Title
Learning, Control and Optimization for Electricity Distribution Networks

Permalink
https://escholarship.org/uc/item/7c10g8mq

Author
Moffat, Keith

Publication Date
2022

Peer reviewed|Thesis/dissertation

Learning, Control and Optimization for Electricity Distribution Networks

by

Keith Moffat

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Claire Tomlin, Co-chair
Alexandra von Meier, Co-chair
Associate Professor Duncan Callaway
Associate Professor Javad Lavaei

Summer 2022

Learning, Control and Optimization for Electricity Distribution Networks

Abstract

Learning, Control and Optimization for Electricity Distribution Networks

by

Keith Moffat

Doctor of Philosophy in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Claire Tomlin, Co-chair

Alexandra von Meier, Co-chair

Traditionally, utility distribution companies have treated electricity distribution networks as passive loads, assuming that the possibility of network constraint violations cannot be managed by controlling distributed energy resources (DERs) such as electric vehicle charging stations. The emergence of edge computing has made it possible for DER management systems (DERMS) to control the power injected/extracted by DERs in real time. By actively managing DERs when necessary, DERMS both increase interconnection availability and reduce the need for infrastructure upgrades. This dissertation presents several tools that can either make power-injection decisions for a DERMS or that support the system's decision-making. These tools include Unsupervised Impedance and Topology Estimation, Multiple Model Adaptive Power System State Estimation, Voltage Phasor Control, Linearized Output Projected Gradient Descent Feedback Optimization, and Nullspace-Based Power Flow Linearization. The focus of the work is to develop theory which can explain when and how to use these tools on real distribution networks. A large portion of this dissertation develops the theory behind Nullspace-Based Power Flow Linearization, a novel interpretation of power flow linearization that explicitly considers constant-voltage buses, phasor angle-reference ambiguity, and power balance. This dissertation also includes a comment that was submitted in August 2022 on CPUC Proceeding R2207005. With Proceeding R2207005, the California Public Utilities Commission is investigating new rules and price structures for utilities which will change the future of DER management in California.

For my family.

Contents

List of Figures

Acknowledgments

First and foremost I would like to thank Alexandra "Sascha" von Meier, who has advised me throughout grad school. Sascha, your combination of technical expertise and perspective on life have made these past six years as rich of a learning experience as I could ask for. I am grateful to have you as my advisor.

I would also like to thank Claire Tomlin, who has also advised me throughout grad school. Claire your talent with advising is obvious to all those lucky enough to experience it. I have found that your enthusiasm for research is contagious and the direction you give is prescient.

Duncan Callaway has also mentored me throughout grad school. Duncan, I continue to benefit from your thoughtful and honest guidance and I always enjoy our conversations. I look forward to continuing them.

Working on the Gemini-XFC team with Tim Lipman, Bryan Palmintier, Andrew Meintz, Matteo Muratori, Nadia Panossian, and Haitam Laarabi has been a great learning opportunity, crossing the boundary between academic research and real-world implementation. Tim, thank you for this opportunity and for your guidance throughout. To the whole team, thank you for your hard work and for the many things that you have taught me.

Florian Dorfler and Saverio Bolognani's work had a positive impact on my research long before I arrived at IFA. I spent the summer of 2021 working at IFA thanks to the ThinkSwiss research grant. Perhaps my favorite experiences were the whiteboard sessions with Florian and Saverio. These whiteboard sessions impacted the work in this dissertation. My continued collaboration with Jean-Sebastien Brouillon has also been a highlight, and I am grateful for the research conversations and friendships with all of the folks I have met at IFA.

Before getting to Berkeley, Jason Stauth gave me the opportunity to get involved in academic research at Dartmouth and made this path that I have taken a possibility. Without his belief in me I would not be writing this. Charlie Sullivan taught be Kirchhoff's Laws at the beginning of my undergraduate education, power electronics at the end of my undergraduate education, and much more in between. Eric Hansen, Simon Shepherd, Steve Peterson, and a number of other Dartmouth professors taught and encouraged me at various points throughout my undergraduate education as well. Thank you.

Grayson Zulauf connected me with Motiv Power Systems, my first engineering job. I have always enjoyed our conversations about engineering and every other topic. Collaborating with Eric Din, Chris Schaef, and Matt Bossart inspired me to be a better engineer. Gunnar Shaw has provided me with steady friendship and support. Moriah Wilson, a true friend, continues to inspire me every day to be my best self and to be kind.

Javad Lavaei's Control and Optimization of Power System's course jump-started my power systems optimization research. Javad, I have also appreciated your feedback as a member of my dissertation committee.

I have also had the benefit of working with Kameshwar Poolla, and have benefited from our impromptu conversations about power systems research. Seth Sanders helped welcome me to Berkeley, and I have enjoyed our conversations since then that blend the boundaries

between power electronic and power systems. I have also greatly enjoyed my conversations with Robert Pilawa and his students since he arrived at Berkeley.

Andy Packard had a large impact on my grad school experience. I continue to draw inspiration from his love for learning and the care he afforded others, without pretense. The courses I have taken with Murat Arcak, Mark Mueller, Ben Recht, Pieter Abbeel, Sergey Levine, Somayeh Sojoudi, and Laurent El Ghaoui have also impacted my research and will continue to shape my work going forward.

Severin Borenstein's Energy and Environmental Markets course introduced me to electricity markets and the opportunities for research in this space. My conversations with Achintya Madduri about the future of California electricity policy have vitalized my research in distribution network control and energy markets.

It was my great fortune to start my Ph.D. at the same time as Mohini Bariya. She has taught me a lot and is a remarkable engineer and friend. Sascha has gathered a group of thoughtful and bright students in her lab that I am proud to be a part of—Jaimie Swartz, Gabriel Colon-Reyes, Kyle Brady, Miles Rusch, Guillame Goujard, Dahlia Saba, and Laurel Dunn, as well as the postdocs Liz Ratnam, Aminy Ostfeld, and Miriam Aczel. I am also grateful that I have gotten to share my work space on SDH4 with Jaimie, Gabriel, Miles, and Mohini for multiple years. It has always been a pleasure to see them in the morning.

I am also grateful for my companions in the Hybrid Systems Research Group including Roel Dobbe, Patricia Hidalgo-Gonzalez, Somil Bansal, Sylvia Herbert, Kene Akametalu, David Fridovich-Kiel, Andrea Bajcsy, Marrius Wiggert, Tyler Westenbroek, Ellis Ratner, Jaime Fernandez Fisac, Margaret Chapman, Vincenc Rubies Royo, and Donggun Lee. Your collaborative approach to research inspired me when I started grad school, and continues to inspire me today.

Demonstrating Voltage Phasor Control on hardware required expertise across domains and I was lucky to work with Jasper Pakshong, Leo Chu, Jaimie Swartz, T.G. Roberts, Maxime Baudette, Christoph Gehbauer, Guna Bharati, Sudipta Chakraborty, Moustafa Abdelbaky, and Gabe Fierro to successfully demonstrate VPC. That was a fun summer and I am proud of what we accomplished together.

Working with Sunash Sharma, Francesca Giardine, Utkarsha Agwan, Joe Palakapilly, Brittany Wais, and Daniel Tutt has been a pleasure and I am excited to see where their careers will take them. I am also grateful for all of the students who made teaching 137A such a positive experience. I hope that they got as much out of it as I did.

Forrest Laine has helped keep me sane throughout grad school and helped me find interesting research directions. Jonathan Lee and Ian Bolliger kept, and continue to keep, the research stoke high. Jonny also influenced my decision to go to Berkeley in a positive way. Taking classes with Kate Schweidel and Charlotte Vallon has made the classes more fun. The coffees with Jose Daniel Lara and Rodrigo Henriquez-Auba have introduced me to a number of exciting research topics and power systems operation practices. Conversations with Anna Brockway, Phillippe Phanivong, and Lucas Spangher are always interesting and remind me that power systems are far more than just an engineering application.

Chapter 1

Introduction

Note to reader:
Readers who would appreciate an explanation of the work in this dissertation that does not assume familiarity with power systems research should start with the introduction in Appendix A.

Climate change, among other considerations, is motivating the electrification of sectors of the economy that have historically been powered by petroleum and natural gas. In particular, large portions of the transportation and heating sectors will be electrified in the coming decades. This electrification will significantly increase—perhaps doubling—the load on the electric grid [1], presenting novel challenges for grid operators.

Traditionally, utility distribution companies (UDCs) have treated distribution networks as passive loads, assuming that the possibility of network constraint violations cannot be managed by controlling distributed energy resources (DERs) such as electric vehicle (EV) charging stations, electric home and water heaters, and rooftop solar panels. Rather than controlling DERs, UDCs have instead avoided network constraint violations by limiting interconnections onto the distribution network. According to this approach, network constraint violations can only be overcome by upgrading the infrastructure to support the largest coincident load, even though the largest coincident load occurs only once a year. This passive, interconnection capacity-based approach is impeding transportation and heating electrification and will incur unnecessary infrastructure expenses.

The emergence of edge computing has made it possible for DER management systems (DERMS) to control the power injected/extracted by DERs in real time. Using DERMS to avoid network constraint violations offers an alternative to the standard UDC practice of using interconnection capacity limits to avoid potential constraint violations. By actively managing DERs when necessary, DERMS both increase interconnection availability and reduce the need for infrastructure upgrades.

This dissertation presents a few tools that can either make power-injection decisions for a DERMS platform or support the decision-making portion of such a platform. "Power injection decisions" are the decisions made for DER steady state load/generation such as limiting

the power of an EV charging station. Tools that support power-injection decisions include power-flow linearization, network parameter estimation, network topology estimation, and state estimation.

Chapter 4 and Chapter 9 of this dissertation present two different methods for avoiding steady state voltage and current-flow constraint violations — "Voltage Phasor Control" (VPC) and "Feedback Optimization" (FO). VPC and FO are different ways of making power-injection decisions that optimize an objective function and avoid network constraint violations. VPC and FO have different capabilities and goals, and different hardware, communication, and measurement requirements. Either (or both) may therefore be appropriate for a given DERMS application.

VPC, presented in Chapter 4, uses voltage phasor measurements to make quasi-local DER power-injection decisions to avoid network constraint violations. The quasi-local nature of VPC offers a way for DERs to make power-injection decisions in immediate response to disturbances on the network before receiving an updated command from a centralized DERMS calculation. Chapter 4 provides a framework for understanding the impact that VPC has on network constraints and theory that can be used to apply VPC to any DERMS application.

FO is an alternative method for optimizing DER power injections that makes decisions in closed loop with the grid. By only taking one gradient step at each point in time and using the grid to "solve" power flow, FO provides a way to make centralized power-injection decisions quickly and account for mismatches between the model used for the calculations and the true grid model. FO requires the network's power-voltage sensitivity—the sensitivity between the power injections and the network voltages.

Chapter 7 presents "Nullspace-based Power Flow Linearization" (NPFL), a new method for determining the power-voltage sensitivity (and other sensitivities) from the network's admittance matrix. NPFL is a general linearization formulation that takes into account that some buses have constant voltage, the angle-reference ambiguity of AC power systems, and that the power injections must obey power balance. Chapter 7 provides the theory for NPFL and uses a simple two-bus DC network to develop intuition. Section 7.1 motivates NPFL and Section 7.2 focuses on the common circumstance in which a distribution network has a single slack bus/substation. Section 7.3 presents a general form of NPFL that can apply to any network. The sensitivities provided Sections 7.2 and 7.3 can be applied to AC networks of arbitrary size and used to produce other sensitivities such as power injection-loss sensitivity, as done in Section 7.5. Section 7.4 compares NPFL with other prevalent power flow linearizations.

Before presenting NPFL in Chapter 7, Chapter 5 provides the necessary background on power flow linearization, and Chapter 6 presents theory on power flow linearization for networks with a slack/constant-voltage bus. After presenting NPFL, Chapters 8 and 9 provide sample applications of NPFL. Chapter 8 presents "Single bus Injection Power-voltage Sensitivity" (SIPS), the linearization of the local power-voltage sensitivity for a given bus. Chapter 8 derives SIPS from NPFL and highlights nuances that arise when learning SIPS from sensor measurements. SIPS can be used to derive the Voltage Feedback Controllers re-

quired for VPC. Chapter 9 demonstrates how NPFL can be used to improve the performance of FO implemented with the Linearized Output Projected Gradient (LOPG) algorithm.

Many DERMS tools, such as VPC and FO, require distribution network models. For example, NPFL requires the network's bus admittance matrix. However, accurate network models are often not available. Chapter 2 presents an "unsupervised" method for learning a grid model from sensor measurements when there is no information about the grid available. Chapter 2 also discusses the practical challenges that arise with noisy sensor measurements. Chapter 3 presents a method for determining the topology of a network when the line admittances are known but the configuration of the network switches is not. The method introduced in Chapter 3 also provides a state estimate for the network.

The final contribution of this dissertation may be found in Appendix B. In Proceeding R2207005, the California Public Utility Commission (CPUC) is investigating new rates that will incentivize DERs to align their electricity consumption with distribution grid infrastructure [2]. I have contributed a public comment that presents my view on how real-time pricing and flexible interconnections facilitate electrification. The tools presented in this dissertation can be used to support real-time pricing and flexible interconnections at scale.

Chapter 2

Unsupervised Network Estimation

Over the last twenty years phasor measurement units (PMUs) have been widely adopted on transmission networks, but remain scarce in distribution systems. The paucity of distribution PMUs (D-PMUs) coincides with the low quality of distribution network information. Distribution network models are often erroneous or nonexistent. Traditionally, such limited information has been adequate as distribution networks behaved as predictable, passive loads. The proliferation of distributed energy resources (DERs), and natural and human threats to the grid, may make this black-box approach inadmissible.

D-PMUs promise to increase system visibility with measurements of current and voltage magnitudes and phase angles, encoded as complex numbers. However, translating D-PMU measurements into actionable information is nontrivial. In this chapter, we address the specific problem of learning the two basic components of a distribution model: the complex-valued impedance, and the topology (network connectivity), jointly parametrized by network admittance Y, from D-PMU measurements alone. This is a challenging task because every nodal voltage is coupled to *all* nodal currents through Y. Thus estimation must address the network as a whole. Sensor noise, commonly ignored in the power system literature, also affects estimation performance. Noise in PMU measurements comes from multiple sources, including the inevitable error of approximating a time domain signal by its fundamental frequency [3] and the error from imperfect voltage and current transformers connecting the PMUs to distribution lines [4], [5].

Many methods have been proposed for impedance and/or topology identification. A large number are heavily supervised, requiring extensive prior information such as precise network models and knowledge of all possible topologies. [6] and [7] compare real PMU measurements to simulated measurements from multiple possible topologies to identify the current network configuration. Similar approaches using non-PMU measurements also exist [8], [9]. Several approaches combine switch status identification with traditional state estimation [10]–[13]. While these methods have the advantage of requiring fewer measurements, they suffer from high prior information requirements and do not estimate impedance.

Existing unsupervised techniques are either intrusive or make restrictive assumptions. [14] proposes injecting harmonic voltage perturbations to measure the network impedance

from the perspective of the perturbation point. [15] uses measurement correlations to estimate topology using a sparse Markov random field, but requires assumptions on line parameters and load statistics. [16] presents two approaches to topology reconstruction from estimates of the network's Kron reduced admittance matrix. The first recovers topology using a semidefinite program by exploiting the problem's sparse-plus-low-rank structure. The second recovers radial topologies using clique rules, however it is not successfully demonstrated. However neither [16] does not establish fundamental limitations on Kron estimation or consider noise sensitivity. Similarly, [17] acknowledges the low rank structure but does not provide fundamental limits for PMU measurement data sets. [17] introduces an adaptive Lasso optimization approach, as well as a regularization technique for low rank matrix estimation. These estimation techniques can be applied to the quantities we introduce here.

Radial (or tree-shaped) topology reconstruction is of special interest in the distribution context, as most networks are radial. It is also the core challenge of phylogenetics, with genetic distances replacing impedances. The baseline genetic distance-based method for phylogenetic tree reconstruction is the neighbor-joining method [18], [19]. However the neighbor-joining method only deals with tree-endpoints and not internal nodes. Radial tree reconstruction is also of interest in machine learning for creating latent variable models that describe large data sets and predict behavior [20], [21]. Applications include data clustering and Hidden Markov Model estimation. Various techniques have been proposed for creating latent tree models, including hierarchical latent class models [22], [23], the BIN model [24], and expectation maximization methods [25], [26]. However these methods also deal only with tree endpoints.

In [27], Choi et al. introduce the Recursive Grouping (RG) algorithm for learning latent tree graphical models from data. They use information distance, calculated as the logarithmic correlation between random variables, as the distance metric. The RG algorithm is applied to radial electric grids in [28], using effective resistance and reactance as the distance metrics in independent, real-valued tree reconstructions. Impressively, this estimation technique requires only time-stamped measurements of voltage and current magnitude, without phasor angles. However the accuracy of the estimates depend on loads being uncorrelated. The analysis does not consider measurement noise sensitivity, and all the demonstrations use perfect measurements.

We propose a method for joint topology and impedance estimation solely from D-PMU measurements. Our method is *unsupervised*—requiring no apriori information about the network structure or user guidance—making it easier to use, widely applicable, and resilient to human error. It makes no assumptions on load behaviour or network parameters. It addresses the network as a whole and specifically addresses noise robustness by formulating the problem to be better conditioned than existing work. Impedances are estimated via regression but the method is agnostic to the optimization method used for regression.

Our method uses nodal voltage and current injection D-PMU measurements at all active network nodes to estimate a Kron reduced network impedance or admittance model. Next, effective impedances between active nodes are extracted from the Kron reduced model. For radial networks, we introduce Complex Recursive Grouping (CRG) to recover network

topology from the complex-valued effective impedance estimates. Fig. 2.1 shows a schematic representation of the proposed method. For the toy network in Fig. 2.2 the proposed method proceeds as follows: 1) The unsuperised regression estimates the admittance or impedance matrix for the Kron reduced network on the right of Fig. 2.2 using ordinary least squares $\big($e.g. Eq. (2.8)$\big)$, or another optimization formulation. 2) The effective impedance estimates between the active nodes $\hat{\zeta}_{\mathcal{A}}$ are calculated using Eq. (2.1). 3) The full network is reconstructed using the CRG algorithm, as illustrated by Fig. 2.4.

The four primary contributions of this chapter are:

1. Fundamental information limits for (D-PMU) measurements made on electrical networks.

2. Establishment of the Kron reduced admittance matrix as the best achievable outcome of unsupervised learning on electrical network measurements.

3. Proposal of the reduced Laplacian ($subKron$) matrix network representation to improve impedance estimation noise-robustness.

4. Generalization of the Recursive Grouping algorithm to complex-valued distances to recover radial topologies from effective impedances.

The chapter is organized as follows: Section 3.1 provides necessary background information. Section 2.2 establishes the challenges of network estimation. Section 2.3 introduces the subKron network model and section 2.4 explains its physical meaning. Section 2.5 discusses noise-robust unsupervised impedance estimation. Section 2.6 presents the CRG algorithm. Finally, section 2.7 evaluates the proposed methods in simulation.

2.1   Background

Notation and Definitions

$M \in \mathbb{C}^{p \times q}$ is a complex-valued p-by-q matrix with i^{th} row M_i^T and element $M(l, m)$ in row l, column m. M^T and M^H are its transpose and conjugate transpose respectively. $^{-1}$ and † denote the matrix inverse and pseudoinverse rspectively. $\mathcal{Q}$ is a set with cardinality $Q = |\mathcal{Q}|$. I denotes current measurements—*not* the identity matrix. 1 is the 1's vector and e_i is the ith standard basis vector. Operators $Re(\bullet)$ and $Im(\bullet)$ respectively return the real and imaginary parts of their arguments. $\ker(M)$ is the span of vectors $\{v\}$ such that $Mv = 0$. $1j = \sqrt{-1}$. Matrices A and B are *informationally equivalent* if we can exactly compute A from B and vice versa.

A graph $\mathcal{G}$ has nodes $\mathcal{N}$ and edges $\mathcal{E}$. If $\epsilon_{ij} \in \mathcal{E}$, nodes i and j are connected by an edge with weight w_{ij}. The degree of node i in $\mathcal{G}$ is the number of nodes to which it is directly connected. A node with degree $= 1$ is a *leaf*. A connected graph has a path (a sequences of edges) between every pair of nodes. In an acyclic graph, this path is unique. A *tree* is a
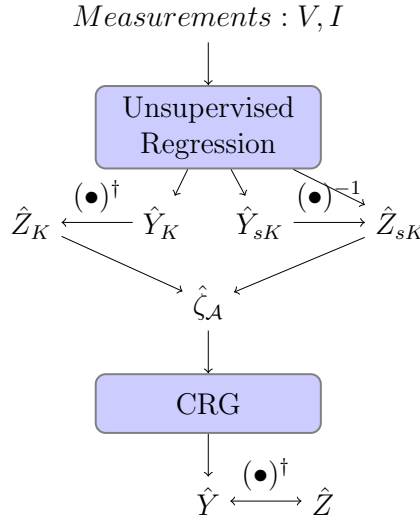
Figure 2.1: Impedance and topology estimation algorithm schematic. Measurements V and I are fed into estimation algorithms to recover various network representations.

connected and acyclic graph. Trees are termed *radial* in power systems. To orient a tree, one node is chosen as the *root*. Node i and j are respectively *parent* and *child* if $e_{ij} \in \mathcal{E}$ and i is closer to the root. Nodes k and k' are *siblings* if they share a parent.

We consider an electrical network with $N = A + P$ nodes and E edges, where N, A, P and E are respectively the cardinality of the set of network nodes $\mathcal{N}$, the set of active nodes, $\mathcal{A}$, the set of passive nodes, $\mathcal{P}$, and the set of edges $\mathcal{E}$. Active nodes are the subset of network nodes where current enters or exits the network. In distribution networks, the substation is an active node that is also a convenient choice for the root. Passive nodes are the subset of network nodes where three or more lines come together and current is rerouted, but does not enter or exit the network. We assume voltage and current phasor measurements are available at *at least* all active nodes. These definitions are visualized in Fig. 2.2.

Complex-valued, synchronized voltage and current phasor measurements from all N nodes at T time points can be collected into matrices $V \in \mathbb{C}^{N \times T}$ and $I \in \mathbb{C}^{N \times T}$. Every row of V and I is a measurement time series at one node. Ohm's law defines a linear relationship between V and I: [Admittance] $I = YV \leftrightarrow V = ZI$ [Impedance]. The two forms are parametrized by either the network *admittance* matrix, $Y \in \mathbb{C}^{N \times N}$, or the network *impedance* matrix, $Z \in \mathbb{C}^{N \times N}$. In graph theory terms, Y is the *weighted Laplacian* of the graph describing the electrical network, with edges corresponding to physical lines and weights corresponding to line admittances. For a network with no shunt admittances, Y is a *loopless* weighted Laplacian [29]. Physically, Z is the inverse of Y, however it cannot be computed as Y^{-1} due to the following well-known lemma [30]:

Lemma 2.1.1. *Loopless Laplacian Y of a connected graph is rank $(N-1)$ with $\mathbf{1} \in \ker(Y)$*

However, we can calculate Z from Y with the matrix psuedoinverse: $Z = Y^\dagger$. We neglect shunt admittances in this work. At active nodes, shunt admittances are irrelevant as both voltage and current measurements are available.

Effective Impedance

In [29] the effective resistance R_{ij} of a resistive network is defined as the voltage drop between nodes i and j when 1 amp of current is injected at i and extracted at j. This is a fundamental quantity for parametrizing network behaviour. We define *effective impedance* ζ_{ij} as the complex-valued voltage phasor difference between nodes i and j of a resistive+reactive network when $(1 + 0j)$ amp of fundamental frequency current is injected at i and extracted at j. ζ_{ij} is computed from Z as:

$$\zeta_{ij} = (e_i - e_j)^T Z (e_i - e_j) \tag{2.1}$$

Effective impedance is symmetric: $\zeta_{ij} = \zeta_{ji}$. All pairwise effective impedances can be collected into $\zeta \in \mathbb{C}^{N \times N}$ such that $\zeta(i,j) = \zeta_{ij}$. $\zeta_{\mathcal{A}} \in \mathbb{C}^{A \times A}$ is the collection of effective impedances between active nodes. We use subscripts rather than parentheses with ζ to emphasize that ζ, unlike Z, isn't structured in a useful manner. The entries of Z can be written in terms of ζ using:

$$Z(i,j) = -\frac{1}{2}\left(\zeta_{ij} - \frac{1}{N}\sum_{k=1}^{N}(\zeta_{ik} + \zeta_{kj}) + \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\zeta_{kl} \right) \tag{2.2}$$

Kron Reduction

While $I = YV$ relates voltages and currents at *all* nodes, the Kron reduction reduces this equation to relate voltages and currents at active nodes *only* [29], [31]. (In general, the Kron reduction can eliminate any subset of nodes $\mathcal{U} \subseteq \mathcal{P}$. We assume $\mathcal{U} = \mathcal{P}$.) The Kron reduction is derived from Ohm's law using $I_i^T = \mathbf{0}$, $\forall i \in \mathcal{P}$. Without loss of generality, we partition the complete V and I matrices into active and passive nodal measurements, plugging in $I_{\mathcal{P}} = 0$:

$$\begin{bmatrix} I_{\mathcal{A}} \\ 0 \end{bmatrix} = \begin{bmatrix} Y_{\mathcal{A}\mathcal{A}} & Y_{\mathcal{P}\mathcal{A}}^T \\ Y_{\mathcal{P}\mathcal{A}} & Y_{\mathcal{P}\mathcal{P}} \end{bmatrix} \begin{bmatrix} V_{\mathcal{A}} \\ V_{\mathcal{P}} \end{bmatrix} \leftrightarrow \begin{bmatrix} V_{\mathcal{A}} \\ V_{\mathcal{P}} \end{bmatrix} = \begin{bmatrix} Z_{\mathcal{A}\mathcal{A}} & Z_{\mathcal{P}\mathcal{A}}^T \\ Z_{\mathcal{P}\mathcal{A}} & Z_{\mathcal{P}\mathcal{P}} \end{bmatrix} \begin{bmatrix} I_{\mathcal{A}} \\ 0 \end{bmatrix}$$

$$I_{\mathcal{A}} = (Y_{\mathcal{A}\mathcal{A}} - Y_{\mathcal{P}\mathcal{A}} Y_{\mathcal{P}\mathcal{P}}^{-1} Y_{\mathcal{P}\mathcal{A}}^T) V_{\mathcal{A}} \leftrightarrow V_{\mathcal{A}} = Z_{\mathcal{A}\mathcal{A}} I_{\mathcal{A}}$$

$$Y_K \triangleq Y_{\mathcal{A}\mathcal{A}} - Y_{\mathcal{P}\mathcal{A}} Y_{\mathcal{P}\mathcal{P}}^{-1} Y_{\mathcal{P}\mathcal{A}}^T \leftrightarrow Z_K \triangleq Z_{\mathcal{A}\mathcal{A}}$$

The Kron reduced admittance matrix, $Y_K \in \mathbb{C}^{A \times A}$, is the Schur complement of $Y \in \mathbb{C}^{N \times N}$, with respect to the passive node set $\mathcal{P}$. For Laplacian matrices, the Schur complement is the cumulative result of Gaussian eliminations of the passive nodes. The Kron reduced
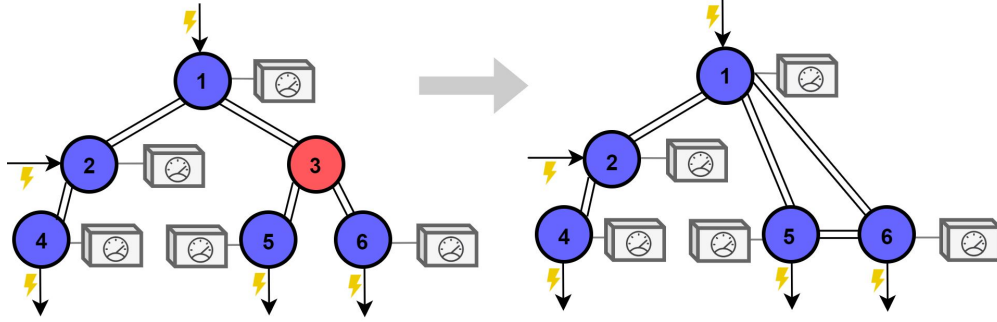
Figure 2.2: Kron reduction of a toy network with 5 active nodes (blue), including the substation, and 1 passive node (red) which does not have any current injection.

impedance matrix, $Z_K \in \mathbb{C}^{A \times A}$, is the Z_{AA} block of $Z \in \mathbb{C}^{N \times N}$. Y_K, like Y, is an admittance matrix and is therefore rank deficient. $Z_K = Y_K^\dagger$ is an impedance matrix.

Physically, Kron reduction eliminates passive nodes, replacing the original network connections with a set of fictitious connections that preserve effective impedances between active nodes: $\forall i, j \in \mathcal{A} : \zeta_{ij} = (e_i - e_j)^T Z_K (e_i - e_j)$. The set of pairwise effective impedances between active nodes are denoted $\zeta_{\mathcal{A}}$. Therefore, the Kron reduction preserves the "input-output" behavior—the relationship between voltages and currents at active nodes—of the complete network.

2.2   Fundamental limitations of Electrical Network Data Sets

Y is not generally recoverable from V and I measurements without knowledge of the network topology, even with full measurement coverage [32]. This is due to fundamental limits, established in Thm. 2.2.1, on the rank of $I \in \mathbb{C}^{M \times T}$ and $V \in \mathbb{C}^{M \times T}$ from network node subset $\mathcal{M}$. Intuitively, a data matrix's rank upper-bounds the information it contains: informationally equivalent matrices must have equal ranks.

Theorem 2.2.1. *(A) For any $V \in \mathbb{C}^{M \times T}$ matrix constructed from time series measurements of voltage phasors at all nodes in the set $\mathcal{M}$ such that $\mathcal{M} \subseteq \mathcal{N}$: $\mathrm{rank}(V) \leq A$*
(B) For any $V \in \mathbb{C}^{M \times T}$ matrix with $T \geq A$ constructed from time series measurements at a subset of nodes $\mathcal{M}$ such that $\mathcal{A} \subseteq \mathcal{M} \subseteq \mathcal{N}$: $\mathrm{rank}(V) = A$

Proof. By Ohm's Law $I = \bar{Y}V$ and by Kirchoff's current law, $\mathrm{rank}(I) \leq A - 1$.
If $\mathcal{M} = \mathcal{N}$, $\bar{Y}$ is the complete network admittance.
If $M < N$, $\bar{Y}$ is the Kron reduction of Y with respect to $\mathcal{M}$.
From lemma 2.1.1, $\mathrm{rank}(\bar{Y}) = M - 1$.

By Sylvester's rank inequality,

$$\text{rank}(I) \geq \text{rank}(V) + \text{rank}(Y) - M \implies \text{rank}(V) \leq A,$$

proving Part A.

Now consider $\mathcal{A} \subseteq \mathcal{M} \subseteq \mathcal{N}$.

By the definition of a graph Laplacian, $\dim(\ker(\bar{Y})) = 1$, with $\ker(\bar{Y}) = \mathbf{1}$.

However, assuming positive voltage measurements, $\mathbf{1}$ is not in the span of $\ker(V)$.

Therefore the kernel dimensions of the matrix product $I = YV$ are additive:

$$\dim(\ker(I)) = \dim(\ker(Y)) + \dim(\ker(V)),$$

and Sylvester's inequality holds with equality:

$$\text{rank}(I) + 1 = \text{rank}(V).$$

Assuming some load variation, the bound on the rank of I holds with equality:

$$\text{rank}(I) = A - 1.$$

Thus, $\text{rank}(V) = A$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Thm. 2.2.1 implies that Y cannot be estimated from V and I without prior knowledge. That is, the information to uniquely specify Y does not exist in V and I alone. Consider when V is not full rank (when the network contains at least one passive node), $\exists y \in \ker(V^T)$ such that: $I = YV = (Y + \mathbf{1}y^T)V$. Similarly, $\exists z \in \ker(I^T)$ such that: $V = ZI = (Z + \mathbf{1}z^T)I$. Thus Y and Z are not uniquely recoverable, as there are multiple possibilities for these matrices that are consistent with the measurements. Therefore, unsupervised Y estimation, with zero prior information, for a network that contains any passive nodes is not generally possible.

Yet, Y contains *two* distinct pieces of information: network connectivity and effective impedances. While Thm. 2.2.1(A) says it is generally impossible to estimate the network connectivity solely from V and I, Thm. 2.2.1(B) implies it is always possible to estimate effective impedances *between active nodes*. That is, the network's input-output behavior is fully contained in active node measurements $V_{\mathcal{A}}$ and $I_{\mathcal{A}}$. Indeed, the matrix relating $V_{\mathcal{A}}$ and $I_{\mathcal{A}}$ *is* the Kron reduced admittance Y_K, and can be *uniquely* recovered from $V_{\mathcal{A}}$ and $I_{\mathcal{A}}$. This is established (as a particular case of a broader result), by Thm. 2.2.2:

Theorem 2.2.2. *Given measurements $V \in \mathbb{C}^{M \times T}$ and $I \in \mathbb{C}^{M \times T}$, at node set $\mathcal{M}$ such that $\mathcal{A} \subseteq \mathcal{M} \subseteq \mathcal{N}$, there is a unique matrix Y relating I to V according to $I = YV$ iff V has full row-rank, and a unique matrix Z relating V to I according to $V = ZI$ iff I has full row-rank. When unique, Y and Z are respectively the network impedance and admittance matrices or an appropriate reduction.*

Proof. Suppose $\ker(V^T) \neq \emptyset$ (not full row rank).

Then, $\exists y \in \mathbb{C}^{1 \times M}$ such that $yV = 0$.

Let $\hat{Y}$ be a solution to $I = YV$.

Then, $I = \hat{Y}V \implies \forall \alpha \in \mathbb{C}: I = (\hat{Y} + \alpha y)V$.

Therefore, there are an infinite number of possible solutions.

Now suppose $\ker(V^T) = \emptyset$ (V does have full row rank).

Let $\hat{Y}_1$ and $\hat{Y}_2$ two possible solutions to $I = YV$.

Then, $I = \hat{Y}_1 V = \hat{Y}_2 V \implies (\hat{Y}_1 - \hat{Y}_2)V = 0$.

This contradicts $\ker(V^T) = \emptyset$, so there can not be two solutions.

A symmetric proof applies for the uniqueness of Z when I has full row rank. □

While Thm. 2.2.1 indicates it isn't generally possible to recover Y from Y_K, in the special case of radial networks it *is* possible.

Radial Networks

It is a property of the Schur complement that Y cannot always be recovered from Y_K. The mapping from a matrix A to its Schur complement, $A_{\mathcal{U}}$, with respect to an arbitrary row and column subset $\mathcal{U}$ isn't generally injective, so recovery of A from $A_{\mathcal{U}}$ isn't guaranteed, even when A is restricted to the set of weighted Laplacians of fully-connected graphs. However when A is further restricted to the set of Laplacians of tree graphs, denoted $\mathcal{T}$, with $\mathcal{U} \subseteq \mathcal{P}$, the following holds:

Theorem 2.2.3. *The Schur complement mapping from $A \in \mathcal{T}$ to $A_{\mathcal{P}}$ is injective. That is, no Schur complement of $A' \in \mathcal{T} \backslash A$ with respect to any set of passive nodes can produce $A_{\mathcal{P}}$.*

Thm. 2.2.3 holds for trees with real or complex edge weights, and is stated here without proof. The proof is based on the CRG, which demonstrates that a radial tree is exactly recoverable from exact distances. For radial electrical networks, effective impedances are a valid distance metric. Together, Thm. 2.2.1-2.2.3 imply Corollary 2.2.3.1:

Corollary 2.2.3.1. *For any radial electric network, Y is fully recoverable from perfect voltage and current PMU measurements at every active node.*

Estimating $\zeta_{\mathcal{A}}$ is a necessary intermediary step in recovering the full network admittance model from active node measurements. In the next section we present a novel network representation which improves effective impedance estimation accuracy from noisy active node PMU measurements.

2.3 subKron Reduction

We introduce the *subKron reduction*, a further reduction of electric networks that extends the Kron reduction. The Kron relationships are: $I_{\mathcal{A}} = Y_K V_{\mathcal{A}} \leftrightarrow V_{\mathcal{A}} = Z_K I_{\mathcal{A}}$. Since relative,

not absolute voltages, drive power flows, we can subtract the substation voltage from all the voltages (Lemma 2.1.1).

$$
\begin{bmatrix} - & I_{\mathcal{A}1}^T & - \\ & \vdots & \\ - & I_{\mathcal{A}A}^T & - \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{1K}^T \\ Y_{K1} & Y_{sK} \end{bmatrix} \begin{bmatrix} - & 0 & - \\ & \vdots & \\ - & (V_{\mathcal{A}A}^T - V_{\mathcal{A}1}^T) & - \end{bmatrix} \tag{2.3}
$$

We define $V_{s\mathcal{A}} \in \mathbb{C}^{(A-1)\times T}$ and $I_{s\mathcal{A}} \in \mathbb{C}^{(A-1)\times T}$ as the data matrices in (2.3) with the first row in each removed. The subKron admittance matrix, $Y_{sK} \in \mathbb{C}^{(A-1)\times(A-1)}$, relates $V_{s\mathcal{A}}$ to $I_{s\mathcal{A}}$. Y_{sK} is equivalent to Y_K with the first row and column removed.

$$
I_{s\mathcal{A}} = Y_{sK} V_{s\mathcal{A}} \tag{2.4}
$$

We can equivalently derive an expression for the subKron impedance matrix which relates $I_{s\mathcal{A}}$ to $V_{s\mathcal{A}}$:

$$
V_{s\mathcal{A}} = Z_{sK} I_{s\mathcal{A}} \tag{2.5}
$$

The elements of Z_{sK} are:

$$
Z_{sK}(i,j) = Z_K(i+1, j+1) - Z_K(1, j+1) - Z_K(1, i+1) + Z_K(1,1) \tag{2.6}
$$

Complete derivations of Eq. (2.4)-(2.6) can be found in Appendix D.1. The subKron representation has lower dimensionality than the Kron representation (by one dimension), but is informationally equivalent:

Lemma 2.3.1. *Y_K is always recoverable from Y_{sK} for a network without shunt admittances.*

Proof. Y_K is a loopless Laplacian.
$Y_K \mathbf{1} = 0$ and $Y_K^T = Y_K$.
Y_{sK} corresponds to Y_K with the first row and column removed.
We can recover the first row and column of Y_K from the elements of Y_{sK} as follows:
$i \in \{2, ..., A\} : Y_K(i,1) = Y_K(1,i) = -\sum_{j=1}^{A-1} Y_{sK}(i,j)$ and $Y_K(1,1) = -\sum_{i=2}^{A} Y_K(i,1)$. $\quad\square$

An important feature of the subKron representation is its rank:

Lemma 2.3.2. *Unlike Y or Y_K, Y_{sK} is always full rank.*

Proof. This follows from Kirchhoff's matrix tree theorem, which states that the total weight of trees in a graph is equal to *any* cofactor of the graph's weighted Laplacian.

$$
T(Y_K) \triangleq \sum_{T \in \mathcal{T}(Y_K)} w(T)
$$

$T(Y_K)$ denotes the total weight of spanning trees in the graph described by Kron impedance matrix Y_K, T is one particular spanning tree of Y_K in the set of all spanning trees $\mathcal{T}(Y_K)$

and $w(T)$ is the weight of tree T (the sum of the weights of the edges in T).
One cofactor of Y_K is $Y_K(1,1)det(Y_{sK})$.
Therefore: $Y_K(1,1)\det(Y_{sK}) = T(Y_K)$.
Since $|\mathcal{T}(Y_K)| \geq 1$, we have $T(Y_K) > 0$.
Therefore, $\det(Y_{sK}) > 0$, proving Y_{sK} is full rank. □

We can convert between the subKron impedance and admittance representations by taking the true matrix inverse: $Z_{sK} = Y_{sK}^{-1}$. In graph theory terms, the subKron matrix Y_{sK} is the reduced Laplacian of Y_K. In the following section, we delve deeper into the physical meaning of the Kron and subKron network matrices, by returning to effective impedances.

2.4   Effective Impedances Between Active Nodes

Effective impedances are useful to concretely establish what is lost and preserved in the Kron and subKron reductions. An electrical network admittance matrix contains two fundamental pieces of information: network connectivity and effective impedances between network node pairs. Kron reduction discards the connectivity information, but preserves the effective impedances between the active nodes, which characterize the network's input-output properties. By Lemma 2.3.1, subKron reduction also preserves inter-active node effective impedances. Therefore, though different in dimensionality and definition, all the network representations introduced in Sections 3.1-2.3 may be used to calculate effective impedances between active nodes.

Lemma 2.4.1 establishes useful relationships between the elements of subKron matrix Z_{sK} and the effective impedances of the original network, ζ.

Lemma 2.4.1. *The $(i,j)^{th}$ element of Z_{sK} is the effective impedance of the <u>shared path</u> between nodes $i, j \in \mathcal{A}$ and the substation.*

Proof. Combining equations (2.2) and (2.6):

$$Z_{sK}(i,j) = \frac{1}{2}\left[\zeta_{1(j+1)} + \zeta_{1(i+1)} - \zeta_{(i+1)(j+1)} - \zeta_{11}\right]$$

□

Corollary 2.4.1.1. *The diagonal elements of Z_{sK} are the effective impedances to the substation.*

Proof. The diagonal elements of Z_{sK} are related to the elements of Z_K by Eq. (2.6): $Z_{sK}(i,i) = Z_K(i+1,i+1) + Z_K(1,1) - 2Z_K(1,i+1) = \zeta_{1(i+1)}$. □

The relationships between the different network models and the effective impedances are summarized in Fig. 2.3. Note that $\zeta_{\mathcal{A}} \subseteq \zeta$.
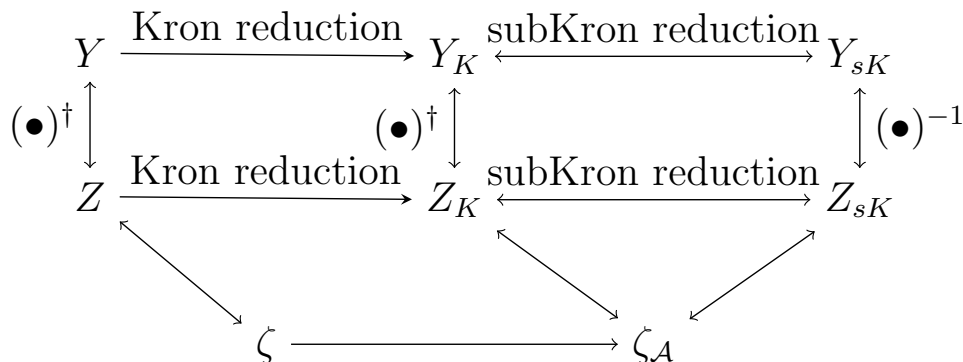
Figure 2.3: Relationships between the different network models. Bidirectional arrows indicate informational equivalence of representations.

2.5   Impedance Estimation

The Kron and subKron representations are informationally equivalent, but estimating them from real-world measurements are not identical challenges. Noise in real world measurements makes network impedance estimation—i.e. "inverting" Ohm's Law to infer Y or Z from V and I—nontrivial. Here, impedance estimation refers to estimation of either the complete (Y or Z), Kron (Y_K or Z_K) or subKron (Y_{sK} or Z_{sK}) network representations. As we establish in the following, the major advantage of the subKron reduction is that it can be estimated with better accuracy than the Kron reduction in the presence of noise.

Ordinary Least Squares

A simple regression method is *ordinary least squares* (OLS). Given data matrices $R \in \mathbb{C}^{n \times T}$, $P \in \mathbb{C}^{n \times T}$ which are known to obey an underlying linear relationship parametrized by matrix $M \in \mathbb{C}^{n \times n}$, we can write: $R \approx MP$ where we use $\approx$ rather than $=$ to reflect that R and P contain noise. OLS estimates M from R and P as $\hat{M} = arg\min_M ||R - MP||_2$. The best linear unbiased estimate of M has a closed form solution: $\hat{M} = (P^H P)^{-1} P^H R$. There are no constraints on the structure of $\hat{M}$ in the above formulation.

Many alternatives to OLS exist including those which regularize the objective, account for noise in both P and R or enforce special structure on M. We focus on OLS estimation because of it's intuitive objective, lack of tuning parameters, and computational efficiency, which makes it suited to online applications of impedance estimation such as fault detection. We also find OLS performs comparably to alternative techniques in practice. OLS can be replaced by any regression technique in our proposed method.

For real-world applicability, we must understand the sensitivity of $\hat{M}$ to noise in the measurements. Depending on the structure of the $P^H P$ matrix, $(P^H P)^{-1}$ can contain very large values that magnify noise in $P^H R$ and produce a poor solution $\hat{M}$. The *condition*

number of $P^H P$ is a metric for understanding the noise magnification of $(P^H P)^{-1}$ and is defined as the ratio of its largest and smallest eigenvalues: $\text{cond}(Q) \triangleq \frac{\sigma_{max}(Q)}{\sigma_{min}(Q)}$. A larger $\text{cond}(P^H P)$ means $(P^H P)^{-1}$ will magnify noise, and $\hat{M}$ will be more noise sensitive. At extremes, $\text{cond}(Q) = 1$ means Q is very well-conditioned, while $\text{cond}(Q) = \infty$ means it is rank deficient and not invertible. $\text{cond}(P^H P)$ is related to the correlation level of the rows of P. If the rows of P are highly correlated, $\text{cond}(P^H P)$ is large, and the estimate is more noise sensitive.

Conditioning of Impedance Models

The OLS network impedance estimation formulation is:

$$\hat{Y} = arg \min_Y ||I - YV||_F = (V^H V)^{-1} V^H I \tag{2.7}$$

$|| \circ ||_F$ indicates the Frobenius norm. Expressions for $\hat{Y}_K$, $\hat{Y}_{sK}$, $\hat{Z}_K$ and $\hat{Z}_{sK}$ are determined by active node measurements alone (Thm. 2.2.2), e.g. $\hat{Y}_K = (V_{\mathcal{A}}^H V_{\mathcal{A}})^{-1} V_{\mathcal{A}}^H I_{\mathcal{A}}$ and $\hat{Y}_{sK} = (V_{s\mathcal{A}}^H V_{s\mathcal{A}})^{-1} V_{s\mathcal{A}}^H I_{s\mathcal{A}}$. The equation for the subKron impedance estimate is:

$$\hat{Z}_{sK} = arg \min_{Z_{sK}} ||V_{s\mathcal{A}} - Z_{sK} I_{s\mathcal{A}}||_F = (I_{s\mathcal{A}}^H I_{s\mathcal{A}})^{-1} I_{s\mathcal{A}}^H V_{s\mathcal{A}} \tag{2.8}$$

Each formulation will have different noise sensitivity since each inverts a different data matrix. The rank constraints of Thm. 2.2.1 immediately indicate an issue with estimating Y or Z from measurements at *every* network node: $V, I \in \mathbb{C}^{N \times T}$. Thm. 2.2.1 establishes I is rank deficient, making computation of $(I^H I)^{-1}$ impossible. When $N \geq P + 1$ (network contains more than one passive node), V will also be rank deficient, making computation of $(V^H V)^{-1}$ impossible. Note Thm. 2.2.1 is for *noiseless* measurement matrices. With noisy V and I, the rank constraints won't hold exactly, but will nonetheless manifest in the matrix conditioning. Therefore, though it may be possible to compute $(I^H I)^{-1}$ and $(V^H V)^{-1}$ for real measurements, these matrices will be highly ill-conditioned, and the resulting impedance estimates will be noise sensitive to the point of impracticality. Similarly, Z_K cannot be estimated directly as $I_{\mathcal{A}}$ is rank deficient.

While Y, Z, and Z_K cannot be estimated uniquely, Thm. 2.2.2 establishes Y_K, Y_{sK} and Z_{sK} can, since $V_{\mathcal{A}}, V_{s\mathcal{A}}$ and $I_{s\mathcal{A}}$ are full rank. This is another perspective from which to see the Kron reduction as the best *possible* estimate of the network impedance given no prior information. Though Y_K, Z_{sK} and Y_{sK} are informationally equivalent, estimating them from real data will have differing accuracies as each entails the inversion of a different data matrix with its own condition number. The following conditioning trend holds for the measurement data: $\text{cond}(I_{s\mathcal{A}}^H I_{s\mathcal{A}}) < \text{cond}(V_{s\mathcal{A}}^H V_{s\mathcal{A}}) < \text{cond}(V_{\mathcal{A}}^H V_{\mathcal{A}})$. The conditioning improvement from $V_{\mathcal{A}}$ to $V_{s\mathcal{A}}$ is intuitive, as taking voltage differences is akin to common mode removal [33]. While nodal currents are predominantly driven by individual loads which often behave independently, nodal voltages are correlated by the network structure itself. Therefore, voltage measurements $V_{s\mathcal{A}}$ tend to be more correlated than currents $I_{s\mathcal{A}}$,
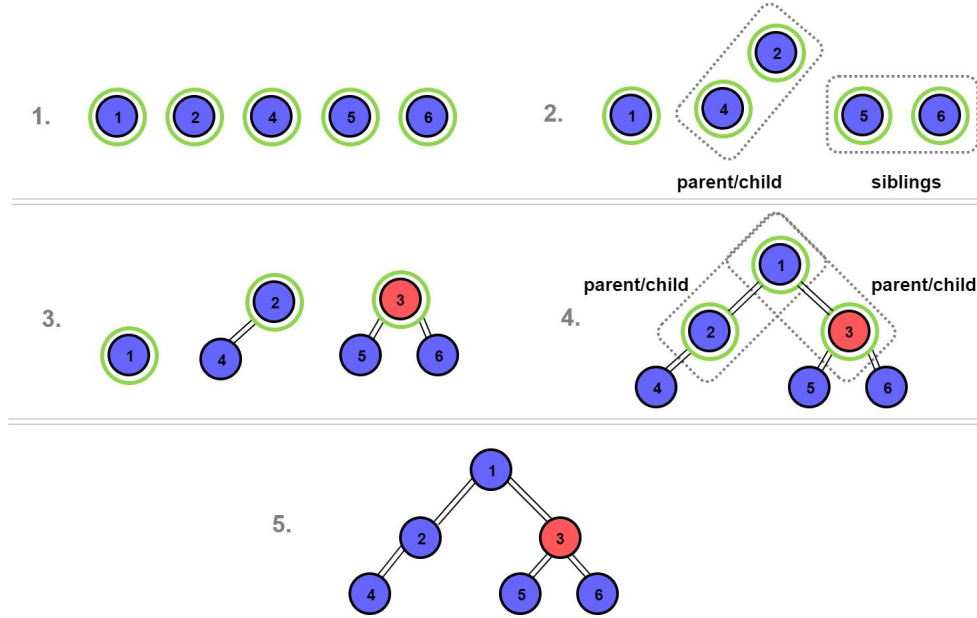
Figure 2.4: CRG on the toy network of Fig. 2.2. The nodes in $\mathcal{O}$ are highlighted in green in each step.

causing $V_{s\mathcal{A}}$ to be more poorly conditioned. Thus, $\hat{Y}_{sK}$ is more noise robust than $\hat{Y}_K$, and $\hat{Z}_{sK}$ estimation is the most noise robust. The improved data conditioning provides better $\zeta_{\mathcal{A}}$ estimates and, in turn, more accurate network reconstruction via the Recursive Grouping algorithm.

2.6   Recursive Grouping Algorithm

The Recursive Grouping (RG) algorithm, presented and described in [27], reconstructs a radial network from real-valued information distances d between the "observed" nodes $\mathcal{O}$, defined here as the set of nodes for which the information distances are known but the parent node is unknown. By comparing the pairwise quantities d_{ij} to the triplet quantities $\Phi_{ijk} \triangleq d_{ik} - d_{jk}, \forall (i,j,k) \in \mathcal{O}$, $\mathcal{O}$ is recursively shrunk until the tree is reconstructed, as demonstrated on the toy network in Fig. 2.4. Using resistance and reactance as the information distances in two separate, real-valued reconstructions, [28] applied RG to radial distribution networks. With noisy data, this separation into real-valued reconstructions is a source of unnecessary reconstruction error.

We introduce *Complex Recursive Grouping* (CRG), a modification of relaxed RG (Appendix D.2), which feeds complex effective impedances directly into the reconstruction. That is, $d_{ij} = \zeta_{ij}$. For complex-valued distances, Lemma D.2.1 becomes:

Lemma 2.6.1. *(i) $\zeta_{ij} = \Phi_{ijk}, \forall k \in \mathcal{O}\backslash(i,j)$ iff i is a leaf node in $\mathcal{O}$, and j is the parent of i.*
(ii) $\Phi_{ijk} = \Phi_{ijk'} \neq \zeta_{ij}, \forall k, k' \in \mathcal{O}\backslash(i,j)$ iff i and j are leaf nodes in $\mathcal{O}$, and siblings.

While CRG uses complex impedances, the inequality tests of CRG must be done on the real number field $\mathbb{R}$, which is ordered (unlike $\mathbb{C}$). Formulating real-valued inequality tests for complex-valued impedances takes care. A natural suggestion is to use the lexicographic ordering (i.e. to compare magnitudes). However, with noise and imperfect $\hat{\zeta}$, this may lead to reconstruction errors. Instead, we use the ϵ-neighborhood of $\hat{\Phi}_{ijk}$ in the complex plane for some parameter ϵ. $z \in \mathbb{C}$ is in the ϵ-neighborhood of $\hat{\Phi}_{ijk}$ if $|z - \hat{\Phi}_{ijk}| \leq \epsilon$. Additionally, we define $\mathcal{K}_{ij}$ as the set of nodes close to nodes i and j, according to the parameter τ: $\mathcal{K}_{ij} \triangleq \{k \in \mathcal{O}\backslash(i,j) | \max(\hat{d}_{ik}, \hat{d}_{jk}) \leq \tau\}$. The CRG parent/child test is:

$$\hat{\beta}_{ij} \triangleq \frac{1}{|\mathcal{K}_{ij}|} \sum_{k \in \mathcal{K}_{ij}} (|\hat{d}_{ij} - \hat{\Phi}_{ijk}|) \leq \epsilon \tag{2.9}$$

and the sibling test is:

$$\hat{\Gamma}_{ij} \triangleq \frac{1}{|\mathcal{K}_{ij}|(|\mathcal{K}_{ij}| - 1)} \sum_{k \in \mathcal{K}_{ij}} \sum_{k' \in \mathcal{K}_{ij}/k} (|\hat{\Phi}_{ijk} - \hat{\Phi}_{ijk'}|) \leq \epsilon, \tag{2.10}$$

where $\hat{\beta}_{ij}$ is the average distance between the complex quantities $\hat{d}_{ij}$ and $\hat{\Phi}_{ijk}$ for $k \in \mathcal{K}_{ij}$, and $\hat{\Gamma}_{ij}$ is the average distance between $\hat{\Phi}_{ijk}$ and $\hat{\Phi}_{ijk'}$ for $k, k' \in \mathcal{K}_{ij}$. (2.10) may incorrectly include parent nodes pairs in the set of siblings. This is resolved by removing the parent node from the set of siblings, once/if the parent is identified by (2.9). Finally, the equations that determine the effective impedances for the new parent nodes (eqns. $(27) - (28)$ in [27]) are linear, and do not need to be modified for complex quantities.

2.7 Simulation and Results

We test our methods on synthetic data generated by simulating seven different radial test feeders with the power flow tool MATPOWER [34]. To emulate realistic conditions, we use minute-resolution real power consumption data of nearby homes from the Pecan Street project, thereby preserving load correlations that exist in the real world. These correlations manifest in voltage and current measurements and can affect estimation performance. The seven cases have 4, 9, 12, 13, 13, 18, and 29 load nodes. The two 13 load cases differ in their topologies. All test cases have 3 passive nodes, except the 12 load case which has 4. Five of the seven test cases are visualized in Fig. 2.5. For each case, we collect time series phasor measurements at the active (load and substation) nodes *only*. To model real D-PMU measurements, we add *complex additive white Gaussian noise* (AWGN) to the voltage and current measurements generated by MATPOWER. This a reasonable noise model choice based

(4)

4 load node test case
5 active nodes, 3 passive nodes

(9)

9 load node test case
10 active nodes, 3 passive nodes

(12)

12 load node test case
14 active nodes, 4 passive nodes

(13b)

13 load node test case
14 active nodes, 3 passive nodes

(13)

13 load node test case
14 active nodes, 3 passive nodes
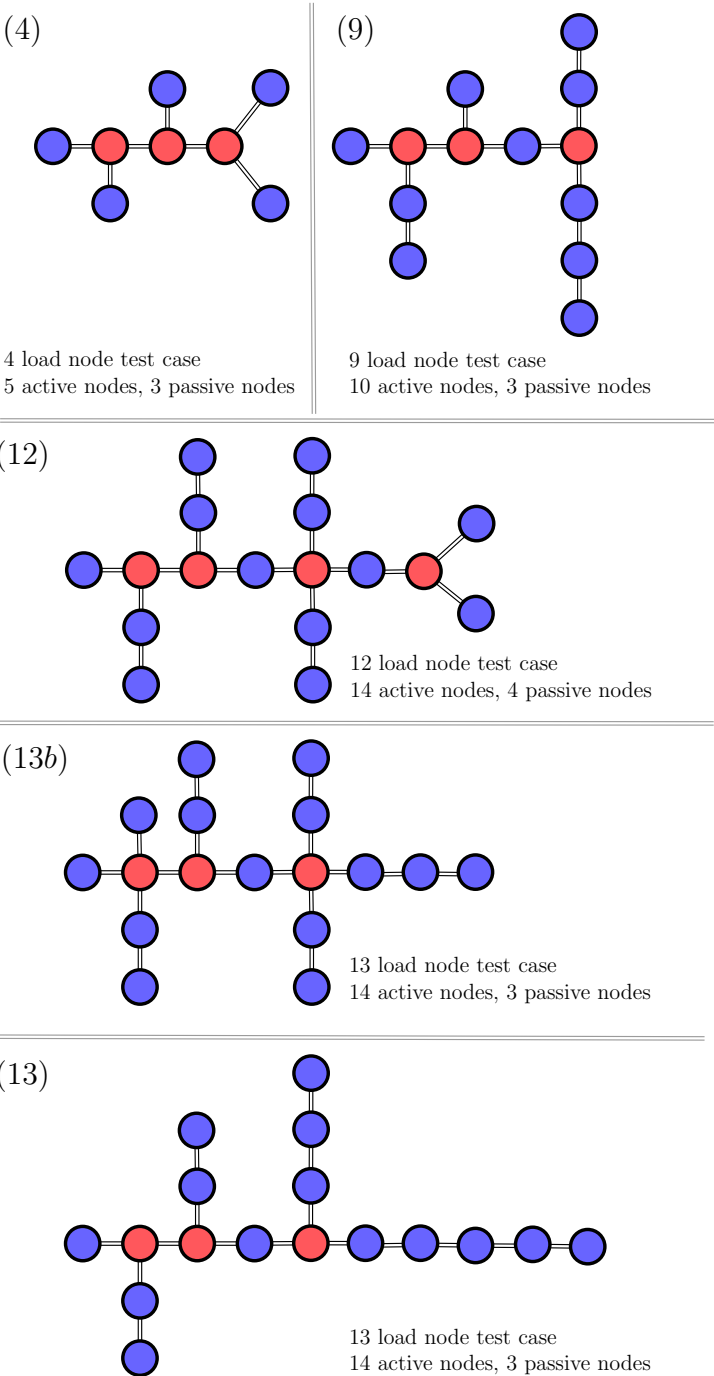
Figure 2.5: Visualization of five of the total seven test networks. Passive nodes are red, active nodes are blue. The set of active nodes is the set of load nodes plus the substation. The measured nodes are all active nodes.
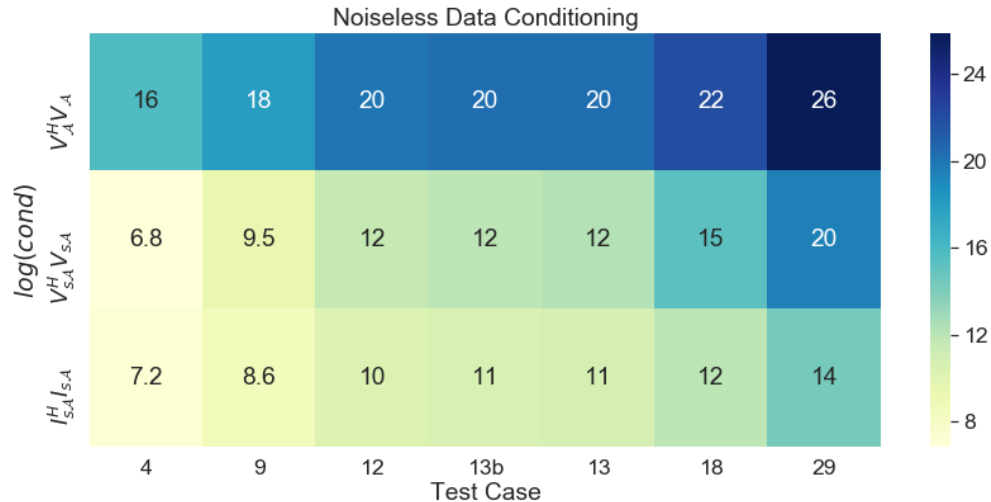
Figure 2.6: Log of the condition number of noiseless measurements for each test case. Larger values lead to greater noise sensitivity.

on empirical PMU noise studies [35]. The noise level is specified as a signal-to-noise ratio (SNR): the mean squared magnitude of the true measurement time series divided by the noise variance. Note that the noise sensitivity of OLS is determined by the condition number. By definition, the condition number reflects estimate sensitivity to the *norm* of the noise, *not* its distribution. Therefore, though we use AWGN in our simulations, the relative performance of the methods will be the same for other noise distributions [36]. Generating voltage measurements with MATPOWER takes on average 0.98 ms per load per time point. That is, simulating a 10 node network over 1000 time points takes 9.8 seconds. OLS network estimation takes 0.7 s. per node per trial.

Effective Impedance Estimation Results

From the noisy measurements of each test case, we estimate $\zeta_{\mathcal{A}}$ via three different approaches. We estimate $\hat{Y}_K$, $\hat{Y}_{sK}$, and $\hat{Z}_{sK}$ and then compute $\zeta_{\mathcal{A}}$ from each using the equations of Section 2.4. We determine the error between the three different estimates—$\hat{\zeta}_{\mathcal{A}}$ from $\hat{Y}_K$, from $\hat{Y}_{sK}$ and from $\hat{Z}_{sK}$—and the true $\zeta_{\mathcal{A}}$. These are all matrix quantities, so we report the normalized *Total Vector Error*:

$$\text{nTVE}(M, \hat{M}) \triangleq \frac{||M - \hat{M}||_2}{||M||_2}$$

$\zeta_{\mathcal{A}}$ estimation accuracy versus SNR is shown in Fig. 2.7. We see errors are larger and decay more slowly with increasing SNR for test cases with more loads. The conditioning trend for the test cases is recorded in Fig. 2.6. As the number of loads (and consequently measurement points) grows, the nodes are more electrically proximate and therefore the voltage measurements more correlated. This leads to poorer conditioning resulting in increased estimation

error and noise sensitivity. This is a compelling and perhaps counter-intuitive result—more PMU measurements do not necessarily improve impedance estimation accuracy if there is noise in the measurements.

Fig. 2.7 also conveys the advantage of the subKron representation over the Kron representation. $\zeta_{\mathcal{A}}$ estimation via $\hat{Y}_K$, $\hat{Y}_{sK}$ and $\hat{Z}_{sK}$ is affected by the conditioning of $V_{\mathcal{A}}$, $V_{s\mathcal{A}}$ and $I_{s\mathcal{A}}$ respectively. Since $V_{\mathcal{A}}$ is the worst conditioned, $\hat{\zeta}_{\mathcal{A}}$ extracted from Kron admittance Y_K has high error at low SNRs and consistently high estimate variance. This represents the performance of existing methods (Assuming that the existing methods use OLS to estimate Y_K). $\hat{\zeta}_{\mathcal{A}}$ extracted from Z_{sK} is clearly the most noise-robust with orders of magnitude lower error at low SNRs and dramatically lower estimate variance across the entire SNR test range. Estimation error drops more rapidly with increasing SNR for $\hat{\zeta}_{\mathcal{A}}$ extracted from Z_{sK} than from Y_K. The improvement is especially pronounced for the larger test cases, as Y_K and Y_{sK} estimation require inversion of the increasingly correlated and ill-conditioned voltage measurements, while Z_{sK} estimation inverts (mostly) uncorrelated current measurements.

While low mean error is obviously critical, low error variance is also important. The high variance of noise sensitive estimation methods has operational consequences. For example, a dramatic change in the network impedance estimate may be due simply to measurement noise, as opposed to real topological changes or faults, leading to frequent false alarms that waste time *and* reduce operator trust in both the sensors and algorithms. The low error variance is an important advantage of the subKron representation over prior work.

Fig. 2.8 plots $\zeta_{\mathcal{A}}$ estimation error for increasing number of measurement data points (T). The SNR is fixed. The Z_{sK} method provides orders of magnitude better performance at small values of T indicating that it is more suitable than existing methods for real time applications. The Z_{sK} method also provides better performance at large values of T, and unlike the Y_K method, increasing T results in a significant reduction of $\hat{Z}_{sK}$ error across all of the test networks.

Tree Reconstruction Results

To evaluate CRG reconstruction performance, we compute the Robinson-Foulds metric (RF) used in computational biology on phylogenetic trees [37]. RF quantifies the difference between two trees that share an active node set. It assesses the validity of each reconstructed edge independent of reconstruction performance upstream or downstream of that edge. We introduce the Normalized Robinson-Foulds metric (nRF), defined as the RF metric divided by the number of active nodes: $\text{nRF} = \frac{\text{RF}}{A}$. Normalization enables rough performance comparison across networks of different sizes. To evaluate the reconstructions, we use nRF to compare the reconstruction with the true network. Thus, the nRF serves as an error metric; the lower scores are better.

CRG performance is evaluated on two input sets. The "baseline" set uses effective impedances $\hat{\zeta}_{\mathcal{A}}$ calculated from the estimate $\hat{Y}_K$—the estimate hypothetically provided by existing methods that do not consider measurement data conditioning. The second set uses $\hat{\zeta}_{\mathcal{A}}$ calculated from $\hat{Z}_{sK}$. Fig. 2.9 shows CRG consistently performs better with the more

accurate effective impedances calculated from $\hat{Z}_{sK}$, than $\hat{Y}_K$, as expected. Performance is significantly better across a range of feeder sizes for SNRs $10 - 1000$.

CRG execution time scales nonlinearly with network size. CRG on $\hat{\zeta}_{\mathcal{A}}$ from $\hat{Y}_K$ runs in .06, .39 and 107 s on 4, 9, and 29 load networks respectively. CRG on $\hat{\zeta}_{\mathcal{A}}$ from $\hat{Z}_{sK}$ is faster—.06, .34 and 65 s, respectively. CRG is faster on more accurate effective impedances because the ϵ tolerance is not expanded as often and in each iteration it is more likely that multiple observed nodes are connected.

2.8 Conclusion

This chapter presents tools for drawing visualizable and actionable information—effective impedances and radial reconstruction—solely from D-PMU data, without prior information. Our primary insights are that:

- It is generally impossible to uniquely recover the complete network even from comprehensive V and I phasor measurements. However, it is always possible to recover pairwise active node effective impedances, $\zeta_{\mathcal{A}}$, from V and I measurements at all active nodes. In radial networks, the complete network can in fact be uniquely recovered from $\zeta_{\mathcal{A}}$.

- The subKron representation, though informationally equivalent to the Kron reduced representation, provides a more noise robust approach for $\zeta_{\mathcal{A}}$ estimation.

- The complex recursive grouping (CRG) algorithm recovers the full network from $\zeta_{\mathcal{A}}$ for radial topologies, and performs well on noisy $\hat{\zeta}_{\mathcal{A}}$.

- Simulations demonstrate the efficacy and practical challenges of the subKron and CRG methods. SubKron-based $\zeta_{\mathcal{A}}$ estimation outperforms Kron-based estimation. All of the impedance estimation techniques become increasingly noise sensitive as network size grows. Also, CRG execution time increases nonlinearly with reconstruction dimension. These practical limitations create SNR-dependent limits on the size of networks that can be estimated.

This chapter also establishes a foundation for future research directions. These research directions include extending the techniques to deal with missing measurement points, 3-phase unbalanced networks, improving the execution time of the CRG algorithm, and applying optimization techniques designed specifically for the impedance estimation problem.
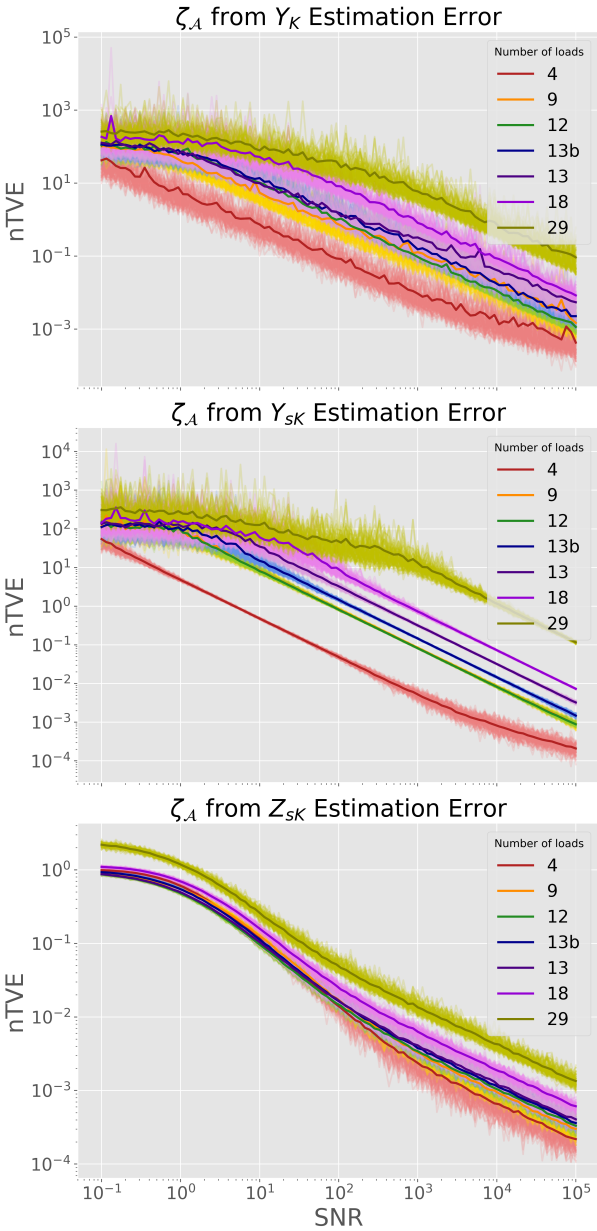
Figure 2.7: OLS $\zeta_{\mathcal{A}}$ estimation error with three approaches for five test cases with varying noise in I and V. Light bands indicate error across 100 trials per SNR value with different noise instances. Dark lines show average error.
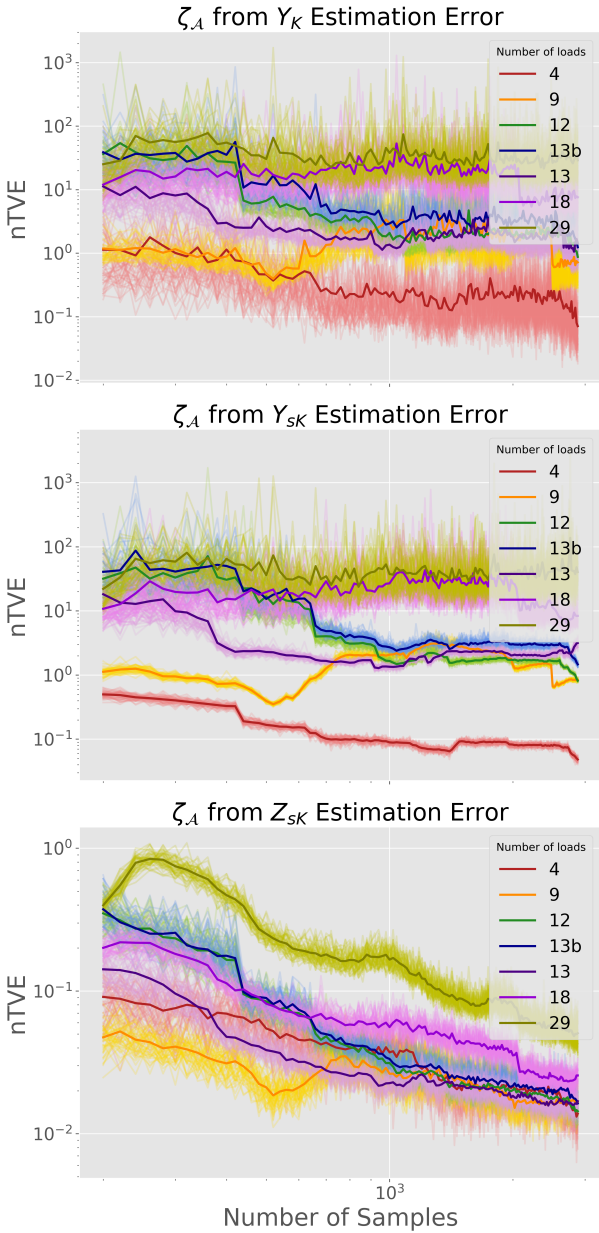
Figure 2.8: OLS $\zeta_{\mathcal{A}}$ estimation error for five test cases with varying number of measurement time points T used for estimation. The noise level in I and V is constant SNR = 100. Light bands indicate error across 100 trials per T choice with different noise instances. Dark lines show average error.
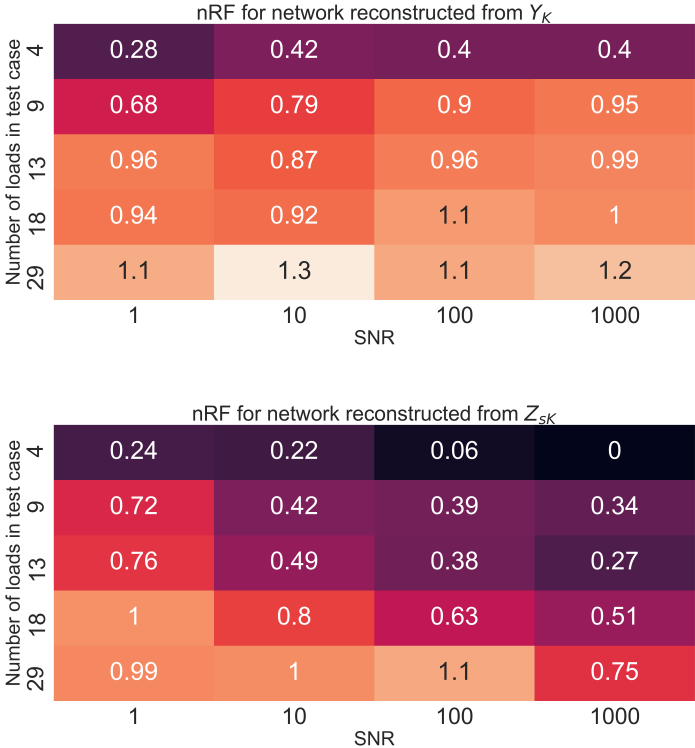
Figure 2.9: nRF of radial network reconstructions from Y_K (top) and Z_{sK} (bottom). The subKron method leads to lower reconstruction error than the Kron method across test case sizes for SNRs $10-1000$. Performance is comparable for SNR $= 1$. nRF values are averaged over 5 trials.

Chapter 3

Power System Hybrid State Estimation

Both human and automated grid operators require an estimate of the power system state to make decisions. The power system state consists of the AC voltage phasors (voltage magnitudes and angles) at all of the nodes in the system. The voltage phasor state vector and the network model fully define all of the power flows on the network.

Power system state estimation (PSSE), originally proposed in 1968 [38], uses the power flow equations and a network model to aggregate noisy voltage magnitude, voltage phasor, bus injection and line flow measurements into a single state estimate. Thus, PSSE requires an accurate network model in order to produce an accurate state estimate. Network models consist of the network topology and the line impedance values. The network topology can change with time as the network is reconfigured by opening/closing the network switches [39].

In conventional state estimation, the network topology and the state estimate are treated separately. The topology is computed offline by a Network Topology Processor (NTP) that receives status updates from the switches throughout the network [39]. However, in real-world operation, these switch signals can be unreliable or unavailable. An alternative to the conventional PSSE + NTP combination is generalized state estimation, which estimates both the topology and the voltage phasor states [40]–[44].

Generalized state estimation methods that do use a grid model [40]–[42] [45]–[47]. [45] Does topology detection using the normalized lagrangian multipliers for the constraints corresponding to the switch configuration. The normalized lagrangian multipliers are used to create a set of suspect switch configurations, with the help of a geometric interpretation of the lagrange multiplier vector. [46] Like above, but can detect both topology errors and bad data [47] Observes the normalized residual of pseudo-measurements correponding to the switch status, Treats the switching status as an additional continuous state variable

In this chapter, we propose modelling the generalized PSSE problem as a hybrid system state estimation problem, treating the network configuration and network voltages as discrete and continuous states, respectively. We note that there are multiple uses of the word "hybrid"
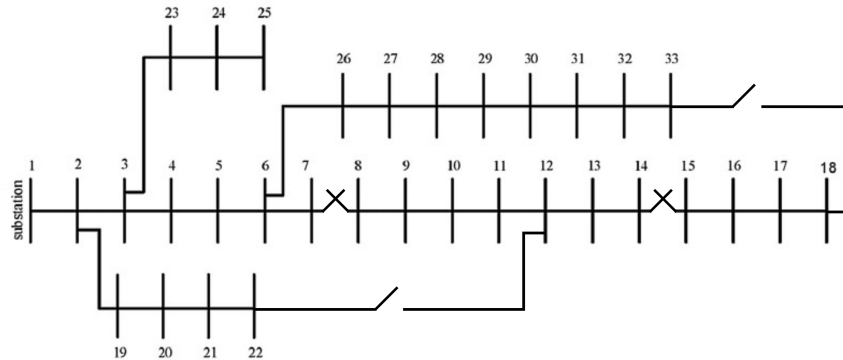
Figure 3.1: Modified 33 node distribution grid to include four switches

to describe PSSE formulations in the power systems literature [48]–[51]. The most common use of "hybrid" in the power systems literature refers to PSSE formulations which include hybrid *data* from both traditional SCADA measurements and Phasor Measurement Units (PMU) [50], [51]. To avoid ambiguity, we introduce two names—"Hybrid Data Power System State Estimation" to refer to PSSE formulations that use both PMU and SCADA *data*, and "Power System Hybrid State Estimation" (PSHSE) to refer to the generalized power system state estimation problem formulated using a hybrid *state*. In this chapter, we focus on the PSHSE problem, with the understanding that the formulation is agnostic to the type of data (PMU, SCADA, or both) that it receives.

In addition to the generalized/non-generalized designation, PSSEs can also be classified as static or dynamic. Conventional PSSEs are static, with Weighted Least Squares (WLS) being the industry standard regression tool [39]. In WLS, the voltage phasor state is estimated by evaluating the WLS normal equations using the measurements at the given timestep [52]. In order to use the WLS normal equations, the nonlinear measurement equations are linearized. Because the standard WLS only uses measurements at the given timestep, the state estimate can be susceptible to errors from the measurement noise in the sensors.

In contrast to static state estimation, dynamic state estimation explicitly models the power system as a dynamic system with a state that evolves over time. The Kalman Filter is an appealing option for dynamic PSSE due to its closed form solution and intuitive tuning parameters—the noise covariances in the dynamics and measurement equations. However, the canonical Kalman Filter is defined for linear systems. The most common approach to applying the Kalman Filter to the nonlinear measurement equations of the PSSE is the Extended Kalman Filter (EKF) [53]. The Unscented Kalman Filter (UKF) is an alternative to the EKF [54]. The UKF has been applied to the PSSE problem in [55] and [56], which demonstrate superior performance when compared with the EKF.

This chapter presents the Multiple model Adaptive Power system State Estimator (MAPSE), a novel approach to the dynamic, generalized PSSE problem that explicitly models an electric grid as a hybrid system. The MAPSE estimates the power system's hybrid state by imple-
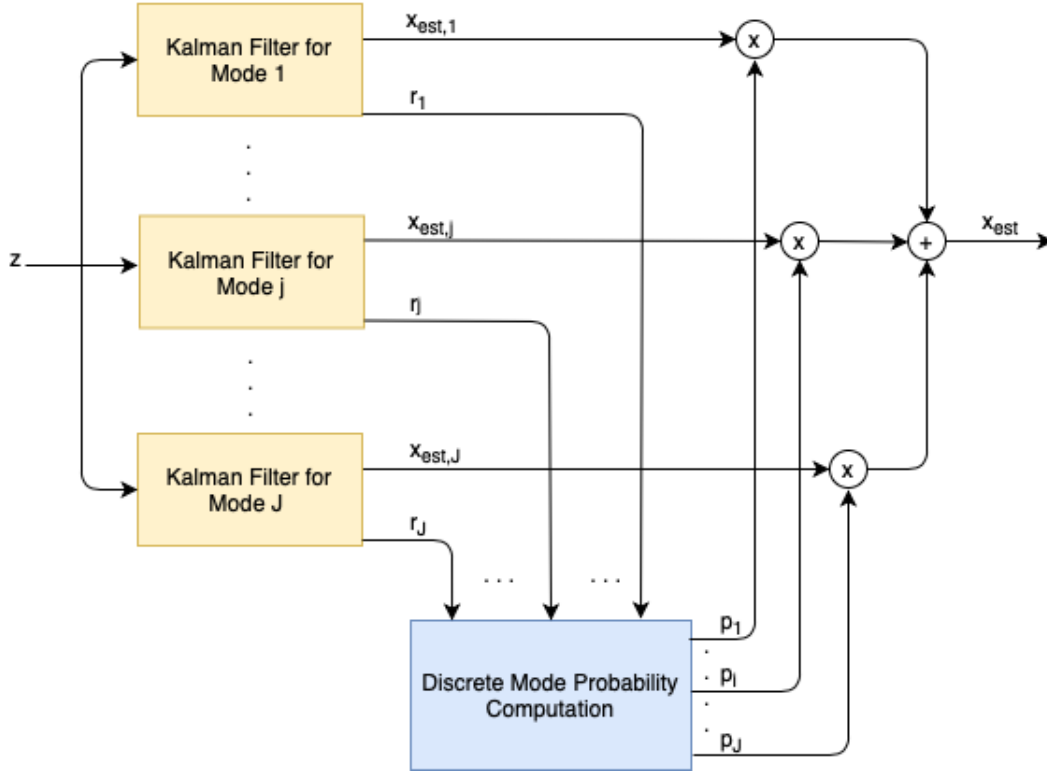
Figure 3.2: MMAE for a linear system with J modes

menting a Multiple Model Adaptive Estimation (MMAE) [57]–[59] with a bank of dynamic state estimators. In this chapter, we use a bank of UKFs as the dynamic state estimators.

3.1 Background

Multiple Model Adaptive Estimation for Linear Systems

The standard, linear system MMAE can be used for linear system, hybrid state estimation applications by creating one continuous state estimator for each of the discrete states/modes [57], [58]. For a linear system with J discrete states, there will be J different modes M_j, $j \in 1, \ldots, J$. $\mathcal{M}$ is the set of discrete states M_j with cardinality J. Each of the modes has its own linear system and corresponding Kalman Filter, constituting one of the filters in the J-long filter bank. For mode j at timestep k, the dynamics matrix $\Phi_j(k)$, measurement matrix $H_j(k)$, continuous state $x_j(k) \in \mathbb{R}^n$, dynamics noise $v_j(k)$ with covariance matrix $Q_j(k)$, and measurement noise $w_j(k)$ with covariance matrix $R_j(k)$, define the linear system

$$x_j(k) = \Phi_j(k-1)x(k-1) + v_j(k-1),$$
$$z(k) = H_j(k)x_j(k) + w_j(k).$$

All of the filters are presented the measurement $z(k) \in \mathbb{R}^m$.

Fig. 3.2 demonstrates how the standard, linear system MMAE is implemented with a bank of Kalman Filters. Each of the Kalman Filters in the bank sends the residual it records at each timestep to the Discrete Mode Probability Computation (DMPC) block. The DMPC block updates the discrete mode probabilities p_j ($p_1 + \cdots + p_J = 1$) using Bayesian recursion. The Bayesian recursion equation (3.12) is given in 3.2.

The Bayesian recursion requires the conditional measurement likelihood $f_z(z(k)|M_j, Z(k-1))$, where $Z(k-1)$ is the set of all measurements prior to timestep k. The MMAE approximates the conditional measurement probability density function $f_z(\cdot|M_j, Z(k-1))$ for mode j with a multivariate Gaussian with measurement covariance $P_{zz,j}(k)$, centered at the predicted measurement $\hat{z}_j(k) = H_j(k)\hat{x}_j(k)$, where $\hat{x}_j(k)$ is the jth model's state estimate at time k. For linear systems, the $P_{zz,j}(k)$ for each $M_j \in \mathcal{M}$ at time k can be calculated offline from the given filter's state covariance $P_j(k)$ using

$$P_{zz,j}(k) = H_j(k)P_j(k)H_j^T(k) + R_j(k).$$

Power System Continuous State and Measurements

For a power system network with $n/2+1$ nodes, $x \in \mathbb{R}^n$ is a vector of voltage magnitudes and angles that constitute the continuous voltage phasor state at all of the non-substation nodes. (We assume that the substation voltage magnitude is known, and that the substation serves as the phasor angle reference node.) $\hat{x}_j(k)$ is the jth mode's voltage phasor state estimate at time k, with voltage magnitude and angle estimates for each of the non-substation nodes. $z(k) \in \mathbb{R}^m$ is the vector of noisy measurements provided to the PSHSE, which could include noisy voltage magnitude, voltage angle, real and/or reactive power injection, real and/or reactive line flow, and current injection/flow measurements.

3.2 Multiple Model Adaptive Power System Estimation

The contribution of this chapter is the Multiple model Adaptive Power system State Estimator (MAPSE) for the PSHSE problem. The MAPSE implements an MMAE on a *nonlinear* system with one voltage phasor vector (continuous state) estimator for each of the possible network topologies (discrete states/modes). The voltage phasor vectors are estimated with UKFs—each candidate network topology has its own UKF [54]. Fig. 3.3 demonstrates the MAPSE design proposed in this chapter. We describe the signal flow of fig. 3.3 in 3.2, after first describing the UKF bank and the DMPC block.

Unscented Kalman Filter Bank

The UKF bank implements a bank of J UKFs, each with a different system model. To reduce the computational burden of implementing the UKF, we implement the version of the UKF that assumes an additive noise dynamic system [60]:

$$x_j(k) = g_j\big(x_j(k-1)\big) + v_j(k-1),$$
$$z(k) = h_j\big(x_j(k)\big) + w_j(k),$$

where g_j and h_j are the nonlinear dynamics and nonlinear measurement equations, respectively, for mode j. Furthermore, in dynamic power system state estimation, it is common to make the martingale assumption, $\mathbb{E}[x(k)] = x(k-1)$ [61]. With the martingale assumption, $g_j\big(x_j(k-1)\big) = x_j(k-1)$.

At each timestep, each UKF in the filter bank calculates the $(2n+1)$ sigma points using the standard selection method for the Unscented Transform (step 1 in algorithm 1). In algorithm 1 the subscript $_p$ corresponds to prediction (or prior) quantities, while $_u$ corresponds to update (or posterior) quantities. Each UKF passes the sigma points $s_{i,x_{u,j}(k-1)}$, $i = 0, \ldots, 2n$, through the state dynamics equations to get the dynamics-transformed sigma points $s_{i,x_{p,j}(k)}$, then passes the dynamics-transformed sigma points through the measurement equations to produce the measurement-transformed sigma points $s_{i,z_j(k)}$. The $s_{i,x_{p,j}(k)}$ and $s_{i,z_j(k)}$ points are combined using the equations in Step 4 of Algorithm 1 to determine the posterior state estimate for the jth filter at time k, $\hat{x}_{u,j}(k)$, and the posterior covariance estimate for the jth filter at time k, $\hat{P}_{u,j}(k)$. $\hat{x}_{u,j}(k)$ and $\hat{P}_{u,j}(k)$ are used to produce the sigma points for the next timestep.

Algorithm 1: The MAPSE Unscented Kalman Filter

Step 1: Select sigma points: For $\Gamma = \sqrt{(n + 1/2)\hat{P}_{u,j}(k-1)}$, and defining $(\cdot)_i$ as the ith row of $(\cdot)$,

$$s_{0,x_{u,j}(k-1)} = \hat{x}_{u,j}(k-1)$$
$$s_{i,x_{u,j}(k-1)} = \hat{x}_{u,j}(k-1) + (\Gamma)_i \text{ for } i = 1, \ldots, n$$
$$s_{i,x_{u,j}(k-1)} = \hat{x}_{u,j}(k-1) - (\Gamma)_i \text{ for } i = n+1, \ldots, 2n$$

Step 2: Pass the sigma points through the dynamics equations:

$$s_{i,x_{p,j}(k)} = g_{k-1}\big(s_{i,x_{u,j}(k-1)}\big) \ \forall i = 1, \ldots, 2n+1 \tag{3.1}$$

Step 3: Pass the sigma points through the measurement equations:

$$s_{i,z_j(k)} = h_k\big(s_{i,x_{p,j}(k)}\big) \ \forall i = 1, \ldots, 2n+1 \tag{3.2}$$

Step 4: Calculate $\hat{x}_{u,j}(k)$ and $\hat{P}_{u,j}(k)$:

$$\hat{x}_{p,j}(k) = \frac{1}{2n+1} \sum_{i=0}^{2n} s_{i,x_{p,j}(k)} \tag{3.3}$$

$$\hat{P}_{p,j}(k) = \frac{1}{2n+1} \sum_{i=0}^{2n} \Big[\big(s_{i,x_{p,j}(k)} - \hat{x}_{p,j}(k)\big)$$

$$\big(s_{i,x_{p,j}(k)} - \hat{x}_{p,j}(k)\big)^T\Big] + Q_j(k-1) \tag{3.4}$$

$$\hat{z}_j(k) = \frac{1}{2n+1} \sum_{i=0}^{2n} s_{i,z_j(k)} \tag{3.5}$$

$$\hat{P}_{zz,j}(k) = \frac{1}{2n+1} \sum_{i=0}^{2n} \Big[\big(s_{i,z_j(k)} - \hat{z}_j(k)\big)$$

$$\big(s_{i,z_j(k)} - \hat{z}_j(k)\big)^T\Big] + R_j(k) \tag{3.6}$$

$$\hat{P}_{xz,j}(k) = \frac{1}{2n+1} \sum_{i=0}^{2n} \Big[\big(s_{i,x_{p,j}(k)} - \hat{x}_{p,j}(k)\big)$$

$$\big(s_{i,z_j(k)} - \hat{z}_j(k)\big)^T\Big] \tag{3.7}$$

$$K_j(k) = \hat{P}_{xz,j}(k)\hat{P}_{zz,j}(k)^{-1} \tag{3.8}$$

$$r_j(k) = z(k) - \hat{z}_j(k) \tag{3.9}$$

$$\hat{x}_{u,j}(k) = \hat{x}_{p,j}(k) + K_j(k)r_j(k) \tag{3.10}$$

$$\hat{P}_{u,j}(k) = \hat{P}_{p,j}(k) - K_j(k)\hat{P}_{zz,j}(k)K_j(k)^T. \tag{3.11}$$

Discrete Mode Probability Computation Block

The DMPC block determines the discrete mode probabilities $p_j(k)$ using an MMAE, which implements the Bayesian recursion

$$
\begin{aligned}
p_j(k) &= f_M\big(M_j|z(k), Z(k-1)\big) \\
&= \frac{f_z\big(z(k)|M_j, Z(k-1)\big) f_M\big(M_j|Z(k-1)\big)}{f_z\big(z(k)|Z(k-1)\big)} \\
&= \frac{f_z\big(z(k)|M_j, Z(k-1)\big) p_j(k-1)}{\sum_{i=1}^{J} f_z\big(z(k)|M_i, Z(k-1)\big) p_i(k-1)}
\end{aligned}
\tag{3.12}
$$

for each $M_j \in \mathcal{M}$. $f_M(\cdot|z(k), Z(k-1))$ is the probability mass function for all of the modes, conditioned on $z(k)$ and $Z(k-1)$. $f_z(\cdot|M_j, Z(k-1))$ is the probability density function for measurement $z(k)$, conditioned on M_j and $Z(k-1)$. The prior probabilities for each $M_j \in \mathcal{M}$ in the Bayesian recursion are taken to be the posterior state probabilities at the previous timestep, $p_j(k-1)$.

The mode probability recursion (3.12) requires the conditional measurement likelihoods $f_z(z(k)|M_j, Z(k-1))$ for each of the modes. As in 3.1, the measurement distribution for each of the modes is approximated by a Gaussian. Thus, $f_z(z(k)|M_j, Z(k-1))$ is approximated by evaluating each mode's Gaussian at time k,

$$
\begin{aligned}
f_z\big(z(k)|M_j, Z(k-1)\big) &\approx \mathcal{N}\big(r_j; 0, \hat{P}_{zz,j}\big) \\
&= \frac{1}{(2\pi)^{m/2}|\hat{P}_{zz,j}|^{1/2}} \exp\Big(-\frac{1}{2} r_j^T(k)\hat{P}_{zz,j}^{-1}(k) r_j(k)\Big)
\end{aligned}
\tag{3.13}
$$

for each mode's residual at time k,

$$
r_j(k) = z(k) - \hat{z}_j(k).
$$

(3.13) requires the measurement covariance $P_{zz,j}$. Unlike the linear system MMAE, for nonlinear systems, $P_{zz,j}$ cannot be computed offline. Fortunately, the jth UKF estimates $\hat{P}_{zz,j}$ using equation (3.6) in order to produce the gain matrix from the sigma points. Thus, each of the UKF blocks can pass $\hat{P}_{zz,j}$ to the DMPC, along with r_j.

Combining the UKF bank and the DMPC Block

Each UKF in the filter bank (fig. 3.3) maintains an internal state estimate $\hat{x}_{u,j}$ using Algorithm 1. Each UKF of the bank sends its measurement residual r_j, as well as its measurement covariance estimate $P_{zz,j}$, to the DMPC block at each timestep. The DMPC block uses the Bayesian recursion (3.12) to maintain a probability mass function for each of the discrete operating states $M_j \in \mathcal{M}$. Regarding the final continuous state estimate at time k—the discrete probabilities can be used to combine the $\hat{x}_j$ estimates from each of the Kalman Filters into a single estimate of $\hat{x}$, as done in Fig. 3.3. Alternatively, the continuous state estimate from the discrete mode with the highest probability can be used.
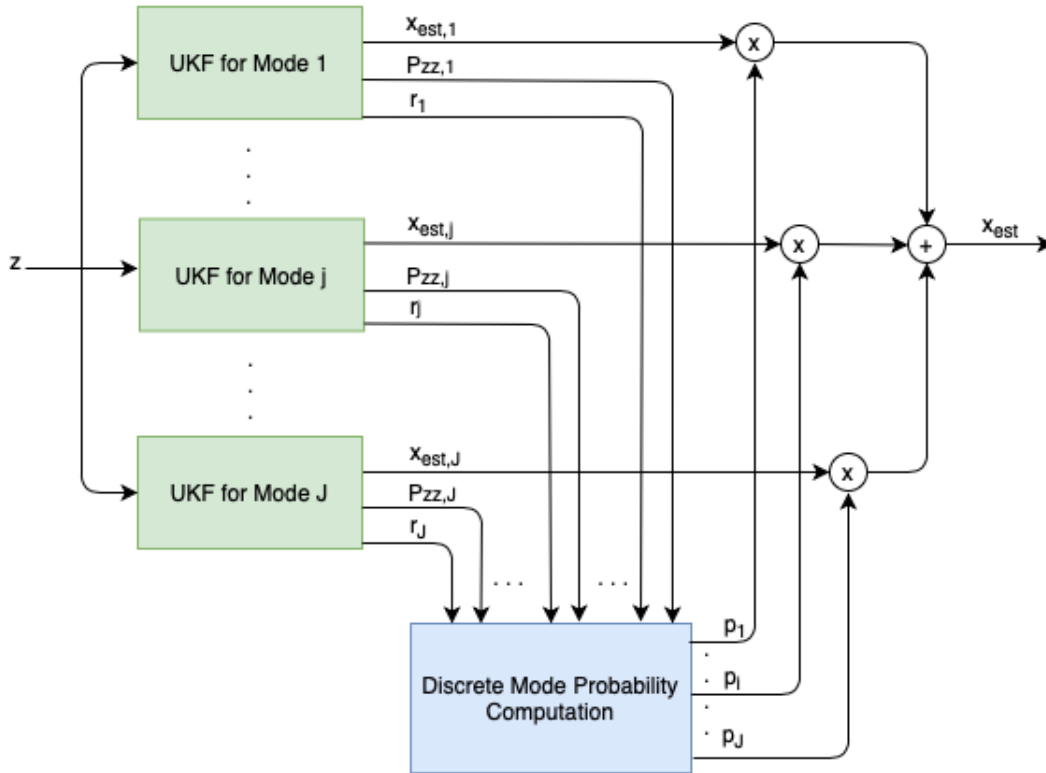
Figure 3.3: Multiple model Adaptive Power system State Estimation implemented with Unscented Kalman Filters

3.3   Results

Test Set Up

This section reports preliminary results demonstrating the efficacy of the MAPSE on a modified version of the single phase, 33 node Baran and Wu test network [62]. We modified the 33 node network to include two of the additional lines listed in the "REDS" (REpository of Distribution Systems) [63] between nodes 18 and 33, and 12 and 22, and put switches between nodes 18 and 33, 12 and 22, 7 and 8, and 14 and 15, as demonstrated in fig. 3.1. We considered four radial candidate topologies, or discrete modes, A, B, C, and D, as shown in fig. 3.4. The MAPSE had access to two sets of measurements:

1. high Signal to Noise Ratio (SNR) PMU measurements on nodes 18, 22, and 33, and

2. low SNR power injection/load pseudo-measurements on all of the load nodes (nodes 2-33).

   To create the time series simulation, we assigned load profiles to each node using the first 32 real power injection/load profiles in the Pecan Street data set [64]. We created reactive

(a) Network Configuration A

(b) Network Configuration B

(c) Network Configuration C

(d) Network Configuration D
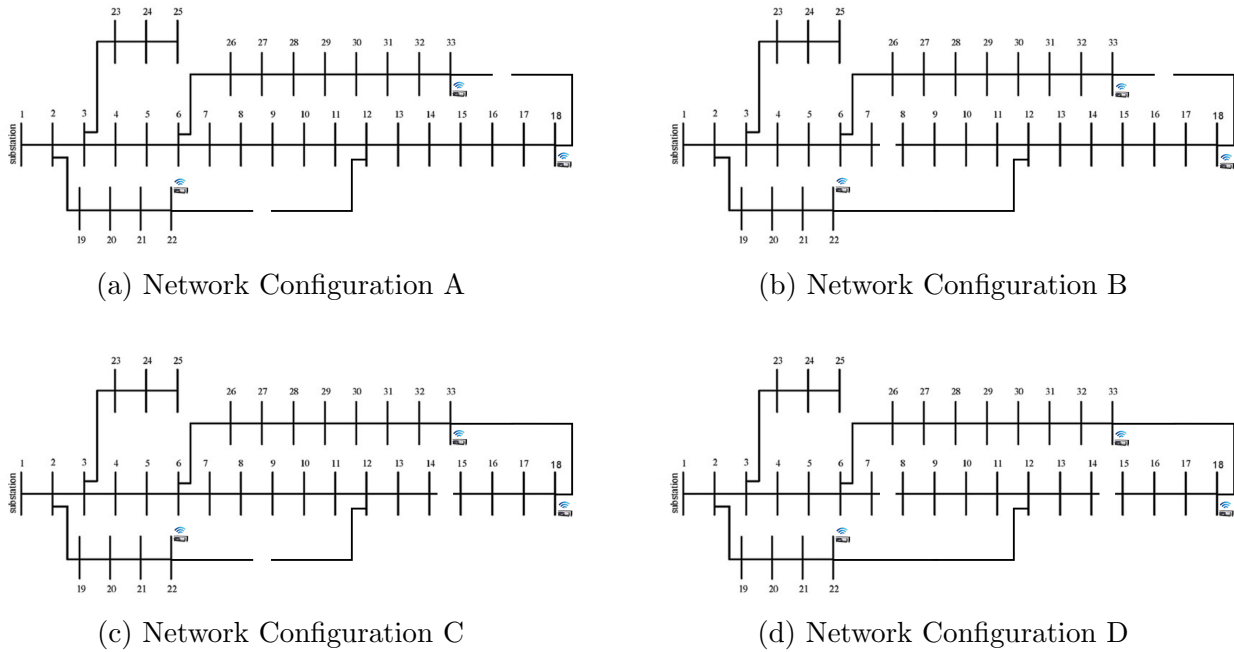
Figure 3.4: Four candidate 33 node reconfigurations with PMUs on nodes 18, 22, and 33

power injection/load data by multiplying the real power data by a constant multiple of .15. We multiplied the Pecan Street load profiles (given in kW) by 20 in order to get a significant voltage drop along the network (multiplying the loads by 20 resulted in the lowest voltage on the network being .95 p.u.). The base quantities for the per unit (p.u.) normalization were chosen according to the REDS model—the base voltage was 12.66 kV, the base power was 5.68 MVA, and the base impedance was 28.2 Ω. We stepped through the Pecan Street load profiles, selecting a load value for each timestep for each node on the network, and holding the voltage at the substation (node 1) equal to 1. At each timestep, we solved power flow using the MATPOWER power flow solver [65] to get the true network state. We then added noise to the measurements. For all three PMU, we added 10^{-6} p.u. Gaussian noise to both the real and imaginary portions of the true voltage phasors, then calculated the magnitudes and the angles to create the noisy phasor measurements. The 10^{-6} p.u. noise quantity is based on the empirical observations of real PMU data in [66]. For the load pseudo-measurements we added white guassian noise to each node's true injection so that the SNR for each power injection/load pseudo-measurement was 0.5. The "signal" for these pseudo-measurements was the variance of the Pecan Street power injection profiles.

Remark 3.3.1. We recognize that the network, load profile, and pseudo-measurement noise model do not represent a true distribution network. In particular, distribution networks are generally larger than 33 nodes and unbalanced three-phase, the aggregation of loads decreases the variance of the power injections, and pseudo-measurement noise is highly correlated be-

tween timesteps and non-Gaussian. The preliminary results reported in this chapter indicate that evaluating the MAPSE on a more realistic test case is warranted.

MAPSE Settings

We ran the MAPSE hybrid state estimator with four UKF blocks—one for each of the A, B, C and D grid configurations. For the dynamic PSSE problem, both the dynamics $g()$ and the (additive) dynamics noise covariance Q can be treated as design choices. We used the martingale assumption for the dynamics model: $g_j(x) = x$. We used $10^{-6}I$ for the dynamics noise covariance matrix Q (I is the identity matrix), after observing the behavior of the MAPSE for Q matrices with different orders of magnitude. We chose R according to the noise in each measurement signal, as described in 3.3. For each of the UKF blocks, we initialized the state covariance P to $10^{-5}I$, the estimates of the voltage magnitudes for all of the nodes to 1, and the estimates of the voltage angles for all of the nodes to 0.

The noise-sensitivity of the likelihood calculation (3.13) used in the DMPC's Bayesian recursion (3.12) increases with the dimension of the measurement residuals r_j. To avoid prohibitive noise-sensitivity, we used only the high SNR PMU measurements (not the load pseudo-measurements) at nodes 18, 22, and 33 in the DMPC, which reduced the dimension of the residuals used by the DMPC from 70 to 6. In addition, we asserted a lower bound of .001 for each discrete mode, $p_j \geq .001 \ \forall M_j \in \mathcal{M}$, so that the discrete mode probabilities could recover in a reasonable amount of time if the configuration switched.

MAPSE Voltage Phasor State Estimation

To evaluate the efficacy of the MAPSE voltage phasor state estimation, we ran a number of simulations on the modified 33 node feeder. We include the results from two indicative simulations here. In the first simulation, we simulated 100 timesteps in configuration A, then 100 timesteps in configuration C. Fig. 3.5 displays the node 14 voltage magnitude estimate produced by the dynamic UKF and static WLS state estimation methods for grid configurations A and C. In configuration A, The PMU at node 18 is close to node 14. In configuration C, there is no PMU close to node 14. Fig. 3.5a demonstrates that both the UKF and the WLS perform well when there is a low-SNR PMU measurement nearby, whereas fig. 3.5b demonstrates that the UKF outperforms the WLS estimation method when the only relevant measurements are the high-SNR load pseudo-measurements.

In the second simulation, we simulated 100 timesteps and switched the grid configuration from A to B at timestep 25, B to C at timestep 50, and C to D at timestep 75. Fig. 3.6 demonstrates the mean Total Vector Error (TVE) for each model's voltage phasor state estimate at each step of the 100 timestep simulation. For model j at timestep k, we define

$$\text{mean TVE} = \frac{1}{n/2} \sum_{l}^{n/2} \left(\frac{|V_{j,l,k}^{\text{est}} - V_{l,k}^{\text{tru}}|}{|V_{l,k}^{\text{tru}}|} \right),$$
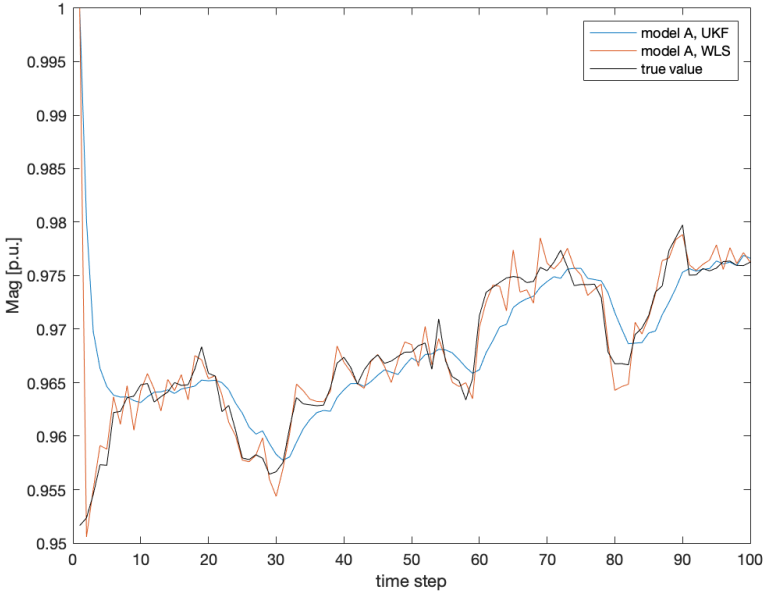
(a) Configuration A
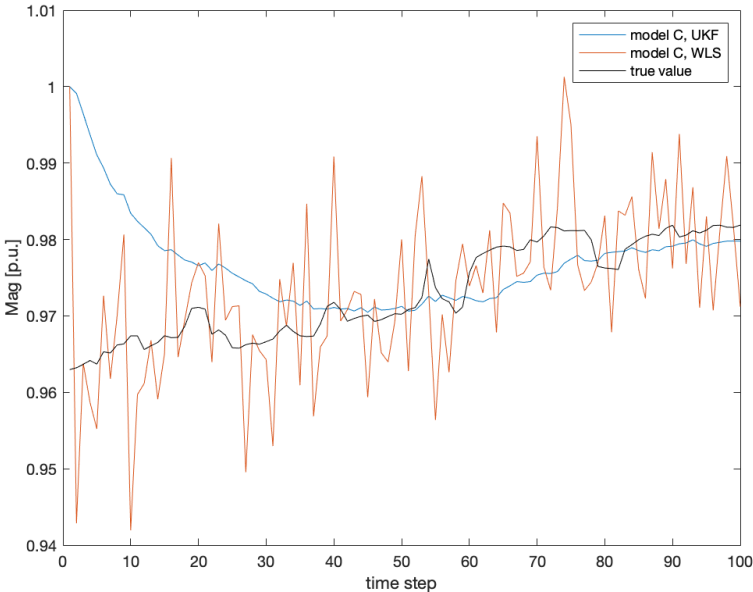
(b) Configuration C

Figure 3.5: A comparison of the node 14 voltage magnitude estimates produced by the dynamic UKF and static WLS state estimation methods. No grid topology changes.
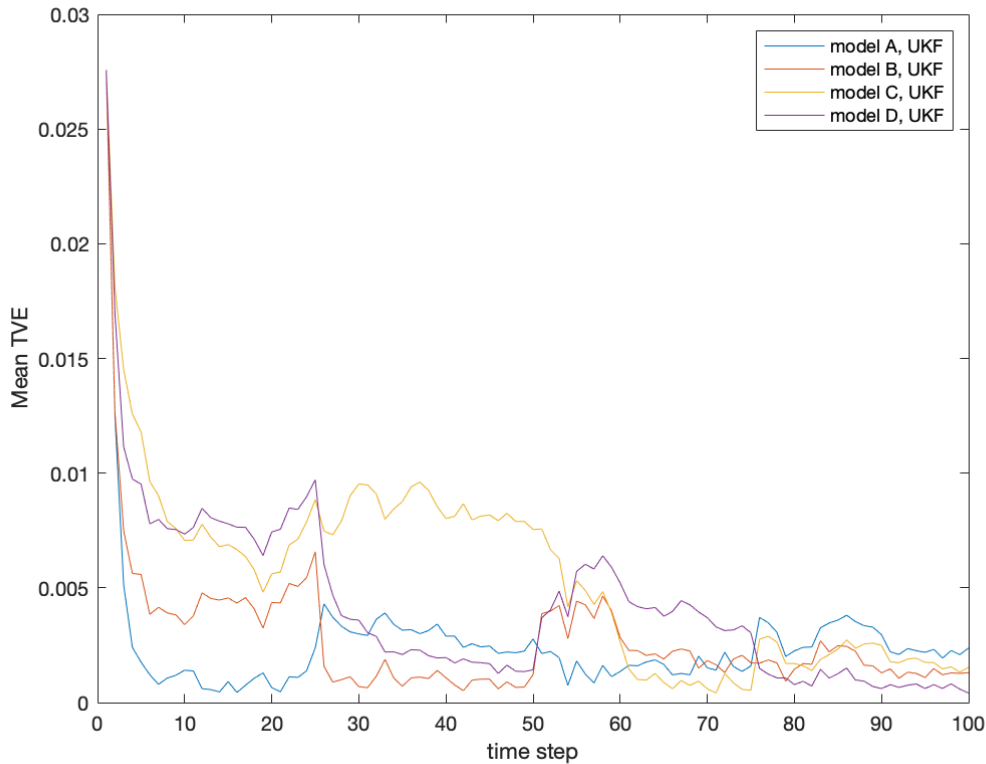
Figure 3.6: The total vector error produced by each UKF in the MAPSE bank in the simulation in which the grid topology was switched from A to B to C to D at timesteps 25, 50, and 75, respectively.

where $V_{l,k}^{\text{tru}} \in \mathbb{C}$ is the true complex voltage phasor (state) of node l at time k, and $V_{j,l,k}^{\text{est}} \in \mathbb{C}$ is the complex voltage phasor (state) of node l at time k, determined from $\hat{x}_j(k)$. Fig. 3.6 demonstrates that, for the simulation described in 3.3, the UKFs described in 3.3 do a reasonable job of estimating the voltage phasor state. Furthermore, the UKF that is using the correct model eventually produces the smallest mean TVE. Note, the UKF dynamics can take a while to settle, as is evident in the transition from configuration B to C at timestep 50.

MAPSE Topology Detection

Fig. 3.7 demonstrates that the MAPSE accurately detects the network topologies A, B, C, and D, and the MAPSE is able to adjust its discrete mode hypothesis when the network topology is changed. The transition between configurations B and C is of particular interest because the DMPC is able to detect the network change from the measurement residuals r
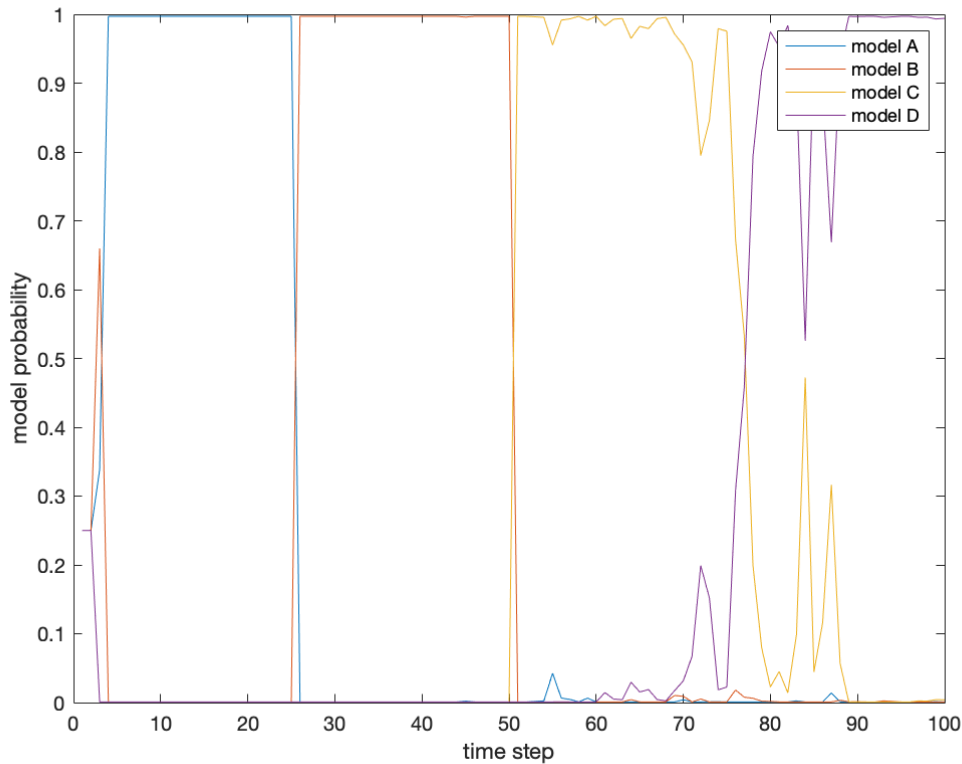
Figure 3.7: The discrete mode probabilities produced by the MAPSE in the simulation in which the grid topology was switched from A to B to C to D at timesteps 25, 50, and 75, respectively.

before the mean TVE for mode C's UKF block decreases to become the smallest of the four configurations.

3.4   Conclusion

This chapter describes the PSHSE problem, presents the MAPSE technique for addressing the PSHSE problem, and provides preliminary results demonstrating the efficacy of the MAPSE. The PSHSE problem provides a number of directions for future work. These directions include rigorous evaluation of the MAPSE's performance on realistic distribution and transmission networks, improving the dynamics equation used by the UKFs, and expanding the PSHSE functionality to include additional network settings and parameters such as transformer tap settings, capacitor bank settings, and other network settings/operation modes that can be modeled as discrete states.

Chapter 4

Voltage Phasor Control for Distribution Networks

The proliferation of distributed solar generation and the electrification of transportation and heating end uses present a novel set of challenges for electricity distribution networks. For many distribution networks/grids/circuits, these challenges include voltage magnitude and line flow constraint violations.

As an alternative to upgrading distribution network infrastructure to accommodate peak power demand/distributed generation, constraint violations can be mitigated by actively managing Distributed Energy Resources (DERs) such as batteries or controllable loads [67], [68]. Both "Distributed Energy Resource Management Systems" (DERMS) and "Active Distribution Networks" (ADNs) [69]–[71] actively manage DERs.

Fig. 4.1 illustrates a simple example using the IEEE 13 Node Feeder. Suppose that large electric vehicle charging loads at nodes 652, 680, and 675 create a line flow/thermal constraint violation between nodes 632 and 671, and an undervoltage constraint violation at node 675. With a sufficiently large DER (e.g. a battery) at node 692, some combination of real and reactive power injections would be able to alleviate both the line flow and undervoltage constraint violations.

In the transmission context, Optimal Power Flow (OPF) is used for economic dispatch, determining how much power each generator should produce subject to physical network constraints. OPF could also be used to actively manage DERs to enhance reliability and economic operation of distribution networks [72]. However, the communication requirements of applying a power flow optimization [73] may be problematic for distribution system applications. Because distribution networks benefit from less statistical aggregation of load than transmission networks, the "disturbances," or unanticipated changes in the real/reactive power injections on the network, occur at a faster rate. Thus, real and reactive power commands determined by an online distribution network OPF calculation may be outdated by the time they are implemented (e.g., if the cloud cover changes, affecting distributed solar generation).

On transmission networks, rather than implementing the centralized OPF solutions in
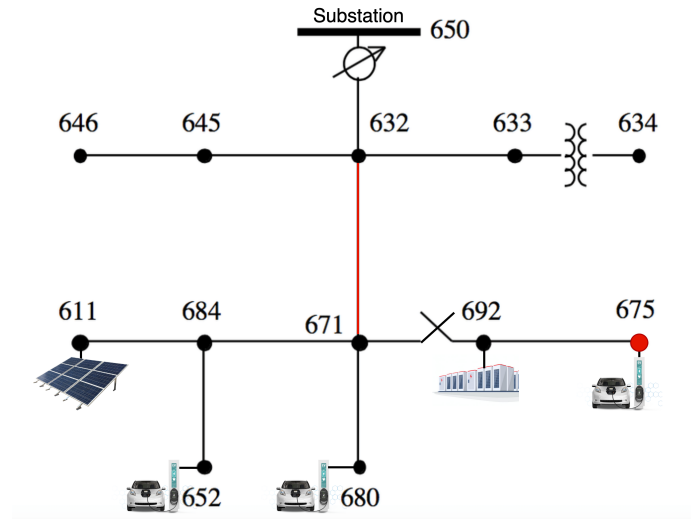
Figure 4.1: The IEEE 13 Node Feeder with a line flow constraint violation on the line between nodes 632 and 671, and a voltage magnitude constraint violation at node 675. (The normally closed switch between nodes 671 and 692 is not relevant to this example.)

open loop, the power generation decisions are adjusted in real time using distributed feedback loops including droop control, automatic voltage regulation, and Automatic Generation Control (AGC) based on the tie line flows [74]–[76]. There are distributed feedback control methods for distribution networks as well, including Volt-Watt and Volt-VAR control [68], [77]–[82], which define real or reactive power injections, respectively, as functions of voltage magnitude. Volt-Watt control can be combined with Volt-VAR control to avoid voltage violations, however the appropriate tradeoff between Volt-Watt and Volt-VAR responsibilities is situational [83]–[85]. [86] provides an optimization-based approach to specifying the Volt-Watt and Volt-VAR curves that minimize voltage deviations and provide bounded-input-bounded-state stability. [68], [79] show that, in some circumstances, Volt-VAR control can alleviate over-voltages on networks with high PV penetration, but can also increase line flow magnitudes. Similarly, Fig. 9 in [87] demonstrates that Volt-VAR control can *reduce* dynamic photovoltaic hosting capacity, if the line flow thermal constraints are binding.

Voltage Phasor Control (VPC), also called Phasor Based Control [88], is a novel paradigm for implementing OPF which incorporates distributed feedback controllers, but in a different manner than AGC or Volt-Watt/Volt-VAR control. With VPC, the optimization broadcasts voltage phasor (magnitude and angle) setpoints, rather than real and reactive power setpoints, to participating nodes equipped with Phasor Measurement Units (PMUs). VPC has been successfully demonstrated in Hardware-In-the-Loop simulations at Lawrence Berkeley Lab's FLEXLAB test facility [89], [90].

In this chapter, we compare the performance of VPC with Voltage Magnitude Control (VMC) [91], [92], where the power flow optimization assigns voltage magnitudes and injection

(a) Centralized OPF implemented with direct power commands

(b) Voltage Phasor Control
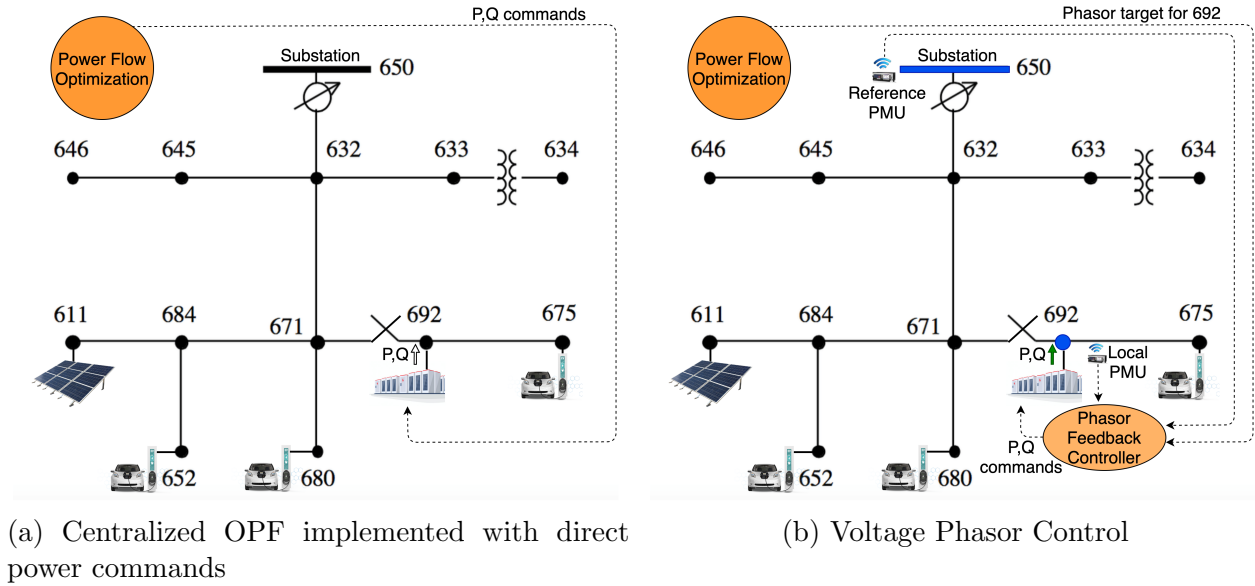
Figure 4.2: Two different control methods for node 692 of the IEEE 13 Node Feeder

power factors to the distributed feedback controllers. When the injection power factor is 1, VMC corresponds to "stiff" Volt-Watt control, where the real power injection at each participating node is modulated to keep the local voltage magnitude equal to the assigned target. When the injection power factor is 0, VMC corresponds to stiff Volt-VAR control.

Specifically, we analyze how effectively the real and reactive power injections under VPC's distributed feedback mitigate the adverse effects of disturbance power injections on the network voltages and upstream line flows. By reducing the impact of disturbance injections, VPC feedback control allows the power flow optimization to be less overcautious than it would be with open loop power commands, and thus attain a more efficient/cost-effective operating point.

Remark 4.0.1. *VPC is agnostic to the OPF implementation. It can be used with centralized OPF, as we assume in this chapter, or with decentralized OPF.*

4.1 Voltage Phasor Control

An OPF solution contains both the optimal power injections and the corresponding voltage phasors. In standard OPF implementation, as demonstrated on the IEEE 13 Node Feeder in Fig. 4.2a, the OPF broadcasts power setpoints to the DER at node 692. In the Introduction, we describe this method as "open loop" because the DERs adhere to the power setpoints regardless of evolving grid conditions.

Control Type	P	Q	V	θ
PQ Control	x	x		
PV Control	x		x	
VPC			x	x

Table 4.1: How different bus types relate to the power flow manifold

VPC is an alternative, "closed loop" method that broadcasts the OPF's voltage phasors to the participating Phasor Controlled Nodes (PCN). Note, the OPF is only able to broadcast two set points because the power flow manifold permits two degrees of freedom for the state vector consisting of voltage magnitude V, voltage angle θ, real power injection P, and reactive power injection Q [93], [94]. Table 4.1 displays how VPC and the standard *PQ* and *PV* control schemes relate to the power flow manifold.

Algorithm 2 describes how VPC is implemented. At each VPC node, a Phasor Feedback Controller (PFC) [95], [96] adjusts the power injections in order to maintain its voltage phasor assignment despite evolving network disturbances. Fig. 4.2b demonstrates VPC with a single PCN at node 692.

Algorithm 2: Voltage Phasor Control

Optimization

At a slower cadence:

1. Run OPF using load/generation predictions.

2. Broadcast phasor targets to the PFCs.

3. Gather measurements, adjust load predictions,
   repeat.

Distributed Phasor Feedback Controllers (PFCs)

Each PFC, at a faster cadence:

1. Measure the local voltage phasor, receive the phasor reference angle from the substation.

2. If a new phasor target has been received from the optimization, update phasor target.

3. Adjust the P and Q of subordinate DERs so that the local voltage phasor is closer to the phasor target,
   repeat.
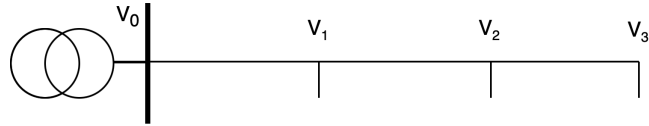
Figure 4.3: Four node circuit

VPC requires synchronized phasor measurements from PMUs, rather than standard voltage magnitude measurements. Furthermore, the phasor measurements require a reference node that defines the "0" angle. For distribution networks, the substation is the natural choice for the phasor reference node as it behaves approximately like a slack bus, supplying the net power imbalance incurred by the load and generation at all of the other nodes on the network.

VPC is well-suited for distribution networks with volatile disturbances because the PFCs make power adjustment decisions in a distributed manner. Each PFC requires only the local phasor measurement and the reference phasor angle measurement, which is constantly broadcast to all of the PFCs, and therefore is able to immediately adjust the power injections of its subordinate DERs without waiting to hear from the OPF computation.

Remark 4.1.1. *While we consider power injection disturbances in this chapter, VPC is applicable to other applications including switch closing, phase balancing, and responding to line outages on mesh networks.*

4.2 VPC and Voltage Disturbance Sensitivity

To clearly demonstrate the effect that disturbance injections have on the voltages on the network, we use the simplest possible circuit—the four node circuit in Fig. 4.3. Node 0 is the substation, which we model as an infinite/slack bus with voltage 1 per unit (p.u.). We consider scenarios in which the circuit has one PCN, one disturbance injection node which is perturbed by i_{dist}, and one "Node of Concern" (NC)—the node at which we are concerned about voltage magnitude constraint violations. We consider the four PCN/NC/disturbance configurations in Table 4.2,[1] and make the assumption that both the pre-disturbance injections and the disturbance injections are constant-current to simplify the circuit analysis.[2] Columns 3 and 4 display the sensitivities of the NC voltage to the disturbance current injection when the system is run in open loop and when VPC is applied, respectively. The

---

[1]We do not include the scenarios which correspond to the circuits in rows 3 and 4 with the NC and disturbance nodes switched because these scenarios are similar to the circuits in rows 3 and 4.

[2]We choose to describe the disturbance as a current injection (rather than a power injection) because it makes the NC voltage linear with respect to the disturbance current injection, even when VPC holds the PCN constant. Note, this linearity is a convenient property of VPC only. It does not apply to circuits with VMC nodes. Linear equations simplify the sensitivity analysis because they are analytic and therefore the complex derivative is defined.
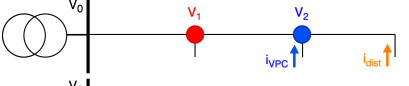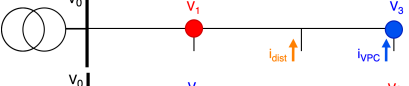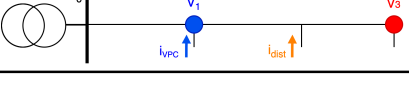
Row	Configuration	Open Loop Sensitivity	Sensitivity With VPC
1		$\frac{\partial v_1}{\partial i_3} = z_{01}$	$\frac{\partial v_1}{\partial i_3} = 0$
2		$\frac{\partial v_3}{\partial i_1} = z_{01}$	$\frac{\partial v_3}{\partial i_1} = 0$
3		$\frac{\partial v_1}{\partial i_2} = z_{01}$	$\frac{\partial v_1}{\partial i_2} = \frac{z_{01}}{z_{23}+z_{12}+z_{23}}z_{01}$
4		$\frac{\partial v_3}{\partial i_2} = z_{01} + z_{12}$	$\frac{\partial v_3}{\partial i_2} = z_{12}$

Table 4.2: Voltage disturbance sensitivities. The phasor controlled node (PCN) is blue and the node of concern (NC) is red.

circuit equations and derivations of the first two sensitivities for Table 4.2 are in Appendix C.1.

For the configurations in rows 1 and 2,[3] for which the disturbance is on the other side of the PCN than the NC, the disturbance has no effect on the voltage at the NC. That is, the PCN injection fully cancels the effect of the disturbance injection on the NC voltage. In the bottom two configurations the NC is on the same side of the VPC node as the disturbance. In these circumstances, for most z_{01} and z_{12} values, the PCN injection reduces the effect of the disturbance on the voltage, but does not eliminate it.

Implementation Insight 1. *When concerned with voltage magnitude(s) of a node or set of nodes, a PCN should be placed close to the node(s) of concern, preferably between the node(s) of concern and the disturbance node(s).*

4.3 VPC and Upstream Line Flow Disturbance Sensitivity

To clearly demonstrate the effect that disturbance injections have on upstream line flows for radial circuits, we use the same four node circuit in Fig. 4.3 and again make the assumption that the injections are constant-current. With this assumption, VPC only affects upstream line flows. In practice, VPC affects downstream line flows as well, due to the voltage dependencies of the downstream loads and generators. However these effects are second order and may not be significant.

---

[3]The circuits in Table 4.2 can be thought of as an equivalent circuit models for larger circuits. For example, the configuration in row 2 corresponds to the IEEE 13 Node Feeder circumstance in which node 675 is the NC, there is a PCN at node 692 and there is a disturbance injection at node 611.

Row	Configuration	Open Loop Sensitivity	Sensitivity With VPC
1		$\frac{\partial i_{01}}{\partial i_3} = -1$	$\frac{\partial i_{01}}{\partial i_3} = 0$
2		$\frac{\partial i_{12}}{\partial i_1} = 0$	$\frac{\partial i_{12}}{\partial i_1} = \frac{z_{01}}{z_{01}+z_{12}}$
3		$\frac{\partial i_{01}}{\partial i_1} = -1$	$\frac{\partial i_{01}}{\partial i_1} = -\frac{z_{12}}{z_{01}+z_{12}}$

Table 4.3: Line flow disturbance sensitivities. The phasor controlled node (PCN) is blue and the line flow of concern (LC) is red.

Similar to the voltage sensitivities, we define the "Line of Concern" (LC) for a given configuration as the line at which we are concerned about line flow constraint violations. We consider the three PCN/LC/disturbance configurations in Table 4.3.[4] The circuit equations and derivations for Table 4.3 are in Appendix C.2.

In the circuit in row 1 of Table 4.3, in which the disturbance is downstream of the PCN, the PCN completely shields the upstream line flows from downstream disturbance injections. In the circuit in row 2, in which the disturbance is upstream of the PCN and the LC, the VPC can either increase or decrease $|i_{12}|$ depending on the circumstance. In the circuit in row 3, in which the disturbance is upstream of the PCN but downstream of the LC, the disturbance affects i_{01} but VPC reduces the likelihood of i_{dist} producing a $|i_{01}|$ constraint violation.

Bounds on changes in the upstream current magnitude

Regarding line flow constraint violations, we are interested in the disturbance-sensitivity of current magnitudes. Unfortunately, generally

$$|\frac{\partial i_{01}}{\partial i_1}| \neq \frac{\partial |i_{01}|}{\partial i_1}.$$

Furthermore, the magnitude operator is not analytic, and therefore we cannot use a standard derivative.

Thus, we are left with conservative bounds on the changes of $|i_{01}|$. To articulate these bounds, we define the following variables for i_{01} and corresponding quantities for i_{12}:

- $\boldsymbol{\Delta i_1}$: The disturbance injection current at node 1.

---

[4]The circuits in Table 4.3 can also be thought of as equivalent circuit models for larger circuits. For example, the configuration in row 3 corresponds to the circumstance in which the line between nodes 671 and 632 is the LC, there is a PCN at 692, and there are disturbance injections at 652 and 680.

- : The i_{01} current before Δi_1 is added to the injection at node 1.

- $i_{01}^{\mathbf{ol}}$: The i_{01} current after Δi_1 is added when the system is run open loop (ol).

- $i_{01}^{\mathbf{VPC}}$: The i_{01} current after Δi_1 is added when VPC holds node 2's voltage phasor constant.

For clarity, Lemmas 4.3.1, 4.3.2, and 4.3.3 assume the pre-disturbance injections are constant-current and the disturbance is defined as a current injection as well.

Lemma 4.3.1. $|i_{01}^{\text{VPC}}| - |i_{01}^{\text{bef}}| \leq |\Delta i_1| |\frac{z_{12}}{z_{01}+z_{12}}|.$

Lemma 4.3.2. $|i_{12}^{\text{VPC}}| - |i_{12}^{\text{bef}}| \leq |\Delta i_1| |\frac{z_{01}}{z_{01}+z_{12}}|.$

Lemmas 4.3.1 and 4.3.2, proved in Appendix C.3, bound the changes in the upstream current magnitudes in terms of the disturbance magnitude $|\Delta i_1|$ and the line impedances.

While Lemmas 4.3.1 and 4.3.2 bound the changes in the upstream line flows, it does not state whether VPC *always* helps or hurts when a disturbance increases the upstream line flows. For the circuit in row 2 of Table 4.3, the effect of VPC depends on the directions of i_{12}^{bef} and Δi_1 in the complex plane, as well as the line impedances. For the circuit in row 3, however, we can bound the difference between $|i_{01}^{\text{VPC}}|^2$ and $|i_{01}^{\text{ol}}|^2$ for all disturbances that increase $|i_{01}^{\text{ol}}|$.

Lemma 4.3.3. *For the circuit in row 3 of Table 4.3, if z_{01} and z_{12} have the same X/R ratio, and if Δi_1 increases $|i_{01}|$, then*

$$|i_{01}^{\text{ol}}|^2 - |i_{01}^{\text{VPC}}|^2 > (1-a)|\Delta i_1|^2 > 0,$$
$$\text{where } a = \frac{z_{12}}{z_{01}+z_{12}} \in [0,1] \subset \mathbb{R}.$$

Lemma 4.3.3, proved in Appendix C.4, states that if the lines between nodes 0 and 1 and nodes 1 and 2 have the same X/R ratio, then the VPC at node 2 reduces the risk that Δi_1 will create an $|i_{01}|$ constraint violation, regardless of the pre-disturbance i_{01}, i_1, and $\hat{v}_2$ values. Furthermore, the intuition derived from the proof of Lemma 4.3.3 is that if the X/R ratios for z_{01} and z_{12} are not pathologically different, then VPC reduces the risk that Δi_1 will create an $|i_{01}|$ constraint violation.

Assessing the conservativeness of the upstream line flow bounds

The bounds in Lemmas 4.3.1, 4.3.2, and 4.3.3 are conservative. To assess the conservativeness of the bounds, we ran a test on the circuit in row 3 of Table 4.3 which compared the Lemma bounds with the observed change in $|i_{01}|$ for a range of disturbance injections at node 1. We set the line impedances $z_{01} = z_{12} = 0.5 + 0.5j$. The current magnitudes were normalized so
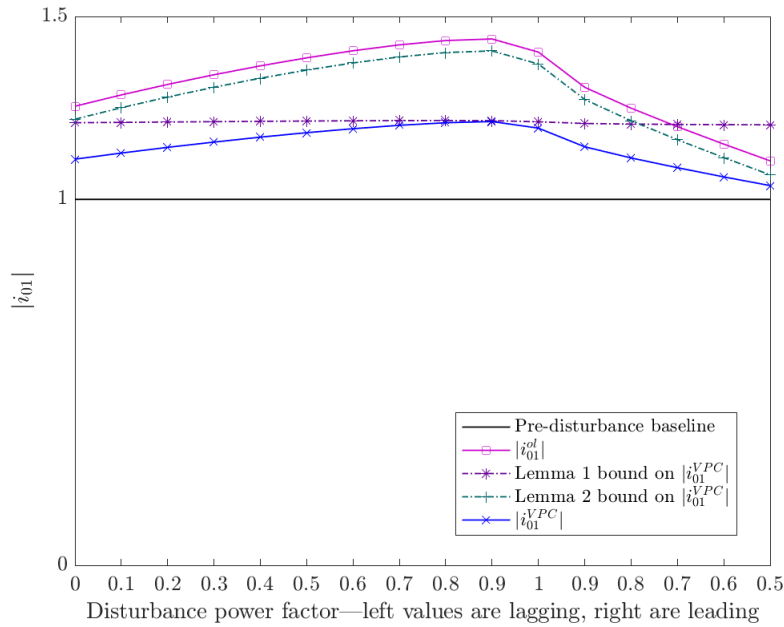
Figure 4.4: Comparison of the bounds in Lemmas 4.3.1 and 4.3.3 for the circuit in row 3 of Table 4.3 for disturbances with a range of power factors. (Lemma 4.3.1 is "Lemma 1" and Lemma 4.3.3 is "Lemma 2.")

that the pre-disturbance $|i_{01}| = 1$. The pre-disturbance voltage magnitudes at nodes 1 and 2 were 0.97 and 0.95 respectively. Disturbance injections with a range of power factors, from 0 to 1 lagging (extracting reactive power), and from 1 to 0.7 leading (injecting reactive power), were injected at node 1. The VPC at node 2 adjusted its injection in order to maintain its assigned voltage phasor.

Fig. 4.4 plots $|i_{01}|$ vs. the power factor of the disturbance at node 1 and demonstrates that the Lemma 4.3.1 bound ranges between conservative and tight. For this example, the Lemma 4.3.1 bound does not assert that VPC reduces the increase in $|i_{01}|$ when the disturbance has a leading power factor of 0.7 or less. Fig. 4.4 also demonstrates that, while the Lemma 4.3.3 bound is generally quite conservative, the Lemma 4.3.3 bound does accurately assert that VPC reduces the increase in the upstream line flow current magnitudes for all disturbance power factors.

Implementation Insight 2. *When concerned with a line flow or set of line flows, a PCN should be placed downstream, but as close as possible, to the line flow(s) of concern.*

4.4   Voltage Magnitude Control

As stated in Section 4.1, the power flow manifold permits two degrees of freedom. Thus, a voltage magnitude command for a given node must be accompanied by an additional state command. We choose the Adjustment Power Factor (APF) of the feedback control as the second state command, which includes stiff Volt-Watt and Volt-VAR control as the APF = 1 and APF = 0 special cases, respectively, and implement VMC with Algorithm 3.

Algorithm 3: Voltage Magnitude Control

Optimization

At a slower cadence,

1. Run OPF using load/generation predictions.

2. Broadcast magnitude targets and APFs to the MFCs.

3. Gather measurements, adjust load predictions,
   repeat.

Distributed Magnitude Feedback Controllers (MFCs)

Each MFC, at a faster cadence,

1. Measure the local voltage magnitude.

2. If a new magnitude target and/or APF has been received, update the magnitude target and/or APF.

3. Adjust the P and Q of subordinate DERs in way that maintains the APF and brings the local voltage magnitude closer to the magnitude target,
   repeat.

While there always exist real and reactive power adjustments at a given node that produce any voltage phasor assignment, there does not always exist real and reactive power adjustments with a given APF that produce any voltage magnitude assignment. That is, a function can be defined from an arbitrary domain of voltage phasors to the appropriate codomain of power adjustment using Ohm's Law. However not all APF/voltage magnitude target combinations can be mapped to power injections on the power flow manifold [94]. For example, consider a predominantly reactive network in which an MCN is assigned APF = 1 (stiff Volt-Watt control). Adjusting the real power injection will have limited effect on the voltage magnitude, thus many voltage magnitude assignments are infeasible.

Implementation Insight 3. *Care must be taken when selecting APFs for VMC. Assigning an infeasible voltage magnitude target/APF will produce voltage instabilities.*

We observed in simulations that VMC is effective in reducing the impact of disturbances on voltage magnitudes in the neighborhood of a given MCN node, as expected. As we highlight in the next section, while VMC usually helps upstream line flow constraint violations, it generally does so less effectively than VPC and, in some cases, exacerbates the increase in upstream line flow caused by the disturbance.

4.5 Comparing VPC and VMC for Upstream line flow Constraint Violations

To demonstrate how VPC and VMC reject the effect of disturbances on upstream line flows, we ran two sets of simulations similar to the set that created Fig. 4.4. We used the circuit in row 3 of Table 4.3 and again set $z_{01} = z_{12} = 0.5 + 0.5j$.

Simulation set I: non-nominal baseline

The first set of simulations use a pre-disturbance baseline with excess distributed generation, similar to the scenarios considered in [68], [79], and [87]. Both nodes 1 and 2 injected .06 W and extracted .02 VAR giving a .95 lagging power factor[5] and pre-disturbance $v_2 = 1.05\angle 1°$. The current magnitudes were normalized so that the pre-disturbance $|i_{01}| = 1$. Disturbance injections with a range of power factors from 0.7 to 1 lagging, and from 1 to 0.7 leading were injected at node 1. The PFC or MFC at node 2 adjusted its injection in order to maintain $\hat{v}_2 = 1.05\angle 1°$ or $|v_2| = 1.05$, respectively. Fig. 4.5 plots $|i_{01}|$ vs. the power factor of the disturbance. The black line is $|i_{01}^{\text{bef}}|$, the magenta line with squares is $|i_{01}^{\text{ol}}|$, and the blue line with x markers is $|i_{01}^{\text{VPC}}|$. The red, orange, and yellow lines with circles represent the post-disturbance $|i_{01}|$ with VMC applied to node 2 with lagging power factors of 0, 0.5, and 1, respectively. Fig. 4.6 plots the real and reactive power injection adjustments (control effort) at node 2 that the PFC and MFCs used to maintain the assigned voltage targets.

Fig. 4.5 demonstrates that VPC reduces the effect of the disturbance on $|i_{01}|$, regardless of the disturbance power factor. With APF = 0, VMC slightly increases the effect of the disturbance on $|i_{01}|$ for all but the most lagging power factor disturbance. This scenario is similar to the scenarios in [79] and [87] in which Volt-VAR control exacerbated the line flow/thermal constraint violations because the Volt-VAR controllers' VAR extractions reduced the power factor of the line flows. With APF = 0.5, VMC slightly reduces the effect of the disturbance on $|i_{01}|$ for leading power factor disturbances. With APF = 1, VMC reduces the effect of the disturbance on $|i_{01}|$ significantly for leading power factor disturbances. VMC's overreaction for leading power factor disturbances is explained by Fig. 4.6—the APF = 1 VMC brings the voltage at node 2 back down by extracting a lot of real power, which

---

[5]We use the convention that a "lagging" power factor extracts VARs, regardless of whether real power is being injected or extracted.
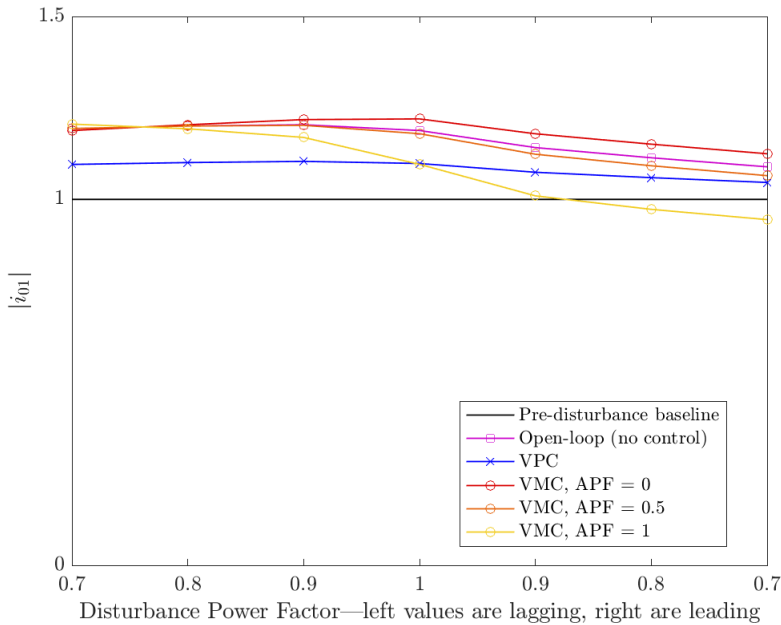
Figure 4.5: Line flows for the circuit in row 3 of Table 4.3 with a $|v_2| = 1.05$ non-nominal pre-disturbance baseline.
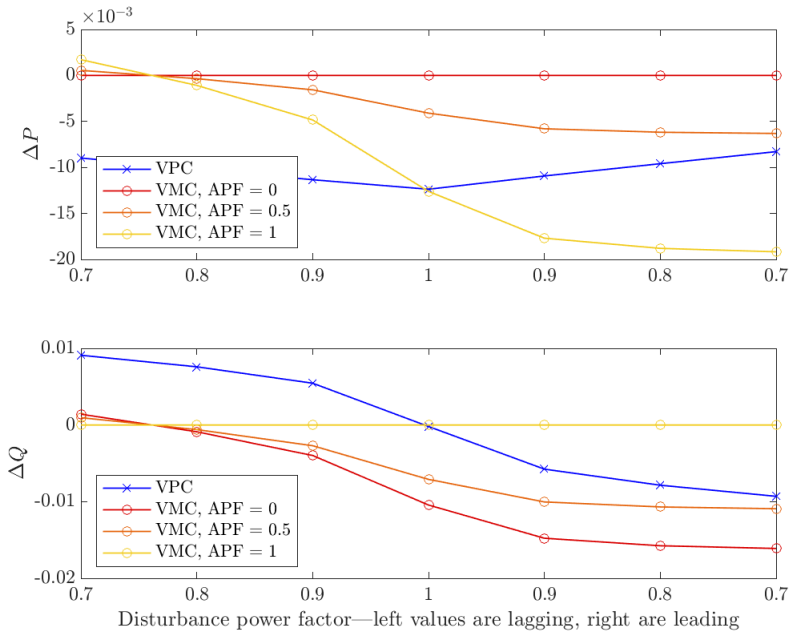
Figure 4.6: The changes in node 2's power injections corresponding to Fig. 4.5.
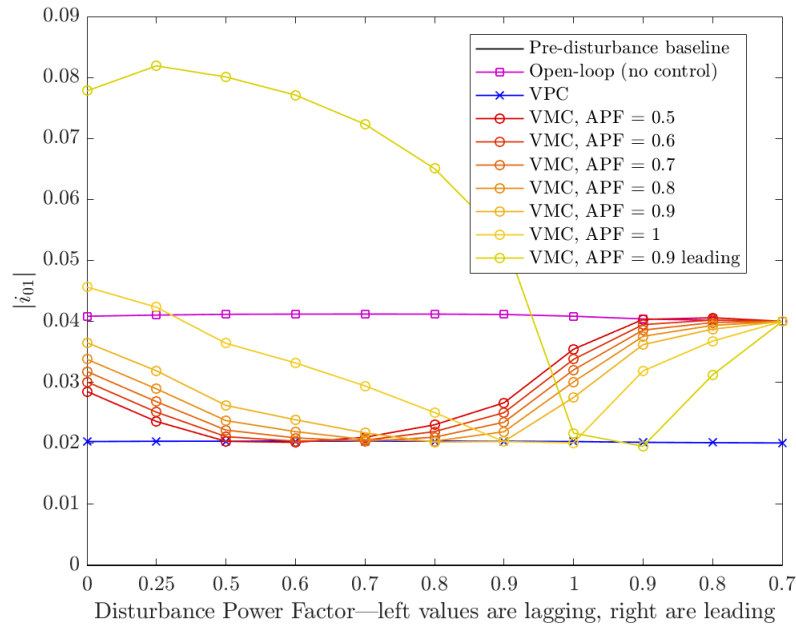
Figure 4.7: Line flows for the circuit in Row 3 of Table 4.3 with the nominal pre-disturbance baseline.

is expensive. This real power extraction cancels some of the pre-disturbance baseline real power flow from node 0 to 1.

Simulation set II: nominal baseline

In the second set of simulations, the pre-disturbance baseline for the test was $i_1 = i_2 = 0$, which results in $v_1 = v_2 = 1$ and $|i_{01}| = 0$ (the "nominal" baseline). Disturbance loads with a range of power factors from 0 to 1 lagging and from 1 to 0.7 leading were extracted from node 1, lowering node 2's voltage below 1. The PFC or MFC at node 2 adjusted its injection in order to maintain $v_2 = 1$. In addition to the standard power factors, we also included the "APF = 0.9 leading" power factor VMC. Unlike the standard power factor VMCs, when the voltage magnitude is higher than the target, the APF = 0.9 leading VMC extracts real power and counter-intuitively *injects* reactive power. When the voltage is lower than the target, it does the opposite. Fig. 4.7 plots $|i_{01}|$ vs. the power factor of the disturbance and provides a number of insights.

First, the 0.7 leading power factor disturbance injection does not result in any change in the voltage magnitude at node 2. Thus, none of the VMCs adjust the power injection at node 2, and therefore do not reduce the increase in the upstream line flow. The VPC, on the other hand, recognizes the change in the voltage phasor angle at node 2, and adjusts

its real and reactive power injections so that the increase in the line flow is one half of the no-control case.

Second, the VMC can exacerbate the effect of the disturbance on $|i_{01}|$. This is demonstrated by the 0.9 leading APF for disturbances with lagging power factors, and by the APF = 1 Volt-Watt controller when the disturbance power factor is close to zero.

Third, Fig. 4.7 demonstrates that the VMC that reduces the effect of the disturbance on the upstream line flow the most is the VMC for which the APF is matched to the power factor of the disturbance.

Fourth, the VPC injection is very similar to the VMC injection when the VMC's APF matches the disturbance injection's power factor. This is true even when the disturbance power factor is leading—when the disturbance power factor is leading and the voltage is low, the VPC *extracts* reactive power and injects real power. Extracting reactive power when the voltage is low is non-intuitive, but is the correct response for reducing the impact of a leading power factor disturbance. This VPC–VMC matching characteristic is not a hard rule for all pre-disturbance baselines and line impedances, however it does provide intuition for how VPC responds to disturbances: VPC injections tend to have power factors similar to the power factors of the disturbance injections that prompted the VPC injections.

Implementation Insight 4. *VPC usually outperforms VMC in reducing the effect of disturbances on line flows that are upstream of both the disturbance and the controlled node in two circumstances:*

1. *when the power factor of VMC does not match the power factor of the disturbance;*

2. *when the disturbance has a leading power factor.*

4.6 Conclusion

In this chapter we introduced Voltage Phasor Control as a novel approach for implementing OPF which incorporates distributed feedback controllers. We described how VPC feedback rejects the effects of disturbance injections on the network's voltage magnitudes and upstream line flows. Through simulations, we compared the performance of VPC and VMC. These results establish VPC as a tool that can help distribution networks operate safely and reliably within their constraints. Ongoing research directions include alternative distributed feedback laws that incorporate the voltage angle measurement, as well as the VPC power flow optimization formulation.

Chapter 5

Power Flow Linearization

In power systems analysis, we are often interested in how changes in the real and reactive power injections will affect the voltages on the network.[1] Unfortunately, power flow equations that explicitly give the network voltages as functions of the power injections do not exist [94].[2] However, at a given operating point, it *is* possible to provide the power-voltage sensitivity—a linear approximation of how changes in power injections change the voltages on the system. This is the goal of power flow linearization and this chapter. An overview of the prevalent power flow linearizations is given in Section 7.4.

5.1 Useful Operations for Power Flow Linearization

We find that the following operations, defined in [93], are useful for determining the power flow linearization.

Complex-to-real Operator

The $\langle \cdot \rangle$ operator converts from n-dimensional complex space to equivalent $2n$-dimensional real space.

$$\langle \cdot \rangle := \begin{cases} \begin{bmatrix} \mathrm{Re}(\cdot) & -\mathrm{Im}(\cdot) \\ \mathrm{Im}(\cdot) & \mathrm{Re}(\cdot) \end{bmatrix}, & \text{if } \cdot \text{ is a matrix,} \\ \begin{bmatrix} \mathrm{Re}(\cdot) \\ \mathrm{Im}(\cdot) \end{bmatrix}, & \text{if } \cdot \text{ is a vector.} \end{cases} \tag{5.1}$$

Appendix E.1 gives some intuition for how $\langle \cdot \rangle$ maintains the relationship between complex numbers in real space.

---

[1]Given the voltage changes, it is easy to derive the changes in the current flows on the network using Ohm's Law.

[2]Solving for the network voltages given the power injections is referred to as the "power flow," or "load flow" problem.

Operators for Differentiating Complex Vectors in Real Space

The following matrices/operations are helpful for calculating the partial derivatives of $\langle \cdot \rangle$:

$$\text{diag}(\cdot) := \text{A diagonal matrix with the entries of the vector } \cdot \text{ on the diagonal,}$$

$$N := \begin{bmatrix} I_n & 0_n \\ 0_n & -I_n \end{bmatrix},$$

$$R(\cdot) := \begin{bmatrix} \text{diag}\big(\cos(\angle\cdot)\big) & -\text{diag}(|\cdot|)\text{diag}\big(\sin(\angle\cdot)\big) \\ \text{diag}\big(\sin(\angle\cdot)\big) & \text{diag}(|\cdot|)\text{diag}\big(\cos(\angle\cdot)\big) \end{bmatrix}.$$

N is the derivative of the complex conjugate of a given variable,

$$N = \frac{\partial \langle \bar{\cdot} \rangle}{\partial \langle \cdot \rangle} = \left\langle \frac{\partial \bar{\cdot}}{\partial \cdot} \right\rangle.$$

$R(\cdot)$ is the derivative of a vector in rectangular coordinates with respect to its polar coordinates,

$$R(\cdot) = \frac{\partial \langle \cdot \rangle}{\partial \begin{bmatrix} |\cdot| \\ \angle\cdot \end{bmatrix}}.$$

Using the chain rule, both N and $R(\cdot)$ can be used to give the derivative of composite functions that contain complex conjugates and rectangular/polar coordinates, respectively.

5.2 The AC Power Flow Manifold

Given the network voltages, it is easy to solve for the network power injections using (5.2). However, given the network power injections, it is not easy to determine the network voltages. This is, in part, due to the fact that the map from the network voltages to the power injections is not a one-to-one function. As stated in the introduction of this chapter, in place of finding equations that give the voltages as functions of the power injections, we are interested in finding the power-voltage sensitivity at a given point. We achieve this by finding the tangent plane to the implicitly-defined power flow manifold [93], [94], [97].

AC power flow describes the nonlinear relationship between the complex voltages u and powers ξ. From Kirchhoff's Current Law, Ohm's Law, and the complex power equation, we get the equation that defines the AC power flow manifold,

$$\xi = \text{diag}(u)\overline{Yu}, \tag{5.2}$$

where Y is the complex admittance matrix, u is the complex representation of (v, θ), and ξ is the complex representation of (p, q):

$$u = v\cos(\theta) + jv\sin(\theta),$$
$$\xi = p + jq.$$

As demonstrated in [93], defining x as the grid state for an n-node network[3],

$$x = (v, \theta, p, q) \in \mathbb{R}^{4n},$$

we can rewrite the power flow manifold in implicit form using

$$F(x) = \langle \mathrm{diag}(u)\overline{Y}u - \xi \rangle = 0_{2n}, \tag{5.3}$$

where $F : \mathbb{R}^{4n} \to \mathbb{R}^{2n}$ constrains the manifold in $\mathbb{R}^{4n}$ such that the power flow manifold is in $\mathbb{R}^{2n}$/there are $2n$ free variables. The manifold is defined as

$$\mathcal{M} := \{x \mid F(x) = 0_{2n}\}. \tag{5.4}$$

5.3   The AC Power Flow Manifold Tangent Plane

The tangent plane at x^* is the best local linear approximation of $\mathcal{M}$ at x^*, and is defined by the normal vector space to $\mathcal{M}$ at x^*. Since $\mathcal{M}$ is a $2n$-dimensional submanifold of $4n$, the normal vector space will be $2n$-dimensional and the tangent plane will be the orthogonal $2n$-dimensional vector space. Lemma 1 in [93] states that $\mathcal{M}$ is a $2n$-dimensional regular submanifold of $\mathbb{R}^{4n}$, thus the $2n$-dimensional tangent plane can be determined at every point of $\mathcal{M}$.

The normal vector space is given by the row space of the Jacobian of F at x^*. The tangent plane is the vector space that is orthogonal to the normal vector space,

$$\mathcal{T}|_{x^*} := \left\{ x \mid \left. \frac{\partial F(x)}{\partial x} \right|_{x^*} \delta x = 0_{2n} \right\}.$$

Appendix E.2 gives a two-dimensional example of a tangent plane to provide some intuition for implicit functions and tangent planes.

We need the Jacobian of F with respect to x to find $\mathcal{T}|_{x^*}$ for the power flow manifold. (5.3), however, is defined in terms of the complex representations of the variables in x. One option for differentiating F is to write (5.3) in real coordinates using the $\langle \cdot \rangle$ complex transformation,

$$0 = \langle \mathrm{diag}(u)\overline{Y}u - \xi \rangle$$

$$0 = \begin{bmatrix} \mathrm{diag}(\mathrm{Re}\,u) & -\mathrm{diag}(\mathrm{Im}\,u) \\ \mathrm{diag}(\mathrm{Im}\,u) & \mathrm{diag}(\mathrm{Re}\,u) \end{bmatrix} N \begin{bmatrix} \mathrm{Re}\,Y & -\mathrm{Im}\,Y \\ \mathrm{Im}\,Y & \mathrm{Re}\,Y) \end{bmatrix} \begin{bmatrix} \mathrm{Re}\,u \\ \mathrm{Im}\,u \end{bmatrix} - \begin{bmatrix} p \\ q \end{bmatrix}. \tag{5.5}$$

---

[3]We use the term "node" rather than "bus" because the power flow manifold equations work for three-phase networks as well as single-phase networks. "node" is intentionally ambiguous—for single-phase networks (or single-phase equivalent networks), each bus on the network is a node. For multiple-phase networks that are not analyzed using a single-phase equivalent network, each phase at each bus is a node. That is, if an unbalanced three-phase network has k buses and three phases at every bus, there will be $n = 3k$ nodes.

Rather than differentiating (5.5), we find that there is a set of equations that is easier to differentiate. Adding the real and imaginary portions of the bus current injection variable i to the state vector $x \in \mathbb{R}^{4n}$ we get the new real-valued state variable z,

$$z = (v, \theta, \operatorname{Re} i, \operatorname{Im} i, p, q) \in \mathbb{R}^{6n}.$$

With z, the power flow manifold is determined by the Ohm's Law equations and the power injection equations:

$$0_{4n} = F(z) = \begin{bmatrix} F^{\text{Ohm's}}(z) \\ F^{\text{power}}(z) \end{bmatrix} = \begin{bmatrix} \langle Yu - i \rangle \\ \langle \operatorname{diag}(u)\bar{i} - \xi \rangle \end{bmatrix}.$$

The same power flow manifold manifold $\mathcal{M}$ can be defined by

$$\mathcal{M} := \left\{ z \mid F(z) = 0_{4n} \right\},$$

which is now a submanifold of $\mathbb{R}^{6n}$. In the $6n$-dimensional z space, the normal vector space at z^* defined by the rows of the Jacobian of $F(z)$ is $4n$-dimensional. Thus, the tangent plane defined as the space orthogonal to the normal vector space,

$$\mathcal{T}|_{z^*} := \left\{ z \mid \left. \frac{\partial F(z)}{\partial z} \right|_{z^*} \delta z = 0_{4n} \right\},$$

is still a $2n$-dimensional plane. Once $\mathcal{T}|_{z^*}$ is calculated for $z \in \mathbb{R}^{6n}$, we eliminate the bus current injections to project $\mathcal{T}|_{z^*}$ onto $x \in \mathbb{R}^{4n}$, providing $\mathcal{T}|_{x^*}$.

The AC Power Flow Manifold Tangent Plane Equation

Lemma 5.3.1. *The power flow manifold tangent plane at $x^* \in \mathbb{R}^{4n}$ is*

$$\mathcal{T}|_{x^*} := \left\{ x \mid A|_{x^*} \delta x = 0_{2n} \right\}, \tag{5.6}$$

where $A|_{x^*}$, the power flow manifold Jacobian with respect to $x \in \mathbb{R}^{4n}$ at x^*, and $\Gamma|_{x^*}$, the voltage-power sensitivity matrix at x^*, are given by

$$A|_{x^*} = \begin{bmatrix} \Gamma|_{x^*} & -I \end{bmatrix} \in \mathbb{R}^{2n \times 4n}, \tag{5.7}$$

$$\Gamma|_{x^*} = \left(\left(\langle \operatorname{diag}(\overline{Yu^*}) \rangle + \langle \operatorname{diag}(u^*) \rangle N \langle Y \rangle \right) R(u^*) \right) \in \mathbb{R}^{2n \times 2n}, \tag{5.8}$$

where u^* is the complex representation of the (v^*, θ^*) portion of x^*.

Proof. Keeping in mind that $\operatorname{diag}(a)b = \operatorname{diag}(b)a$, the partial derivatives of the manifold are

$$\frac{\partial F^{\text{Ohm's}}}{\partial \upsilon} = \langle Y \rangle R(u^*) \qquad \frac{\partial F^{\text{Ohm's}}}{\partial \iota} = -I \qquad \frac{\partial F^{\text{Ohm's}}}{\partial s} = 0$$

$$\frac{\partial F^{\text{power}}}{\partial v} = \langle \text{diag}(\overline{i^*}) \rangle R(u^*) \qquad \frac{\partial F^{\text{power}}}{\partial \iota} = \langle \text{diag}(u^*) \rangle N \qquad \frac{\partial F^{\text{power}}}{\partial s} = -I,$$

and the Jacobian is

$$\begin{bmatrix} \frac{\partial F^{\text{Ohm's}}}{\partial x} \\ \frac{\partial F^{\text{power}}}{\partial x} \end{bmatrix} = \begin{bmatrix} \langle Y \rangle R(u^*) & -I & 0 \\ \langle \text{diag}(\overline{i^*}) \rangle R(u^*) & \langle \text{diag}(u^*) \rangle N & -I \end{bmatrix}.$$

The tangent plane is given by the right-nullspace, or kernel, of the Jacobian,

$$\begin{bmatrix} \frac{\partial F^{\text{Ohm's}}}{\partial x} \\ \frac{\partial F^{\text{power}}}{\partial x} \end{bmatrix} \Bigg|_{z^*} \delta z = 0_{4n}, \text{ where } \delta z = (z - z^*).$$

Eliminating the bus current injections using

$$\begin{bmatrix} \delta \operatorname{Re} i \\ \delta \operatorname{Im} i \end{bmatrix} \Bigg|_{z^*} = \langle Y \rangle R(u^*) \begin{bmatrix} \delta v \\ \delta \theta \end{bmatrix} \Bigg|_{z^*}$$

from the first $2n$ rows of the Jabobian and substituting $Y u^*$ for i^* gives the equations for $A|_{x^*}$ and $\Gamma|_{x^*}$:

$$A|_{x^*} = \begin{bmatrix} \Gamma|_{x^*} & -I \end{bmatrix} \in \mathbb{R}^{2n \times 4n},$$
$$\Gamma|_{x^*} = \left(\left(\langle \text{diag}(\overline{Y u^*}) \rangle + \langle \text{diag}(u^*) \rangle N \langle Y \rangle \right) R(u^*) \right) \in \mathbb{R}^{2n \times 2n}.$$

<div style="text-align: right">□</div>

$\mathcal{T}|_{x^*}$ is the right-nullspace, or kernel, of $A|_{x^*}$. Lemma 5.3.2 states that the right-nullspace of A is orthogonal to the vector space of the rows of A.

Lemma 5.3.2. $\text{image}(A^T) \perp \ker(A)$

Proof. Assume $A^T c \not\perp x$ for some vector c and some $x \in \ker(A)$.

$$0 \neq (A^T c)^T x$$
$$0 \neq c^T A x$$
$$0 \neq 0$$

Thus we have derived a contradiction and the assumption must be wrong.   □

Two-Bus DC Demonstration of the Power Flow Manifold

In this section we derive and plot the power flow manifold for a two-bus DC network. While almost all power systems are AC not DC, two-bus DC networks are useful for developing intuition because they have only four dimensions. Thus, this two-bus DC demonstration, as well as the two-bus DC demonstrations in Sections 6.1 and 7.2, are intended to help

develop understanding. AC, not DC, networks are the primary application of NPFL for power systems analysis.

The DC voltages, currents, and powers are

$$\mathbf{v} \in \mathbb{R}^n,\ \ell \in \mathbb{R}^n,\ \text{and } \mathbf{p} \in \mathbb{R}^n.$$

We define the following DC states

$$\mathbf{x} = (\mathbf{v}, \mathbf{p}) \in \mathbb{R}^{2n},\ \text{and } \mathbf{z} = (\mathbf{v}, \ell, \mathbf{p}) \in \mathbb{R}^{3n}$$

The DC power flow manifold is defined as

$$\mathcal{M} = \left\{ \mathbf{x} \mid \mathbf{F}(\mathbf{x}) = 0_n \right\},\ \text{or}$$

$$\mathcal{M} = \left\{ \mathbf{z} \mid \begin{bmatrix} \mathbf{F}^{\text{Ohm's}} \\ \mathbf{F}^{\text{power}} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}\mathbf{v} - \ell \\ \text{diag}(\mathbf{v})\ell - \mathbf{p} \end{bmatrix} = 0_{2n} \right\}.$$

The partial derivatives give the linearization of the power flow manifold:

$$\begin{bmatrix} \frac{\partial \mathbf{F}^{\text{Ohm's}}}{\partial \mathbf{z}} \\ \frac{\partial \mathbf{F}^{\text{power}}}{\partial \mathbf{z}} \end{bmatrix}\Bigg|_{\mathbf{z}^*} = \begin{bmatrix} \mathbf{Y} & -I & 0 \\ \text{diag}(\ell^*) & \text{diag}(\mathbf{v}^*) & -I \end{bmatrix},$$

where $\mathbf{Y}$ is the real-valued admittance matrix. Eliminating the current injection equations and ℓ and using $\delta\ell^* = \mathbf{Y}\delta\mathbf{v}^*$ gives the expression for the tangent plane to the power flow manifold in $\mathbf{x}$-space ($\mathbb{R}^{2n}$ for DC circuits),

$$\mathcal{T}|_{\mathbf{x}^*} := \left\{ \mathbf{x} \mid A|_{\mathbf{x}^*}\, \delta\mathbf{x} = 0_n \right\}, \tag{5.9}$$

where

$$A|_{\mathbf{x}^*} = \begin{bmatrix} \Gamma|_{\mathbf{x}^*} & -I \end{bmatrix} \in \mathbb{R}^{n \times 2n},$$
$$\Gamma|_{\mathbf{x}^*} = \left(\text{diag}(\mathbf{Y}\mathbf{v}^*) + \text{diag}(\mathbf{v}^*)\mathbf{Y} \right) \in \mathbb{R}^{n \times n}.$$

Fig. 5.1 plots the power flow manifold and the tangent plane for a two-bus DC network in ($\delta\mathbf{v}_1$, $\delta\mathbf{p}_1$, $\delta\mathbf{p}_2$) space. The origin corresponds to the point at which $\delta\mathbf{v}_1 = \delta\mathbf{p}_1 = \delta\mathbf{p}_2 = 0$. The power flow manifold $\mathcal{M}$ is a curved surface that touches tangent plane at the origin. The tangent plane is a flat surface that touches the tangent plane only at the origin, and corresponds to the right-nullspace of A. The vector/space A^T is orthogonal to both the power flow manifold and the tangent plane at the origin, That is, the row space of A is perpendicular to the right-nullspace of A, as stated in Lemma 5.3.2. Figure 5.1 demonstrates that $\mathcal{T}|_{\mathbf{x}^*}$ is a good approximation of $\mathcal{M}$ in the vicinity of $\mathbf{x}^*$.

The points x that are on the AC power flow manifold tangent plane defined in (5.6) satisfy the equation

$$A|_{x^*}\, \delta x = 0_{2n}.$$

(a) $-0.5 \leq \delta\mathbf{p} \leq 0.5$

(b) $-0.1 \leq \delta\mathbf{p} \leq 0.1$

(c) $-0.01 \leq \delta\mathbf{p} \leq 0.01$
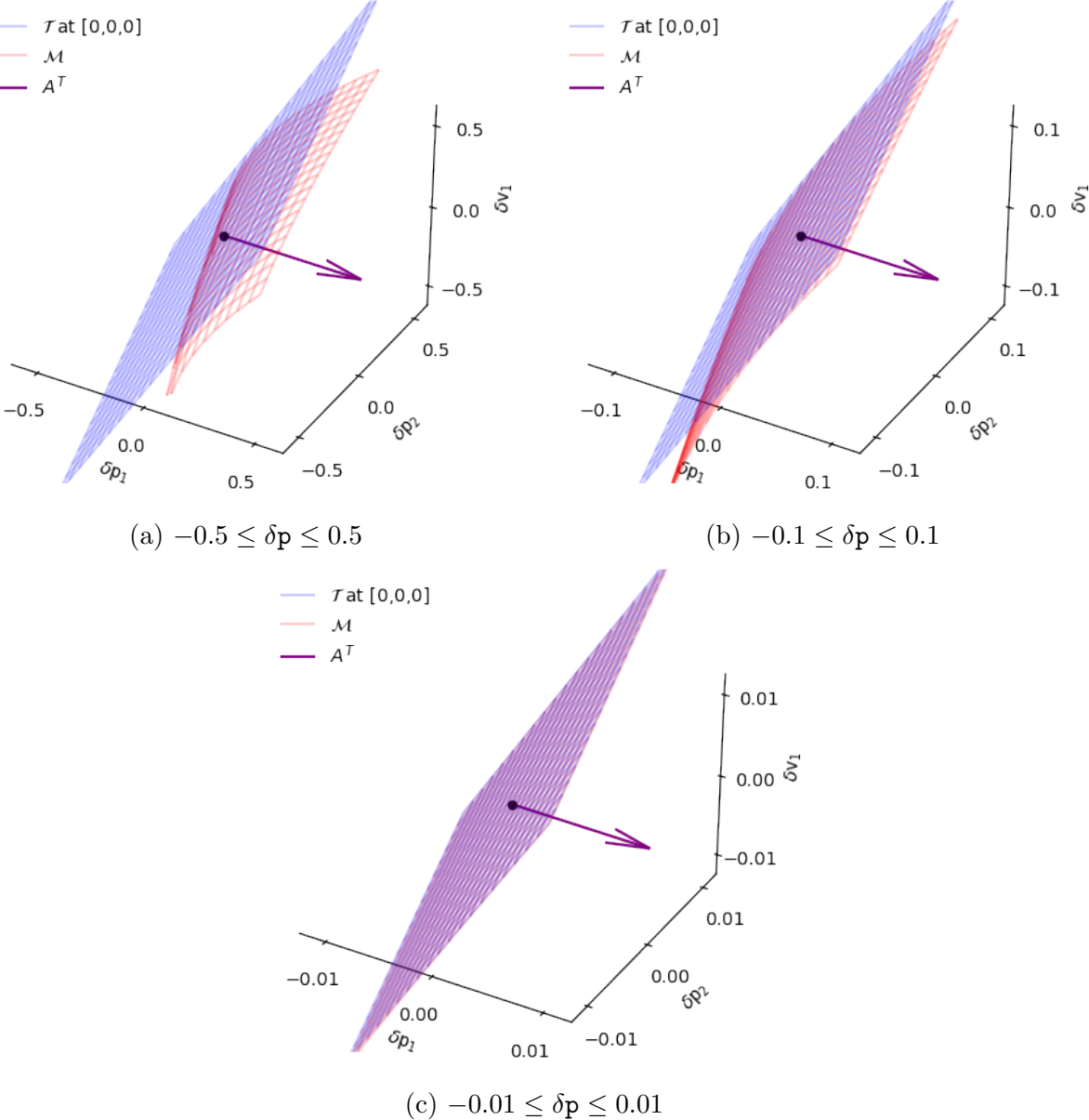
Figure 5.1: The power flow manifold and the tangent plane for a two-bus DC system. $\delta\mathbf{v}_2$ is not plotted because it would require four dimensions.

Rearranging this equation and dropping $|_{x^*}$ for notation clarity gives the voltage-power sensitivity equation

$$\begin{bmatrix} \delta p \\ \delta q \end{bmatrix} = \Gamma \begin{bmatrix} \delta v \\ \delta \theta \end{bmatrix}, \tag{5.10}$$

$$\begin{bmatrix} \delta s \end{bmatrix} = \Gamma \begin{bmatrix} \delta v \end{bmatrix}. \tag{5.11}$$

5.4 The Power-Voltage Sensitivity Matrix

In power systems analysis, we are often interested in how changes in the real and reactive power injections will affect the voltages on the network. We define this relationship as the "power-voltage" sensitivity and use the matrix Ψ,

$$\begin{bmatrix} \delta v \\ \delta \theta \end{bmatrix} = \Psi \begin{bmatrix} \delta p \\ \delta q \end{bmatrix}, \tag{5.12}$$

$$\begin{bmatrix} \delta v \end{bmatrix} = \Psi \begin{bmatrix} \delta s \end{bmatrix}.$$

(7.14) gives the inverse of the power-voltage sensitivity. Due to the structure and non-injectivity of the power flow equations (5.4), in order to get the power-voltage sensitivity, we must invert (7.14). Unfortunately, Γ is not full rank and is therefore not invertible. Γ will always be rank deficient becuase of the ambiguity in the voltage angle reference. Furthermore, if there are not shunt admittances on the network the bus admittance matrix Y will have a zero eigenvalue and Γ will have two zero eigenvalues—one corresponding to the common mode voltage magnitude [98], and one corresponding the angle-reference ambiguity.

The rank deficiency of Γ can be addressed by using the Moore-Penrose pseudoinverse of Γ, $\Gamma^\dagger$, rather than the inverse. However the following two issues arise when $\Gamma^\dagger$ is used as the power-voltage sensitivity in real power system applications.

1. $\Gamma^\dagger$ does not hold the voltage at any buses constant. In power systems analysis there is almost always at least one (approximately) constant voltage bus. For distribution network modeling, the substation acts like a constant voltage bus.

2. The power injections at all n nodes δs do not necessarily obey power balance. In power systems at steady state, power injected into the network must be equal to the power extracted from the network. Furthermore, the (change in) power injections at all n nodes are often not known a priori. For example, on a distribution network with one substation, the change in the substation power injection that accompanies the change in the power injections elsewhere on the network is not known ahead of time.

Chapter 6 introduces a power flow linearization that addresses Issue 1. Chapter 7 builds on the power flow linearization introduced in Chapter 6 to produce a power flow linearization that addresses both Issue 1 and Issue 2.

Chapter 6

Power Flow Linearization for Networks with a Slack Bus

Often, in power systems, the voltage magnitude at (at least) one of the nodes on the network is held constant. For transmission networks, the excitation of some generators are controlled to maintain a constant voltage.[1] For distribution networks with a single substation that is attached to the transmission network, the substation acts like a constant voltage bus.[2]

In addition to the constant voltage magnitude, it is necessary to establish an angle reference for the network. While any node can technically be chosen as the angle reference, choosing a/the node with constant voltage is the logical choice. The angle of the bus that is used as the angle reference is set to zero at all times, and is therefore also constant.

In the load flow problem the voltage phasors are solved for as a function of the power injections on the network. The power injections are not given for every bus on the network—load flow requires that one bus is a "slack bus," for which the power injection is not specified.[3] The slack bus has two attributes in load flow—first, the slack bus is constant voltage regardless of what the generation/load profile of the system is, and second, the slack bus adjusts its power injection so that power balance (power in = power out) is maintained for the network.[4]

Using a slack bus to solve load flow is particularly appropriate for distribution networks with a single substation because the substation provides both the functionalities of the slack

---

[1] In load flow, these generators are the "PV" buses.

[2] The substation can be approximated as a voltage source or "infinite bus" because the impedance of the transmission network seen from the distribution network is very small compared with the distribution network impedances. The impedance of the transmission network is small because the transmission lines have less impedance than the distribution lines and because of the substation transformer—the per unit base impedance on the transmission side of the transformer will be greater by a factor of the transformer's turns ratio squared.

[3] The term "slack bus" goes back to the 1950s [99].

[4] It is possible to separate the slack bus's constant-voltage and power-balance attributes, and it is possible to divide the power-balance responsibility among multiple nodes on the network [100]–[102]. Section 7.3 provides a formulation that can support multiple buses participating in power-balance, rather than just a single constant-voltage bus.

bus—the substation acts like a constant voltage bus and injects enough power to match the net power demand of the rest of the distribution network. Thus, a power flow linearization that takes into account both the constant-voltage and power-balance attributes of a slack bus is useful for distribution network modeling. Most power flow linearizations, such as DC Power Flow [101], [103], implicitly assume that the network has a slack bus [104].

6.1 The Power Flow Manifold Tangent Plane with a Slack Bus

Without loss of generality, let us assume that bus 1 is the slack bus. Constant slack bus voltages correspond to setting the change in the voltage at that bus equal to zero $\delta v_1 = 0$ and $\delta \theta_1 = 0$. We define the column-elimination matrix for a network with a slack bus χ_1 as

$$\chi_1 = \begin{bmatrix} 0 & 0 \\ I_{n-1} & 0 \\ 0 & 0 \\ 0 & I_{n-1} \end{bmatrix} \in \mathbb{R}^{2n \times (2n-2)},$$

We name the δv vector with the slack bus entries removed

$$\chi_1^T \delta v = \chi_1^T \delta v,$$

and the state vector x with the slack bus voltage entries removed

$$\chi_1^T x = \begin{bmatrix} \chi_1 & 0 \\ 0 & I_{2n} \end{bmatrix}^T x.$$

We define the "power flow equations with a slack bus" $F_{\chi_1} : \mathbb{R}^{4n-2} \to \mathbb{R}^{2n-2}$ to be the same as F in (5.3) but with the first and $(n+1)$'th entries of x, corresponding to the slack bus voltage magnitude and angle, removed from x and set to 1 and 0 in (5.3), respectively.[5] With a slack bus, the power flow manifold is a $(2n-2)$-dimensional subspace of $\mathbb{R}^{4n-2}$ that satisfies F_{χ_1},[6]

$$\mathcal{M}_{\chi_1} := \left\{ \chi_1^T x \mid F_{\chi_1}(\chi_1^T x) = 0_{2n} \right\}. \tag{6.1}$$

The dimension of a manifold corresponds to the degrees of freedom for the function that the manifold implicitly defines.

---

[5]We assume that the slack bus voltage is 1 per unit (p.u.).

[6]We have dropped the $|_{x^*}$ notation, which indicates the point at which the tangent plane is calculated and touches the power flow manifold, for notational clarity.

Lemma 6.1.1. *The power flow manifold $\mathcal{M}_{\chi_1}$ is given by the intersection of $\mathcal{M}$ and the*

$$x \mid \begin{bmatrix} 1 & 0_{n-1} & 0 & 0_{3n-1} \\ 0 & 0_{n-1} & 1 & 0_{3n-1} \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{6.2}$$

plane in $\mathbb{R}^{4n}$ when the slack bus voltage magnitude is 1.

Proof. By definition, $\mathcal{M}_{\chi_1}$ is $\mathcal{M}$ with the first and $(n+1)$'th entries of x set to 1 and 0, respectively. Thus, $\mathcal{M}_{\chi_1}$ is given by the intersection of $\mathcal{M}$ and the aforementioned plane in $\mathbb{R}^{4n}$. $\qquad \square$

Next we define the tangent plane for a power system with a constant voltage node. This tangent plane is useful because it provides intuition for what the voltage-power and power-voltage sensitivity mean when there is a constant-voltage node on the network.

Theorem 6.1.2. *The power flow manifold tangent plane at $x^*_{\chi_1} \in \mathbb{R}^{4n-2}$ is*

$$\mathcal{T}_{\chi_1} := \left\{ \chi_1^T x \mid A_{\chi_1} \chi_1^T \delta x = 0_{2n} \right\}, \tag{6.3}$$

where A_{χ_1}, the Jacobian for the power flow manifold for a network with a constant-voltage node, is

$$A_{\chi_1} = \begin{bmatrix} \Gamma_{\chi_1} & -I \end{bmatrix}, \tag{6.4}$$

and Γ_{χ_1}, the voltage-power sensitivity for a network with a constant voltage node, is

$$\Gamma_{\chi_1} = \Gamma \chi_1. \tag{6.5}$$

Proof. Lemma 5.3.1 defines the power flow manifold for $x \in \mathbb{R}^{4n}$ with (5.6). The power flow manifold tangent plane $\mathcal{T}_{\chi_1}$ at $x^*_{\chi_1} \in \mathbb{R}^{4n-2}$ is given by the intersection of $\mathcal{T}$ and the plane defined in (6.2). Thus, $\mathcal{T}_{\chi_1}$ is the projection onto $\mathbb{R}^{4n-2}$ of the set of vectors in $\mathbb{R}^{4n}$ that are orthogonal to both the rows of A and the rows of the matrix that define the plane in (6.2),

$$\mathcal{T}_{\chi_1} = \left\{ \begin{bmatrix} \chi_1 & 0 \\ 0 & I_{2n} \end{bmatrix}^T x \mid A\delta x = 0_{2n} \text{ and } \begin{bmatrix} 1 & 0_{n-1} & 0 & 0_{3n-1} \\ 0 & 0_{n-1} & 1 & 0_{3n-1} \end{bmatrix} \delta x = 0_2 \right\}. \tag{6.6}$$

The set of vectors in $\mathbb{R}^{4n}$ that are orthogonal to the rows of the matrix in (6.2) is

$$image\left(\begin{bmatrix} \chi_1 & 0 \\ 0 & I_{2n} \end{bmatrix} \right).$$

Thus the second constraint in (6.6) constrains the permissible vectors δx to be the set of vectors with 0s in the 1st and $(n+1)^{\text{th}}$ entries.

The intersection of the set of vectors δx that are in $image\left(\begin{bmatrix} \chi_1 & 0 \\ 0 & I_{2n} \end{bmatrix}\right)$ and satisfy $A\delta x = 0$ is the set of vectors with 0s in the 1st and $(n+1)^{\text{th}}$ entries that satisfy satisfy $A\delta x = 0$.

Projecting this set onto $\mathbb{R}^{4n-2}$ gives the set of vectors that satisfy $A_{\chi_1} {}_{\chi_1^T}\delta x = 0$, where

$$A_{\chi_1} = A \begin{bmatrix} \chi_1 & 0 \\ 0 & I_{2n} \end{bmatrix}.$$

$\square$

The row space of A_{χ_1}, or $image(A_{\chi_1}^T)$, is the $2n$-dimensional vector space that is normal to $\mathcal{M}_{\chi_1}$ and $\mathcal{T}_{\chi_1}$ at ${}_{\chi_1^T}x^*$. The right-nullspace of A_{χ_1} is the power flow manifold tangent plane when there is a constant voltage bus (Lemma 5.3.2). $image(\Gamma_{\chi_1}) \in \mathbb{R}^{2n}$ is the vector space of permissible power flow injections according to the power flow manifold tangent plane approximation. The right-nullspace of A_{χ_1} and $image(\Gamma_{\chi_1})$ are related, but exist in different dimensional spaces—$ker(A_{\chi_1}) \in \mathbb{R}^{4n-2}$ and $image(\Gamma_{\chi_1}) \in \mathbb{R}^{2n}$. If the right-nullspace of A_{χ_1} is projected onto $\delta s \in \mathbb{R}^{2n}$ space, it is equivalent to $image(\Gamma_{\chi_1})$.

Two-Bus DC Demonstration of the Power Flow Manifold Tangent Plane with a Slack Bus

In this section we demonstrate the power flow manifold tangent plane for a two-bus DC network in order to provide intuition. As stated for the two-bus DC demonstration in Section 5.3, two-bus DC networks are useful for developing intuition because $\mathbf{x}$ is four-dimensional. Thus, by removing one dimension of $\mathbf{x}$, the manifolds and tangent planes can be plotted in three dimensions.

The two-bus DC demonstration in Section 5.3 introduced the tangent plane of the power flow manifold for a two-bus DC network (5.9) without a slack bus. For a two-bus DC network with a slack bus at bus 1, the tangent plane of the power flow manifold is given by

$$\mathcal{T}_{\chi_1} := \left\{ {}_{\chi_1^T}\mathbf{x} \mid \mathbf{A}_{\chi_1} {}_{\chi_1^T}\delta\mathbf{x} = 0_n \right\}, \quad \text{where}$$
$$\mathbf{A}_{\chi_1} = \begin{bmatrix} \Gamma_{\chi_1} & -I \end{bmatrix} \in \mathbb{R}^{n \times (2n-1)}$$

and Γ_{χ_1} is Γ with the first column removed.

Fig. 6.1 demonstrates both the tangent planes for the DC power flow manifold in $(\delta\mathbf{p}_1, \delta\mathbf{p}_2, \delta\mathbf{v}_1)$ space. The standard tangent plane (no slack bus) is a plane in $(\delta\mathbf{p}_1, \delta\mathbf{p}_2, \delta\mathbf{v}_1)$ space corresponding to the right-nullspace of $\mathbf{A}$. The tangent plane to the DC power flow manifold with bus 1 as a slack bus is the line that is the right-nullspace of $\mathbf{A}_{\chi_1}$. This line is given by the intersection of the standard power flow linearization and the $\delta\mathbf{v}_1 = 0$ plane.
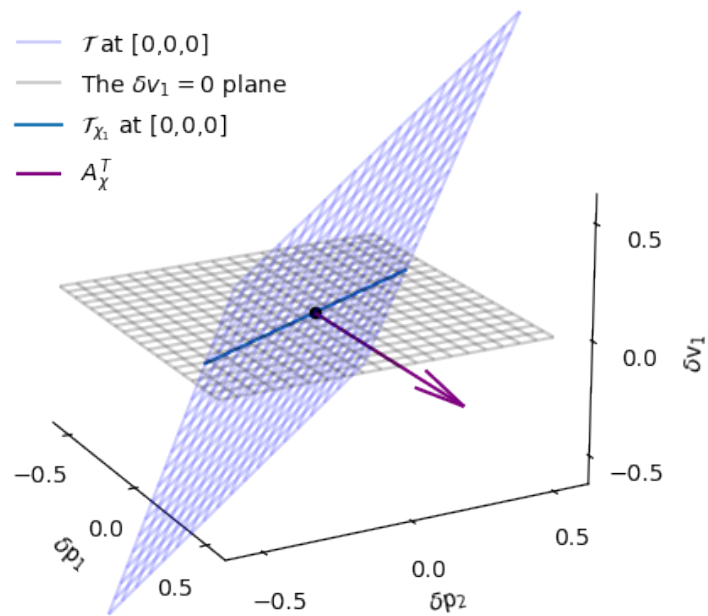
Figure 6.1: The intersection of the two-bus DC power flow manifold tangent plane and the $\delta \mathbf{v}_1 = 0$ plane. The intersection is $\mathtt{ker}(A_\chi)$, the power flow manifold tangent plane when bus 1 is the slack bus.
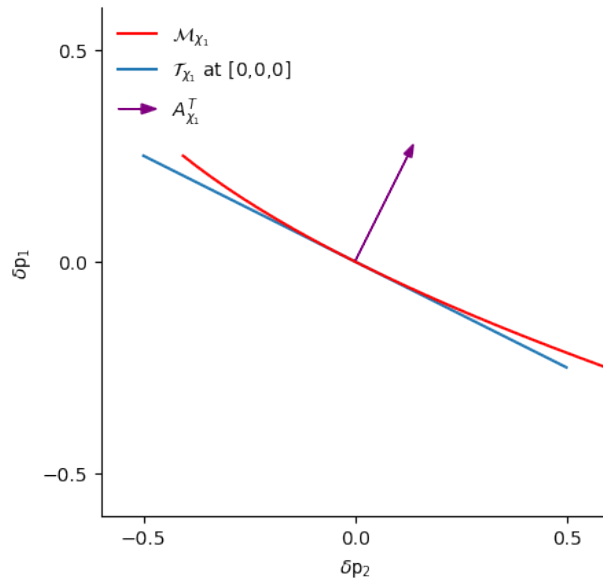
Figure 6.2: The power flow manifold and the manifold tangent plane in two-dimensional $(\delta \mathbf{p}_1, \delta \mathbf{p}_2)$ space for the two-bus DC network with a slack bus.
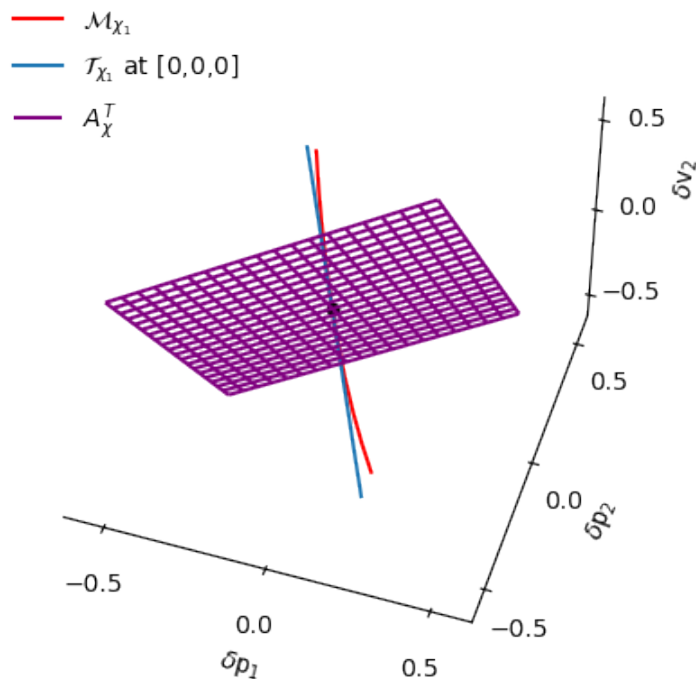
Figure 6.3: The two-bus DC power flow manifold and its tangent plane when bus 1 is a slack bus. δv_2 is plotted on the vertical axis.

Fig. 6.2 plots the tangent plane for the slack-bus network in Fig. 6.1 in $(\delta p_1, \delta p_2)$ space. The power flow manifold tangent plane of a network with a slack bus is given by the right-nullspace or kernel of A_{χ_1}. In $(\delta p_1, \delta p_2)$ space, the right-nullspace of $\mathbf{A}_{\chi_1}$ is equivalent to $image(\Gamma_{\chi_1})$.

Fig. 6.3 plots the two-bus DC power flow manifold when bus 1 is a slack bus in $(\delta p_1, \delta p_2, \delta v_2)$ space. Since $\delta v_1 = 0$, this is the full power flow manifold for the two-bus DC system. Note, $\mathbf{A}_{\chi_1}^{\mathsf{T}}$ is a plane and both the power flow manifold and the tangent plane are lines in $(\delta p_1, \delta p_2, \delta v_2)$ space. Thus, setting δp_1, δp_2, or δv_2 determines the remaining two variables.

6.2 Naive Power-Voltage Sensitivity with a Slack Bus

We would like a power-voltage sensitivity expression/linearization for $_{\chi_1^T}\delta v$ in terms of δs. However, Γ_{χ_1} is not a square matrix and therefore it is not invertible. One way to get an expression for $_{\chi_1^T}\delta v$ in terms of δs is by taking the Moore-Penrose pseudoinverse of Γ_{χ_1},

giving[7]

$$\chi_1^T \delta v = \Gamma_{\chi_1}^\dagger ds. \tag{6.7}$$

The problem with using (6.7) as a power-voltage sensitivity is that it requires $\delta s \in \mathbb{R}^{2n}$ as an input. That is, (6.7) requires the power injections at all n buses as the input. For networks with a slack bus, the power injection at the slack bus is not known a priori. (6.7) cannot be used without a guess for the slack bus power injection. Chapter 7 introduces Nullspace-based Power Flow, which does not require a guess for the slack bus power injection.

---

[7]We define $\Gamma_{\chi_1}^\dagger$ as the pseudoinverse of Γ_{χ_1}. That is, the pseudoinverse is taken after Γ is right-multiplied by χ_1.

Chapter 7

Nullspace-based Power Flow Linearization

Traditionally, the goal of power flow linearization has been to produce a matrix that approximates the relationship between changes in the power injections (inputs) and changes in the voltages on the system (outputs).[1] In this chapter we introduce a modified goal for power flow linearization—produce a matrix that takes a vector of changes in power injections as the input and produces both the changes in the voltages on the system *and* an adjusted vector of power injections that satisfies an approximation of power balance for the system.[2]

This chapter introduces Nullspace-based Power Flow Linearization (NPFL), a new power flow linearization that addresses the modified power flow linearization goal above and incorporates the knowledge that the voltage at at least one node is held constant. For a network with a single slack bus, NPFL incorporates the additional knowledge that the slack bus will adjust its power injection so that the total power into the network is equal to the total power out. For a network with a single slack bus, NPFL answers the following two questions.

1. How do the changes in the power injections at the non-slack buses change the slack bus power injections, given that the slack bus is entirely responsible for maintaining power balance?

2. How do changes in the power injections change the voltages at the non-slack bus nodes, given that the slack bus is entirely responsible for maintaining power balance?

For distribution networks with a single substation, NPFL provides both an accurate power-voltage sensitivity matrix and the sensitivities between each power injection (e.g., each distributed energy resource) and the substation power injection. The substation power

---

[1]Reminder: throughout this dissertation, "power injections" refer to both generators and loads. For a load, the power injection is negative.

[2]"Power balance" is defined as asserting that the power injected into the network is equal to the power that leaves the system, either as power to loads or as losses.

injection sensitivity includes the voltage-dependencies of the shunt admittances and an approximation of the network power losses. The substation power injection sensitivity is useful if, for example, the load on a given distribution network has overloaded the substation transformer and the power injections of the DERs on the network must be adjusted to avoid a power outage. For distribution networks with a *multiple* substations, NPFL provides the sensitivities between the power injections and each substation's power injection. To the author's knowledge, an explicit equation for the substation/slack bus injection sensitivities were not previously available.

While this dissertation focuses on distribution network applications of NPFL, NPFL can be applied to transmission networks as well. Section 7.5 provides another application of NPFL—improving locational marginal price (LMP) accuracy—that will be explored in future research.

It is also important to state the limitations of NPFL. The accuracy of NPFL is entirely based on the accuracy of the bus admittance matrix used to construct Γ. Thus, any inaccuracies in the bus admittance matrix will manifest as inaccuracies in the NPFL for a given network. Furthermore, any network properties that cannot be included in the bus admittance matrix, such as loads with nonlinear voltage dependencies or transformer nonlinearities, cannot be captured by the NPFL network model. NPFL is also dependent upon the accuracy of the state estimate at which the power flow manifold tangent plane is calculated.

Section 7.1 describes why NPFL is not obvious, by describing why simply pseudoinverting Γ_{χ_1} does not answer Questions 1 and 2 above. Section 7.2 gives the NPFL for a network with a single slack bus and Section 7.3 gives the NPFL for a general network. Section 7.4 gives an overview of the prevalent power flow linearizations, and their relationship to NPFL. Section 7.5 gives a sample application of NPFL—determining the aggregate loss sensitivity for a network at a given operating point.

7.1 Motivation for Nullspace-based Power Flow Linearization

Chapter 6 concluded by giving the naive power-voltage sensitivity, $\Gamma_{\chi_1}^\dagger$ in (6.7). $\Gamma_{\chi_1}^\dagger$ takes δs, the vector of the real and reactive power injections at all of the nodes on the network (the input), and produces $\chi_1^T x$, the vector of the voltage magnitudes and angles at the non-slack bus nodes (the output).

We define

$$\delta s_1 := \text{ the slack bus's power injection.}$$

Let us also define the permutation matrix Ξ_1 as

$$\Xi_1 := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & I_{n-1} & 0 & 0 \\ 0 & 0 & 0 & I_{n-1} \end{bmatrix}.$$

Ξ_1 moves the $(n+1)^{\text{th}}$ row to the second row and moves all of the other rows down. We define Ξ_1-permuted Γ_{χ_1} matrix and δs vector as

$$_{\Xi_1}\Gamma_{\chi_1} = \Xi_1 \Gamma_{\chi_1} \text{ and}$$
$$_{\Xi_1}\delta s = \Xi_1 ds.$$

$_{\Xi_1}\delta s$ can be partitioned into δs_1 and $_{\chi_1^T}\delta s$, the power injections for the rest of the buses on the network,

$$_{\Xi_1}\delta s = \begin{bmatrix} \delta s_1 \\ _{\chi_1^T}\delta s \end{bmatrix}, \text{ where}$$
$$_{\chi_1^T}\delta s = \chi_1^T \delta s.$$

(6.7) requires δs_1 as an input. This δs_1 requirement is problematic because the slack bus's power injection in generally not known a priori. (6.7) can be used if $\delta\tilde{s}_1$, a guess for the slack bus's injection, is provided by an external source. We define $_{\Xi_1}\delta\tilde{s}$ as the $_{\Xi_1}\delta s$ vector with $\delta\tilde{s}_1$ inserted in the first two entries—the entries corresponding to the slack bus. $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger \in \mathbb{R}^{(2n-2)\times 2n}$ converts a $_{\Xi_1}\delta\tilde{s} \in \mathbb{R}^{2n}$ vector to a $_{\chi_1^T}\delta v \in \mathbb{R}^{(2n-2)}$ vector,

$$_{\chi_1^T}\delta v = \left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger {}_{\Xi_1}\delta\tilde{s}.$$

Because $_{\Xi_1}\Gamma_{\chi_1}$ has two more rows than columns, $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$ will have a right-nullspace that is at least two-dimensional.

Lemma 7.1.1. $rank(\Gamma_{\chi_1}) = 2n - 2$ *almost surely.*

Proof. From (5.8) and (6.5),

$$\Gamma_{\chi_1} = \left[\left(\langle \text{diag}(\overline{Yu^*}) \rangle + \langle \text{diag}(u^*) \rangle N \langle Y \rangle \right) R(u^*) \right] \chi_1 \in \mathbb{R}^{2n \times (2n-2)}$$
$$= \left[\langle \text{diag}(\overline{Yu^*}) \rangle \chi_1 + \langle \text{diag}(u^*) \rangle N \langle Y \rangle \chi_1 \right] \left[\chi_1^T R(u^*) \chi_1 \right]$$

$rank\left(\langle \text{diag}(\overline{Yu^*}) \rangle \right) = 2n$ almost surely (a.s.).
Therefore $rank\left(\langle \text{diag}(\overline{Yu^*}) \rangle \right) \chi_1 = 2n - 2$ a.s.,
and $\langle \text{diag}(\overline{Yu^*}) \rangle \chi_1 + \langle \text{diag}(u^*) \rangle N \langle Y \rangle \chi_1 = 2n - 2$ a.s..
Also, $rank\left(\chi_1^T R(u^*) \chi_1 \right) = 2n - 2$.
Therefore $rank\left(\left[\langle \text{diag}(\overline{Yu^*}) \rangle \chi_1 + \langle \text{diag}(u^*) \rangle N \langle Y \rangle \chi_1 \right] \left[\chi_1^T R(u^*) \chi_1 \right] \right) = 2n - 2$ a.s.. $\quad\square$

Lemma 7.1.1 states that the right-nullspace of $\Gamma_{\chi_1}^\dagger$ will be two-dimensional in practice. This is important for NPFL because NPFL requires a two-dimensional right-nullsapce for $\Gamma_{\chi_1}^\dagger$. Since the right-nullspace of $\Gamma_{\chi_1}^\dagger$ is never more than two dimensions in practice, there is no ambiguity in selecting the two-dimensional right-nullsapce for NPFL.

The right-nullspace of $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$ maps a portion of the $_{\Xi_1}\delta\tilde{s}$ vector to zero. Taking the singular value decomposition of $_{\Xi_1}\Gamma_{\chi_1}$ gives

$$_{\Xi_1}\Gamma_{\chi_1} = {}_{\Xi_1}\mathsf{U}\Sigma\mathsf{V}^{\mathrm{T}} = \left[_{\Xi_1}\mathsf{U}_{+}, {}_{\Xi_1}\mathsf{U}_0\right]\Sigma\mathsf{V}^{\mathrm{T}}, \tag{7.1}$$

where $_{\Xi_1}\mathsf{U}_0 \in \mathbb{R}^{2n\times 2}$ is the matrix whose columns span the left-nullspace of $_{\Xi_1}\Gamma_{\chi_1}$ and the right-nullspace of $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$.

Remark 7.1.2.

$$\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger {}_{\Xi_1}\delta\tilde{s} = \left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger \left[I \quad \left[-_{\Xi_1}\mathsf{U}_0\right]\left[_{\Xi_1}\mathsf{U}_0\right]^T\right] {}_{\Xi_1}\delta\tilde{s}.$$

Remark 7.1.2 states that $_{\Xi_1}\delta\tilde{s}$ produces the same $_{\chi_1^T}\delta\upsilon$ as $_{\Xi_1}\delta\tilde{s}$ projected onto $image(_{\Xi_1}\Gamma_{\chi_1})$ in the direction $_{\Xi_1}\mathsf{U}_0$. $_{\Xi_1}\mathsf{U}_0$ is orthogonal to $image(_{\Xi_1}\Gamma_{\chi_1})$ and not parallel to the slack bus injection axis. Therefore, the projection of $_{\Xi_1}\delta\tilde{s}$ onto $image(_{\Xi_1}\Gamma_{\chi_1})$ is dependent upon $_{\Xi_1}\delta\tilde{s}_1$, the slack bus injection guess. This projection statement is demonstrated in two dimensions by the difference between the projections in Fig. 7.1a and Fig. 7.1b.

Projecting $_{\Xi_1}\delta\tilde{s}$ onto $image(_{\Xi_1}\Gamma_{\chi_1})$ changes the non-slack bus entries of $_{\Xi_1}\delta\tilde{s}$ as well as the slack bus entries of $_{\Xi_1}\delta\tilde{s}$. This is also demonstrated by Fig. 7.1. Changing both the slack bus and non-slack bus power injections contradicts what we have asserted about the system—that the slack bus is entirely responsible for maintaining power balance on the system. Therefore, $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$ does not answer Questions 1 or 2.

7.2 Nullspace-based Power Flow Linearization for a Network with a Slack Bus

The introduction of this chapter introduced Questions 1 and 2 for a network with a slack bus. The Nullspace Power Balance in Section 7.2 answers Question 1 and the power-voltage sensitivity in Section 7.2 answers Question 2.

Nullspace Power Balance for a Network with a Slack Bus

As stated at the beginning of this chapter, the modified goal of power flow linearization is to produce both the changes in the voltages on the system *and* an adjusted vector of power injections that satisfies an approximation of power balance. We define

$$\delta\hat{s} := \text{ the adjusted vector of power injections that satisfies}$$
$$\text{an approximation of power balance, and}$$
$$_{\Xi_1}\delta\hat{s} = \Xi_1\delta\hat{s}.$$

(a) $\delta\tilde{p}_1 = 0$ (b) $\delta\tilde{p}_1 = \delta p_2$
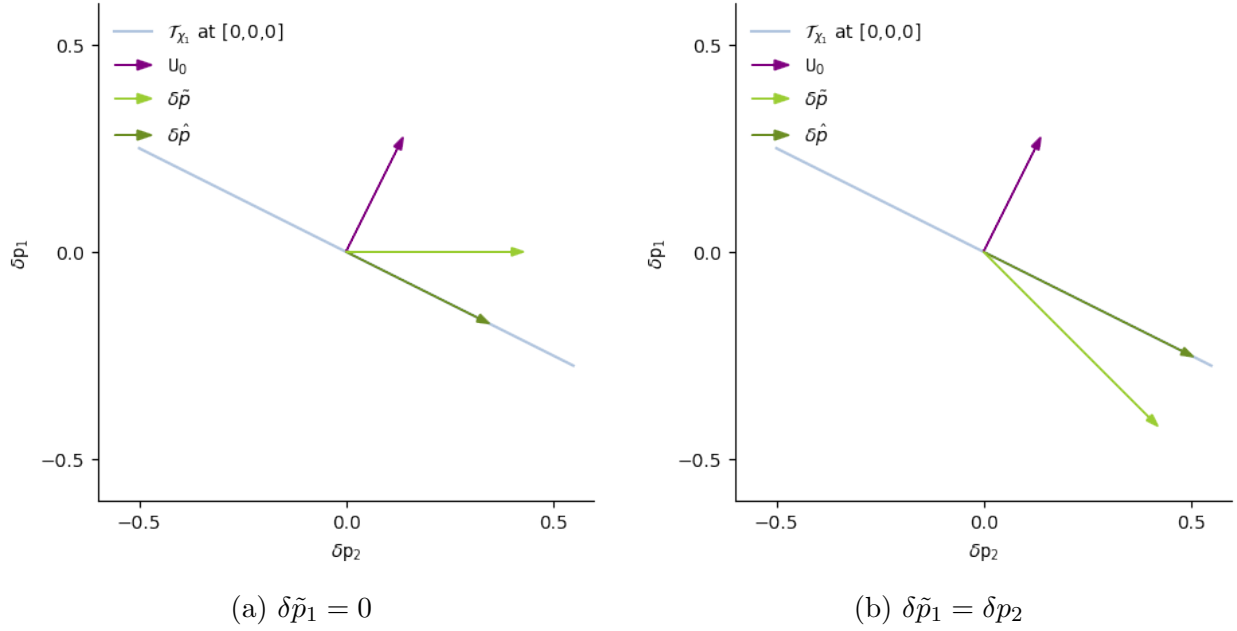
Figure 7.1: Two-dimensional demonstration of the orthogonal projection onto $image(\Gamma_{\chi_1})$ (or $\mathcal{T}_{\chi_1}$) for two different $\delta\tilde{p}_1$ guesses. This plot includes only $\delta\tilde{p}_1$ and $\delta\tilde{p}_2$, however the two-dimension intuition can be applied to higher dimension $\delta\tilde{s}$ vectors.

For a network with a slack bus, we know that power balance is maintained by adjusting only the slack bus injection. Thus, for a network with a slack bus, $_{\Xi_1}\delta\hat{s}$ can be partitioned into $\delta\hat{s}_1$ and $_{\chi_1^T}\delta s$,

$$_{\Xi_1}\delta\hat{s} = \begin{bmatrix} \delta\hat{s}_1 \\ _{\chi_1^T}\delta s \end{bmatrix}.$$

Note, the non-slack bus portion of $_{\Xi_1}\delta\hat{s}$ is $_{\chi_1^T}\delta s$. That is, the non-slack injection in $_{\Xi_1}\delta\hat{s}$ are the given power injections and are not adjusted.

To answer Question 1 posed at the beginning of this chapter, "How do the changes in the power injections at the non-slack buses change the slack bus power injections, given that the slack bus is entirely responsible for maintaining power balance?," we observe that in order for $_{\Xi_1}\delta\hat{s}$ to obey the power-balance rules that are approximated by $\mathcal{T}_{\chi_1}$, $_{\Xi_1}\delta\hat{s}$ must obey the equation

$$_{\Xi_1}\Gamma_{\chi_1}\,_{\chi_1^T}\delta\upsilon = \,_{\Xi_1}\delta\hat{s}. \tag{7.2}$$

Lemma 7.2.1. *For a network with a slack bus,* $_{\Xi_1}\delta\hat{s} = \begin{bmatrix} \delta\hat{s}_1 \\ _{\chi_1^T}\delta s \end{bmatrix}$ *is on* $\mathcal{T}_{\chi_1}$ *when* $\delta\hat{s}_1$ *is set*

using

$$\delta \hat{s}_1 = - \left[{}_{\Xi_1} \mathsf{U}_{0,1} \right]^{-T} \left[{}_{\Xi_1} \mathsf{U}_{0,\chi_1} \right]^{T} {}_{\chi_1^T} \delta s, \tag{7.3}$$

where ${}_{\Xi_1} \mathsf{U}_0 \in \mathbb{R}^{2n \times 2}$ *is the matrix describing the left-nullspace of* ${}_{\Xi_1} \Gamma_{\chi_1}$, ${}_{\Xi_1} \mathsf{U}_{0,1} \in \mathbb{R}^{2 \times 2}$ *is the matrix consisting of the first two rows of* ${}_{\Xi_1} \mathsf{U}_0$, *and* ${}_{\Xi_1} \mathsf{U}_{0,1} \in \mathbb{R}^{(2n-2) \times 2}$ *is matrix consisting of the other rows of* ${}_{\Xi_1} \mathsf{U}_0$. *That is,*

$$_{\Xi_1} \mathsf{U}_0 = \begin{bmatrix} {}_{\Xi_1} \mathsf{U}_{0,1} \\ {}_{\Xi_1} \mathsf{U}_{0,\chi_1} \end{bmatrix}.$$

Proof. In order for ${}_{\Xi_1} \delta \hat{s}$ to obey (7.2), ${}_{\Xi_1} \delta \hat{s}$ must be orthogonal to the left-nullspace of Γ_{χ_1},

$$\left[{}_{\Xi_1} \mathsf{U}_0 \right]^{T} {}_{\Xi_1} \delta s = \begin{bmatrix} {}_{\Xi_1} \mathsf{U}_{0,1} \\ {}_{\Xi_1} \mathsf{U}_{0,\chi_1} \end{bmatrix}^{T} \begin{bmatrix} \delta \hat{s}_1 \\ {}_{\chi_1^T} \delta s \end{bmatrix} = 0.$$

Solving for $\delta \hat{s}_1$ gives

$$\delta \hat{s}_1 = - \left[{}_{\Xi_1} \mathsf{U}_{0,1} \right]^{-T} \left[{}_{\Xi_1} \mathsf{U}_{0,\chi_1} \right]^{T} {}_{\chi_1^T} \delta s.$$

$\square$

Lemma 7.2.1 describes how to use the power flow manifold tangent plane to give an estimate for how changes in the non-slack bus power injections change the slack bus power injection. Fig. 7.2 demonstrates $\delta \hat{s}_1$ estimation in two dimensions. Since β_{slack} points in the direction of the δp_1 axis, the guess provided for δp_1, $\delta \hat{p}_1$, is irrelevant.

(7.3) uses $\mathcal{T}_{\chi_1}$ to answer Question 1 that was posed at the beginning of this chapter.[3] $\mathcal{T}_{\chi_1}$ is the best local linear approximation of the combination of Ohm's law, the complex power equation, and current balance at the given operating point x^*. (7.3) approximates the change in the slack bus power injection as a linear function of the changes in the power injections at all of the other buses, including the effects of the shunt admittances and an approximation of the network power losses.[4]

Power-Voltage Sensitivity for a Network with a Slack Bus

To answer Question 2 posed at the beginning of this chapter, "How do changes in the power injections change the voltages at the non-slack bus nodes, given that the slack bus is entirely responsible for maintaining power balance?," we must find the power-voltage

---

[3](7.3) determines $\delta \hat{s}_1$ from ${}_{\chi_1^T} \delta s$ by asserting that ${}_{\Xi_1} \delta \hat{s}$ is in $image({}_{\Xi_1} \Gamma_{\chi_1})$. Since ${}_{\Xi_1} \delta \hat{s}$ is in $image({}_{\Xi_1} \Gamma_{\chi_1})$, ${}_{\Xi_1} \delta \hat{s}$ is mapped to by some voltage vector ${}_{\chi_1^T} \delta v$.

[4]Section 7.5.
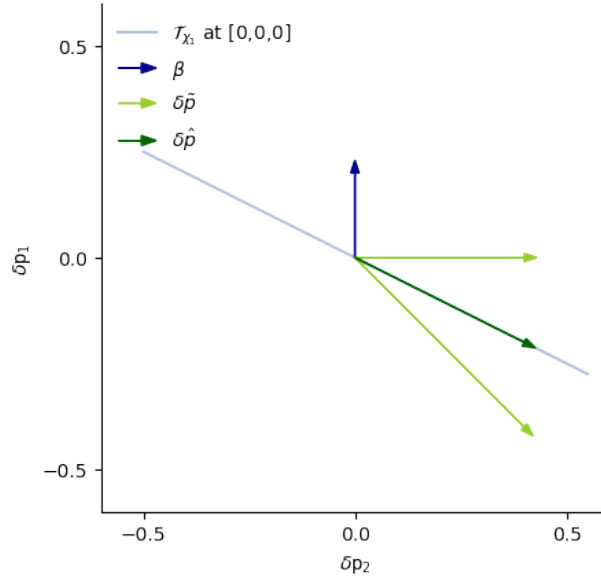
Figure 7.2: Two-dimensional demonstration of the projection onto $image(\Gamma_\chi)$ (or $\mathcal{T}_{\chi_1}$) in the β_{slack} direction from two different $\delta\hat{p}$ guesses.

sensitivity matrix which maps $_{\Xi_1}\delta\tilde{s}$ to $_{\chi_1^T}\delta\upsilon$ such that the slack bus is entirely responsible for maintaining power balance.

Remark 7.1.2 demonstrates that $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$ provides a $_{\chi_1^T}\delta\upsilon$ which shifts all of the entries of $_{\Xi_1}\delta\tilde{s}$, not just the slack bus entries. $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$ does not meet Question 2's criteria.

Theorem 7.2.2. *For a network with a single slack bus,*

$$_{\chi_1^T}\delta\upsilon = \left[_{\chi_1^T}\Psi_{\Xi_1^T}\right]{}_{\Xi_1}\delta\tilde{s} \tag{7.4}$$

gives the changes in the non-slack bus voltages given that the slack bus is entirely responsible for maintaining power balance if the power-voltage sensitivity matrix

$$\left[_{\chi_1^T}\Psi_{\Xi_1^T}\right] = \left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger \begin{bmatrix} 0 & \left[-_{\Xi_1}U_{0,1}\right]^{-T} \left[_{\Xi_1}U_{0,\chi_1}\right]^T \\ 0 & I_{2n-2} \end{bmatrix} \in \mathbb{R}^{(2n-2)\times 2n}. \tag{7.5}$$

Proof. $_{\Xi_1}\delta\hat{s} = \begin{bmatrix} 0 & \left[-_{\Xi_1}U_{0,1}\right]^{-T} \left[_{\Xi_1}U_{0,\chi_1}\right]^T \\ 0 & I_{2n-2} \end{bmatrix}{}_{\Xi_1}\delta\tilde{s}$ is on the power flow manifold (Lemma 7.2.1) and

$$_{\Xi_1}\delta\hat{s} \perp ker\left(\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger\right).$$

Thus, the kernel of $\left[_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger$ does not change $_{\Xi_1}\delta\hat{s}$.

$_{\Xi_1}\delta\hat{s}$ changes only the slack bus portion of $_{\Xi_1}\delta\tilde{s}$ to maintain power balance. □

(7.5) gives the power-voltage sensitivity and (7.4) answers Question 2 for a network with a slack bus.

Corollary 7.2.2.1. *For a network with a single slack bus,*

$$\chi_1^T \delta \upsilon = \left[_{\chi_1^T} \Psi_{\chi_1} \right] {}_{\chi_1^T} \delta s \tag{7.6}$$

gives the same changes in the non-slack bus voltages as (7.4) *using the power-voltage sensitivity matrix*

$$\left[_{\chi_1^T} \Psi_{\chi_1} \right] = \left[_{\chi_1^T} \Gamma_{\chi_1} \right]^{-1} \in \mathbb{R}^{(2n-2) \times (2n-2)}. \tag{7.7}$$

Proof. $_{\chi_1^T} \mathsf{U \Sigma V^T}$ is the Singular Value Decomposition of $\left[_{\chi_1^T} \Gamma_{\chi_1} \right]$.

$_{\Xi_1} \mathsf{U_+ \Sigma V^T}$ is the thin Singular Value Decomposition of $\left[_{\Xi_1} \Gamma_{\chi_1} \right]$.

$_{\chi_1^T} \mathsf{U} = (\Xi_1 \chi_1)^T {}_{\Xi_1} \mathsf{U_+}$ because Ξ_1 is an orthogonal matrix.

Taking (7.4),

$$
\begin{aligned}
{}_{\chi_1^T} \delta \upsilon &= \left[_{\Xi_1} \Gamma_{\chi_1} \right]^\dagger \begin{bmatrix} 0 & \left[-_{\Xi_1} \mathsf{U}_{0,1} \right]^{-T} \left[_{\Xi_1} \mathsf{U}_{0,\chi_1} \right]^T \\ 0 & I_{2n-2} \end{bmatrix} {}_{\Xi_1} \delta \tilde{s} \\
&= \mathsf{V \Sigma^{-1}} {}_{\Xi_1} \mathsf{U_+}^T \left[\begin{bmatrix} -_{\Xi_1} \mathsf{U}_{0,1} \end{bmatrix}^{-T} \left[_{\Xi_1} \mathsf{U}_{0,\chi_1} \right]^T \\ I_{2n-2} \end{bmatrix} {}_{\chi_1^T} \delta \tilde{s} \\
&= \mathsf{V \Sigma^{-1}} \left[_{\Xi_1} \mathsf{U_+}^T \Xi_1 \chi_1 \right] {}_{\chi_1^T} \delta \tilde{s} \\
&= \mathsf{V \Sigma^{-1}} {}_{\chi_1^T} \mathsf{U^T} {}_{\chi_1^T} \delta \tilde{s} \\
&= \left[_{\chi_1^T} \Psi_{\chi_1} \right] {}_{\chi_1^T} \delta s
\end{aligned} \tag{7.8}
$$

(7.8) is true because $_{\Xi_1} \mathsf{U_+} \perp {}_{\Xi_1} \mathsf{U_0}$. □

$\left[_{\chi_1^T} \Gamma_{\chi_1} \right]$ is Γ with the rows corresponding to the real and reactive power injections and the columns corresponding to the slack bus voltage and angle removed. (7.7) requires fewer calculations than (7.5) and no pseudoinverse. Thus, for a network with a slack bus, it is more efficient to determine the power-voltage sensitivity using (7.6) than using (7.4).

Two-Bus DC Demonstration of Nullspace-based Power Flow Linearization for a Network with a Slack Bus

The NPFL for a two-bus DC network with a slack bus at bus 1 corresponds to the power flow manifold tangent plane in Fig. 6.3. In this section we work out the power-voltage sensitivity and slack bus power injection sensitivity for specific values to check that the values agree with intuition.

For a two-bus DC network with line conductance g_{12} and bus two shunt conductance g_{20},

$$Y = \begin{bmatrix} g_{12} & -g_{12} \\ -g_{12} & g_{12} + g_{20} \end{bmatrix}, \; \delta p = \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix}, \text{ and } \delta v = \begin{bmatrix} \delta v_1 \\ \delta v_2 \end{bmatrix},$$

where g_{12} and g_{20} are the line conductance and shunt conductance at bus 2, respectively. Γ gives the voltage-power sensitivity

$$\Gamma \begin{bmatrix} \delta v_1 \\ \delta v_2 \end{bmatrix} = \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix},$$

$$\Gamma = \begin{bmatrix} 2g_{12} - v_2^* g_{12} & -g_{12} \\ -v_2^* g_{12} & -g_{12} + 2v_2^* g_{12} + 2v_2^* g_{20} \end{bmatrix}.$$

Setting bus 1 to be the slack bus, $v_1 = 1$,

$$\chi_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

χ_1 is 2×1 rather than 4×2 because this example is for a DC network not an AC network. The voltage-power sensitivity with a slack bus is

$$\Gamma_{\chi_1} \delta v_2 = \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix}, \tag{7.9}$$

$$\Gamma_{\chi_1} = \begin{bmatrix} -g_{12} \\ -g_{12} + 2v_2^* g_{12} + 2v_2^* g_{20} \end{bmatrix}.$$

The basis vector for the left-nullspace of Γ_{χ_1} is

$$U_0 = \begin{bmatrix} -g_{12} + 2v_2^* g_{12} + 2v_2^* g_{20} \\ g_{12} \end{bmatrix}.$$

We use (7.3) to answer Question 1,

$$\delta \hat{p}_1 = \frac{-g_{12}}{-g_{12} + 2v_2^* g_{12} + 2v_2^* g_{20}} \delta p_2 \tag{7.10}$$

There is one slack bus so we use (7.6) to answer Question 2,

$$\delta v_2 = \begin{bmatrix} \chi_1^T \Psi_{\chi_1} \end{bmatrix} \delta p_2, \tag{7.11}$$

$$\chi_1^T \Psi_{\chi_1} = [-g_{12} + 2v_2^* g_{12} + 2v_2^* g_{20}]^{-1} \delta p_2.$$

Consider the example admittance values $g_{12} = 1$ and $g_{20} = 0.5$. Calculating the NPFL at $v_1 = 1$, $v_2 = 1$ we get the following slack bus power injection sensitivity and power-voltage

sensitivity,[5]

$$\delta\hat{\mathbf{p}}_1 = -\frac{1}{2}\delta\mathbf{p}_2, \text{ and}$$

$$\delta\mathbf{v}_2 = \frac{1}{2}\delta\mathbf{p}_2.$$

These sensitivities agree with intuition—the change in the power injection at bus 2 will flow both into the line to bus 1 and into the shunt admittance at bus 2. Thus, the change in the power injection at bus 1 will be $-\frac{1}{2}\delta\mathbf{p}_2$. Since the voltage at bus 1 is held constant, the voltage at bus 2 is determined by the drop in voltage that occurs on the line to bus 1. Since only one half of the change in the power injection at bus 2 changes the power flow on the line to bus 1, the change in bus 2's voltage will be given by $\frac{1}{2}\delta\mathbf{p}_2$.

7.3 General Nullspace-based Power Flow Linearization

The Nullspace Power Balance introduced in Section 7.2 applies to a network with a single slack bus/constant voltage bus. In this section, we introduce a general form of Nullspace Power Balance which can apply to networks with more than one slack bus/constant voltage bus. The general form of Nullspace Power Balance applies to distribution networks with more than one substation.

Recall the modified goal of power flow linearization that we stated at the beginning of this chapter—"produce a matrix that takes a vector of changes in power injections as the input and produces both the changes in the voltages on the system and an adjusted vector of power injections that satisfies an approximation of power balance for the system." Questions 1 and 2 introduced at the beginning of Chapter 7 incorporated the knowledge that, for a network with a single slack bus, the slack bus will adjust its power injection so that the total power into the network is equal to the total power out. In this section we replace this single slack bus approximation with the knowledge of "the set of permissible power perturbation directions."[6] The set of permissible power perturbation directions allows/requires the person using the NPFL to specify which nodes adjust their power injections in order to maintain power balance.

The general Nullspace Power Balance and the general form of NPFL (still) assume that the voltage at at least one node is held constant. The general form of NPFL, which satisfies the general form of Nullspace Power Balance, answers the following two questions.

---

[5]$\mathbf{g}_{12} = 1$, $\mathbf{g}_{20} = 0.5$, $\mathbf{v}_1 = 1$, and $\mathbf{v}_2 = 1$ were used to generate all of the two-bus DC network plots in this dissertation.

[6]The set of permissible power perturbation directions is the mathematically rigorous term for "the ways in which NPFL is allowed to adjust the given power injections so that the power injections satisfy power flow."

1. What are the changes in the power injections that satisfy power balance that are obtained by adjusting the given set of power injections according to the set of permissible power perturbation directions?

2. How do the changes in the power injections that satisfy power balance given by Question 1 change the voltages at the non-constant-voltage buses?

These new questions incorporate the "set of permissible power perturbation directions" concept into the modified goal of power flow linearization that we stated at the beginning of Chapter 7.

To develop intuition for the new Questions 1 and 2, consider the circumstance of a distribution network with two substations and a set of permissible power perturbation directions defined as "the two substations on the network will adjust their real power and reactive power injections to maintaining power balance."

General Nullspace Power Balance

To answer the new Question 1 given above, we start by defining M as the set of elements of υ which are held constant and m as the cardinality of M. We express the set of permissible power perturbation directions mathematically by stacking the vectors β_k, $k \in \begin{bmatrix} 1 & \dots & m \end{bmatrix}$, for each permissible power perturbation direction together to make the matrix β,

$$\beta := \begin{bmatrix} \beta_1 & \dots & \beta_m \end{bmatrix}.$$

Remark 7.3.4, presented later in this section, describes why β must have m columns.

Remark 7.3.1. *β's vector space is the permissible projection directions.*

We define $\chi \in \mathbb{R}^{2n \times (2n-m)}$ as the concatenation of the indicator vectors for each voltage magnitude or angle entry of υ that is *not* held constant.[7] Mathematically, this definition corresponds to

$$\chi = \begin{bmatrix} e_{k_1} & \dots & e_{k_{2n-m}} \end{bmatrix}, \ k_i \notin \text{M}. \tag{7.12}$$

For some $i \in [1, \dots, 2n - m]$, i is the column count and k_i is the index of the voltage magnitude or angle in υ. This formulation allows the voltage magnitude and/or the voltage angle at each node to be held constant.

We define M_{slack} as a special case of M which corresponds to the circumstance in which the voltage magnitude and angle at each slack bus are both held constant. The voltage magnitude indices in the first half of M_{slack} match the voltage angle indices in the second half of M_{slack}

$$\text{M}_{\text{slack}} = [k_1, \dots, k_m] = [k_1, \dots, k_{\frac{m}{2}}, k_{1+n}, \dots, k_{\frac{m}{2}+n}]. \tag{7.13}$$

---

[7](7.12) assumes that υ is defined according to (5.11). That is, all of the voltage magnitudes then all of the voltage angles.

We define the change-in-voltage magnitude and change-in-angle vector with the constant entries removed as

$$\chi^T \delta v := \chi^T \delta v,$$

and the voltage-power sensitivity matrix with the columns corresponding to the constant entries removed

$$\Gamma_\chi := \Gamma \chi.$$

The voltage-power sensitivity equation for a general network is

$$\delta s = \Gamma_\chi \chi^T \delta v. \tag{7.14}$$

We intentionally do not define a Ξ permutation matrix for the general NPFL because the NPFL does not require that the set of permissible power perturbation directions is the concatenation of the power injections from individual nodes. That is, NPFL does not require that β is constructed from $e_i \in \mathbb{R}^{2n}$ vectors.[8]

We define $\mathcal{M}_\chi$ as done in (6.3) but using the general χ instead of the specific χ_1. Similarly, we define $\mathcal{T}_\chi$ and Γ_χ as done in (6.1) and (6.5) respectively using χ instead of χ_1. Taking the Singular Value Decomposition of Γ_χ we get

$$\Gamma_\chi = \mathsf{U}\Sigma\mathsf{V}^{\mathrm{T}} = [\mathsf{U}_+, \mathsf{U}_0]\,\Sigma\mathsf{V}^{\mathrm{T}}, \tag{7.15}$$

where $\mathsf{U}_0 \in \mathbb{R}^{2n-m}$ is the left-nullspace of Γ_χ.

Lemma 7.3.2. *$rank(\Gamma_\chi) = 2n - m$ almost surely.*

Proof. The proof follows the same steps as the proof for Lemma 7.1.1. $\square$

We define $\delta\tilde{s}$ as any vector in $\mathbb{R}^{2n}$. $\delta\tilde{s}$ may or may not satisfy power balance for a given network.

Lemma 7.3.3. *If $rank(\beta) = m$, $\delta\hat{s}$ is the projection of $\delta\tilde{s}$ onto $\mathcal{T}_\chi$ in the direction specified by β when $\delta\hat{s}$ is defined using*

$$\delta\hat{s} = \Omega\delta\tilde{s} \tag{7.16}$$

$$\Omega := I - \beta\left(\mathsf{U}_0^T\beta\right)^{-1}\mathsf{U}_0^T \tag{7.17}$$

where $\mathsf{U}_0 \in \mathbb{R}^{2n\times m}$ is the matrix describing the left-nullspace of Γ_χ.[9]

---

[8]In the derivation of the single-slack NPFL, the Ξ_1 permutation matrix helped keep the notation organized, but was not necessary.

[9](7.16) and (7.17) can be derived as follows,

$$\delta\hat{s} = \delta\tilde{s} - \beta a \text{ for some a.}$$
$$0 = \mathsf{U}_0^T(\delta\tilde{s} - \beta a) \text{ for } \delta\hat{s} \text{ to be on } \mathcal{T}_\chi.$$
$$a = (\mathsf{U}_0^T\beta)^{-1}\mathsf{U}_0^T\delta\tilde{s}.$$
$$\text{Therefore, } \delta\hat{s} = \left(I - \beta\left(\mathsf{U}_0^T\beta\right)^{-1}\mathsf{U}_0^T\right)\delta\tilde{s}.$$
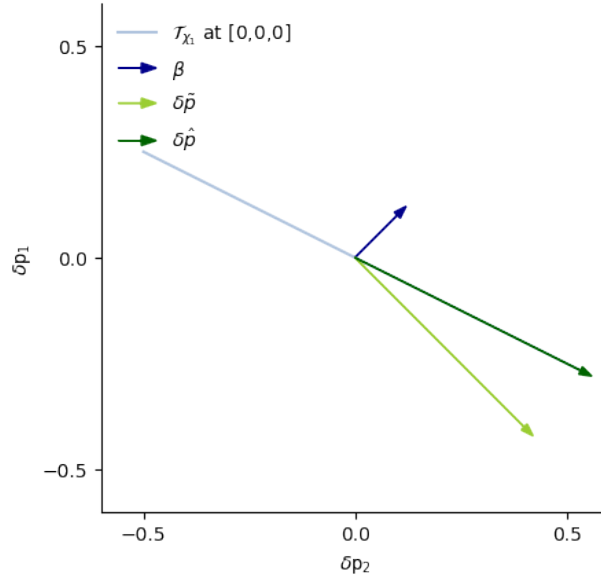
Figure 7.3: Two-dimensional demonstration of the projection onto $image(\Gamma_\chi)$ (or $\mathcal{T}_{\chi_1}$) in the β direction.

Proof. $\delta\hat{s} \in \mathcal{T}_\chi$ if $\delta\hat{s} \in image(\Gamma_\chi)$.

$$\mathsf{U}_0^T \delta\hat{s} = \mathsf{U}_0^T \left(I - \beta(\mathsf{U}_0^T \beta)^{-1}\mathsf{U}_0^T\right)\delta\tilde{s}$$
$$= 0.$$

Therefore $\delta\hat{s} \in image(\Gamma_\chi)$.

$\delta\hat{s} - \delta\tilde{s} = \beta\left(\mathsf{U}_0^T\beta\right)^{-1}\mathsf{U}_0^T$ is the projection direction.

Therefore the projection onto $\mathcal{T}_\chi$ is in the direction β. □

Ω is the power injection projection matrix. It projects any power injection vector $\delta\tilde{s}$ onto the power flow manifold according to the set of permissible power perturbation directions specified by the person using NPFL. The projection $\delta\hat{s}$ satisfies $\mathcal{T}_\chi$'s approximation of power balance, and (7.16) answers Question 1 given at the beginning of this section.

The projection onto $\mathcal{T}_{\chi_1}$ in direction β is demonstrated in Fig. 7.3. In Fig. 7.3, $\beta \neq \beta_{\mathrm{slack}}$ and $\delta\tilde{p}_1$ affects $\delta\hat{p}$. This is generally true—when $\beta \neq \beta_{\mathrm{slack}}$ then all of the entries of $\delta\tilde{s}$, including the entries that correspond to the constant voltage bus, affect $\delta\hat{s}$.

Remark 7.3.4. *If $rank(\beta) < m$ and $\delta\check{s}$ is set using an alternate equation that uses the pseudoinverse of $\left(\mathsf{U}_0^T\beta\right)$,*

$$\delta\check{s} = \left(I - \beta(\mathsf{U}_0^T\beta)^\dagger\mathsf{U}_0^T\right)\delta\tilde{s},$$

$\delta\check{s}$ is not necessarily on $\mathcal{T}_\chi$.

Thus, NPFL should not be used unless the matrix β that codifies the set of permissible power perturbation directions has rank equal to m, the number of entries in $\delta\upsilon$ held constant.

Comparing the Nullspace Power Balance in Section 7.2 with Ω_1

Lemma 7.2.1 introduced (7.3), which gives $\delta\hat{s}_1$ for a network with a single slack bus. For a network with a single slack bus,

$$
\beta_1 = \begin{bmatrix} 1 & 0 \\ 0_n & 0_n \\ 0 & 1 \\ 0_{n-2} & 0_{n-2} \end{bmatrix} \in \mathbb{R}^{2n\times 2}
\tag{7.18}
$$

and $\Xi_1\beta_1 = \begin{bmatrix} I_2 & 0 \end{bmatrix}^T \in \mathbb{R}^{2n\times 2}$.

Remark 7.3.5. *For a network with a single slack bus and Ω_1 built with β_1 using (7.17),*

$$
\Xi_1\Omega_1\Xi_1^T = \begin{bmatrix} 0 & \left[-\Xi_1 U_{0,1}\right]^{-T} \left[\Xi_1 U_{0,\chi_1}\right]^T \\ 0 & I_{2n-2} \end{bmatrix}.
\tag{7.19}
$$

Proof.

$$
\begin{aligned}
\Xi_1\Omega\Xi_1^T &= \Xi_1 \left(I - \beta \left(U_0^T\beta\right)^{-1} U_0^T \right) \Xi_1^T \\
&= I - \Xi_1\beta \left(U_0^T\Xi_1^{-1}\Xi_1\beta\right)^{-1} U_0^T\Xi_1^T \\
&= I - \beta_1 \left(\Xi_1 U_0^T\beta_1\right)^{-1} \Xi_1 U_0^T \\
&= \begin{bmatrix} 0 & \left[-\Xi_1 U_{0,1}\right]^{-T} \left[\Xi_1 U_{0,\chi_1}\right]^T \\ 0 & I_{2n-2} \end{bmatrix}
\end{aligned}
$$

$\square$

Thus, (7.16) and (7.3) give the same estimate for $\delta\hat{s}$ for a network with a single slack bus.

General Power-Voltage Sensitivity

To answer Question 2 posed at the beginning of this section for a general network, "How do the changes in the power injections that satisfy power balance given by Question 1 change the voltages at the non-constant-voltage buses?," we must find the power-voltage sensitivity matrix which maps $\delta\tilde{s}$ to $\chi^T\delta\upsilon$ such that $\chi^T\delta\upsilon$ is also mapped to by $\delta\hat{s} = \Omega\delta\tilde{s}$. That is, the power-voltage sensitivity matrix gives the changes in the voltages at the non-constant-voltage buses that correspond to the given power injections $\delta\tilde{s}$ adjusted to satisfy power balance according to the allowable power injection adjustments β.

Similar to $\Gamma_{\chi_1}^\dagger$, we define

$$\Gamma_\chi^\dagger := [\Gamma_\chi]^\dagger.$$

Theorem 7.3.6. *If* $rank(\beta) = m$,

$$\chi_1^T \delta\upsilon = {}_{\chi^T}\Psi\delta\tilde{s} \tag{7.20}$$

gives the changes in the voltages at the non-constant-voltage buses that correspond to the given changes in the power injections $\delta\tilde{s}$ *adjusted to satisfy power balance according to* β *if the power-voltage sensitivity matrix*

$$\chi^T\Psi = \Gamma_\chi^\dagger\Omega \in \mathbb{R}^{(2n-m)\times 2n}. \tag{7.21}$$

Proof. $\delta\hat{s} = \Omega\delta\tilde{s} \in image\,(\Gamma_\chi)$ is on the power flow manifold (Lemma 7.3.3) and

$$\delta\hat{s} \perp ker\left(\Gamma_\chi^\dagger\right).$$

Thus, the kernel of $\Gamma_\chi^\dagger$ does not change $\delta\hat{s}$.

$\delta\hat{s}$ adjusts $\delta\tilde{s}$ to satisfy power balance according to β (Lemma 7.3.3). $\qquad\square$

(7.21) gives the power-voltage sensitivity and (7.20) answers Question 2.

Remark 7.3.7. (7.4) *is a special case of* (7.20).

Proof. When there is a single slack bus and $\beta = \beta_1$ from (7.18), (7.20) gives

$$
\begin{aligned}
\chi_1^T \delta\upsilon = {}_{\chi_1^T}\Psi\delta\tilde{s} &= \Gamma_{\chi_1}^\dagger\Omega_1\delta\tilde{s} \\
&= \Gamma_{\chi_1}^\dagger\Xi_1^T\Xi_1\Omega_1\Xi_1^T\Xi_1\delta\tilde{s} \\
&= \left[{}_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger \begin{bmatrix} 0 & \left[-{}_{\Xi_1}\mathsf{U}_{0,1}\right]^{-T} \\ 0 & \begin{matrix}\left[{}_{\Xi_1}\mathsf{U}_{0,\chi_1}\right]^T \\ I_{2n-2}\end{matrix} \end{bmatrix} {}_{\Xi_1}\delta\tilde{s}.
\end{aligned}
$$

$\left[{}_{\Xi_1}\Gamma_{\chi_1}\right]^\dagger = \Gamma_{\chi_1}^\dagger\Xi_1^T$ because Ξ_1 is an orthogonal matrix. The substition for $\Xi_1\Omega_1\Xi_1^T$ is given by Remark 7.3.5. $\qquad\square$

7.4 Relationship between NPFL and other Power Flow Linearizations

Power flow linearizations are ubiquitous in power system analysis and the literature on power flow linearizations is vast. This section contextualizes NPFL in relation to the other prevalent power flow linearizations.

In order to describe the linearizations, we define the conductance matrix G and susceptance matrix B according to

$$Y =: G + Bj,$$

and the network impedance matrix Z and its constituent resistance matrix R and reactance matrix X according to

$$Z := R + Xj := Y^{-1}.$$

If Y does not include shunt admittances it is not full rank and the pseudoinverse $Y^{\dagger}$ is use. We also introduce the $_{\mathcal{R}}$ subscript corresponding to the "reduced Laplacian" for networks with a single slack bus. We also use $_{\mathcal{R}}$ to indicate that the slack bus entry of a vector has been removed:

$$[\cdot]_{\mathcal{R}} := \begin{cases} \begin{bmatrix} 0 & I_{n-1} \end{bmatrix} [\cdot] \begin{bmatrix} 0 \\ I_{n-1} \end{bmatrix}, & \text{if } [\cdot] \text{ is a matrix,} \\ \begin{bmatrix} 0 & I_{n-1} \end{bmatrix} [\cdot], & \text{if } [\cdot] \text{ is a vector.} \end{cases}$$

That is, matrices with the $_{\mathcal{R}}$ subscript have their first row and column removed, and vectors with the $_{\mathcal{R}}$ subscript have their first entry removed. Once again, we assume without loss of generality that the substation bus is the first bus in the power injection and voltage vectors.

Other Power Flow Linearizations

<u>DC Power Flow</u> (DCPF)[10] [101], [103], [105]–[108] is derived with the following approximations:

Approximation 1. *$u \approx 1$, thus $s \approx \bar{i}$.*

Approximation 2. *$\theta \approx 0$, thus $\sin(\theta) \approx \theta$ and $\cos(\theta) \approx 1$.*

Approximation 3. *The network conductances are all 0.*

Given these assumptions, DCPF gives an expression for the real power injections at the non-slack buses in terms of the voltage angles relative to the slack bus angle.

$$\delta p_{\mathcal{R}} = -B_{\mathcal{R}} \delta\theta_{\mathcal{R}} \tag{7.22}$$

The reduced Laplacian $B_{\mathcal{R}}$ is full rank even if there are no shunt admittances on the network. Thus, (7.22) can be inverted to give the real power-voltage angle sensitivity

$$\delta\theta_{\mathcal{R}} = X_{\mathcal{R}} \delta p_{\mathcal{R}}.$$

---

[10]The term "DC Power Flow" is an artifact of the DC works-like grid models built in the control rooms from DC circuit components for system operators in the mid-twentieth century.

Decoupled Linearized Power Flow (DLPF) [109] uses the same approximations, but provides an equation for the voltage magnitudes as well,

$$\begin{aligned} \delta p_{\mathcal{R}} &= -B_{\mathcal{R}}\delta\theta_{\mathcal{R}} \\ \delta q_{\mathcal{R}} &= -B_{\mathcal{R}}\delta v_{\mathcal{R}}. \end{aligned} \tag{7.23}$$

Inverting (7.23) gives the reactive power-voltage magnitude sensitivity

$$\delta v_{\mathcal{R}} = X_{\mathcal{R}}\delta q_{\mathcal{R}}.$$

Coupled Linearized Power Flow (CLPF) [86], [110]–[113] uses approximations 1 and 2 and replaces Approximation 3 with

Approximation 4. *The product of difference terms are small, thus $(v_i - v_k)(\theta_i - \theta_k) \approx 0$.*

The CLPF equations are

$$\begin{bmatrix} \delta p_{\mathcal{R}} \\ \delta q_{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} G_{\mathcal{R}} & -B_{\mathcal{R}} \\ -B_{\mathcal{R}} & -G_{\mathcal{R}} \end{bmatrix} \begin{bmatrix} \delta v_{\mathcal{R}} \\ \delta\theta_{\mathcal{R}} \end{bmatrix}, \tag{7.24}$$

derived in the first portion of Appendix F.1. Inverting the CLPF matrix gives the power-voltage sensitivity,

$$\begin{bmatrix} \delta v_{\mathcal{R}} \\ \delta\theta_{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} R_{\mathcal{R}} & X_{\mathcal{R}} \\ X_{\mathcal{R}} & -R_{\mathcal{R}} \end{bmatrix} \begin{bmatrix} \delta p_{\mathcal{R}} \\ \delta q_{\mathcal{R}} \end{bmatrix}, \tag{7.25}$$

as demonstrated in the second portion of Appendix F.1.

Voltage Magnitude Squared Linearized Edge Flow (SLEF), first introduced for radial networks in [114], is an alternative power flow linearization formulation that relates the squared voltage magnitudes to the power flows on each edge/line/branch [97], [115], [116]. [114] introduced this linearization under the name "simplified DistFlow," which has subsequently been changed to "LinDistFlow," because [114] focused on radial distribution network applications. While the LinDistFlow name has become a part of the power systems lexicon, we use the term "Voltage Magnitude Squared Linearized Edge Flow"/SLEF to emphasize the fact that it can apply to mesh networks as well as radial networks.[11] SLEF is derived for each individual edge using the following approximation.

Approximation 5. *There are no power losses on the lines.*

---

[11]We use "edge" rather than "branch" because "branch" implies a radial network.

Defining the set of network edges/lines/branches as $\mathcal{E}$, r_{ik} and x_{ik} as the resistance and reactance of edge (i,k), and p_{ik} and q_{ik} as the real and reactive power flowing from node i onto edge (i,k) respectively, SLEF is given by

$$v_i^2 - v_k^2 = 2r_{ik}p_{ik} + 2x_{ik}q_{ik} \ \forall (i,k) \in \mathcal{E}, \tag{7.26}$$

$$\sum_j p_{ij} = p_i, \ \sum_j q_{ij} = q_i, \ \forall i \in \{1, \ldots, n\}. \tag{7.27}$$

Standard Linearized Edge Flow (LEF) includes the additional approximation

Approximation 6. $v_i^2 - v_k^2 \approx 2(v_i - v_k)$.

This approximation gives the following LEF equations for voltage magnitudes

$$v_i - v_k = r_{ik}p_{ik} + x_{ik}q_{ik} \ \forall (i,k) \in \mathcal{E}, \tag{7.28}$$

$$\sum_j p_{ij} = p_i, \ \sum_j q_{ij} = q_i, \ \forall i \in \{1, \ldots, n\}. \tag{7.29}$$

[117], [118] added the voltage angle equation

$$\theta_i - \theta_k = x_{ik}p_{ik} - r_{ik}q_{ik} \ \forall (i,k) \in \mathcal{E} \tag{7.30}$$

to the LEF formulation, based on Approximations 1, 2 and 4.

Remark 7.4.1. *LEF and CLPF are equivalent.*

The proof is given in Appendix F.2. This statement is in conceptual agreement with the statement in [119] that the convex relaxations of the bus-injection and branch-flow power flow formulations are equivalent, however neither statement implies the other.

NPFL for a Single-Slack, No-Shunt Admittance Network at Nominal Voltage

Lemma 7.4.2. *For a network with a single slack bus and no shunt admittances, when $u^* = 1_n$,[12] the voltage-power sensitivity is given by*

$$\chi_1^T \Gamma \chi_1 = \begin{bmatrix} G_{\mathcal{R}} & -B_{\mathcal{R}} \\ -B_{\mathcal{R}} & -G_{\mathcal{R}} \end{bmatrix}, \tag{7.31}$$

---

[12]When there are shunt admittances but no additional load on the network,

$$\chi_1^T \Gamma \chi_1 = \chi_1^T \langle \text{diag}(u_{\text{noLoad}}) \rangle N \langle Y \rangle R(u_{\text{noLoad}}) \chi_1,$$

where u_{noLoad} is the "No Load" voltage vector [104], [120] that produces no current injections ($u_{\text{noLoad}} = ker(Y)$) in addition to the current that flows into the shunt admittances.

the power-voltage sensitivity is given by

$$\chi_1^T \Psi_{\chi_1} = \begin{bmatrix} R_{\mathcal{R}} & X_{\mathcal{R}} \\ X_{\mathcal{R}} & -R_{\mathcal{R}} \end{bmatrix}, \tag{7.32}$$

and the slack bus power injection is given by

$$\Xi_1 \Omega_1 \Xi_1^T = \begin{bmatrix} 0 & \begin{bmatrix} -1_{n-1}^T & 0 \\ 0 & -1_{n-1}^T \end{bmatrix} \\ 0 & I_{2n-2} \end{bmatrix}. \tag{7.33}$$

See Appendix F.3 for the proof of Lemma 7.4.2. Note, when the conditions for Lemma 7.4.2 apply, the slack bus real and reactive power injections are equal to the negative sum of all of the other real and reactive power injections on the network, respectively, (7.33). When the conditions for Lemma 7.4.2 do not hold, this is not generally the case.

NPFL and Other Power Flow Linearizations

Remark 7.4.3. *DCPF, DLPF, CLPF, and LEF assume that the constant-voltage/slack bus power injection maintains power balance, but do not provide the constant-voltage/slack bus power injection sensitivity.*

Remark 7.4.4. *DCPF, DLPF, CLPF, and LEF are special cases of NPFL power-voltage sensitivity. They are equivalent to $\chi_1^T \Psi_{\chi_1}$ when $\chi_1^T \Psi_{\chi_1}$ is calculated for a single-slack bus network with $u^* = 1_n$ (7.32).*

The equivalence between $\chi_1^T \Psi_{\chi_1}$ in (7.32) and DCPF, DLPF, CLPF, and LEF is evident from inspection and Remark 7.4.1, and leads to the following hypothesis.

Hypothesis 1. *Defining the power flow manifold using the voltage magnitude squared, that is v^2 rather than v, provides a more accurate power flow linearization.*

Hypothesis 1 is based on the fact that LEF is derived using Assumptions 5 and 6, while LEF is derived using just Assumption 5.

(7.6) provides the power-voltage sensitivity for voltages in rectangular coordinates [104], [121], [122] if the $R(u^*)$ matrix is removed from the Γ definition in (5.8). It is worth noting that voltages in rectangular coordinates are less useful than voltages in polar coordinates because voltage constraints apply to voltage magnitudes.

There are power flow linearizations that are not special cases of NPFL power-voltage sensitivity. Linearizations derived from fixed points [120], [123] and linearizations produced by a scenario-specific optimization [124], [125] cannot be derived from the power flow manifold tangent plane and are not special cases of NPFL. Linearizations that consider constant current/ZIP load characteristics [104], [126], [127] cannot be derived from NPFL in its current form. An NPFL formulation that allows for constant current loads/power injections that depend linearly on voltage is the subject of current research.

7.5 Nullspace-based Power Flow Linearization for Loss Sensitivity

This section provides an example of how NPFL can be used to find sensitivities other than power-voltage sensitivity and slack bus power-injection sensitivity that may also be of interest. The sum of the power injections at all of the nodes on the grid is equal to the sum of the losses on the network and the power that leaves the network through the shunt admittances. If the network has no shunt conductances (the real part of the shunt admittances are all zero), then the sum of the real power injections at all of the nodes on the grid is equal to the sum of the real power losses on the network.[13] The power injection–real power loss sensitivity may be a useful tool for power system operators for applications such as determining the loss component of locational marginal prices (LMPs).

The change in real power losses $\delta\mathfrak{L}$ are given by the equation

$$\delta\mathfrak{L} = \left| \sum_{i=1}^{n} \delta p_i \right| = \left| 1^T \delta p \right|.$$

Remark 7.5.1 gives the loss sensitivity when a single slack bus is responsible for power balance. We define $\Omega_{1,p}$ as the rows of Ω_1 that give $\delta\hat{p}_1$, the estimate of the real power injection at the slack bus.

Remark 7.5.1. $\left(\Omega_{1,p} + \begin{bmatrix} 1_{n-1}^T & 0_{n-1}^T \end{bmatrix} \right)$ *is the loss sensitivity that corresponds to the power flow manifold tangent plant at x^* when a single slack bus is responsible for maintaining power balance. That is,*

$$\delta\mathfrak{L} = \left| \left(\Omega_{1,p} + \begin{bmatrix} 1_{n-1}^T & 0_{n-1}^T \end{bmatrix} \right) \delta s_\chi \right| \tag{7.34}$$

Proof.

$$\begin{aligned}
\delta\mathfrak{L} &= \left| 1_n^T \delta p_{\text{true}} \right| \\
&= \left| \delta\hat{p}_1 + \begin{bmatrix} 1_{n-1}^T & 0_{n-1}^T \end{bmatrix} \delta s_\chi \right| \\
&= \left| \left(\Omega_{1,p} + \begin{bmatrix} 1_{n-1}^T & 0_{n-1}^T \end{bmatrix} \right) \delta s_\chi \right|
\end{aligned}$$

$\square$

Remark 7.5.2 gives the loss sensitivity when power balance responsibility is determined by a general β.

Remark 7.5.2. $\begin{bmatrix} 1_n^T & 0_n^T \end{bmatrix} \Omega$ *is the loss sensitivity that corresponds to the power flow manifold tangent plant at x^* and the power balance responsibility vector β. That is,*

$$\delta\mathfrak{L} = \left| \begin{bmatrix} 1_n^T & 0_n^T \end{bmatrix} \Omega \delta\tilde{s} \right| \tag{7.35}$$

---

[13]In power system operation we are often interested in the real power losses rather than the reactive power losses, as the real power losses are physically meaningful, describing the additional real power that must be injected into the network.

Proof.

$$\delta \mathfrak{L} = \left| 1_n^T \delta p_{\text{true}} \right|$$
$$= \left| \begin{bmatrix} 1_n^T & 0_n^T \end{bmatrix} \delta \hat{s} \right|$$
$$= \left| \begin{bmatrix} 1_n^T & 0_n^T \end{bmatrix} \Omega \delta \tilde{s} \right|.$$

□

When $\beta = \begin{bmatrix} I_2 & 0 \end{bmatrix}$, (7.35) is equivalent to (7.34). It is also important to note that the loss sensitivity expressions (7.34) and (7.35) do not distinguish between line-flow power losses and shunt admittance "losses." We use quotation marks for the shunt admittance losses because the shunt admittances can also be used to model constant-impedance loads. If shunt admittances are used to model constant impedance loads, care must be taken when using (7.34) and (7.35) to model the network losses. One way to do this would be to subtract the shunt admittance loads from the loss estimates given by (7.34) and (7.35).

Two-Bus DC Demonstration of Nullspace-based Power Flow Linearization for Loss Sensitivity

For a two-bus DC network in which Bus 1 acts as a slack bus with $v_1 = 1$, the loss sensitivity is given by the two-bus DC equivalent of (7.34),

$$\delta \mathfrak{L} = \left| \left(\Omega_{1,p} + 1 \right) \delta p_2 \right|,$$

where $\Omega_{1,p}$ is the fraction in (7.10),

$$\Omega_{1,p} = \frac{-g_{12}}{-g_{12} + 2v_2^* g_{12} + 2v_2^* g_{20}}.$$

Fig. 7.4 plots the marginal loss/loss sensitivity for changes in the power injection at Bus 2. We use the following expressions for the losses $\mathfrak{L}$ when $g_{12} = 1$ and $g_{20} = 0$[14] to help explain Fig. 7.4,

$$\mathfrak{L} = (v_2 - 1)(v_2 - 1)$$
$$= p_2 - \sqrt{p_2 + \frac{1}{4}} + \frac{1}{2}.$$

These equations are plotted in Fig. 7.5. The second equation was derived using $p_2 = v_2(v_2 - 1)$.

When $v_2^* = 1$, the (marginal) loss sensitivity is 0. At $v_2^* = 1$, $p_2^* = \mathfrak{L} = 0$ and the slope of the loss curve in Fig. 7.5b is 0. At this operating point, there are no losses on the network and a marginal (infinitesimally small) increase in load does not increase the losses.

---

[14]Note, unlike the previous two-bus DC network demonstrations which used $g_{20} = \frac{1}{2}$, this example uses $g_{20} = 0$.
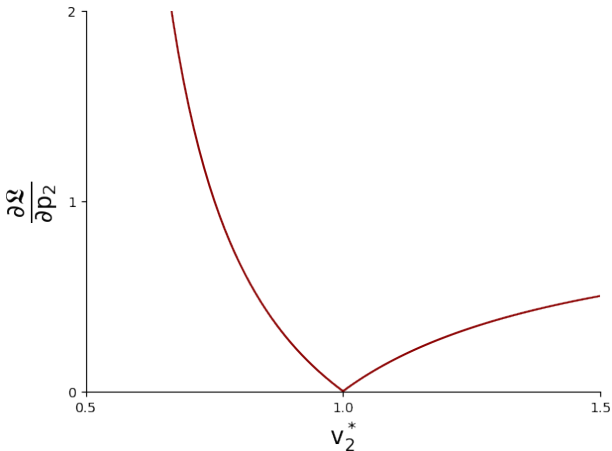
Figure 7.4: Two-bus DC network loss sensitivity for $g_{12} = 1$, $g_{20} = 0$, and $v_1 = 1$
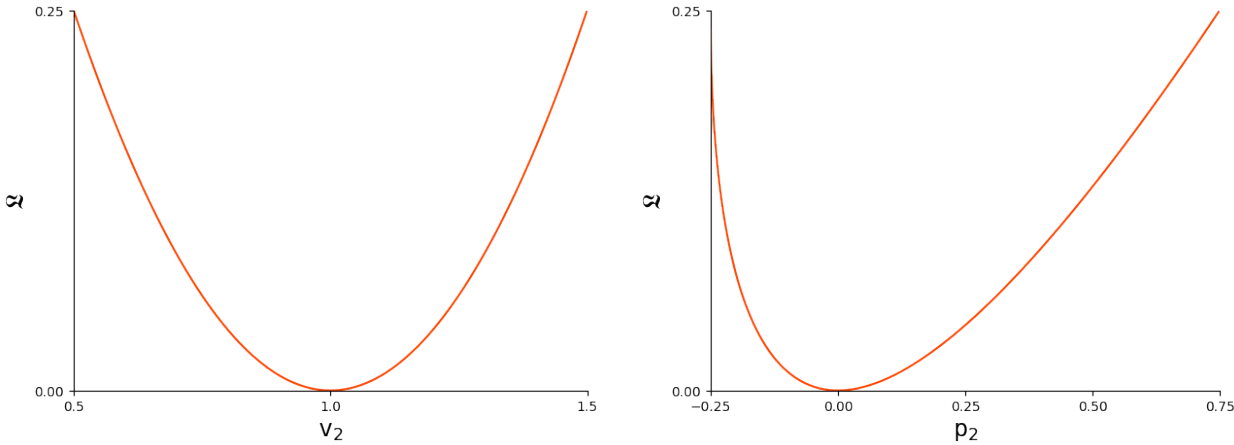
(a) Two-bus DC network total losses as a function of v_2

(b) Two-bus DC network total losses as a function of p_2

Figure 7.5: Total loss plots for $g_{12} = 1$, $g_{20} = 0$, and $v_1 = 1$

When $v_2^* = \frac{3}{4}$, the loss sensitivity is $|-1| = 1$. At $v_2^* = \frac{3}{4}$, $p_2^* = -0.1875$ and the slope of the loss curve in Fig. 7.5b is -1. At this operating point, a marginal increase in the load at Bus 2 results in losses equal to the marginal increase in load.

When $v_2^* = \frac{1}{2}$, the loss sensitivity is $|-\infty| = \infty$. At $v_2^* = \frac{1}{2}$, $p_2^* = -0.25$, $\mathfrak{L} = 0.25$ and the slope of the loss curve in Fig. 7.5b is infinite. At this operating point, a marginal increase in the load at Bus 2 would theoretically result in infinite losses. This statement corresponds to the fact that it is not possible to increase the load past $\frac{1}{4}$, because $\frac{1}{4}$ is the maximum power that can be transferred from Bus 1 to Bus 2 when $v_1 = 1$ and $g_{12} = 1$ [128].

Nullspace-based Power Flow Linearization for Locational Marginal Prices

This section makes it clear that the NPFL takes into account the marginal network losses at the linearization point. This is an important aspect of power flow linearization that has a number of currently-unused applications. For example, if LMPs were calculated using NPFL rather than using DC Power Flow or even Coupled Power Flow Linearization (Section 7.4), then it would not be necessary to add explicit loss-approximation terms to the LMP—the marginal losses would be baked into the LMPs.

Take, for example, a distribution network with a single substation (or, equivalently, a network with a single generator that behaves like a slack bus). The first row of the Ω_1 matrix in (7.19) gives the sensitivity of the substation injection to any injection on the network, taking into account the marginal losses on the network. If the LMP at a given node is calculated based on how much a change in the power injection at the given node affects the change in the substation power injection (where power is "purchased"), and the change in the substation power injection is calculated using Ω_1, then the marginal losses are included in the LMP for the given node. Thus, including additional loss-approximations adders in the LMP is unnecessary.

Transmission networks are more complicated because they have multiple generators, but the same logic for incorporating losses applies to transmission network LMPs—if the linearization is taken at the current operating point, and power balance is asserted by setting the voltage at at least one bus to be constant, then the LMPs given by the dual variables of the cost-minimizing optimal power flow (OPF) will include the marginal losses on the network. Including both δs and δv in the cost-minimizing OPF decision variables and including (7.14) in the constraint equations will assert power balance for the δs power injection vector and include the marginal losses in the LMP prices.

Chapter 8

Single Bus Injection Power-Voltage Sensitivity

Chapters 5, 6, and 7 described the power flow manifold and the accompanying tangent plane/linearization for a full network. The equations derived in Chapters 5, 6, and 7 require the full bus admittance matrix Y and can be used determine the impact of a power injection anywhere on the network on any voltage in the network.

In this chapter we use the NPFL power-voltage sensitivity (the answer to Question 2 that motivated NPFL) to derive the power flow linearization that relates *only* the power injections at a given bus to the voltages at that given bus. We title this linearization "Single bus Injection Power-voltage Sensitivity" (SIPS). SIPS can be used for a number of local control objectives, such as Voltage Phasor Control [90], [129].

SIPS uses a Thevenin Equivalent model for the electric grid [90], [129]. From the perspective of the given bus, the Thevenin Equivalent model has the same admittance as the grid.

8.1 Relationship between SIPS and Full-Network NPFL

For a Thevenin network, we are interested in the linearization of the power flow manifold for a two bus network with admittance matrix

$$Y = \begin{bmatrix} y_{12} & -y_{12} \\ -y_{12} & y_{12} + y_{20} \end{bmatrix}.$$

The power-voltage sensitivity matrix for bus two is

$$_{\chi^T}\Psi_\chi = \left[\left(\langle \overline{(u_2^* - 1)y_{12} + u_2^* y_{20}} \rangle + \langle \operatorname{diag}(u_2^*) \rangle N \langle y_{12} + y_{20} \rangle \right) R(u_2^*) \right]^{-1}. \tag{8.1}$$

[129] introduced a closely related power-voltage sensitivity matrix for a Thevenin-equivalent network model of a network,

$$\Psi_{\text{th}} = \left[\left(\langle \overline{(u_2^* - 1)y_{\text{th}}} \rangle + \langle \text{diag}(u_2^*) \rangle N \langle y_{\text{th}} \rangle \right) R(u_2^*) \right]^{-1}. \tag{8.2}$$

We are interested in describing the differences between (8.1) and (8.2). The two formulations are equivalent if the network does not have shunt admittances, $y_{20} = 0$. This can be seen by substituting y_{12} for y_{th}. If the network *does* have shunt admittances, $y_{20} \neq 0$, and $_{\chi^T} \Psi_\chi$ will be slightly different than Ψ_{th}. If both y_{12} and y_{20} are available, (8.1) should be used. If $y_{20} = 0$ or if y_{20} is not available and is assumed to be zero, (8.1) and (8.2) are equivalent.

Online admittance estimation can be used if y_{12} and y_{20} are unknown. [129] describes a method for estimating the Thevenin impedance online by applying recursive least squares to the changes in the local voltage phasor measurement and the current injection phasor measurement. The local voltage phasor measurement requires a voltage phasor measurement unit. The current injection phasor measurement can be estimated from the voltage phase and the real and reactive power injections.

We define $\hat{y}_{\text{th}}$ as the admittance estimate that is produced by applying least squares to the changes in the local voltage and current injection measurements. If there are no shunt admittances/voltage dependencies on the network (and the substation acts as an infinite bus), then $\hat{y}_{\text{th}} \approx y_{12}$. If $\hat{y}_{\text{th}} = y_{12}$, plugging $\hat{y}_{\text{th}}$ in for y_{th} in (8.2) gives (8.1) with $y_{20} = 0$.

On the other hand, if there are shunt admittances, $\hat{y}_{\text{th}} \approx (y_{12} + y_{20})$. If $\hat{y}_{\text{th}} = (y_{12} + y_{20})$, plugging $\hat{y}_{\text{th}}$ in for y_{th} in (8.2) gives

$$\Psi_{\text{th}} = \left[\left(\langle \overline{(u_2^* - 1)y_{12} + u_2^* y_{20} - y_{20}} \rangle + \langle \text{diag}(u_2^*) \rangle N \langle y_{12} + y_{20} \rangle \right) R(u_2^*) \right]^{-1},$$

which is accurate except for the $-y_{20}$ in the left-term. In order to avoid this source of error, it is necessary to have an estimate of both y_{12} and y_{20}, which can be plugged directly into (8.1).

y_{12} and y_{20} can be estimated separately by estimating the full $Y_{\text{th}} \in \mathbb{C}^{2\times 2}$ bus admittance matrix for the Thevenin equivalent network. Estimating Y_{th} requires a current phasor measurement on either the substation current or the shunt admittance.[1] Thus, the admittance estimation could no longer be done locally. Also, the substation's current measurement will be influenced by all of the injections on the network, and therefore be very noisy. In practice, just using $\hat{y}_{\text{th}}$ in (8.2) may be sufficiently accurate. Shunt admittances are often small compared with line admittances and therefore the error introduced into the linearization by the erroneous y_{20} term is likely to be benign.

---

[1]If the shunt admittance represents a "conceptual" current path such as the half-line charging admittances for Pi line models or the voltage dependence of local loads, the current phasor measurement would have to be on the substation transformer, which behaves like the the voltage source for the Thevenin equivalent model.

8.2 Bus Injection Power-Voltage Sensitivity Sign Changes

Figure 8.1: The power-voltage sensitivities for bus 2 of a two bus network with $y_{12} = 1j$, plotted against the voltage setpoint for bus 2 in polar coordinates. $V_1 = 1$.

Circumstances in which the sign of the power-voltage sensitivity is state-dependent are of particular interest for power system control applications. That is, at one voltage, injecting real or reactive power results in the voltage increasing, while at another voltage, the same change in real or reactive power injection results in voltage decreasing. We investigate this phenomenon for a two bus network, considering three different values for the line admittance—1) $y_{12} = 1j$, 2) $y_{12} = 0.5 + 0.5j$, and 3) $y_{12} = 0.1 + 1j$. The two bus network can be thought of as a Thevenin-equivalent network for a bus in a larger network. We use (8.1) to determine the SIPS.

We focus on the relationship between real power injections and voltage magnitude (the upper-left images in Fig.s 8.1–8.3. When the line is purely inductive (Fig. 8.1), the sign of the power-voltage sensitivity can be either positive or negative, depending on the linearization
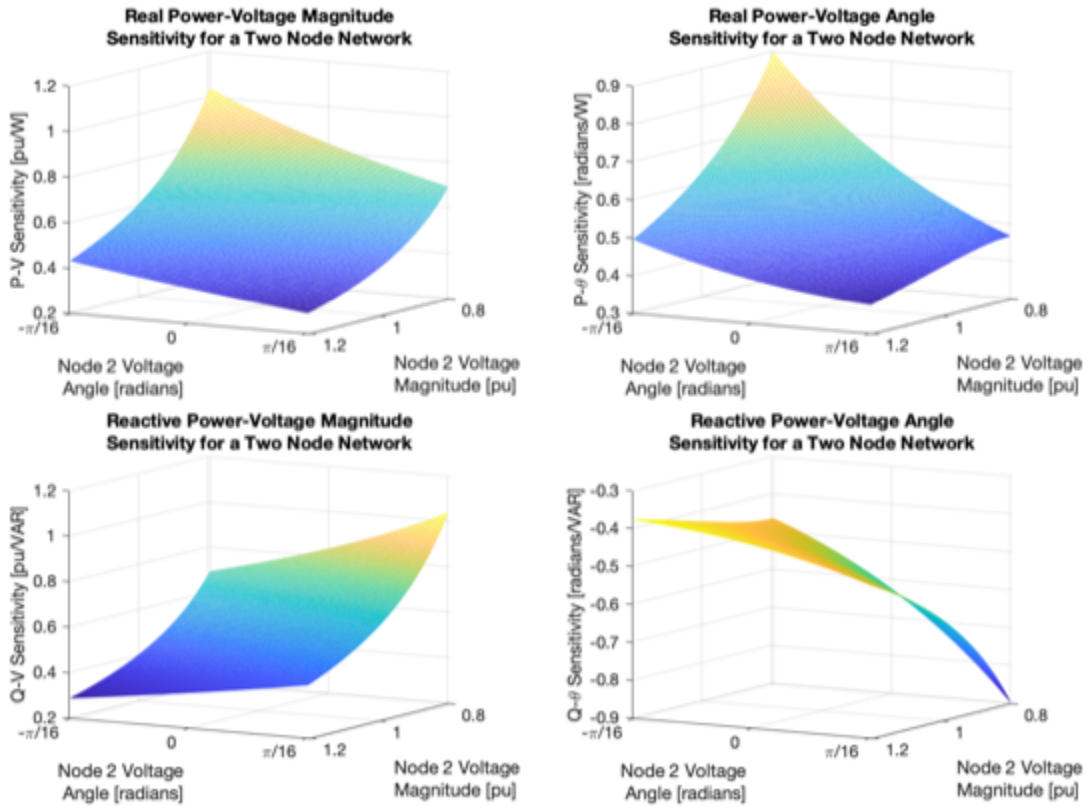
Figure 8.2: The power-voltage sensitivities for bus 2 of a two bus network with $y_{12} = 0.5 + 0.5j$, plotted against the voltage setpoint for bus 2 in polar coordinates. $V_1 = 1$.

setpoint. When the line is equally resistive and inductive (Fig. 8.2), the power-voltage sensitivity is always positive. And when the line is significantly more resistive than inductive (Fig. 8.3), the sign of the power-voltage sensitivity can be either positive or negative.

These plots support general standard grid-operation practices—for transmission networks, in which the line impedances are generally dominated by the inductance, the voltage magnitude is primarily controlled using reactive power injections. As shown by Fig.s 8.1–8.3, the voltage magnitude is more sensitive to reactive power injections than real power injections. Furthermore, the reactive power-voltage magnitude sensitivity is positive regardless of Bus 2's voltage magnitude or angle.

Fig.s 8.1–8.3 give insight into circumstances in which real power injections are used to control voltage magnitude—if the network is primarily inductive, it is a good idea to estimate the power-voltage sensitivity using (8.1) or (8.2) before making control decisions. The upper-left diagram in Fig. 8.3 demonstrates that the real power-voltage magnitude sensitivity for a primarily inductive network can be negative when the local bus voltage is low and the local

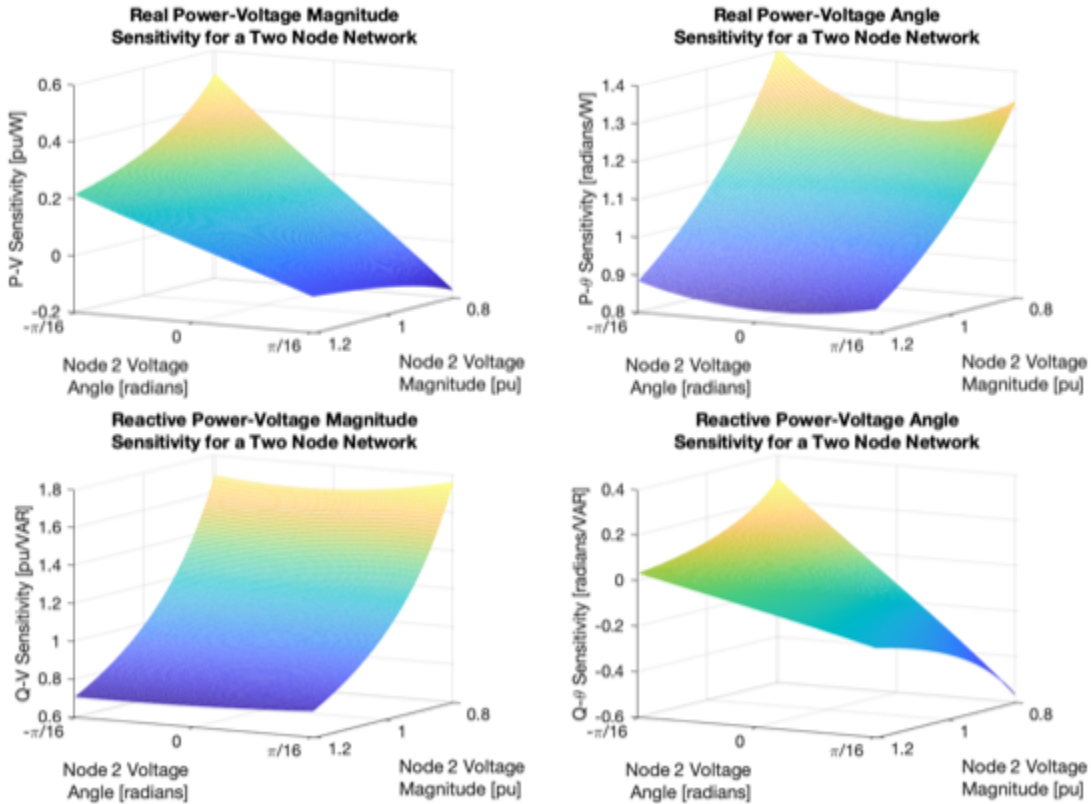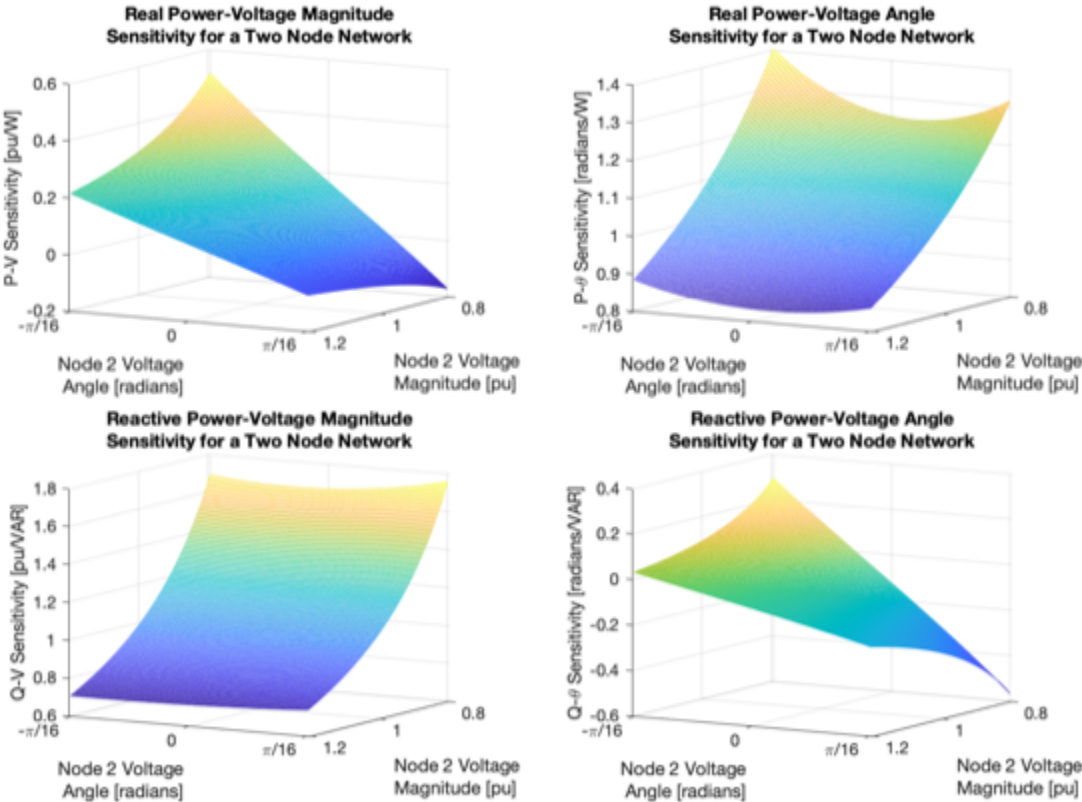Figure 8.3: The power-voltage sensitivities for bus 2 of a two bus network with $y_{12} = 0.1 + 1j$, plotted against the voltage setpoint for bus 2 in polar coordinates. $V_1 = 1$.

voltage angle is leading the substation angle. If this circumstance is encountered and the negative sensitivity is not take into account, using real power to control the local voltage magnitude could destabilize the network.

Chapter 9

Distribution Network Feedback Optimization with Nullspace-based Power Flow Linearization

As outlined in the introduction, traditionally distribution networks have not been actively controlled. Distribution network operators have avoided constraint violations by limiting the connections onto the network. Actively managing distribution networks will reduce the cost of electrifying the transportation and heating sectors by using grid capacity in a dynamic and efficient manner.

Optimal Power Flow (OPF) is the industry standard for controlling electric networks. Given a cost function, the grid state, and the system constraints, OPF determines the optimal operating state using an optimization method such as gradient descent. On transmission networks, OPF determines how much power each generator produces at every moment in the year.

Theoretically, OPF could also be applied to distribution networks. However, OPF is not well-suited for distribution networks because of the following challenges:

1. distribution network models are often inaccurate or unavailable,

2. distribution networks have many constraints and many DERs to control, and

3. the distribution network state changes quickly and in a manner that is hard to accurately predict.

Challenge 1 is addressed in Chapters 2 and 3. Challenges 2 and 3 highlight the computational challenges of applying OPF to distribution networks. 2 points out that distribution networks have many constraints (e.g. the maximum power that can flow through the substation transformer, the line flow constraints, and the voltage magnitude constraints at the end of the feeder lines) and could have many controlled distributed resources (e.g. the charging power for each EV charging station on the network) that control decisions must be made

for. Challenge 3 points out that distribution networks are stochastic systems and, therefore, online control will outperform open-loop control decisions made in advance. Challenge 3 also points out that distribution network states shift rapidly. In order for online control decisions to be implemented while they are still relevant, the decisions must be made quickly.

Full, centralized OPF is not well suited for distribution networks because of the computational challenges described by Challenges 2 and 3. A number of alternatives to standard centralized OPF have been proposed [130]. Feedback Optimization (FO), recently proposed and developed in [93], [131]–[145], addresses the computational issues of OPF by only taking one optimization step at each timestep. Once the optimization step is determined, the output is sent to the actuators and the resultant measurements are "fed back" to the FO engine, which takes another optimization step to determine the actuation for the next timestep. FO saves computation by 1) just taking one optimization step, not solving the full optimization at each timestep, and 2) using the physical infrastructure to "solve" power flow. In addition to the computational complexity concerns, the feedback nature of FO provides robustness to external disturbances and model mismatch that standard open loop optimization (e.g. OPF) does not. Thus, FO is well-suited for distribution networks.

9.1 Feedback Optimization implemented with Linearized Output Projected Gradient Descent

There are a number of algorithms which can be used to implement FO. We focus on Linearized Output Projected Gradient descent (LOPG) [132], [135], [146], [147] because it is simple and effective. We define the following variables:

ν is the output,

ρ is the input,

∇ is the gradient operator with respect to ρ,

$\phi(\rho)$ is the input-output equation, $\nu = \phi(\rho)$,

Φ is the input-output sensitivity matrix,

C is the input inequalities constraint matrix,

c is the input inequalities constraint vector,

D is the output inequalities constraint matrix, and

d is the output inequalities constraint vector.

LOPG finds the input adjustment $\alpha\delta\rho$ such that $\rho + \alpha\delta\rho$ does not violate the system constraints and $\delta\rho$ is as close to $\nabla J(\rho, \nu)$ as possible. $\delta\rho$ is found using the following convex optimization:

$$\begin{aligned}
\arg\min_{\omega} \quad & \left\| \omega + \begin{bmatrix} I & \Phi^T \end{bmatrix} \nabla J(\rho, \nu)^T \right\| \\
\text{subject to} \quad & C(\nu + \alpha\omega) \leq c, \\
& D(\nu + \alpha\Phi\omega) \leq d,
\end{aligned} \tag{9.1}$$

Using (9.1), we define the LOPG algorithm.

Algorithm 4: Linearized Output Projected Gradient Descent (LOPG) with NPFL

At each timestep t:

1. Measure/estimate the system inputs $\rho(t)$ and outputs $\nu(t)$.

2. Calculate Φ using NPFL.

3. Run (9.1) to determine $\delta\rho(t+1)$.

4. Send $\rho(t+1) = \rho(t) + \alpha\delta\rho(t+1)$ to the actuators.

LOPG requires three externally-provided inputs—the cost function $J(\rho, \nu)$, measurements/estimates of the outputs ν (and sometimes the inputs ρ) from the previous timestep, and the sensitivity between the inputs and outputs of the system Φ. For a distribution network with a single slack bus, the power-voltage sensitivity is given by (7.6). (7.6) can be used for voltage magnitude constraints. (7.3) can be used if the substation transformer power flow is constrained. If the line flows are constrained, the power injection-to-line flow sensitivities[1] can be built from the power-voltage sensitivity matrix (7.6) and the weighted bus incidence matrix [153]. (7.34) can be used if the aggregate network losses are included in the objective function.

Ongoing research is implementing LOPG with all of these sensitivities to control EV charging on a full-scale model of Bay Area [154]. This research includes HELICS cosimulation [155], [156] of both a BEAM transportation network simulation [157], [158] and a PyDSS simulation of the distribution and subtransmission networks [159], [160]. The LOPG Feedback Optimization Controller is implemented on the full, unbalanced, three-phase network model.

Linearized Output Projected Gradient Feedback Optimization for Slack Bus Power Injection Constraints

We demonstrate one way in which Algorithm 4 can be used to optimize the service provided by a distribution network, while not damaging the distribution network infrastructure. In particular, we consider a distribution network with a substation power flow constraint, which could be determined by the substation transformer or another piece of equipment.

The substation power flow constraint manifests as a slack bus power injection constraint for the distribution network model. Fig. 9.1 demonstrates a circumstance in which the

---

[1]The power injection-to-line flow sensitivities are equivalent to Injection Shift Factors (ISFs) and closely related to Power Transfer Distribution Factors (PTDFs) commonly used to control transmission networks [103], [148]–[152].

substation power injection minimum is set to -0.3. This constraint limits the maximum amount of power that is allowed to flow into Bus 1.[2]

LOPG requires a sensitivity between the input δp_2, and the constrained output, δp_1. The naive power balance assumption, which applies only to a network without shunt admittances at the nominal voltage (Section 7.4), is that $\delta p_1 = -\delta p_2$. Fig. 9.1a plots the LOPG step that makes this assumption. As is evident from the figure, with this incorrect input-output sensitivity, the LOPG will be overly-cautious and find a suboptimal point.

Alternatively, the LOPG algorithm could also project the $\omega \, (\delta p_2)$ vector onto the power flow manifold orthogonally. However, as demonstrated by Fig. 9.1b, this method also creates issues. The orthogonal projection onto the power flow manifold also adjusts δp_2, reducing it slightly. Thus, if the $\omega \, (\delta p_2)$ vector is actuatued, the resultant δp_1 injection will violate the constraint.

Finally, Fig. 9.1c demonstrates an LOPG step which uses NPFL to determine the sensitivity between the input δp_2 and the output δp_1. The β direction is vertical, which corresponds to the behavior of the true system—the slack bus adjusts its power injection in order to maintain power balance. Also, for this stepsize, the tangent plane is close to the true power flow manifold. Thus, the LOPG step is neither too conservative nor too aggressive when LOPG uses NPFL.

Of course, the intuition is simple for two-bus networks with a single slack bus, and hardcoded rules could be applied to achieve the same results. Algorithm 4 can, however, be extended to networks with arbitrary numbers of actuators and constraints, and arbitrary numbers of slack buses or β power-balancing directions, without requiring hardcoded rules or external information.

---

[2]The analysis in this section applies to a positive injection constraint as well. The negative constraint fit on the plot better than the positive constraint.

(a) Assuming $\delta\hat{p}_1 = -\delta p_2$

(b) Orthogonally projecting ω onto the power flow manifold

(c) Using NPFL to determine the slack bus power injection sensitivity
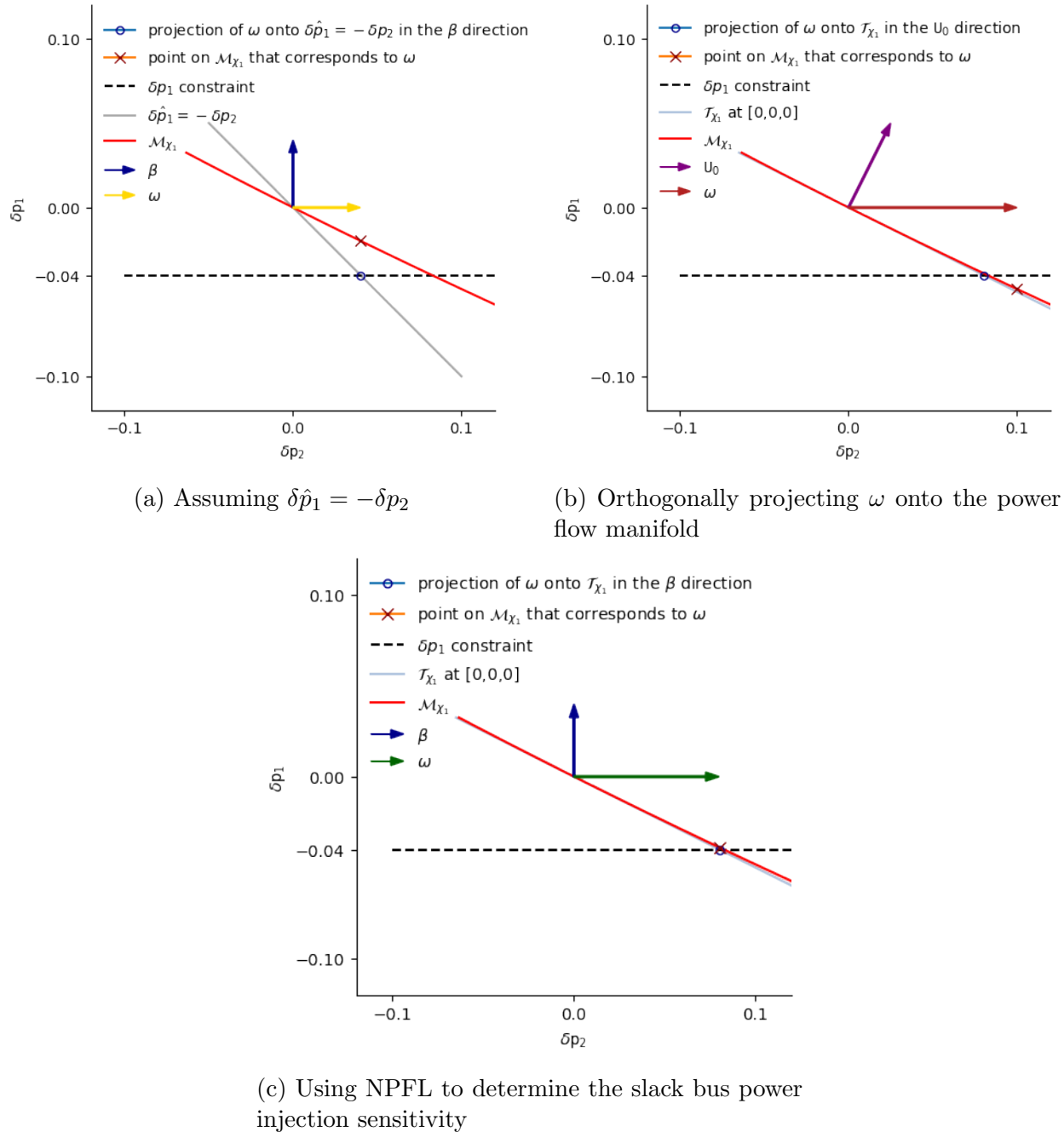
Figure 9.1: Linearized Output Projected Gradient Feedback Optimization for a network with power injection constraints on the slack bus

Bibliography

[1] E. Larson, C. Greig, J. Jenkins, *et al.*, *The report — net-zero america project*, `https://netzeroamerica.princeton.edu/the-report`, (Accessed on 08/25/2022).

[2] A. Madduri, M. Foudeh, P. Phillips, and A. Gupta, *Advanced strategies for demand flexibility management and customer der compensation*, `https://www.cpuc.ca.gov/-/media/cpuc-website/divisions/energy-division/documents/demand-response/demand-response-workshops/advanced-der---demand-flexibility-management/ed-white-paper---advanced-strategies-for-demand-flexibility-management.pdf`, (Accessed on 08/25/2022).

[3] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Transactions on Smart Grid*, vol. 8, pp. 2926–2936, 6 2017.

[4] C. M. Roberts, C. M. Shand, K. W. Brady, E. M. Stewart, A. W. Mcmorran, and G. A. Taylor, "Improving Distribution Network Model Accuracy using Impedance Estimation from Micro-Synchrophasor Data," pp. 1–5, 2016.

[5] K. Brady, "Uses and limitations of micro-synchrophasor measurements in distribution grid management," *PhD diss., Master's thesis, EECS Department, University of California, Berkeley*, 2016.

[6] R. Arghandeh, M. Gahr, A. von Meier, G. Cavraro, M. Ruh, and G. Andersson, "Topology detection in microgrids with micro-synchrophasors," *arXiv*, Jan. 2015.

[7] G. Cavraro, R. Arghandeh, K. Poolla, and A. von Meier, "Data-driven approach for distribution network topology detection," *2015 IEEE Power Energy Society General Meeting*, Jan. 2015.

[8] R. Sevlian, J. Yu, Y. Liao, *et al.*, "Vader: Visualization and analytics for distributed energy resources," *arXiv*, Aug. 2017.

[9] A. Sharon, R. Levy, Y. Cohen, A. Haiut, A. Stroh, and D. Raz, "Automatic network topology analysis," *U.S. Patent 6,205,122*, Mar. 2001.

[10] F. Wu and W. Liu, "Detection of topology errors by state estimation (power systems)," *IEEE Transactions on Power Systems*, vol. 4, pp. 176–183, 1 1989.

[11] A. Gomez-Exposito and A. Abur, "Power system state estimation: Theory and implementation," *CRC press*, 2004.

[12] A. Monticelli, "Fast decoupled state estimator," *State Estimation in Electric Power Systems*, pp. 313–342, 1999.

[13] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," *2012 IEEE Power and Energy Society General Meeting*, pp. 1–8, 2012.

[14] L. Marco, F. Blaabjerg, and R. Teodorescu, "Grid impedance estimation via excitation of lcl-filter resonance," *IEEE Transactions on Industry Applications*, vol. 43, pp. 1401–1407, 5 2007).

[15] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," *In Decision and Control (CDC), 2013 IEEE 52nd Annual Conference*, pp. 1659–1664, 2013.

[16] Y. Yuan, O. Ardakanian, S. Low, and C. Tomlin, "On the Inverse Power Flow Problem," vol. 022, pp. 1–32, 2016. arXiv: `1610.06631`. [Online]. Available: `http://arxiv.org/abs/1610.06631`.

[17] O. Ardakanian, Y. Yuan, V. Wong, *et al.*, "On identification of distribution grids," *IEEE Transactions on Control of Network Systems*, 2019.

[18] N. Saitou and M. Nei, "The neighbor-joining method: A new method for reconstructing phylogenetic trees," *Molecular biology and evolution*, vol. 4, pp. 406–425, 4 1987.

[19] R. D. Durbin and S. R. Eddy, "Biological sequence analysis: Probabilistic models of proteins and nucleic acids," 1998.

[20] R. Dechter and J. Pearl, "Network-based heuristics for constraint-satisfaction problems," *Artificial Intelligence*, pp. 370–425, 1988.

[21] P. Zwiernik, "Latent tree models," *arXiv*, 2017.

[22] N. Zhang, "Hierarchical latent class models for cluster analysis," *Journal of Machine Learning Research*, vol. 5, pp. 697–723, 6 2004.

[23] Y. Wang, N. Zhang, and T. Chen, "Latent tree models and approximate inference in bayesian networks," *Journal of Artificial Intelligence Research*, vol. 32, pp. 879–900, 2008.

[24] S. Harmeling and C. Williams, "Greedy learning of binary latent trees," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 1087–1097, 6 2011.

[25] G. Elidan and N. F., "Learning hidden variable networks: The information bottleneck approach," *Journal of Machine Learning Research*, vol. 6, pp. 81–127, Jan. 2005.

[26] T. Griffiths, C. Kemp, and J. Tenenbaum, "Bayesian models of cognition," 2008.

[27] M. Choi, V. Tan, A. Anandkumar, and A. Willsky, "Learning latent tree graphical models," *Journal of Machine Learning Research*, May 2011.

[28] S. Park, D. Deka, and M. Chertkov, "Exact topology and parameter estimation in distribution grids with minimal observability," *2018 Power Systems Computation Conference (PSCC)*, Jun. 2018.

[29] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I*, vol. 60, pp. 150–163, 1 Jan. 2013.

[30] L. Lovasz, "Eigenvalues of graphs," 2007. [Online]. Available: `http://web.cs.elte.hu/~lovasz/eigenvals-x.pdf`.

[31] G. Kron, "Tensor analysis of networks," 1939.

[32] O. Ardakanian, Y. Yuan, R. Dobbe, A. von Meier, S. Low, and C. Tomlin, "Event detection and localization in distribution grids with phasor measurement units," *2017 IEEE Power and Energy Society General Meeting*, Jul. 2017.

[33] K. Moffat, M. Bariya, and A. von Meier, "Network impedance estimation for microgrid control using noisy synchrophasor measurements," *2018 IEEE Workshop on Control and Modeling for Power Electronics (COMPEL)*, Jun. 2018.

[34] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, pp. 12–19, 1 Feb. 2011.

[35] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, and H. Cao, "Characterizing and quantifying noise in pmu data," *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5, 2016.

[36] J. F. Grcar, "Optimal sensitivity analysis of linear least squares.," *Lawrence Berkeley National Laboratory, Report*, 2003.

[37] D. Robinson and L. Foulds, "Comparison of phylogenetic trees," *Mathematical Biosciences*, vol. 53, pp. 131–147, 1-2 Feb. 1981.

[38] F. C. Schweppe and J. Wildes, "Power system static-state estimation, part i: Exact model," *IEEE Transactions on Power Apparatus and systems*, no. 1, pp. 120–125, 1970.

[39] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC press, 2004.

[40] A. Monticelli, *State estimation in electric power systems: a generalized approach*. Springer Science & Business Media, 2012.

[41] O. Alsac, N. Vempati, B. Stott, and A. Monticelli, "Generalized state estimation," *IEEE Transactions on power systems*, vol. 13, no. 3, pp. 1069–1075, 1998.

[42] A. Abur, H. Kim, and M. Celik, "Identifying the unknown circuit breaker statuses in power networks," *IEEE Transactions on Power Systems*, vol. 10, no. 4, pp. 2029–2037, 1995.

[43] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875–3883, 2016.

[44] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312–2322, 2018.

[45] E. Lourenço, K. Clements, and A. S. Costa, "Geometrically-based hypothesis testing for topology error identification," in *Proc. of the 14th PSCC*, sn, 2002, pp. 24–28.

[46] E. M. Lourenço, E. P. Coelho, and B. C. Pal, "Topology error and bad data processing in generalized state estimation," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3190–3200, 2014.

[47] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *2012 IEEE Power and Energy Society General Meeting*, IEEE, 2012, pp. 1–8.

[48] A. Monticelli, C. Murari, and F. F. Wu, "A hybrid state estimator: Solving normal equations by orthogonal transformations," *IEEE Transactions on Power Apparatus and Systems*, no. 12, pp. 3460–3468, 1985.

[49] L. Lin, L. Jasa, E. Ambikairajah, *et al.*, "A hybrid state estimation scheme for power systems," in *Asia-Pacific Conference on Circuits and Systems*, IEEE, vol. 1, 2002, pp. 555–558.

[50] A. Phadke, J. Thorp, R. Nuqui, and M. Zhou, "Recent developments in state estimation with phasor measurements," in *2009 IEEE/PES Power Systems Conference and Exposition*, IEEE, 2009, pp. 1–7.

[51] I. Džafić, R. A. Jabr, and T. Hrnjić, "Hybrid state estimation in complex variables," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5288–5296, 2018.

[52] F. F. Wu, "Power system state estimation: A survey," *International Journal of Electrical Power & Energy Systems*, vol. 12, no. 2, pp. 80–87, 1990.

[53] N. Shivakumar and A. Jain, "A review of power system dynamic state estimation techniques," in *2008 Joint International Conference on Power System Technology and IEEE Power India Conference*, IEEE, 2008, pp. 1–6.

[54] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[55] A. Dubey and S. Chakrabarti, "An unscented kalman filter based hybrid state estimator considering conventional and pmu measurements," in *2016 IEEE 6th International Conference on Power Systems (ICPS)*, IEEE, 2016, pp. 1–6.

[56] G. Valverde and V. Terzija, "Unscented kalman filter for power system dynamic state estimation," *IET generation, transmission & distribution*, vol. 5, no. 1, pp. 29–37, 2011.

[57] P. S. Maybeck, *Stochastic models, estimation, and control.* Academic press, 1982, ch. 10.

[58] P. D. Hanlon and P. S. Maybeck, "Multiple-model adaptive estimation using a residual correlation kalman filter bank," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 393–406, 2000.

[59] I. Hwang, H. Balakrishnan, and C. Tomlin, "State estimation for hybrid systems: Applications to aircraft tracking," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 5, pp. 556–566, 2006.

[60] Y. Hao, Z. Xiong, F. Sun, and X. Wang, "Comparison of unscented kalman filters," in *2007 International Conference on Mechatronics and Automation*, IEEE, 2007, pp. 895–899.

[61] A. S. Debs and R. E. Larson, "A dynamic estimator for tracking the state of a power system," *IEEE Transactions on Power Apparatus and Systems*, no. 7, pp. 1670–1678, 1970.

[62] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.

[63] *Reds: Repository of distribution systems*, `http://www.dejazzer.com/reds.html`, Accessed: 2021-09-26.

[64] *Pecan street data set*, `https://www.pecanstreet.org/dataport/about/`, Accessed: 2020-09-30.

[65] R. D. Zimmerman, C. E. Murillo-Sánchez, and D. Gan, "Matpower: A matlab power system simulation package," *Manual, Power Systems Engineering Research Center, Ithaca NY*, vol. 1, pp. 10–7, 1997.

[66] M. Bariya, K. Moffat, and A. Von Meier, "Empirical noise estimation in distribution synchrophasor measurements," in *2019 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA)*, IEEE, 2019, pp. 1–7.

[67] I. J. Perez-Arriaga, J. D. Jenkins, and C. Batlle, "A regulatory framework for an evolving electricity sector: Highlights of the mit utility of the future study," *Economics of Energy & Environmental Policy*, vol. 6, no. 1, pp. 71–92, 2017.

[68] K. A. Horowitz, A. Jain, F. Ding, B. Mather, and B. Palmintier, "A techno-economic comparison of traditional upgrades, volt-var controls, and coordinated distributed energy resource management systems for integration of distributed photovoltaic resources," *International Journal of Electrical Power & Energy Systems*, vol. 123, p. 106 222, 2020.

[69] S. Gill, I. Kockar, and G. W. Ault, "Dynamic optimal power flow for active distribution networks," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 121–131, 2013.

[70] M. Nick, R. Cherkaoui, and M. Paolone, "Optimal allocation of dispersed energy storage systems in active distribution networks for energy balance and grid support," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2300–2310, 2014.

[71] R. Hidalgo, C. Abbey, and G. Joós, "A review of active distribution networks enabling technologies," in *IEEE PES General Meeting*, IEEE, 2010, pp. 1–9.

[72] M. Z. Liu, L. F. Ochoa, and S. H. Low, "On the implementation of opf-based setpoints for active distribution networks," *IEEE Transactions on Smart Grid*, 2021.

[73] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "On the need for communication for voltage regulation of power distribution grids," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1111–1123, 2019.

[74] N. Cohn, "Some aspects of tie-line bias control on interconnected power systems [includes discussion]," *Transactions of the American institute of electrical engineers. Part III: power apparatus and systems*, vol. 75, no. 3, pp. 1415–1436, 1956.

[75] P. Kumar, D. P. Kothari, *et al.*, "Recent philosophies of automatic generation control strategies in power systems," *IEEE transactions on power systems*, vol. 20, no. 1, pp. 346–357, 2005.

[76] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A probabilistic framework for reserve scheduling and n-1 security assessment of systems with high wind power penetration," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3885–3896, 2013.

[77] I. S. C. C. 21, "Ieee standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std*, pp. 1547–2018, 2018.

[78] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *52nd IEEE Conference on Decision and Control*, IEEE, 2013, pp. 4329–4334.

[79] A. T. Procopiou and L. F. Ochoa, "On the limitations of volt-var control in pv-rich residential lv networks: A uk case study," in *2019 IEEE Milan PowerTech*, IEEE, 2019, pp. 1–6.

[80] W. Nacmanson, "Advanced planning of pv-rich distribution networks–deliverable 6: Consolidation of findings," 2021.

[81] A. F. Hoke, P. Gotseff, M. Emmanuel, N. D. Wunder, and J. I. Giraldez Miner, "Estimating customer impact of volt-watt using only smart meter voltage data," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2019.

[82] D. Chathurangi, U. Jayatunga, S. Perera, A. Agalgaonkar, and T. Siyambalapitiya, "Comparative evaluation of solar pv hosting capacity enhancement using volt-var and volt-watt control strategies," *Renewable Energy*, 2021.

[83] M. Rylander, M. J. Reno, J. E. Quiroz, *et al.*, "Methods to determine recommended feeder-wide advanced inverter settings for improving distribution system performance," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, IEEE, 2016, pp. 1393–1398.

[84] M. Bello, D. Montenegro-Martinez, B. York, and J. Smith, "Optimal settings for multiple groups of smart inverters on secondary systems using autonomous control," *IEEE Transactions on Industry Applications*, vol. 54, no. 2, pp. 1218–1223, 2017.

[85] S. Yoshizawa, Y. Yanagiya, H. Ishii, *et al.*, "Voltage-sensitivity-based volt-var-watt settings of smart inverters for mitigating voltage rise in distribution systems," *IEEE Open Access Journal of Power and Energy*, vol. 8, pp. 584–595, 2021.

[86] K. Baker, A. Bernstein, E. Dall'Anese, and C. Zhao, "Network-cognizant voltage droop control for distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2098–2108, 2017.

[87] A. K. Jain, K. Horowitz, F. Ding, *et al.*, "Dynamic hosting capacity analysis for distributed photovoltaic resources—framework and case study," *Applied Energy*, vol. 280, p. 115 633, 2020.

[88] A. von Meier, E. L. Ratnam, K. Brady, K. Moffat, and J. Swartz, "Phasor-based control for scalable integration of variable energy resources," *Energies*, vol. 13, no. 1, p. 190, 2020.

[89] K. Moffat, J. Pakshong, L. Chu, *et al.*, "Phasor based control with the distributed, extensible grid control platform," in *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, IEEE, 2021, pp. 1–5.

[90] K. Moffat and A. von Meier, "Linear quadratic phasor control of unbalanced distribution networks," in *2021 IEEE Madrid PowerTech*, IEEE, 2021, pp. 1–6.

[91] D. K. Molzahn and L. A. Roald, "Towards an ac optimal power flow algorithm with robust feasibility guarantees," in *2018 Power Systems Computation Conference (PSCC)*, IEEE, 2018, pp. 1–7.

[92] D. Molzahn and L. A. Roald, "Grid-aware versus grid-agnostic distribution system control: A method for certifying engineering constraint satisfaction," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[93] S. Bolognani and F. Dörfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2015, pp. 402–409.

[94] I. A. Hiskens and R. J. Davy, "Exploring the power flow solution space boundary," *IEEE transactions on power systems*, vol. 16, no. 3, pp. 389–395, 2001.

[95] K. Moffat, "Local power-voltage sensitivity and thevenin impedance estimation from phasor measurements," in *2021 IEEE Madrid PowerTech*, IEEE, 2021, pp. 1–6.

[96] K. Moffat and A. von Meier, "Linear quadratic phasor control of unbalanced distribution networks," in *2021 IEEE Madrid PowerTech*, IEEE, 2021, pp. 1–6.

[97] D. K. Molzahn, I. A. Hiskens, *et al.*, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.

[98] K. Moffat, M. Bariya, and A. von Meier, "Network impedance estimation for microgrid control using noisy synchrophasor measurements," in *2018 IEEE 19th Workshop on Control and Modeling for Power Electronics (COMPEL)*, IEEE, 2018, pp. 1–6.

[99] J. Ward and H. Hale, "Digital computer solution of power-flow problems [includes discussion]," *Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems*, vol. 75, no. 3, pp. 398–404, 1956.

[100] K. Yamane, "New methods for load flow calculation without any swing bus," Ph.D. dissertation, 1971.

[101] R. Baldick, "Course notes for ee394v restructured electricity markets: Locational marginal pricing," 2018.

[102] S. V. Dhople, Y. C. Chen, A. Al-Digs, and A. D. Domınguez-Garcıa, "Reexamining the distributed slack bus," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4870–4879, 2020.

[103] B. Stott, J. Jardim, and O. Alsaç, "Dc power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.

[104] S. V. Dhople, S. S. Guggilam, and Y. C. Chen, "Linear approximations to ac power flow in rectangular coordinates," in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2015, pp. 211–217.

[105] T. J. Overbye, X. Cheng, and Y. Sun, "A comparison of the ac and dc power flow models for lmp calculations," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, IEEE, 2004, 9–pp.

[106] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of dc power flow for active power flow analysis," in *IEEE Power Engineering Society General Meeting, 2005*, IEEE, 2005, pp. 454–459.

[107] Y. Qi, D. Shi, and D. Tylavsky, "Impact of assumptions on dc power flow model accuracy," in *2012 North American Power Symposium (NAPS)*, IEEE, 2012, pp. 1–6.

[108] K. Dvijotham and D. K. Molzahn, "Error bounds on the dc power flow approximation: A convex relaxation approach," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, IEEE, 2016, pp. 2411–2418.

[109] R. Kaye and F. Wu, "Analysis of linearized decoupled power flow approximations for steady-state security assessment," *IEEE transactions on circuits and systems*, vol. 31, no. 7, pp. 623–636, 1984.

[110] D. Shirmohammadi, H. W. Hong, A. Semlyen, and G. Luo, "A compensation-based power flow method for weakly meshed distribution and transmission networks," *IEEE Transactions on power systems*, vol. 3, no. 2, pp. 753–762, 1988.

[111] G.-X. Luo and A. Semlyen, "Efficient load flow for large weakly meshed networks," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1309–1316, 1990.

[112] D. Rajicic, R. Ackovski, and R. Taleski, "Voltage correction power flow," *IEEE Transactions on Power Delivery*, vol. 9, no. 2, pp. 1056–1062, 1994.

[113] C. S. Cheng and D. Shirmohammadi, "A three-phase power flow method for real-time distribution system analysis," *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 671–679, 1995.

[114] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.

[115] J. A. Taylor, *Convex optimization of power systems*. Cambridge University Press, 2015.

[116] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Voltage regulation algorithms for multiphase power distribution grids," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3913–3923, 2015.

[117] M. Sankur, "Optimal control of commercial office battery systems, and grid integrated energy resources on distribution networks," Ph.D. dissertation, UC Berkeley, 2017.

[118] M. D. Sankur, R. Dobbe, E. Stewart, D. S. Callaway, and D. B. Arnold, "A linearized power flow model for optimization in unbalanced distribution systems," *arXiv preprint arXiv:1606.04492*, 2016.

[119] S. H. Low, "Convex relaxation of optimal power flow—part i: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.

[120] S. Bolognani and S. Zampieri, "On the existence and linear approximation of the power flow solution in power distribution networks," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 163–172, 2015.

[121] R. Baldick, "Dc power flow in rectangular coordinates," in *Proc. DIMACS Workshop Energy Infrastruct., Des. Stability Resilience*, 2013.

[122] S. S. Guggilam, E. Dall'Anese, Y. C. Chen, S. V. Dhople, and G. B. Giannakis, "Scalable optimization methods for distribution networks with high pv integration," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2061–2070, 2016.

[123] A. Bernstein, C. Wang, E. Dall'Anese, J.-Y. Le Boudec, and C. Zhao, "Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 5832–5843, 2018.

[124] S. Misra, D. K. Molzahn, and K. Dvijotham, "Optimal adaptive linearizations of the ac power flow equations," in *2018 Power Systems Computation Conference (PSCC)*, IEEE, 2018, pp. 1–7.

[125] M. Hohmann, J. Warrington, and J. Lygeros, "Optimal linearizations of power systems with uncertain supply and demand," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1504–1512, 2018.

[126] J. R. Martı, H. Ahmadi, and L. Bashualdo, "Linear power-flow formulation based on a voltage-dependent load model," *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1682–1690, 2013.

[127] H. Ahmadi, J. R. Martı, and A. von Meier, "A linear power flow formulation for three-phase distribution systems," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 5012–5021, 2016.

[128] *Keith moffat's introduction to power systems notes*, `https://www.keithmoffat.com/IntrEPS/IntrEPS_Lecture8_MaxPowerTransfer.pdf`, Accessed: 2022-09-07.

[129] K. Moffat, "Local power-voltage sensitivity and thevenin impedance estimation from phasor measurements," in *2021 IEEE Madrid PowerTech*, IEEE, 2021, pp. 1–6.

[130] D. K. Molzahn, F. Dörfler, H. Sandberg, *et al.*, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

[131] D. Bertsekas and J. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Athena Scientific, 2015.

[132] V. Häberle, A. Hauswirth, L. Ortmann, S. Bolognani, and F. Dörfler, "Non-convex feedback optimization with input and output constraints," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 343–348, 2020.

[133] M. Picallo, S. Bolognani, and F. Dörfler, "Closing the loop: Dynamic state estimation and feedback optimization of power grids," *Electric Power Systems Research*, vol. 189, p. 106 753, 2020.

[134] E. Dall'Anese, A. Simonetto, S. Becker, and L. Madden, "Optimization and learning with information streams: Time-varying algorithms and applications," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 71–83, 2020.

[135] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Optimization algorithms as robust feedback controllers," *arXiv preprint arXiv:2103.11329*, 2021.

[136] G. Bianchin, J. Cortés, J. I. Poveda, and E. Dall'Anese, "Time-varying optimization of lti systems via projected primal-dual gradient flows," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 1, pp. 474–486, 2021.

[137] G. Belgioioso, D. Liao-McPherson, M. H. de Badyn, S. Bolognani, J. Lygeros, and F. Dörfler, "Sampled-data online feedback equilibrium seeking: Stability and tracking," in *2021 60th IEEE Conference on Decision and Control (CDC)*, IEEE, 2021, pp. 2702–2708.

[138] D. Liao-McPherson, E. C. Balta, A. Rupenyan, and J. Lygeros, "On robustness in optimization-based constrained iterative learning control," *IEEE Control Systems Letters*, 2022.

[139] S. H. Low and D. E. Lapsley, "Optimization flow control. i. basic algorithm and convergence," *IEEE/ACM Transactions on networking*, vol. 7, no. 6, pp. 861–874, 1999.

[140] C. Zhao, U. Topcu, and S. H. Low, "Optimal load control via frequency measurement and neighborhood area communication," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3576–3587, 2013.

[141] N. Li, C. Zhao, and L. Chen, "Connecting automatic generation control and economic dispatch from an optimization view," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 254–264, 2015.

[142] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 942–952, 2016.

[143] S. Menta, A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Stability of dynamic feedback optimization with applications to power systems," in *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2018, pp. 136–143.

[144] M. Colombino, E. Dall'Anese, and A. Bernstein, "Online optimization as a feedback controller: Stability and tracking," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 422–432, 2019.

[145] A. Bernstein and E. Dall'Anese, "Real-time feedback-based optimization of distribution grids: A unified approach," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1197–1209, 2019.

[146] L. Ortmann, A. Hauswirth, I. Caduff, F. Dörfler, and S. Bolognani, "Experimental validation of feedback optimization in power distribution grids," *Electric Power Systems Research*, vol. 189, p. 106 782, 2020.

[147] G. Hotz and l. ortman lukas, "Online feedback optimization for emergency power system operation," M.S. thesis, ETH Zurich, Automatic Control Laboratory, 2021.

[148] R. Baldick, "Variation of distribution factors with loading," *IEEE Transactions on Power Systems*, vol. 18, no. 4, pp. 1316–1323, 2003.

[149] M. Liu and G. Gross, "Effectiveness of the distribution factor approximations used in congestion modeling," in *Proceedings of the 14th Power Systems Computation Conference, Seville, 24–28 June 2002*, Citeseer, 2002.

[150] C. Duthaler, M. Emery, G. Andersson, and M. Kurzidem, *Analysis of the use of ptdf in the ucte transmission grid 2018*.

[151] Y. C. Chen, A. D. Domınguez-Garcıa, and P. W. Sauer, "Measurement-based estimation of linear sensitivity distribution factors and applications," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1372–1382, 2013.

[152] Y. C. Chen, S. V. Dhople, A. D. Domınguez-Garcıa, and P. W. Sauer, "Generalized injection shift factors," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2071–2080, 2016.

[153] A. Bernstein and E. Dall'Anese, "Linear power-flow models in multiphase distribution networks," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, IEEE, 2017, pp. 1–6.

[154] N. Panossian, M. Muratori, B. Palmintier, A. Meintz, T. Lipman, and K. Moffat, "Challenges and opportunities of integrating electric vehicles in electricity distribution systems," *Current Sustainable/Renewable Energy Reports*, pp. 1–14, 2022.

[155] B. Palmintier, D. Krishnamurthy, *et al.*, *Hierarchical engine for large-scale infrastructure co-simulation (helics)*, `https://helics.org/`.

[156] B. Palmintier, D. Krishnamurthy, P. Top, S. Smith, J. Daily, and J. Fuller, "Design of the helics high-performance transmission-distribution-communication-market co-simulation framework," in *2017 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, IEEE, 2017, pp. 1–6.

[157] C. Sheppard, R. Waraich, A. Campbell, A. Pozdnukov, and A. R. Gopal, "Modeling plug-in electric vehicle charging demand with beam: The framework for behavior energy autonomy mobility," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Tech. Rep., 2017.

[158] M. Muratori, P. Jadun, B. Bush, *et al.*, "Future integrated mobility-energy systems: A modeling perspective," *Renewable and Sustainable Energy Reviews*, vol. 119, p. 109 541, 2020.

[159] A. Latif, *Pydss, national renewable energy laboratory*, `https://nrel.github.io/PyDSS/index.html`.

[160] N. Panossian, T. Elgindy, B. Palmintier, and D. Wallison, "Synthetic, realistic transmission and distribution co-simulation for voltage control benchmarking," in *2021 IEEE Texas Power and Energy Conference (TPEC)*, IEEE, 2021, pp. 1–5.

Appendix A

The No-Jargon Introduction

Note to reader:
This "Introduction" is intended to be exactly that—an introduction to the work in this dissertation for a reader who is not already familiar with power systems research. It provides an explanation for the work in this dissertation without assuming that the reader has a prior understanding of power systems, or a familiarity with the jargon of the field of power systems research.

Power systems, or "the grid," use electricity[1] to transfer energy from generators to loads so the loads can work. "Generators" are anything that injects power into the grid.[2] "Loads" are anything that intentionally takes power from the grid. Examples of loads include a light that is on, a toaster that is toasting, or an electric vehicle that is charging. "Work" has a technical meaning in physics ("the application of force along a displacement") that is

---

[1]Water and mountains are an imperfect-but-helpful analogy for electricity. The grid can be thought of as a hypothetical mountain that is used to transmit work from one location to another location. Generators use an external source of power to push water up the mountain to a higher elevation. The water then flows laterally across the mountain through a river to the location where the load is. When it gets to the location where the load is, the water flows down the mountain through a wheel/turbine, powering the load. The power delivered to the load is given by multiplying the river's elevation on the mountain where the load is by the amount of water flowing. In this explanation, the riverbed is analogous to the wires on the grid, the elevation of the riverbed is analogous to electric voltage, and the river is analogous to electric current.

This analogy is intended to help develop intuition for why electric power is the product of voltage and current. In the water-and-mountain analogy, if the load wanted more power, the generator could push more water up the mountain to flow through the riverbed to the load (i.e., send more current at the same voltage), or it could push the same amount of water to a riverbed at a higher elevation (i.e., send the same amount of current at a higher voltage).

[2]Traditionally, generators have been large machines that produce electric power for distribution by spinning a magnet inside coils of wires that are attached to a grid. The magnet spins because it is attached to turbine blades that are pushed by steam or water. Most of the electric power on the grid today is provided by power plants that burn natural gas or coal to produce steam that pushes turbine blades. But there are other types of generators that do not rotate magnets inside of coils of wire, such as solar panels and batteries. Solar panels and batteries are also "generators."

different fom its meanings in everyday language. The technical distinctions of "work" are not important for power systems analysis, however; what is important is that loads take power from the grid. "Power" is the rate at which energy is transferred from generators to loads. Thus, power is a measurement that is made at each instant, whereas energy is the cumulative sum of power over a period of time.[3] All the power that loads take from the grid comes from generators. The total generation and total load[4] must be approximately equal at all moments in time. If this balance is disrupted for a significant amount of time, the grid will have a "blackout" and not be able to transfer power from the generators to the loads until the grid is restarted.

Power systems deliver power from generators to loads using power lines.[5] The electric power flowing into, out of, or through any location on the grid is given by multiplying the voltage at the location by the current flowing into, out of, or through the location at each moment in time.[6] Thus, the amount of electric power flowing into, out of, or through a given location on the grid can be increased in two ways: by increasing the voltage at the location, or by increasing the current flowing into, out of, or through the location.

The grid's physical infrastructure constrains the amount of current that can flow on a line as well as the maximum voltage at which it is safe to operate the grid.[7] If there is too much current flowing in the lines, the lines become too hot. These are called "thermal," or "flow" constraints. If the line voltage is too high, then the voltage difference between the power lines can damage equipment. If the grid voltage is too low, then the voltage on the grid can suddenly collapse to zero, causing a blackout. Thus, electric grids are operated in such a way that both the line flows and the voltages stay within upper and lower bounds. The grid is also operated so that the grid frequency stays within upper and lower bounds, but the work in this dissertation focuses on thermal and voltage constraints, not frequency constraints.

---

[3]To illustrate the difference between power and energy, consider the example of two generators with one attached to a solar panel and the other attached to a battery. The two generators have the same capacity, but the generator attached to the solar panel only generates electricity when the sun is shining on the solar panel, whereas the battery generates power whenever it is told to (if the battery has not been fully discharged). Thus, if the sun shines for the first five minutes of an hour and then is blocked by rain clouds, the solar panel will generate full electricity for only the first five minutes then turn off for the rest of the hour. The battery, however, may generate electricity for the full hour. The solar panel and battery may thus be able to provide the same amount of electric power for the grid, but the solar panel will provide less energy over the course of the hour than the battery.

[4]In addition to the loads, some power is "lost" as heat on the power lines when power flows from the generators to the loads. We define the "total load" on the grid to be the sum of all the loads attached to the network and all of the losses on the network.

[5]Power lines are large metal (aluminum) wires that can transfer large amounts of electricity.

[6]See footnote 1.

[7]Continuing the water analogy to understand grid constraints, the flow/thermal constraints are the maximum amount of current that can flow through a riverbed/wire before bad things happen (i.e., overheating the wire). Voltage constraints are defined as the range of elevations at which it is safe for a riverbed to be. For reasons specific to electric grids, there are both upper and lower limits on voltage.

Electric power cannot be told where to flow—once power is on the grid, it flows from the generators to the loads according to the laws of physics. If the loads on the grid are not managed by grid operators, as is traditionally assumed, the power/current flows on the grid can only be altered by changing which generators supply power to the grid at each moment in time.

The Internet and relatively inexpensive, powerful computers have made it possible to manage loads in real time, upending the traditional assumption that only generators can be controlled. It is now possible to avoid thermal and voltage constraints by reducing the power drawn by certain loads on the system in real time. For example, it is now possible to tell electric vehicles in each neighborhood that there is too much load in that neighborhood at a given moment and the power lines have hit their thermal limit. The utility company that manages the grid can offer to pay the electric vehicle owners if they reduce their vehicle charging rate until the power/current flows on the lines are below their limits.

In the future there might be thousands of loads on the grid which can be managed to avoid constraint violations. Thus, it will be necessary to have tools that automatically make decisions for each load based on high-level objectives that are set by the utility company/grid operator. It is worth noting that the same tools that automatically make decisions for loads can also be used to automatically make decisions for generators distributed across the grid, such as rooftop solar panels or Tesla Powerwall batteries. By thinking of the distributed generators as "negative loads," the tools used to automatically manage loads can also be used to control the power output of distributed generators without any modification to the tools.

The final distinction that will help in understanding the application of this dissertation is the distinction between transmission networks and distribution networks.[8] Transmission networks span large areas, such as the western United States and parts of Canada and Mexico, with thick wires that transfer a lot of power at "very high" voltage (e.g., hundreds of kilovolts (kV)). Distribution networks attach to the transmission grid via substations, which step the voltage down to just "high" voltage (e.g., tens of kV). A single transmission system can provide power to thousands of distribution networks.

While the same physics apply to transmission and distribution networks, the two types of systems are operated in different ways. Typically, transmission systems have been actively managed to avoid constraint violations—that is, care is taken when grid operators decide how much power to generate at each generator so that the power/current flows on the transmission lines do not violate thermal constraints.

Distribution networks, on the other hand, are not actively managed. The power flows from the substation to the loads without any coordination of the loads and generators on the distribution network. Thermal and voltage constraints are avoided on distribution networks by 1) building the distribution infrastructure for the most challenging possible circumstance,

---

[8]In this dissertation I use "distribution network," "distribution grid," and "distribution system" interchangeably. There are subtle differences between some of these terms, but the distinctions are not important for the work in this dissertation.

and 2) limiting new connections onto the grid.[9] Requiring distribution networks to support the most challenging possible circumstance without any coordination of applicable loads and generators results in unnecessary interconnection restrictions and the need to install larger, more expensive equipment than would be necessary if the loads and generators were coordinated.

This dissertation is focused on the emerging practice of managing/controlling the power consumption or generation of distributed energy resources to avoid violating distribution network thermal and voltage constraints. "Managing distributed energy resources" could mean managing how much power an electric vehicle charging station uses, how much power a Tesla Powerwall battery injects back into the grid, or reducing the power generated by a rooftop solar panel. Managing distributed energy resources to avoid network constraints is not something that has been done historically,[10] but it may become commonplace in the coming years as the amount of electric power the grid delivers from generators to loads increases. Managing distributed energy resources to avoid network constraints will result in more affordable, more accessible, and more efficient electricity service.

In the future, we expect there to be many electric vehicles, batteries, rooftop solar panels, and other distributed energy resources attached to the grid. Managing these resources will be challenging because the decisions for many resources must be made at once, and they must be made in a "fair" manner.[11] Furthermore, distribution network operators do not generally know exactly how a change in the power consumption or generation from one resource will affect a given voltage or line flow that is in danger of violating a constraint. Consider an example distribution network which supplies power to both an electric vehicle charging station in town and homes outside of town. If the voltage is too low at the homes outside of town, it would be helpful for distribution network operators to know exactly how sensitive the voltage at those homes is to changes in the power consumption at the electric vehicle charging station. At present, distribution network operators do not know the sensitivities

---

[9]For example, if an individual household or business wanted to build a powerful electric vehicle charging station, the utility that owns and operates the distribution network would limit if/where the household/business could build the charging station based on where the grid has capacity for large new loads. The capacity for new loads is based on the single moment of the year that is most challenging for the thermal and voltage constraints. For the rest of the year, the line flows and network voltages are below thermal and voltage constraints.

Similarly, not every house is allowed to install new rooftop solar panels that will inject power back into the grid. The decision to install new solar panels that will inject power back into the grid must be cleared with the utility, which is allowed to say that a household/business is not allowed to inject power back into the grid.

[10]There are a number of utility programs and companies that control distributed energy resources, but these programs/companies focus on reducing load or increasing generation so that the total load and the total generation on the grid are equal, not avoiding distribution network constraints. "Virtual Power Plant" is a term commonly used to describe controlling distributed energy resources so that the total generation matches the total load.

[11]The question of what is "fair" allocation of grid resources is critically important, but outside of the scope of this dissertation, which focuses on the engineering challenges. The work in this dissertation seeks to provide tools to operate the grid more efficiently, and to make the trade-offs of policy decisions transparent.

that describe how a change in the power injections will change the voltages and line flows on a distribution network.[12] This lack of knowledge prevents distribution network operators from actively managing the resources on distribution networks.

This dissertation introduces new equations that describe how changes in the power injections of the resources attached to the grid will affect the voltages, line flows, and other quantities of interest at any location on a given distribution network. These equations that describe the "sensitivity" of the grid require a model of the grid that includes how the grid wires are connected, what the wires are made of, and how long the wires are.

This dissertation also introduces methods for learning grid models from sensor measurements. As it turns out, it is not always possible to learn a useful grid model from sensor measurements. This dissertation explains why, and it proposes a method for determining how accurate the learned grid model can be expected to be based on the quality of the sensors.

Finally, this dissertation introduces two different tools for automatically managing distributed energy resources to avoid thermal and voltage constraint violations on distribution networks. The two tools, "Voltage Phasor Control" and "Feedback Optimization," make automatic decisions based on an objective function. An objective function is a mathematical expression that codifies the grid-operation goals.[13] Thus, the objective function can be crafted so that the Voltage Phasor Control tool and/or the Feedback Optimization tool control distributed energy resources in a fair manner. The tools themselves do not specify an objective function—that is left up to the policy-makers, grid operators, and people who use the grid.

---

[12]Distribution network operators have an estimate or intuition for the sensitivities of the voltages and line flows, but do not generally have an accurate mathematical expression for them because there are challenging nuances in accurately determining the sensitivities. Also, traditional distribution network operation has not required exact sensitivities. Accurate mathematical expressions for the sensitivities are helpful for automatically managing the power injections of distributed energy resources to avoid distribution network constraint violations.

[13]Examples of goals that can be expressed as objective functions include "deliver as much electric power as possible," "give all participants the same amount of electric power," or some combination of the two.

Appendix B

Public Comment on the California Public Utility Commission's Proceeding R2207005

California Public Utilities Commission Proceeding
Public Comments

Proceeding Number: R2207005
Filed Date: 14-JUL-22
Status: Active
Filer Name List: CPUC
Description: Order Instituting Rulemaking to Advance Demand Flexibility Through Electric Rates.
Assignment List: ALJ: Stephanie Wang (Assigned Jul 19, 2022)
COMMISSIONER: Alice Reynolds (Assigned Jul 19, 2022)

Total Comments: 1

Keith Moffat   **Berkeley, CA   94703**

Dear Commission,

I am a PhD student in Electrical Engineering at UC Berkeley. My research focuses on distribution network optimization and electricity pricing. In my research for my Ph.D. at UC Berkeley I have come to the independent conclusion that real-time electricity pricing is the correct way to engage customer flexibility and to provide customers with a cheaper, more equitable electricity service. Thus, I am pleased to see Proceeding R2207005 from the California Public Utility Commission (CPUC) and would like to offer my full support for the proposed "CalFUSE roadmap" in the CPUC Energy Division's Demand Flexibility Whitepaper.

The CalFUSE roadmap is forward-looking and will create growth in California's electricity sector. Growth of the electricity sector is necessary to meet California's greenhouse gas emission goals. Currently, the growth of the electricity sector is impeded by the static electricity rates that are presented to customers by the Utility Distribution Companies (UDCs) and other Load Serving Entities (LSEs)?see the "Replacing Demand Charges" Section in my extended comment, which can be found at the link at the end of this comment. The CalFUSE roadmap outlines sequential steps for UDCs and LSEs to provide customers with real-time electricity rates that will benefit the entire electricity sector.

The CalFUSE roadmap's call for using real-time electricity rates to spur electrification in California is prescient. Real-time electricity rates, however, only address electricity consumption/production inefficiencies for consumers/producers that are already connected to the electric grid. In addition to real-time electricity rates, including flexible interconnection agreements for certain subsectors (DC fast charging stations and medium-to-large renewable energy plants) will benefit electricity consumers, electricity producers, and UDCs/LSEs by enabling new grid connections. The "Flexible Interconnections for Certain Subsectors" Section in my extended comment provides more detail for why real-time electricity rates should be accompanied by flexible interconnections for certain subsectors.

The time is now for real-time electricity rates and flexible interconnections. Emerging technologies have matured to the point that real-time electricity rates and flexible interconnections are both necessary and practical. The technologies that make real-time electricity rates and flexible interconnections necessary include electric vehicles and electric home heating/cooling, which are posed to double electricity consumption in California, as well as intermittent renewable generation such as solar and wind power. The technologies that make real-time electricity rates and flexible interconnections practical include home energy management products, as well as internet and cloud computation infrastructure, that make the determination and dissemination of the real-time prices and power limits practical at scale.

Regarding the technical aspects of implementing real-time electricity rates and flexible interconnections, I would like to make the point that there is active research in this field. Research resources may be tapped to sort out the technical challenges that arise with

real-time electricity rates and flexible interconnections.

My extended comment, which can be found at the link below, describes
1. the case for flexible interconnections for certain subsectors,
2. the case for replacing demand charges with scarcity pricing for capacity cost recovery,

3. the technical details of implementing delivery scarcity pricing.

Regards,

Keith Moffat

Ph.D. Candidate
Electrical Engineering and Computer Science Department
UC Berkeley
e: keithm@berkeley.edu
w: keithmoffat.com

Extended comment link:

https://www.keithmoffat.
com/KeithMoffat_Comment_CPUCproceedingR2207005_DemandFlexibilityThroughElect
ricRates_8.15.22.pdf

Aug 15, 2022 1:51 pm

B.1 Extended Comment on Proceeding R2207005

8/15/22

Dear Commission,

I, Keith Moffat, offer the following comments on **Proceeding R2207005**, "**Order Instituting Rulemaking to Advance Demand Flexibility Through Electric Rates**" related to the recent CalFUSE proposal [2].

Executive Summary

I am a PhD student in Electrical Engineering at UC Berkeley. My research focuses on distribution network optimization and electricity pricing. In my research for my Ph.D. at UC Berkeley I have come to the independent conclusion that real-time electricity pricing is the correct way to engage customer flexibility and to provide customers with a cheaper, more equitable electricity service. Thus, I am pleased to see Proceeding R2207005 from the California Public Utility Commission (CPUC) and would like to offer my full support for the proposed "CalFUSE roadmap" in the CPUC Energy Division's Demand Flexibility Whitepaper.

The CalFUSE roadmap is forward-looking and will create growth in California's electricity sector. Growth of the electricity sector is necessary to meet California's greenhouse gas emission goals. Currently, the growth of the electricity sector is impeded by the static electricity rates that are presented to customers by the Utility Distribution Companies (UDCs) and other Load Serving Entities (LSEs)—see the "Replacing demand charges" comment below. The CalFUSE roadmap outlines sequential steps for UDCs and LSEs to provide customers with real-time electricity rates that will benefit the entire electricity sector.

The CalFUSE roadmap's call for using real-time electricity rates to spur electrification in California is prescient. Real-time electricity rates, however, only address electricity consumption/production inefficiencies for consumers/producers that are already connected to the electric grid. In addition to real-time electricity rates, including flexible interconnection agreements for certain subsectors (DC fast charging stations and medium-to-large renewable energy plants) will benefit electricity consumers, electricity producers, and UDCs/LSEs by enabling new grid connections. The "Flexible Interconnections for certain subsectors" comment below provides more detail for why real-time electricity rates should be accompanied by flexible interconnections for certain subsectors.

The time is now for real-time electricity rates and flexible interconnections. Emerging technologies have matured to the point that real-time electricity rates and flexible interconnections are both necessary and practical. The technologies that make real-time electricity rates and flexible interconnections necessary include electric vehicles and electric home heating/cooling, which are posed to double electricity consumption in California, as well as intermittent renewable generation such as solar and wind power. The technologies that

make real-time electricity rates and flexible interconnections practical include home energy management products, as well as internet and cloud computation infrastructure, that make the determination and dissemination of the real-time prices and power limits practical at scale.

Regarding the technical aspects of implementing real-time electricity rates and flexible interconnections, I would like to make the point that there is active research in this field. Research resources may be tapped to sort out the technical challenges that arise with real-time electricity rates and flexible interconnections.

The rest of this comment discusses 1. the case for flexible interconnections for certain subsectors, 2. the case for replacing demand charges with scarcity pricing for capacity cost recovery, 3. the technical details of implementing delivery scarcity pricing. I submit these comments because I have been studying the best way to implement distribution scarcity pricing and flexible interconnections over the course of my Ph.D. While I focus on technical details in this comment, I would like to reiterate that the broader goals of implementing real-time pricing and flexible interconnections are what is important at this time. The implementation specifics, which I focus on in comment 3, can be determined later in the implementation process.

Comment 1: Flexible Interconnections for Certain Subsectors

Real-time electricity rates only address electricity consumption/production inefficiencies for consumers/producers that are already connected to the electric grid. In addition to real-time electricity rates, including flexible interconnection agreements for certain subsectors (e.g. the DC fast charging station subsector and the medium-to-large renewable energy plant subsector) will benefit electricity consumers, electricity producers, and UDCs/LSEs by enabling new grid connections.

It is necessary to build a significant amount of new infrastructure to support the electrification of the transportation sector. For example, California will have to build many DC Fast Charging stations, which pull orders of magnitude more electricity from the grid than standard Level 1 and Level 2 charging stations. The current Interconnection Capacity Analysis (ICA) process for siting new charging stations impedes electrification because the ICA process rejects new charging station sites that should not be rejected.

The current ICA process only allows new connections to the grid if it is impossible for the new connection to overload the grid infrastructure. Thus, the current ICA process is based on the premise that any load on the network should be allowed to extract as much power as it would like at any moment in time. Operating a grid with this max-power-at-any-time premise leads to overbuilt grid infrastructure, as the aggregate max load (generation) on the system occurs at very few moments in the year.

While appropriate for many types of electricity customers such as residential or standard commercial customers, the max-power-at-any-time premise is not appropriate for electricity subsectors that 1) use (produce) a lot of electricity, and 2) are flexible in when they can consume (produce) electricity. DC fast charging stations are an example of an electricity

subsector that matches criteria (1) and (2), as are medium-to-large intermittent renewable generation plants.

Flexible interconnection is an alternative to max-power-at-any-time interconnection. Flexible interconnection allows the utility to reduce the amount of power that a consumer can take from (producer can inject onto) the grid to avoid grid constraint violations. The flexible interconnection can be based on real time grid measurements or time of use.

Flexible interconnection has grassroots support in the research community and has been implemented in industry. Smarter Grid Solutions released a flexible interconnection product in 2010 that enabled the UK grid to connect more wind power, without requiring expensive infrastructure upgrades.

In addition to introducing new electricity rates, the CalFUSE roadmap should include a requirement that UDCs support flexible interconnections for the DC fast charging station subsector and medium-to-large renewable energy plant subsector. Incorporating flexible interconnections for these sectors will increase electrification, reduce carbon emissions, use grid infrastructure more efficiently, and provide a cheaper electricity service to customers.

Comment 2: Replacing Demand Charges

Demand charges have distorted the electricity market, prevented electrification, and encouraged customers to pursue inefficient consumption patterns. Thus, demand charges result in unnecessary costs for electricity consumers and unnecessary greenhouse gas emissions. Replacing demand charges with scarcity pricing for capacity cost recovery, as proposed in CalFUSE, will provide significant value to customers by allowing customers to save money by aligning their consumption patterns with grid infrastructure constraints.

Demand charges are per-kW (power-based, rather than energy-based) rates that are intended to be a proxy for the infrastructure investment incurred by a grid-connected load or generator. Unfortunately, demand charges are a poor proxy for infrastructure investment and severely impede the construction of new electric resources such as electric vehicle charging stations.

Regarding cost causation, demand charges are correlated with the capacity/maximum power flow of "local" grid infrastructure. However, individual peak energy use is not well-correlated with the peak energy use for non-local grid infrastructure. That is, as you get further from the customer, it becomes less likely that the customer's peak power consumption coincides with the peak power flowing through a given grid component. Most of the grid infrastructure falls into this "non-local" category.

Consider the example of an electric vehicle DC fast charging station whose peak energy use occurs at noon. The peak power flowing through the charging station's service drop will be at noon. However, if the charging station is in an area with a lot of solar power, the charging station's power use at noon might reduce the amount of power flowing through the main power lines or substation transformer at noon. Thus, the time of customer power consumption is critical and should be considered when determining the best way to recover capacity costs.

The CalFUSE roadmap proposes replacing demand charges with scarcity pricing for capacity cost recovery. This switch will create significant value for the entire electricity sector, including electricity customers, generators, UDCs, and LSEs. By allowing (a subset of) electricity customers to respond to real-time prices, the (subset of) customers may align their consumption with both the real-time generation cost and the grid infrastructure, which will avoid unnecessary infrastructure replacements and result in cheaper electricity prices for everyone. Cheaper prices will, in turn, lead to the growth of the electricity sector in California.

Comment 3: Implementing Delivery Scarcity Pricing

It will be important to implement scarcity pricing for capacity cost recovery in the correct manner. The CalFUSE roadmap splits the scarcity price framework up into delivery scarcity price, capacity scarcity price, and ramp scarcity price. The delivery scarcity price is focused on recovering the fixed costs that were used to build the distribution network infrastructure (the power lines, transformers, and other pieces of grid equipment). The following comments are focused on the specifics of implementing the delivery scarcity pricing for distribution network capacity cost recovery.

Comment 3a: On Delivery Scarcity Pricing—the Pricing Mechanism

The mechanism that determines the delivery scarcity price must be transparent and fair. The best way to create transparency and fairness is with an algorithm. The algorithm rules and settings can be published publicly and adjusted as necessary to meet the needs of society regarding equity, optimality, and other considerations. Note, it is the algorithm rules and settings that are adjusted by policy makers, not the prices themselves.

Pages 58-60 of the CalFUSE whitepaper outline a method of determining scarcity prices using a quadratic function with hand-picked parameters. While simple, this quadratic-function-with-hand-picked-parameter method has several downsides. One downside is that the process for choosing the parameters could be corrupted and would be the basis for endless debate among stakeholders. A second downside is that the hand-picked parameters cannot be adjusted in real-time to match the grid conditions as they evolve.

Lagrangian-based optimization is an alternative method for determining the delivery scarcity price. Lagrangian-based optimization is used ubiquitously in optimization. For example, the Locational Marginal Prices (LMPs) for the CAISO transmission network are calculated using Lagrangian-based optimization. A Lagrangian-based optimization algorithm would be a better way of determining the delivery scarcity prices than a quadratic-function-with-hand-picked-parameter method because Lagrangian-based optimization automatically adjusts the prices, rather than requiring an exogenous entity to hand-pick parameters.

Comment 3b: On Delivery Scarcity Pricing—Capacity Utilization vs. Constraint-based Pricing

The delivery scarcity pricing method proposed on pages 58-60 of the CalFUSE whitepaper implements a rate adder based on capacity utilization. A rate adder based on capacity utilization would work well when individual customer electricity consumption can be mapped directly to the capacity utilization for an existing piece of grid infrastructure. For example, individual customer energy use can be easily mapped to substation transformer capacity utilization on a radial network that is served by that single substation transformer. Such a scenario corresponds to the example given in the CalFUSE whitepaper.

However, a rate adder based on capacity utilization would not be easy to implement when there is not a clear mapping from individual customer energy use to capacity utilization. For example, if a feeder is served by multiple substation transformers in different locations, it would be more difficult to map individual customer energy use to substation transformer capacity utilization for each transformer.

Another example is related to voltage constraint violations—the voltage on the grid must be kept between minimum and maximum values, according to grid codes. While voltage violations cannot be mapped to capacity utilization of specific, existing infrastructure, voltage violations also demonstrate delivery resource scarcity.

If delivery scarcity pricing is implemented based on (proximity to) constraint violations, rather than the capacity utilization of existing infrastructure, then the delivery scarcity pricing will apply to all types of delivery resource scarcity.

Lagrangian-based optimization, as proposed in Comment 3a, supports constraint-based pricing. Implementing delivery scarcity pricing with Lagrangian-based optimization will produce a system that automatically adjusts prices based on (proximity to) constraint violations, applies to all types of delivery resource scarcity, and does not require hand-picked parameters.

Regards,

Keith Moffat
Ph.D. Candidate
Electrical Engineering and Computer Science Department
UC Berkeley
e: keithm@berkeley.edu
w: keithmoffat.com

Appendix C

Voltage Phasor Control Proofs

C.1 Derivations of the voltage sensitivities in Table 4.2

The sensitivities of the circuits without VPC are straightforward derivatives with respect to the disturbance injection current. With VPC, the PCN current injection at the PCN is a function of the disturbance node injection.

Circuit in row 1 of Table 4.2

The circuit equations are

$$v_1 = 1 - i_{01}z_{10}, \text{ and}$$
$$i_{01} = -(i_1 + i_2 + i_3).$$

Without VPC:

$$\frac{\partial v_1}{\partial i_3} = z_{10}.$$

With VPC:

$$i_2 = \frac{\hat{v}_2 - i_3(z_{10} + z_{21}) - i_1 z_{10} - 1}{z_{10} + z_{21}}. \tag{C.1}$$

Therefore, $\frac{\partial i_2}{\partial i_3} = -1$. From KCL,

$$\frac{\partial i_{01}}{\partial i_3} = -\frac{\partial i_2}{\partial i_3} - 1 = 0.$$

Using the chain rule, we find that

$$\frac{\partial v_1}{i_3} = \frac{\partial v_1}{\partial i_{01}} \frac{\partial i_{01}}{i_3} = 0.$$

Circuit in row 2 of Table 4.2

The circuit equation is

$$v_3 = 1 + i_3 z_{32} + (i_2 + i_3) z_{21} + (i_3 + i_2 + i_1) z_{10}. \tag{C.2}$$

Without VPC:

$$\frac{\partial v_3}{\partial i_1} = z_{10}.$$

With VPC, Eqn. (C.1) still holds. Therefore,

$$\frac{\partial i_2}{\partial i_1} = -\frac{z_{10}}{z_{10} + z_{21}},$$

and, from the circuit equation (C.2),

$$\frac{\partial v_3}{\partial i_1} = \frac{\partial i_2}{\partial i_1} z_{21} + \frac{\partial i_2}{\partial i_1} z_{10} + z_{10}$$
$$= 0.$$

C.2   Derivations of the line flow sensitivities in Table 4.3

The circuit equations are

$$1 - v_2 = i_{01} z_{10} + i_{12} z_{12},$$
$$i_{01} = -i_1 + i_{12}, \text{ and} \tag{C.3}$$
$$i_{12} = -i_2 - i_3. \tag{C.4}$$

When VPC holds v_2 at $\hat{v}_2$, i_{01} and i_{12} are functions of i_1:

$$i_{01} = \frac{1 - \hat{v}_2}{z_{10} + z_{12}} - i_1 \frac{z_{12}}{z_{10} + z_{12}}. \tag{C.5}$$

$$i_{12} = \frac{1 - \hat{v}_2}{z_{10} + z_{12}} + i_1 \frac{z_{01}}{z_{10} + z_{12}}, \tag{C.6}$$

The sensitivities in Table 4.3 come from the partial derivatives of equations (C.3), (C.4), (C.5), and (C.6).

C.3   Proof of Lemma 4.3.1

Proof. From (C.5), $\frac{\partial i_{01}}{\partial i_1} = -a$.

Using the Triangle Inequality,

$$|i_{01}^{\mathrm{VPC}}| - |i_{01}^{\mathrm{bef}}| \leq |\Delta i_1 \frac{\partial i_{01}}{\partial i_1}| = |\Delta i_1||a|.$$

$\square$

A similar proof applies to Lemma 4.3.2.

C.4   Proof of Lemma 4.3.3

Proof. Without VPC, $i_{01}^{\mathrm{ol}} = (i_{01}^{\mathrm{bef}} - \Delta i_1)$, and

$$\begin{aligned}
|i_{01}^{\mathrm{ol}}|^2 &= (i_{01}^{\mathrm{bef}} - \Delta i_1)(i_{01}^{\mathrm{bef}} - \Delta i_1)^* \\
&= |i_{01}^{\mathrm{bef}}|^2 - 2\cos(\theta_{i_{01}^{\mathrm{bef}}} - \theta_{\Delta i_1})|i_{01}^{\mathrm{bef}}||\Delta i_1| + |\Delta i_1|^2.
\end{aligned}$$

With VPC, $i_{01}^{\mathrm{VPC}} = (i_{01}^{\mathrm{bef}} - a\Delta i_1)$, and

$$\begin{aligned}
|i_{01}^{\mathrm{VPC}}|^2 &= (i_{01}^{\mathrm{bef}} - a\Delta i_1)(i_{01}^{\mathrm{bef}} - a\Delta i_1)^* \\
&= |i_{01}^{\mathrm{bef}}|^2 - 2a\cos(\theta_{i_{01}^{\mathrm{bef}}} - \theta_{\Delta i_1})|i_{01}^{\mathrm{bef}}||\Delta i_1| + a^2|\Delta i_1|^2.
\end{aligned}$$

Because Δi_1 increases i_{01} and z_{01} and z_{12} have the same X/R ratio, $|i_{01}^{\mathrm{VPC}}|^2 > |i_{01}^{\mathrm{bef}}|^2$, and

$$0 < -2a\cos(\theta_{i_{01}^{\mathrm{bef}}} - \theta_{\Delta i_1})|i_{01}^{\mathrm{bef}}||\Delta i_1| + a^2|\Delta i_1|^2.$$

Using the fact that $0 < a < 1$,

$$0 < 2(a-1)\cos(\theta_{i_{01}^{\mathrm{bef}}} - \theta_{\Delta i_1})|i_{01}^{\mathrm{bef}}||\Delta i_1| + (a - a^2)|\Delta i_1|^2. \tag{C.7}$$

Taking the difference between $|i_{01}^{\mathrm{ol}}|^2$ and $|i_{01}^{\mathrm{VPC}}|^2$, we get

$$\begin{aligned}
|i_{01}^{\mathrm{ol}}|^2 - |i_{01}^{\mathrm{VPC}}|^2 &= 2(a-1)\cos(\theta_{i_{01}^{\mathrm{bef}}} - \theta_{\Delta i_1})|i_{01}^{\mathrm{bef}}||\Delta i_1| \\
&\quad + (a - a^2)|\Delta i_1|^2 + (1-a)|\Delta i_1|^2. \tag{C.8}
\end{aligned}$$

Substituting (C.7) into (C.8) gives

$$|i_{01}^{\mathrm{ol}}|^2 - |i_{01}^{\mathrm{VPC}}|^2 > (1-a)|\Delta i_1|^2,$$

which is positive because $0 < a < 1$. $\square$

Appendix D

Unsupervised Impedance and Topology Estimation of Distribution Networks

D.1 The subKron Reduction Derivation

The subKron admittance matrix in Eq. (2.4) is derived in section 2.3. We can equivalently derive an expression for the subKron impedance matrix as follows: $V_{\mathcal{A}} - \mathbf{1}V_{\mathcal{A}1}^T = Z_K I_{\mathcal{A}} - \mathbf{1}Z_{K1}^T I_{\mathcal{A}}$. Z_{K1}^T is the first row of Z_K. Since Z_K is symmetric, Z_{K1} is also the first column of Z_K. By conservation of current, $I_{\mathcal{A}1} = -\sum_{i=2}^A I_{\mathcal{A}i}$, so we can rewrite the above as follows:

$$V_{\mathcal{A}} - \mathbf{1}V_{\mathcal{A}1}^T = (Z_K - \mathbf{1}Z_{K1}^T) \begin{bmatrix} -\mathbf{1}^T \\ \mathbb{I}_{(A-1)} \end{bmatrix} \begin{bmatrix} - & I_{\mathcal{A}2} & - \\ & \vdots & \\ - & I_{\mathcal{A}A} & - \end{bmatrix}$$

where $\mathbb{I}_{(A-1)}$ is the $(A-1) \times (A-1)$ identity matrix. Define $\bar{Z} = (Z_K - \mathbf{1}Z_{K1}^T) \begin{bmatrix} -\mathbf{1}^T \\ \mathbb{I}_{(A-1)} \end{bmatrix}$. Then the elements of $\bar{Z} \in \mathbb{C}^{A \times (A-1)}$ are related to the values in Z_K by: $\bar{Z}(i,j) = Z_K(i, j+1) - Z_K(1, j+1) - Z_K(1, i) + Z_K(1, 1)$. Using the symmetry of Z_K, we can see that $\forall j : \bar{Z}(1, j) = 0$: $\bar{Z}(1, j) = Z_K(1, j+1) - Z_K(1, j+1) - Z_K(1, 1) + Z_K(1, 1) = 0$ Therefore, the first row of $\bar{Z}$ is 0, corresponding to the row of 0s in $V_{\mathcal{A}} - \mathbf{1}V_{\mathcal{A}1}^T$. Discarding both rows of zeros reduces the equation dimensionality and leaves us with a square matrix $Z_{sK} \in \mathbb{C}^{(A-1) \times (A-1)}$.

$$\begin{bmatrix} - & (V_{\mathcal{A}2} - V_{\mathcal{A}1})^T & - \\ & \vdots & \\ - & (V_{\mathcal{A}A} - V_{\mathcal{A}1})^T & - \end{bmatrix} = Z_{sK} \begin{bmatrix} - & I_{\mathcal{A}2} & - \\ & \vdots & \\ - & I_{\mathcal{A}A} & - \end{bmatrix}$$

Z_{sK} is the subKron impedance matrix with elements defined by Eq. (2.6). Eq. (2.6) comes directly from the values in $\bar{Z}$. The two subKron forms of Ohm's Law are then: $I_{s\mathcal{A}} = Y_{sK}V_{s\mathcal{A}} \leftrightarrow V_{s\mathcal{A}} = Z_{sK}I_{s\mathcal{A}}$.

D.2 Recursive Grouping and Relaxed Recursive Grouping

RG constructs the radial network recursively, progressively shrinking $\mathcal{O}$ as parent nodes are identified/introduced using lemma D.2.1. For distribution networks, the *initial* set $\mathcal{O}$ is the set $\mathcal{A}$, which includes at least all of the leaf nodes. Once a parent is identified/introduced, the children nodes are removed from $\mathcal{O}$. A toy demonstration of RG is given in Fig. 2.4.

RG determines connectivity relationships between observed nodes by comparing d_{ij} with $\Phi_{ijk} \triangleq d_{ik} - d_{jk}$, for all possible combinations $i, j, k \in \mathcal{O}$. The following Lemma D.2.1 [27] describes the relationship between d_{ij} and Φ_{ijk} for nodes which are parent/child pairs or siblings.

Lemma D.2.1. *(i)* $d_{ij} = \Phi_{ijk}, \forall k \in \mathcal{O} \backslash (i, j)$ *iff i is a leaf node in $\mathcal{O}$, and j is the parent of i.*
(ii) $-d_{ij} < \Phi_{ijk} = \Phi_{ijk'} < d_{ij}, \forall k, k' \in \mathcal{O} \backslash (i, j)$ *iff i and j are leaf nodes in $\mathcal{O}$, and siblings.*

If two or more nodes are siblings without a parent node, a new parent node is added. As new nodes are added, the effective impedances between them and the rest of the network are calculated using linear operations (eqns. $(13) - (14)$ in [27]). Thus, RG is a linear operator.

[27] also introduces "Relaxed Recursive Grouping" for scenarios where the inter-node distance estimates ($\hat{d}$) contain noise/errors. The parent/child criteria in Lemma D.2.1 is replaced with:

$$|\hat{d}_{ij} - \hat{\Phi}_{ijk}| \leq \epsilon, \forall k \in \mathcal{K}_{ij}, \tag{D.1}$$

and the sibling group criteria in Lemma D.2.1 with:

$$\hat{\Lambda}_{ij} \triangleq \max_{k \in \kappa_{ij}}(|\hat{\Phi}_{ijk}|) - \min_{k \in \kappa_{ij}}(|\hat{\Phi}_{ijk}|) \leq \epsilon. \tag{D.2}$$

Eqs. (D.1) and (D.2) introduce a bias, because it is more likely that pair (i, j) will be recognized as a parent/child or sibling pair if $\mathcal{K}_{ij}$ contains fewer nodes. To avoid this, we introduce alternative relaxed RG tests which consider the average $\hat{\Phi}_{ijk}$ value over $k \in \mathcal{K}_{ij}$, rather than the max. The proposed relaxed RG parent/child and sibling tests are precisely the CRG tests (2.9) and (2.10).

In order for relaxed RG to converge, it must select at least two nodes in $\mathcal{O}$ to be a parent/children pair or siblings at each iteration. But with imperfect $\hat{d}$ estimates, it is possible that no nodes in $\mathcal{O}$ satisfy the parent/child or sibling tests for a given ϵ. This can be addressed using the expanding ϵ technique presented in [28]—ϵ (for the current comparison round) is increased when no nodes satisfy the parent/child or sibling tests.

Appendix E

Mathematical Background for Power Flow Manifold Tangent Planes

E.1 Converting Complex-Valued Equations to Real-Valued Equations

$$\langle x \rangle := \begin{bmatrix} \mathrm{Re}(x) \\ \mathrm{Im}(x) \end{bmatrix} \text{ if } x \text{ is a vector,} \tag{E.1}$$

$$\langle M \rangle := \begin{bmatrix} \mathrm{Re}(M) & -\mathrm{Im}(M) \\ \mathrm{Im}(M) & \mathrm{Re}(M) \end{bmatrix} \text{ if } M \text{ is a matrix.} \tag{E.2}$$

The $\langle \cdot \rangle$ conversion works for matrix-vector multiplications. Consider the product of a complex matrix $M = A + Bj$ and a complex vector $x = y + zj$,

$$
\begin{aligned}
Mx &= [A + Bj][y + zj] \\
&= [Ay - Bz] + [Az + By]j.
\end{aligned}
$$

Alternatively, we can compute the product in real space,

$$
\begin{aligned}
\langle Mx \rangle &= \langle M \rangle \langle x \rangle \\
&= \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} \\
&= \begin{bmatrix} Ay - Bz \\ By + Az \end{bmatrix},
\end{aligned}
$$

arriving at the same answer as the complex multiplication after taking $\langle \cdot \rangle$ of the complex multiplication.

The $\langle \cdot \rangle$ conversion works for matrix-matrix multiplications as well. Consider the product of a complex matrix $M = A + Cj$ and another complex matrix $D = R + Ej$,

$$
\begin{aligned}
MD &= [A + Cj][R + Ej] \\
&= [AR - CE] + [AE + CR]j.
\end{aligned}
$$

Alternatively, we can compute the product in real space,

$$
\begin{aligned}
\langle MD \rangle &= \langle M \rangle \langle D \rangle \\
&= \begin{bmatrix} A & -C \\ C & A \end{bmatrix} \begin{bmatrix} R & -E \\ E & R \end{bmatrix} \\
&= \begin{bmatrix} AR - CE & -AE - CR \\ AE + CR & AR - CE \end{bmatrix},
\end{aligned}
$$

once again arriving at the same answer as the complex multiplication after taking $\langle \cdot \rangle$ of the complex multiplication.

Finally, we define the following rules for complex conjugates,

$$\langle \overline{x} \rangle = N \langle x \rangle \text{ if } x \text{ is a vector}, but \tag{E.3}$$

$$\langle \overline{M} \rangle = N \langle M \rangle N \text{ if } M \text{ is a matrix}. \tag{E.4}$$

For $N \in \mathbb{R}^{2n}$ defined as

$$
N := \begin{bmatrix} I_n & 0_n \\ 0_n & -I_n \end{bmatrix}.
$$

The (E.3) and (E.4) equations follow from (E.2) and (E.1).

E.2 Implicit Functions, Manifolds, and Tangent Planes

An implicit function of x is a relation of x defined by setting a function equal to zero. For example,

$$\Phi(x) = 0.$$

Φ defines the submanifold of the implicit function,

$$\mathcal{M} := \{x \mid \Phi(x) = 0\}.$$

The dimensions of x and Φ define the dimension of the submanifold. The Jacobian of $\Phi(x)$ with respect to x is the matrix of partial derivatives,

$$J = \frac{\partial \Phi(x)}{\partial x}.$$

The rows of the Jacobian evaluated at x^* define vector space that is normal to $\mathcal{M}$ at x^*. The tangent plane at x^* is the space of vectors that are orthogonal to the rows of the Jacobian at x^*,

$$\mathcal{T}|_{x^*} := \{x \mid J|_{x^*}\delta x = 0\}.$$

Two-dimensional Example

Let us take an example implicit function Φ of $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$ to provide some intuition,

$$\Phi(x) = x_1 - x_2^2 = 0.$$

The 1-dimensional submanifold of $\mathbb{R}^2$ is defined as

$$\mathcal{M} := \{x \mid \Phi(x) = 0\}.$$

The Jacobian J is given by

$$J = \frac{\partial \Phi(x)}{\partial x} = \begin{bmatrix} 1 & -2x_2^* \end{bmatrix}.$$

The tangent plane at $x^* = \begin{bmatrix} x_1^* & x_2^* \end{bmatrix}$ is given by

$$\mathcal{T}|_{x^*} := \left\{ x \mid \begin{bmatrix} 1 & -2x_2^* \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} = 0 \right\},$$

where $\delta x = x - x^*$. The normal vector (space) to the manifold $\mathcal{M}$ and the tangent vector (plane) is the row vector (space) of J, $\begin{bmatrix} 1 & -2x_2^* \end{bmatrix}^T$. The tangent vector can be written explicitly as

$$\delta x_1 - 2x_2^* \delta x_2 = 0. \tag{E.5}$$

This tangent vector can be compared with the derivative of the explicit function $\phi(x_1)$ derived by setting $x_2 = \phi(x_1)$,

$$\phi(x_1) = \pm\sqrt{x_1}.$$

The derivative of $\phi(x_1)$ with respect to x_1 evaluated at x_1^* is

$$\frac{\partial \phi(x_1)}{\partial x_1} = \pm \frac{1}{2\sqrt{x_1^*}}. \tag{E.6}$$

Plugging $(x_2^*)^2$ in for x_1^* in (E.6) demonstrates that the derivative (E.6) describes the same tangent vector in $\mathbb{R}^2$ as (E.5).

Appendix F

Power Flow Equivalence proofs

F.1   Coupled Linear Power Flow Derivation

We define $\delta v_{lk} = v_l - v_k$, $\delta v_{k1} = v_k - 1$, and $\delta \theta_{lk} = \theta_l - \theta_k$. We also define the $_{l,:}$ subscript to mean "all of the entries corresponding to i." For the set of edge real power flows, $p_{l,:}$ is all of the real power flows from bus i. For a matrix M, $\mathsf{M}_{l,:}$, to be all of the entries of the ith row of M.

CLPF Voltage-Power Sensitivity

Using Approximation 1,

$$s = \text{diag}(v)\bar{i}$$
$$s \approx \bar{i}.$$

Assuming that the network has no shunt elements,[1] we can write an expression for $\text{Re}(i_l)$ as

$$\text{Re}(i_l) = \sum_{k \neq l}^{n} [v_l(\cos(\theta_l)g_{lk} - \sin(\theta_l)b_{lk}) - v_k(\cos(\theta_k)g_{lk} - \sin(\theta_k)b_{lk})].$$

Plugging in the small angle approximations and expanding the voltage magnitudes that are multiplied by angles into $v_l = 1 + \delta v_{l1}$ and $v_k = 1 + \delta v_{k1}$ gives

$$\text{Re}(i_l) \approx \sum_{k \neq l}^{n} [v_l(g_{lk} - \theta_l b_{lk}) - v_k(g_{lk} - \theta_k b_{lk})]$$
$$= \sum_{k \neq l}^{n} \left[(v_l - v_k)g_{lk} - (1 + \delta v_{l1})\theta_l b_{lk} + (1 + \delta v_{k1})\theta_k b_{lk} \right].$$

---

[1]A similiar derivation applies to networks with shunt elements. We include the no-shunt derivation here because it is simpler.

Discarding products of δv and θ leaves

$$\text{Re}(i_l) \approx \sum_{k \neq l}^{n} \left[g_{lk}(v_l - v_k) - b_{lk}(\theta_l - \theta_k) \right]$$

$$= G_{l,:}v - B_{l,:}\theta.$$

Similarly, we derive the linearization for $\text{Im}(i_l)$,

$$\text{Im}(i_l) = \sum_{k \neq l}^{n} [v_l(\sin(\theta_l)g_{lk} + \cos(\theta_l)b_{lk}) - v_k(\sin(\theta_k)g_{lk} + \cos(\theta_k)b_{lk})]$$

$$\approx \sum_{k \neq l}^{n} [v_l(\theta_l g_{lk} + b_{lk}) - v_k(\theta_k g_{lk} + b_{lk})]$$

$$= \sum_{k \neq l}^{n} \left[(v_l - v_k)b_{lk} + (1 + \delta v_{l1})\theta_l g_{lk} - (1 + \delta v_{k1})\theta_k g_{lk} \right]$$

$$\approx \sum_{k \neq l}^{n} \left[b_{lk}(v_l - v_k) + g_{lk}(\theta_l - \theta_k) \right]$$

$$= B_{l,:}v + G_{l,:}\theta.$$

Combining these results gives the CLPF approximation of the power injections in terms of the voltage phasors,

$$\begin{bmatrix} p \\ q \end{bmatrix} \approx \begin{bmatrix} G & -B \\ -B & -G \end{bmatrix} \begin{bmatrix} v \\ \theta \end{bmatrix}.$$

Taking the difference from a given operating point[2] and eliminating the power injections and voltages for the constant voltage/slack bus gives the CLPF voltage-power sensitivity,

$$\begin{bmatrix} \delta p_R \\ \delta q_R \end{bmatrix} = \begin{bmatrix} G_R & -B_R \\ -B_R & -G_R \end{bmatrix} \begin{bmatrix} \delta v_R \\ \delta \theta_R \end{bmatrix}.$$

CLPF Power-Voltage Sensitivity

Taking the inverse of the CLPF voltage-power sensitivity using the complex-matrix equivalence in Appendix E.1,

$$\begin{bmatrix} G_R & -B_R \\ -B_R & -G_R \end{bmatrix}^{-1} = (N\langle Y_R \rangle)^{-1}$$

$$= \langle Z_R \rangle N$$

$$= \begin{bmatrix} R_{\mathcal{R}} & X_{\mathcal{R}} \\ X_{\mathcal{R}} & -R_{\mathcal{R}} \end{bmatrix}.$$

---

[2]Commonly-used operating points include the Flat voltage profile and the "No Load" voltage profile [104], [120].

F.2 Equivalence of Branch Flow and Bus Injection Linearizations

The linearized edge flow equation in v for a edge from node i to node j is

$$
\begin{bmatrix} \delta v_{lk} \\ \theta_{lk} \end{bmatrix} \approx \begin{bmatrix} r_{lk} & x_{lk} \\ x_{lk} & -r_{lk} \end{bmatrix} \begin{bmatrix} p_{lk} \\ q_{lk} \end{bmatrix}
$$

The inverse of this equation is

$$
\begin{bmatrix} p_{lk} \\ q_{lk} \end{bmatrix} \approx \begin{bmatrix} g_{lk} & -b_{lk} \\ -b_{lk} & -g_{lk} \end{bmatrix} \begin{bmatrix} \delta v_{lk} \\ \theta_{lk} \end{bmatrix}
$$

The edge flow equations for all the edges connected to node i are given by:

$$
\begin{bmatrix} p_{l1} \\ \vdots \\ p_{ln} \\ q_{l1} \\ \vdots \\ q_{ln} \end{bmatrix} \approx \begin{bmatrix} g_{l1} & \dots & 0 & -b_{l1} & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & g_{ln} & 0 & \dots & -b_{ln} \\ -b_{l1} & \dots & 0 & -g_{l1} & \dots & 0 \\ & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -b_{ln} & 0 & \dots & -g_{ln} \end{bmatrix} \begin{bmatrix} v_1 - v_l \\ \vdots \\ v_n - v_l \\ \theta_1 - \theta_l \\ \vdots \\ \theta_n - \theta_l \end{bmatrix} \tag{F.1}
$$

$$
\implies \begin{bmatrix} p_{l,:} \\ q_{l,:} \end{bmatrix} \approx \begin{bmatrix} \mathrm{diag}(\mathsf{G}_{l,:}) & -\mathrm{diag}(\mathsf{B}_{l,:}) \\ -\mathrm{diag}(\mathsf{B}_{l,:}) & -\mathrm{diag}(\mathsf{G}_{l,:}) \end{bmatrix} \begin{bmatrix} v^{\delta i} \\ \theta^{\delta i} \end{bmatrix}. \tag{F.2}
$$

Note the difference between G and G. G is *not* the network conductance matrix. Instead, G contains conductances of each network line: $\mathsf{G}(i,k) = g_{lk}$, the conductance of line (i,k). Similarly, B differs from B. If a connection does not exist between nodes i and j, $g_{lk} = b_{lk} = 0$. $\mathsf{G}(i,i)$ and $\mathsf{B}(i,i)$ are both 0 as well, not the shunt admittance.

Using the definiition of the $_{l,:}$ subscript from the beginning of F.1, We have the following relationship between the edge flows from node i and the nodal power injection at node i,

$$
\begin{bmatrix} p_l \\ q_l \end{bmatrix} = -\begin{bmatrix} \mathbf{1}^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{1}^T \end{bmatrix} \begin{bmatrix} p_{l,:} \\ q_{l,:} \end{bmatrix}. \tag{F.3}
$$

We also have the following relationship between the voltage magnitude and angle differences along the edges emanating from node i and the nodal voltage magnitudes and angles,

$$
\begin{bmatrix} v^{\delta i} \\ \theta^{\delta i} \end{bmatrix} = \begin{bmatrix} \mathcal{E}^{(i)} & 0 \\ 0 & \mathcal{E}^{(i)} \end{bmatrix} \begin{bmatrix} v \\ \theta \end{bmatrix}. \tag{F.4}
$$

Here $\mathcal{E}^{(i)} \in \mathbb{Z}^{n \times n}$ is related to the incidence matrix. However, it doesn't only describe the existing edges in the graph, but every possible edge. The jth row of $\mathcal{E}^{(i)}$ contains only two

nonzero entries. Namely, $\mathcal{E}^{(i)}(j,i) = -1$ and $\mathcal{E}^{(i)}(j,j) = 1$. Therefore, the ith column of $\mathcal{E}^{(i)}$ contains all -1's. We can write $\mathcal{E}^{(i)}$ as the sum of two matrices,

$$\mathcal{E}^{(i)} = I - 1^{(i)}.$$

where $1^{(i)} \in \mathbb{Z}^{n \times n}$ is a matrix of all zeros, except the ith column which contains ones. plugging (F.1) and (F.4) into (F.3), allows us to write an equation for the matrix relating the power injection at node i to the nodal voltage magnitudes and angles.

$$\begin{bmatrix} p_l \\ q_l \end{bmatrix} \approx - \begin{bmatrix} \mathbf{1}^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{1}^T \end{bmatrix} \begin{bmatrix} \text{diag}(\mathsf{G}_{l,:}) & -\text{diag}(\mathsf{B}_{l,:}) \\ -\text{diag}(\mathsf{B}_{l,:}) & -\text{diag}(\mathsf{G}_{l,:}) \end{bmatrix} \begin{bmatrix} \mathcal{E}^{(i)} & 0 \\ 0 & \mathcal{E}^{(i)} \end{bmatrix} \begin{bmatrix} v \\ \theta \end{bmatrix}$$

$$\approx - \begin{bmatrix} \mathsf{G}_{l,:} & -\mathsf{B}_{l,:} \\ -\mathsf{B}_{l,:} & -\mathsf{G}_{l,:} \end{bmatrix} \begin{bmatrix} \mathcal{E}^{(i)} & 0 \\ 0 & \mathcal{E}^{(i)} \end{bmatrix} \begin{bmatrix} v \\ \theta \end{bmatrix}.$$

Recall that $\mathsf{G}_{l,:} \in \mathbb{R}^n$ and $\mathsf{B}_{l,:} \in \mathbb{R}^n$ contain conductances and susceptances respectively for each edge (i,k). If the edge (i,k) doesn't exist, the corresponding element of $\mathsf{G}_{l,:} \in \mathbb{R}^n$ and $\mathsf{B}_{l,:} \in \mathbb{R}^n$ will be zero.

$$\begin{aligned} \mathsf{G}_{l,:}\mathcal{E}^{(i)} &= \mathsf{G}_{l,:}I - \mathsf{G}_{l,:}1^{(i)} \\ &= \begin{bmatrix} g_{l1} & \cdots & g_{ln} \end{bmatrix} - \begin{bmatrix} 0 & \cdots & -\sum_{j=1}^n g_{lk} & \cdots & 0 \end{bmatrix} \\ &= \begin{bmatrix} g_{l1} & \cdots & -\sum_{j=1}^n g_{lk} & \cdots & g_{ln} \end{bmatrix} = -G_{l,:}. \end{aligned}$$

Where $G_{l,:}$ is the ith row of network conductance matrix G without any shunt elements. Similarly, we have

$$\mathsf{B}_{l,:}\mathcal{E}^{(i)} = -B_{l,:}$$

putting these results together gives:

$$\begin{bmatrix} p_l \\ q_l \end{bmatrix} \approx \begin{bmatrix} G_{l,:} & -B_{l,:} \\ -B_{l,:} & -G_{l,:} \end{bmatrix} \begin{bmatrix} v \\ \theta \end{bmatrix}.$$

Stacking together p_l and q_l for every node gives CLPF in bus injection form

$$\begin{bmatrix} p \\ q \end{bmatrix} \approx \begin{bmatrix} G & -B \\ -B & -G \end{bmatrix} \begin{bmatrix} v \\ \theta \end{bmatrix}.$$

F.3 NPFL for a Network without Shunt Admittances at Flat voltage

power-voltage Sensitivity

If $u^* = u_{\text{flat}} = 1_n$,

$$R(u_{\text{flat}}) = 1_{2n}$$
$$\langle \text{diag}(u_{\text{flat}})\rangle = 1_{2n}.$$

If, in addition, there are no shunt admittances,

$$Y u_{\text{flat}}) = 0_n.$$

plugging these values into Γ in (5.8) gives

$$\Gamma = \begin{bmatrix} G_{\mathcal{R}} & -B_{\mathcal{R}} \\ -B_{\mathcal{R}} & -G_{\mathcal{R}} \end{bmatrix}.$$

Calculating $_{\chi_1}\Gamma_{\chi_1}$ for (7.6) gives

$$_{\chi_1}\Gamma_{\chi_1} = \begin{bmatrix} G_{\mathcal{R}} & -B_{\mathcal{R}} \\ -B_{\mathcal{R}} & -G_{\mathcal{R}} \end{bmatrix}.$$

Using the result from Appendix F.1,

$$_{\chi_1}^T \psi_{\chi_1} = \begin{bmatrix} R_{\mathcal{R}} & X_{\mathcal{R}} \\ X_{\mathcal{R}} & -R_{\mathcal{R}} \end{bmatrix}.$$

Slack Bus power Inejction

Because $ker(G^T) = 1_n$ and $ker(B^T) = 1_n$

$$_{\Xi_1}\Gamma_{\chi_1} = \Xi_1 \begin{bmatrix} U_+ & U_0 \end{bmatrix} \Sigma v^T$$
$$_{\Xi_1}U_{0,1} = I_2$$
$$_{\Xi_1}U_{0,\chi_1} = \begin{bmatrix} 1_{n-1} & 0 \\ 0 & 1_{n-1} \end{bmatrix}.$$

plugging into (7.3) to calculate $\Xi_1\Omega_1\Xi_1^T$ from Remark 7.3.5 gives

$$\Xi_1\Omega_1\Xi_1^T = \begin{bmatrix} 0 & \begin{bmatrix} -1_{n-1}^T & 0 \\ 0 & -1_{n-1}^T \end{bmatrix} \\ 0 & I_{2n-2} \end{bmatrix}.$$