, $p_i^{\omega}(x)$, for the upper level objective in (3.1), thus precipitating the introduction of the equilibrium model (3.3)-(3.10) which is an enhanced version of the traditional DC optimal power flow model

(OPF) commonly found in the power systems literature. The OPF is an operational model solved by the ISO and is necessary for acquiring the wholesale prices (LMP), $p_i^\omega(x)$. Given that transmission expansion is a long timescale decision however, we introduce the influence of firms' short term competitive behavior which represents the dynamic nature of the network corresponding to the time lag between the decision to expand and implementation processes. Thus in this augmented formulation, firms may choose to upgrade generators, for example by changing fuel type, increasing efficiency or output limits [8] [37], or alternatively allow financially unattractive power plants to be retired. Since upgrade requests may be processed through expedited procedures [61], the decision time frame is similar to that of generator retirements [63] [62]. Therefore, the above-mentioned equilibrium is made up of the optimality conditions of an OPF with unit commitment (3.3)-(3.8) solved by the ISO in each scenario, and of an investment model (3.9)-(3.10) for each generator firm.

The OPF^{UC} is a cost minimization operational model with unit commitment decisions that is characterized by power balance constraints at each node (3.4), Kirchoff's law (3.5), and operating limits of the system (3.6,3.7). The cost function $C_j(q_j^\omega, y_j)$ consists of a fixed cost component as well as variable costs based on the quantity of energy generation, q_j^ω . Unit commitment is described by integer variables o_j^ω . The OPF produces and returns the feedback values p_j^ω to the upper hierarchy in response to the line capacity changes as a function of x (3.7).

The firm's goal on the other hand is to balance the minimization of its own operational cost $C_j(q_j^\omega, y_j)$ and investment cost $\eta r(y_j)$. The interaction between the firm's investment decision variable(y) and the ISO's OPF implementation (q, o, x, θ) is facilitated through the cost function $C_j(q_j^\omega, y_j)$ that appears in both objective functions (3.3,3.9). While q_j^ω is a variable in the OPF, it becomes a parameter to the firm's model, while y_j is conversely a variable in the firm's model but a parameter in the OPF. This clear

separation of models for each individual agent provides a good macro view of the driving forces within the electricity delivery system, while clearly exposing their goals and relationships.

Equilibrium

OPF^{UC} ($\forall \omega \in \Omega$):

$$\min_{o,q,z,\theta} \sum_j o_j^\omega C_j(q_j^\omega, y_j) \quad (3.3)$$

$$\text{s.t. } q_j^\omega - d_j^\omega = \sum_{i \in \mathcal{N}^j} z_{ij}^\omega \quad \perp p_j^\omega \quad \forall j \in N \quad (3.4)$$

$$S_{ij}(\theta_i^\omega - \theta_j^\omega) = z_{ij}^\omega \quad \forall ij \in A \quad (3.5)$$

$$o_j^\omega \underline{u}_j \leq q_j^\omega \leq o_j^\omega \bar{u}_j \quad \forall j \in N \quad (3.6)$$

$$z_{ij}^\omega \in [-b_{ij}(x), b_{ij}(x)] \quad \forall ij \in A \quad (3.7)$$

$$\theta \text{ free, } o \in \{0, 1\} \quad (3.8)$$

Firm ($\forall f \in F$):

$$\min_{y \geq 0} \sum_\omega \pi_\omega \sum_{j \in J^f} o_j^\omega C_j(q_j^\omega, y_j) + \eta r(y_j) \quad (3.9)$$

$$\text{s.t. } h_f(y) \leq 0 \quad (3.10)$$

Now, if we were to approximate the integer formulation by using some smooth function $\hat{C}_j(q_j^\omega, y_j)$ in place of $o_j^\omega C_j(q_j^\omega, y_j)$, we are able to connect these models as one set of complementarity constraints. The mixed complementarity program (MCP*) formed by the first order conditions of the approximated model (3.11)-(3.17) circumvents the intractability of a solution process caused by unit commitment in the ISO dispatch model but replaces this with a nonconvexity in the MCP formulation. We present one such approximation in the Appendix which enabled us to solve MCP*

given reasonable starting points. We will return to discuss this MCP again in Section 3.2, where insights gained from our solution process enable further exploitation of the structure presented here.

MCP* (1st order conditions of:)

OPF* ($\forall \omega \in \Omega$):

$$\min_{z, \theta, q} \sum_f \sum_{j \in G_f} \hat{C}_j(q_j^\omega, y_j) \quad (3.11)$$

$$\text{s.t. } q_j^\omega - d_j^\omega = \sum_{i \in \mathcal{N}^j} z_{ij}^\omega \quad \perp p_j^\omega \quad \forall j \in N \quad (3.12)$$

$$S_{ij}(\theta_i^\omega - \theta_j^\omega) = z_{ij}^\omega \quad \forall ij \in A \quad (3.13)$$

$$q_j^\omega \in [0, \bar{u}_j], \quad \theta \text{ free} \quad \forall j \in N \quad (3.14)$$

$$z_{ij}^\omega \in [-b_{ij}(x), b_{ij}(x)] \quad \forall ij \in A \quad (3.15)$$

Firm* ($\forall f \in F$):

$$\min_{y \geq 0} \sum_{\omega} \pi_{\omega} \sum_{j \in J_f} \hat{C}_j(q_j^\omega, y_j) + \eta r(y_j) \quad (3.16)$$

$$\text{s.t. } h_f(y) \leq 0 \quad (3.17)$$

3.1.3 Model Assumptions

As noted at the start, generator bidding processes play a significant role in consumer price determination and discussion regarding short term bidding strategy have been explored in [23], [42] and [41]. Our formulation however assumes that such market advantages will be inconsequential over extended periods of time by reasoning that a generator's true cost curve can be inferred via the bidding process in the long run. Under this assumption, the ISO may be viewed as a price setter because it controls the electricity

dispatch process, although firms have control over the parametrization of the cost curve through variable y_j .

Note that as a result, the objective function listed in (3.18) which models a firm's maximum expected profit instead of minimum cost is mathematically no different from (3.16), because the above assumption results in p_j^ω becoming a parameter in the firm's objective function.

$$\max_{y \geq 0} \sum_{\omega} \pi_{\omega} \sum_{j \in J_f} p_j^{\omega} q_j^{\omega} - \hat{C}_j(q_j^{\omega}, y_j) - \eta r(y_j) \quad (3.18)$$

3.1.4 Generator upgrades

Based on standard power engineering literature and datasets, a quadratic function (3.19) is used in calculating the total cost incurred by generator j when generating q_j units of electricity, where γ_j defines the unit commitment (fixed) cost while the rest of the equation defines a per unit variable cost. In order to account for the time lag between the decision and implementation processes of transmission line expansion, the introduction of generator investment variable y_j allows current generators to compete with future technology by upgrading current power plants to reduce operational costs. For our purposes, we define the relationship between investment y_j and cost function $C_j(q_j, y_j)$ with a diminishing rewards function, meaning each additional unit of investment, y_j , will have a diminishing effect on the cost of generation as shown in (3.20).

$$C_j^*(q_j^{\omega}) = \alpha_j (q_j^{\omega})^2 + \beta_j q_j^{\omega} + \gamma_j \quad (3.19)$$

$$C_j(q_j^{\omega}, y_j) = \alpha_j^*(y_j) q_j^2 + \beta_j q_j + \gamma_j,$$

$$\text{where } \alpha_j^*(y_j) = \underbrace{\alpha_j}_{\text{max. benefit}} + \underbrace{\frac{\alpha_j}{4} \left(\overbrace{\exp\{-\lambda_j y_j\} - 1}^{\text{diminishing rewards fcn}} \right)}_{\text{gen. size}} \quad (3.20)$$

As illustrated in figure 3.2, positive values of investment shift the curve

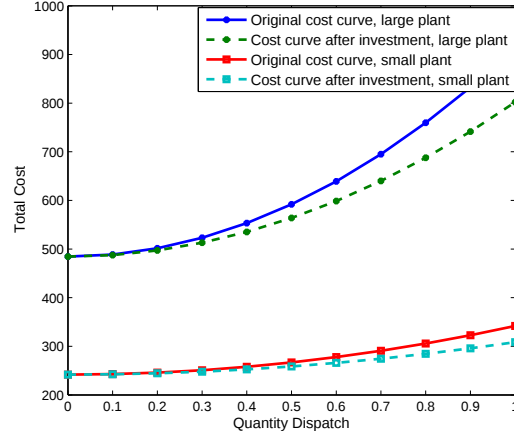


Figure 3.2: Illustration of investment variable's effect on cost curve

to the right, thereby making cost of dispatch cheaper. During testing, we associate λ_j with generator size, meaning larger generators require bigger investments to increase efficiency. This affords the modeler flexibility when defining the investment requirements for different classes of generators, according to elements such as size, location and type of fuel. The maximum efficiency benefit is $-\frac{\alpha_{0,j}}{4}$ and in practice, we define λ_j as $-\log(0.001)\psi_j$, based on the assumption that the first unit of y_j uses ψ_j monetary units to reduce $\frac{\alpha_{0,j}}{4}$ by 0.1%.

The advantage of this function is that it is simple in application by only affecting the quadratic term, and it is also smooth, continuous and convex in the (q, y) space. The modeler however is free to consider more complex upgrade functions such as separating y into multiple components that allow the firm to impact the generator operation with different upgrade features, for example to alter a plant's generation limits $[\underline{u}, \bar{u}]$, and/or to alter its fuel source [8] [37].

3.2 Solution Method

3.2.1 Optimization Model

The RTO decision problem, which at first glance appears to be a simple model, suffers from the lack of derivative information and good theoretical properties of the LMP values, $p_i^\omega(x)$. Specifically, unit commitment in the lower hierarchy affects the continuity of the function, and the possibility of multiple equilibria may have an adverse effect on its smoothness. Thus, we are inclined to treat the objective (3.1) as a blackbox function, and contend that the use of derivative-free optimization (DFO) is a reasonable approach in which the solver allows an external process to solve the optimization based solely on the availability of objective function values [64]. In order to accommodate the limitation of including modeling constraints in DFO solvers, model (3.1)-(3.2) is replaced with a penalty function, where $x \in X$ represent simple bound constraints:

Penalty Function:

$$\min_{x \in X} \left\{ \mu c(x) + \sum_{\omega} \pi_{\omega} \sum_i d_i^{\omega} p_i^{\omega}(x) + \sigma \sum_k \max \{g_k(x), 0\} \right\} \quad (3.21)$$

The bulk of the solution process however lies in the methodology behind obtaining a function evaluation of $p_i^{\omega}(x)$ to provide to the optimization model. This constitutes the focus of the remainder of Section 3.2, whereby we solve the MCP in the lower level of the hierarchy, utilizing the Extended Mathematical Programming Framework (EMP) in conjunction with GAMS, as discussed in [32].

3.2.2 MCP

Recall MCP* which is based upon the first order conditions of approximation model (3.11)-(3.17) where $\hat{C}_j(q_j^\omega, y_j)$ is documented in the Appendix. While smooth and continuous, the model still has non-convex objective functions with respect to q_j^ω . It is therefore necessary to provide the MCP with good starting points in order to find an appropriate equilibrium solution, and we do so based on the solution of the OPF, solved independently and to global optimality for each scenario. Selecting a good solution from multiple equilibria is also a concern, and Procedure 1 addresses these concerns in the following manner.

Procedure 5: MCP Solution Process

Initialization: $q_j^\omega = 0, y_j = 0 \quad \forall j \in N, \forall \omega \in \Omega$

1 Loop

// Obtain q_j^ω starting points with y_j fixed, eqns.(3.11-3.15)

2 $\hat{q}_j^\omega \leftarrow \text{OPF}^*(y_j) \quad \forall \omega \in \Omega$

3

4 **if** $\hat{q}_j^\omega = q_j^\omega$ **then** *// convergence*

5 break;

6 **else** *// Solve equilibrium with $\hat{q}_j^\omega$ starting points*

7 $(q_j^\omega, y_j) \leftarrow \text{MCP}^*$ with $\hat{q}_j^\omega$ starting points

The method starts with $\text{OPF}^*(y_j)$ which solves the dispatch model OPF^* with input parameters y_j initialized to the current conditions, that is with zero investment. This recognizes attractive sites for generator upgrades and provides good starting points to MCP^* in the next step. After solving MCP^* , the whole process is repeated by using the investment solutions y_j obtained from the equilibrium as input parameters (feedback) into OPF^* , until we arrive at a stable solution. This method makes the problem tractable and our tests show it typically converging within 4 to 5 iterations, and in no more than 10.

During numerical testing, we observed that prior to convergence when

$q_j^\omega \neq \hat{q}_j^\omega$, the unit commitment decisions did not change because the non-convexity in the approximation function $\hat{C}_j(q_j^\omega, y_j)$ limited the space searched by a Newtonian based solver. That is to say that the equilibrium solution did not alter the active set of generators (unit commitment decisions) selected for dispatch and that this set only changed when the algorithm looped back to solve OPF* with the updated investment values y_j provided by MCP*. This brings us to the next section in which we exploit this solution structure to replace the approximation function $\hat{C}_j(q_j^\omega, y_j)$.

3.2.3 Restricted MCP

Based on the observations above, we introduce the following “Restricted MCP” (RMCP) equilibrium where $\tilde{C}_j(y_j, q_j^\omega, o_j^\omega) = o_j^\omega C_j(y_j, q_j^\omega)$. This updated version exploits the solution structure by replacing the approximation model MCP* with a restricted version of the original equilibrium model (3.3)-(3.10).

The key idea here is to fix integer variables o_j^ω , which leaves us with an MCP based on the original equilibrium model (3.3)-(3.10), thus allowing us to bypass the approximation function. The benefit of using RCMP is that our ability to directly form an MCP independent of any non-convex approximation function results in a more efficient solution process. In addition, the function $\tilde{C}_j(q_j^\omega, y_j, o_j^\omega | o_j^\omega)$ is now convex with respect to both q and y , and this eliminates the need to provide the MCP with starting values $\hat{q}_j^\omega$. Thus, the updated algorithm based on the RCMP is presented in Procedure 6.

RMCP (1st order conditions of:)

ROPF ($\forall \omega \in \Omega$):

$$\min_{q,z,\theta} \sum_j \tilde{C}_j(q_j^\omega, y_j, o_j^\omega | o_j^\omega) \quad (3.22)$$

$$\text{s.t.} \quad (3.4) - (3.8) \quad (3.23)$$

RFirm ($\forall f \in F$):

$$\min_{y \geq 0} \sum_{\omega} \pi_{\omega} \sum_{j \in J^f} \tilde{C}_j(q_j^\omega, y_j, o_j^\omega | o_j^\omega) + \eta r(y_j) \quad (3.24)$$

$$\text{s.t.} \quad h_f(y) \leq 0 \quad (3.25)$$

Procedure 6: Restricted MCP Solution Process

Initialization: $q_j^\omega = 0, y_j = 0 \quad \forall j \in N, \forall \omega \in \Omega$

1 Loop

 // Obtain $(\hat{q}_j^\omega, o_j^\omega)$ with y_j fixed, eqns.(3.3-3.8)

2 $(\hat{q}_j^\omega, o_j^\omega) \leftarrow \text{OPF}^{\text{UC}}(y_j) \quad \forall \omega \in \Omega$

3

4 **if** $\hat{q}_j^\omega = q_j^\omega$ **then** // convergence

5 break;

6 **else** // Solve RMCP with o_j^ω fixed

7 $(q_j^\omega, y_j) \leftarrow \text{RMCP}(o_j^\omega)$

Like before, the methods starts by the optimal dispatch problem based on the current state of the system, with $y_j = 0$. In this revised procedure, OPF* is replaced with the original OPF^{UC} model (3.3)-(3.8), and we then solve the RMCP which has input variables o_j^ω fixed, and the process is repeated until a stable solution is found.

With each successive iteration, note that the RMCP produces a new equilibrium, each of which is arrived at via different active sets of o_j^ω . The criteria for convergence however is a globally optimal dispatch and OPF solution given the investment decisions y_j for each generator. Since

$\tilde{C}_j(q_j^\omega, y_j, o_j^\omega | o_j^\omega)$ is strongly convex with respect to y at a fixed solution q , the firm's investment model is likewise globally optimal given the ISO's operational decision variables q_j^ω , provided that $\eta r(y_j)$ is also convex. Hence, the solution at convergence is a Nash equilibrium, from which we obtain the price response function, $p_i(x)$.

3.2.4 Comparison and Stability Analysis

A numerical comparison reveals that procedure 5 gives the same structural solutions as procedure 6, meaning that given initial starting points $\hat{q}_j^\omega$, both methods produce the same active set when solving MCP* and RMCP respectively. Slight perturbations are observed in the (q_j^ω, y_j) values returned by both algorithms, due to the approximation function $C_j^*(q_j^\omega, y_j)$ that is used in procedure 5, which could have cascading effects when the optimization problems do not have unique solutions. Procedure 6 was also found to converge slightly quicker, typically within 3-4 iterations, and in no more than 6. Therefore, we posit the use of procedure 6 over 5 because the elimination of approximation error makes the algorithm computationally more efficient.

A stability analysis using the 14 bus test case [70] introduced in Section 3.3 also revealed the need to regularize the lower level in order to manage issues with non-uniqueness of solution. The original IEEE dataset is comprised of 3 identical generators (3,6,8) which in certain situations can lead to multiple Nash equilibria and non-unique $p_i^\omega(x)$, as illustrated below.

Solution A: OPF^{UC}($y_j = 0$):

Scen	q_1	q_2	q_3	q_6	q_8
ω_1	2.17	0.42			
ω_2	3.09	0.53			
ω_3	2.52	0.43	0.85		0.85
ω_4	2.71	0.47	1.00	1.00	

ω_5	2.53	0.44	0.91	0.91	0.91
------------	------	------	------	------	------

Nash Equilibrium A:

Firm	y_1	y_2	y_3	y_6	y_8
f_1	253.95		46.05		
f_2		167.03		32.97	

Scen	q_1	q_2	q_3	q_6	q_8	$\sum_i d_i^\omega p_i^\omega$
ω_1	2.17	0.42				8811.06
ω_2	3.09	0.53				14500.07
ω_3	3.29	0.57	0.81			19234.14
ω_4	3.32	0.86	1.00			27020.60
ω_5	3.26	0.57	0.97	0.89		23637.04

Solution B: OPF^{UC}($y_j = 0$):

Scen	q_1	q_2	q_3	q_6	q_8
ω_1	2.17	0.42			
ω_2	3.09	0.53			
ω_3	2.52	0.43		0.85	0.85
ω_4	2.71	0.47		1.00	1.00
ω_5	2.53	0.44	0.91	0.91	0.91

Nash Equilibrium B:

Firm	y_1	y_2	y_3	y_6	y_8
f_1	243.25		40.42		16.33
f_2		167.24			

Scen	q_1	q_2	q_3	q_6	q_8	$\sum_i d_i^\omega p_i^\omega$
ω_1	2.17	0.42				8812.03
ω_2	3.09	0.53				14501.72
ω_3	3.29	0.57	0.81			19239.83
ω_4	3.32	0.86	1.00			27020.43

ω_5	3.30	0.57	0.96	0.86	23650.07
------------	------	------	------	------	----------

As a direct result of identical generator units, solutions A and B which have different dispatch decisions are both global solutions to the model $\text{OPF}^{\text{UC}}(y_j=0)$ of the same system. Consequently, they yield different Nash equilibriums with different $p_i^\omega(x)$ values. Implicitly, DFO algorithms require $p_i^\omega(x)$ to be a continuous function of x , and while this topic requires further study, we leave it to be explored in future work. Instead, Section 3.3 addresses this by adding slight perturbations to the dataset, thus ensuring a unique global solution to each system $\text{OPF}^{\text{UC}}(y_j)$.

3.3 Numerical Examples

The proposed methodology was applied on an adapted version of the 14 bus test case obtained from the University of Washington's Power Systems Test Case Archive [70]. Figure 3.3 provides a graphical depiction of this example, with generators located at buses 1, 2, 3, 6 and 8. In adapting this dataset for our purposes, we have defined active upper bounds on the transmission line arcs b_{ij} and divided the 5 available generators between the portfolio of two firms; Firm F1 controls the generators at buses 1, 3

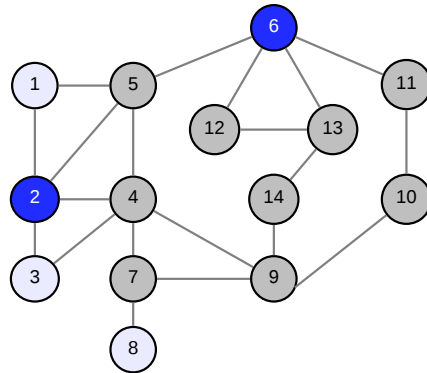


Figure 3.3: Graphical depiction of 14 bus example

and 8 while firm F2 operates the generators at buses 2 and 6, each firm with its own budget. Since we are considering unit commitment, each generator has also been assigned a non-zero minimum generating capacity as a fraction of the given maximum generating capacity, as well as a non-zero fixed cost parameter γ_j as a fraction of the unit's minimum operating expense (that is operation cost at its minimum generating capacity). Given the original dataset, it seems natural to extract information regarding generator size (large, medium, small) and based on this, appropriate investment parameters ψ_j have been assigned. The small generators 3, 6 and 8 have also slight perturbations added to their cost data in order to maintain uniqueness of OPF solution, as we had discussed in Section 3.2. Table 3.5 details the data enhancements that have been kept standard for all following numerical examples.

Table 3.5: Equilibrium Model Parameters

Description	Parameter	Value
Coefficient for $\underline{u}_j$ as a fraction of $\bar{u}_j$		0.3
Coefficient for γ_j as a fraction of $C_j(q_j = \underline{u}_j, y_j = 0)$		0.2
Firm investment budget (\$ units)	Firm F1 (Gens 1, 3, 8)	300
	Firm F2 (Gens 2, 6)	200
Generator upgrade parameter, ψ (\$ units)	Large generator (Gen 1)	200
	Medium generator (Gen 2)	150
	Small generator (Gens 3, 6, 8)	100

Table 3.6: Example 1: Scenario information

Scenario	ω_1	ω_2	ω_3	ω_4	ω_5
π_ω probability	0.3	0.3	0.2	0.1	0.1
d_j^ω coefficient	1	1.4	1.8	2	2.2

In addition, five distinct scenarios have been added to the dataset, and these are listed in Table 3.6. Since the existing data set provides us with only a single set of demand values for each node, each scenario ω is some

multiple of the original demand data, based on the coefficients provided in Table 3.6. These coefficients can be considered representative aggregations for seasonal fluctuations in electricity usage.

3.3.1 A function evaluation of $p_j^\omega(x)$ for fixed x

To illustrate the solution process of Procedure 6 which solves the equilibrium in the lower level hierarchy (3.3)-(3.8), we solve the model with a generous line limit of 2.0, or mathematically speaking $b_{ij}(x) = 2.0$ for all ij arcs in constraint (3.7) of the OPF^{UC} and ROPF models.

At iteration 1, the “OPF_UC” table contains the global solution to the OPF^{UC} model for the system’s current state (firm investment $y = 0$) in each scenario ω . The solutions are obtained using Baron [65] and from here, the active set o_j^ω is passed to the RMCP. Based on the resulting KKT conditions, an equilibrium (q_j^ω, y_j) is then returned by the PATH solver [26] [33] in conjunction with the GAMS EMP framework as discussed in [32]. This equilibrium is listed in “Table RMCP, iteration 1”. Following the algorithm outlined in Procedure 6, the steps are repeated, beginning with a new OPF^{UC} evaluation based on the updated firm response y_j , and if this results in a different active set o_j^ω , a new set of KKT conditions is generated for the RMCP. This iterative process continues until convergence is met, which in this case is at the third iteration.

Notice that in the first iteration, generator 3 is only dispatched in scenario ω_5 which has a low probability (0.1) of occurring. Firm f_1 ’s reluctance to invest in generator 3, combined with investment in all other generators, leads to a correction in the unit commitment dispatch in the next iteration. At this point, generator 3 is removed from the dispatch lineup, retired, as a result of its inability to compete in terms of cost efficiency.

Variation in the LMP values at different nodes (not provided) indicates that congestion occurs in scenarios ω_2 to ω_5 on line 1-2. This is determined

OPF_UC, iteration 1:

Scenario	q_1	q_2	q_3	q_6	q_8
ω_1	2.17	0.42			
ω_2	2.21	0.42			1.00
ω_3	2.50	0.43		0.74	1.00
ω_4	2.71	0.47		1.00	1.00
ω_5	2.84	0.50	0.37	1.00	1.00

RMCP, iteration 1:

Scenario	q_1	q_2	q_3	q_6	q_8
ω_1	2.17	0.42			
ω_2	2.61	0.45		0.57	
ω_3	2.87	0.50		0.30	1.00
ω_4	2.86	0.59		0.73	1.00
ω_5	2.86	0.60	0.30	0.94	1.00
Firm	y_1	y_2	y_3	y_6	y_8
f_1	248.87				51.13
f_2		153.55		33.14	

by the non-zero multiplier of z_{12}^ω , which is real power flowing along line (1-2) in scenario ω .

Finally, the Nash Equilibrium at convergence provides us with LMP function evaluations p_j^ω that can be returned to the upper hierarchy. Evaluating the expected total consumer cost at this solution, that is $\sum_\omega \pi_\omega \sum_i d_i^\omega p_i^\omega(x)$, yields 16414.

3.3.2 Transmission Line Expansion

With the understanding of what a p_i^ω function evaluation consists of, we now return to the problem of transmission line expansion in the long timescale hierarchy. Based on equations (3.1)-(3.2), we present the follow-

OPF_UC, iteration 2 and 3 (convergence):

Scenario	q_1	q_2	q_3	q_6	q_8
ω_1	2.17	0.42			
ω_2	3.01	0.62			
ω_3	2.96	0.70			1.00
ω_4	2.86	0.59		0.73	1.00
ω_5	2.88	0.82		1.00	1.00

RMCP, iteration 2 and 3 (convergence):

Scenario	q_1	q_2	q_3	q_6	q_8
ω_1	2.17	0.42			
ω_2	3.01	0.62			
ω_3	2.96	0.70			1.00
ω_4	2.86	0.59		0.73	1.00
ω_5	2.88	0.82		1.00	1.00
Firm	y_1	y_2	y_3	y_6	y_8
f_1	253.57				46.43
f_2		168.03		31.97	

Equilibrium:

Scenario	ω_1	ω_2	ω_3	ω_4	ω_5
OPF objective	7773	11658	16071	18403	20552
Wholesaler cost	8810	18207	19347	21354	23779

ing optimization model:

Transmission Expansion Model

$$\min_{0 \leq x \leq \delta} \quad \overbrace{\|x\|_1}^{\text{investment}} + \overbrace{\sum_{\omega} \pi_{\omega} \sum_i d_i^{\omega} p_i^{\omega}(x)}^{\text{consumer welfare}} \quad (3.26)$$

$$\text{s.t. } \|x\|_1 - \kappa \leq 0 \quad (3.27)$$

Variable x describes the amount of line expansion on a specific line with bound constraints δ (physical limitations) optimized over a small dimensional set X , and is subject to a budget of κ (3.27) where investment cost $\mu c(x)$ is simply the 1st norm of x with multiplier $\mu = 1$. Equation (3.28) is a result of applying the penalty function (3.21) to this model.

$$\min_{\substack{0 \leq x \leq \delta \\ x \in X}} \left\{ \|x\|_1 + \sum_{\omega} \pi_{\omega} \sum_i d_i^{\omega} p_i^{\omega}(x) + \sigma \max \left(\|x\|_1 - \kappa, 0 \right) \right\} \quad (3.28)$$

Based on its reported superior performance [64], we have selected the multi-coordinate search (MCS) algorithm [43] as our DFO solver for the optimization model, using equation (3.28) as the objective function and the parameter settings outlined in Table 3.7.

Table 3.7: Optimization Model Parameters

Description/Parameter	Value
Physical bound, δ	1.5
Investment budget, κ	2
Line limits, $b_{ij}(0)$	1.0
Penalty value, σ	1500
Set of decision variables, X	$\{x_{12}, x_{15}, x_{23}, x_{78}, x_{79}\}$

Here, the starting transmission line limits, $b_{ij}(0)$ have been lowered from the previous example to 1.0 for all lines, and the algorithm was run on a 5 dimensional problem defined by set X with a suitable penalty value of $\sigma = 1500$ that results in the satisfaction of constraint (3.27). When evaluating response variable $p_i^\omega(x)$, variable x is passed to the equilibrium model (3.3)-(3.10) where the transmission line limits in the OPF and RMCP (3.7) are updated by setting $b_{ij}(x_{ij}) = b_{ij}(0) + x_{ij}$. On its default settings, the MCS algorithm yields the solution X^* listed below.

$$X^* = \begin{pmatrix} x_{12}^* \\ x_{15}^* \\ x_{23}^* \\ x_{78}^* \\ x_{79}^* \end{pmatrix} = \begin{pmatrix} 1.04 \\ 9.9e^{-6} \\ 0 \\ 0 \\ 1.4e^{-5} \end{pmatrix} \quad \begin{array}{l} f(x^*) = 14944.63 \\ \text{Num. iterations} = 5 \\ \text{Num. evaluations} = 1298 \end{array}$$

For practicality purposes however, the solution we would implement in reality is X^{**} and although it yields a slightly worse objective value, the difference between $f(x^*)$ and $f(x^{**})$ is negligible.

$$X^{**} = \begin{cases} 1.04, & \text{for } x_{12} \\ 0 & \text{otherwise} \end{cases} \quad f(x^{**}) = 14944.75$$

Both objective values are clearly an improvement over the consumer cost evaluated in example A. An analysis of the OPF model indicates the existence of line congestion at the solution x^{**} , even though the budget constraint (3.27) is non-active. This seems counter-intuitive since network congestion restricts the ability of generation at low-priced locations to supply power to customers at high priced locations. This is further investigated below.

3.3.3 Benchmark Comparison

If we expect an unrestricted system (a network with inactive line limits) to allow for energy dispatch at its highest efficiency, it seems reasonable to

use the equilibrium solution of this unrestricted system as a benchmark comparison against the optimization solution. That is to say we take an evaluation of p_j^ω while disregarding the bounds of the b_{ij} line limits, and use it to calculate expected total consumer cost $\sum_\omega \pi_\omega \sum_i d_i^\omega p_i^\omega(x)$. This evaluates to 15803.82, which is inferior to the solution X^{**} .

This result reinforces the notion that easing congestion in transmission lines does not always benefit the monetary interests of the wholesale consumer, pointing towards a non-alignment of goals in the OPF operations with respect to consumer welfare. An analysis of the solution space reveals that the reason behind this non-intuitive behavior in the LMP prices p_j^ω , can be traced to a combination of the unit commitment variables and the cost function used in the dataset. Recall that the IEEE dataset provides quadratic functions $C_j^*(q)$ which describe a generator's total operational cost in producing q units of power. This translates to a linear function of increasing marginal cost as q increases. Thus, when an increase in transmission limits allow a generator to be dropped from the active dispatch set $\mathcal{O}_j^\omega$, the generators which increase production to pick up the slack will have an increased marginal cost. This directly impacts the LMP prices, as they are a reflection of the marginal cost of the marginal generator at that location. Therefore when line limits are sufficiently relaxed to affect a change in the active dispatch set $\mathcal{O}_j^\omega$, it is possible for the LMP prices p_j^ω to either increase or decrease based on the updated marginal cost of the updated marginal generator at that location.

With this understanding, we can conclude that when considering the monetary interests of the wholesale consumer, the MCS solution X^{**} , is superior to the solution provided by the unrestricted system. If we maintain this assumption of quadratic cost curves in a much larger system, it is possible that the non-alignment effect will be diminished due to an increase in the number of possible dispatch options. That is to say that any potential downside in altering the active set (i.e. dropping a generator

from the dispatch solution) can now be averaged over a larger number of generators. This specifically translates to a lesser increase in the marginal operating cost of any generator, when compared to the impact of a similar change in a smaller system.

3.4 Conclusion

In conclusion, this paper introduces a hierarchical and scenario based framework for informing long-term power planning decisions, and is designed to capture the richness of the model through transparent interactions between multiple agents who are making decisions on different timescales. By maintaining a clear partition of these timelines, and defining the interactions between the different agents and their decision processes in a transparent manner, this model is not only able to capture the complexity of the entire system, but is nimble enough to facilitate adaptation as the real process evolves.

Thus, we have implemented this framework to inform the transmission line planning process, and it incorporates grid operational constraints as well as short term firm behavior in response to planned changes. The long-term planning model was solved using a derivative-free solver [43], and we have proposed an iterative procedure to solve the shorter-timescale equilibrium which provides feedback to the long timescale optimization problem. The iterative process allows us to make the equilibrium (3.3)-(3.10) tractable, and provides us with a Nash equilibrium at convergence, which is typically in 3-4 iterations. Numerical tests run on the 14 bus IEEE dataset have provided a global solution to the expansion process, and analysis of the results indicate a non-alignment in the objectives of grid operation and consumer welfare, which we discuss in detail in subsection 3.3.3.

The general approach of this framework leaves open a number of

promising directions for future enhancement as its multi-agent and hierarchical nature allows for adaptation to different objectives, levels of detail, and to evolving systems. First of all, the objective of the planning model could be enhanced to include other social welfare measurements, to study their impact, if any, on the goal alignment of different agents. The growing impact of renewable energy, energy storage and ISO demand response policies are also other points of interest that may be studied further for their long term effect on grid operation and layout as a whole.

Complex models such as this are tricky to solve in very large instances, and a drawback of the current formulation is tractability in a realistic setting when considering a much greater number of nodes, scenarios and participants. Additionally, more accurate representations of the power system process will not just require greater detail, but increased complexity. Follow-up work could investigate ways to exploit the problem structure for more efficient formulations and solution methods, working toward a system that is able to demonstrate the benefits and drawbacks of a particular transmission line siting decision at an aggregate level.

While a lot of work from previous chapters focused on the accuracy of daily operations and implementing those tools in a realistic setting which is a known setting, this chapter attempts to address uncertainties far into the future while considering operational impact. It would be interesting to broaden the scope of the toolbox in such a manner that bridges the gap between these chapters by allowing the user to take their research of operations and implement it into a planning setting such as this. This may not necessarily mean augmenting the datasets with fabricated data, but rather create a mechanism that helps the user include data addressing specific uncertainties or situations that interest them.

4 REAL-TIME CO-OPTIMIZATION OF ENERGY AND RESERVE MARKETS WITH SECURITY CONSTRAINTS

4.1 Introduction

Determining an efficient dispatch of generation resources is a complex question which can necessitate balancing multiple goals depending on the situation. Two most common themes of grid management are economic efficiency and system reliability. Many grid reliability solutions fall under the general umbrella of security constrained models, such as security constrained optimal power flow (SCOPF) [19, 75], security constrained economic dispatch (SCED) [53, 38] and security constrained unit commitment (SCUC). Generally, the objective of such models is to minimize operating cost while maintaining solution feasibility after the loss of a major component, including failure of line, generator or transformer. An “N-1” security criterion takes into account all major components and is met when the dispatch solution allows the network to maintain system stability in the event any one of those failures occur.

Security constraints can be preventive or corrective. Preventive constraints result in a conservatively operated system with a base-case dispatch that can survive the occurrence of the modeled contingencies. Alternatively, corrective constraints reconfigure the post-contingency system with fast-acting equipment [18], by taking recourse action to return the system to either a steady or a sustainable state. Therefore, security constrained models with corrective action are closely associated with some ancillary services, namely the reserve market. In real-time dispatch models, the regulation market deals with real-time frequency control while the real-time reserve market supports contingency recovery, especially under loss of generation resources [78]. With reference to energy reserves, co-optimization has been extensively discussed in the literature as a way

of reducing grid costs and promoting economic efficiency. Although the methods and motivations differ, co-optimization is well documented to have a positive effect on system efficiency and pricing [4, 20], and has real application purposes among Independent System Operators [54, 78, 54].

While co-optimization and security constrained models are clearly both highly studied topics, both separately and in conjunction with one another [5, 11], they are rarely combined in a single paper when dealing with real-time applications. Most likely reasons for this are due to model complexity and size. In these cases, we can assume that unit commitment is not a factor and that the time limits for solving the model is very strict, roughly a 5 minute window [54, 74]. Work exploring these themes for real-time markets [6] tend to struggle with the dimensionality of the model and rely on smaller datasets to test their theories. Recent papers by Street *et. al.* [68] and Ding *et. al.* [24] circumvent the issue of tractability using robust optimization which avoids explicitly modeling each contingency by using the worst-case scenario as a solution guideline.

In this chapter, we focus on combining the concerns of reserve co-optimization with grid reliability by solving an “N-1” security constrained model designed for real-time energy dispatch. A key difference between this and previously mentioned work is the explicit modeling of contingencies, and the ability of our solution process to neatly deal with potentially infeasible scenarios. This feature is particularly useful because the topology of a real-time system may change throughout the day with switching events and changing status of circuit breakers [48, 15, 35]. As such, pre-processing of a contingency list may prove difficult unless it can be implemented quickly when responding to topology changes. As our algorithm is not based on any pre-conceived assumptions about network topology, it is well suited for this environment.

Our model, which is based on Benders decomposition techniques, is detailed in Section 4.2. Clearly, a traditional “N-1” SCED with correc-

tive action is too large to solve using traditional means for realistic sized datasets, and is difficult to solve even when using classical decomposition. By using a modified Benders algorithm and harnessing the power of parallel processing as outlined in Section 4.3, we are able to solve realistic sized datasets in real-time. Beyond the obvious computational advantage of solving a large model in real-time, the model presented in Section 4.2 benefits from its clear outline of the relationship between energy reserves and contingency management. Numerical examples in Section 4.4 show that this setup results in a more efficient dispatch of reserves that only targets locations necessary to handle network vulnerabilities. Section 4.6 concludes with some future points of debate.

4.2 Model Formulation

Set	Description	Indices
$\mathcal{G}$	Generators	g
$\mathcal{T}$	Time (discrete timepoints)	t
$\mathcal{K}$	Contingencies	k
$\mathcal{I}$	Iteration (for Benders heuristic)	i, j
$(\mathcal{K}, \mathcal{I})_{\text{cut}}$	Cuts generated by violated subproblems	ki, kj
$\mathcal{K}_{\text{M}}$	Set of subproblems added to the Master model	k
$\mathcal{K}_{\text{L}}$	Set of line failure scenarios	k
$\mathcal{K}_{\text{G}}$	Set of generator failure scenarios	k

Table 4.1: Description of Sets

4.2.1 Background

Referencing the model in [53], the general form of SCED with post-contingency corrective action (or recourse) can be written as (4.1)-(4.7). Variables (x_0, u_0) represent the state and control variables of the base case (i.e. pre-contingency) dispatch solution to the optimal power flow (OPF) problem,

Variables	Description
x_0	State variables for basecase solution (e.g. lineflow)
u_0	Control variables for basecase solution (e.g. generation)
x_{kt}	State variables for contingency k at timepoint t
u_{kt}	Control variables for contingency k at timepoint t
v_{ki}	Violation of cut at contingency k and iteration i
s_{kt}	Constraint violation of subproblem k at timepoint t
Functions	Description
c_0, f_0, h_0	Objective and constraint functions for basecase
f_k, h_{kt}	Constraint functions for contingency k at timepoint t
Parameters	Description
$\tilde{\lambda}_{ki}$	Lagrangian multiplier of (4.31)
$\tilde{w}_{ki}$	Objective value of subproblem k at iteration i
$\hat{u}_i$	Solution of Master control variable at iteration i
δ_t	Physical constraints on control variables at time t
M	Penalty variable for constraint violations of v_{ki}

Table 4.2: Description of Parameters and Variables

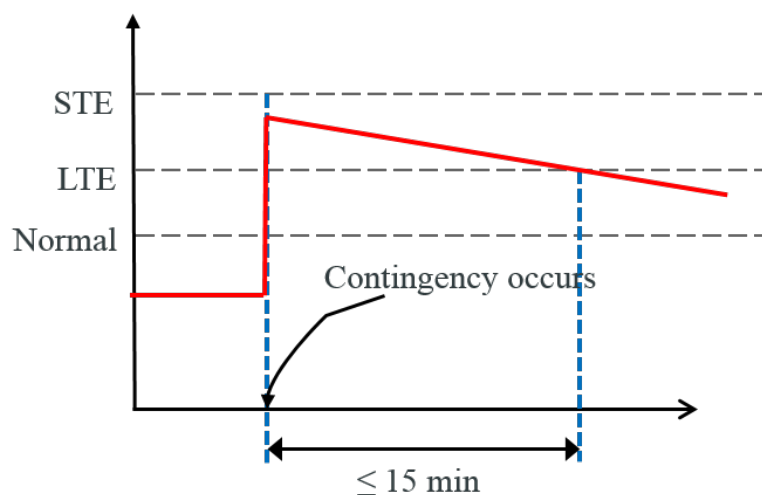


Figure 4.1: Example of post-contingency line limit requirements

which we base on DC powerflow equations (4.2)-(4.3), a reasonable assumption when exploring a model for real-time dispatch. A minimum cost objective function $c_0(x_0, u_0)$ using the base case dispatch is solved (4.1). Equations (4.4)-(4.7) model the recourse solution after a contingency solution where time plays an explicit role with discrete checkpoints, $t \in \mathcal{T}$. (4.4)-(4.5) describe the updated DCOPF equations after a contingency occurs, while (4.6) provides physical limits on available corrective action. Generators for example are constrained by their ramping capabilities, and line limits may also be relaxed due to the exigent circumstances as shown in 4.1. Linking both pre and post-contingency dispatch solutions is (4.7) where $t = t_0$ represents the immediate checkpoint post-contingency and (4.7) indicates that control variables u_{k,t_0} cannot abruptly change. Note however that the state variables x_{k,t_0} will instantaneously change according to physical laws.

$$\min_{x,u} \quad c_0(x_0, u_0) \quad (4.1)$$

$$\text{s.t.} \quad f_0(x_0, u_0) = 0 \quad (4.2)$$

$$h_0(x_0, u_0) \leq 0 \quad (4.3)$$

$$f_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K} \quad (4.4)$$

$$h_{kt}(x_{kt}, u_{kt}) \leq 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K} \quad (4.5)$$

$$|u_{kt} - u_{k,t-1}| \leq \Delta_t \quad t = 1, \dots, T, k \in \mathcal{K} \quad (4.6)$$

$$u_{k,t_0} - u_0 = 0 \quad \forall k \in \mathcal{K} \quad (4.7)$$

There is generally some flexibility in post-contingency states, and in some cases, power flow along lines may temporarily exceed normal line ratings, so long as they are returned to regular operating rules within an allotted time. ISO New England uses four levels of thermal capacity ratings for transmission facilities: Normal, Long Time Emergency (LTE),

Short Time Emergency (STE) and Drastic Action Limit (DAL), with increasing rating numbers [53]. An unexpected line failure for example will inevitably result in instant line flow surges (a state variable) on other parts of the network, until generation levels (a control variable) are adjusted accordingly. In the example illustrated by Figure 4.1, it is safe for a line to temporarily operate at the STE limit, as long as it is returned to the LTE limit within 15 minutes. Eventually of course, all lines should return to their normal operation level.

With an “N-1” security criterion, it is easy to see how the dimension of this model can very quickly become intractable, especially in real-life applications. In addition, equations (4.4)-(4.7) assume that a feasible solution exists for all contingencies $k \in \mathcal{K}$. Based on that assumption, a generalized Benders decomposition scheme reformulates (4.1)-(4.7) into (4.8)-(4.17). Note that s_k is an artificially added variable to model constraint violations resulting from the base case solution (x_0, u_0) , and it is easy to see that for a solution to be feasible, all values for $w_k(u_0)$ must equal 0.

Benders Master:

$$\min_{x_0, u_0} c_0(x_0, u_0) \quad (4.8)$$

$$\text{s.t. } f_0(x_0, u_0) = 0 \quad (4.9)$$

$$h_0(x_0, u_0) \leq 0 \quad (4.10)$$

$$w_k(u_0) \leq 0 \quad \forall k \in \{\mathcal{K}\} \quad (4.11)$$

Benders Subproblem ($\forall k \in \mathcal{K}$):

$$w_k(u_0) = \min_{x_{kt}, u_{kt}, s_{kt}} \|s_k\| \quad (4.12)$$

$$\text{s.t. } g_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T} \quad (4.13)$$

$$h_{kt}(x_{kt}, u_{kt}) \leq 0 \quad \forall t \in \mathcal{T} \quad (4.14)$$

$$|u_{kt} - u_{k,t-1}| - s_{kt} \leq \Delta_t \quad t = 1, \dots, T \quad (4.15)$$

$$u_{k,t_0} - u_0 = 0 \quad \perp \tilde{\lambda}_k \quad (4.16)$$

$$s_{kt} \geq 0 \quad \forall t \in \mathcal{T} \quad (4.17)$$

Unfortunately, it may not be realistic to assume or impose feasibility for all possible contingencies, especially when considering the “N-1” security case. Figure 4.2 illustrates two types of infeasibilities - (a) intrinsically infeasible contingencies which are unable to be satisfied with any base case solution, and (b) conflicting cuts whereby two or more contingencies may not be simultaneously feasible. The general Benders decomposition method (4.8)-(4.17) is clearly unable to identify conflicting contingencies (Figure 4.2b). As for intrinsically infeasible contingencies (Figure 4.2a), we would expect the Benders subproblem to always be infeasible for any solution of (x_0, u_0) , were it not for slack variable s_k in the subproblem. While some intrinsically infeasible contingencies will still result in infeasible subproblems, the relaxed subproblem (4.12)-(4.17) does not guarantee this.

It should also be noted that (4.11) is difficult to impose directly. However, based on LP duality, we can show that for any point $\bar{u}_0^i$, we can provide a linear function that underestimates $w_k(u_0)$, where $\tilde{\lambda}_k$ is the Lagrangian multiplier of (4.16) at $\bar{u}_0^i$.

$$\tilde{w}_k(\bar{u}_0^i) + \tilde{\lambda}_k(u_0 - \bar{u}_0^i) \leq w_k(u_0) \quad (4.18)$$

Hence, the below statement is a necessary condition to satisfy (4.11) and can be used in its place.

$$\tilde{w}_{ki} + \tilde{\lambda}_{ki}(u_0 - \bar{u}_0^i) \leq 0 \quad (4.19)$$

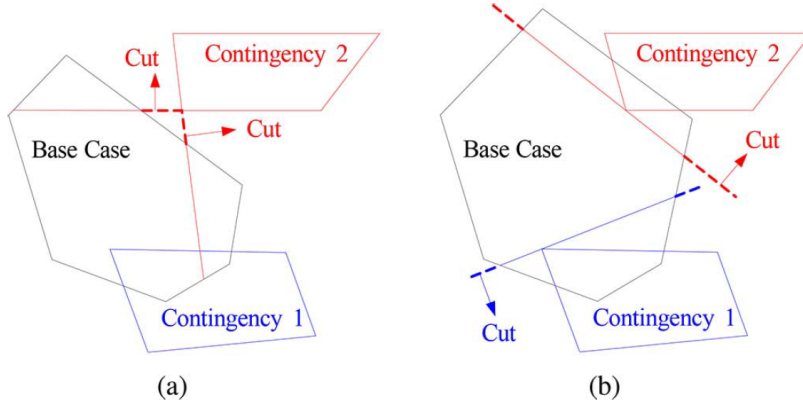


Figure 4.2: (a) Contingency 2 is intrinsically infeasible as there is no overlap in the feasible region of the base case with its corresponding subproblem. (b) Both contingencies 1 and 2 are not simultaneously feasible with the base case and create conflicting cuts.

With this and potential contingency infeasibilities in mind, Liu and Ferris [53] propose an iterative heuristic that solves the following modified Benders decomposition model (4.20)-(4.32) which we discuss in the next subsection.

4.2.2 Iterative Modified Benders

At iteration j :

Modified Benders Master:

$$\min_{x_0, u_0} c_0(x_0, u_0) + \sum_{k \in \mathcal{K}} M v_k \quad (4.20)$$

$$\text{s.t. } f_0(x_0, u_0) = 0 \quad (4.21)$$

$$h_0(x_0, u_0) \leq 0 \quad (4.22)$$

$$\tilde{w}_{ki} + \tilde{\lambda}_{ki}(u_0 - \bar{u}_0^i) - v_k \leq 0 \quad \forall ki \in \{\mathcal{K}, \mathcal{I}\}_{\text{cut}} \quad (4.23)$$

$$f_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K}_M \quad (4.24)$$

$$h_k(x_{kt}, u_{kt}) - v_k \leq 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K}_M \quad (4.25)$$

$$v_k \geq 0 \quad \forall k \in \mathcal{K} \quad (4.26)$$

Modified Benders Subproblem ($\forall k \in \mathcal{K}$):

$$\tilde{w}_{kj} = \min_{x_{kt}, u_{kt}, s_{kt}} \|s_{kt}\|_\infty \quad (4.27)$$

$$\text{s.t. } g_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T} \quad (4.28)$$

$$h_{kt}(x_{kt}, u_{kt}) \leq 0 \quad \forall t \in \mathcal{T} \quad (4.29)$$

$$|u_{kt} - u_{k,t-1}| - s_{kt} \leq \Delta_t \quad t = 1, \dots, T \quad (4.30)$$

$$u_{k0} - \bar{u}_0^j = 0 \quad \perp \tilde{\lambda}_{kj} \quad (4.31)$$

$$s_{kt} \geq 0 \quad \forall t \in \mathcal{T} \quad (4.32)$$

In this iterative algorithm, solving the Master model (4.20)-(4.26) yields a solution for state and control variables (x_0, u_0) in the base case at iteration $j \in \mathcal{I}$. Like (4.2)-(4.3) in the generalized SCED model, equations (4.21)-(4.22) represent network DCOPF constraints, including variable bounds.

The Master is solved, and the solution is passed to the set of subproblems by setting $\bar{u}_0^j = u_0$. The input $\bar{u}_0^j$ is then used to solve the subproblems $k \in \mathcal{K}$ at iteration j . If some subproblem returns a positive objective, $\tilde{w}_{kj} > 0$, this indicates that the basecase solution $\bar{u}_0^j$ found in iteration j , was not feasible for contingency k , leading to the feasibility cut in (4.23).

Based on Liu and Ferris' numerical experiments [53], they found that a simple iterative Benders algorithm in this case did not always converge quickly, and attributed the problem to specific contingencies deemed "difficult to resolve". More accurately, these contingencies required a higher number of iterations and added cuts in order for the Master to converge, if at all. Equations (4.24)-(4.25) directly add these contingencies $k \in \mathcal{K}_M$ to the master model. We leave the details regarding the identification of these contingencies to Section 4.3.

Finally, the master objective function (4.20) also includes a penalty term Mv_k , where v_k is a positive slack variable that allows the model to identify and remove infeasible contingencies. Conflicting contingencies are identified using (4.23) while the intrinsically infeasible contingencies that fail to yield an infeasible subproblem, as discussed in Section 4.2.1, are identified in (4.25). On the other hand, infeasible subproblems are easily identified and removed in the first iteration of the heuristic. A solution is found when the base case solution is able to satisfy the subproblems of all remaining feasible contingencies.

4.2.3 Differentiating contingency events

As previously mentioned in Section 4.1, co-optimization and security constrained models are rarely combined. To clarify the roll of ancillary services in the case of equipment failure, it is important to distinguish between the different contingencies. For the purposes of this chapter, we will only consider two types of contingency events: line and generator failure, and how they relate the the reserve market. Reserves are typically

included in day-ahead planning in case of system failures, and the set-aside capacity is used to handle recourse actions in case of equipment failures. There are multiple types of reserves, and common examples include spinning, non-spinning and backup reserves. Spinning reserves are extra generation capacity from generators already connected to the grid, while non-spinning reserves are extra generation capacity that are not connected to the grid but can be brought online with slight delays (e.g. 5-15 minutes). Backup reserves require a longer startup time (e.g. 30-60 minutes), but can be brought online to alleviate other short-term reserves when faced with longer system outages. In addition to different types of reserves, ISOs may also impose zonal constraints, defining for example a minimum reserve amount within a densely populated zone, in addition to a network-wide minimum. These zonal and reserve type requirements differ among ISOs and is necessary to maintaining a reliable grid.

The co-optimization of reserves goes hand in hand with considering generator failure contingencies as in practice, both line and generator failures are treated differently. At ISO New England, post-contingency action for line failures provide higher flexibility in line limits, but continue to reallocate reserves in the recourse action. This allows the system an extra reliability fallback, where the solution is able to accommodate a second failure (of a generator). Unexpected generator failures result in a sudden system imbalance, where total system demand exceeds total system supply. Therefore, these types of contingencies are allowed to tap into the pre-allocated base-case reserves as part of their post-contingency action in order to quickly make up the loss of power, but should return to a stable operating point in a much quicker time period. Since reserve amounts are generally used to accommodate generator failures, the combination of those these elements is a natural extension. Based on this, we extend the model in Section 4.2.2 by passing reserve information to generator related subproblems, but not line related subproblems. The updated model that

includes reserves and generator contingencies is shown in (4.33)-(4.46).

Here, the vector of control variables u_0, u_k also cover (where necessary) allocation of reserves. The reserve solution from the basecase does not affect the line contingency subproblem, but does affect the generator contingency subproblem, as indicated by equation (4.43). Recall that (4.34)-(4.35) are post-contingency DCOPF equations including variable bounds. In order to ensure reserves are properly allocated where they are needed (i.e. contingency recovery), the reserve solution in basecase becomes a variable bound for corrective action in the subproblem. In other words, post-contingency corrective action is limited as such: $u_{kt}^{\text{Gen}} \leq u_0^{\text{Gen}} + u_0^{\text{Reserve}}$. By adding this detail and updating the generator contingency cut in (4.37), the model guarantees that reserves are well distributed for the sole purpose of contingency handling.

At iteration j :

Modified Benders Master:

$$\min_{x_0, u_0} c_0(x_0, u_0) + \sum_{k \in \mathcal{K}} M v_k \quad (4.33)$$

$$\text{s.t. } f_0(x_0, u_0) = 0 \quad (4.34)$$

$$h_0(x_0, u_0) \leq 0 \quad (4.35)$$

$$\tilde{w}_{ki} + \tilde{\lambda}_{ki}(u_0 - \bar{u}_0^i) - v_k \leq 0 \quad \forall ki \in \{\mathcal{K}_L, \mathcal{I}\}_{\text{cut}} \quad (4.36)$$

$$\begin{aligned} & \tilde{w}_{ki} + \tilde{\lambda}_{ki}(u_0 - \bar{u}_0^i) \\ & + \sum_t \tilde{\gamma}_{kit}(u_0 - \bar{u}_0^i) - v_k \leq 0 \quad \forall ki \in \{\mathcal{K}_G, \mathcal{I}\}_{\text{cut}} \end{aligned} \quad (4.37)$$

$$f_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K}_M \quad (4.38)$$

$$h_k(x_{kt}, u_{kt}) - v_k \leq 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K}_M \quad (4.39)$$

$$v_k \geq 0 \quad \forall k \in \mathcal{K} \quad (4.40)$$

Line contingency Subproblem ($\forall k \in \mathcal{K}_L$):

Solve (4.27)-(4.32)

Generator contingency Subproblem ($\forall k \in \mathcal{K}_G$):

$$\tilde{w}_{kj} = \min_{x_{kt}, u_{kt}, s_{kt}} \|s_{k\cdot}\|_\infty \quad (4.41)$$

$$\text{s.t. } g_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T} \quad (4.42)$$

$$h_{kt}(x_{kt}, u_{kt}, \bar{u}_0^j) \leq 0 \quad \forall t \in \mathcal{T} \quad (4.43)$$

$$|u_{kt} - u_{k,t-1}| - s_{kt} \leq \Delta_t \quad t = 1, \dots, T \quad (4.44)$$

$$u_{k0} - \bar{u}_0^j = 0 \quad \perp \tilde{\lambda}_{kj} \quad (4.45)$$

$$s_{kt} \geq 0 \quad \forall t \in \mathcal{T} \quad (4.46)$$

An observation of this model is that if the reserve information was not communicated between master and subproblem in (4.37) and (4.43), no reserve would be allocated based on the minimum objective function (4.33), unless zonal reserve requirements were added. Without zonal requirements, the minimum amount of reserve is distributed only as needed to handle post-contingency action. This could lead to huge operational savings. Another observation is that care must be taken in determining the correct feasibility cut in the master because different contingencies have different post-contingency rules. The cut applied by line contingency subproblems (4.36) only affects generation levels for the master model, while the cut derived from generator contingency subproblems (4.37) affects both generation level and reserve allocation.

The following section details our proposed heuristic for solving this model, which despite the decomposition, can still be complex to solve in real-time especially when applied to realistic sized datasets.

4.3 Solution method

4.3.1 Algorithm Overview

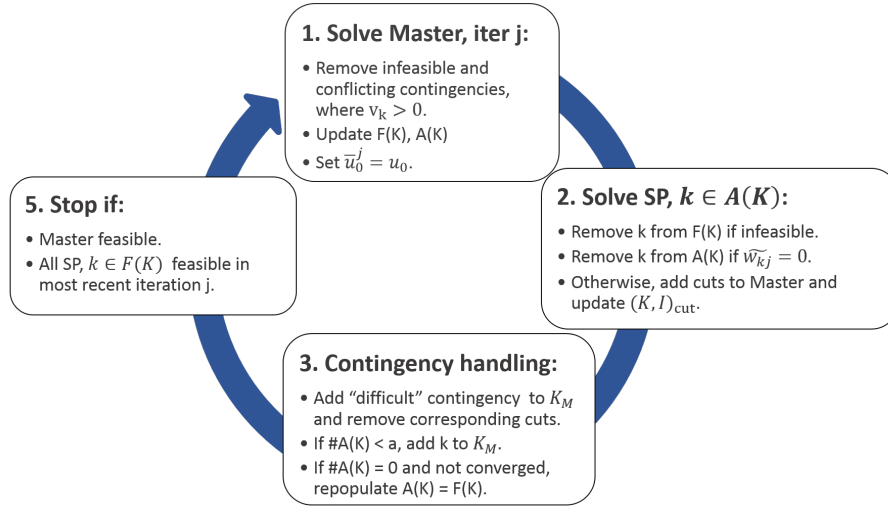


Figure 4.3: Flowchart showing an overview of the Parallel Benders algorithm

For our proposed iterative process, we introduce the notion of a feasible contingency set $k \in F(K)$, and an active contingency set $k \in A(K)$, where $A(K) \subseteq F(K)$. The feasible contingency set only includes feasible contingencies (see example of infeasible contingencies in Figure 4.2) while the active contingency set refers to contingencies which are actively considered in iteration j . Also recall the master contingency set K_M found in (4.38)-(4.39), which is populated with "difficult to resolve" contingencies.

The flowchart in Figure 4.3 provides a birds eye illustration of our proposed heuristic. It begins with finding a feasible Master solution. At the first iteration, $j = 1$, both feasible and active sets are set to the entire contingency list, $A(K) = F(K) = K$. In future iterations, $j > 1$, when the master finds infeasible or conflicting contingencies ($v_k > 0$), it proceeds to remove them and resolve the master model in order to get a feasible (x_0, u_0) solution. This master solution is then passed to the subproblems

by setting $\bar{u}_0^j = u_0$.

The contingency subproblems within the active set are then solved using parallel processing and based on their outcome, are handled accordingly. It is obvious that infeasible subproblems should be removed from the feasible set $F(\mathcal{K})$, while subproblems with violated constraints, $\tilde{w}_{kj} > 0$, should result in cuts added to the master model. The notion of an active set $A(\mathcal{K})$ however, came from an observation in [53], where the majority of subproblems found to be feasible at an iteration ($\tilde{w}_{kj} = 0$) had a tendency to remain feasible in subsequent iterations (e.g. $\tilde{w}_{k,j+1} = 0$). Removing these subproblems from the active set in future iterations results in huge computational advantages as the algorithm does not have to process all contingencies $k \in F(\mathcal{K})$ for each iteration j . In the contingency handling portion of the algorithm, the following is achieved:

1. Identify and deal with “difficult” contingencies.
2. Ensure that all valid contingencies $k \in F(\mathcal{K})$ are feasible at the master base case solution $w_k(u_0) = 0$.

Details on contingency handling of “difficult” contingencies is provided in the upcoming section, 4.3.2. As for the second item on that list, it is satisfied by repopulating the active set $A(\mathcal{K}) = F(\mathcal{K})$ when the active set is empty. By doing so, all of the subproblems that were temporarily removed from the active set in earlier iterations in Step 2 are re-evaluated to ensure feasibility at the final solution. The algorithm converges when the master model and all valid subproblems are found to be feasible in the final iteration, $w_k(u_0) = 0, \forall k \in F(\mathcal{K})$.

4.3.2 Contingency Handling

As mentioned earlier, the majority of subproblems found to be feasible at an iteration tended to remain feasible in subsequent iterations. On the flip

side, it was also observed that a small number of “difficult” contingencies tended to stay in the active list for a long time, creating algorithmic inefficiencies. The obvious inefficiency is when a contingency needs to generate multiple cuts in order to generate a feasible base case solution, which leads to many extra iterations. In the case of line failure contingencies, “difficult” contingencies were identified by looking at the size of the active set $A(\mathcal{K})$. Due to the iterative process, feasibility cuts resulted in a smaller active set with each subsequent iteration. When the number of active elements was small enough, these “difficult to resolve” contingencies are added directly to the master model, using set $\mathcal{K}_M$ in equations (4.38)-(4.39). Tests show that this drastically improves computational time.

However, supplementary numerical testing found that our addition of generator failures had a huge adverse effect on the existing process, not just due to it being a larger model, but also because these subproblems reacted differently compared to the original line failure contingencies. We observed that feasibility cuts from generator failure subproblems were a lot less effective at improving the base case solution when compared to cuts from line failure subproblems, and each additional cut did very little to improve the base case solution. As a result, iterations were beginning to produce parallel cuts, and thus failing to converge. The results also indicated another difference between line and generator failures. In most of the testcases, it was possible to satisfy all, if not most of the generator failure contingencies. This is in stark difference to solving for line failure contingencies, where 16%-18% of those scenarios are found to either be intrinsically infeasible or conflicting with one another in the final solution (recall Figure 4.2).

The latter observation leads to the belief that given sufficient and well allocated resources, generator failures as we have modeled them, have greater flexibility when taking corrective actions. In other words, there is a larger overlap between the feasible regions of the master model and a

generator contingency subproblem. This also lends credence to our first observation, where cuts added by generator contingency models were less effective than line contingency models. Recall that due to decomposition methods, the subproblems are relaxations of the original (4.4)-(4.7). This leads to feasibility cuts that are not as tight as they could be, even when the subproblem was only minimally relaxed.

By extending the same trick of including generator contingencies directly to the master model, it was observed that the addition of a single generator contingency resulted in many other generator contingencies becoming feasible. More careful analysis revealed that feasibility for all generator subproblems could be found by adding a small subset of generator contingencies to $\mathcal{K}_M$. The challenge was to identify which subset would be the most useful. Our solution was to select the contingency with the highest violation at the end of each iteration, $k : \max_k \tilde{w}_{kj}, \forall k \in A(\mathcal{K})$ and add that to the master model contingency list by updating set $\mathcal{K}_M$. This essentially identifies the subproblem in the worst position with respect to the basecase solution at the end of every iteration and denotes it to be a “difficult contingency”, then adds it to the master model. The effect of this policy is that a “spatially distributed subset” of generator subproblems is added to the Master model, which significantly speeds up the solution process. When a generator contingency k is added to the Master model at the end of iteration j , the Master model at iteration $j + 1$ finds a solution that is beneficial to the contingency analysis of other nearby generators. This is because adding contingency k results in the reallocation of reserves to where they are needed (close by to the failure defined by contingency k). Thus in the iteration $j + 1$, we expect not only a reduction in the number of active generator subproblems, but for contingency $k : \max_k \tilde{w}_{k,j+1}$ to identify a different problem area based on the most current base case solution.

Note that the subproblem with the largest violation is not guaranteed

Contingency type	Time(min)		Line Limits		Reserves
	t_0	t_1	t_0	t_1	
Line failure	0	15	STE	LTE	reallocated used
Generator failure	0	10	LTE	LTE	

Table 4.3: Differences in post-contingency corrective actions

# Cont	GAMS runtime (s)			JAVA runtime (s)		
	Num threads			Num threads		
	4	8	16	4	8	16
400	146.76	117.33	110.1	97.88	65.23	46.40
1000	284.85	181.68	138.48	142.91	119.22	79.98
2896	755.86	386.25	238.41	419.00	233.01	161.20

Table 4.4: Comparing runtimes of JAVA and GUSS implementations, based on size of contingency list and number of threads

to be a generator contingency, and while we do allow the algorithm to add line failure scenarios to the Master model, as established earlier, numerical testing indicates that cuts generated by line contingencies are more effective in improving basecase solutions. To manage the differences between both types of contingencies, we also propose the two following changes:

1. Find the largest violated generator subproblem at the end of iteration $j = 1$ to add to the Master.
2. Assigning a greater weight to generator violations when determining the largest violation. For example, using $m\tilde{w}_{kj}$ where
$$\begin{cases} m > 1 & \text{if } k \in \mathcal{K}_G \\ m = 1 & \text{o/w} \end{cases}$$

4.3.3 Algorithmic enhancements

In addition to the changes discussed in Sections 4.2.3 and 4.3.2, we also made the following alterations to the material in [53], with the express pur-

pose of increasing algorithmic efficiency and applying the methodology in real-time.

1. *Emphasizing practical implementation:* Our current model only considers one future recourse timepoint t_1 compared to the original 3 timepoints in [53]. As a result, our models do not consider the Drastic Action Limit (DAL) which is higher than the short term emergency (STE) limit. Our reasons for this is to better translate the original SCED implementation for real-time use. Given constant system changes and alternative avenues for corrective action, we believe that in this (real-time) context, it is sufficient to find a solution that helps return the system to stability. We define a stable solution as one that is able to (a) meet demand and (b) operate at long term emergency (LTE) line limits. Depending on the line, most LTE limits can typically be sustained between 30 minutes to 4 hours. Once a stable solution has been implemented, manual action can be taken by the system operator to eventually return the system to its “steady state” and the SCED cycle restarted.

Table 4.3 compares the differences in our assumptions and handling of post-contingency actions for both line and generator failures. In the case of line failures, line flows are allowed to reach STE limits at the immediate occurrence of a contingency, to account for the fact that state variables will instantly change according to physical laws. By the next time checkpoint, $t_1 = 15$ minutes, the line limit allowances are reduced to long term limits (LTE), and reserves are reallocated as needed. Generator failures on the other hand do not incur drastic state variable changes that will exceed the default line limits. Instead, the focus is on quickly restoring balance such that total system supply is able to meet total system demand. Therefore, we set the next time checkpoint to $t_1 = 5$ minutes and allow line flow

Testcase	#(Line,Gen)	Basecase Obj	(N-1) Line Failure				(N-1) Line & Gen Failure		
			Time(s)	Objective	#Feas(Lines)	Iter	Time(s)	Objective	#Feas(Line,Gen)
2383wp	2896, 327	1794184.67	178.41	1817855.19	2359	6	332.87	1866367.92	2359, 327
2736sp	3269, 420	1276033.67	81.04	1288259.79	2749	3	259.368	1292701.45	2749, 420
2737sop	3269, 399	764008.59	89.76	764733.43	2755	3	252.91	768277.41	2755, 399
2746wop	3307, 514	1178163.98	115.53	1184835.63	2796	3	597.82	1188258.55	2796, 514
2746wp	3279, 520	1587849.30	179.12	1618225.95	2729	4	471.07	1621963.48	2729, 520

Table 4.5: Comparison of Numerical Results using 24 threads

to increase to LTE limits. As discussed earlier, reserves are utilized instead of reallocated in this case.

2. *Parallel processing improvements*: In addition to implementing the Benders decomposition framework, the crux of the original algorithm was the speed up gained by solving the large number of subproblems in parallel batches by using the GUSS facility [16] within GAMS, and harnessing the multi-core and multi-threading capability of our systems. By exploiting our knowledge of GUSS' background processing, we were able to significantly reduce the time and hardware requirements of the original algorithm by programming better thread handling functions in GAMS.

However, in order to accommodate the increased complexity of the added reserves and generator contingencies, we switched the implementation to using JAVA API for GAMS [2]. This change yielded time saving benefits as high as 40% in some cases. A more thorough analysis of the time breakdown suggests that much of this is attributed to better load balancing when solving the subproblems. The GUSS facility requires that we pre-assign the batches of subproblems to each thread. On the contrary, JAVA API for GAMS immediately allocates remaining subproblems to the next available thread. Thus, long running jobs do not unnecessarily hold up or blocks the remaining jobs in queue. While the results in Table 4.4 clearly show a benefit in using the new JAVA API implementation, the benefit of doing so

is strongly correlated to the size of contingency list as well as the number of threads. The JAVA speed up is less pronounced in small instances or when there are a higher number of threads. In the latter case, the load balancing is less likely to be badly unbalanced because a higher number of available threads equates to smaller batches of pre-assigned jobs.

3. *Improved feasibility checks and handling:* In the original work [53], a feasibility checker was initiated on each contingency before it was added to the master contingency set $\mathcal{K}_M$. This was an extra precaution to ensure that a contingency was not intrinsically feasible before being added to the master model. In our more recent experiments, we find that a more effective approach is to skip the feasibility checker, and allow variable slack v_k to directly identify any infeasible contingencies (4.38)-(4.39). This action is initiated when the active list is small, $|A(\mathcal{K})| \leq a$. In our experiments, we find that a good choice for parameter a is $3 \leq a \leq 4$.

In the case of generator contingencies however, our methods necessitate a slightly altered approach for feasibility checking. Recall from Section 4.3.2, that the largest violated contingency was added to set $\mathcal{K}_M$ in the master model at the end of each iteration j . Typically, this corresponds to a generator failure contingency. We also discussed the difficulty that the relaxed generator subproblem had in generating efficient cuts, compared to the relaxed line subproblem. This difficulty also extends to the subproblem identifying intrinsically infeasible contingencies. Therefore, a feasibility checker for generator subproblems is launched only once, at the end of Step 2 of iteration 1, for all generator subproblems left in the active set.

4.4 Numerical Results

4.4.1 14-bus IEEE Testcase

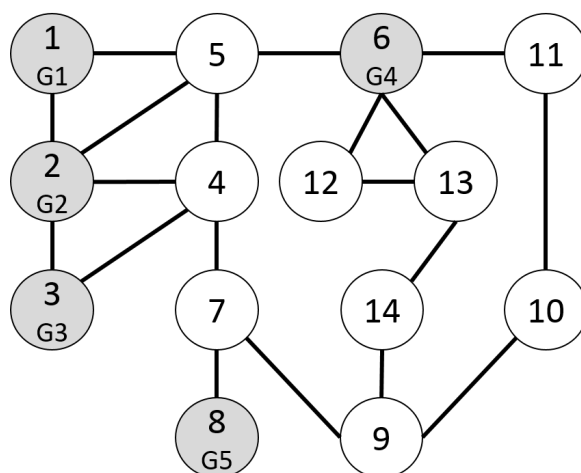


Figure 4.4: Graphical depiction of 14-bus testcase

To better understand the heuristic, we apply it to the 14-bus IEEE test-case. It is common knowledge that the original Matpower datasource lacks meaningful values for generator ramp rates and different line limits. Therefore we make use of the augmented data from the testcases provided through the GAMS OPF Toolbox suite [69]. Specifically, we used the “uw-calc” computations of ‘RampUp’, ‘rateA’ and ‘rateB’ from the augmented datasets to estimate the limits for ramping, STE and LTE in our tests. As for reserve requirements, due to the variances of practice and without access to reliable test data, we opted to forgo including zonal requirements. Instead, we use the final solution to provide insight to the model. As generator reserve bids and/or limits are not provided, we have assumed that any available (committed) generator is able to provide both operational and reserve power generation, up to their regular generation limit. Objective cost functions are piecewise linear approximations provided in the aforementioned toolbox, and reserve cost is estimated at 10% of

regular generation cost. Finally, the contingency list consists of all line and generator failure scenarios. Additionally, we added minimum generator limits as all those values were originally set at 0, as doing so provides us with a more useful analysis.

Iter	Objective	Solution
1	7743.659	$P_G = [189.6, 39.4, 10, 10, 10]$ $P_R = [0, 0, 0, 0, 0]$ - Active subproblems: ALL - Cont G_1 found infeasible in subproblem. - Cont G_2 added to master (largest violated genSP).
2	7843.584	$P_G = [205.4, 23.6, 10, 10, 10]$ $P_R = [15.8, 0, 5.7, 2.1, 0]$ - Cont L_{7-8} found infeasible in master. - Active subproblems: G_3, G_4, G_5 - All active subproblems found feasible.
3	7843.584	- All subproblems rechecked and found feasible. - CONVERGED.

Table 4.6: Iteration analysis of 14-bus testcase

As shown in Table 4.6, the initial basecase solution does not take into account any of the failure scenarios, and allocates zero reserves. This is fixed in iteration 2, where cuts and adding contingency G_2 to the master model result in a more expensive solution with reserve allocation. The algorithm also identifies two scenarios which it cannot satisfy, specifically equipment failure of line L_{7-8} and generator G_1 . A quick glance at Figure 4.4 clearly shows that a failure of line L_{7-8} will isolate generator G_5 at bus 8, which has an operating minimum of 10 MWh. This results in an imbalanced system, and is hence identified as an infeasible contingency. Regarding generator G_1 , it has the highest Rampup rate, and if it fails, the system is unable to compensate appropriately within the 10-minute corrective action timespan. This is identified in iteration 1. Note that the final reserve solution adds up to 23.6MWh, which is the operating point of

generator G_2 which is the next largest generator after generator G_1 . This is the minimum amount of reserve needed for corrective action.

4.4.2 IEEE Polish Cases

As the goal of this study is to make security constrained modeling more accessible for real-time use, our study focused on testing 5 of the IEEE Polish testcases (2383wp, 2736sp, 2737sop, 2746wop, 2746wp) as we find the network size in these cases to be similar to that of ISO New England. All numerical experiments were run on a Dell R810 server with two 256GB memory, 4 2G E7-4850 Xeon chips and 40 cores, using 24 threads, and as previously mentioned, is programmed using JAVA API for GAMS, and uses the CPLEX solver.

Table 4.5 compares the results of running the security constrained model with just “N-1” line contingencies versus “N-1” line and generator contingencies. As discussed in 4.2.3, adding generator failures to the contingency list does increase model complexity and the added computational difficulty is obvious in the results, showing large time increases of at least double in most cases. Interestingly, the results also show that almost all generator contingencies can be made feasible (in contrast to line contingencies). Time wise, we are able to solve 3 of the 5 models in under 5 minutes, with the two largest datasets involving case2746 taking between 7-10 minutes. Further observation reveals that a fair amount of time is spent going through all the still active generator subproblems in iteration 2, because the basecase solution at iteration 1 returns a solution of 0 reserve. This is clearly exhibited in the 14-bus testcase (Table 4.6). We find that by adding a simple system-wide reserve requirement where $|P_R| > 1\%$ demand, we are able to reduce the runtime in most cases.

Table 4.7 summarizes the differences in runtime. The second last column compares the runtime (in seconds) of the algorithm considering a 0% reserve requirement to the last column which shows the runtimes of

Testcase	Demand (D)	Reserve Soln		Runtime (s)	
		Total	% of D	0% of D	1% of D
case2383wp	245.584	6.556	3.92%	322.87	246.27
case2736sp	180.75	3.7	2.10%	259.37	214.04
case2737sop	112.672	3.7	3.28%	252.91	278.44
case2746wop	189.21	3.6	1.95%	597.82	428.15
case2746wp	248.73	4.0	1.61%	471.07	451.88

Table 4.7: Reserve allocation results

same model, but using a simple 1% reserve requirement. This update resulted in significant time savings in 4 of the 5 cases, and the only time it did not in case2737sop, it was because an extra iteration was added to the solution time. Given the size and complexity of the original model, we believe this to be really good time, and with higher computing capabilities, should be able to solve these reliably under 5 minutes. One nice feature that assists in implementing this heuristic in a true real-time setting is that a feasible base-cases solution is generated at the beginning of every iteration. While these non-converged solutions may not cover all necessary N-1 contingencies, it will cover most.

Table 4.7 also shows that depending on the testcase, we are able to satisfy post-contingency action requirements by allocating a just a small percentage of total system demand as system reserves, between 1.6% to slightly under 4%. What these results reveal is the importance of linking security constrained models to reserve allocation. In doing so, reserve allocation is targeted and also done at cost and operational efficiency. In fact, the amount of reserve needed is quite a small percentage of total demand. Note that our 1% system-wide reserve requirement is lower than this bare minimum, testing reveals that using a higher system-wide reserve requirement does not yield any further runtime improvements, and without knowing the minimum needed, a higher requirement factor may result in over-allocation when it is not necessary.

4.5 Contingency Modeling Extensions in the OPF Toolbox

Tying this back into the OPF toolbox, a simpler GAMS-only version of the above model has been added to the OPF Toolbox. To accommodate models with contingency-handling features, the following data definitions have also been added to the Toolbox's .gdx data description process. These definitions are optional, and are not required for regular operation of the standard OPF models.

GDX Name	Description
Set contingency	Set of contingencies, $\omega \in \Omega$
Parameter cont_gen(ω , 'type', $g \in \mathcal{G}$)	Describes if a generator goes on or offline during a contingency ω
Parameter cont_line(ω , 'type', $ijc \in \mathcal{E}$)	Describes if a line goes on or offline during contingency ω

Table 4.8: Data Definitions for Contingency Analysis

This form of data definition allows the user to define multi-equipment failures for a contingency, and is also flexible enough to accommodate cases where a user may wish to specify if equipment comes online as opposed to going offline.

4.6 Conclusion

In conclusion, this chapter introduces a model framework and algorithm to solve security constrained models for co-optimized dispatch models in real-time. The original model which is based on post-contingency action and "N-1" security constraints is intractable for realistic datasets and cannot be easily solved. Our proposed decomposition breaks the model up into smaller pieces while our proposed heuristic uses parallel processing and better contingency handling procedures to increase algorithmic efficiency.

The program uses JAVA API for GAMS and makes use of its enhanced thread handling abilities.

This work contributes to the existing security constrained literature by directly applying a security constrained model to a co-optimized dispatch model, complete with OPF constraints. This process also better aligns the goals of security and economic efficiency by clearly detailing the relationship between ancillary services and grid reliability, and is designed to run on reasonably sized datasets in real-time. In addition, it provides an elegant way of dealing with potential contingency infeasibilities which does not require pre-processing. This means that no assumptions are made about prior topology, and the algorithm easily adapts to changing topologies in a true real-time system. Numerical experiments show that in many cases, it is possible to solve such instances in under 5 minutes, which we believe to be a good indicator of real-world applicability.

The general approach of this model and algorithm does leave open a number of promising enhancements. Based on its results, ISOs may use the results to reconsider existing zonal reserve requirements. While this model does account for an “N-1” security system, using an “N-k” security list may make this direct modeling approach less appealing. As such, one may not want to completely get rid of zonal reserve requirements, but rather use the analysis to reconsider their distribution to improve economic efficiency. Follow-up work could also investigate the explicit modeling of different reserve types, which may lead to an integer based model. In this case however, it would require more careful thought as to its intended purpose, because solving an integer model may not be appropriate in a real-time solution process.

5 STOCHASTIC MODELING WITH RENEWABLE ENERGY

With the many uncertainties currently faced by the Power Systems grid, many models require support for stochastic data. One example of these uncertainties comes from renewable energy. Alternative renewable electric generation, primarily wind and solar, is revolutionizing the power industry. The EPA's Clean Power Plan was one of many stimuli that began driving the transition. The policy, which was first proposed by the Environmental Protection Agency (EPA), required individual states to take responsibility for reducing carbon dioxide emissions by a formula-driven percentage. Overall, the Clean Power Plan aims to reduce carbon pollution from the US power sector by 30 % (based on 2005 levels) by the year 2030 [12]. Recent political developments have seen the proposed legislation defunded and placed under agency review, but this has not dampened the prospective growth of renewable energy in the United States. In response, the United States Climate Alliance was formed, with 13 states and Puerto Rico signaling their intention to uphold the objectives of the 2015 Paris Agreement on climate change by meeting their state targets of the Clean Power Plan.

Despite setbacks, the general consensus for the future of renewable energy is positive as continuing research in the area contribute to improved technology, forecasts and operational procedures. In Nevada, state legislation was recently passed to accelerate the retirement of its largest coal plant [73], citing the increasing cost of aging plants as the driving force behind the decision. Examples like this, when taken into consideration with the rapidly declining cost of wind and solar installations, indicate that renewable energy will continue to see growth. For example, the global levelized cost of electricity (LCOE) for solar photovoltaic (PV) generation has fallen to \$100/MWh in 2016, an almost 70% decline in just 7 years, while the global LCOE of offshore wind has fallen at least 22% in just 1 year

[57]. The downward spiral of costs shows no signs of letting up as many countries, with both developing and developed economies, are investing in both wind and solar, which were considered “alternative energies” not so long ago. Recently, India saw the construction of the Ramanathapuram solar complex, billed as the world’s largest ever PV project [34], while funding for solar-powered mini-grids to bring electricity to remote villages, has become popular all over the world, especially across the African, Asian and South American continents [66]. As a result, research into energy storage has also been pushed to the forefront and is considered as a necessary step to transition into a complete green energy portfolio. The most recent Annual Energy Outlook report of 2021 [72] published by the U.S. Energy Information Administration (EIA) agrees that renewables will have a significant presence in a future not too long away. The report’s predictions indicate that by 2030, just 9 years from now, renewables will collectively surpass natural gas to be the predominant source of generation in the United States. Their analysis and statistics project that the share of renewables in the U.S. electricity generation mix will increase from a proportion of 21% in 2020 to 42% in 2050. Wind and solar generation are expected to be responsible for most of that growth.

Yet as wind and solar production increase, the variations in wind speed and cloud cover lead to significant volatility in power supply. Disruptive power outages are avoided by setting aside generation capacity, so called reserves, to cover this volatility since electricity storage is currently too expensive or too limited to handle them. With the projected growth of renewables, there is an increasing urgency for a more reliable reserves solution. But whether we consider reserves, storage, or a combination, as the ultimate recourse action in the light of this stochasticity, it is clear that datasets for these types of models require scenario data enhancement.

In the following section, we first discuss a typical day-ahead operational model with reserve co-optimization. Then we discuss data sources and

provenance used to craft a day-ahead stochastic model incorporating solar power generation. From there, we utilize our knowledge of that data to come up with a moving window solution methodology that maintains future knowledge of power generation requirements with look-ahead information, and discuss numerical results. Finally, we tie this back in to the OPF Toolbox from Chapter 1 by describing data definitions that can be used to detail stochastic data within the framework of the toolbox.

5.1 Day-Ahead Operational Model with Reserve Co-optimization

A fundamental Power Systems modeling concern is solving the interconnected OPF model over multiple time periods (24 hours for example) with stochastic elements. A simple multi-period stochastic DC model would look something like this.

$$\min_{x,u} c_0(x_0, u_0) \quad (5.1)$$

Power Flow Constraints for base case:

$$\text{s.t.} \quad f_0(x_0, u_0) = 0 \quad (5.2)$$

$$h_0(x_0, u_0) \leq 0 \quad (5.3)$$

Power Flow Constraints for scenarios over time:

$$f_k(x_{kt}, u_{kt}) = 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K} \quad (5.4)$$

$$h_{kt}(x_{kt}, u_{kt}) \leq 0 \quad \forall t \in \mathcal{T}, k \in \mathcal{K} \quad (5.5)$$

Ramping Constraints:

$$|u_{kt} - u_{k,t-1}| \leq \Delta_t \quad t = 1, \dots, T, k \in \mathcal{K} \quad (5.6)$$

Initial state:

$$u_{k,t_0} - u_0 = 0 \quad \forall k \in \mathcal{K} \quad (5.7)$$

Note that the generalized form looks identical to the contingency model in Section 4.2.1. An important difference is that the underlying network may change in case of a contingency failure, for example, if a line failure disconnect parts of the network. In a scenario based model however, this is not necessarily true, and oftentimes is unlikely to be the case.

In the example stochastic model we have tested and included in the OPF toolbox however, the scenario data is built off solar irradiance, which we discuss in Section 5.2. While the scenarios represent physical manifestations of sun exposure onto the earth's surface, the realization of a real life scenario does not manifest along a single scenario line throughout time, which is a common stochastic model assumption. As we move throughout the day, the realization of solar insolation can change between the different scenarios, or fall somewhere in between. If we were to attempt a traditional stochastic model for this situation, we could end up with n^{24} scenarios. Solving a reserves model with a time dimension and such a large number of scenarios is of course intractable.

We next discuss the provenance of our data for considering solar renewables in the above-mentioned framework, and attempt to leverage our understanding of that data into a reasonable sized model and heuristic.

5.2 Data sources for modeling exercise

5.2.1 New England Transmission System

The IEEE 39 bus system is a well-known 39-bus and 10-generator New-England Power System, first published by T. Athay et. al. [7]. We have chosen to focus on this dataset as our initial test case for multiple reasons. First, knowledge of the network's location lends itself well to estimating actual geographic locations for each of the nodes, and the East Coast interest in renewable energy is not insignificant. Connecticut, Massachusetts and Vermont are all members of the U.S. Climate Alliance, while the latter two rank highly clean energy efforts [3]. In addition to the location information, adapting a well-known testcase allows us to leverage previous work from Chapter 1, namely the enhanced IEEE dataset library and accompanying GAMS toolbox for Power Systems analysis. Based on that, we were able to easily adapt the IEEE39 dataset with minor alterations, in order to carry out an exploratory phase with model insights.

The first step was to attach geographical coordinates to each of the buses in the dataset, which facilitates a direct link between the transmission network and solar insolation forecasts. By providing the latitude and longitude to our collaborators at NOAA and SSEC, we were able to obtain solar insolation forecasts for each of those locations. Each of the forecasts can then be used as a potential scenario dictating the solar energy contribution present at each node. Details of these weather observations and forecast streams is provided in the following subheading.

Note that the estimated geographic coordinates, while corresponding to the transmission network, do not necessarily correspond to existing solar installations. The underlying assumptions here are two-fold, and both based on the flexibility of solar array installations. First, we assume (possibly through installing multiple smaller arrays) that any bus within the transmission network is able to contribute a reasonable amount of

solar and that geographical location is not a impediment. The second assumption is that this flexibility allows the solar contribution to either originate ‘at’ the node, or close enough that we can aggregate the new power source with the closest node.

Finally, these tests utilize two of the enhanced data features from the aforementioned IEEE data library, specifically “uwcalc” generator ramp rates and the unit commitment solution.

5.2.2 CAPS Ensemble Surface Insolation Forecast

Surface, or solar, irradiance is the radiant flux (or power) received on the surface per unit area, in our case, measured in watts per square meter (W/m^2). This is a result from the sun exposure onto the earth’s surface. Irradiance over time is called insolation, and insolation is affected, among other things, by the solar zenith angle, cloud optimal depth, single scatter albedo and the asymmetry factor. Such technical details of computing surface irradiance are beyond the scope of this writeup, but can be found outlined in [13] and [55]. Importantly, the papers state that surface irradiance is equal to the clear sky irradiance multiplied by the cloud transmittance, thereby requiring also a cloud cover forecast. It is easy to see that cloud cover reduces surface irradiance, and thus the amount of solar power a plant can produce.

High spatial (4km) and temporal (hourly) ensemble cloud forecasts over the continental US were used to provide 14 hour predictions of cloud cover, cloud transmittance, and surface irradiance for potential New England solar power plants locations during a May 15-24, 2012 case study period. The case study period was conducted during the 2012 NOAA Hazardous Weather Testbed (HWT)¹ and the cloud forecasts were obtained from the Oklahoma University Center for Analysis and Prediction of Storms (CAPS)

¹<http://hwt.nssl.noaa.gov>

². During the case study period, we were able to obtain a 16 ensemble Cloud Optical Thickness (COT) forecast, at a 4-km resolution (4-km Weather Research and Forecasting). This process can provide the same data at any location in the continental US, and is programmatically converted to a form that can be processed directly by state-of-the-art optimization software.

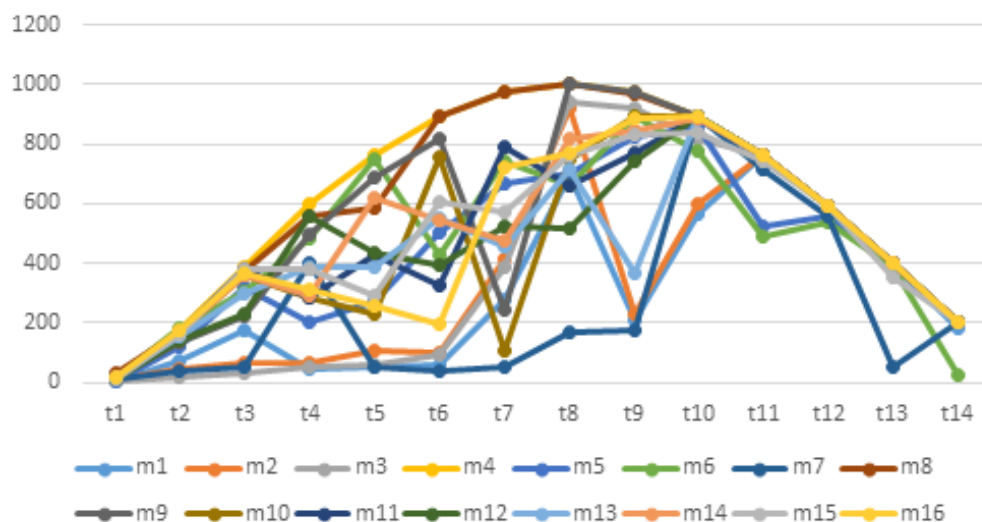


Figure 5.1: Example of Solar Ensemble Forecast

What this translates to for our model, is the ability to obtain 16 forecasted scenarios of solar insolation at any location in the continental US over 8 days during the case study period, specifically May 15-18, and May 21-24 of 2012. Figure 5.1 shows an example of the solar ensemble forecasts. Examples of the ensemble irradiance forecasts ($m_1 - m_{16}$), measured in W/m^2 that is used in the optimization are given below. Notice the distinct differences in the day 3 and day 6 data in terms of the volatility of the ensemble predictions. The conditions on Day 3 were mostly cloud-free and much easier to forecast than the comparatively cloudy conditions present on Day 6.

²<http://www.caps.ou.edu>

Given reasonable technical approximations of a solar installation, we are able to estimate potential solar energy at that any of the locations in our network.

5.2.3 Hourly Demand Data

In order to provide a more realistic dataset to match the 14-hour forecasts, the original 39-bus dataset was enhanced with historical data from ISO New England. We extracted “hourly day-ahead cleared demand” data, during 8 dates: May 15-18,2012 and May 21-24 2012. This data is publicly available at the ISO New England archives ³. The data was then used to build a (different) daily profile for each of the aforementioned dates.

5.3 Solution Heuristic

As previously mentioned, while each CAPS ensemble forecast represents physical manifestations of sun exposure onto the earth’s surface, the realization of a real life scenario does not manifest along a single scenario line throughout time. In fact, we would have to consider n^{14} scenarios based on this alone. To minimize the number of scenarios, we first considered modeling it as a rolling horizon 2-stage stochastic model. This provides us with a ‘base case’ solution at every time period that allows us to react to the spread of scenarios that could happen at any specific hour. If we approach this problem naively and only solve the stochastic OPF for a single timeperiod in each iteration, we are likely to eventually end up with an infeasible solution midway through the day. This is because the solutions at the initial time-steps do not “look-ahead” and the future demand profile needs, and the current timeperiod solution will not take future needs into account. These future model infeasibilities will result from not being able

³<https://www.iso-ne.com/isoexpress/web/reports/load-and-demand>

to satisfy the ramping constraints within the operational bounds of the generators.

Through discussions with our collaborators at NOAA, they have indicated that the expected value of the prediction models (which in many geographical reasons is the simple average of all ensembles) is a reasonable estimate of the final realized product. Therefore, the proposed solution heuristic uses that initial concept but expands it with look-ahead information. How it works specifically is we solve a two stage stochastic model iteratively, beginning with time t_0 , and using that solution to inform the solution in the next time-step. The solutions for each time-step are linked by ramping constraints to the periods before and after.

Since our collaborators at NOAA have indicated that they believe the expected value of the prediction models is a reasonable estimate of the final realized product, we utilized this fact to calculate an expected demand profile (which is static) for the remainder of the future time-steps beyond the current one we are solving for. As such, we are only solving the stochastic model for n scenarios at each timestep. The n scenarios correspond to the scenario information for the current timestep t being solved for, but each scenario also includes (static) future demand information

from $t + 1$ until t_{24} .

Procedure 7: Solution methodology for a 24-hour Day-Ahead Operational Model with Reserve Co-optimization

- 1 Compute the expected value of scenarios for future timeperiods, $t_2 - t_{24}$
 - 2 Solve the stochastic model for the first timeperiod. This will provide operating solutions for each future timeperiod, but we only save the base case solution of the first timeperiod.
 - 3 Using the solution from the first timeperiod, solve the stochastic model for the following timeperiod with additional ramping constraints linking the previous timeperiod's solution (which is fixed in this model) and the current one. Similarly, only the base case solution of this timeperiod is saved and passed on.
 - 4 Continue until all time-periods have been accounted for.
-

5.4 Scenario extension to GDX datafile definition

To accommodate the stochastic model data requirements, we add the following definitions to the OPF Toolbox's .gdx data description process. These definitions are optional, and are not required for regular operation of the standard OPF models.

GDX Name	Description
Set scenario	Set of Scenarios, $\omega \in \Omega$
Parameter sample (ω , bus $i \in \mathcal{N}$, 'datatype', $t \in \mathcal{T}$)	Describes scenario ω using changes in bus data.

Table 5.1: Data Definitions for Stochastic Modeling

In the sample parameter, 'datatype' can be one of the following [P_d , 'pScen'], where P_d is the GAMS file notation for real demand at a bus,

and 'pScen' represents the probability of scenario ω occurring. In this format, the scenario data definition can be used to represent the following instances for stochastic modeling.

1. Uncertainty in the demand profile.
2. A forecast of extra power injection into the network due to renewable energy. The policy on how to treat the renewables (e.g. must take or spilling) must be decided by the modeler.

The above mentioned day-ahead operational model has been added, utilizing the scenario data definitions, as a stochastic modeling example to the OPF Toolbox.

5.5 Observed Numerical Results and Conclusions

Table 5.2 compares the solutions of two models, differentiated by their ramping limit capabilities. Columns 'Gen' and 'Reserves' refer to the total generation power of the base case solution at time t . EVPI refers to the Expected Value of Perfect information, which calculates the value of having perfect information (knowing exactly what scenario will occur) versus solving for the stochastic solution. 'Curtail' refers to the amount of Solar power that the model chooses to curtail (or waste). In the first set of results, we see that amount of reserves plateaus at 2.842. The reason for this is because that is the total limit of all ramping capability across the network. Therefore, the model (with a cost minimization objective) does not allocate any more reserves than that simply because the network would not be able to handle a higher amount of reserves. Note that the amount of EVPI is correlated to the spread of the solar predictions. Figure 5.1 shows an example of this. Case 2, with highly relaxed ramping capabilities (in

fact there is more ramping capability than needed) does not suffer from this plateau. In fact, the total generation solution is smaller overall, as the solution relies on the allocated reserves to manage the solar scenarios. Note that EVPI is minimal in this case, and that solar curtailing only happens in Case 1.

Table 5.2: Solution Comparison with different ramping capabilities

Time	Curtail	EVPI	Gen	Reserves	EVPI ²	Gen ²	Reserves ²
t_1		0.001	61.707	0.593	0.001	61.707	0.593
t_2		7.534	59.27	2.507	0.003	59.27	2.507
t_3		102.229	57.681	2.842	0.007	56.852	3.671
t_4		291.474	56.083	2.842	0.009	53.928	4.997
t_5		416.744	54.887	2.842	0.01	51.843	5.887
t_6	0.387	502.952	53.704	2.842	0.011	50.123	6.422
t_7	0.398	567.943	52.75	2.842	0.011	49.159	6.433
t_8		328.051	51.644	2.842	0.006	49.542	4.943
t_9		422.099	52.478	2.842	0.007	50.314	5.006
t_{10}		32.784	52.296	2.842	0.005	51.859	3.278
t_{11}		83.776	54.069	2.842	0.007	53.079	3.832
t_{12}		53.596	55.722	2.842	0.004	55.365	3.199
t_{13}		8.32	57.595	2.38	0.006	57.595	2.38
t_{14}		0.948	60.311	1.216	0.002	60.311	1.216

The stochastic modeling extension of the OPF toolbox provides the research community with a standardized template with which to share scenario data and modeling efforts, thus easing collaboration efforts. This provided example can be adapted to consider wind generation, compare different types of policies for renewable generation, and also compare the solution quality of a solution heuristic such as the one proposed in this Chapter against a multi-stage model.

6 CONCLUSIONS AND FUTURE WORK

In summary, the work in this thesis document has resulted in an OPF toolbox, which was motivated by wanting to bridge the gap between different models, software, solvers, data formats, industry and academics. As a collaborative and unified tool, it provides a structured method of analyzing different models, software, solvers in power systems research and encourages effective collaboration between researchers and industry. Future extensions could look at improving the ACOPF performance in these large models and explore the feasibility of using Procedure 4 in a realistic setting (i.e. solving the complete 24 hour UCAC in “real-time” for day-ahead planning). This may involve using alternative starting point methods, different linearized approximations of the AC model and more carefully exploring the relationship between unit commitment and ease of AC solution finding. We could also consider pursuing different starting points for the STF OPF model as the default starting point does not perform well in large scale settings.

The realistic and large-scale datasets provided as part of the data archive of the OPF Toolbox allows researchers to test models on realistic datasets, helping take research from pedagogical to practical. In the author’s opinion, an important future extension would be to bridge gap between OPF and Power World solutions, especially in the case of the EPIGRIDS datasets. This would further enhance collaboration between different academic researchers as well as industry.

The Hierarchical Framework from Chapter 3 showcases an example where we combine multi-agent/period models into tractable, flexible framework. A possible future extension would be to apply a similar framework to different situations. This thesis also provides an exemplar of incorporating stochastic data into operational model as well as an effective and practical (as opposed to pedagogical) algorithm to solve the “N-1”

security model quickly.

Finally, the OPF toolbox has been enhanced with stochastic and contingency analysis extensions, and compelling future work would be to define a standardized format for inputting scenario information (for the stochastic modeling extension), which can then be easily translated with the toolbox into the `.gdx` format. As an example, the input format file could be in the form of a `.csv` file.

A.1 Model descriptions

A.1.1 Power Flow Models

1. Optimal DC power flow models
 - a) `dcopf.gms`
Regular DC power flow model.
 - b) `dcopf_shift.gms`
DC power flow model with shift matrices.
2. Optimal AC power flow models
 - a) `polar_acopf.gms`
Polar-power voltage formulation.
 - b) `rect_acopf.gms`
Rectangular power-voltage formulation.
 - c) `iv_acopf.gms`
Rectangular current-voltage formulation.
3. Decoupled AC power flow models
 - a) `polar_decoupled.gms`
Polar-power voltage formulation.
 - b) `ybus_polar_decoupled.gms`
Ybus Polar-power voltage formulation.
4. Ybus AC power flow models
 - a) `ybus_polar_acopf.gms`
Polar-power voltage formulation.

b) `ybus_rect_acopf.gms`

Rectangular power-voltage formulation.

c) `ybus_iv_acopf.gms`

Rectangular current-voltage formulation.

5. `uc_dc.gms`

Optimal DC power flow model with unit commitment.

6. `uc_ac.gms`

Optimal AC power flow model with unit commitment.

The GAMS models take as input a GDX file written by the `to_gdx.gms` utility. Currently, this utility supports reading from a Matpower-formatted ‘.m’ structure file using an awk script and from a PSSE-formatted raw text file using an awk script for comma delimited files or a C++ application for comma or space delimited files. Additional information on data utilities can be found in the Data Utilities section of this project.

A.2 Power Flow Model usage

The basic model can be run from the command line with a single option:

```
gams model.gms --case=/path/case.gdx
```

Note that running `dcopf_shift.gms` requires an additional dataset, namely the shift matrices. Please refer to Section A.3 for more details.

Sections A.2.1 and A.2.2 provide general options based on whether they are single period power flow or (up to) multi-period unit commitment models. Additional options specific to the DC and AC formulation are listed in Section A.2.3 and A.2.4 respectively.

A.2.1 OPF model options

These options apply to models that do not incorporate unit commitment, such as `dcopf.gms`, `polar_acopf.gms`, `iv_acopf.gms` etc. These models solve the power flow model for a single time period and rely on the dataset to provide the on/off status of generators. Do not include square braces [] in the options.

1. `--timeperiod=[#]`
Select which timeperiod to solve in a specific dataset. (*Default* = 1)
2. `--obj=[objective]`
 - `pwl`: Piecewise linear objective
 - `quad`: Quadratic objective (*Default*)
 - `linear`: Simplified linear objective that uses only linear components of the quadratic objective function. Useful for finding feasibility.
 - `0`: 0 objective function. Useful for solving feasibility models.
3. `--linelimits=[datatype]`
 - `given`: Uses original source data on line limits (*Default*)
 - `uwcalc`: Inferred line limits calculated according to the System Characteristic Inference section of this study
 - `inf`: Removes all line limits, sets them to infinity
4. `--genPmin=[datatype]`
 - `0`: Removes lower bound on generator operating limits, i.e. set to 0
 - `given`: Uses original source data for minimum operating limit of generators (*Default*)

- `uwcalc`: Inferred line limits calculated according to the System Characteristic Inference section of this study

5. `--allon=[option]`

- `gens`: All generators are on/active and abide by operating bounds
- `lines`: All valid lines are on/active in the network
- `both`: All valid lines and generators are active
- *(Default)* This option is ignored

6. `--savesol=[#]`

- 0: Solution is not saved *(Default)*
- 1: Solution is saved into `gdx` file, in similar format as input file

** For more a thorough save of the environment, the GAMS command line option `gdx=out` will save all data, variable and equation information at the end of the run into `outfile.gdx`. See GAMS manual for further information.*

7. `--verbose=[#]`

- 0: Regular listing file (`.lst`) output *(Default)*
- 1: Listing file output is suppressed before model solve

A.2.2 Unit commitment model options

These options apply to models incorporating unit commitment, `uc_dc.gms` and `uc_ac.gms`. They are able to solve the power flow model for single or multiple time periods, and include ramping constraints which are not considered in single period models. Note the difference in the default objective function used here, compared to the regular powerflow models in Section A.2.1. Do not include square braces [] in the options.

1. `--times=[timeframe]`: Select which timeframe to solve the model.
 - `t`: Solve a single time period, `t`.
 - `"t1*tn"`: Solve multiple time periods, from `t1` to `tn`
 - (*Default*) Time frame is read in from the input datafile
2. `--obj=[objective]`
 - `pwl`: Piecewise linear objective (*Default*)
 - `quad`: Quadratic objective
3. `--demandbids=[#]`
 - `0`: Demand is fixed (*Default*)
 - `1`: Incremental elastic demand bidding is allowed, if data available
4. `--linelimits=[datatype]`
 - `given`: Uses original source data on line limits (*Default*)
 - `uwcalc`: Inferred line limits calculated according to the System Characteristic Inference section of this study
 - `inf`: Removes all line limits, sets them to infinity
5. `--genPmin=[datatype]`
 - `0`: Removes lower bound on generator operating limits, i.e. set to 0
 - `given`: Uses original source data for minimum operating limit of generators (*Default*)
 - `uwcalc`: Inferred line limits calculated according to the System Characteristic Inference section of this study

6. `--ramprates=[datatype]`

- `given`: Uses original source data for ramp up and ramp down rates (*Default*)
- `uwcalc`: Inferred ramping limits calculated according to the System Characteristic Inference section of this study

7. `--allon=[option]`

- `lines`: All valid lines are on/active in the network
- (*Default*) This option is ignored

8. `--savesol=[#]`

- 0: Solution is not saved (*Default*)
- 1: Solution is saved into.gdx file, in similar format as input file

9. `--verbose=[#]`

- 0: Regular listing file (.lst) output (*Default*)
- 1: Listing file output is suppressed before model solve

10. `--relax=[#]`

- 0: Model solved regularly (*Default*)
- 1: Relaxed integer models, e.g. solved with `rmip/rmiqcp`

A.2.3 DC specific model options

1. `--lineloss=[#]`: Approximates line loss by increasing demand.

- 0: No changes to provided demand profile. (*Default*)
- 1: Increase active demand values by 5.5%.

A.2.4 AC specific model options

1. `--qlim=[#]`: Whether to enforce reactive power limits as D-curve circle constraints

- 0: Ignore D-curves, instead just use rectangle constraints. (*Default*)
- 1: Include D-curve constraints.

*** Does not apply to polar_decoupled*

2. `--slim=[#]`: Whether to enforce apparent power limits instead of simple bound active line power constraints

- 0: Use simple bounds on active line power. (*Default*)
- 1: Use apparent power limits.

*** Does not apply to iv_acopf, polar_decoupled or ybus models*

3. `--ic=[#]`: Choose method for generating initial conditions, i.e. NLP starting point

- 0: **[Midpoint]** Begin with all voltage magnitudes, voltage angles, real power, and reactive power variables at the midpoint of their bounds, calculating line flow variables from these values. (*Default*)
- 1: **[Random]** All variables initialized using random draws between variable bounds.
- 2: **[Flat]** Flat start, where all initial guesses for voltage magnitude and voltage angle are set to 1.0 and zero, respectively, and power flow initial guesses are set to zero.
- 3: **[Random/AC]** Voltage magnitude and voltage angle variables are initialized using random draws between variable bounds.

Real and reactive power variables are initialized using AC transmission line model (applied to each line separately) and the initialized voltage magnitude and voltage angle values.

- 4: **[DC/AC]** Real power and voltage angle values are initialized using a DCOPF model. Voltage magnitudes are initialized at 1.0. Reactive power is initialized using relevant equations from the AC transmission line model (applied to each line separately) and the initialized voltage magnitude and voltage angle values.
- 5: **[DC-/AC]** Voltage angle values are initialized using a DCOPF model (real power values are obtained in the DCOPF, but discarded). Voltage magnitudes are initialized at 1.0. Real and reactive power variables are initialized using the AC transmission line model (applied to each line separately) and the initialized voltage magnitude and voltage angle values.
- 6: **[Decoupled]** Voltage magnitude, voltage angle, real power and reactive power variables are initialized using a decoupled ACOPF model.
- 7: **[DCLoss]** Real power and voltage angle values are initialized using a DCOPF model with line loss approximation (`--lineLoss=1.055`). Voltage magnitudes are initialized at 1.0. Reactive power is initialized using relevant equations from the AC transmission line model (applied to each line separately) and the initialized voltage magnitude and voltage angle values.
- 8: **[Matpower]** Use voltage magnitude, voltage angle, real power, and reactive power values given in Matpower solutions (if available).
- 9: **[inputFile]** Use voltage magnitude, voltage angle, real power, and reactive power values given in the GDX file.

*** Does not apply to polar_decoupled*

A.3 Additional Notes

1. Shift matrices

In order to use `dcopf_shift.gms`, a `gdx` file containing the shift matrix needs to be at the same location as the input case file. In other words, `gams dcopf_shift.gms --case=/path/case.gdx` will search for `/path/case_Shift_Matrix.gdx`.

If the shift matrix does not already exist, use `calc_S_matrix.gms` in the Data utilities section to generate the necessary file.

*** Currently, the LMP is not computed for `dcopf_shift.gms`.*

B COLLECTION OF DATA UTILITIES

B.1 Model descriptions

B.1.1 Format conversion tools

1. `to_gdx.gms`
Converts matpower and psse files to gdx format.
2. `to_matpower.gms`
Converts a gdx file (of above format) to a Matpower-formatted `.m` file.
3. `to_psse.gms`
Converts a gdx file (of above format) to a PSSE-formatted `.raw` file.
4. `gdx2xls.gms`
Converts a gdx file (of above format) to an `.xls` or `.xlsx` output file, depending on the version of Excel installed. Versions prior to 2007 use `.xls` while later versions use `.xlsx`. To write `.xls` files when using later versions of Excel (2007 and after), the `--out` option must be specified with an `.xls` file extension. The `.xls` extension is not recommended due to sheet size limitations which were increased beginning in Excel 2007.
All Excel files start with a “Table_of_Contents” sheet that lists all set and parameter entries of the `.gdx` inputfile in the first column. Each entry is a clickable link that guides the user to the appropriate sheet with the stored data values. All other data sheets are named after GAMS sets and parameters from the input file, and have a clickable link, “TOC”, in the A1 cell that returns the user to the main “Table_of_Contents” sheet.

B.1.2 Other tools

1. `calc_S_matrix.gms`
Outputs a file containing the shift matrix, for use in `dcopf_shift.gms`
2. `make_demand_curves`
Adds demandbid information to a datafile

B.2 Model Options

B.2.1 Format conversion model options

These options apply to models that deal with format conversion as listed in B.1.1. The basic model can be run from the command line with a single option, `gams model.gms --in=/path/inputfile`, as explained below. Do not include square braces [] in the options.

1. `--in=[/path/to/inputfile]` (Required)
Select the input file to be converted. With the exception of `to_gdx.gms`, all input files refer to GAMS input files with `.gdx` extensions. For `to_gdx.gms`, the extension can either be a PSSE-formatted `.raw` or a Matpower-formatted `.m` file.
2. `--out=[/path/to/outfile]`
Specify a destination for the final GDX file output. If not given, this is assumed to be the same as the input file but with the appropriate extension, i.e. `inputfile(.m,.raw,.xlsx,.gdx)`.

B.2.2 Model to_gdx.gms options

These options apply specifically to the data utility `to_gdx.gms`. Currently, this utility supports reading from a Matpower-formatted `'m'` structure file using an awk script and from a PSSE-formatted raw text file using an

awk script for comma delimited files or a C++ application for comma or space delimited files. Do not include square braces [] in the options.

Note that if line symmetry is encountered in the input datafile (i.e. two separate lines *ijc* and *jic* exist), this is replaced with lines *ijc* and *ijc₂* in the .gdx file. This avoids issues with defining line losses in the ACOPF models. More information on how lines are treated in the GAMS models can be found in the Model Formulation document.

Note: Space delimited PSSE files cannot to our knowledge be parsed by awk, as they allow quoted strings with spaces for some of the fields, which makes parsing the rows using regular expressions potentially impossible.

1. `--type=[datatype]`: Specify the format of the input file.
 - `psse`: Input is a PSSE-formatted file, `inputfile.raw`.
 - `matpower`: Input is a Matpower-formatted file, `inputfile.m` (Default).
2. `--mode=[conversion_mode]`: Specify the reader to be used
 - `awk`: Awk script (Default)
Note: PSSE awk script does not currently support some features of PSSE raw data.
 - `cpp`: Compiled C++ executable (Only supported for `--type=psse`).
Note: C++ executable requires GAMS IDE and appropriate environment variables set to compile.
3. `--Sbus=[#]`
 - 0: Do not calculate shift factor matrix (Default)
 - 1: Calculate the matrix of linear shift factors for DCOPF using GAMS model `calc_S_matrix.gms`.
4. `--monitorall=[#]`: (valid only if `--Sbus=1`)

- 0: Only calculate shift factor matrix rows for highest voltage lines and transformers
- 1: Calculate shift factor matrix rows for all lines and transformers (*Default*)

B.2.3 Model `calc_S_matrix.gms` options

The data utility `calc_S_matrix.gms` is called with a single option:

```
gams model.gms --case=/path/case.gdx
```

The resulting output file is `/path/case_Shift_Matrix.gdx`. Note that this model may take a while as it currently produces a dense matrix which requires $|N|$ linear program solves, where $|N|$ is the number of nodes in the network.

B.2.4 Model `make_demand_curves.gms` options

The data utility `make_demand_curves.gms` is a utility for adding auxiliary demandbid information to `.gdx` datafiles that does not have this information. Note that this may not work accurately if a datafile already has demandbid information. It requires two files, the original input file (without demandbid information) and a solution file. The solution file should be generated using the `--savesol` option from one of the opf models in the model archive, and more specifically, it should provide LMP values as this is used to compute appropriate values for demand bidding. The `make_demand_curves.gms` model has the following options. Do not include square braces `[ ]` in the options.

1. `--in=[/path/to/inputfile]` (Required)
Select the (gdx) input file.
2. `--lmp=[/path/to/solfile]` (Required)
Specify a solution file with LMP values.

3. --out=[/path/to/outfile]

Specify an output file name. If not required, this function will overwrite the inputfile. Note that the only difference between the two is the addition of demandbid information in the output file.

B.3 Prime Mover Abbreviations

Prime mover information is contained in the `geninfo` parameter of the GDX file structure, where `geninfo(bus, 'PrimeMover', 'PM_ABBR')=1` if it applies to a specific generator. Below is a list of available abbreviations that are used in the data files.

Abbreviation	Description
ST	Steam Turbine, including nuclear, geothermal and solar steam (does not include Combined Cycle)
GT	Combustion (Gas) Turbine
IC	Internal Combustion (diesel, piston) Engine
CA	Combined Cycle Steam Part
CC	Combined Cycle Total Unit (plants that are in planning stage, specific generator details cannot be provided)
CS	Combined Cycle Single Shaft (combustion turbine and steam turbine share a single generator)
CT	Combined Cycle Combustion Turbine Part (type of coal must be reported as energy source for integrated coal)
HY	Hydraulic Turbine (includes turbines associated with delivery of water by pipeline)
PS	Hydraulic Turbine - Reversible (pumped storage)
PV	Photovoltaic
WT	Wind Turbine
CE	Compressed Air Energy Storage

FC	Fuel Cell
OT	Other
NA	Unknown at this time (plants that are in planning stage, specific generator details cannot be provided)

Table B.1: Prime Mover Abbreviations

B.4 Fuel Info Abbreviations

Fuel information is contained in the `geninfo` parameter of the GDX file structure, where `geninfo(bus, 'Fuel', 'FUEL_ABBR')=1` if it applies to a specific generator. Below is a list of available abbreviations that are used in the data files.

Abbreviation	Description
BIT	Anthracite Coal, Bituminous Coal
LIG	Lignite Coal
SUB	Subbituminous Coal
WC	Waste/Other Coal (Anthracite Culm, Bituminous Gob, Fine Coal, Lignite Waste, Waste Coal)
SC	Coal-based Synfuel and include briquettes, pellets, or extrusions, which are formed by binding materials and processes that recycle material
DFO	Distillate Fuel Oil (includes all Diesel and No. 1, No. 2, and No. 4 Fuel Oils)
JF	Jet Fuel
KER	Kerosene
RTO	Residual Fuel Oil (includes No. 5 and No. 6 Fuel Oils and Bunker C Fuel Oil)
WO	Oil-Other and Waste Oil (Butane (Liquid), Crude Oil,

	Liquid Byproducts, Oil Waste, Propane (Liquid), Re-Refined Motor Oil, Sludge Oil, Tar Oil)
PC	Petroleum Coke
NG	Natural Gas
BFG	Blast-Furnace Gas
OG	Other Gas (Butane, Coal Processes, Coke-Oven, Refinery, and other processes)
PG	Propane
NUC	Nuclear (Uranium, Plutonium, Thorium)
AB	Agriculture Crop Byproducts/Straw/Energy Crops
BLQ	Black Liquor
GEO	Geothermal
LFG	Landfill Gas
MSW	Municipal Solid Waste
OBS	Other Biomass Solids (Animal Manure and Waste, Solid Byproducts, and other solid biomass not specified)
OBL	Other Biomass Liquids (Ethanol, Fish Oil, Liquid Acetonitrile Waste, Medical Waste, Tall Oil, Waste Alcohol, and other biomass liquids not specified)
OBG	Other Biomass Gases (Digester Gas, Methane, and other biomass gases)
OTH	Other (Batteries, Chemicals, Coke Breeze, Hydrogen, Pitch, Sulfur, Tar Coal, and miscellaneous technologies)
PUR	Purchased Steam
SLW	Sludge Waste
SUN	Solar (Photovoltaic, Thermal)
TDF	Tires
WAT	Water (Conventional, Pumped Storage)
WDS	Wood/Wood Waste Solids (Paper Pellets, Railroad

	Ties, Utility Poles, Wood Chips, and other wood solids)
WDL	Wood Waste Liquids (Red Liquor, Sludge Wood, Spent Sulfite Liquor, and other wood related liquids not specified)
WND	Wind
NA	Not Available

Table B.2: Fuel Abbreviations

C DECOUPLED AC MODEL

Due to the nature of the non-convexity in the power flow equations, large scale ACOPF models are challenging to solve, and research continues to seek algorithms that can provide fast, scalable, and robust solution techniques. Approximations, decompositions and engineering judgment are commonly used to obtain practical solutions. Given that realistic sized models depend on “good” initial conditions to satisfy reasonable computational time and robust convergence, this suggests that if we have a starting point that is fairly close to a feasible ACOPF solution, it may be possible to improve solution time and convergence. The decoupled ACOPF model (`polar_decoupled.gms`) may potentially serve this purpose by providing an initial starting point for the full ACOPF. Note that the model archive also contains `ybus_polar_decoupled.gms` which uses the y-bus admittance matrix as discussed in Section 1.2.2.

C.1 The Power Flow Equations

The power flow equations describe the power system network operating point in steady state and hence are based on complex phasor representation of voltage-current relationships at each bus. Using polar coordinates for complex voltages and rectangular representation for complex power, the active and reactive power flow node balance equations at bus $i \in \mathcal{N}$ are formulated by (C.1 - C.2) respectively.

$$\begin{aligned}
P_i &= P_i - d_i^P \\
&= \sum_{k=1}^n \left(|V_i| |V_k| G_{ik} \cos(\theta_i - \theta_k) \right. \\
&\quad \left. + |V_i| |V_k| B_{ik} \sin(\theta_i - \theta_k) \right) \quad \forall i \in \mathcal{N} \quad (C.1)
\end{aligned}$$

$$\begin{aligned}
Q_i &= Q_i - d_i^Q \\
&= \sum_{k=1}^n \left(|V_i| |V_k| G_{ik} \sin(\theta_i - \theta_k) \right. \\
&\quad \left. - |V_i| |V_k| B_{ik} \cos(\theta_i - \theta_k) \right) \quad \forall i \in \mathcal{N} \quad (C.2)
\end{aligned}$$

The most commonly used technique to solve the non-linear power flow equations above is Newton-Raphson. This method requires derivatives of the power flow equations with respect to voltage magnitudes and angles, and is referred to as the power flow Jacobian matrix. The power flow formulation using the Jacobian matrix using (C.3).

$$\begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \begin{bmatrix} \Delta\theta \\ \Delta V \end{bmatrix} = \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix}$$

Where

$$J_{11} = \frac{\partial P^{bus}}{\partial \theta}, J_{12} = \frac{\partial P^{bus}}{\partial V}, J_{21} = \frac{\partial Q^{bus}}{\partial \theta}, J_{22} = \frac{\partial Q^{bus}}{\partial V} \quad (C.3)$$

In most real world applications, other assumptions can be made on the range of both parameters and solution variables. Two inequalities that commonly appear in many power flows applications are:

$$\frac{x}{r} \gg 1 \quad \text{and also} \quad |\theta_i - \theta_k| < 20^\circ$$

Both of these relationships prove to have small values in J_{12} and J_{21} submatrix, and these small values suggest a weak link between active power and voltage magnitude, as well as reactive power and voltage angle. Therefore, if J_{12} and J_{21} in (C.3) are neglected in the power flow calculation, one obtains the Fast Decoupled Load Flow, for which the linearized relations are:

$$\begin{aligned} J_{11} \times \Delta\theta &= \Delta P \\ J_{22} \times \Delta V &= \Delta Q \end{aligned} \tag{C.4}$$

We employ this idea to decompose the full AC optimal power flow into two subproblems.

Since voltage magnitudes have weak impact on active power injections, voltage magnitudes may be treated as constant values when solving the active power subproblem. Similarly, voltage angles are assumed to weakly impact reactive power injections. This suggests employing voltage angles as constant values for reactive power subproblem. The constant values (voltage magnitudes) for active power subproblem are needed to be specified to start P-Q decoupling. These starting points are a solution of power flow problem having voltage phasors and apparent power injections of each bus. The number of rows for bus, branch and generator vectors are $|\mathcal{N}|$, $|\mathcal{E}|$ and $|\mathcal{G}|$ respectively.

C.2 $P - \theta$ subproblem

For the $P - \theta$ subproblem, the voltage magnitude at each bus is fixed as a constant value, denoted as $\hat{V}$, while reactive power balance equations are neglected in the constraint set. The solution vector for the standard $P - \theta$ OPF problem consists of an $|\mathcal{N}| \times 1$ vector of voltage angles θ_i , $\forall i \in \mathcal{N}$, and a $|\mathcal{G}| \times 1$ vector of generator active injections P_i , $i \in \mathcal{G}$. The decision

variable $x_{P\theta}$ has dimension $(|\mathcal{N}| + |\mathcal{G}|) \times 1$.

$$x_{P\theta} = \begin{bmatrix} \theta \\ P \end{bmatrix}$$

The objective function is to minimize a summation of each generator quadratic or piecewise-linear cost function (C.5) subject to (C.6 - C.10).

$$\min_{Q,V} \sum_{i \in \mathcal{N}} \tilde{c}_i(P_i) \quad (\text{C.5})$$

$$\text{s.t. } g_1(x) = P_i(\theta_i, \hat{V}_i) + d_i^P + (V_i)^2 g_i^s - P_i = 0 \quad \forall i \in \mathcal{N} \quad (\text{C.6})$$

$$h_f(x) = |F_{ijc}(\theta_{ij}, \hat{V}_{ij})| - l_{ijc} \leq 0 \quad \forall ij \in \mathcal{E} \quad (\text{C.7})$$

$$h_t(x) = |F_{jic}(\theta_{ij}, \hat{V}_{ij})| - l_{jic} \leq 0 \quad \forall ij \in \mathcal{E} \quad (\text{C.8})$$

$$\frac{-\pi}{3} \leq \theta(i) - \theta(j) \leq \frac{\pi}{3} \quad \forall (ij) : ij \in \mathcal{E} \quad (\text{C.9})$$

$$\underline{P}_i \leq P_i \leq \bar{P}_i \quad \forall i \in \mathcal{G} \quad (\text{C.10})$$

C.3 $Q - V$ subproblem

For the $Q - V$ subproblem, the voltage angles are fixed as a constant value denoted $\hat{\theta}$. The equality power flow equations for the active power are neglected in the constraint set. The optimization vector x_{QV} for the standard $Q - V$ OPF problem consists of the $|\mathcal{N}| \times 1$ vector of voltage magnitude V_i , $i \in \mathcal{N}$ and the $|\mathcal{G}| \times 1$ vectors of generator reactive injections Q_i , $i \in \mathcal{G}$. The dimension of the decision variable x_{QV} is $(|\mathcal{N}| + |\mathcal{G}|) \times 1$.

$$x_{QV} = \begin{bmatrix} V \\ Q \end{bmatrix}$$

The objective function is to minimize a transmission line loss which is a function of voltage phasors (C.11) subject to (C.12-C.15).

$$\min \sum_{i \in \mathcal{N}} \left(V_i^2 g_i^s + \sum_{(jc): ijc, jic \in \mathcal{E}} F_{ijc} \right) \quad (\text{C.11})$$

$$\text{s.t. } g_2(x) = Q_i(\hat{\theta}_i, V_i) + d_i^Q - (V_i)^2 b_i^s - Q_i = 0 \quad \forall i \in \mathcal{N} \quad (\text{C.12})$$

$$h_f(x) = |F_{ijc}(\hat{\theta}_{ij}, V_{ij})| - l_{ijc} \leq 0 \quad \forall ij c \in \mathcal{E} \quad (\text{C.13})$$

$$h_t(x) = |F_{jic}(\hat{\theta}_{ij}, V_{ij})| - l_{jic} \leq 0 \quad \forall ij c \in \mathcal{E} \quad (\text{C.14})$$

$$\underline{Q}_i \leq Q_i \leq \overline{Q}_i \quad \forall i \in \mathcal{G} \quad , \quad \underline{V}_i \leq V_i \leq \overline{V}_i \quad \forall i \in \mathcal{N} \quad (\text{C.15})$$

Note that inequality constraints for line flow limit are included in both problems to represent the security requirements on the system.

D GENERATOR CAPABILITY CURVE APPROXIMATION

Using machine rating standards and the rectangle constraints commonly specified in power system data sets, we developed approximations of the capability curves for typical generators using the provided rectangle constraints and knowledge of typical intersection points from machine rating standards. The rectangle constraints are denoted as P^{min} and P^{max} for limits on active power output and Q^{min} and Q^{max} for limits on reactive power output, and map to $(\underline{P}, \bar{P}, \underline{Q}, \bar{Q})$ respectively in the nomenclature. These approximations rely on the rectangle constraints representing a single round-rotor synchronous generator. This analysis is not applicable to rectangle constraints that represent aggregations of generators and related equipment (e.g., switched capacitors, load demand at the generator bus, etc.).

D.1 Armature Current Limit

The armature current limit defines the right side of the capability curve. This limit is defined using the rated MVA of the generator. The armature current limit is described by a circle with center at the origin and radius equal to the rated MVA of the generator. Lacking more detailed information, we define

$$R^{max} = \max(P^{max}, Q^{max}) \quad (\text{D.1})$$

as the R^{max} , maximum radius of the circle. The constraint for the armature current limit is

$$P^2 + Q^2 \leq (R^{max})^2 \quad (\text{D.2})$$

where P and Q denote the active and reactive power outputs of the gener-

ator.

A value of R^{max} that is less than P^{max} is interpreted as a maximum mechanical input power limit that is below the maximum electrical power generation limit of the synchronous generator. For such cases, a maximum active power limit of P^{max} must also be imposed on the capability curve.

D.2 Field Current Limit

The field current limit defines the top portion of the capability curve. An equation for this limit in terms of generator parameters is derived in [49]. However, typical power flow data sets do not have sufficient information for direct use of the equations from this derivation. Alternatively, we use knowledge of standard machine ratings to approximate this limit.

A circle has two degrees of freedom in the location of the center $(P_0^{field}, Q_0^{field})$ and one degree of freedom in the radius (r^{field}) . We therefore need three pieces of information to define the field current limit.

We first use the fact that the field current limit is a circle with center $(P_0^{field}, Q_0^{field})$ on the Q -axis; that is, $P_0^{field} = 0$ [49]. We then assume that the maximum reactive power output Q^{max} is achieved at zero active power output, which indicates that the point $(0, Q^{max})$ is on the circle. Finally, we use the fact that standard machine specifications use operation at rated power factor as the intersection between the field current limit and the armature current limit. Assuming a rated power factor of 0.80 lagging, this gives a second point on the circle: $(0.8R^{max}, 0.6R^{max})$. The parameters for the field current limit circle are then given by a solution to

$$(0.8R^{max})^2 + (0.6R^{max} - Q_0^{field})^2 = (r^{field})^2 \quad (D.3a)$$

$$(Q^{max} - Q_0^{field})^2 = (r^{field})^2 \quad (D.3b)$$

Solving (D.3) yields

$$Q_0^{field} = \frac{(Q^{max})^2 - (R^{max})^2}{2(Q^{max} - 0.6R^{max})} \quad (D.4a)$$

$$r^{field} = Q^{max} - Q_0^{field} \quad (D.4b)$$

The resulting field current limit is

$$P^2 + (Q - Q_0^{field})^2 \leq (r^{field})^2 \quad (D.5)$$

The relative values of Q^{max} and P^{max} result in three cases of interest:

1. From (D.1), if $Q^{max} \geq P^{max}$, then $R^{max} = Q^{max}$ and $Q_0^{field} = 0$. For this case, the armature current and field current constraints are identical. Recall that a maximum active power limit (i.e., $P \leq P^{max}$) is specified. A typical curve for this case is shown in Fig. D.1.

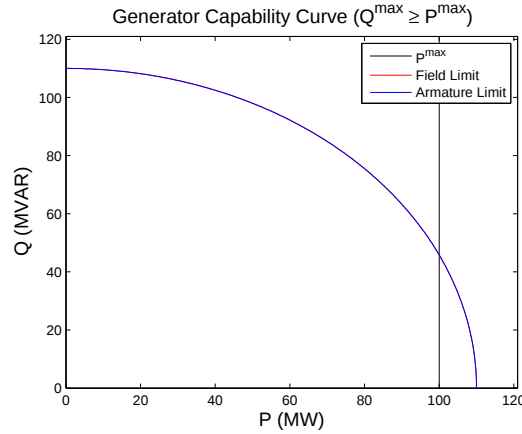


Figure D.1: Generator Capability Curve ($Q^{max} \geq P^{max}$)

2. If $Q^{max} \leq 0.6P^{max}$, then Q_0^{field} is non-negative. Only negative values of Q_0^{field} are physically meaningful (see [49]). Accordingly, if $Q_0^{field} \geq 0$, we will choose to represent the upper portion of the capability

curve as a horizontal line (i.e., $Q \leq Q^{max}$) and disregard the value of Q_0^{field} . A typical curve for this case is shown in Fig. D.2.

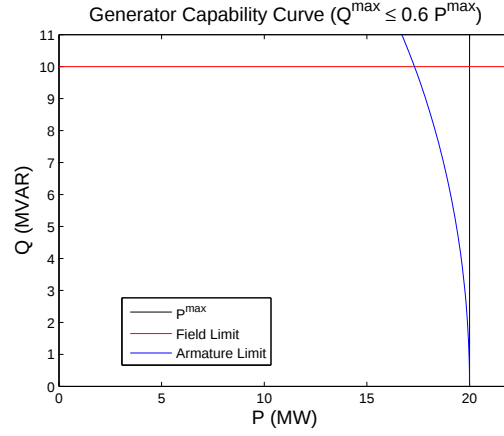


Figure D.2: Generator Capability Curve ($Q^{max} \leq 0.6P^{max}$)

3. If $P^{max} \geq Q^{max} > 0.6P^{max}$, which is expected to be the case for typical generators, the armature and field current limits impose distinct circle constraints. A typical curve is shown in Fig. D.3.

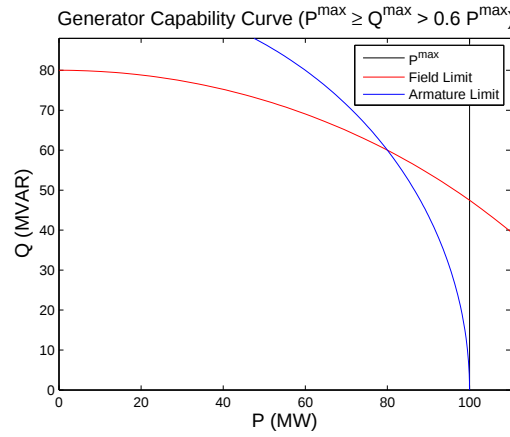


Figure D.3: Generator Capability Curve ($P^{max} \geq Q^{max} > 0.6P^{max}$)

D.3 End Region Heating Limit

The lower limits of the capability curve are due to heating of the end regions of the synchronous generator. There is not a detailed derivation for these limits from the generator parameters as in the case of the field current limits [49].

We use several assumptions to approximate these limits in the same manner as the field current limits. Specifically, we assume:

1. The end region heating limit takes the form a circle with center (P_0^{end}, Q_0^{end}) on the Q -axis (i.e., $P_0^{end} = 0$) and radius r^{end} .
2. The point $(0, Q^{min})$ is on this circle.
3. The intersection of this limit with the armature current limit occurs at 0.95 power factor leading as in Fig. 1.1.

1) , 2) the point $(0, Q^{min})$ is on this circle, and 3) Using these assumptions to follow a similar development as for the field current limits, we obtain

$$Q_0^{end} = \frac{(Q^{min})^2 - (R^{max})^2}{2(Q^{min} + 0.31R^{max})} \quad (D.6a)$$

$$r^{end} = Q_0^{end} - Q^{min} \quad (D.6b)$$

The resulting end region heating limit is

$$P^2 + (Q - Q_0^{end})^2 \leq (r^{end})^2 \quad (D.7)$$

The relative values of Q^{min} and R^{max} result in three cases of interest:

1. If $|Q^{min}| < 0.31R^{max}$, then Q_0^{end} is non-positive. Only positive values of Q_0^{end} are physically meaningful (see [49]). Accordingly, if $Q_0^{end} \leq 0$,

we will choose to represent the upper portion of the capability curve as a horizontal line (i.e., $Q \geq Q^{min}$) and disregard the value of Q_0^{end} . A typical curve for this case is shown in Fig. D.4.

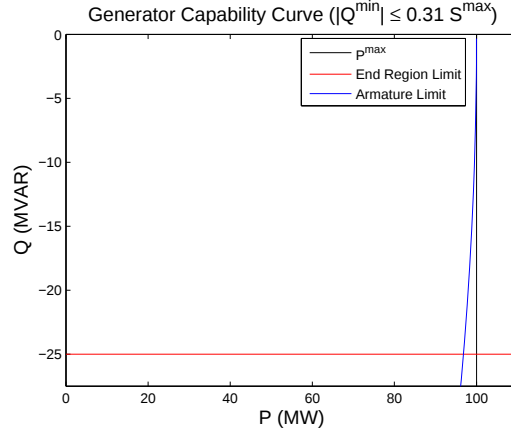


Figure D.4: Generator Capability Curve ($|Q^{min}| \leq 0.31 R^{max}$)

2. If $S^{max} \geq |Q^{min}| > 0.31 R^{max}$, which is expected to be the case for typical generators, the armature and field current limits impose distinct circle constraints. A typical curve is shown in Fig. D.5.

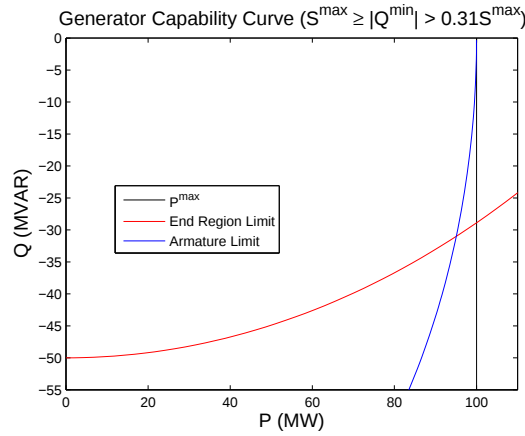


Figure D.5: Generator Capability Curve ($R^{max} > |Q^{min}| > 0.31 R^{max}$)

3. The case $|Q^{min}| > R^{max}$ is atypical for synchronous generators. For this case, the armature current limit is binding before the specified

Q^{min} ; that is, the synchronous generator cannot actually reach Q^{min} due to the armature current limit. If this case does occur, use a horizontal line as the lower limit (i.e., $Q \geq Q^{min}$) and ignore the armature current limit for negative values of Q . In other words, use the original rectangle constraints for the lower half of the generator capability curve. Warn the user that an atypical case has occurred.

D.4 Prime Mover Limits

Limits on the mechanical input power from the prime mover impose constraints on the active power generation. The maximum and minimum active power generation P^{max} and P^{min} are determined directly from the rectangle constraints.

For typical coal, natural gas, oil, and wood fired generators, minimum active power generations that are set to zero (i.e., $P^{min} = 0$) likely indicate missing or incorrect data. A statistical study of the minimum economic operating point (“eco-min”) of these generators is presented in [40]. This study uses the values of P^{min} from a dataset containing generators in PJM. To estimate P^{min} , we use the results of [40] corresponding to the data that is often available in power flow data sets.

Power flow data sets specify generators’ nameplate capacity (i.e., P^{max}) and often provide generators’ prime mover type (i.e., steam turbine, combustion turbine, combined cycle combustion turbine, etc.). If both of these data fields are available, we use the median eco-min data specified in Figures 15, 23, and 25 of [40] to approximate P^{min} . This data is reproduced in Table D.1, which considers steam turbines and combined cycle prime movers, and Table D.2, which considers combustion turbine prime movers operated both independently and as part of a combined cycle plant. Values of P^{min} are specified as a percentage of P^{max} . Note that [40] indicates the possibility of substantial variance around these median data.

P^{max}	Steam Turbine	Combined Cycle
0-200 MW	38%	80%
200-400 MW	39%	46%
400-600 MW	49%	41%
600-800 MW	60%	48%
> 800 MW	64%	42%

Table D.1: Typical P^{min} for Steam Turbines and Combined Cycle Prime Movers (Figures 15 and 23 in [40])

P^{max}	Single CT	CT in Combined Cycle Plant
0-50 MW	76%	80%
50-100 MW	66%	95%
100-150 MW	59%	63%
150-200 MW	81%	63%
200-250 MW	71%	58%
250-300 MW	–	64%

Table D.2: Typical P^{min} for Combustion Turbine Prime Movers (Figure 25 in [40])

If the power flow data does not include the prime mover type, we use the averages among all prime mover types from Figure 10 of [40] (reproduced as Table D.3) to specify P^{min} based on the nameplate capacity data only. Note that [40] indicates the possibility of substantial variance around these median data.

P^{max}	P^{min}
0-200 MW	69%
200-400 MW	42%
400-600 MW	45%
600-800 MW	48%
> 800 MW	69%

Table D.3: Typical P^{min} Without Prime Mover Data (Figure 10 in [40])

D.5 Limitation of the D-curve

A nonlinear generator D-curve constraint for use in AC models were imposed at all generator buses, but some polish cases produce infeasibility since there are generators that contain no sufficient ratio of $\frac{P_{max}}{P_{min}}$ or $\frac{Q_{max}-Q_{min}}{Q_{max}}$. To cross off generators that are not suitable to impose the D-curve constraint, heuristic criterions for the ratio $\frac{P_{max}}{P_{min}}$ and $\frac{Q_{max}-Q_{min}}{Q_{max}}$ are added and the D-curve constraint is not applied for generators that have $\frac{P_{max}}{P_{min}} \leq 1.1$ or $\frac{Q_{max}-Q_{min}}{Q_{max}} \leq 0.1$. When this happens, it shows EPS in the relevant fields to indicate that the D-curve constraint will not be imposed on those generators. Figure D.6 below explains when to impose the D-curve constraint. The figure (a) shows the example of generator that as worst case produces a fixed output. So it is not applicable to impose the D-curve, whereas the figure (b) describes the generator that has resonable ratio of $\frac{P_{max}}{P_{min}}$ and $\frac{Q_{max}-Q_{min}}{Q_{max}}$ so that the D-curve constraint is implemented.

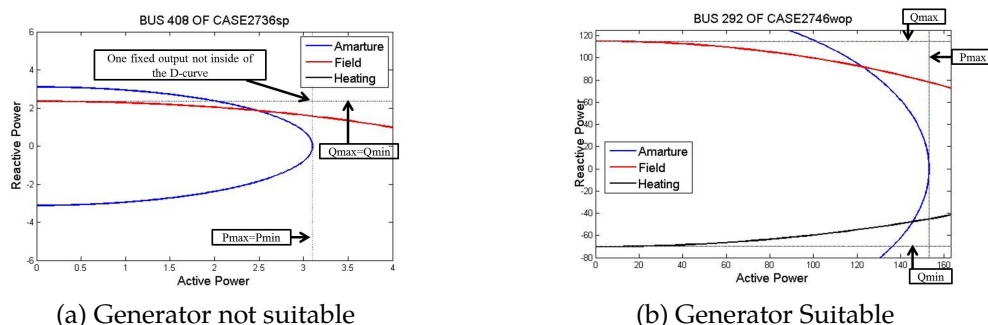


Figure D.6: Example of when to impose the D-curve constraint

D.6 Nonphysical Data

Because these curves are computed from the given values of P^{max} , P^{min} , Q^{max} , and Q^{min} , in the case of atypical or unrealistic generator profiles (for example in the IEEE testcases) the computed constraints may be physically

incorrect or meaningless. When this happens, the data reflects this by showing NA in the relevant fields so the modeler will know not to impose constraints based on improper values.

An additional “uwcalc” set of values are also provided that update the information so the all the profiles include tighter D-curve constraints than the original data generates.

E UNIT COMMITMENT APPROXIMATION

Looking at the original quadratic cost function, $\alpha_j q_j^2 + \beta_j q_j + \gamma_j$, the intercept term, γ_j is split into two components for this unit commitment approximation, as reflected in equation E.1. The first component is the addition of a costly and prohibitive barrier to operating generator j within the bounds $(0, \underline{u}_j)$ and tends to zero as it approaches $\underline{u}_j$, while the second component multiplies the original intercept value, now redefined as $\gamma_{0,j}$, with the cumulative distribution function (cdf) from a Normal distribution, ϕ . Figure E.1 provides an illustration of both components, the approximation function, and its comparison to a integer based unit commitment function.

$$\gamma_j = q_j M_j \exp\left\{\frac{-(q_j - \mu_j)^2}{2\sigma_j^2}\right\} + \gamma_{0,j} \phi(q_j, \mu_j, \sigma_j) \quad (\text{E.1})$$

In the first component, μ_j is viewed as controlling the mean of the peak, M_j as controlling the height of the peak, and σ_j as controlling the width of the curve. The multiplication with q_j ensures a zero cost when $q_j = 0$. The second component in equation E.1 also helps achieve this while retaining the original cost component of γ_j when $q_j \geq \underline{u}_j$. When applying this to varied datasets with different parameter settings and possibly different scalings, in practice, we select these parameters based upon a combination of data and necessary premises for the approximation to be effective. Specifically, it is ideal for the approximation function to approach the true cost function slightly beyond the point $q_j = \underline{u}_j$ as depicted in Figure E.1, which is simultaneously a conservative measure to discourage solutions within the infeasible space of a true unit commitment model, as well as an almost exact approximation when $q_j \geq \underline{u}_j$. Additionally, selection of μ_j and M_j are important in ensuring a sufficiently steep slope at $q_j = 0$, and the height of the peak also ensures the cost of operating some generator j within the $[0, \underline{u}_j]$ bounds is prohibitive. Equations E.2 to E.4

provide the formulations used for the above parameters in our numerical test cases which adheres to the necessary conditions and makes use of the provided data.

$$\mu_j = \underline{u}_j / 2.5; \quad (\text{E.2})$$

$$M_j = 100(\alpha_j \bar{u}_j^2 + \beta_j \bar{u} + \gamma_j) \quad (\text{E.3})$$

$$\sigma^2 = \frac{-9\underline{u}_j^2}{50[\log(k_j) - \log(M_j) - \log(\underline{u}_j)]},$$

$$k_j = (\alpha_j \underline{u}_j^2 + \beta_j \underline{u} + \gamma_j) * 0.025 \quad (\text{E.4})$$

Thus the final cost function with unit commitment and investment effect, $C(y_j, q_j)$, is approximated as follows:

$$C(q_j, y_j) = \alpha_j q_j^2 + \beta_j q_j + \gamma_j,$$

$$\alpha_j = \alpha_{0,j}(y_j) + \frac{\alpha_{0,j}}{4}(\exp(-\lambda_j y_j) - 1),$$

$$\gamma_j = q_j M_j \exp\left\{-\frac{(q_j - \mu_j)^2}{2\sigma_j^2}\right\} + \gamma_{0,j}\phi(q_j, \mu_j, \sigma_j) \quad (\text{E.5})$$

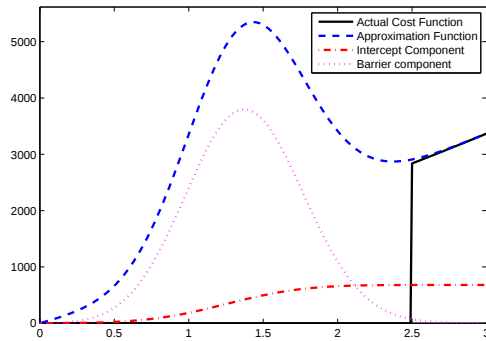


Figure E.1: Approximation of Unit Commitment

REFERENCES

- [1] Epigrids transmission system models. <https://arpa-e.energy.gov/technologies/projects/epigrids-transmission-system-models>. Accessed: 2021-08-11.
- [2] 2013. *Gams java api documentation version 24.0*. GAMS Development Corporation, Washington D.C., USA.
- [3] 2017. Clean energy momentum, ranking state progress.
- [4] Alvey, Trevor, Doug Goodwin, Xingwang Ma, Dan Streiffert, and David Sun. 1998. A security-constrained bid-clearing system for the new zealand wholesale electricity market. *IEEE Transactions on Power Systems* 13(2):340–346.
- [5] Amjady, Nima, Jamshid Aghaei, and Heidar Ali Shayanfar. 2009. Stochastic multiobjective market clearing of joint energy and reserves auctions ensuring power system security. *IEEE Transactions on Power Systems* 24(4):1841–1854.
- [6] Arroyo, José M, and Francisco D Galiana. 2005. Energy and reserve pricing in security and network-constrained electricity markets. *IEEE Transactions on Power Systems* 20(2):634–643.
- [7] Athay, T, R Podmore, and S Virmani. 1979. A practical method for the direct analysis of transient stability. *IEEE Transactions on Power Apparatus and Systems* (2):573–584.
- [8] Averill, Andrew. 2011. Walker announces Charter street plant will convert to natural gas. *The Badger Herald*.
- [9] Bergen, A. R., and V. Vittal. 2000. *Power System Analysis 2nd*. Prentice Hall.

- [10] Bergen Arthur, R, and Vittal Vijay. 1999. *Power systems analysis*. 2nd ed. NJ: Prentice Hall.
- [11] Bertsimas, Dimitris, Eugene Litvinov, Xu Andy Sun, Jinye Zhao, and Tongxin Zheng. 2013. Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems* 28(1):52–63.
- [12] Binz, Ron, Richard Sedano, Denise Furey, and Dan Mullen. 2014. Practicing risk aware electricity regulation: 2014 update. *CERES [Internet]*. April.
- [13] Bird, Richard E, and Roland L Hulstrom. 1981. Simplified clear sky model for direct and diffuse insolation on horizontal surfaces. Tech. Rep., Solar Energy Research Inst., Golden, CO (USA).
- [14] Borenstein, Severin, and James Bushnell. 2000. Electricity restructuring: deregulation or reregulation. *Regulation* 23:46–53.
- [15] Bose, Anjan, and Kevin A Clements. 1987. Real-time modeling of power networks. *Proceedings of the IEEE* 75(12):1607–1622.
- [16] Bussieck, Michael R, Michael C Ferris, and Timo Lohmann. 2012. Guss: Solving collections of data related models within gams. In *Algebraic modeling systems*, 35–56. Springer.
- [17] Cain, Mary B., Richard P. O’Neill, and Anya Castillo. 2012. History of optimal power flow and formulations. Washington DC: Federal Energy Regulatory Commission.
- [18] Cain, Mary B, Richard P O’Neill, and Anya Castillo. 2012. History of optimal power flow and formulations. *Federal Energy Regulatory Commission* 1–36.

- [19] Capitanescu, Florin, and Louis Wehenkel. 2007. Improving the statement of the corrective security-constrained optimal power flow problem. *IEEE Transactions on Power Systems* 22(2):887–889.
- [20] Caramanis, Michael, Elli Ntakou, William W Hogan, Aranya Chakraborty, and Jens Schoene. 2016. Co-optimization of power and reserves in dynamic t&d power markets with nondispatchable renewable generation and distributed energy resources. *Proceedings of the IEEE* 104(4):807–836.
- [21] Castillo, Anya, and Richard P. O’Neill. 2013. Survey of approaches to solving the acopf. Washington DC: Federal Energy Regulatory Commission.
- [22] Corporation, GAMS. 2015. *Examiner manual*. Washington, DC.
- [23] Day, Christopher J., Benjamin F. Hobbs, and Jong-Shi Pang. 2002. Oligopolistic Competition in Power Networks: A Conjectured Supply Function Approach. *IEEE Transactions on Power Systems* 17(3):597–607.
- [24] Ding, Tao, Zhong-Yu Wu, Jiajun Lv, Zhaohong Bie, and Xianjun Zhang. 2016. Robust co-optimization to energy and ancillary service joint dispatch considering wind power uncertainties in real-time electricity markets. *IEEE transactions on sustainable energy* 7(4):1547–57.
- [25] Dirkse, Steven P. 2007. Quality Assurance, Performance Analysis, and The GAMS/COIN-OR Solvers. Presented at INFORMS International.
- [26] Dirkse, Steven P, and Michael C Ferris. 1994. The path solver for complementarity problems. *GAMS-The Solver Manuals*. Washington DC: GAMS Development Corporation.

- [27] Energy Information Administration. 2007. International Energy Outlook: 2007. Tech. Rep., U.S. Department of Energy, Washington, DC.
- [28] Faruqui, Ahmad, and Sanem Sergici. 2009. Household Response to Dynamic Pricing of Electricity—A Survey of the Experimental Evidence. *Whitepaper from The Brattle Group*.
- [29] Federal Energy Regulatory Commission. 2000. Regional Transmission Organizations (Final Rule, Order on Rehearing). Docket No. RM99-2-001; Order No. 2000-A.
- [30] ———. 2005. Reactive power supply and consumption. Staff Report, Docket no. AD05-1-000.
- [31] ———. 2009. The Strategic Plan, FY 2009-2014. Washington DC. Revised Mar. 2013.
- [32] Ferris, M.C., S. P. Dirkse, J.H. Jagla, and A. Meeraus. 2009. An extended mathematical programming framework. *Computers and Chemical Engineering* 33:1973–82.
- [33] Ferris, Michael C., and Todd S. Munson. 1998. Complementarity problems in gams and the path solver. *Journal of Economic Dynamics and Control* 24:2000.
- [34] Finance, Bloomberg New Energy. 2017. Global trends in renewable energy investment 2017. *Frankfurt School of Finance and Management gGmbH, Frankfurt, Germany*.
- [35] Fisher, Emily B, Richard P O'Neill, and Michael C Ferris. 2008. Optimal transmission switching. *IEEE Transactions on Power Systems* 23(3): 1346–1355.
- [36] G.E. Energy. NEOS server. <http://neos-dev-web.neosserver.org/guide/?q=node/116>.

- [37] ———. Mar 2011. GE upgrade project to expand output, increase efficiency at progress energy plant. Raleigh, NC: Business Wire.
- [38] Guo, Xian, and James McCalley. 2015. Risk-based constraint relaxation for security constrained economic dispatch. In *North american power symposium (naps)*, 2015, 1–6. IEEE.
- [39] Hedman, Kory W, Richard P O'Neill, and Shmuel S Oren. 2009. Analyzing valid inequalities of the generation unit commitment problem. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, 1–6. IEEE.
- [40] Higgins, Michael. 2013. Eco-Min Estimation. Tech. Rep., Federal Energy Regulatory Commission, Washington, DC. Unpublished.
- [41] Hobbs, Benjamin F. 2001. Linear Complementarity Models of Nash-Cournot Competition in Bilateral and POOLCO Power Markets. *IEEE Transactions on Power Systems* 16(2):194–201.
- [42] Hobbs, Benjamin F., Carolyn B. Metzler, and Jong-Shi Pang. 2000. Strategic Gaming Analysis for Electric Power Systems: An MPEC Approach. *IEEE Transactions on Power Systems* 15(2):638–45.
- [43] Huyer, W., and A. Neumaier. 1999. Global optimization by multilevel coordinate search. *Journal of Global Optimization* 14:331–55.
- [44] Initiative, MIT Energy, et al. 2011. The future of the electric grid. *Massachusetts Institute of Technology*.
- [45] Jackson, J.Y. 1971. Interpretation and use of generator reactive capability diagrams. *Industry and General Applications, IEEE Transactions on IGA*-7(6):729–732.
- [46] J.Choi, T.Tran, A. A. El-Keib, R. Thomas, H. Oh, and R. Billinton. 2005. A Method for Transmission System Expansion Planning Considering

- Probabilistic Reliability Criteria. *IEEE Transactions on Power Systems* 20(3):1606–15.
- [47] Jin, Shan, and S.M Ryan. 2011. Capacity Expansion in the Integrated Supply Network for an Electricity Market. *IEEE Transactions on Power Systems* 26(4):2275–84.
 - [48] Kezunovic, Mladen. 2006. Monitoring of power system topology in real-time. In *Proceedings of the 39th annual hawaii international conference on system sciences (hicss'06)*, vol. 10, 244b–244b. IEEE.
 - [49] Kundur, P., N.J. Balu, and M.G. Lauby. 1994. *Power System Stability and Control*. New York: McGraw-hill New York.
 - [50] L. O. Chua, C. A. Desoer, and E. S. Kuh. 1987. *Linear and Nonlinear Circuit*. McGraw-Hill.
 - [51] Le Xie, Marija Ilic, Steve Puller, and Shmuel Oren. 2013. Quantifying Benefits of Demand Response and Look-ahead Dispatch in Systems with Variable Resources. Tech. Rep., Power Systems Engineering Research Center (PSERC), Tempe, Arizona.
 - [52] Liu, H., Y. Shen, Z. B. Zabinsky, C. Liu, A. Courts, and S. Joo. 2008. Social Welfare Maximization in Transmission Enhancement Considering Network Congestion. *IEEE Transactions on Power Systems* 23(3): 1105–14.
 - [53] Liu, Yanchao, Michael C Ferris, and Feng Zhao. 2015. Computational study of security constrained economic dispatch with multi-stage rescheduling. *Power Systems, IEEE Transactions on* 30(2):920–929.
 - [54] Ma, Xingwang, Yonghong Chen, and Jie Wan. 2009. Midwest iso co-optimization based real-time dispatch and pricing of energy and ancillary services. In *2009 ieee power & energy society general meeting*, 1–6. IEEE.

- [55] Meador, WE, and WR Weaver. 1980. Two-stream approximations to radiative transfer in planetary atmospheres: A unified description of existing methods and a new improvement. *Journal of the atmospheric sciences* 37(3):630–643.
- [56] Milano, Federico. 2010. *Power System Modelling and Scripting*. New York, NY, USA: Springer-Verlag.
- [57] Motyka, Marlene. 2016. 2017 outlook on renewable energy.
- [58] Park, Byungkwon, Michael C. Ferris, and Christopher L. DeMarco. 2019. Benefits of sparse tableau over nodal admittance formulation for power-flow studies. *IEEE Transactions on Power Systems* 34(6):5023–5032.
- [59] Park, Byungkwon, Jesse Holzer, and Christopher L. DeMarco. 2019. A sparse tableau formulation for node-breaker representations in security-constrained optimal power flow. *IEEE Transactions on Power Systems* 34(1):637–647.
- [60] Park, Byungkwon, Jayanth Netha, Michael C. Ferris, and Christopher L. DeMarco. 2018. Sparse tableau approach for power system analysis and design. In *2018 north american power symposium (naps)*, 1–6.
- [61] PJM Interconnection. 2011. *PJM Manual 14A: Generation and Transmission Interconnection Process*, 17th ed. Effective Date: January 22, 2015.
- [62] ———. 2015. Future deactivations. <http://www.pjm.com/~media/planning/gen-retire/pending-deactivation-requests.ashx>.
- [63] ———. 2015. PJM generator deactivation summary sheet. <http://www.pjm.com/~media/planning/gen-retire/generator-deactivations.ashx>.

- [64] Rios, L.M., and N. V. Sahinidis. 2011. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization* 1–47.
- [65] Sahinidis, N., and M. Tawarmalani. 2012. *Baron solver manual*. GAMS.
- [66] Sawin, Janet L, Kristen M. Seyboth, Freyr Sverrisson, Hannah E. Murdock, Rana Adib, Christine Lins, et al. 2017. Renewables 2017 global status report.
- [67] Shrestha, G.B., and P. A. J. Fonseka. 2004. Congestion-Driven Transmission Expansion in Competitive Power Markets. *IEEE Transactions on Power Systems* 19(3):1658–65.
- [68] Street, Alexandre, Alexandre Moreira, and José M Arroyo. 2014. Energy and reserve scheduling under a joint generation and transmission security criterion: An adjustable robust optimization approach. *IEEE Transactions on Power Systems* 29(1):3–14.
- [69] Tang, Lisa, Christopher DeMaro, Micahel Ferris, Bernard Lesieutre, and Byungkwon Park. A toolbox for exploring ac opf formulations, datasets and solution methods.
- [70] University of Washington. 2011. Power Systems Test Case Archive. <http://www.ee.washington.edu/research/pstca/>. Electronic resource provided by Dept. of Electrical Engineering.
- [71] U.S. Energy Information Administration. 2015. Eia faq: How much electricity is lost in transmission and distribution in the united states. <http://www.eia.gov/tools/faqs/faq.cfm?id=105&t=3>. Accessed on Aug 4, 2015.
- [72] ———. 2021. Annual Energy Outlook 2021 with projections to 2050. https://www.eia.gov/outlooks/aeo/pdf/AEO_Narrative_2021.pdf. Accessed: August 10, 2021.

- [73] Walton, Robert. 2015. Utilities vote to close 2,250 mw navajo plant, largest coal generator in western us.
- [74] Wang, Qi, Chunyu Zhang, Yi Ding, George Xydis, Jianhui Wang, and Jacob Østergaard. 2015. Review of real-time electricity markets for integrating distributed energy resources and demand response. *Applied Energy* 138:695–706.
- [75] Wang, Qin, James D McCalley, Tongxin Zheng, and Eugene Litvinov. 2013. A computational strategy to solve preventive risk-based security-constrained opf. *IEEE Transactions on Power Systems* 28(2):1666–1675.
- [76] Wong, Paul, P Albrecht, R Allan, Roy Billinton, Qian Chen, C Fong, Sandro Haddad, Wenyuan Li, R Mukerji, Diane Patton, et al. 1999. The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *Power Systems, IEEE Transactions on* 14(3):1010–1020.
- [77] Zhao, J.H., Z.Y. Dong, P. Lindsay, and K.P. Wong. 2009. Flexible Transmission Expansion Planning With Uncertainties in an Electricity Market. *IEEE Transactions on Power Systems* 24(1):479–88.
- [78] Zheng, Tongxin, and Eugene Litvinov. 2008. Contingency-based zonal reserve modeling and pricing in a co-optimized energy and reserve market. *IEEE Transactions on Power Systems* 23(2):277–286.
- [79] Zimmerman, Ray D, and Carlos E Murillo-Sánchez. 2011. Matpower 4.1 user’s manual. *Power Systems Engineering Research Center, Cornell University, Ithaca, NY*.