Load Learning and Topology Optimization for Power Networks

Siddharth Bhela

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

Vassilis Kekatos, Chair

Chen-Ching Liu

Virgilio A. Centeno

Ryan K. Williams

Weijun Xie

May 10, 2019

Blacksburg, Virginia

Keywords: Smart inverters, power flow, state estimation, load learning, grid probing, generic rank, topological observability, topology design, mixed-integer programming

Load Learning and Topology Optimization for Power Networks

Siddharth Bhela

(ABSTRACT)

With the advent of distributed energy resources (DERs), electric vehicles, and demand-response programs, grid operators are in dire need of new monitoring and design tools that help improve efficiency, reliability, and stability of modern power networks. To this end, the work in this thesis explores a generalized modeling and analysis framework for two pertinent tasks: i) learning loads via grid probing; and ii) optimizing power grid topologies for stability. Distribution grids currently lack comprehensive real-time metering. Nevertheless, grid operators require precise knowledge of loads and renewable generation to accomplish any feeder optimization task. At the same time, new grid technologies, such as solar panels and energy storage units are interfaced via inverters with advanced sensing and actuation capabilities. In this context, we first put forth the idea of engaging power electronics to probe an electric grid and record its voltage response at actuated and metered buses to infer non-metered loads. Probing can be accomplished by commanding inverters to momentarily perturb their power injections. Multiple probing actions can be induced within a few tens of seconds. Load inference via grid probing is formulated as an implicit nonlinear system identification task, which is shown to be topologically observable under certain conditions. The analysis holds for single- and multi-phase grids, radial or meshed, and applies to phasor or magnitude-only voltage data. Using probing to learn non-constant-power loads is also analyzed as a special case. Once a probing setup is deemed topologically observable, a methodology for designing probing injections abiding by inverter and network constraints to improve load estimates is provided. The probing task under noisy phasor and non-phasor

data is tackled using a semidefinite-program relaxation. As a second contribution, we also study the effect of topology on the linear time-invariant dynamics of power networks. For a variety of stability metrics, a unified framework based on the H_2-norm of the system is presented. The proposed framework assesses the robustness of power grids to small disturbances and is used to study the optimal placement of new lines on existing networks as well as the design of radial topologies for new networks.

Load Learning and Topology Optimization for Power Networks

Siddharth Bhela

(GENERAL AUDIENCE ABSTRACT)

Increased penetration of distributed energy resources such as solar panels, wind farms, and energy storage systems is forcing utilities to rethink how they design and operate their power networks. To ensure efficient and reliable operation of distribution networks and to perform any grid-wide optimization or dispatch tasks, the system operator needs to precisely know the net load (energy output) of every customer. However, due to the sheer extent of distribution networks (millions of customers) and low investment interest in the past, distribution grids have limited metering infrastructure. Nevertheless, data from grid sensors comprised of voltage and load measurements are readily available from a subset of customers at high temporal resolution. In addition, the smart inverters found in solar panels, energy storage units, and electric vehicles can be controlled within microseconds. The work in this thesis explores how the proliferation of grid sensors together with the controllability of smart inverters can be leveraged for inferring the non-metered loads i.e., energy output of customers that are not equipped with smart inverters/sensors. In addition to the load learning task, this thesis also presents a modeling and analysis framework to study the optimal design of topologies (how customers are electrically inter-connected) for improving stability of our power networks.

Dedication

To my wife, Tithi,

my parents, Sham and Anita

and my brother, Siddheshvar

Acknowledgments

This dissertation [1] has culminated due to the help and kindness of many people that I have had the good fortune of coming in contact with over the years. First and foremost, I owe my deepest gratitude to my Ph.D. advisor, Dr. Vassilis Kekatos, for making me a better researcher, writer, and presenter; for always inspiring me, and for allowing me to benefit from his knowledge and expertise. His mentorship and energy has made my Ph.D. experience productive and stimulating. I would like to thank the members of my Ph.D. advisory committee, Dr. Chen-Ching Liu, Dr. Virgilio Centeno, Dr. Ryan K. Williams, and Dr. Weijun Xie. Their varied perspectives and valuable feedback on my work has helped me gain new insights into my research. I am grateful to all the professors at Virginia Tech whose classes have enriched my understanding of complex subjects.

Without the generous support of Dr. Harsha Veeramachaneni the work in this thesis would be incomplete. Dr. Deepjyoti Deka and Dr. Harsha Nagarajan with whom I had the opportunity to work with at Los Alamos National Laboratory have had great intellectual impact on my research as well. I thank all my friends and colleagues at the PEC lab and at Virginia Tech for their input on my work. The journey of being a graduate student would not have been as fun without their companionship. I would also like to thank my parents, Sham and Anita, and my brother, Siddheshvar, for their unconditional love and never ceasing emotional support. Lastly, this dissertation would not have been possible without my wife, Tithi. I owe her for always believing in me and for being my source of strength.

[1] This research was supported by the NSF-EPCN-CAREER grant 1751085, Windlogics Inc., LDRD-ER (POD), and from the Advanced Grid Modeling Program in the Office of Electricity in U.S. Dept. of Energy

Contents

List of Figures

List of Tables

List of Abbreviations

CRLB Cramer-Rao Lower Bound

DERs Distributed Energy Resources

DSSE Distribution System State Estimation

LDF Linearized Distribution Flow

LTI Linear Time-invariant

MILP Mixed-integer Linear Program

MISDP Mixed-Integer Semidefinite Program

MSD Max-sum Diversity

P2L Probing-to-learn

PF Power Flow

PMUs Phasor Measurement Units

PSSE Power System State Estimation

QP Quadratic Program

RMSE Root Mean Square Error

SDP Semidefinite Program

SNR Signal-to-noise Ratio

WLAV Weighted Least Absolute Value

WLS Weighted Least-squares

ZIP Non-constant-power

Chapter 1

Introduction

With the advent of distributed energy resources (DERs), electric vehicles, and demand-response programs, grid operators are in dire need of new monitoring and design tools that help improve efficiency, reliability, and stability of modern power networks. To this end, the work in this thesis explores a generalized modeling and analysis framework for two pertinent tasks: i) learning loads via grid probing; and ii) optimizing power grid topologies for stability. While the former task utilizes the steady-state power flow model, the latter relies on the linear dynamic model. Nevertheless, both problems utilize similar graph-theoretic and optimization toolboxes to establish and validate our claims.

1.1 Grid probing for load learning

Due to limited instrumentation, low investment interest in the past, and the sheer scale of residential electricity networks, low-voltage grids have limited observability [6]. Traditionally, utility operators monitor distribution grids by collecting load, voltage and current magnitude measurements only at a few buses. This mode of operation has been functional due to the stationarity of conventional loads, the availability of historical data, and the under-utilization of distribution grids. Nevertheless, with the advent of smart power networks, there is a need for enhancing grid observability both in space and time to accomplish grid dispatch objectives, such as voltage regulation, power loss minimization, or optimal dispatch.

Determining whether a transmission network is observable, given a set of specifications, is performed using topological or numerical methods [64]. Topological methods rely on graph-theoretic principles, whereas the numerical ones study the rank of the Jacobian matrix related to the decoupled power flow model. An analysis of power flow (PF) problems under different specification sets is conducted in [45], again under the decoupled model. Given the higher resistance-to-reactance ratios, the decoupled grid model does not apply to distribution systems.

In networks that are not completely observable, techniques for placing limited meters to improve distribution system state estimation (DSSE) have been reported in [7], [71], [72]. A heuristic rule aiming at reducing the variance of voltage magnitude estimates at non-metered buses is suggested in [71], and is extended to voltage angles too in [72]. Pseudo-measurements have also been used in restoring observability in distribution grids [23]. Historical load profiles and real-time measurements are input to a neural network-based state estimator in [60], while a robust DSSE using pseudo-measurements is developed in [85]. Because psuedo-measurements may not be easily available or accurate, the aforementioned approaches cannot be used for real-time estimation and dispatch of power grids.

To overcome the aforementioned drawbacks, the work in this thesis exploits: i) grid sensor data from metered buses that are readily available at high temporal resolution; and ii) controllability of smart inverters found in solar panels, energy storage units, and electric vehicles to infer non-metered loads. Probing inverters to change their active injections and power factors momentarily can drive the system to a new operating point and elicit additional grid data that can be exploited for the load inference task. The proposed probing-to-learn (P2L) technique is not limited to load learning and has been recently used for other inference tasks such as topology learning and identification [21], [20], [19].

1.2 Optimal topology design

Increased penetration of renewables coupled with the proliferation of grid sensors and inverter-interfaced devices will allow utilities to have better observability of their power networks. At the same time, higher variability associated with DERs, renewables, and changing load patterns will pose an enormous challenge for design and stable operation of power networks [77]. The inherent uncertainty associated with renewable energy sources and active loads is likely to produce more frequent and higher amplitude disturbances [77]. In addition, owing to the lower aggregate inertia of systems with high penetration of renewables, the capability of power networks to handle such disturbances may be significantly reduced [83]. Without detailed simulations, it is usually hard to infer how a change in network topology influences the overall grid behavior and performance. Recent work in [44] shows that the impact of network topology on the power system can be quantified through the network Laplacian matrix eigenvalues. In addition, grid robustness against low frequency disturbances is mostly determined by network connectivity [44], further motivating this study. Past studies in the power and control systems communities have also looked at designing network topologies for specific goals using system theoretic tools. Such goals include reduction of transient line losses [77], improvement in feedback control [76], [58], coherence based network design [32] and augmentation [74]. Semidefinite programming (SDP) based tools have also been utilized to design and augment network topologies for dynamic control [39], [32], [65].

Reference [28] presented a unified framework to study topology design based on the $\mathcal{H}_2$-norm, for a variety of objective functions, such as line loss reduction, fast damping of oscillations, and network coherence. In [28], the focus was on topology reconfiguration rather than topology design. Further, the work in [28] developed suboptimal algorithms, albeit with guarantees on optimality gap, to tackle the combinatorial design problems involved. Different

from the work in [28], here we are interested in studying the optimal placement of new lines on existing networks as well as the design of radial (tree) and meshed (loopy) topologies for new networks with the goal of improving stability.

1.3 Contributions

To improve efficiency and reliability of our power networks and to aid utilities in performing grid-wide optimization tasks, we first present the novel idea of grid probing for learning loads. Secondly, to improve stability of power networks we develop a generalized modeling and analysis framework for optimally designing grid topologies. The key contributions of this thesis are as follows:

1. We put forth our novel P2L technique in Section 2.2. Exploiting the stationarity of non-metered loads over the probing period and assuming noiseless sensor data, the P2L problem is posed as a PF task coupled over multiple time instances.

2. We provide intuitive and easily verifiable graph-theoretic conditions under which probing succeeds in recovering unmonitored loads under phasor (Section 2.3) and non-phasor sensor data (Section 2.4). The established criteria capture the effects of the number of probing actions and the graph connectivity between metered and non-metered buses.

3. While several existing works exploit the radial (tree-like) structure of distribution grids for various learning tasks, our approach is generalized to cover meshed grids, including loopy multiphase distribution grids as well as transmission systems.

4. We cover the simple yet pertinent case of the single-slot problem (assuming passively collected grid data without probing) in Section 2.5. This setup makes no assumptions

on the underlying load model and is pertinent to power networks that don't have the necessary infrastructure to control smart inverters and obtain synchronized grid data.

5. The aforementioned identifiability criteria decide whether a probing setup is successful. If a setup is deemed successful, the utility could judiciously design inverter injections to improve load estimation. Nonetheless, these injections have to be implementable by inverters; compliant with voltage regulation and line flow limits; and drive the electric grid to sufficiently different states for better numerical conditioning. This injection design task is further challenged by the nonlinearity of the power flow equations and the fact that non-metered loads are not known *a priori*. In Section 3.2 we have developed an algorithmic approach for designing probing injections that conform to grid safety requirements and improve load estimates.

6. The proposed P2L tasks are tackled through our SDP-based solvers described in Section 3.3.

7. We present a comprehensive modeling and analysis framework for the topology design problem to optimize a $\mathcal{H}_2$-norm based performance metric subject to budget constraints in Sections 4.2 and 4.3.

8. In Section 4.4 we show that although the topology design task is inherently non-convex, it is possible to exactly reformulate the problem in tractable form using McCormick relaxation (or linearization). This can then be used with off-the-shelf solvers to determine the optimal solution.

9. By exploiting graph-theoretic properties to tighten bounds on the continuous optimization variables in Section 4.5, we show significant improvements in computation time.

For the load learning task our research efforts have resulted in three conference papers [15], [11], [10], and three journal publications [12], [13], [14]. Our papers [15] and [12] presented preliminary results on the grid probing framework and associated SDP-based solvers. The analysis there was restricted to two probing actions and the results were confined to radial grids. Moreoever, the probing actions were not designed, but rather relied on the natural variability of load/generation at metered buses. Load learning via passive data collection (without probing) was analyzed in [10]. These initial results were significantly extended in [11], [13]. More specifically, i) the analysis was extended non-trivially to multiple rather than only two probing actions; ii) probing setups with voltage magnitude and/or angle data were studied in a unified fashion; iii) we used the feeder connectivity to upper bound the number of probing actions beyond which there is no identifiability benefit; iv) a new proving technique generalized the analysis from radial to meshed grids, thus covering the timely topic of loopy multiphase distribution grids and transmission systems; v) grid probing to infer non-constant-power (ZIP) loads was explored as special case. The design of efficient probing setpoints that conform to inverter and network constraints while improving load estimates was further explored in [14]. For the topology design task our research efforts have led to a flagship controls conference paper [9].

Note that Chapters 2 and 3 have been largely reproduced from our papers [15], [11], [10], [12], [13], and [14], while Chapter 4 has been reproduced from [9].

Chapter 2

Smart Inverter Grid Probing for Learning Loads

2.1 Introduction

Low-voltage distribution grids have been plagued with limited observability, due to limited instrumentation, low investment interest in the past, and their sheer extent [6]. Traditionally, utility operators monitor distribution grids by collecting measurements infrequently and only from a few critical buses. This mode of operation has been functional due to the under-utilization of distribution grids and the availability of historical data. Nevertheless, with the advent of DERs, electric vehicles, and demand-response programs, there is a critical need to reliably estimate the system state and learn non-metered loads to optimally dispatch the grid on a frequent basis (say 20 min). To this end, the communication capabilities of grid sensors together with the actuation and sensing features of power inverters found in solar panels, energy storage units, and electric vehicles could be utilized toward unveiling loads.

Although estimating loads or the grid state has heavily relied on pseudo-measurements, such measurements may not be available or accurate under the current mode of operation [51], [41]. On the other hand, the widespread deployment of digital relays, phasor measurement units (PMUs), and inverter-interfaced DERs provide excellent opportunities for improving distribution grid observability [69], [38]. In addition, regular polling and on-demand reads

of customer loads and voltages via smart meters have enhanced the accuracy of distribution system state estimation [5], [42]. Since the previous schemes collect data on a hourly basis, they are of limited use for real-time optimization.

Rather than passively collecting grid readings to infer non-metered loads, this works advocates engaging inverters to probe the grid and thus actively collect feeder data. We define *probing* as the technique of perturbing an electric grid for the purpose of finding unknown parameters. The idea of probing has been previously suggested towards estimating the electro-mechanical oscillation modes in power transmission systems [87], [79]. Perturbing the voltage and/or current of a single inverter has been adopted in the power electronics community to determine the grid-equivalent Thevenin impedance of inverters [48]. Moreover, modulating the primary droop control loop of inverters has been recently suggested for learning loads and topologies in direct-current grids [2]. Graph algorithms and identifiability conditions for recovering feeder topologies using inverter probing data have been devised in [20], [19].

Beyond their standard energy conversion functionality, smart inverters are being utilized for reactive power control and other feeder optimization tasks [81]. In fact, the grid voltage response to inverter injection changes has been used as a means to solve optimal power flow tasks in a decentralized and/or communication-free fashion; see for example [33], [49], [25], [75], [3]. Leveraging exactly this voltage response, grid probing attributes smart inverters a third functionality towards monitoring rather than grid control.

The contribution of this chapter is on three fronts. First, we formulate our *Probing-to-Learn* (P2L) technique in Section 2.2. Exploiting the stationarity of non-metered loads during probing and assuming noiseless inverter readings, the P2L problem is posed as a coupled power flow task. Second, we provide intuitive and easily verifiable graph-theoretic conditions under which probing succeeds in finding non-metered loads under phasor (Section 2.3) and

non-phasor data (Section 2.4). Third, Section 2.5 extends probing to infer ZIP loads.

Regarding *notation*, column vectors (matrices) are denoted by lower- (upper-) case boldface letters and sets by calligraphic symbols. The cardinality of set $\mathcal{X}$ is denoted by $|\mathcal{X}|$, and its complement by $\bar{\mathcal{X}}$. The operators $(\cdot)^\top$ and $(\cdot)^H$ stand for (complex) transposition; the floor and ceiling functions are denoted by $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$; $\text{dg}(\mathbf{x})$ defines a diagonal matrix having $\mathbf{x}$ on its main diagonal; and $\mathbf{I}_N$ is the $N \times N$ identity matrix. The notation $\mathbf{x}_\mathcal{A}$ denotes the sub-vector of $\mathbf{x}$ indexed by $\mathcal{A}$; and $\mathbf{X}_{\mathcal{A},\mathcal{B}}$ is the matrix obtained by sampling the rows and columns of $\mathbf{X}$ indexed respectively by $\mathcal{A}$ and $\mathcal{B}$. Symbol $\mathbf{1}$ denotes the all-one vector and $\mathbf{e}_k$ is the k-th canonical vector; their dimensions would be clear from the context. The notation $\mathbf{V} \succeq \mathbf{0}$ means that $\mathbf{V}$ is a Hermitian (complex and conjugate symmetric) positive semidefinite matrix; the matrix trace is denoted by $\text{Tr}(\cdot)$; and $\|\mathbf{a}\|_2$ is the ℓ_2-norm of vector $\mathbf{a}$. The notation $k = 1 : K$ is a shorthand to $k = 1, \ldots, K$.

2.2 Grid Probing

Albeit not every bus is metered in a distribution grid, some buses are equipped with sensors recording voltage magnitudes and/or angles, actual powers, and power factors. Moreover, the power injections in solar panels and energy storage devices can be instantly controlled using advanced power electronics. Building on the physical law that perturbing power injections at different buses is reflected on voltage changes across the grid, the key idea here is to engage power electronics to probe the grid with the purpose of learning non-metered loads.

To formally describe grid probing, let us briefly review a feeder model. Consider a feeder represented by a graph $\mathcal{G} = (\mathcal{N}^+, \mathcal{L})$ where the nodes in $\mathcal{N}^+ := \{0, \ldots, N\}$ correspond to buses, and the edges in $\mathcal{L}$ to distribution lines. Let $\mathbf{Y} := \mathbf{G} + j\mathbf{B}$ be the grid bus admittance matrix and $\mathbf{G}$ (resp. $\mathbf{B}$) be the bus conductance (resp. susceptance) matrix. By definition,

the entries B_{nm} and G_{nm} for $n \neq m$ are non-zero only if $(n, m) \in \mathcal{L}$. Let us express the voltage phasor at bus $n \in \mathcal{N}^+$ in Cartesian and polar coordinates as

$$v_n = v_{r,n} + jv_{i,n} = u_n e^{j\theta_n}.$$

The substation is indexed by $n = 0$, its voltage remains fixed at $1 + j0$, and the remaining buses comprise the set $\mathcal{N}$. If $\mathbf{v}_r := [v_{r,0} \ \cdots \ v_{r,N}]^\top$ and $\mathbf{v}_i := [v_{i,0} \ \cdots \ v_{i,N}]^\top$, define the system state as $\mathbf{v} := [\mathbf{v}_r^\top \ \mathbf{v}_i^\top]^\top$. Apparently, for each bus $n \in \mathcal{N}^+$, the squared voltage magnitude and the net power injections are quadratic functions of $\mathbf{v}$, whereas the voltage angle is a trigonometric function of $\mathbf{v}$ [40]

$$u_n(\mathbf{v}) = u_n^2 = v_{r,n}^2 + v_{i,n}^2 \tag{2.1a}$$

$$p_n(\mathbf{v}) = v_{r,n} \sum_{m=0}^{N} (v_{r,m} G_{nm} - v_{i,m} B_{nm}) + v_{i,n} \sum_{m=0}^{N} (v_{r,m} B_{nm} + v_{i,m} G_{nm}) \tag{2.1b}$$

$$q_n(\mathbf{v}) = v_{i,n} \sum_{m=0}^{N} (v_{r,m} G_{nm} - v_{i,m} B_{nm}) - v_{r,n} \sum_{m=0}^{N} (v_{r,m} B_{nm} + v_{i,m} G_{nm}) \tag{2.1c}$$

$$\theta_n(\mathbf{v}) = \arctan\left(\frac{v_{r,n}}{v_{i,n}}\right). \tag{2.1d}$$

With the proliferation of grid sensors and inverters, the distribution grid operator may have access to all four quantities $(u_n, \theta_n, p_n, q_n)$ on a subset of buses. Different from the conventional PF setup with PQ and PV buses, we partition $\mathcal{N}^+$ into the subsets:

- The set $\mathcal{M}$ of metered buses for which $(u_n, \theta_n, p_n, q_n)$ are known and their power injections are possibly controllable. This set includes the substation and buses equipped with smart sensors and/or inverters. Its cardinality is denoted by $M := |\mathcal{M}|$.

- The set $\mathcal{O}$ of non-metered buses where no information is available. Its cardinality is

denoted by $O := |\mathcal{O}|$, and apparently, $N + 1 = M + O$.

The inverters interfacing DERs are typically modeled as constant-power generators [81], [33], [25]: Internal control loops can reach setpoints for (re)-active power injections within microseconds. The setpoints should comply with solar irradiance and the rating of the inverter.

Remark 2.1. We emphasize $p_n + jq_n$ is the net complex injection. If bus n hosts a smart inverter and a non-controllable load, it is henceforth assumed that the operator measures $p_n + jq_n$ and the voltage at the point of common coupling, and controls the complex injection from the inverter. This assumption is reasonable since smart inverters are usually equipped with sensors; e.g., the Pecan Street project measures both the net and inverter injections [67].

Given the feeder topology captured in $\mathbf{Y}$ and the specifications $\{(u_n, \theta_n, p_n, q_n)\}_{n\in\mathcal{M}}$, our goal is to recover the power injections at non-metered buses $\{(p_n, q_n)\}_{n\in\mathcal{O}}$. Lacking a direct mapping from $\{(u_n, \theta_n, p_n, q_n)\}_{n\in\mathcal{M}}$ to $\{(p_n, q_n)\}_{n\in\mathcal{O}}$, the problem of finding the non-metered loads boils down to the task of recovering the underlying state $\mathbf{v}$ first. Collecting the grid data $\{(u_n^t, \theta_n^t, p_n^t, q_n^t)\}_{n\in\mathcal{M}}$ at time t and assuming for now these data are noiseless, we get the specifications

$$u_n(\mathbf{v}_t) = u_n^t, \qquad \forall n \in \mathcal{M} \tag{2.2a}$$

$$\theta_n(\mathbf{v}_t) = \theta_n^t, \qquad \forall n \in \mathcal{M} \tag{2.2b}$$

$$p_n(\mathbf{v}_t) = p_n^t \qquad \forall n \in \mathcal{M} \tag{2.2c}$$

$$q_n(\mathbf{v}_t) = q_n^t \qquad \forall n \in \mathcal{M} \tag{2.2d}$$

which involve $4M$ equations over $2(N+1)$ unknowns. A necessary condition for solving (2.2)

is $4M \geq 2(N+1)$. Since $N+1 = M+O$, the condition simplifies to

$$M \geq O. \tag{2.3}$$

In other words, the metered buses must be at least as many as the non-metered ones.

To relax this condition on M, one may consider jointly processing the data $\{(u_n^t, \theta_n^t, p_n^t, q_n^t)\}_{n\in\mathcal{M}}$ collected across multiple times $t \in \mathcal{T}$ with $\mathcal{T} := \{1, \ldots, T\}$. This approach does not improve the observability of the equations in (2.2), simply because the equations are independent over $\mathcal{T}$. Moreover, both the $4MT$ equations and the $2(N+1)T$ state variables $\{\mathbf{v}_t\}_{t=1}^T$ scale with T. One way to relate power flow specifications across time is to assume that the non-metered loads remain invariant across $\mathcal{T}$, that is

$$p_n(\mathbf{v}_t) = p_n(\mathbf{v}_{t+1}), \qquad \forall n \in \mathcal{O}, t \in \mathcal{T}' \tag{2.4a}$$

$$q_n(\mathbf{v}_t) = q_n(\mathbf{v}_{t+1}), \qquad \forall n \in \mathcal{O}, t \in \mathcal{T}' \tag{2.4b}$$

where $\mathcal{T}' := \{1, \ldots, T-1\}$. In this way, we obtain the additional $2O(T-1)$ equations and couple the states $\{\mathbf{v}_t\}_{t=1}^T$.

Even though there may be an observability advantage in coupling specifications across time, the timespan of $\mathcal{T}$ is critical: For non-metered loads to remain unchanged, the timespan of $\mathcal{T}$ should be relatively short. But if the duration of $\mathcal{T}$ is too short, the metered injections in the buses of $\mathcal{M}$ may not change either. In this case, the grid state remains identical over $\mathcal{T}$, the scheme degenerates to the setup of (2.2) for $T = 1$, and there is no advantage by coupling specifications.

At this point, smart inverters come to our rescue: The timespan of $\mathcal{T}$ can be made sufficiently short so that the non-metered loads in the buses of $\mathcal{O}$ remain invariant over $\mathcal{T}$,

whereas the power injections from smart inverters vary. The key point here is to couple power flow specifications through what we term *grid probing*. Probing can be accomplished by commanding inverters to change their power injections for one second. An inverter can curtail its solar output; (dis)-charge an energy storage unit; and/or change its power factor. Multiple probing actions can be instructed within tens of seconds. By intentionally perturbing inverter injections, the grid transitions across different states $\{\mathbf{v}_t\}_{t=1}^T$ depending on the probing injections and non-metered loads. Recording voltages $\{u_n^t, \theta_n^t\}_{t=1}^T$ over $n \in \mathcal{M}$ could unveil non-metered loads.

The metered buses in $\mathcal{M}$ can be classified into probing buses and metered but non-controllable buses. Although grid data $(u_n^t, \theta_n^t, p_n^t, q_n^t)$ are collected on both probed and metered buses, the operator can control only the probing buses. To simplify the presentation, we will henceforth assume that all metered buses are probing buses, although the analysis and algorithms apply to the more general setup.

Probing postulates two assumptions on non-metered loads: *a1)* They remain constant throughout $\mathcal{T}$; and *a2)* are modeled as of constant power. Assumption *a1)* may be reasonable over the short duration of probing. Regarding *a2)*, one could alternatively adopt a ZIP load model for bus $n \in \mathcal{O}$ [50]

$$-p_n^t(u_n^t) = \alpha_{p_n}(u_n^t)^2 + \beta_{p_n} u_n^t + \gamma_{p_n} \tag{2.5a}$$

$$-q_n^t(u_n^t) = \alpha_{q_n}(u_n^t)^2 + \beta_{q_n} u_n^t + \gamma_{q_n}. \tag{2.5b}$$

The parameters $(\alpha_{p_n}, \beta_{p_n}, \gamma_{p_n})$ correspond to the constant-impedance, constant-current, and constant-power components of active load; likewise $(\alpha_{q_n}, \beta_{q_n}, \gamma_{q_n})$ for reactive load. A non-metered ZIP load is then described by six rather than two parameters. Moreover, despite the model for load, n does not change across $\mathcal{T}$, its power injection $p_n + jq_n$ does change for

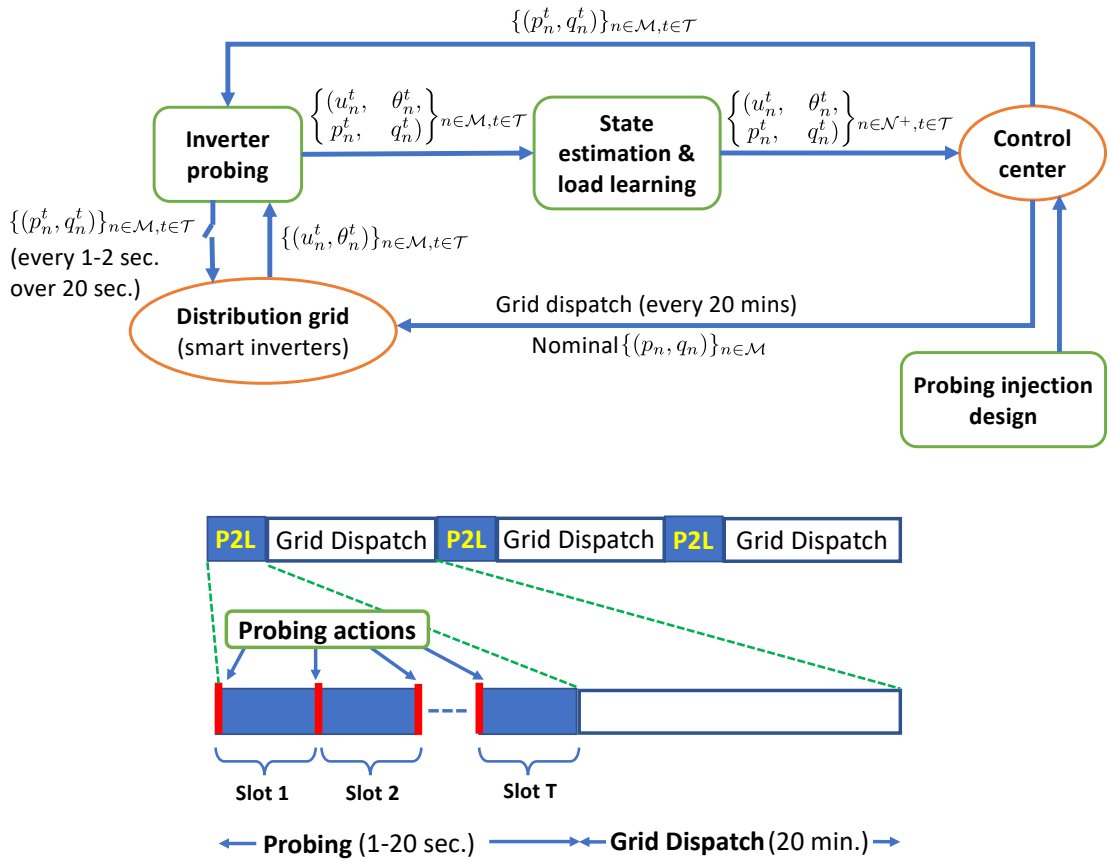

Figure 2.1: Overview of the P2L framework: (a) block diagram depicting the P2L task with phasor data (top panel); (b) temporal organization of grid operation (bottom panel).

varying u_n. Then, the coupling equations are not valid in the form of (2.4) anymore. If the ZIP parameters are assumed invariant over $\mathcal{T}$, the power flow equations can still be coupled across $\mathcal{T}$, yet the identifiability analysis and the associated solvers become perplex. To bypass this complexity, Section 2.5 copes with ZIP loads by resorting to single-slot probing.

Figure 2.1 depicts how probing can be incorporated into grid operation: Suppose a utility operates a demand-response program; manages energy storage; or controls smart inverters for reactive power control on a 20-min basis. To solve the optimal power flow problem, the operator needs to know the injections at non-metered buses. To do so, a probing interval lasting few tens of seconds precedes the feeder dispatch. This interval $\mathcal{T}$ is divided into T

probing slots indexed by $t = 1, \ldots, T$. During each probing slot t, every inverter $n \in \mathcal{M}$ changes its injections to the setpoints (p_n^t, q_n^t) and reads voltage data (u_n^t) or (u_n^t, θ_n^t). At the end of interval $\mathcal{T}$, each inverter $n \in \mathcal{M}$ sends the collected data $\{u_n^t\}_{t\in\mathcal{T}}$ or $\{u_n^t, \theta_n^t\}_{t\in\mathcal{T}}$ back to the utility, and switches its setpoints back to their nominal values. The utility processes the collected data, infers the non-metered loads, and dispatches the grid for the next 20-min period. Some implementation details follow.

Remark 2.2. The probing setpoints (p_n^t, q_n^t) for all t and $n \in \mathcal{M}$ are decided by the utility prior to $\mathcal{T}$ and communicated to all inverters via two-way communication links. This is to ensure that probing complies with voltage constraints and for improved load estimation accuracy; see Chapter 3.2. The commanded setpoints are attained by simple PID controllers. Further, the inverters act synchronously along probing slots. Since potential delays may raise synchronization issues, developing protocols where inverters probe asynchronously is of interest.

Remark 2.3. The proposed probing scheme aims at recovering loads assuming the feeder topology is known. The topology includes bus connectivity and line impedances, phase assignments, and the statuses of capacitors and voltage regulators; see Remark 2.4. Although small errors in line impedances and regulator tap settings could be modeled as measurement noise, grid probing is as sensitive to topology errors as power flow equations are. However, probing can be also used for inferring grid topologies and line parameters without knowing non-metered loads [20], [19]. Moreover, phase assignments can be inferred from smart meter data; see e.g., [53]. Such techniques could precede P2L to find or calibrate feeder models.

Remark 2.4. Feeders are equipped with voltage-control devices, such as regulators and capacitor banks, which respond to voltage excursions by changing their taps and switching on/off with time delays of around 30-90 seconds [50]. Since $\mathcal{T}$ lasts 20 sec or less, probing is not expected to trigger voltage control actions per se. Nonetheless, there are still chances for

these actions to occur during $\mathcal{T}$ due to load fluctuations. If the utility does not monitor these devices in real-time or it cannot override their settings during probing, the topology learning techniques of Remark 2.3 could be possibly used. Voltage control actions and topology reconfigurations will be ignored in this work. Interestingly though, such actions could be used towards grid probing too.

Grid probing can be now formally stated as follows.

Definition 2.5 (Probing-to-Learn task with phasor data)**.** Given $\mathbf{Y}$ and probing data $(u_n^t, \theta_n^t, p_n^t, q_n^t)$ for all $n \in \mathcal{M}$ and $t \in \mathcal{T}$, the probing-to-learn (P2L) task entails solving the equations in (2.2) for $t \in \mathcal{T}$ jointly with the coupling equations in (2.4).

The P2L task involves $4MT + 2O(T-1)$ equations in $2(N+1)T$ unknowns. A necessary condition for solving it is

$$M \geq \frac{O}{T} \tag{2.6}$$

which coincides with the condition in (2.3) for $T = 1$. For $T \geq 2$ however, it improves upon (2.3) if probing over multiple time instances is allowed. In [12], we have derived conditions under which the P2L task recovers non-metered loads for $T = 2$. The analysis there was further confined to non-phasor grid data $\{(u_n^t, p_n^t, q_n^t)\}_{n\in\mathcal{M}, t\in\mathcal{T}}$ and radial grids. The conference work of [11] extended the previous claims (without proofs) to meshed networks. Here, we broaden the scope to study the identifiability of the P2L task with phasor data $\{(u_n^t, \theta_n^t, p_n^t, q_n^t)\}_{n\in\mathcal{M}}$ over $\mathcal{T}$, and show that the analysis with non-phasor data can be seen as a special case of the former [13].

2.3 Identifiability of P2L with Phasor Data

As customary in identifiability analysis, data will be assumed noiseless; noisy data are considered in Chapter 3.2 [14]. The relationship between the inputs $\{u_n^t, \theta_n^t, p_n^t, q_n^t\}_{n\in\mathcal{M}}$ and the outputs $\{p_n^t, q_n^t\}_{n\in\mathcal{O}}$ of the P2L task is implicit since the PF equations involve $\{\mathbf{v}_t\}_{t=1}^T$ as nuisance variables. Because of this, P2L is tackled in two steps. The first step of finding $\{\mathbf{v}_t\}_{t=1}^T$ is the challenging one. In the second step, one simply evaluates $(p_n(\mathbf{v}_t), q_n(\mathbf{v}_t))$ for all $n \in \mathcal{O}$ and $t = 1$. For numerical stability, one can recover the unknown injections by averaging as $\frac{1}{T}\sum_{t=1}^T p_n(\mathbf{v}_t)$ and $\frac{1}{T}\sum_{t=1}^T q_n(\mathbf{v}_t)$ for all $n \in \mathcal{O}$. Hence, if the system states $\{\mathbf{v}_t\}_{t=1}^T$ can be recovered by solving (2.2) and (2.4), the P2L task is deemed successful.

Granted the P2L equations are non-linear, identifiability can be ensured only within a neighborhood of the nominal $\{\mathbf{v}_t\}_{t=1}^T$. Upon invoking the inverse function theorem, a necessary and sufficient condition for locally solving P2L is that the Jacobian matrix $\mathbf{J}(\{\mathbf{v}_t\})$ related to the nonlinear equations of (2.2) and (2.4) is full rank. Because $\mathbf{J}(\{\mathbf{v}_t\})$ depends on $\{\mathbf{v}_t\}$, characterizing its column rank for any $\{\mathbf{v}_t\}$ is challenging.

To tackle this issue, we resort to the *generic rank* of a matrix defined as the maximum possible rank attained if the non-zero entries of the matrix are allowed to take arbitrary real values [82], [80]. If the generic rank of an $M \times N$ matrix $\mathbf{E}$ with $M \geq N$ equals N, matrix $\mathbf{E}$ is said to be of *full generic rank*. The generic rank of a matrix is related to a graph constructed by the sparsity pattern of the matrix, that is the locations of its (non)-zero entries. To explain this link, some graph-theoretic concepts are needed.

A graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ is bipartite if $\mathcal{N}$ can be partitioned into disjoint subsets $\mathcal{N}_1$ and $\mathcal{N}_2$, such that $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$, and every $\ell \in \mathcal{L}$ connects a node in $\mathcal{N}_1$ to a node in $\mathcal{N}_2$. A subset of edges $\mathcal{L}' \subseteq \mathcal{L}$ is termed a perfect matching of $\mathcal{N}_1$ to $\mathcal{N}_2$, if every vertex in $\mathcal{N}_1$ is incident to exactly one edge in $\mathcal{L}'$. The degree $\delta_n(\mathcal{G})$ of node n is defined as the number of edges

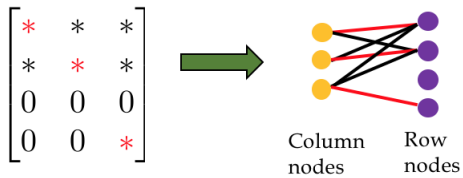

Figure 2.2: The sparsity pattern of $\mathbf{E}$ and its bipartite graph $\mathcal{G}_E$: Column nodes are linked to row nodes depending on the entries of $\mathbf{E}$. The perfect matching is marked in red. From Lemma 2.6, any matrix with this sparsity pattern is generically full rank. Had $E_{4,3} = 0$, no perfect matching would exist.

incident to node n in $\mathcal{G}$. Given a matrix $\mathbf{E} \in \mathbb{R}^{M \times N}$, construct a bipartite graph $\mathcal{G}_E$ having $M + N$ nodes: Each column of $\mathbf{E}$ is mapped to a column node and each row of $\mathbf{E}$ to a row node. An edge runs from the n-th column node to the m-th row node only if $E_{mn} \neq 0$; see Fig. 2.2. Based on $\mathcal{G}_E$, we will use next claim.

Lemma 2.6 ([82], [80])**.** *An $M \times N$ matrix $\mathbf{E}$ has full generic rank if and only if the bipartite graph $\mathcal{G}_E$ features a perfect matching from the column nodes to its row nodes.*

According to Lemma 2.6 (proved in [80, Th. 12.10]), the generic identifiability of P2L relies on the sparsity pattern of $\mathbf{J}(\{\mathbf{v}_t\})$. The goal is to match every column node (state) of $\mathbf{J}(\{\mathbf{v}_t\})$ to a unique row node (equation). The non-zero entries of $\mathbf{J}(\{\mathbf{v}_t\})$ are the available links.

To characterize the sparsity pattern of $\mathbf{J}(\{\mathbf{v}_t\})$, consider the Jacobian matrices

$$\mathbf{J}^u(\mathbf{v}) = \begin{bmatrix} 2\,\mathrm{dg}(\mathbf{v}_r) & 2\,\mathrm{dg}(\mathbf{v}_i) \end{bmatrix} \tag{2.7a}$$

$$\mathbf{J}^\theta(\mathbf{v}) = \begin{bmatrix} 2\,\mathrm{dg}(\frac{\mathbf{v}_i}{\mathbf{v}_r^2+\mathbf{v}_i^2}) & 2\,\mathrm{dg}(\frac{-\mathbf{v}_r}{\mathbf{v}_r^2+\mathbf{v}_i^2}) \end{bmatrix} \tag{2.7b}$$

$$\mathbf{J}^p(\mathbf{v}) = \begin{bmatrix} -\mathbf{G}\,\mathrm{dg}(\mathbf{v}_r) - \mathbf{B}\,\mathrm{dg}(\mathbf{v}_i) & -\mathbf{G}\,\mathrm{dg}(\mathbf{v}_i) + \mathbf{B}\,\mathrm{dg}(\mathbf{v}_r) \\ -\mathrm{dg}(\mathbf{G}\mathbf{v}_r) + \mathrm{dg}(\mathbf{B}\mathbf{v}_i) & -\mathrm{dg}(\mathbf{B}\mathbf{v}_r) - \mathrm{dg}(\mathbf{G}\mathbf{v}_i) \end{bmatrix} \tag{2.7c}$$

$$\mathbf{J}^q(\mathbf{v}) = \begin{bmatrix} \mathbf{B}\,\mathrm{dg}(\mathbf{v}_r) - \mathbf{G}\,\mathrm{dg}(\mathbf{v}_i) & \mathbf{B}\,\mathrm{dg}(\mathbf{v}_i) + \mathbf{G}\,\mathrm{dg}(\mathbf{v}_r) \\ +\mathrm{dg}(\mathbf{B}\mathbf{v}_r) + \mathrm{dg}(\mathbf{G}\mathbf{v}_i) & -\mathrm{dg}(\mathbf{G}\mathbf{v}_r) + \mathrm{dg}(\mathbf{B}\mathbf{v}_i) \end{bmatrix} \tag{2.7d}$$

associated accordingly with the squared voltage magnitudes and voltage angles, and the (re)active power injections over all buses. Matrix $\mathbf{J}(\{\mathbf{v}_t\})$ consists of stacked row-sampled submatrices of $\mathbf{J}^u(\mathbf{v}_t)$, $\mathbf{J}^\theta(\mathbf{v}_t)$, $\mathbf{J}^p(\mathbf{v}_t)$, and $\mathbf{J}^q(\mathbf{v}_t)$ corresponding to (2.2) and (2.4) for $t \in \mathcal{T}$. The matrices obtained by selecting the rows of $\mathbf{J}^u(\mathbf{v}_t)$ associated with buses in $\mathcal{M}$ and $\mathcal{O}$ are respectively denoted by $\mathbf{J}^u_{\mathcal{M}}(\mathbf{v}_t)$ and $\mathbf{J}^u_{\mathcal{O}}(\mathbf{v}_t)$. Similar notation is used for $\mathbf{J}^\theta(\mathbf{v}_t)$, $\mathbf{J}^p(\mathbf{v}_t)$, and $\mathbf{J}^q(\mathbf{v}_t)$. Let us define

$$\mathbf{J}_{\mathcal{M}}(\mathbf{v}_t) := \begin{bmatrix} \mathbf{J}^u_{\mathcal{M}}(\mathbf{v}_t) \\ \mathbf{J}^\theta_{\mathcal{M}}(\mathbf{v}_t) \\ \mathbf{J}^p_{\mathcal{M}}(\mathbf{v}_t) \\ \mathbf{J}^q_{\mathcal{M}}(\mathbf{v}_t) \end{bmatrix} \text{ and } \mathbf{J}_{\mathcal{O}}(\mathbf{v}_t) := \begin{bmatrix} \mathbf{J}^p_{\mathcal{O}}(\mathbf{v}_t) \\ \mathbf{J}^q_{\mathcal{O}}(\mathbf{v}_t) \end{bmatrix}.$$

Every $\mathbf{J}_{\mathcal{M}}(\mathbf{v}_t)$ corresponds to $4M$ metering equations, and every $\mathbf{J}_{\mathcal{O}}(\mathbf{v}_t)$ to $2O$ coupling equations. Having defined $\mathbf{J}_{\mathcal{M}}(\mathbf{v}_t)$ and $\mathbf{J}_{\mathcal{O}}(\mathbf{v}_t)$, the entire Jacobian matrix $\mathbf{J}(\{\mathbf{v}_t\})$ can be row-permuted as

$$\left[\begin{array}{c:c:c:c:c} \mathbf{J}_{\mathcal{M}}(\mathbf{v}_1) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{J}_{\mathcal{O}}(\mathbf{v}_1) & -\mathbf{J}_{\mathcal{O}}(\mathbf{v}_2) & \mathbf{0} & \cdots & \mathbf{0} \\ \hdashline \mathbf{0} & \mathbf{J}_{\mathcal{M}}(\mathbf{v}_2) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathcal{O}}(\mathbf{v}_2) & -\mathbf{J}_{\mathcal{O}}(\mathbf{v}_3) & \cdots & \mathbf{0} \\ \hdashline \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathcal{M}}(\mathbf{v}_3) & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{J}_{\mathcal{O}}(\mathbf{v}_T) \\ \hdashline \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{\mathcal{M}}(\mathbf{v}_T) \end{array}\right]. \tag{2.9}$$

This row-permuted version of $\mathbf{J}(\{\mathbf{v}_t\})$ will be denoted by $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$, and has been obtained by interleaving block rows of metering and coupling equations.

Matrix $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$ features the sparsity pattern of a block tridiagonal matrix. To reveal this structure, split each block row of coupling equations into two block rows. The top block row will be grouped with the previous block row of metering equations. The bottom block row will be grouped with the next block row of metering equations as in Fig. 2.3.

Focus now on the blocks lying on the main diagonal of $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$. These blocks will be denoted by $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ for $t \in \mathcal{T}$. If for each $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$, its columns can be perfectly matched to its rows, then a perfect bipartite matching for the entire $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$ has been obtained. Then, Lemma 2.6 guarantees that $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$ and $\mathbf{J}(\{\mathbf{v}_t\})$ are generically full rank.

Our goal is to assign coupling equations to blocks so that every block $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ enjoys a perfect bipartite matching. There are $2O(T-1)$ coupling equations to be assigned to T blocks. A uniform allocation should assign $\frac{2O(T-1)}{T}$ coupling equations per block. With this allocation, block t will have $4M$ metering equations and $\frac{2O(T-1)}{T}$ coupling equations over its $2(N+1) = 2M + 2O$ states in $\mathbf{v}_t$. For a perfect bipartite matching to exist, we need $4M + \frac{2O(T-1)}{T} \geq 2M + 2O$.

The last requirement coincides with the necessary condition of (2.6) for $T \geq 2$; but it is not enough: Every coupling equation can be assigned to exactly one between two specific blocks; see Fig. 2.3. For example, a coupling equation in the block row involving $\mathbf{J}_{\mathcal{O}}(\mathbf{v}_2)$ and $-\mathbf{J}_{\mathcal{O}}(\mathbf{v}_3)$ can be grouped either with $\mathbf{J}_{\mathcal{M}}(\mathbf{v}_2)$ or $\mathbf{J}_{\mathcal{M}}(\mathbf{v}_3)$. Partitioning the coupling equations into groups of $\frac{2O(T-1)}{T}$ while adhering to the latter requirement is the crux of the identifiability analysis. To allocate coupling equations, let us first define the bipartite grid graph $\mathcal{G}_b$.

Definition 2.7 (Bipartite grid graph)**.** Consider the graph obtained from $\mathcal{G}$ upon maintaining only the edges between $\mathcal{M}$ and $\mathcal{O}$. Replicate the node set $\mathcal{M}$ to form $\mathcal{M}'$, and connect the nodes in $\mathcal{M}'$ to nodes in $\mathcal{O}$ by replicating the $\mathcal{M}$–$\mathcal{O}$ edges. The obtained bipartite graph will be denoted by $\mathcal{G}_b$.

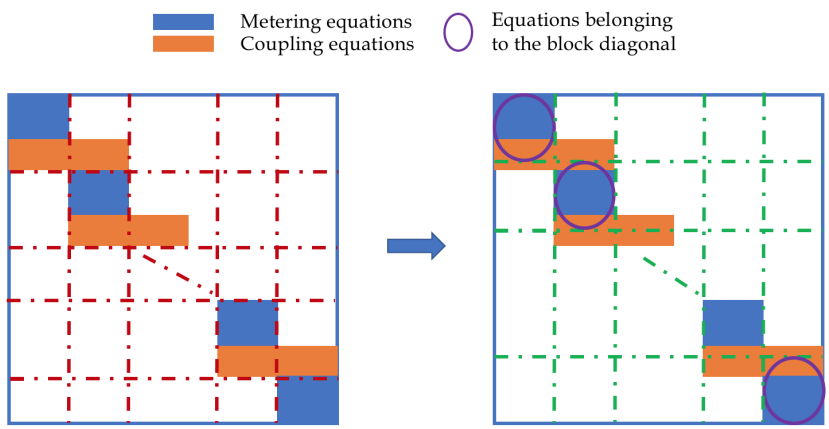

Figure 2.3: Sparsity pattern of $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$ [cf. (2.9)] (left panel). Block tridiagonal $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$ revealed after splitting each block row of coupling equations (right panel).

The identifiability of P2L relies on a matching in $\mathcal{G}_b$.

Theorem 2.8. *If $\mathcal{O}$ can be partitioned into $\{\bar{\mathcal{O}}_k\}_{k=1}^{\lceil T/2 \rceil}$ so that each one of them independently can be perfectly matched to $\mathcal{M} \cup \mathcal{M}'$ on $\mathcal{G}_b$, the Jacobian matrix $\mathbf{J}(\{\mathbf{v}_t\})$ related to the P2L task with phasor data is generically full rank.*

In essence, Theorem 2.8 provides sufficient conditions for successful probing. The proof of Theorem 2.8 relies on two lemmas shown next: Lemma 2.9 provides sufficient conditions for the coupling equations assigned to block t, so that $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ enjoys a bipartite matching. Lemma 2.10 explains when these conditions can be met simultaneously for all $t \in \mathcal{T}$. The analysis uses the concept of a *multi-set*. Different from a conventional set that contains unique elements, a multi-set is allowed to have multiple instances of elements. For example, we will override the definition of set union, so that $\{a, b\} \cup \{a, b\}$ does not yield $\{a, b\}$, but the multi-set $\{a, a, b, b\}$.

Lemma 2.9. *Partition $\mathcal{O}$ into $\mathcal{O}_t \cup \bar{\mathcal{O}}_t$ so that $|\mathcal{O} \cup \mathcal{O}_t| = 2O(T-1)/T$. Assume block $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ is assigned some coupling equations related to $\mathcal{O} \cup \mathcal{O}_t$. If the vertices in $\bar{\mathcal{O}}_t$ can be matched to the vertices in $\mathcal{M} \cup \mathcal{M}'$ on $\mathcal{G}_b$, the block $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ features a bipartite matching from its columns to its rows.*

Proof. It can be easily verified from (2.7) that the sparsity patterns of $\mathbf{J}^u(\mathbf{v}_t)$ an $\mathbf{J}^\theta(\mathbf{v}_t)$ coincide with the sparsity pattern of $[\mathbf{I}_{N+1}\ \mathbf{I}_{N+1}]$. The sparsity patterns of $\mathbf{J}^p(\mathbf{v}_t)$ and $\mathbf{J}^q(\mathbf{v}_t)$ coincide with the sparsity pattern of $[\mathbf{G}\ \mathbf{G}]$ where $\mathbf{G}$ is the bus conductance matrix; see [40, Table 3.2]. From (2.9), the sparsity pattern of $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ is

$$\begin{bmatrix} \mathbf{I}_{\mathcal{M},\mathcal{N}^+} & \mathbf{I}_{\mathcal{M},\mathcal{N}^+} \\ \mathbf{I}_{\mathcal{M},\mathcal{N}^+} & \mathbf{I}_{\mathcal{M},\mathcal{N}^+} \\ \mathbf{G}_{\mathcal{M},\mathcal{N}^+} & \mathbf{G}_{\mathcal{M},\mathcal{N}^+} \\ \mathbf{G}_{\mathcal{M},\mathcal{N}^+} & \mathbf{G}_{\mathcal{M},\mathcal{N}^+} \\ \mathbf{G}_{\mathcal{O},\mathcal{N}^+} & \mathbf{G}_{\mathcal{O},\mathcal{N}^+} \\ \mathbf{G}_{\mathcal{O}_t,\mathcal{N}^+} & \mathbf{G}_{\mathcal{O}_t,\mathcal{N}^+} \end{bmatrix} \tag{2.10}$$

where the first block row relates to voltage magnitudes; the second to voltage angles; the third and fourth to probing injections; while the fifth and sixth to coupled injections.

To create a bipartite matching for block $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$, unfold the sparsity pattern in (2.10) column-wise using $\mathcal{N}^+ = \mathcal{M} \cup \mathcal{O}$ as

$$\begin{bmatrix} \boxed{\mathbf{I}_{\mathcal{M},\mathcal{M}}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} & \mathbf{I}_{\mathcal{M},\mathcal{M}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} \\ \mathbf{I}_{\mathcal{M},\mathcal{M}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} & \boxed{\mathbf{I}_{\mathcal{M},\mathcal{M}}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} \\ \mathbf{G}_{\mathcal{M},\mathcal{M}} & \mathbf{G}_{\mathcal{M},\mathcal{O}} & \mathbf{G}_{\mathcal{M},\mathcal{M}} & \boxed{\mathbf{G}_{\mathcal{M},\mathcal{O}}} \\ \mathbf{G}_{\mathcal{M},\mathcal{M}} & \mathbf{G}_{\mathcal{M},\mathcal{O}} & \mathbf{G}_{\mathcal{M},\mathcal{M}} & \boxed{\mathbf{G}_{\mathcal{M},\mathcal{O}}} \\ \mathbf{G}_{\mathcal{O},\mathcal{M}} & \boxed{\mathbf{G}_{\mathcal{O},\mathcal{O}}} & \mathbf{G}_{\mathcal{O},\mathcal{M}} & \mathbf{G}_{\mathcal{O},\mathcal{O}} \\ \mathbf{G}_{\mathcal{O}_t,\mathcal{M}} & \mathbf{G}_{\mathcal{O}_t,\mathcal{O}} & \mathbf{G}_{\mathcal{O}_t,\mathcal{M}} & \boxed{\mathbf{G}_{\mathcal{O}_t,\mathcal{O}}} \end{bmatrix}. \tag{2.11}$$

The first block column in (2.11) relates to variables $\{v^t_{r,n}\}_{n\in\mathcal{M}}$, and can be matched to the first block row via $\mathbf{I}_{\mathcal{M},\mathcal{M}}$. Similarly, the third block column relates to variables $\{v^t_{i,n}\}_{n\in\mathcal{M}}$, and can be matched to the second block row. The second block column relates to variables

$\{v_{r,n}^t\}_{n\in\mathcal{O}}$, and can be matched to the fifth block row via the main diagonal of $\mathbf{G}_{\mathcal{O},\mathcal{O}}$.

To achieve a bipartite matching, the fourth block column related to variables $\{v_{i,n}^t\}_{n\in\mathcal{O}}$ has to be matched to the union of the third, fourth, and sixth block rows. Lacking a simple diagonal matching now, we leverage the sparsity pattern of $\mathbf{G}$. It suffices to match the column nodes in $\mathcal{O}$ to the row nodes in $\mathcal{M}\cup\mathcal{M}\cup\mathcal{O}_t$. Because $\mathcal{O}=\mathcal{O}_t\cup\bar{\mathcal{O}}_t$, the column nodes $\mathcal{O}_t$ can be matched to the row nodes $\mathcal{O}_t$ via some diagonal entries of $\mathbf{G}_{\mathcal{O}_t,\mathcal{O}}$. Then, the column nodes $\bar{\mathcal{O}}_t$ have to be matched to the row nodes $\mathcal{M}\cup\mathcal{M}$. This can be accomplished based on the hypothesis of this Lemma, thus completing its proof. □

Lemma 2.10. *Under the condition of Theorem 2.8, the coupling equations for two successive blocks $\tilde{\mathbf{J}}_t(\mathbf{v}_t)$ with $t=2k-1$ and $t=2k$ share the same sparsity pattern of $\mathcal{O}\cup\mathcal{O}_k$ for $k=1,\ldots,\lceil T/2\rceil$.*

Proof. The pair of blocks $\tilde{\mathbf{J}}_{2k-1}(\mathbf{v}_{2k-1})$ and $\tilde{\mathbf{J}}_{2k}(\mathbf{v}_{2k})$ will be jointly indexed by k. Define also

$$\mathcal{R}_k := \bigcup_{\tau=1}^{k}\bar{\mathcal{O}}_\tau. \tag{2.12}$$

In addition to the claim of this lemma, we will also prove that when passing from pair $k-1$ to pair k, a set of coupling equations represented by $\mathcal{R}_{k-1}\cup\mathcal{R}_{k-1}$ have not been assigned to block $2k-2$, and are free to be assigned to block $2k-1$.

Proving by induction, we start with the base case. The pair indexed by $k=1$ consists of $\tilde{\mathbf{J}}_1(\mathbf{v}_1)$ and $\tilde{\mathbf{J}}_2(\mathbf{v}_2)$. The active and reactive equations coupling these two blocks can be represented by $\mathcal{O}\cup\mathcal{O}=\mathcal{O}_1\cup\mathcal{O}_1\cup\bar{\mathcal{O}}_1\cup\bar{\mathcal{O}}_1$. Let us assign $\mathcal{O}_1\cup\mathcal{O}_1\cup\bar{\mathcal{O}}_1$ to block 1. With this assignment, the coupling equations for $\tilde{\mathbf{J}}_1(\mathbf{v}_1)$ get the sparsity pattern of $\mathcal{O}\cup\mathcal{O}_1$. The remaining coupling equations in $\bar{\mathcal{O}}_1$ are assigned to block 2.

Block 2 shares with block 3 the coupling equations $\mathcal{O}\cup\mathcal{O}$, which are again expressed as

$\mathcal{O}_1 \cup \mathcal{O}_1 \cup \bar{\mathcal{O}}_1 \cup \bar{\mathcal{O}}_1$. From this new set of coupling equations, assign $\mathcal{O}_1 \cup \mathcal{O}_1$ to block 2. Hence, the coupling equations for $\tilde{\mathbf{J}}_2(\mathbf{v}_2)$ have the sparsity pattern of $\mathcal{O}_1 \cup \mathcal{O}_1 \cup \bar{\mathcal{O}}_1 = \mathcal{O} \cup \mathcal{O}_1$. The unused coupling equations are represented by $\bar{\mathcal{O}}_1 \cup \bar{\mathcal{O}}_1 = \mathcal{R}_1 \cup \mathcal{R}_1$.

Suppose the claim holds for the block pair $k-1$. It is next shown that the claim holds for the block pair k consisting of blocks $2k-1$ and $2k$. Starting with the odd block $2k-1$, the unused equations $\mathcal{R}_{k-1} \cup \mathcal{R}_{k-1}$ that couple blocks $2k-2$ and $2k-1$ are assigned to block $2k-1$. Block $2k-1$ is also coupled to block $2k$ via $\mathcal{O} \cup \mathcal{O}$ equations, which can be expressed as $\mathcal{O}_k \cup \mathcal{O}_k \cup \bar{\mathcal{O}}_k \cup \bar{\mathcal{O}}_k$. The key point here is that by the definition of $\mathcal{R}_{k-1}$ and because $\bar{\mathcal{O}}_k$'s are mutually exclusive by the hypothesis of this lemma, it holds that

$$\mathcal{R}_{k-1} \cap \bar{\mathcal{O}}_k = \emptyset \text{ and } \mathcal{R}_{k-1} \subset \mathcal{O}, \text{ so that } \mathcal{R}_{k-1} \subseteq \mathcal{O}_k.$$

Therefore, the set $\mathcal{O}_k$ can be partitioned into $\mathcal{R}_{k-1}$ and $\mathcal{O}_k \setminus \mathcal{R}_{k-1}$. From the equations coupling blocks $2k-1$ and $2k$, the equations $(\mathcal{O}_k \setminus \mathcal{R}_{k-1}) \cup (\mathcal{O}_k \setminus \mathcal{R}_{k-1}) \cup \bar{\mathcal{O}}_k$ are assigned to block $2k-1$. In this way, the coupling equations for block $2k-1$ have the sparsity pattern

$$\underbrace{\mathcal{R}_{k-1} \cup \mathcal{R}_{k-1}}_{\text{with block } 2k-2} \cup \underbrace{(\mathcal{O}_k \setminus \mathcal{R}_{k-1}) \cup (\mathcal{O}_k \setminus \mathcal{R}_{k-1}) \cup \bar{\mathcal{O}}_k}_{\text{with block } 2k} = \mathcal{O} \cup \mathcal{O}_k.$$

The unused equations coupling blocks $2k-1$ and $2k$ are $\mathcal{R}_{k-1} \cup \mathcal{R}_{k-1} \cup \bar{\mathcal{O}}_k$.

Moving to block $2k$ of pair k, the unused equations $\mathcal{R}_{k-1} \cup \mathcal{R}_{k-1} \cup \bar{\mathcal{O}}_k$ coupling block $2k$ with block $2k-1$ are assigned to block $2k$. Block $2k$ is also coupled with block $2k+1$ through $\mathcal{O} \cup \mathcal{O} = \mathcal{O}_k \cup \mathcal{O}_k \cup \bar{\mathcal{O}}_k \cup \bar{\mathcal{O}}_k$. From this new set of coupling equations, assign equations $(\mathcal{O}_k \setminus \mathcal{R}_{k-1}) \cup (\mathcal{O}_k \setminus \mathcal{R}_{k-1})$ to block $2k$. Hence, the coupling equations assigned to block $2k-1$

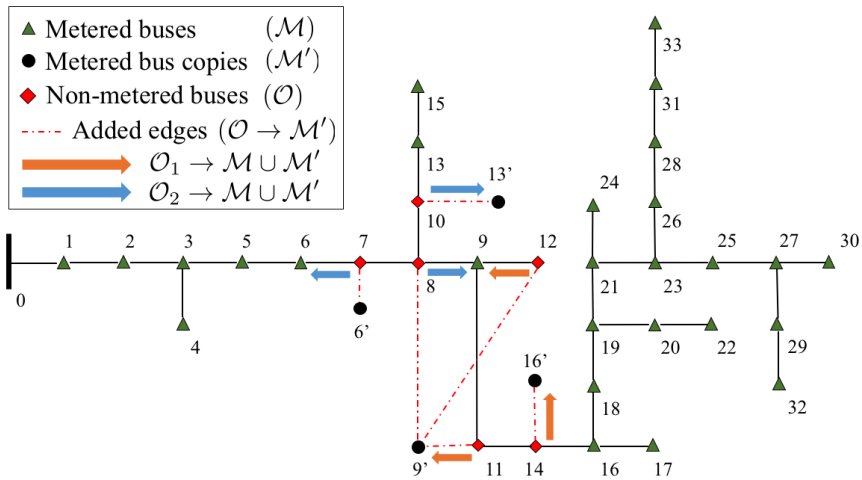

Figure 2.4: Matchings on the IEEE 34-bus grid for $T = 4$ and $O = 6$ for the P2L task with phasor data.

have the sparsity pattern

$$\underbrace{\mathcal{R}_{k-1} \cup \mathcal{R}_{k-1} \cup \bar{\mathcal{O}}_k}_{\text{with block } 2k-1} \cup \underbrace{(\mathcal{O}_k \setminus \mathcal{R}_{k-1}) \cup (\mathcal{O}_k \setminus \mathcal{R}_{k-1})}_{\text{with block } 2k+1} = \mathcal{O} \cup \mathcal{O}_k.$$

The unused equations coupling blocks $2k$ and $2k+1$ are

$$\mathcal{R}_{k-1} \cup \mathcal{R}_{k-1} \cup \bar{\mathcal{O}}_k \cup \bar{\mathcal{O}}_k = \mathcal{R}_k \cup \mathcal{R}_k.$$

For the last block pair, the coupling equations already assigned to block $T-1$ have the sparsity pattern $\mathcal{O} \cup \mathcal{O}_{T/2}$. The last block T differs from the previous blocks as it only gets the $\mathcal{R}_{T/2-1} \cup \mathcal{R}_{T/2-1} \cup \bar{\mathcal{O}}_{T/2}$ unused coupling equations between blocks $T-1$ and T. Under the condition of Th. 2.8, $\mathcal{R}_{T/2-1} \cup \bar{\mathcal{O}}_{T/2} = \mathcal{O}$ and because $\mathcal{O} = \mathcal{O}_{T/2} \cup \bar{\mathcal{O}}_{T/2}$, we also have $\mathcal{R}_{T/2-1} = \mathcal{O}_{T/2}$. Hence, the sparsity pattern of the last block is also given by $\mathcal{O} \cup \mathcal{O}_{T/2}$. Since every pair of successive blocks has the same structure, the diagonal blocks of $\tilde{\mathbf{J}}(\{\mathbf{v}_t\})$ will exhibit $\lceil T/2 \rceil$ distinct sparsity patterns. $\square$

Algorithm 1 Test for Successful Probing (phasor data)

Assign unit capacity to edges in $\mathcal{G}_b$ to define graph $\tilde{\mathcal{G}}_b$.
In $\tilde{\mathcal{G}}_b$, add source node n_s, and connect it to all nodes in $\mathcal{O}$. These edges are assigned unit capacity.
In $\tilde{\mathcal{G}}_b$, add destination node n_d, and connect it to all nodes in $\mathcal{M} \cup \mathcal{M}'$.
Initialize $T = 2$.
while $T \leq T_{\max}$ **do**
  The edges running between $\mathcal{M} \cup \mathcal{M}'$ and n_d are assigned capacities of $T/2$.
  Run a max-flow problem between n_s and n_d.
  if obtained n_s–n_d flow equals O **then**
    return *Probing setup is deemed successful for T.*
  else
    $T := T + 2$
  end if
  return *Probing setup is deemed unsuccessful.*
end while

Theorem 2.8 follows as a direct consequence of Lemmas 2.9 and 2.10. To simplify the exposition, we will henceforth assume even T. To appreciate the conditions of Theorem 2.8, examine the *probing setup*, that is the placement of non-metered and probed buses, of Figure 2.4. The black circles denote the copies $\mathcal{M}'$ of nodes in $\mathcal{M}$, and the dashed red lines show the added edges from $\mathcal{O}$ to $\mathcal{M}'$. The operator needs to infer the loads at the $O = 6$ non-metered buses marked by red diamonds. To study if probing this feeder over $T = 4$ slots is successful, the set $\mathcal{O}$ has to be partitioned into two subsets $\mathcal{O}_1$ and $\mathcal{O}_2$, so that the buses of each subset are matched to buses in $\mathcal{M} \cup \mathcal{M}'$ on $\mathcal{G}_b$. The orange and blue arrows show precisely these matchings. If the feeder were to be probed over $T = 2$ slots instead, probing would fail since buses $\{8, 11, 12\}$ cannot be uniquely matched to any buses in $\mathcal{M} \cup \mathcal{M}'$.

As illustrated through this example, to check the condition of Theorem 2.8 for a particular $(\mathcal{O}, \mathcal{M})$ probing setup, first one has to construct the bipartite graph $\mathcal{G}_b$ from $\mathcal{G}$. Then, given a number of probing actions T: *i)* the set $\mathcal{O}$ has to be partitioned into the subsets $\{\bar{\mathcal{O}}_k\}_{k=1}^{T/2}$; and *ii)* the nodes within each $\bar{\mathcal{O}}_k$ have to be mapped to the nodes in $\mathcal{M} \cup \mathcal{M}'$ on $\mathcal{G}_b$. Albeit

these steps may seem computationally hard, they can be solved by a linear program as detailed in Algorithm 1.

Given a probing setup, Algorithm 1 finds the maximum flow between nodes n_s and n_d over graph $\tilde{\mathcal{G}}_b$ constructed from $\mathcal{G}_b$. The edges in $\tilde{\mathcal{G}}_b$ are organized in three layers: The edges of the first layer connect n_s to $\mathcal{O}$ and have unit capacities. The edges of the second layer connect $\mathcal{O}$ to $\mathcal{M} \cup \mathcal{M}'$ and have unit capacities as well. The edges of the third layer connect $\mathcal{M} \cup \mathcal{M}'$ to n_d and have capacities of $T/2$. This is to ensure that each node in $\mathcal{M} \cup \mathcal{M}'$ is mapped to at most $T/2$ nodes in $\mathcal{O}$ through the second layer. If the maximum n_s–n_d flow equals O, all first-layer edges have been used to their capacity to map every node in $\mathcal{O}$ to exactly one node in $\mathcal{M} \cup \mathcal{M}'$.

The max-flow problem can be solved using the Ford-Fulkerson algorithm, whose complexity scales linearly with the number of graph nodes and edges [35]. Moreover, if all edge capacities are integers, the algorithm finds an integral maximal flow. If the maximum n_s–n_d flow is smaller than O, there is no matching for the tested T. Then, the edge capacities at the third layer can be increased and the process is repeated. Theorem 2.8 asserts that the chances of successful probing improve for larger T. This is because progressively smaller subsets of $\mathcal{O}$ need to be mapped to $\mathcal{M} \cup \mathcal{M}'$. Yet this gain in T is limited by the bus placement $(\mathcal{M}, \mathcal{O})$ as quantified next.

Lemma 2.11. *If $\delta_{\mathcal{M}}$ is the maximum node degree over $\mathcal{M}$ on $\tilde{\mathcal{G}}_b$, a probing setup with phasor data cannot turn into successful beyond $T_{\max} = \delta_{\mathcal{M}} - 1$.*

Proof. Consider node $m \in \mathcal{M}$ in $\tilde{\mathcal{G}}_b$ with degree δ_m. This node is connected to node n_d via an edge having capacity $T/2$, and to $\delta_m - 1$ nodes in $\mathcal{O}$ via unit-capacity edges. The maximum flow that can pass through the second-layer edges to m is $\delta_m - 1$. This flow will be funneled through edge (m, n_d). Then, there is no advantage for this edge to have capacity

larger than $\delta_m - 1$, so that $T/2 \leq \delta_m - 1$. Considering all $m \in \mathcal{M} \cup \mathcal{M}'$, there is no point in testing for values of T beyond $T \leq 2(\delta_{\mathcal{M}} - 1)$.

The bound can be improved, since the previous argument assumed that all $\delta_m - 1$ edges between $\mathcal{O}$ and $m \in \mathcal{M}$ have reached their capacity. That will not happen since the $\mathcal{O}$ nodes adjacent to m on the feeder, can be shared between m and its copy $m' \in \mathcal{M}'$ on $\tilde{\mathcal{G}}_b$. Hence, the flow passing jointly through m and m' cannot exceed $\delta_m - 1$. Then, the capacity of edge (m, n_d) plus the capacity of edge (m', n_d) can be safely limited to $\delta_m - 1$, implying $T \leq \delta_m - 1$. □

Lemma 2.11 implies that increasing T beyond $T_{\max}$ has no hope in making probing successful for a specific placement, and Algorithm 1 terminates with a negative answer. Once a $(\mathcal{M}, \mathcal{O})$ placement is deemed successful, there are two questions to be answered: *i)* how to select probing injections; and *ii)* how to recover the non-metered loads. Both questions along with numerical tests are deferred to Chapter 3.2.

2.4 Identifiability of P2L with Non-phasor Data

Since PMUs have limited penetration in distribution grids, requiring voltage phasor data at probing buses may be unrealistic. This section studies probing with non-phasor data.

Definition 2.12 (P2L task with non-phasor data)**.** Given $\mathbf{Y}$ and probing data (u_n^t, p_n^t, q_n^t) for $n \in \mathcal{M}$ and $t \in \mathcal{T}$, the P2L task entails solving the equations in (2.2a), (2.2c) and (2.2d) for $t \in \mathcal{T}$, jointly with the coupling equations in (2.4).

A simple count of equations and unknowns dictates $M \geq \frac{2O}{T}$, which is clearly more restrictive than (2.6). We next provide a sufficient condition under which this task is solvable.

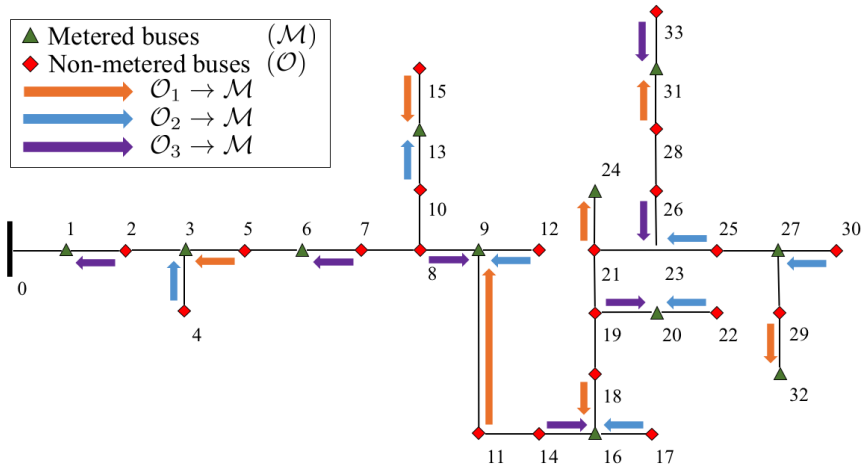

Figure 2.5: Matchings on the IEEE 34-bus grid for $T = 6$ and $O = 21$ for the P2L task with non-phasor data.

Theorem 2.13. *If $\mathcal{O}$ can be partitioned into $\{\bar{\mathcal{O}}_k\}_{k=1}^{\lceil T/2 \rceil}$ such that each one of them independently can be perfectly matched to $\mathcal{M}$ on $\mathcal{G}$, the Jacobian matrix $\mathbf{J}(\{\mathbf{v}_t\})$ related to the P2L task with non-phasor data is generically full rank.*

Dropping the voltage angle metering equations, matrix $\mathbf{J}_{\mathcal{M}}(\mathbf{v}_t)$ in (2.9) is replaced by

$$\mathbf{J}_{\mathcal{M}}(\mathbf{v}_t) := \begin{bmatrix} \mathbf{J}^u_{\mathcal{M}}(\mathbf{v}_t) \\ \mathbf{J}^p_{\mathcal{M}}(\mathbf{v}_t) \\ \mathbf{J}^q_{\mathcal{M}}(\mathbf{v}_t) \end{bmatrix}.$$

Similar to Theorem 2.8, it is not hard to see that the nodes in $\bar{\mathcal{O}}_t$ have to be matched to the nodes in $\mathcal{M}$, rather than $\mathcal{M} \cup \mathcal{M}'$.

Consider for example the probing setup of Figure 2.5. To infer the loads at $O = 21$ non-metered buses with $T = 6$ probing slots, the set $\mathcal{O}$ has to be partitioned into three subsets $\mathcal{O}_1$, $\mathcal{O}_2$, and $\mathcal{O}_3$, so that the buses of each subset are matched to $\mathcal{M}$. The orange, blue, and purple arrows in the figure show these matchings. Because non-metered buses are divided into three subsets, up to three non-metered buses can be matched to the same probed bus.

Algorithm 2 Test for Successful Probing (non-phasor data)

```
Connect n_s to all nodes in O with unit-capacity edges.
Connect O to M based on G with unit-capacity edges.
Connect all nodes in M to n_d.
Initialize T = 2.
while T ≤ T_max do
    Assign capacity T/2 to edges between M and n_d.
    Run a max-flow problem between n_s and n_d.
    if obtained n_s–n_d flow equals O then
        return  Probing setup is deemed successful for T.
    else
        T := T + 2
    end if
    return  Probing setup is deemed unsuccessful.
end while
```

For example, buses $\{14, 17, 18\}$ are all matched to the probed bus 16. Probing the same feeder over $T = 2$ or $T = 4$ rather than $T = 6$ slots would fail.

The condition of Theorem 2.13 can be easily tested by Algorithm 2 and up to the value of $T_{\max}$ provided next.

Corollary 2.14. *If $\delta_{\mathcal{M}}$ is the maximum degree of the nodes in $\mathcal{M}$ on the graph constructed by Alg. 2, a probing setup with non-phasor data cannot turn into successful beyond $T_{\max} = 2(\delta_{\mathcal{M}} - 1)$.*

Corollary 2.14 is proved as part of the proof of Lemma 2.11. Compared to Theorem 2.13, the condition of Theorem 2.8 provided more flexibility towards attaining a bipartite matching since probed buses can be used twice. If a probing setup is successful for non-phasor data, it is also successful for phasor data. Interestingly, the matchings in Theorems 2.8 and 2.13 depend solely on the sparsity pattern of $\mathbf{G}$ and the probing setup, so the claims here apply to even meshed (e.g., multiphase) grids.

2.5 Single-Slot Probing

The analysis so far depends on assumption *a2*) of constant-power loads. Under the ZIP load model of (2.5), the coupling equations in (2.4) are no longer valid, and thus, the metering equations decouple across $\mathcal{T}$. Can the non-metered loads $p_n + jq_n$ for $n \in \mathcal{O}$ still be recovered upon collecting data on $\mathcal{M}$? This answer can be on the affirmative with *single-slot probing*, that is $T = 1$. Leveraging the tools of Sections 2.3 and 2.4, we next study the observability of single-slot probing. The ensuing two results provide conditions for successful load recovery using (non)-phasor data.

Theorem 2.15. *If each bus in $\mathcal{O}$ can be matched to one unique bus in $\mathcal{M}$ on $\mathcal{G}$, the Jacobian matrix $\mathbf{J}(\mathbf{v}_1)$ related to single-slot probing ($T = 1$) with phasor data has full generic rank.*

Proof. The sparsity pattern of $\mathbf{J}(\mathbf{v}_t)$ can be derived from (2.10)–(2.11) by eliminating the blocks related to coupling equations

$$\begin{bmatrix} \boxed{\mathbf{I}_{\mathcal{M},\mathcal{M}}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} & \mathbf{I}_{\mathcal{M},\mathcal{M}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} \\ \mathbf{I}_{\mathcal{M},\mathcal{M}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} & \boxed{\mathbf{I}_{\mathcal{M},\mathcal{M}}} & \mathbf{I}_{\mathcal{M},\mathcal{O}} \\ \mathbf{G}_{\mathcal{M},\mathcal{M}} & \boxed{\mathbf{G}_{\mathcal{M},\mathcal{O}}} & \mathbf{G}_{\mathcal{M},\mathcal{M}} & \mathbf{G}_{\mathcal{M},\mathcal{O}} \\ \mathbf{G}_{\mathcal{M},\mathcal{M}} & \mathbf{G}_{\mathcal{M},\mathcal{O}} & \mathbf{G}_{\mathcal{M},\mathcal{M}} & \boxed{\mathbf{G}_{\mathcal{M},\mathcal{O}}} \end{bmatrix}. \tag{2.14}$$

Variables $\{v_{r,n}^t\}_{n\in\mathcal{M}}$ corresponding to the first block column and $\{v_{i,n}^t\}_{n\in\mathcal{M}}$ to the third block column can be matched respectively to the first and second block row via $\mathbf{I}_{\mathcal{M},\mathcal{M}}$. To complete the bipartite matching, the second and fourth block columns (variables $\{v_{r,n}^t\}_{n\in\mathcal{O}}$ and $\{v_{i,n}^t\}_{n\in\mathcal{O}}$) can be matched to the third and fourth block rows, accordingly. Hence, $\mathbf{J}(\mathbf{v}_t)$ is generically full rank if there exists a perfect matching in $\mathbf{G}_{\mathcal{M},\mathcal{O}}$, that is every non-metered node in $\mathcal{O}$ is mapped to a unique node in $\mathcal{M}$; see also Lemma 2.6. $\square$

Theorem 2.16. *If each bus in $\mathcal{O}$ can be matched to two unique buses in $\mathcal{M}$ on $\mathcal{G}$, the Jacobian $\mathbf{J}(\mathbf{v}_1)$ related to single-slot probing ($T = 1$) with non-phasor data has full generic rank.*

Proof. Given non-phasor data, the second block row related to voltage angles in (2.14) is dropped. Following the arguments to the proof for Theorem 2.15, matrix $\mathbf{J}(\mathbf{v}_t)$ can be shown to be generically full rank if there exists a perfect matching in $\mathbf{G}_{\mathcal{M},(\mathcal{O}\cup\mathcal{O})}$, that is every node in $\mathcal{O}$ is mapped to two unique nodes in $\mathcal{M}$; see also Lemma 2.6. □

The conditions of Th. 2.15 and 2.16 can be tested by Algorithm 2 by fixing $T = 2$ and $T = 1$, respectively. Figure 2.6 shows successful placements per Theorems 2.15 and 2.16.

If the conditions of Th. 2.15 and 2.16 are met, the non-metered loads $p_n + jq_n$ for $n \in \mathcal{O}$ can be recovered using single-slot probing, regardless if these loads are constant-power or not. However, the operator may also want to estimate their ZIP parameters in (2.5). Estimating these parameters directly with multi-slot probing becomes complicated. Instead, one could adopt multi-slot probing in a two-step process as follows: First, the feeder is probed over $\mathcal{T}$ with $|\mathcal{T}| = T > 3$. Under Th. 2.15 and 2.16, the operator obtains estimates $(\hat{u}_n^t, \hat{p}_n^t, \hat{q}_n^t)$ for all non-metered buses $n \in \mathcal{O}$ and $t \in \mathcal{T}$. Secondly, the ZIP parameters for active load n can be estimated through the least-squares (LS) fit

$$[\hat{\alpha}_{p_n} \ \hat{\beta}_{p_n} \ \hat{\gamma}_{p_n}]^\top := (\hat{\mathbf{U}}_n^\top \hat{\mathbf{U}}_n)^{-1} \hat{\mathbf{U}}_n^\top \hat{\mathbf{p}}_n \tag{2.15}$$

where $\hat{\mathbf{p}}_n := [\hat{p}_n^1 \ \ldots \ \hat{p}_n^T]^\top$ and the t-th row of matrix $\hat{\mathbf{U}}_n$ is $[(\hat{u}_n^t)^2 \ \hat{u}_n^t \ 1]$ for $t = 1, \ldots, T$. Similar LS fits can be performed for the reactive ZIP load parameters. A major concern here is that all entries of the Vandermonde matrix $\hat{\mathbf{U}}_n$ are close to unity in compliance with voltage regulation. For $T = 3$, the determinant of $\hat{\mathbf{U}}_n$ is calculated as $(\hat{u}_n^1 - \hat{u}_n^2)(\hat{u}_n^1 - \hat{u}_n^3)(\hat{u}_n^2 - \hat{u}_n^3)$ [47], which yields $|\hat{\mathbf{U}}_n| = -2 \cdot 10^{-3}$ even for bus voltages as widely spread as $\hat{u}_n^1 = 0.9$, $\hat{u}_n^2 = 1.0$,

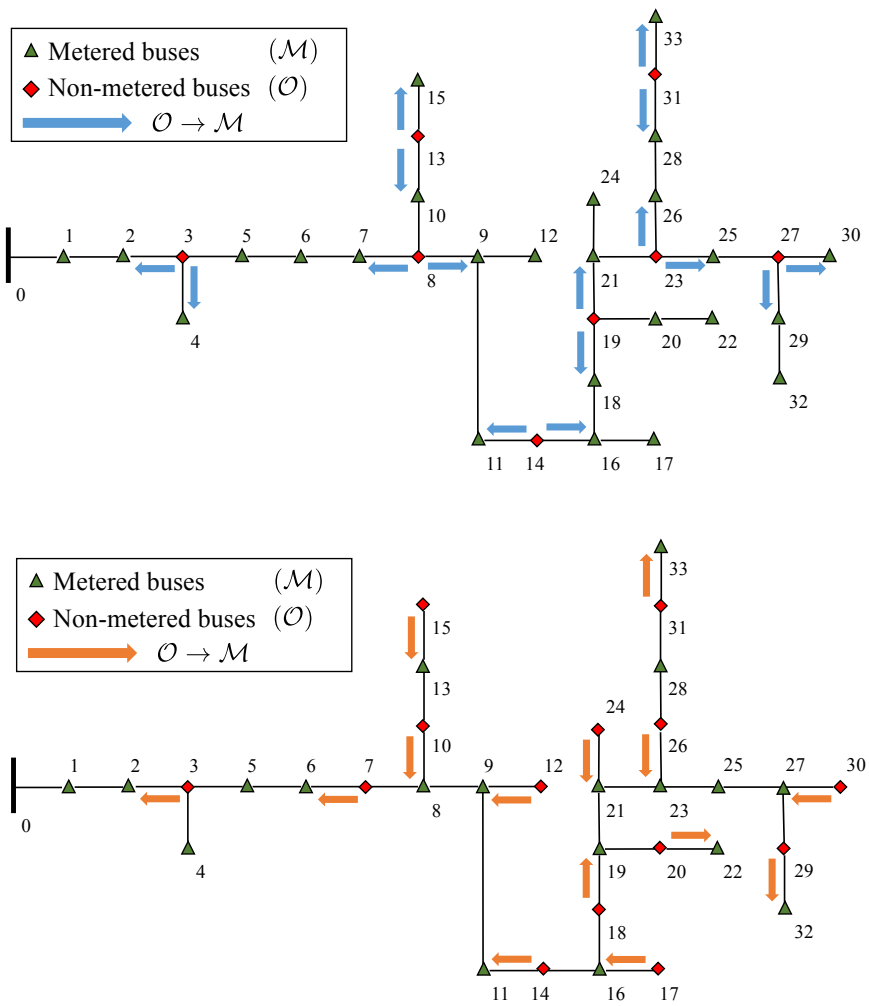

Figure 2.6: Matchings on the IEEE 34-bus grid for single-slot probing ($T = 1$) with: non-phasor data and $O = 8$ (top panel); phasor data and $O = 14$ (bottom panel). Single-slot probing waives assumption *a2)* on exclusively constant-power loads.

and $\hat{u}_n^3 = 1.1$. This reveals that the task of estimating ZIP parameters from voltage/power data is *ill-posed*. This is germane to the task itself rather than the method (here probing) used to collect the data.

Finally, note that Th. 2.15 and 2.16 hold even when data are not collected via probing, e.g., smart meter data. Therefore, our observability analysis covers the general setup where voltage and active/reactive power data or specifications are given only for $\mathcal{M}$. Under the linearized power flow model similar conditions were derived in [10] for radial grids. Note that

the tools and analysis presented in this chapter can be easily applied to check observability of networks with different measurement types. In fact, the presented framework can be also be utilized to derive the traditional topological observability conditions for transmission systems [1].

2.6 Conclusions

In this chapter we presented the novel technique of intentionally probing an electric grid using inverters to recover non-metered loads. The technique leverages the actuation capabilities of smart inverters, the data collected at probed buses, and the stationarity of non-metered loads, to formulate a power flow problem coupled over multiple times. We provided sufficient conditions for solvability that can be easily verified by solving a max-flow problem on a grid graph to test if a probing placement is successful. Beyond probing, we also cast the pertinent task of finding loads using data from a subset of buses as a special case. Assuming a probing setup satisfies these conditions, we explore next how inverter probing setpoints can be designed to improve load estimation accuracy while adhering to grid safety requirements.

Chapter 3

Probing Injection Design and Solvers

3.1 Introduction

Chapter 2 put forth the novel data acquisition scheme of probing-to-learn (P2L). The P2L scheme leverages smart inverters to probe an electric grid with the purpose of finding the values of non-metered loads. It also provided conditions under which a particular probing setup is successful. In particular, it was shown that given the feeder graph $\mathcal{G}$, the locations of non-metered buses $\mathcal{O}$ and probing buses $\mathcal{M}$, and the number of probing actions T, a simple linear program could tell whether non-metered loads could be recovered or not. Assuming noiseless data, this test relied on the generic rank of the Jacobian matrix $\mathbf{J}\left(\{\mathbf{v}_t\}\right)$ related to the P2L equations. It is thus a *topological* rather than a *numerical observability* guarantee [40, Ch. 4.6].

Even for the standard PF and power system state estimation (PSSE) setups, topological observability relates to the sparsity structure of the associated Jacobian matrix. This structure alone however cannot adequately capture the numerical column rank of the Jacobian matrix. There exist specification or measurement sets whose Jacobian is full column-rank in general, but becomes ill-conditioned or even singular under specific state values (including the boundaries for voltage collapse); see e.g., [40, Ch. 10]. In addition, once a probing setup is deemed topologically observable, the power injections of probing inverters could be judiciously selected to improve load or state estimates. This is challenging since P2L is an

implicit nonlinear identification task, and probing injections should be comply to network constraints without knowing the non-metered loads.

The contribution of this chapter is on two practical aspects of grid probing. First, a systematic approach to design probing setpoints that conform to grid safety and improve numerical accuracy is developed in Section 3.2. Second, the proposed P2L task is tackled through SDP-based solvers presented in Section 3.3. The conditions of Chapter 2 along with the probing setpoint design and the solver, are numerically validated using actual residential load data from the Pecan Street project on the IEEE 34-bus benchmark feeder in Section 3.4. Conclusions and current research efforts are outlined in Section 3.5.

3.2 Designing Probing Setpoints

Suppose probing has been deemed successful for a particular $(\mathcal{M}, \mathcal{O})$ placement of probed and non-metered buses, i.e., the setup $(\mathcal{M}, \mathcal{O})$ has passed the test of Algorithm 1 or 2. The next question is how to select probing setpoints that are implementable by inverters; compliant to feeder constraints; and at the same time, improve estimation accuracy. This chapter deals with the design of inverter setpoints during probing interval $\mathcal{T}$ with slots $t = 1, \ldots, T$, for a given T.

To facilitate the exposition, let us stack the power injections at all probing buses $\{(p_n, q_n)\}_{n\in\mathcal{M}}$ in vectors $\mathbf{p}_\mathcal{M}$, $\mathbf{q}_\mathcal{M}$, and $\mathbf{s}_\mathcal{M} := [\mathbf{p}_\mathcal{M}^\top \ \mathbf{q}_\mathcal{M}^\top]^\top$. Likewise, the injections at all non-metered buses $\{(p_n, q_n)\}_{n\in\mathcal{O}}$ are collected in $\mathbf{p}_\mathcal{O}$, $\mathbf{q}_\mathcal{O}$, and $\mathbf{s}_\mathcal{O} := [\mathbf{p}_\mathcal{O}^\top \ \mathbf{q}_\mathcal{O}^\top]^\top$. The injections at slot t will be denoted by a superscript t.

In search of a meaningful metric to design the probing injections $\{\mathbf{s}_\mathcal{M}^t\}_{t=1}^T$, one could consider the minimum mean square estimation error for non-metered loads $\mathbf{s}_\mathcal{O}$ or states $\{\mathbf{v}_t\}_{t=1}^T$. The

former is hard to derive given the implicit estimation task involved. The latter exhibits the Cramer-Rao lower bound (CRLB) of $[\mathbf{J}^\top(\{\mathbf{v}_t\})\,\mathbf{J}(\{\mathbf{v}_t\})]^{-1}$; a proof for this CRLB can be obtained by adopting the result in [84]. Since $\mathbf{J}(\{\mathbf{v}_t\})$ depends linearly on $\{\mathbf{v}_t\}_{t=1}^T$, the CRLB depends inverse quadratically on the unknown states.

To arrive at a practical solution, we resort to selecting probing setpoints so that the electric grid is driven to the most diverse states $\{\mathbf{v}_t\}_{t=1}^T$ while abiding by inverter and feeder operational constraints. We conjecture that probing the grid to effect larger state variations across $\mathcal{T}$ would yield smaller condition numbers for $\mathbf{J}(\{\mathbf{v}_t\})$ and $\mathbf{J}^\top(\{\mathbf{v}_t\})\,\mathbf{J}(\{\mathbf{v}_t\})$.

Hence, the goal is to design $\{\mathbf{s}_{\mathcal{M}}^t\}_{t=1}^T$ that yield the most diverse system states $\{\mathbf{v}_t\}_{t=1}^T$. Since the system states depend on both $\{\mathbf{s}_{\mathcal{M}}^t\}_{t=1}^T$ and the unknown $\mathbf{s}_{\mathcal{O}}$ in a non-linear fashion, our design adopts a linearized power flow model. The latter can be obtained by taking the first-order Taylor's series approximation of the PF equations with respect to nodal voltages expressed in polar coordinates [16], [27]. Unless a reference system state is available, the linearization occurs at the flat voltage profile of $\tilde{\mathbf{v}} = u_0\mathbf{1} + j\mathbf{0}$, and yields the so termed *linearized distribution flow* (LDF) model [8], [27], which can be rearranged for our analysis as

$$\mathbf{y} := \begin{bmatrix} \mathbf{u} - u_0\mathbf{1} \\ \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{K} & \mathbf{L} \\ \mathbf{M} & \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{\mathcal{M}} \\ \mathbf{s}_{\mathcal{O}} \end{bmatrix}. \tag{3.1}$$

The vectors $\mathbf{u}$ and $\boldsymbol{\theta}$ collect the voltage magnitudes and angles at all buses excluding the substation; and matrices $(\mathbf{K}, \mathbf{L}, \mathbf{M}, \mathbf{N})$ depend on the bus admittance matrix $\mathbf{Y}$; see [27], [21]. Armed with a linear mapping between power injections and voltages, the design of setpoints $\{\mathbf{s}_{\mathcal{M}}^t\}_{t=1}^T$ is accomplished next in three steps.

3.2.1 Library of Inverter Implementable Probing Setpoints

The first step of the setpoint design builds a library $\mathcal{S}$ of $K \gg T$ *candidate* injection vectors indexed by k

$$\mathcal{S} := \{\mathbf{s}_{\mathcal{M}}^{k}\}_{k=1}^{K}. \tag{3.2}$$

The entries of each $\mathbf{s}_{\mathcal{M}}^{k}$ should be *implementable*, in the sense that each probing inverter should be able to inject the requested value of complex power.

To characterize the allowable range of inverter injections (p_n, q_n) with $n \in \mathcal{M}$, two inverter classes are identified. The first class consists of inverters interfacing solar panels. When inverter n interfaces a solar panel, its complex injection is limited by its apparent power capacity $\bar{s}_n$ as

$$p_n^2 + q_n^2 \leq \bar{s}_n^2. \tag{3.3}$$

Moreover, if the maximum active power that can be generated given the solar irradiance at the current probing period is $\bar{p}_n$, then its active power injection is limited by

$$0 \leq p_n \leq \bar{p}_n. \tag{3.4}$$

The second class consists of inverters interfacing energy storage units. The apparent power constraint of (3.3) should still be enforced. If the power rate of energy storage unit n is $\bar{p}_n$, the active injection from inverter n should lie within

$$-\bar{p}_n \leq p_n \leq \bar{p}_n \tag{3.5}$$

since the battery can be charged or discharged. Given the short duration of probing, limits on the state of charge have been ignored for simplicity.

Given the limitations for each inverter class, a candidate probing injection $\mathbf{s}_{\mathcal{M}}^k \in \mathcal{S}$ can be constructed by sampling uniformly at random p_n^k within (3.4)–(3.5) for all $n \in \mathcal{M}$. Upon fixing active injections, the reactive injections can be sampled again uniformly at random within $|q_n^k| \leq \sqrt{\bar{s}_n^2 - (p_n^k)^2}$ to comply with (3.3). Scenarios where a single bus hosts multiple inverters belonging to the previous two or additional classes can be incorporated in the sampling process.

As explained in Remark 2.1, a probing bus $n \in \mathcal{M}$ may be hosting controllable inverters and non-controllable assets (non-probing inverters and non-controllable loads). The process of sampling implementable injections through (3.3)–(3.5) can be repeated for all controllable inverters. The net injection from non-controllable assets is assumed to be metered; that is the case for the Pecan Street dataset [67]. The complex powers injected into bus n are summed up and used in the P2L. To keep the notation uncluttered, we will slightly abuse notation and denote this net injection at bus n as $p_n + jq_n$.

The sampling process is repeated K times to construct library $\mathcal{S}$. Although each candidate probing vector $\mathbf{s}_{\mathcal{M}}^k \in \mathcal{S}$ can be implemented by inverters, the aggregate effect of probing injections may be violating feeder constraints. To handle this concern, we next reduce library $\mathcal{S}$ to only those probing injections abiding by feeder constraints.

3.2.2 Maintaining only Network-Compliant Probing Setpoints

Even though a probing action lasts for one second or two, the operator may still want to guarantee that it does not violate any feeder constraints. For example, voltage regulation standards dictate voltage magnitudes to remain within a pre-specified range as $\underline{u} \leq u_n \leq \overline{u}$ for all $n \in \mathcal{N}^+$. A probing injection vector $\mathbf{s}_{\mathcal{M}}^k \in \mathcal{S}$ is deemed *network-compliant* if the incurred voltage deviations are maintained within the allowable range $\underline{u}\mathbf{1} \leq \mathbf{u} - u_0\mathbf{1} \leq \overline{u}\mathbf{1}$

with the inequalities applied entry-wise. Thanks to (3.1), these voltage constraints can be expressed as linear inequality constraints on $\mathbf{s}_{\mathcal{M}}^k$

$$\underline{u}\mathbf{1} \leq \mathbf{K}\mathbf{s}_{\mathcal{M}}^k + \mathbf{L}\mathbf{s}_{\mathcal{O}} \leq \overline{u}\mathbf{1}. \tag{3.6}$$

One cannot directly check whether $\mathbf{s}_{\mathcal{M}}^k$ is network-compliant, since $\mathbf{s}_{\mathcal{O}}$ is unknown. To bypass this complication, non-metered loads are assumed to lie within a known range

$$\underline{\mathbf{s}}_{\mathcal{O}} \leq \mathbf{s}_{\mathcal{O}} \leq \overline{\mathbf{s}}_{\mathcal{O}}. \tag{3.7}$$

The bounds $(\underline{\mathbf{s}}_{\mathcal{O}}, \overline{\mathbf{s}}_{\mathcal{O}})$ can be derived from historical data, the confidence intervals of load forecasts, or the load estimates obtained during the previous probing period.

Adopting a robust design, we would like to comply with the voltage constraints in (3.6) for all possible values of non-metered loads in (3.7). To do so, we leverage the next version of Farka's lemma on the containment of polytopes.

Lemma 3.1 ([30], [59], [86])**.** *The non-empty polytope* $\mathcal{P}_1 := \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ *with* $\mathbf{A} \in \mathbb{R}^{M\times N}$ *is contained within the polytope* $\mathcal{P}_2 := \{\mathbf{x} : \mathbf{C}\mathbf{x} \leq \mathbf{d}\}$ *with* $\mathbf{C} \in \mathbb{R}^{K\times N}$ *if and only if there exists matrix* $\mathbf{E} \geq \mathbf{0}$ *satisfying* $\mathbf{E}\mathbf{A} = \mathbf{C}$ *and* $\mathbf{E}\mathbf{b} \leq \mathbf{d}$.

Based on Lemma 3.1, to ensure that the polytope over $\mathbf{s}_{\mathcal{O}}$ defined in (3.7) is contained within the polytope of (3.6), we need to solve the feasibility problem

$$\begin{aligned} \text{find} \quad & \mathbf{E} \\ \text{s.to} \quad & \mathbf{E} \geq \mathbf{0} \\ & \mathbf{E}\begin{bmatrix} -\mathbf{I}_{2O} \\ \mathbf{I}_{2O} \end{bmatrix} = \begin{bmatrix} -\mathbf{L} \\ \mathbf{L} \end{bmatrix} \end{aligned} \tag{3.8}$$

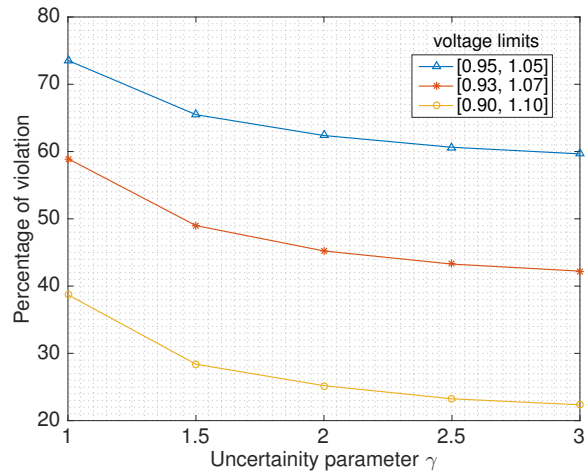

Figure 3.1: Percentage of candidate vectors that violate (3.6) for varying voltage bounds. The percentage of non-admissible vectors decreases with smaller load uncertainty and/or looser voltage regulation bounds.

$$\mathbf{E}\begin{bmatrix} -\underline{\mathbf{s}}_{\mathcal{O}} \\ \overline{\mathbf{s}}_{\mathcal{O}} \end{bmatrix} \leq \begin{bmatrix} \mathbf{K}\mathbf{s}^k_{\mathcal{M}} + (u_0 - \underline{u})\mathbf{1} \\ -\mathbf{K}\mathbf{s}^k_{\mathcal{M}} - (u_0 - \overline{u})\mathbf{1} \end{bmatrix}.$$

Given $\mathbf{s}^k_{\mathcal{M}}$, if the linear program in (3.8) is feasible, the candidate vector $\mathbf{s}^k_{\mathcal{M}}$ is deemed network-compliant and is copied to the *reduced library* $\mathcal{S}_r$. Otherwise, the candidate vector is not copied to $\mathcal{S}_r$ since there exist load values within $[\underline{\mathbf{s}}_{\mathcal{O}}, \overline{\mathbf{s}}_{\mathcal{O}}]$ that violate the voltage constraints in (3.6). The test of (3.8) is repeated for all $\mathbf{s}^k_{\mathcal{M}} \in \mathcal{S}$ to get the reduced library $\mathcal{S}_r := \{\mathbf{s}^\ell_{\mathcal{M}}\}_{\ell=1}^L$ of L candidate injection vectors with $L \leq K$.

To demonstrate the importance of this library reduction step, we ran a numerical test on the IEEE 34-bus feeder for $T = 4$ and $O = 10$; see Fig. 3.1. Load uncertainty in (3.7) was confined within $\underline{\mathbf{s}}_{\mathcal{O}} = (1 - \frac{1}{\gamma})\mathbf{s}_{\mathcal{O}}$ and $\overline{\mathbf{s}}_{\mathcal{O}} = (1 + \frac{1}{\gamma})\mathbf{s}_{\mathcal{O}}$ for $\gamma > 0$. The candidate inverter injections in $\mathcal{S}$ were randomly drawn from ± 0.2 pu and tested against (3.8). For increasing γ, the uncertainty bounds in (3.7) became tighter and progressively more candidate vectors were rendered admissible. Even for loose voltage regulation limits of $\pm 10\%$ and tight load uncertainty, more than 20% of the candidates in $\mathcal{S}$ violated (3.6).

The reduction from $\mathcal{S}$ to $\mathcal{S}_r$ via (3.8) can be generalized. For example, limits on line and transformer flows can be expressed as linear functions of power injections and appended to (3.6). Moreover, correlations in load forecasts across buses, or power factor limitations applied on a per-bus basis, both can be directly captured as linear inequalities and appended to (3.7). Finally, if the library has been reduced significantly so that $L < T$, the operator could broaden the voltage interval $[\underline{u}, \overline{u}]$ and/or tighten the load uncertainty range in (3.6) if grid probing is still needed to recover non-metered loads.

3.2.3 Finding Probing Setpoints with Most Diverse System States

Given the reduced library $\mathcal{S}_r = \{\mathbf{s}_{\mathcal{M}}^{\ell}\}_{\ell=1}^{L}$ of implementable and network-compliant candidates, the last step is to select the T candidates yielding the most diverse states. Recall that the system state $\mathbf{v}^{\ell}$ related to probing injection $\mathbf{s}_{\mathcal{M}}^{\ell}$ depends also on the unknown loads $\mathbf{s}_{\mathcal{O}}$. Moreover, the dependence on both $\mathbf{s}_{\mathcal{M}}^{\ell}$ and $\mathbf{s}_{\mathcal{O}}$ is non-linear and implicit. The approximate LDF model of (3.1) can help us circumvent these technical challenges.

The Euclidean distance between the system states induced by injections $\mathbf{s}_{\mathcal{M}}^{\ell}, \mathbf{s}_{\mathcal{M}}^{\ell'} \in \mathcal{S}_r$ will be surrogated by the Euclidean distance between the approximate states of (3.1) as

$$\|\mathbf{v}^{\ell} - \mathbf{v}^{\ell'}\|_2 \simeq \|\mathbf{y}_{\ell} - \mathbf{y}_{\ell'}\|_2$$

for all $\ell, \ell' = 1, \ldots, L$. The latter simplifies as

$$\begin{aligned}\|\mathbf{y}_{\ell} - \mathbf{y}_{\ell'}\|_2 &= \left\| \begin{bmatrix} \mathbf{K} & \mathbf{L} \\ \mathbf{M} & \mathbf{N} \end{bmatrix} \left(\begin{bmatrix} \mathbf{s}_{\mathcal{M}}^{\ell} \\ \mathbf{s}_{\mathcal{O}} \end{bmatrix} - \begin{bmatrix} \mathbf{s}_{\mathcal{M}}^{\ell'} \\ \mathbf{s}_{\mathcal{O}} \end{bmatrix} \right) \right\|_2 \\ &= \left\| \begin{bmatrix} \mathbf{K} \\ \mathbf{M} \end{bmatrix} \left(\mathbf{s}_{\mathcal{M}}^{\ell} - \mathbf{s}_{\mathcal{M}}^{\ell'} \right) \right\|_2\end{aligned}$$

where we have exploited the linearity in (3.1) together with the fact that non-metered loads remain roughly invariant during probing. We define the distance between $\mathbf{s}_{\mathcal{M}}^{\ell}, \mathbf{s}_{\mathcal{M}}^{\ell'} \in \mathcal{S}_r$ as

$$\begin{aligned} d(\ell, \ell') &:= \|\mathbf{y}_{\ell} - \mathbf{y}_{\ell'}\|_2^2 \\ &= (\mathbf{s}_{\mathcal{M}}^{\ell} - \mathbf{s}_{\mathcal{M}}^{\ell'})^{\top}(\mathbf{K}^{\top}\mathbf{K} + \mathbf{M}^{\top}\mathbf{M})(\mathbf{s}_{\mathcal{M}}^{\ell} - \mathbf{s}_{\mathcal{M}}^{\ell'}). \end{aligned} \tag{3.9}$$

Based on this metric, we would like to select a subset $\mathcal{A}$ of T out of the L candidate vectors in $\mathcal{S}_r$ so that the sum of their pairwise distances is maximized

$$\begin{aligned} \max_{\mathcal{A} \subset \mathcal{S}_r} \quad & \sum_{\ell \in \mathcal{A}} \sum_{\ell' \in \mathcal{A}} d(\ell, \ell') \\ \text{s.to} \quad & |\mathcal{A}| = T. \end{aligned} \tag{3.10}$$

The task in (3.10) is known as the *max-sum diversity* (MSD) problem, and appears frequently in information retrieval, computational geometry, and operations research [22]. In fact, MSD can be reformulated as a binary quadratic program (QP) after introducing the $L \times L$ distance matrix $\mathbf{D}$ with entries $D_{\ell,\ell'} := d(\ell, \ell')$ as

$$f^{\star} := \max_{\mathbf{x} \in \{0,1\}^L} \quad \mathbf{x}^{\top}\mathbf{D}\mathbf{x} \tag{3.11a}$$

$$\text{s.to} \quad \mathbf{x}^{\top}\mathbf{1} = T. \tag{3.11b}$$

Despite its simple form, the MSD task is NP-hard [22]. However, thanks to the properties of $\mathbf{D}$, the problem in (3.11) enjoys a polynomial-time approximate scheme (PTAS) [22].

Although $\mathbf{D}$ is indefinite, the objective in (3.11a) can be shown to be concave under constraint (3.11b). To see this, define the $2N \times L$ matrix $\tilde{\mathbf{Y}} := [\mathbf{y}_1 \ \cdots \ \mathbf{y}_L]$ and use the definition of

$d(\ell, \ell')$ to rewrite the objective of (3.11) as

$$
\begin{aligned}
f(\mathbf{x}) := \mathbf{x}^\top \mathbf{D}\mathbf{x} &= \sum_{\ell=1}^{L}\sum_{\ell'=1}^{L} x_\ell x_{\ell'} D_{\ell,\ell'} \\
&= \sum_{\ell=1}^{L}\sum_{\ell'=1}^{L} x_\ell x_{\ell'} \left(\|\mathbf{y}_\ell\|_2^2 + \|\mathbf{y}_{\ell'}\|_2^2 - 2\mathbf{y}_\ell^\top \mathbf{y}_{\ell'}\right) \\
&= \sum_{\ell'=1}^{L} x_{\ell'} \|\mathbf{y}_{\ell'}\|_2^2 \left(\sum_{\ell=1}^{L} x_\ell\right) + \sum_{\ell=1}^{L} x_\ell \|\mathbf{y}_\ell\|_2^2 \left(\sum_{\ell'=1}^{L} x_{\ell'}\right) \\
&- 2\sum_{\ell=1}^{L}\sum_{\ell'=1}^{L} x_\ell x_{\ell'} \mathbf{y}_\ell^\top \mathbf{y}_{\ell'} \\
&= 2T\mathbf{c}^\top \mathbf{x} - 2\mathbf{x}^\top \tilde{\mathbf{Y}}^\top \tilde{\mathbf{Y}}\mathbf{x}.
\end{aligned}
$$

where $\mathbf{c} := [\|\mathbf{y}_1\|_2^2 \ \cdots \ \|\mathbf{y}_L\|_2^2]^\top$. Since $\tilde{\mathbf{Y}}^\top \tilde{\mathbf{Y}} \succeq \mathbf{0}$, the objective $f(\mathbf{x})$ equals a concave quadratic function.

For moderate L (a few hundreds), the task in (3.11) can be handled by a mixed-integer QP solver. For $T = 2$, the MSD solution can be found by an exhaustive search. For larger T, we will use a randomized rounding approach, as adopted from [70] in [22, Remark 2]. The approach is briefly reviewed here for completeness. Its first step solves the relaxed problem

$$
\hat{\mathbf{x}} := \arg\min_{\mathbf{0}\le\mathbf{x}\le\mathbf{1}} \quad 2\mathbf{x}^\top \tilde{\mathbf{Y}}^\top \tilde{\mathbf{Y}}\mathbf{x} - 2T\mathbf{c}^\top \mathbf{x} \tag{3.12a}
$$

$$
\text{s.to} \quad \mathbf{x}^\top \mathbf{1} = T. \tag{3.12b}
$$

Since the binary constraints of (3.11) are related to box constraints in (3.12), it holds that $f(\hat{\mathbf{x}}) \ge f^\star$. To construct a point $\tilde{\mathbf{x}}$ that is feasible for (3.11), draw L-dimensional vectors $\{\tilde{\mathbf{x}}_i\}$ whose entries are independent Bernoulli random variables with mean $(1-\beta)\hat{\mathbf{x}}$ for some $\beta > 0$, say $\beta = 0.1$. The so constructed binary vectors $\tilde{\mathbf{x}}_i$'s satisfy $\mathbb{E}[\tilde{\mathbf{x}}_i^\top \mathbf{1}] = (1-\beta)T$ and $\mathbb{E}[\tilde{\mathbf{x}}_i^\top \mathbf{D}\tilde{\mathbf{x}}_i] = (1-\beta)^2 \hat{\mathbf{x}}^\top \mathbf{D}\hat{\mathbf{x}}$. The purpose of scaling $\hat{\mathbf{x}}$ by $(1-\beta)$ is to ensure $\tilde{\mathbf{x}}_i$'s are both

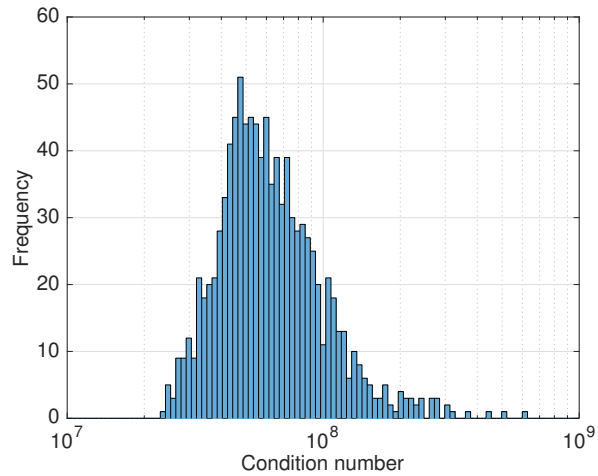

Figure 3.2: Histogram of condition numbers for the Jacobian matrix $\mathbf{J}(\{\mathbf{v}_t\}_{t=1}^4)$ obtained by randomly sampling quadruplets of $\mathbf{s}_{\mathcal{M}}$'s from $\mathcal{S}_r$.

feasible for (3.11) and yield relatively high cost with significant probability [22].

Let us now comment on the complexity for designing probing setpoints. The first step described in Section 3.2.1 is computationally inexpensive. The second step of Section 3.2.2 involves solving the linear program in (3.8) K times, once for each candidate setpoint vector. The third step of Section 3.2.3 entails solving the linearly-constrained quadratic program of (3.12), whose complexity is cubic in the number of variables L. As detailed later in Section 3.4, running this design process for the IEEE 34-bus feeder and $K = 100$ candidate setpoints took $1 - 1.5$ min depending on (M, O). The tests were run on a laptop computer using generic off-the-shelf solvers.

To justify the need for this third step in probing design, we conducted a test on the IEEE 34-bus feeder for $T = 4$ and $O = 6$. For this test, load uncertainty was confined within a factor of ± 1 times the nominal loads. Given library $\mathcal{S}$ of randomized injections drawn from ± 0.2 pu and obeying (3.3)–(3.5), we constructed the reduced library $\mathcal{S}_r$ based on (3.8) for $[\underline{u}, \overline{u}] = [0.90, 1.10]$. We then solved (3.12) and followed the randomized rounding process to construct 100 binary $\tilde{\mathbf{x}}_i$'s. We evaluated the cost $f(\tilde{\mathbf{x}}_i)$ for those $\tilde{\mathbf{x}}_i$'s satisfying $\tilde{\mathbf{x}}_i^\top \mathbf{1} = T$,

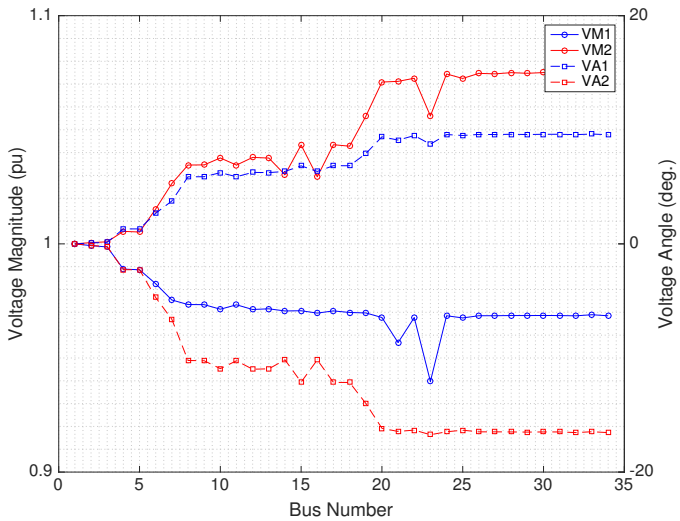

Figure 3.3: Probing setpoint design on the IEEE 34-bus grid for $T = 2$ and $O = 6$. The blue (red) lines correspond to the system states induced by the first (second) probing setpoints $\{\mathbf{s}^t_{\mathcal{M}}\}_{t=1}^2$. Solid lines depict voltage magnitudes, and dashed lines voltage angles.

and returned the $\tilde{\mathbf{x}}_i$ yielding the largest cost. The condition number of the Jacobian matrix evaluated at the so obtained $\tilde{\mathbf{x}}_i$ was $3 \cdot 10^6$. We also calculated the condition number of the Jacobian matrix evaluated at random candidate quadruplets in $\mathcal{S}_r$. The latter condition numbers ranged within $10^7 - 10^9$; see Fig. 3.2. Hence, albeit MSD adds computational complexity, it is an important part of the probing design process.

To show that this MSD step provides diverse system states, we performed another test on the IEEE 34-bus feeder under the same setup, but for $T = 2$. The states induced by the designed probing and nominal loads are shown in Figure 3.3.

Upon solving (3.10) near optimally, we have obtained T probing injection vectors $\{\mathbf{s}^t_{\mathcal{M}}\}_{t=1}^T$ that are: *i)* implementable by inverters; *ii)* network-compliant; and *iii)* yield diverse system states. In the process of probing design, the first step (Section 3.2.1) operates on the entries of $\mathbf{s}^t_{\mathcal{M}}$'s; the second step (Section 3.2.2) considers each vector $\mathbf{s}^t_{\mathcal{M}}$ as a whole; and the third step (Section 3.2.3) accounts for the joint effect of probing injections $\{\mathbf{s}^t_{\mathcal{M}}\}_{t=1}^T$.

3.3 Solving the P2L tasks

Recall that the P2L task with phasor data involves solving the set of non-linear equations

$$u_n(\mathbf{v}_t) = u_n^t \qquad \forall n \in \mathcal{M}, t \in \mathcal{T} \tag{3.13a}$$

$$\theta_n(\mathbf{v}_t) = \theta_n^t \qquad \forall n \in \mathcal{M}, t \in \mathcal{T} \tag{3.13b}$$

$$p_n(\mathbf{v}_t) = p_n^t \qquad \forall n \in \mathcal{M}, t \in \mathcal{T} \tag{3.13c}$$

$$q_n(\mathbf{v}_t) = q_n^t \qquad \forall n \in \mathcal{M}, t \in \mathcal{T} \tag{3.13d}$$

$$p_n(\mathbf{v}_t) = p_n(\mathbf{v}_{t+1}) \qquad \forall n \in \mathcal{O}, t \in \mathcal{T}' \tag{3.13e}$$

$$q_n(\mathbf{v}_t) = q_n(\mathbf{v}_{t+1}) \qquad \forall n \in \mathcal{O}, t \in \mathcal{T}' \tag{3.13f}$$

where $\mathbf{v}_t$'s are the system states across $\mathcal{T} = \{1, \ldots, T\}$; $\{(u_n^t, \theta_n^t, p_n^t, q_n^t)\}_{n \in \mathcal{M}}$ are the probing data collected at time t; (3.13a)–(3.13d) are the $4MT$ metering equations; and (3.13e)–(3.13f) are the $2O(T-1)$ coupling equations with $\mathcal{T}' := \{1, \ldots, T-1\}$. For the P2L task with non-phasor data, the angle information in (3.13b) is unavailable.

Having characterized the local identifiability for the P2L tasks, this section presents solvers for tackling P2L. If grid specifications are noiseless and the power injections in $\mathcal{O}$ remain unaltered during probing, the P2L tasks boil down to solving the equations in (3.13). The latter can be tackled by adopting the SDP-based solvers developed in [88], [51], [56], [55].

The key idea of the SDP-based solver is that the PF specifications in (2.1b)–(2.1c) are quadratic functions of $\mathbf{v}$ and therefore can be expressed as $\mathbf{v}^H \mathbf{M}_k \mathbf{v} = \hat{s}_k$ for a specific Hermitian matrix $\mathbf{M}_k$ for all k [51], [56]. By introducing the matrix variable $\mathbf{V} = \mathbf{v}\mathbf{v}^H$, the specifications can be equivalently written as $\mathrm{Tr}(\mathbf{M}_k \mathbf{V}) = \hat{s}_k$ and the PF task can be posed

as the feasibility problem

$$\text{find} \quad (\mathbf{V}, \mathbf{v}) \tag{3.14a}$$

$$\text{s.to} \quad \mathrm{Tr}(\mathbf{M}_k \mathbf{V}) = \hat{s}_k, \quad k = 1, \ldots, 2N+2 \tag{3.14b}$$

$$\mathbf{V} = \mathbf{v}\mathbf{v}^H. \tag{3.14c}$$

Although the constraints in (3.14b) are linear with respect to $\mathbf{V}$, the constraint in (3.14c) is non-convex. To see this, note that $\mathbf{V} = \mathbf{v}\mathbf{v}^H$ is equivalent to $\mathbf{V} \succeq \mathbf{0}$ and $\mathrm{rank}(\mathbf{V}) = 1$. Problem (3.14) can be relaxed by dropping the non-convex rank constraint and replacing $\mathbf{V} = \mathbf{v}\mathbf{v}^H$ with $\mathbf{V} \succeq \mathbf{0}$. To locate feasible points of rank-one, the feasibility problem is turned into the SDP minimization task [56]

$$\min_{\mathbf{V} \succeq \mathbf{0}} \quad \mathrm{Tr}(\mathbf{M}\mathbf{V}) \tag{3.15a}$$

$$\text{s.to} \quad \mathrm{Tr}(\mathbf{M}_k \mathbf{V}) = \hat{s}_k, \quad k = 1, \ldots, 2N+2. \tag{3.15b}$$

The matrix $\mathbf{M}$ is judiciously selected to provide rank-one minimizers if the power system operates close to the flat voltage profile. Two practical choices for $\mathbf{M}$ are $\mathbf{G}$ and $-\mathbf{B}$ yielding PF states with minimal system losses and avoiding low-voltage solutions. If problem (3.15) yields a rank-one minimizer with eigenvalue decomposition $\mathbf{V}^* = \lambda \mathbf{u}\mathbf{u}^H$, a PF solution is recovered as $\mathbf{v} = \sqrt{\lambda}\mathbf{u}$.

Building on (3.15), the P2L task can be posed as

$$\min_{\mathbf{V}_t \succeq \mathbf{0}} \quad \sum_{t=1}^{T} \mathrm{Tr}(\mathbf{M}\mathbf{V}_t) \tag{3.16a}$$

$$\text{s.to} \quad \mathrm{Tr}(\mathbf{M}_k \mathbf{V}_t) = \hat{s}_k^t, \quad k = 1:3M, t \in \mathcal{T} \tag{3.16b}$$

$$\mathrm{Tr}(\mathbf{M}_l \mathbf{V}_t) = \mathrm{Tr}(\mathbf{M}_l \mathbf{V}_{t+1}), \quad l = 1:2O, t \in \mathcal{T}' \tag{3.16c}$$

where the matrix variables $\mathbf{V}_t \succeq \mathbf{0}$ have been obtained upon relaxing the rank-one constraint $\mathbf{V}_t = \tilde{\mathbf{v}}_t \tilde{\mathbf{v}}_t^H$ on the original system states for $t \in \mathcal{T}$. The measurements $\hat{s}_k^t$ relate to state $\mathbf{v}_t$ in (3.16b); and the constraints in (3.16c) couple the T states.

3.3.1 Noisy Measurements

Probing data are inexact due to measurement noise and modeling inaccuracies in the metering equations of (3.13a)–(3.13d). To account for small fluctuations in non-metered loads during probing, a noise term is added to the RHS of the coupling equations in (3.13e)–(3.13f). To cope with noisy data, we extend the penalized SDP-based state estimator of [55] to the P2L setting as follows

$$\min \ \alpha \sum_{t=1}^{T} \mathrm{Tr}(\mathbf{M}\mathbf{V}_t) + \sum_{t=1}^{T} \sum_{k=1}^{3M} f_k(\epsilon_k^t) + \sum_{t=1}^{T-1} \sum_{l=1}^{2O} f_l(\xi_l^t) \tag{3.17a}$$

$$\text{over} \ \ \mathbf{V}_t \succeq \mathbf{0}, \{\epsilon_k^t\}_{k=1}^{3M}, \quad t \in \mathcal{T} \tag{3.17b}$$

$$\{\xi_l^t\}_{l=1}^{2O}, \quad t \in \mathcal{T}' \tag{3.17c}$$

$$\text{s.to} \ \ \mathrm{Tr}(\mathbf{M}_k \mathbf{V}_t) + \epsilon_k^t = \hat{s}_k^t, \quad k = 1:3M, t \in \mathcal{T} \tag{3.17d}$$

$$\mathrm{Tr}(\mathbf{M}_l \mathbf{V}_t) = \mathrm{Tr}(\mathbf{M}_l \mathbf{V}_{t+1}) + \xi_l^t, \quad l = 1:2O, t \in \mathcal{T}' \tag{3.17e}$$

The auxiliary variables ϵ_k^t can be substituted from (3.17d)–(3.17e) into the objective of (3.17); they are introduced here only to simplify notation. The data fitting penalties f_k can be either a weighted squared or absolute value, that is

$$f_k(\epsilon_k^t) = \left(\frac{\epsilon_k^t}{\sigma_k^2}\right)^2 = \frac{(\hat{s}_k^t - \mathrm{Tr}(\mathbf{M}_k \mathbf{V}_t))^2}{\sigma_k^2} \quad \text{or}$$
$$f_k(\epsilon_k^t) = \frac{|\epsilon_k^t|}{\sigma_k} = \frac{|\hat{s}_k^t - \mathrm{Tr}(\mathbf{M}_k \mathbf{V}_t)|}{\sigma_k}$$

with different σ_k's depending on the uncertainty of the k-th datum. Likewise, the auxiliary variables ξ_l^t's capture variations of non-metered loads and are penalized through f_l's, which are defined as f_k's.

The first summand in (3.17a) corresponds to a regularizer promoting rank-one minimizers for $\mathbf{V}_t$. The second and third summands in (3.17a) are data-fitting terms. The tuning parameter $\alpha > 0$ governs the balance between the regularizer and the data-fitting terms: For $\alpha = 0$, the P2L cost involves only the data-fitting terms; whereas for increasing α, more emphasis is placed on the regularizer [55]. If one or more of the minimizers $\mathbf{V}_t^\star$ of (3.17) is not rank-one, the heuristic for constructing a system state $\mathbf{v}_t^\star$ proposed in [55] is used.

Additional constraints can be added to strengthen the SDP relaxation. For example, if non-metered buses are known to host exclusively loads, the constraints $\mathrm{Tr}(\mathbf{M}_l\mathbf{V}_t) \leq 0$ for $l = 1 : 2O$, and $t \in \mathcal{T}$ can be appended to (3.17). Additional information on loads, such as the uncertainty range of (3.7), can be readily incorporated. As in [55], if bus n is known to be a zero-injection bus, then $\tilde{i}_n = \mathbf{e}_n^\top \mathbf{Y}\tilde{\mathbf{v}}$ has to be zero. Therefore, the constraint $\tilde{\mathbf{v}}\tilde{i}_n^\star = \mathbf{V}\mathbf{Y}^\star\mathbf{e}_n = \mathbf{0}$ can be added.

Given phasor data, the metering equations corresponding to voltage magnitudes can be dropped. If the vectors of voltage phasors $\{\tilde{\mathbf{v}}_t\}_{t=1}^T$ are included as optimization variables, the direct measurements on the voltage phasors of $\mathcal{M}$ can be simply expressed as

$$\tilde{v}_{t,k} + \epsilon_k^t = \hat{s}_k^t, \quad k = 1 : M,\ t \in \mathcal{T}. \tag{3.19}$$

To capture the dependence between $\tilde{\mathbf{v}}_t$ and $\tilde{\mathbf{V}}_t$, the non-convex constraint

$$\mathrm{rank}\left(\begin{bmatrix} \mathbf{V}_t & \tilde{\mathbf{v}}_t \\ \tilde{\mathbf{v}}_t^H & 1 \end{bmatrix}\right) = 1$$

can be surrogated by the next SDP constraint as in [88]

$$\begin{bmatrix} \mathbf{V}_t & \tilde{\mathbf{v}}_t \\ \tilde{\mathbf{v}}_t^H & 1 \end{bmatrix} \succeq \mathbf{0}, \quad t \in \mathcal{T}. \tag{3.20}$$

Since the $\tilde{\mathbf{v}}_t$'s are optimization variables now, there is no need to use the heuristic of [55] to recover the system states.

3.4 Numerical Tests

The topological observability criteria for the P2L task and the SDP-based solvers were numerically tested using the IEEE 34-bus feeder. The original multi-phase grid was converted to an equivalent single-phase grid [37]. The numerical tests were run on a 2.7 GHz Intel Core i5 laptop computer with 8 GB RAM using the Sedumi solver on YALMIP and MATLAB [73], [54].

3.4.1 Numerical Observability

Since Theorems 2.8 and 2.13 rely on the sparsity pattern rather than the exact values of $\mathbf{J}(\{\mathbf{v}_t\})$, we evaluated $\mathbf{J}(\{\mathbf{v}_t\})$ for 1,000 random state sequences $\{\mathbf{v}_t\}_{t=1}^T$. The scenarios of phasor and non-phasor data were tested under four probing setups. For each setup, the placement of non-metered $\mathcal{O}$ and probing buses $\mathcal{M}$ were fixed. We generated 1,000 random state sequences by randomly drawing voltage magnitudes in the range $[0.90, 1.10]$ per unit and voltage angles in the range $[-1.5, 1.5]^\circ$. Assuming non-phasor data first, the following four setups were constructed according to the condition of Th. 2.13:

- Setup A meets the condition for $O = 16$ and $T = 2$.

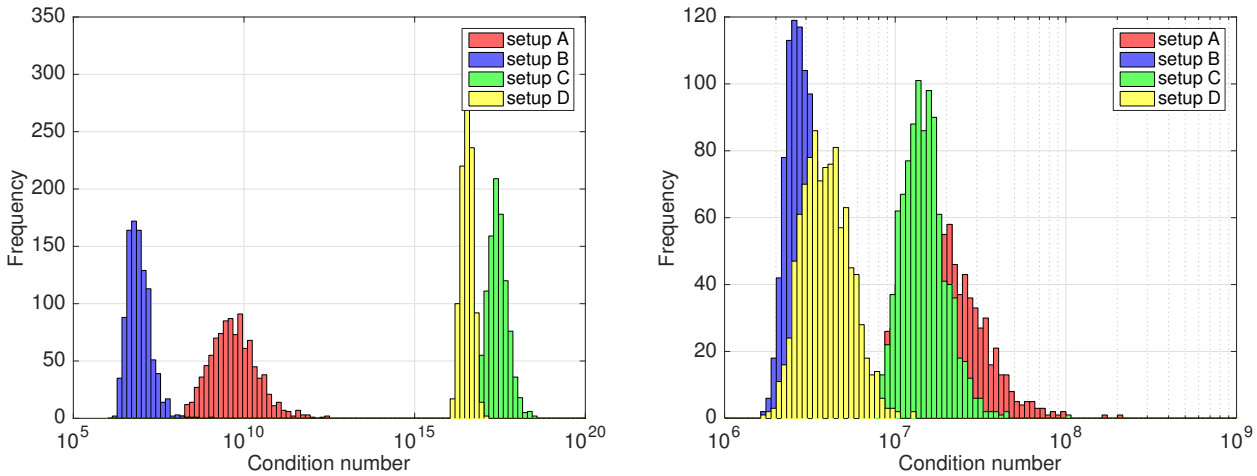

Figure 3.4: Histograms of the condition numbers for the P2L Jacobian matrices with non-phasor data for $T = 2$ (left) and $T = 4$ (right) probing actions.

- Setup B meets the condition for $O = 6$ and $T = 2$.
- Setup C does not meet the condition for $O = 16$ and $T = 2$, but it does for $T = 4$.
- Setup D does not meet the condition for $O = 6$ and $T = 2$, but it does for $T = 4$.

The same setups were considered for phasor data. As discussed in Sections 2.3-2.4, setups A and B meet also the condition of Th. 2.8. Additionally, setups C and D were constructed such that they meet the condition of Th. 2.8 for $T = 2$.

Non-phasor data: Figure 3.4 depicts the condition number histograms obtained under the four setups for $T = 2$ and 4. Under setups A and B, although the dimensions of $\mathbf{J}(\{\mathbf{v}_t\})$ increase with T, the condition numbers did not. In fact, the condition number was sometimes reduced, especially in networks with large $\mathcal{O}$. For setups C and D, there was a significant shift in the histograms from $T = 2$ to $T = 4$, which validates Th. 2.13. By and large, the condition number improves for decreasing O and increasing T. Hence, when more loads are to be recovered, longer probing periods should be used. Of course, longer probing periods may violate the stationarity assumption on loads.

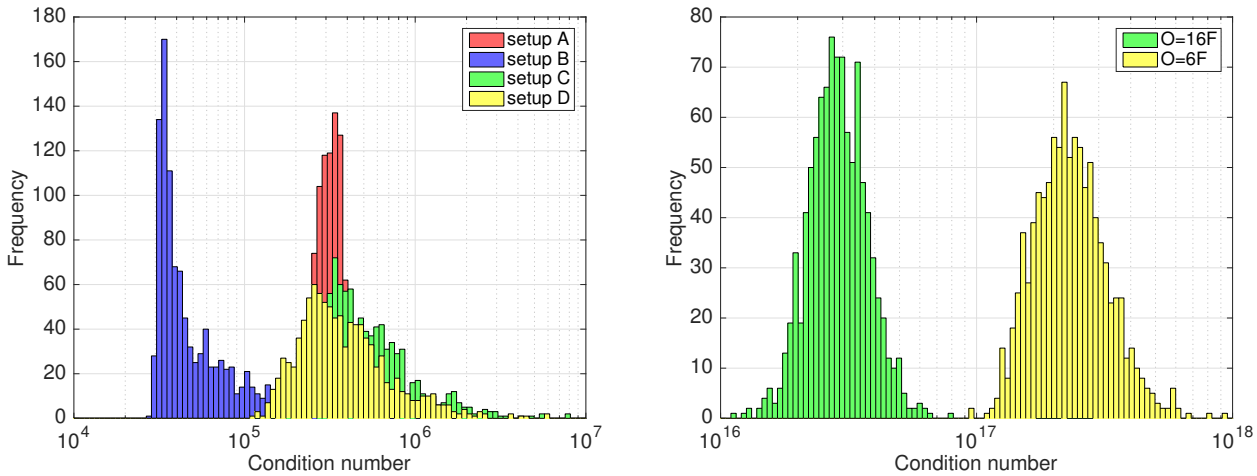

Figure 3.5: Histograms of the condition numbers for the P2L Jacobian matrices with phasor data for $T = 2$ (left) and $T = 4$ (right) probing actions.

Phasor data: Figure 3.5 displays the condition number histograms of $\mathbf{J}(\{\mathbf{v}_t\})$ again for $T = 2$ and 4. As expected, due to the value added of phasor data, the condition numbers decrease significantly. In addition, setups C and D that failed for $T = 2$ with non-phasor data, become successful with phasor probing data. The tests corroborate the criteria of Th. 2.8. The right panel of Figure 3.5 displays the condition number histograms under the following two setups that did not satisfy the condition of Th. 2.8: *i)* for $T = 4$ and $O = 6$) (yellow histogram); and *ii)* for $T = 4$ and $O = 16$ (green histogram).

Single-slot probing scenario: We also tested the special case of $T = 1$, where one fixes voltages and injections on a subset of buses $\mathcal{M}$ and tries to find the loads at the remaining buses $\mathcal{O}$. This setup is pertinent to learning ZIP loads as discussed in Section 2.5. We tested two fixed placements of non-metered buses that met the conditions of Theorems 2.15 and 2.16, respectively. We then evaluated $\mathbf{J}(\mathbf{v}_1)$ at $1,000$ random system states. Figure 3.6 shows the histograms for the condition numbers of $\mathbf{J}(\mathbf{v}_1)$. Bus placements that did not meet the criteria of Th. 2.15 and 2.16 exhibited condition numbers similar to those in the right panel of Figure 3.5.

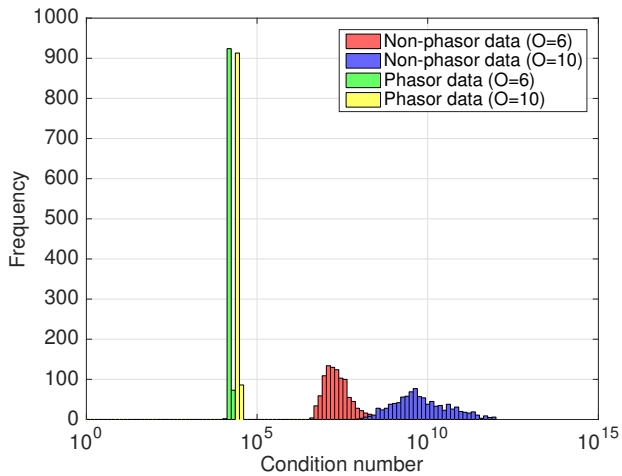

Figure 3.6: Histograms of the condition numbers for the Jacobian matrix $\mathbf{J}(\mathbf{v}_1)$ with phasor and non-phasor data for single-slot probing ($T = 1$).

The condition number of the Jacobian matrices in PSSE tasks for transmission systems is known to depend heavily on the specification set [43], [31]: A larger number of voltage magnitude and line flow measurements tends to yield a lower condition number. It is thus expected that adding line flow measurements would improve load and state estimation.

3.4.2 SDP-based P2L

Given noisy specifications, the P2L tasks were tackled using actual data and the SDP-based solver of (3.17)–(3.20). The loads on the IEEE 34-bus grid were taken from the Pecan Street dataset [67], between 10:00 a.m. and 01:40 p.m. on January 1, 2013, and in 10-minute intervals. Load sequences were scaled so that the peak active load over the tested period was 0.5 pu. Lacking values for reactive loads, a lagging power factor of 0.9 was simulated for all loads.

To simulate probing injections at buses in $\mathcal{M}$, we first created a data library $\mathcal{S}$ of $K = 100$ randomized injection vectors as described in Section 3.2.1 for $\bar{p}_n = 0.2$ pu. The library $\mathcal{S}$ was then reduced to $\mathcal{S}_r$ to ensure that voltage magnitudes lie within $[0.90, 1.10]$ pu for non-

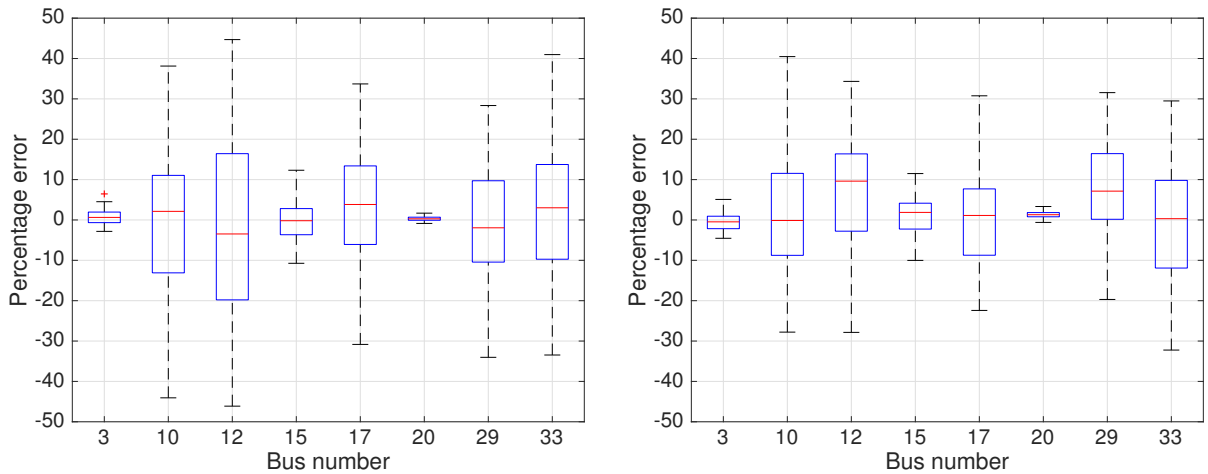

Figure 3.7: Percentage error in active power injection estimates with phasor data for $T = 4$ and $O = 8$ without MSD (left) and with MSD (right).

metered loads within $[\mathbf{0}, 2\mathbf{s}_{\mathcal{O}}]$ as described in Section 3.2.2. For all tests, the regularization parameter was set to $\alpha = 20,000$, and the functions f_k and f_l in (3.17) were selected as the WLS costs. To simulate measurement noise, the probing data recorded for an actual quantity x (e.g., voltage magnitude or power injection) was modeled as $\hat{x} = x(1+\epsilon)$, where ϵ is a zero-mean Gaussian random variable. The variance σ^2 of ϵ was selected to yield the desired value of signal-to-noise ratio (SNR)

$$10 \cdot \log_{10} \frac{x^2}{\mathbb{E}[x^2\epsilon^2]} = -20 \cdot \log_{10} \sigma. \tag{3.21}$$

This variance is the same variance appearing in (3.17) as σ_k^2. Likewise, to capture small load variations, non-metered loads were simulated by perturbing their nominal value p_n as $\hat{p}_n^t = (1+\epsilon_n)p_n$ for $t \in \mathcal{T}$, and similarly for q_n's.

To check whether the MSD step of Section 3.2.3 improves estimation, we tested P2L with and without this step. The test considered 100 Monte Carlo realizations for the loads at 10:00 a.m. The P2L task was run for $T = 4$, $O = 8$, and using phasor data. The SNR values were set respectively to 80 and 60 dB for metered and non-metered buses. PMUs are

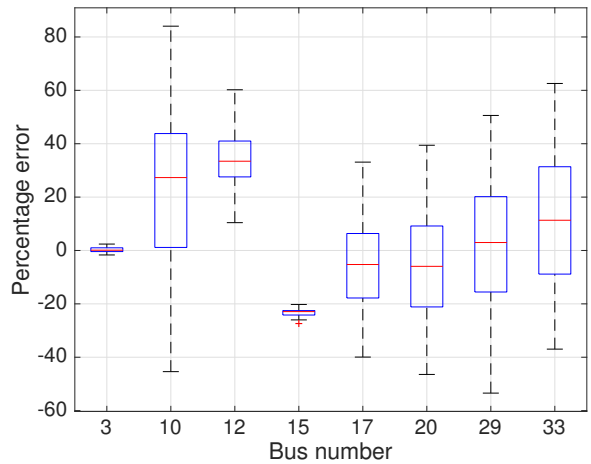

Figure 3.8: Percentage error in active power injection estimates with non-phasor data for $T = 4$ and $O = 8$.

expected to have such high accuracy [36]. The range of percentage errors was reduced from $[-50, +50]\%$ to $[-30, +40]\%$ by selecting the T most diversifying setpoints.

To verify the improvement of using phasor over non-phasor probing data, we repeated the previous MSD setup but now for non-phasor data. The obtained percentage errors are depicted in Figure 3.8 and are of worse accuracy compared to those in the right panel of Figure 3.7. We also tested the single-slot probing scenario of $T = 1$ under slightly different probing setups for (non)-phasor data. Figure 3.9 illustrates the statistics of the obtained percentage errors.

Remark 3.2. Based on the numerical tests, we have observed that load estimates generally improve when: a) the MSD step is implemented; b) phasor data are utilized; c) the duration T is increased; and d) O is decreased.

We next evaluated how the estimation accuracy of system states depends on the SNR. For this test, the SNRs for non-metered loads and probing data were identical. The state estimation accuracy was evaluated in terms of the root mean square error (RMSE) defined as $\sqrt{\sum_{t=1}^{T} \|\mathbf{v}_t - \hat{\mathbf{v}}_t\|_2^2/(NT)}$ averaged over 20 Monte Carlo tests. Figure 3.10 shows how the

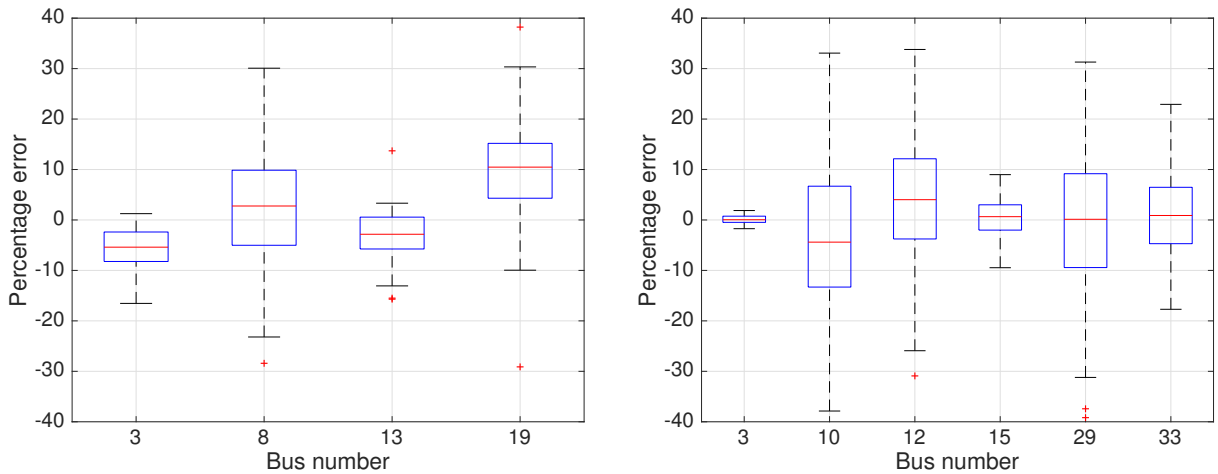

Figure 3.9: Percentage error in active power injection estimates for $T = 1$ with non-phasor data and $O = 4$ (left); and with phasor data and $O = 6$ (right).

RMSE decreases for increasing SNR.

To validate P2L over different loading conditions, we ran numerical tests for the period of 10:00 a.m. and 01:40 p.m. and every 10 min using phasor data. The SNRs for probing data and non-metered loads were again fixed to 80 and 60dB, respectively. Figure 3.11 presents the actual and estimated non-metered (re)active loads on buses $\{4, 6, 15, 20, 27, 31\}$ for $T = 4$ and $T = 6$. The plots show the load estimation improvement by increasing T.

Regarding the runtime of our algorithms, each P2L task took between $95 - 180$ sec, which were allocated as follows:

- The linear programs of (3.8) took 70 sec overall, to check the feeder compliance of $K = 100$ candidate probing setpoints.
- The quadratic program of (3.12) needed to select the T most diversifying setpoints was solved in less than 10 sec.
- The SDP formulation of (3.17) together with the heuristic of [55] to obtain a rank-one solution took $25 - 100$ sec.

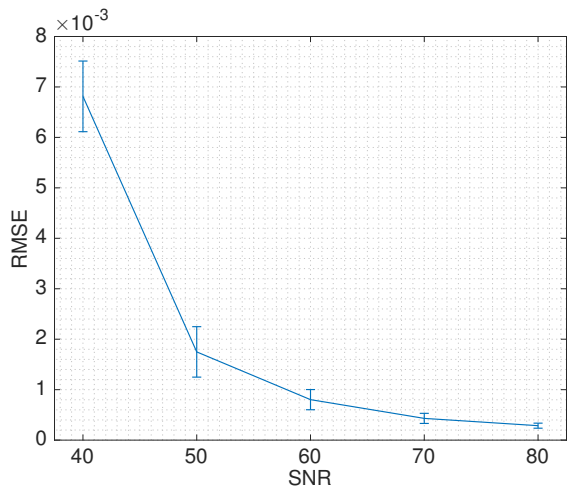

Figure 3.10: RMSE for the system state with non-phasor data for $T = 2$.

The load learning task for the single-slot probing setup ($T = 1$) was solved in less than 15 sec.

3.5 Conclusions

We presented the novel data acquisition scheme of probing an electric grid via smart inverters to infer non-metered loads. In Chapter 2 we studied the topological observability of grid probing using (non)-phasor data in potentially meshed networks. Given a probing setup is deemed topologically observable, in this chapter we presented a systematic methodology for designing probing injections. The goal was improved estimation accuracy and adherence to inverter and feeder constraints even without knowing non-metered loads. We also tackled the computational tasks involved in grid probing via penalized SDP-based solvers that accounted for noisy measurements and non-stationary loads.

Numerical tests using synthetic and real-world data on benchmark feeders demonstrate the ensuing take-away simulation findings: *i)* High-accuracy phasor data are better for load recovery than non-phasor data; *ii)* Having the most diverse system states during probing

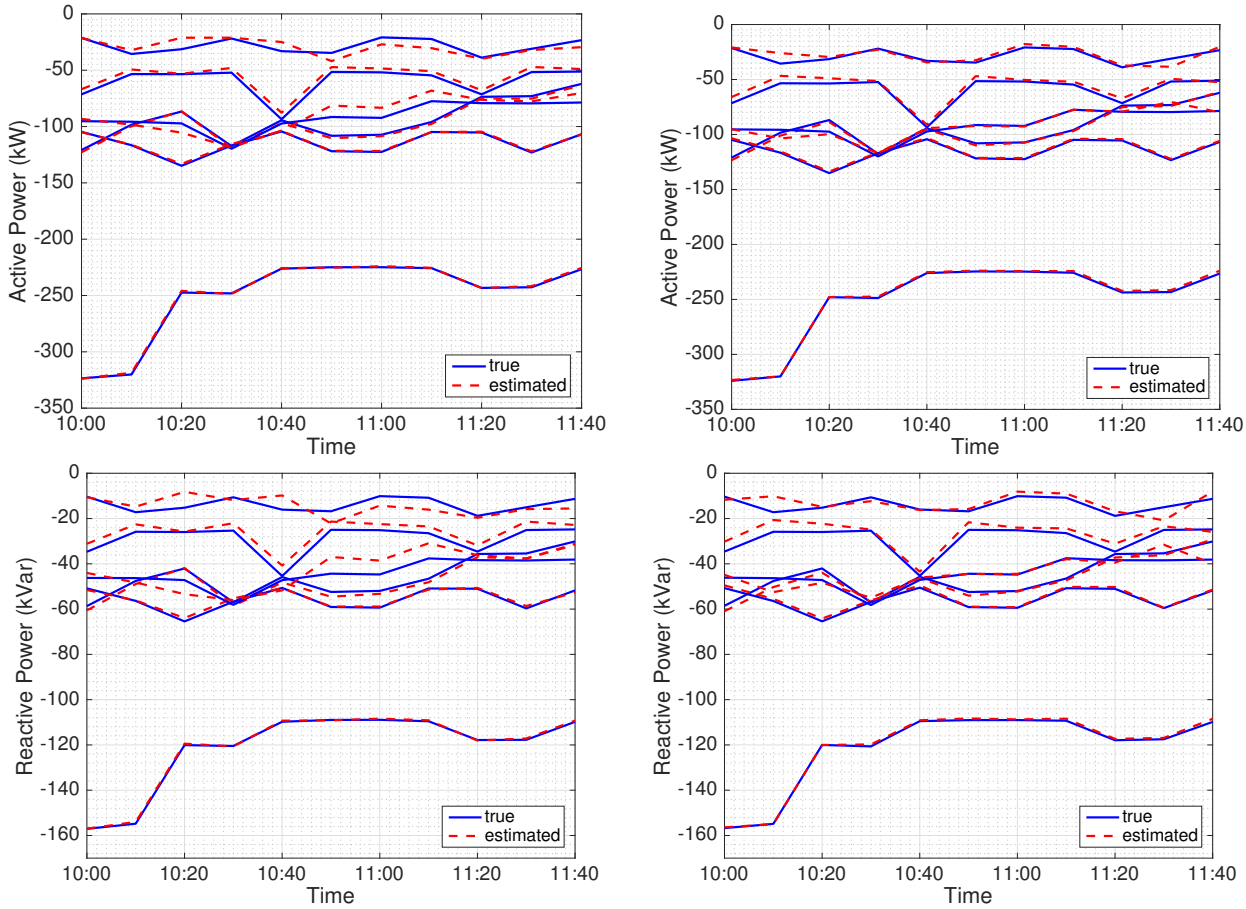

Figure 3.11: Active power injection estimates using probing with phasor data for $T = 4$ (top left) and $T = 6$ (top right); similarly for reactive power injection estimates.

yields better load estimates; *iii)* Probing seemed to yield better estimates under broad voltage regulation range and tight load uncertainty *iv)* Although increasing T improved the system state accuracy, the obtained load estimates were not always better, especially for larger O. Nevertheless, we were able to recover a reasonable number of loads; and *v)* Including the extra constraints to strengthen the SDP relaxation provided better numerical accuracy.

Chapter 4

Optimal Topology Design for Disturbance Minimization

4.1 Introduction

The electric power system is continually changing. It is expected that the grid of the future will have higher variability due to renewables, changing load patterns, and distributed energy sources [77]. This paradigm shift will pose an enormous challenge for design and stable operation of power networks. The inherent uncertainty associated with renewable energy sources and active loads is likely to produce more frequent and higher amplitude disturbances [77]. In addition, owing to the lower aggregate inertia of systems with high penetration of renewables, the capability of power networks to handle such disturbances may be significantly reduced [83].

Thus, improving the dynamic performance of the power grid is of importance and has received greater attention from academia and industry. Efforts in this direction include development of payment structures and novel markets, as well as analysis of techniques to incentivize load-side participation [63], [78], [34], [61]. The benefits of load-side controllers has motivated a series of recent works to understand how different system parameters and controller designs impact the transient response of the network [57], [44]. Recent work in [68] has also explored the optimal placement of virtual inertia to improve stability.

Compared to the previously mentioned system parameters, the effect of network topology on transient stability is less well understood. Without detailed simulations, it is usually hard to infer how a change in network topology influences the overall grid behavior and performance. Recent work in [44] shows that the impact of network topology on the power system can be quantified through the network Laplacian matrix eigenvalues. In addition, grid robustness against low frequency disturbances is mostly determined by network connectivity [44], further motivating this study. Past studies in the power and control systems communities have also looked at designing network topologies for specific goals using system theoretic tools. Such goals include reduction of transient line losses [77], improvement in feedback control [76], [58], coherence based network design [32] and augmentation [74]. Semidefinite programming (SDP) based tools have also been utilized to design and augment network topologies for dynamic control [39], [32], [65].

While the primary focus here is topology design, we recognize that there is line of related work dealing with learning network topologies and line parameters [21], [20], [27]. Schemes that rely on passive data have been used in [27] and [21] for learning radial topologies. Different from them, the work in [20] actively probes the grid to recover radial topologies and verify line statuses.

In this chapter we are interested in studying the effect of topology on the power grid dynamics. For a variety of objective functions, such as line loss reduction, fast damping of oscillations, and network coherence, reference [28] presented a unified framework to study topology design based on the $\mathcal{H}_2$-norm. In [28], the focus was on topology reconfiguration rather then topology design. Further, the work in [28] developed suboptimal algorithms, albeit with guarantees on optimality gap, to tackle the combinatorial design problems involved. Here, we present reformulations of the topology design task that allow us to solve the problem to optimality.

Our contributions are as follows. First, we provide a comprehensive modeling and analysis framework for the topology design problem to optimize a $\mathcal{H}_2$ norm based performance metric subject to budget constraints in Sections 4.2 and 4.3. Second, in Section 4.4 we show that although the topology design task is inherently non-convex, it is possible to exactly reformulate the problem in tractable form using McCormick relaxation (or linearization). This can then be used with off-the-shelf solvers to determine the optimal solution. Further, we show that exploiting graph-theoretic properties to tighten bounds on the continuous optimization variables yields significant improvements in computation time. Section 4.6 discusses numerical tests based on the IEEE 39−bus test case followed by conclusions and future directions in Section 4.7.

Notation: Column vectors (matrices) are denoted by lower- (upper-) case letters and sets by calligraphic symbols, unless noted otherwise. The cardinality of set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. Given a real-valued sequence $\{x_1, x_2, \ldots, x_N\}$, $x \in \mathbb{R}^N$ is the vector obtained by stacking the scalars x_i and $\mathrm{dg}(\{x_i\})$ is the corresponding diagonal matrix. The operator $(\cdot)^\top$ stands for transposition. The N-dimensional all ones vector is denoted by 1_N; I_N is the $N \times N$ identity matrix; and e_i is the canonical vector with a 1 at the i-th entry and zero everywhere else. The notation $\dot{\theta}_i$ denotes its time derivative $\frac{\delta \theta_i}{\delta t}$.

4.2 Dynamic Power System Model

A power network having $N+1$ buses can be modeled as a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, whose nodes $\mathcal{V} := \{0, 1, \ldots, N\}$ correspond to buses, and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ to undirected lines. Bus $i = 0$ is selected as the reference; all other buses constitute set $\mathcal{V}_r := \mathcal{V} \backslash \{0\}$. Let $b_{ij} > 0$ be the susceptance of line $(i, j) \in \mathcal{E}$ connecting nodes i and $j \in \mathcal{V}$. Then, the susceptance

Laplacian matrix $L \in \mathbb{R}^{(N+1)\times(N+1)}$ associated with the grid graph $\mathcal{G}$ can be defined as

$$L_{ij} := \begin{cases} -b_{ij} & , \text{ if } (i,j) \in \mathcal{E} \\ \sum_{(i,j)\in\mathcal{E}} b_{ij} & , \text{ if } j = i \\ 0 & , \text{ otherwise.} \end{cases}$$

Each node $i \in \mathcal{V}$ is associated with a phase angle θ_i, frequency $\omega_i = \dot{\theta}_i$, inertia constant M_i, and damping coefficient D_i; see [52]. Since bus i may host an ensemble of devices such as synchronous machines, renewable or energy storage sources, frequency-dependent or actively controlled frequency-responsive loads, the parameters M_i and D_i characterize their aggregate behavior [68].

Focusing on small-signal stability, the quantities (θ_i, ω_i) will henceforth refer to the deviations of nodal voltage angles and frequencies from their nominal values. The grid dynamics at bus i can then be described by the linearized swing equation [52]

$$M_i\dot{\omega}_i + D_i\omega_i = P_i^m - P_i^e + u_i \tag{4.1}$$

where P_i^m denotes the mechanical power input; P_i^e is the electric power flowing from bus i to the grid; and u_i is the power consumed at bus i. Again, the aforementioned quantities refer to the deviations from their scheduled values. Under the linearized DC model, the power flowing from bus i to the grid can be approximated as [52]

$$P_i^e \simeq \sum_{(i,j)\in\mathcal{E}} b_{ij}(\theta_i - \theta_j). \tag{4.2}$$

Combining (4.1) and (4.2), the state-space representation of the power grid is

$$\begin{bmatrix} \dot{\theta} \\ \dot{\omega} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & I_N \\ -M^{-1}L & -M^{-1}D \end{bmatrix}}_{A:=} \begin{bmatrix} \theta \\ \omega \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ M^{-1} \end{bmatrix}}_{B:=} u \tag{4.3}$$

where $M := \mathrm{dg}(\{M_i\})$ and $D := \mathrm{dg}(\{D_i\})$ are diagonal matrices containing the inertia and damping coefficients; the states $\omega \in \mathbb{R}^{N+1}$ and $\theta \in \mathbb{R}^{N+1}$ are accordingly the stacked vectors of nodal frequencies and angles; and $u \in \mathbb{R}^{N+1}$ is the vector of local power disturbances. The subsequent analysis relies on the ensuing assumption.

Assumption 4.1. *The inertia coefficients are strictly positive and damping coefficients are identical for all buses, that is $M_i > 0$ and $D_i = d$ for all $i \in \mathcal{V}$.*

The non-zero inertia assumption is not necessary, but simplifies our presentation. If $M_i = 0$ for a bus $i \in \mathcal{V}$, then a Kron-reduced network containing only nodes with inertia can be obtained. The second assumption of constant damping has been previously used in [77], [76]. When it is not satisfied, the stability metric defined in the next section does not enjoy a closed-form expression, and only bounds can be derived. In our future work, we plan to extend our approach to the case with variable D.

4.3 Generalized Network Coherence Metrics

Given the state-space model in (4.3), our goal is to design network topologies or augment existing ones to minimize the voltage angle deviations caused by load disturbances. These angle deviations are formally captured by the metric of *network coherence* [4]. The latter is defined as $\lim_{t\to\infty} \mathbb{E}[f_c(t)]$ where $f_c(t)$ is the steady-state deviation of the angles from their

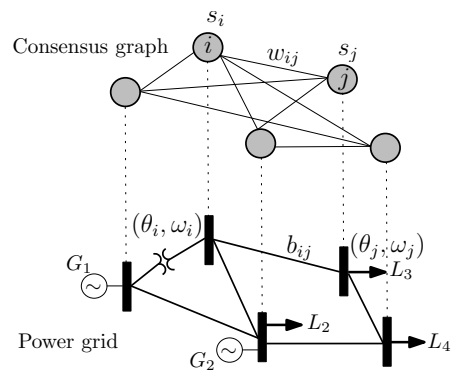

Figure 4.1: Illustration of power grid and its associated coherence graph.

grid-average

$$f_c(t) := \sum_{i=1}^{N} \left(\theta_i(t) - \frac{1}{N} \sum_{j=1}^{N} \theta_j(t) \right)^2 . \tag{4.4}$$

In other words, network coherence quantifies how tightly bus angles drift together. Larger variances in angle deviations reflect a more disordered network [77].

Instead of network coherence, the system operator may be interested in minimizing the expected steady-state value of a generalized function combining both voltage angle and frequency excursions [28], [68]

$$f(t) := \sum_{\forall i \neq j} w_{ij} (\theta_i(t) - \theta_j(t))^2 + \sum_{i \in \mathcal{V}} s_i \omega_i^2(t) \tag{4.5}$$

where w_{ij} and s_i are given non-negative scalars. The weights w_{ij} induce a connected weighted graph $\mathcal{G}_w$ that is not necessarily identical to $\mathcal{G}$; see Fig. 4.1. Let W be the Laplacian matrix of graph $\mathcal{G}_w$ and $S := \mathrm{dg}(\{s_i\})$. Then, it is not hard to verify that $f(t) = \|y(t)\|_2^2$ where

$$y(t) := \underbrace{\begin{bmatrix} W^{1/2} & 0 \\ 0 & S^{1/2} \end{bmatrix}}_{C:=} \begin{bmatrix} \theta(t) \\ \omega(t) \end{bmatrix} . \tag{4.6}$$

Being a Laplacian matrix, matrix W is positive semidefinite, and so its matrix square root $W^{1/2}$ is well-defined. The matrix $S^{1/2}$ is well-defined too. The importance of the generalized coherence metric of (4.5) is that for different choices of (W, S), one can represent different grid performance metrics and study them in a unified manner [28]: For instance, if the system operator is only interested in minimizing frequency excursions, one can set $W = 0$ and $S = I_{N+1}$. Similarly, if the goal is to reduce transient line losses, one can choose $W = L$ and $S = 0$. Lastly, the network coherence metric of (4.4) corresponds to the case of $W = I_{N+1} - \frac{1}{N+1} 1_{N+1} 1_{N+1}^\top$ and $S = 0$. Figure 4.1 shows the physical power network and its associated coherence graph. Note that the choice of (W, S) for network coherence penalizes global deviations, whereas that for line loss reduction penalizes local deviations.

4.3.1 Relation to Stability Analysis

The expected steady-state value of $f(t)$ can be interpreted as the squared $\mathcal{H}_2$-norm of the linear time-invariant (LTI) system described by (4.3) and (4.6). This system will be compactly denoted by $H := (A, B, C)$. Leveraging this link, the generalized network coherence is amenable to a closed-form expression [28].

The $\mathcal{H}_2$ norm is widely used as a stability performance metric and has several interpretations [18]: For unit-variance stochastic white noise $u(t)$, the $\mathcal{H}_2$-norm of an LTI system equals the steady-state output variance [18, Ch. 5], [77], [68]

$$\|H\|_{\mathcal{H}_2}^2 := \lim_{t \to \infty} \mathbb{E}\left[\|y(t)\|_2^2\right]. \tag{4.7}$$

For unit-impulse disturbances $u_i(t) = e_i \delta(t)$ for $i = 1, \ldots, N$, the $\mathcal{H}_2$-norm can be equiva-

lently written as

$$\|H\|_{\mathcal{H}_2}^2 := \sum_{i=1}^{N} \int_0^\infty \|y_i(t)\|_2^2 dt \tag{4.8}$$

where $y_i(t)$ is the system output corresponding to disturbance vector $u_i(t)$. Instead of evaluating the expectation in (4.7) or the time integral in (4.8), the $\mathcal{H}_2$-norm for system H can be expressed as [18, Ch. 5]

$$\|H\|_{\mathcal{H}_2}^2 = \mathrm{Tr}(B^\top Q B) \tag{4.9}$$

where $Q := \int_0^\infty e^{A^\top t} C^\top C e^{At} dt$ is the observability Gramian matrix of the LTI system H. In fact, the matrix Q can be computed as the solution to the linear Lyapunov equation [18, Ch. 5]

$$A^\top Q + QA = -C^\top C. \tag{4.10}$$

Matrix Q is known to be symmetric positive semidefinite, so it can be partitioned as

$$Q = \begin{bmatrix} Q_1 & Q_0 \\ Q_0^\top & Q_2 \end{bmatrix}. \tag{4.11}$$

Based on the reformulation in (4.9), we next design topologies attaining minimum generalized network coherence.

4.4 Grid Topology Design as an MILP

Among other criteria, the topology of a grid can be designed to minimize the generalized network coherence metric of (4.5). Given a graph $\hat{\mathcal{G}} = (\mathcal{V}, \hat{\mathcal{E}})$, where $\hat{\mathcal{E}}$ is the set of candidate lines weighted by their susceptances, the goal is to find the subset $\mathcal{E} \subseteq \hat{\mathcal{E}}$ of cardinality $|\mathcal{E}| \leq K$ with $K \geq N$ attaining the minimum generalized network coherence. Given the

equivalences of the previous section, this task can be posed as the optimization problem

$$\arg\min_{\mathcal{E}\in\hat{\mathcal{E}}} \quad \mathrm{Tr}(B^\top QB) \tag{4.12a}$$

$$\text{s.to} \quad |\mathcal{E}| \leq K \tag{4.12b}$$

$$Q \text{ satisfies (4.10)} \tag{4.12c}$$

$$\mathcal{E} \text{ is connected.} \tag{4.12d}$$

The constraint in (4.12b) reflects the budget on the number of edges. For $K = N$, the problem in (4.12) yields a tree topology, which is important for typically radially operated distribution grids. Interestingly, leveraging the problem structure under Assumption 4.1, it can be shown that the objective of (4.12) simplifies as [68], [28], [77]

$$\mathrm{Tr}(B^\top QB) = \frac{\mathrm{Tr}(WL^+) + \mathrm{Tr}(SM^{-1})}{2d}. \tag{4.13}$$

If Assumption 4.1 is not met, i.e., the damping coefficients are not identical ($D \neq dI_N$), then it may not be possible to find a closed-form expression for the objective in (4.12). If the damping coefficients are known to lie within $[d_{\min}, d_{\max}]$, then the objective of (4.12) can be bounded as; see [77], [68]

$$\frac{\mathrm{Tr}(WL^+) + \mathrm{Tr}(SM^{-1})}{2d_{\max}} \leq \mathrm{Tr}(B^\top QB) \leq \frac{\mathrm{Tr}(WL^+) + \mathrm{Tr}(SM^{-1})}{2d_{\min}}.$$

Therefore, as the range $[d_{\min}, d_{\max}]$ becomes narrower, minimizing the numerator of (4.13) approaches the optimal solution to (4.12). From Assumption 4.1, we consider $d_{\min} = d_{\max}$ here.

The second summand in the right-hand side of (4.13) is independent of the grid topology,

and can thus be ignored. Problem (4.12) can then be reformulated as

$$\begin{aligned} \arg\min_{\mathcal{E}\in\hat{\mathcal{E}}} \quad & \mathrm{Tr}(WL^{+}) \\ \text{s.to} \quad & |\mathcal{E}| \leq K \\ & \mathrm{rank}(L) = N \end{aligned} \tag{4.14}$$

where the rank constraint ensures that the graph induced by $\mathcal{E}$ is connected. Problem (4.14) could be tackled with brute-force algorithms over all the possible topologies of budget K and below; though that would incur exponential complexity.

The objective in (4.14) can be written in terms of the inverse of the *reduced Laplacian* matrix of $\mathcal{G}$ as explained next. The claim has appeared in [77, Lemma 3.2], albeit for the restricted case where W and L have the same structure.

Lemma 4.2. *If $\tilde{W}$ and $\tilde{L}$ are the $N \times N$ matrices obtained after removing the first row and column from W and L, respectively, then*

$$\mathrm{Tr}(WL^{+}) = \mathrm{Tr}(\tilde{W}\tilde{L}^{-1}).$$

Proof. The Laplacian matrices can be described in terms of their reduced counterparts as

$$W = R\tilde{W}R^{\top} \quad \text{and} \quad L = R\tilde{L}R^{\top} \tag{4.15}$$

where $R := [1_N \ \ -I_N]^{\top}$. This is because a Laplacian matrix satisfies $L1_{N+1} = 0_{N+1}$. If we define matrix $J := I_N - \frac{1}{N+1}1_N 1_N^{\top}$, then it is not hard to see that

$$R^{\top}R = I_N + 1_N 1_N^{\top} = J^{-1}. \tag{4.16}$$

Then the pseudoinverse of L can be expressed as

$$L^{+} = RJ\tilde{L}^{-1}JR^{\top}. \tag{4.17}$$

The latter can be shown by simply verifying that $LL^{+}L = L$ and $L^{+}LL^{+} = L^{+}$. From (4.15) and (4.17), we get that

$$\mathrm{Tr}(WL^{+}) = \mathrm{Tr}\left(R\tilde{W}R^{\top}RJ\tilde{L}^{-1}JR^{\top}\right) = \mathrm{Tr}(\tilde{W}\tilde{L}^{-1})$$

where the last equality follows from (4.16) and the cyclic property of the trace. □

To express the optimization in (4.14) over $\hat{\mathcal{E}}$ in a more convenient form, let us introduce a binary variable z_m for every line $m \in \hat{\mathcal{E}}$. This variable is $z_m = 1$ if line m is selected, i.e., $m \in \mathcal{E}$; and $z_m = 0$, otherwise. If we stack variables $\{z_m\}_{m\in\hat{\mathcal{E}}}$ in vector z, then z has to lie in the set

$$\mathcal{Z} := \left\{ z : z^{\top}1_{|\hat{\mathcal{E}}|} \leq K,\ z \in \{0,1\}^{|\hat{\mathcal{E}}|} \right\}. \tag{4.18}$$

Based on the line selection vector z, the reduced susceptance Laplacian of $\mathcal{G}$ can be expressed as

$$\tilde{L}(z) = \sum_{(i,j)\in\hat{\mathcal{E}}} z_{ij} b_{ij} a_{ij} a_{ij}^{\top} \tag{4.19}$$

where each vector a_{ij} corresponds to line $(i,j) \in \hat{\mathcal{E}}$, and its n-th entry is defined as

$$[a_{ij}]_n := \begin{cases} +1 & , \text{ if } n = i \\ -1 & , \text{ if } n = j \\ 0 & , \text{ otherwise.} \end{cases}$$

Given Lemma 4.2 and (4.19), the optimization in (4.14) can be equivalently written as the mixed-integer semidefinite program (MI-SDP)

$$\arg\min_{X,z\in\mathcal{Z}} \quad \mathrm{Tr}(\tilde{W}X) \tag{4.20a}$$

$$\text{s.to} \quad \begin{bmatrix} X & I_N \\ I_N & \tilde{L}(z) \end{bmatrix} \succeq 0. \tag{4.20b}$$

To see this, the constraint in (4.20b) is equivalent to $X \succeq 0$ and $X \succeq \tilde{L}^{-1}(z)$; see [17, Sec. A.5.5]. Since $\tilde{W} \succ 0$, the optimal X can be shown to be $X = \tilde{L}^{-1}(z)$. In fact, constraint (4.20b) waives the possibility of the optimal $\tilde{L}$ being singular, and thus, ensures the connectedness of the graph.

To overcome the computational complexity of the MI-SDP in (4.20), one may relax the binary variables to box constraints as $z \in [0,1]^{|\hat{\mathcal{E}}|}$ to get an ordinary SDP, which can be handled by off-the-shelf solvers for moderately-sized networks. Being a relaxation, the SDP solution provides a lower bound on the cost of (4.20). If the obtained solution of the SDP turns out to be binary, then this z minimizes the MI-SDP in (4.20) as well. Otherwise, (randomized) rounding heuristics can be adopted to acquire a feasible z.

Aiming at an exact solver, we will next show how the MI-SDP of (4.20) can be equivalently formulated as an MILP. To this end, we first rewrite (4.20) as

$$(X^*, z^*) \in \arg\min_{X,z\in\mathcal{Z}} \quad \mathrm{Tr}(\tilde{W}X) \tag{4.21a}$$

$$\text{s.to} \quad \tilde{L}(z)X = I_N. \tag{4.21b}$$

Note that for constraint (4.21b) to hold, $\tilde{L}(z^*)$ must be non-singular and $X^* = \tilde{L}^{-1}(z^*)$. Although its cost is linear, problem (4.21) is non-convex due to the bilinear constraints in

(4.21b) and because vector z is binary. To handle the former, we adopt the McCormick relaxation technique [62], which is briefly reviewed next.

Constraint (4.21b) involves the bilinear terms $z_m X_{ij}$ for $m \in \hat{\mathcal{E}}$ and $i, j \in \mathcal{V}_r$. For each such term, introduce a new variable

$$y_{mij} = z_m X_{ij}. \tag{4.22}$$

Suppose that the entries X_{ij}^* are known to lie within the range $[\underline{X}_{ij}, \overline{X}_{ij}]$. Since $z_m \in [0, 1]$, the ensuing inequalities hold trivially [62]

$$z_m(X_{ij} - \underline{X}_{ij}) \geq 0 \tag{4.23a}$$

$$(z_m - 1)(X_{ij} - \overline{X}_{ij}) \geq 0 \tag{4.23b}$$

$$z_m(X_{ij} - \overline{X}_{ij}) \leq 0 \tag{4.23c}$$

$$(z_m - 1)(X_{ij} - \underline{X}_{ij}) \leq 0. \tag{4.23d}$$

Substituting $z_m X_{ij}$ by y_{mij} in (4.23) provides

$$y_{mij} \geq z_m \underline{X}_{ij} \tag{4.24a}$$

$$y_{mij} \geq X_{ij} + z_m \overline{X}_{ij} - \overline{X}_{ij} \tag{4.24b}$$

$$y_{mij} \leq z_m \overline{X}_{ij} \tag{4.24c}$$

$$y_{mij} \leq X_{ij} + z_m \underline{X}_{ij} - \underline{X}_{ij}. \tag{4.24d}$$

One can replace the bilinear terms in (4.21b) by y_{mij}'s and enforce (4.22) and (4.24) as additional constraints for all $m \in \hat{\mathcal{E}}$ and $i, j \in \mathcal{V}_r$. In that case, the constraints (4.24) are apparently redundant. However, one may simplify the problem by dropping (4.22) to get an MILP reformulation of (4.21). Interestingly, this reformulation is *exact* due to the binary nature of z. To see this, if $z_m^* = 1$ for some m, then (4.24b) and (4.24d) imply $y_{mij}^* = X_{ij}^*$ for

all $i, j \in \mathcal{V}_r$. Otherwise, if $z_m^* = 0$, then (4.24a) and (4.24c) imply $y_{mij}^* = 0$ for all $i, j \in \mathcal{V}_r$.

Through the aforementioned process, problem (4.21) has been converted to an MILP over the variables $\{X_{ij}\}$, $\{z_m\}$, and $\{y_{mij}\}$. MILPs can be handled by state-of-the-art solvers such as Gurobi [46], and are known to scale better than MI-SDPs. Recall that the MILP reformulation of (4.21) requires the bounds $(\underline{X}_{ij}, \overline{X}_{ij})$ on each (i, j)-th entry of X^*. Lacking prior information on X^*, one could select an arbitrarily wide range $[\underline{X}_{ij}, \overline{X}_{ij}]$. However, this could slow down the MILP solver significantly. In the other extreme, if X^* is known, that is $\underline{X}_{ij} = \overline{X}_{ij}$ for all $i, j \in \mathcal{V}_r$, then the binary vector z can be recovered by simply solving the linear equations in (4.21b). To improve the run time of the involved MILP, we next derive tighter, non-trivial bounds on the entries of X_{ij}^*.

4.5 Graph-theoretic Bounds on the Optimization Variables

Depending on the problem structure, different bounds can be derived on X_{ij}^*s. This section considers two classes of topology design tasks. In the first task, some lines are already energized and the operator would like to augment a connected network by additional lines to further improve its stability. In the second task, a network topology is designed from scratch with the additional requirement of a radial grid.

4.5.1 Augmenting Existing Networks

Given the graph $\hat{\mathcal{G}} = (\mathcal{V}, \hat{\mathcal{E}})$, this problem setup considers a pre-existing connected network described by $\mathcal{G}_e = (\mathcal{V}, \mathcal{E}_e)$, and the goal is to energize additional lines from $\hat{\mathcal{E}} \setminus \mathcal{E}_e$ to improve its stability. In essence, this corresponds to the problem in (4.21) with the entries of z

corresponding to the lines in $\mathcal{E}_e$ being set to one. Then, based on (4.19), the reduced Laplacian matrix of the existing network is obviously

$$\tilde{L}_e := \sum_{(i,j)\in\mathcal{E}_e} b_{ij} a_{ij} a_{ij}^\top.$$

Under this setup, the entries of X^* minimizing (4.21) under the additional constraints $z_\ell = 1$ for all $\ell \in \mathcal{E}_e$, can be bounded as follows.

Lemma 4.3. *The entries of X^* are bounded by*

$$0 < X^*_{ij} \leq \frac{[\tilde{L}_e^{-1}]_{ii} + [\tilde{L}_e^{-1}]_{jj}}{2}, \quad \forall (i,j) \in \mathcal{V}_r. \tag{4.25}$$

Proof. Observe that $\tilde{L}_e$ can be written as $\tilde{L}(z)$ by fixing the entries of z corresponding to the lines in $\mathcal{E}_e$ to 1. Since the same entries will remain 1 in z^*, it readily follows that $\tilde{L}(z^*) \succeq \tilde{L}_e \succ 0$, where $\tilde{L}_e$ is non-singular since the existing network is already connected. Then, it follows that $X^* \preceq \tilde{L}_e^{-1}$ and $v^\top(X^* - \tilde{L}_e^{-1})v < 0$ for any $v \neq 0$. Setting $v = e_i$, the diagonal entries of X^* can be bounded as $X^*_{ii} \leq [\tilde{L}_e^{-1}]_{ii}$ for all $i \in \mathcal{V}_r$. Since $X^* \succ 0$ by constraint (4.21b), we have that $(e_i - e_j)^\top X^*(e_i - e_j) > 0$ for all $i, j \in \mathcal{V}_r$. The latter provides $X^*_{ij} < \frac{1}{2}(X^*_{ii} + X^*_{jj})$. Upper bounding the diagonal entries with the bounds obtained earlier proves the upper bound in (4.25). The lower bound in (4.25) can be trivially obtained since $L(z^*)$ is an M-matrix, and so its inverse has positive entries. □

The reduced Laplacian matrix $\tilde{L}_e$ of the existing network $\mathcal{G}_e$ is invertible as long as $\mathcal{G}_e$ is connected. If that is not the case, one could obtain bounds on X^*_{ij}'s by imposing a radial structure on the sought topology as discussed next.

4.5.2 Radial Topology Design

The setup considered here designs a network afresh, but under the requirement that it is radial. The analysis simplifies under the following mild assumption.

Assumption 4.4. *There exists a node in $\mathcal{V}$ that is incident to exactly one edge.*

To derive bounds on the entries of X^* minimizing (4.21) for the special case of $K = N$ (radial network), let us first construct the graph $\mathcal{G}_r = (\mathcal{V}, \hat{\mathcal{E}}_r)$, where $\hat{\mathcal{E}}_r$ consists of the edges in $\hat{\mathcal{E}}$, but with inverse weights $x_{ij} := b_{ij}^{-1}$ for all $(i,j) \in \hat{\mathcal{E}}$. Based on $\mathcal{G}_r$, we define some additional properties that will be useful later. If one of the nodes in $\mathcal{V}$ satisfying Assumption 4.4 is selected as the reference node, then the weight (inverse susceptance) of its incident line is denoted by x_0. Let us also define the minimum of the inverse line susceptances

$$x_{\min} := \min_{(i,j)\in\hat{\mathcal{E}}_r} x_{ij}.$$

Before solving (4.21), we find the maximum spanning tree of graph $\mathcal{G}_r$, and denote the sum of its edge weights by f. The maximum spanning tree can be found efficiently by finding the minimum spanning tree on $\mathcal{G}_r$ upon negating its edge weights [24]. Lastly, for each node $i \in \mathcal{V}_r$, we find its shortest path to the reference node 0 in $\mathcal{G}_r$. The sum of edge weights along this shortest path will be denoted by h_i.

Lemma 4.5. *Under Assumption 4.4, the entries of X^*, minimizing (4.21) for $K = N$, are bounded as*

$$h_i \leq X^*_{ii} \leq f, \quad \forall i \in \mathcal{V}_r \tag{4.26a}$$

$$x_0 \leq X^*_{ij} \leq f - x_{\min}, \quad \forall i,j \in \mathcal{V}_r. \tag{4.26b}$$

Proof. The bounds rely on a fundamental property of the inverse Laplacian matrix of a radial network: If $\tilde{L}$ is the reduced susceptance Laplacian of a radial network and $X := \tilde{L}^{-1}$, then the entry X_{ij} equals the sum of the inverse susceptances that are common to the paths of nodes i and j to the reference bus; see [26, Lemma 1]. Under Assumption 4.4, the common path of any pair of nodes $(i, j) \in \mathcal{V}_r$ must include at least the line incident to the reference bus, and hence, $X^*_{ij} \geq x_0$. By the definition of shortest path, the entry X^*_{ii} is lower bounded by h_i for all $i \in \mathcal{V}_r$, thus providing the lower bound in (4.26a).

Regarding the upper bound in (4.26a), recall that the (i, i)-th entry of X^* equals the sum of weights on the path from i to the reference node 0. The sum of weights on the longest such path is still upper bounded by the sum weight f of all edges in the maximum spanning tree. Because the entry X^*_{ij} for $i \neq j$ must have at least one less edge than the longest path, the upper bound in (4.26b) holds as well. □

Note that the upper bounds can be tight in the setting where the maximum spanning and optimum trees are the same line graph.

4.5.3 Model Simplification and Bound Tightening

Exploiting the structure of $\hat{\mathcal{G}} = (\mathcal{V}, \hat{\mathcal{E}})$ can provide additional information on the bounds of X^* matrix entries to accelerate solving (4.21) for the special case of $K = N$ (radial grid). For example, if $\hat{\mathcal{G}}$ gets disconnected upon removing edge $\ell \in \hat{\mathcal{E}}$, this edge ℓ belongs to the sought tree topology and $z^*_\ell = 1$ before solving (4.21). To identify such edges, we resort to the notion of a *graph cutset* $\mathcal{C} \subset \hat{\mathcal{E}}$, defined as a subset of edges that once removed, splits the graph $\hat{\mathcal{G}}$ into two or more connected components. The edges in this cutset will be termed as *critical edges*. Now, we present a simple algorithm to enumerate all the single critical edges ($|\mathcal{C}| = 1$) by initializing the graph $\hat{\mathcal{G}}$'s weights to unit values.

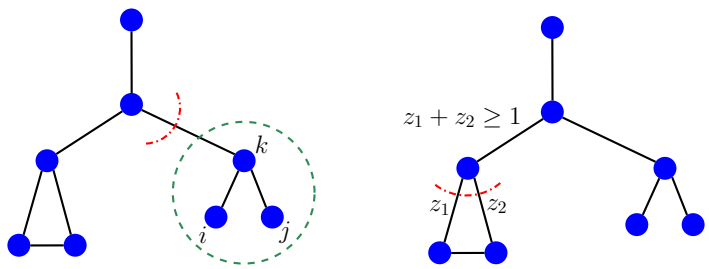

Figure 4.2: Grid graph $\hat{\mathcal{G}} = (\mathcal{V}, \hat{\mathcal{E}})$ with candidate location of edges. The lower bounds on X^*_{ki}, X^*_{kj}, and X^*_{ij} can be tightened using the shortest path weight h_k due to the critical edge colored in red (left panel). The size–2 cutset shown in red implies the constraint $z_1 + z_2 \geq 1$ (right panel).

1. Solve the min-cut problem on $\hat{\mathcal{G}}$ with unit edge weights by using the standard max-flow min-cut algorithm [35].

2. Increase the weight for the edge labelled as critical to $1 + \epsilon$ for some $\epsilon > 0$.

3. If $|\mathcal{C}| > 1$, quit; else, go to Step 1.

The second step ensures that every time we are identifying a new critical edge. Upon completing this process, the entries of z^* corresponding to the critical edges can be safely set to 1. This process not only reduces the binary search for z^* in (4.21), but it further tightens the lower bounds on certain X^*_{ij}'s as discussed in Lemma 4.6; see also Fig. 4.2.

Lemma 4.6. *Suppose a critical edge $\ell = (i, j) \in \hat{\mathcal{E}}$ partitions the nodes of $\hat{\mathcal{G}}$ into two disjoint connected components $\mathcal{V}_\ell$ and its complement $\bar{\mathcal{V}}_\ell$. If $\mathcal{V}_\ell$ contains node i as well as the reference bus, then (4.26b) can be tightened as*

$$h_j \leq X^*_{kj}, \quad \forall k \in \bar{\mathcal{V}}_\ell.$$

Proof. The edge $(i, j) \in \hat{\mathcal{E}}$ is the only edge that connects the nodes in $\bar{\mathcal{V}}_\ell$ to the rest of the network. Hence, the common path to the reference bus of any two nodes in $\bar{\mathcal{V}}_\ell$ must include

the path of node j to the reference. It follows that the shortest path weight h_j is a valid lower bound for all nodes in $\bar{\mathcal{V}}_\ell$. $\square$

Identifying cutsets of cardinality larger than 1 offers additional information to tighten the bounds of entries of the X matrix. If graph $\hat{\mathcal{G}}$ gets disconnected upon removing lines $\ell_1, \ell_2 \in \hat{\mathcal{E}}$, then at least one of these lines should be active. This logical conclusion translates to the constraint $z_{\ell_1} + z_{\ell_2} \geq 1$, which can be augmented to (4.21) to tighten the McCormick reformulation and possibly accelerate the MILP solver. Cutsets of larger cardinality, say $|\mathcal{C}| = k, k > 1$, can be identified by iterating Steps 1 through 3 of the algorithm described earlier. In this case, we assign the weights of $k + \epsilon$ on the critical edges, where $\epsilon > 0$.

4.6 Numerical Tests

All tests were run on a 2.7 GHz Intel Core i5 laptop with 8GB RAM. The MILP formulations were solved using Gurobi v8.0.1 optimizer, written in Julia/JuMP [29, 46].

The performance of the MILP in (4.21) was tested for augmenting an existing network as well as for designing a radial one afresh. For the augmentation setup, the IEEE 39-bus system benchmark was used as the pre-existing connected network [66]. The set $\hat{\mathcal{E}}$ was selected by 10 randomly picked additional lines. From these lines, we solved the restricted version of (4.21) for $K = \{2, 3, 4, 5\}$. To satisfy Assumption 4.1, we assumed $M_i = 10^{-4}$ on all buses that did not host generators, and $D_i = d = 0.025$ for all $i \in \mathcal{V}$. To evaluate an arbitrarily constructed network, we compared its squared $\mathcal{H}_2$ norm to that of the optimal network of the same edge cardinality. The results are summarized in Table 4.1. Additional lines are useful in minimizing disturbances, and so the budget constraint in (4.21) was always met with equality. For all cases, the augmentation design problem was solved in less than 5 seconds.

Table 4.1: Cost of designed grids using the IEEE 39-bus network.

# of lines	Optimal Topology	Suboptimal Topology
2	0.570	0.585
3	0.564	0.582
4	0.557	0.576
5	0.552	0.570

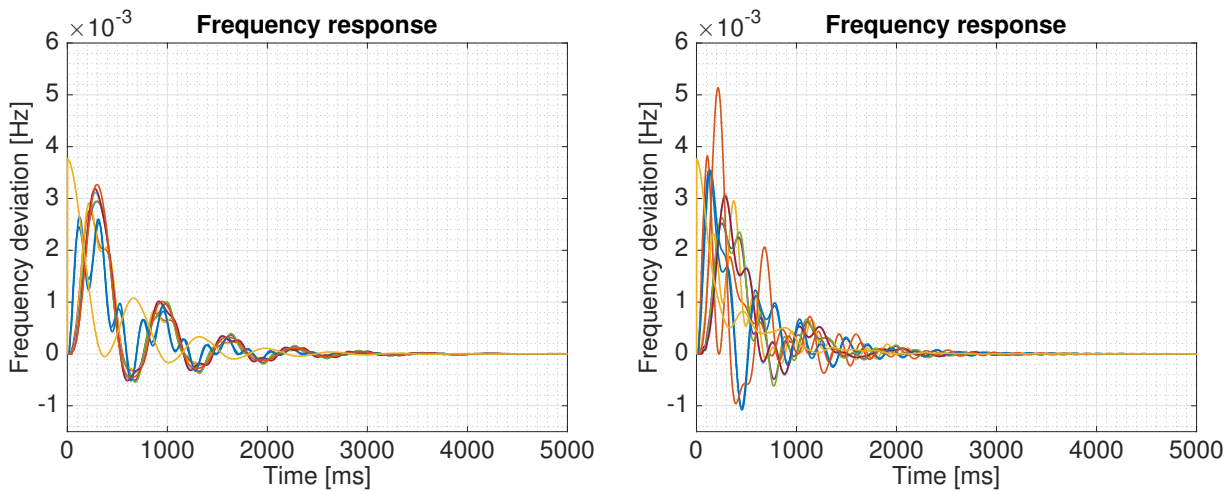

Figure 4.3: Frequency response at generator buses 30–39 for a impulse input on bus 39: (a) Optimal tree (left panel); (b) Sub-optimal tree (right panel).

We next considered the radial topology design problem with $\hat{\mathcal{E}}$ composed of all edges in the IEEE 39-bus network. The optimal cost obtained after solving the MILP in this case was 1.669, and the time required to find the optimal tree was close to 2 hours. Considering that the problem needs to be solved once off-line, this running time may not be of concern. Figure 4.4 shows the optimal radial topology that was identified.

Instead of using the bounds of Lemma 4.5 and the bound tightening procedure of Section 4.5.3, we attempted to solve the same MILP with the relatively looser bounds of $0 \leq X^*_{ij} \leq 10$ for all $(i,j) \in \mathcal{V}$. In this case, the solver reached the optimality gap of 60% after running for 3 hours. Clearly, having tighter bounds improves the computation time.

For a 1 per unit impulse input at bus 39, Figure 4.3 compares the frequency response of the

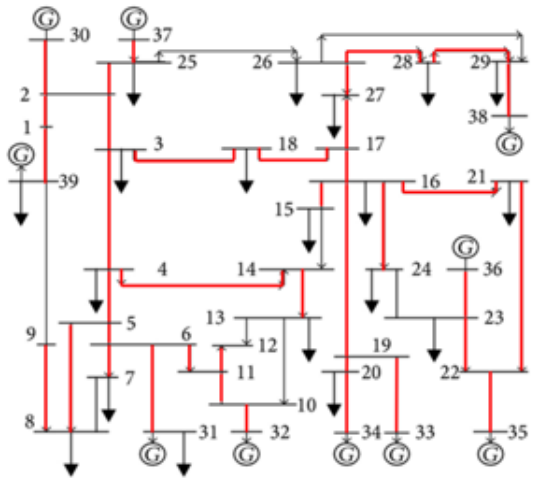

Figure 4.4: Optimal radial topology (bold red lines) considering all possible edges of the IEEE 39-bus feeder.

optimal tree (left panel) with an arbitrarily selected suboptimal tree (right panel). Notice that not only is the amplitude of oscillations lower in the left panel, but the generators also seem to drift together. These observations indicate that the optimal tree selected is much more effective at minimizing the effect of small disturbances.

4.7 Conclusions

We presented a general framework to study the effect of network topology on power grid dynamics. Using a system-theoretic approach, we showed that the original topology design problem (MI-SDP) can be reformulated in tractable form as an MILP. To improve the computation time, we exploited graph-theoretic properties to simplify the model and provide tighter bounds on the continuous optimization variables involved. Numerical tests on the IEEE 39-bus benchmark suggest that meshed networks exhibit improved coherence behavior.

Chapter 5

Summary and Future Work

We have presented a modeling and analysis framework for learning loads and optimizing power grid topologies. We hope that this body of work will aid utilities in improving efficiency, reliability, and stability of modern power networks. As a first step, we studied the novel data acquisition scheme of probing an electric grid via smart inverters to infer non-metered loads in Chapter 2. In particular, topological observability of grid probing using (non)-phasor data in single- and multi-phase grids was established. Using probing to infer non-constant-power loads was studied as a special case. Given a probing setup is deemed topologically observable, in Chapter 3 we presented a systematic methodology for designing probing injections. The goal was improved estimation accuracy and adherence to inverter and feeder constraints even without knowing non-metered loads. The computational tasks involved in grid probing were tackled using penalized SDP-based solvers that accounted for noisy measurements and non-stationary loads. Several extensions of this work remain open. Developing scalable solvers perhaps along the lines of [84]; incorporating measurement from distribution lines and transformers; and applying our topological observability framework to detect data attacks in distribution grids; all constitute pertinent research directions.

We have also developed a generalized framework to study the effect of network topology on power grid dynamics in Chapter 4 . Using a system-theoretic approach, we were able to show that the original topology design problem (MI-SDP) can be reformulated in tractable form as an MILP. To improve the computation time, graph-theoretic properties were exploited

to simplify the model and provide tighter bounds on the continuous optimization variables involved. Current research efforts are focused on tackling the topology design task with non-uniform damping, exploring conditions for submodularity, and considering the design problem from an $\mathcal{H}_\infty$-norm perspective.

Different from the aforementioned problems, we are also exploring data-driven algorithms for solving the linearized power grid dynamical equations without relying on numerical discretization and/or time integration. For large scale systems, this approach will aim to provide a closed-form solution to the differential equations by utilizing sparse measurement data and knowledge of the network model.

Bibliography

[1] A. Abur and A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*. New York, NY: Marcel Dekker, 2004.

[2] M. Angjelichinoski, A. Scaglione, P. Popovski, and Ĉ. Stefanović, "Decentralized DC microgrid monitoring and optimization via primary control perturbations," *IEEE Trans. Signal Processing*, vol. 66, no. 12, pp. 3280–3295, Jun. 2018.

[3] D. B. Arnold, M. D. Sankur, M. Negrete-Pincetic, and D. S. Callaway, "Model-free optimal coordination of distributed energy resources for provisioning transmission-level services," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 817–828, Jan. 2018.

[4] B. Bamieh, M. R. Jovanovic, P. Mitra, and S. Patterson, "Coherence in large-scale networks: Dimension-dependent limitations of local feedback," *IEEE Trans. Automat. Contr.*, vol. 57, no. 9, pp. 2235–2249, Sep. 2012.

[5] M. Baran and T. E. McDermott, "Distribution system state estimation using ami data," in *IEEE/PES Power Systems Conference and Exposition*, Seattle, WA, Mar. 2009.

[6] M. E. Baran, "Challenges in state estimation on distribution systems," in *Power Engineering Society Summer Meeting*, Vancouver, British Columbia, Jul. 2001.

[7] M. E. Baran, J. Zhu, and A. W. Kelley, "Meter placement for real-time monitoring of distribution feeders," *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 332–337, Feb. 1996.

[8] M. Baran and F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Syst.*, vol. 4, no. 1, pp. 735–743, Jan. 1989.

[9] S. Bhela, D. Deka, H. Nagarajan, and V. Kekatos, "Designing power grid topologies for minimizing network disturbances: An exact MILP formulation," in *Proc. IEEE American Control Conf.*, Philadelphia, PA, Jul. 2019.

[10] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Power distribution system observability with smart meter data," in *Proc. IEEE Global Conf. on Signal and Information Process.*, Montreal, Canada, Nov. 2017.

[11] ——, "Power grid probing for load learning: Identifiability over multiple time instances," in *Proc. IEEE Workshop on Comp. Adv. in Multi-Sensor Adaptive Proc.*, Curaçao, Dutch Antilles, Dec. 2017.

[12] ——, "Enhancing observability in distribution grids using smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5953–5961, Nov. 2018.

[13] ——, "Smart inverter grid probing for learning loads: Part I – Identifiability analysis," *IEEE Trans. Power Syst.*, Mar. 2019, (early access).

[14] ——, "Smart inverter grid probing for learning loads: Part II – Probing injection design," *IEEE Trans. Power Syst.*, Mar. 2019, (early access).

[15] S. Bhela, V. Kekatos, L. Zhang, and S. Veeramachaneni, "Enhancing observability in power distribution grids," in *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Process.*, New Orleans, LA, Mar. 2017.

[16] S. Bolognani and F. Dorfler, "Fast power system analysis via implicit linearization of the power flow manifold," in *Proc. Allerton Conf. on Comm., Control, and Computing*, Allerton, IL, Sep. 2015, pp. 402–409.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization.* New York, NY: Cambridge University Press, 2004.

[18] S. P. Boyd and C. Barratt, *Linear Controller Design: Limits of Performance*. Upper Saddle River, NJ: Prentice-Hall, 1991.

[19] G. Cavraro and V. Kekatos, "Graph algorithms for topology identification using power grid probing," *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 689–694, Oct. 2018.

[20] G. Cavraro and V. Kekatos, "Inverter probing for power distribution network topology processing," *IEEE Trans. Control of Network Systems*, pp. 1–1, 2019, (early access).

[21] G. Cavraro, V. Kekatos, and S. Veeramachaneni, "Voltage analytics for power distribution network topology verification," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 1058–1067, Jan. 2019.

[22] A. Cevallos, F. Eisenbrand, and R. Zenklusen, "Max-sum diversity via convex programming," in *Intl. Symp. on Comp. Geometry*, Dagstuhl, Germany, 2016, pp. 1–14.

[23] K. A. Clements, "The impact of pseudo-measurements on state estimator accuracy," in *Proc. IEEE PES General Meeting*, Detroit, MI, Jul. 2011.

[24] T. H. Cormen, *Introduction to algorithms*. MIT press, 2009.

[25] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 942–952, Mar. 2018.

[26] D. Deka, S. Backhaus, and M. Chertkov, "Learning topology of distribution grids using only terminal node measurements," in *Proc. IEEE Intl. Conf. on Smart Grid Commun.*, Sydney, NSW, Australia, Nov. 2016.

[27] D. Deka, M. Chertkov, and S. Backhaus, "Structure learning in power distribution networks," *IEEE Trans. Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, Sep. 2018.

[28] D. Deka, H. Nagarajan, and S. Backhaus, "Optimal topology design for disturbance minimization in power grids," in *Proc. IEEE American Control Conf.*, Seattle, WA, May 2017.

[29] I. Dunning, J. Huchette, and M. Lubin, "Jump: A modeling language for mathematical optimization," *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.

[30] B. C. Eaves and R. M. Freund, "Optimal scaling of balls and polyhedra," *Mathematical Programming*, vol. 23, no. 1, pp. 138–147, Dec. 1982.

[31] R. Ebrahimian and R. Baldick, "State estimator condition number analysis," *IEEE Trans. Power Syst.*, vol. 16, no. 2, pp. 273–279, May 2001.

[32] M. Fardad, F. Lin, and M. R. Jovanović, "Design of optimal sparse interconnection graphs for synchronization of oscillator networks," *IEEE Trans. Automat. Contr.*, vol. 59, no. 9, pp. 2457–2462, Jan. 2014.

[33] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *Proc. IEEE Conf. on Decision and Control*, Florence, Italy, Dec. 2013, pp. 4329–4334.

[34] FERC Order No. 755, "Frequency regulation compensation in the organized wholesale power markets," Oct. 2011.

[35] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network," *Canadian Journal of Mathematics*, vol. 8, pp. 399–404, 1956.

[36] G. Frigo, C. Narduzzi, D. Colangelo, M. Pignati, and M. Paolone, "Definition and assessment of reference values for PMU calibration in static and transient conditions," in *IEEE Intl. Workshop on Applied Measurements for Power Systems*, Aachen, Germany, Sep. 2016.

[37] L. Gan, N. Li, U. Topcu, and S. Low, "On the exactness of convex relaxation for optimal power flow in tree networks," in *Proc. IEEE Conf. on Decision and Control*, Maui, HI, Dec. 2012, pp. 465–471.

[38] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power systems synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1006–1013, May 2016.

[39] A. Ghosh, S. Boyd, and A. Saberi, "Minimizing effective resistance of a graph," *SIAM review*, vol. 50, no. 1, pp. 37–66, Feb. 2008.

[40] A. Gómez-Expósito, A. J. Conejo, and C. Canizares, Eds., *Electric Energy Systems, Analysis and Operation.* Boca Raton, FL: CRC Press, 2009.

[41] A. Gómez-Expósito, C. Gómez-Quiles, and I. Džafić, "State estimation in two time scales for smart distribution systems," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 421–430, Jan. 2015.

[42] G. R. Gray, J. Simmins, G. Rajappan, G. Ravikumar, and S. A. Khaparde, "Making distribution automation work: Smart data is imperative for growth," *IEEE Power Energy Mag.*, vol. 14, no. 1, pp. 58–67, Jan. 2016.

[43] J. W. Gu, K. A. Clements, G. R. Krumpholz, and P. W. Davis, "The solution of ill-conditioned power system state estimation problems via the method of Peters and Wilkinson," *IEEE Trans. Power App. Syst.*, vol. 102, no. 10, pp. 3473–3480, Oct. 1983.

[44] L. Guo, C. Zhao, and S. H. Low, "Graph Laplacian spectrum and primary frequency regulation," in *Proc. IEEE Conf. on Decision and Control*, Miami Beach, FL, Dec. 2018.

[45] Y. Guo, B. Zhang, W. Wu, Q. Guo, and H. Sun, "Solvability and solutions for bus-type

extended load flow," *Intl. Journal of Electrical Power & Energy Systems*, vol. 51, pp. 89–97, 2013.

[46] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2018. [Online]. Available: http://www.gurobi.com

[47] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 2013.

[48] M. Jaksic, Z. Shen, I. Cvetkovic, D. Boroyevich, R. Burgos, and P. Mattavelli, "Wide-bandwidth identification of small-signal dq impedances of ac power systems via single-phase series voltage injection," in *European Conf. on Power Electronics and App.*, Geneva, Switzerland, Sep. 2015.

[49] V. Kekatos, L. Zhang, G. B. Giannakis, and R. Baldick, "Voltage regulation algorithms for multiphase power distribution grids," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3913–3923, Sep. 2016.

[50] W. H. Kersting, *Distribution System Modeling and Analysis*. New York, NY: CRC Press, 2001.

[51] C. Klauber and H. Zhu, "Distribution system state estimation using semidefinite programming," in *Proc. North American Power Symposium*, Charlotte, NC, Oct. 2015.

[52] P. Kundur, *Power system stability and control*. New York, NY: McGraw-Hill, 1994.

[53] Y. Liao, Y. Weng, G. Liu, Z. Zhao, C. woo Tan, and R. Rajagopal, "Unbalanced three-phase distribution grid topology estimation and bus phase identification," 2018. [Online]. Available: https://arxiv.org/abs/1809.07192

[54] J. Lofberg, "A toolbox for modeling and optimization in MATLAB," in *Proc. of the CACSD Conf.*, 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip/

[55] R. Madani, A. Ashraphijuo, J. Lavaei, and R. Baldick, "Power system state estimation with a limited number of measurements," in *Proc. IEEE Conf. on Decision and Control*, Las Vegas, NV, Dec. 2016.

[56] R. Madani, J. Lavaei, and R. Baldick, "Convexification of power flow problem over arbitrary networks," in *Proc. IEEE Conf. on Decision and Control*, Osaka, Japan, Dec. 2015.

[57] E. Mallada, "iDroop: A dynamic droop controller to decouple power grid's steady-state and dynamic performance," in *Proc. IEEE Conf. on Decision and Control*, Las Vegas, NV, Dec. 2016.

[58] E. Mallada and A. Tang, "Improving damping of power networks: Power scheduling and impedance adaptation," in *Proc. IEEE Conf. on Decision and Control*, Orlando, FL, Dec. 2011.

[59] O. Mangasarian, "Set containment characterization," *Journal of Global Optimization*, vol. 24, no. 4, pp. 473–480, Dec. 2002.

[60] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1888–1896, Nov. 2012.

[61] J. Matevosyan, S. Sharma, S.-H. Huang, D. Woodfin, K. Ragsdale, S. Moorty, P. Wattles, and W. Li, "Proposed future ancillary services in Electric Reliability Council of Texas," in *PowerTech*, Eindhoven, Netherlands, Jun. 2015.

[62] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part I- convex underestimating problems," *Mathematical programming*, vol. 10, no. 1, pp. 147–175, 1976.

[63] S. P. Meyn, P. Barooah, A. Bušić, Y. Chen, and J. Ehren, "Ancillary service to the grid using intelligent deferrable loads," *IEEE Trans. Automat. Contr.*, vol. 60, no. 11, pp. 2847–2862, 2015.

[64] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[65] H. Nagarajan, P. R. Pagilla, S. Darbha, R. Bent, and P. P. Khargonekar, "Optimal configurations to minimize disturbance propagation in manufacturing networks," in *Proc. IEEE American Control Conf.*, Seattle, WA, May 2017.

[66] A. Pai, *Energy function analysis for power system stability.* New York, NY: Springer Science & Business Media, 2012.

[67] Pecan Street Inc., 2013. [Online]. Available: dataport.pecanstreet.org/

[68] B. K. Poolla, S. Bolognani, and F. Dörfler, "Optimal placement of virtual inertia in power grids," *IEEE Trans. Automat. Contr.*, vol. 62, no. 12, pp. 6209–6220, Dec. 2017.

[69] A. Primadianto and C. N. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.

[70] P. Raghavan and C. D. Tompson, "Randomized rounding: A technique for provably good algorithms and algorithmic proofs," *Combinatorica*, vol. 7, no. 4, pp. 365–374, Dec. 1987.

[71] A. Shafiu, N. Jenkins, and G. Strbac, "Measurement location for state estimation of distribution networks with generation," in *IEE Proceedings - Gen., Trans. and Dist.*, Mar. 2005, pp. 240–246.

[72] R. Singh, B. Pal, and R. Vinter, "Measurement placement in distribution system state estimation," in *Proc. IEEE PES General Meeting*, Calgary, AB, Jul. 2009.

[73] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization Methods Software*, vol. 11–12, pp. 625–653, Aug. 1999. [Online]. Available: http://sedumi.ie.lehigh.edu

[74] T. Summers, I. Shames, J. Lygeros, and F. Dörfler, "Topology design for optimal network coherence," in *Proc. IEEE European Control Conf.*, Linz, Austria, Jul. 2015.

[75] Y. Tang, K. Dvijotham, and S. Low, "Real-time optimal power flow," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2963–2973, May 2017.

[76] E. Tegling, M. Andreasson, J. W. Simpson-Porco, and H. Sandberg, "Improving performance of droop-controlled microgrids through distributed PI-control," in *Proc. IEEE American Control Conf.*, Boston, MA, Jul. 2016.

[77] E. Tegling, B. Bamieh, and D. F. Gayme, "The price of synchrony: Evaluating the resistive losses in synchronizing power networks," *IEEE Trans. Control of Network Systems*, vol. 2, no. 3, pp. 254–266, 2015.

[78] D. Trudnowski, M. Donnelly, and E. Lightner, "Power-system frequency and stability control using decentralized intelligent loads," in *IEEE/PES Transmission and Distribution Conference and Exhibition*, Dallas, TX, May 2006.

[79] D. Trudnowski and J. Pierre, "Signal processing methods for estimating small-signal dynamic properties from measured responses," in *Inter-area Oscillations in Power Systems*, ser. Power Electronics and Power Systems, A. R. Messina, Ed. Springer, 2009, pp. 1–36.

[80] A. B. Tucker, *Computer Science Handbook, Second Edition.* London, UK: Chapman & Hall/CRC, 2004.

[81] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proc. IEEE*, vol. 99, no. 6, pp. 1063–1073, Jun. 2011.

[82] W. T. Tutte, "The factorization of linear graphs," *Journal of the London Mathematical Society*, vol. 22, no. 2, pp. 107–111, 1947.

[83] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7290–7297, 2014.

[84] G. Wang, A. S. Zamzam, G. B. Giannakis, and N. D. Sidiropoulos, "Power system state estimation via feasible point pursuit: Algorithms and Cramer-Rao bound," *IEEE Trans. Signal Processing*, vol. 66, no. 6, pp. 1649–1658, Mar. 2018.

[85] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1008–1016, May 2013.

[86] L. Zhao, W. Zhang, H. Hao, and K. Kalsi, "A geometric approach to aggregate flexibility modeling of thermostatically controlled loads," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4721–4731, Nov. 2017.

[87] N. Zhou, D. J. Trudnowski, J. W. Pierre, and W. Mittelstadt, "Electromechanical mode online estimation using regularized robust RLS methods," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1670–1680, Nov. 2008.

[88] H. Zhu and G. B. Giannakis, "Power system nonlinear state estimation using distributed semidefinite programming," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1039–1050, Dec. 2014.