# Submodular Optimization for Power System Control and Stability

Zhipeng Liu

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

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Radha Poovendran, Chair

Linda Bushnell, Chair

Daniel Kirschen

Baosen Zhang

Program Authorized to Offer Degree:
Electrical and Computer Engineering

University of Washington

**Abstract**

Submodular Optimization for Power System Control and Stability

Zhipeng Liu

Co-Chairs of the Supervisory Committee:
Professor Radha Poovendran
Electrical and Computer Engineering

Professor Linda Bushnell
Electrical and Computer Engineering

Due to the increasing demand for electricity and unpredictable supplies from renewable energy, power systems are being operated close to their stability limits. Maintaining power system stability in the presence of disturbances is a challenging task over decades. Power system instability often arises from disturbances and the corrective controls are often combinatorial optimization problems which are NP-hard to solve.

Submodularity is a diminishing-return property of set functions which is analogous to the convexity of continuous functions. A combinatorial optimization problem that possesses submodularity can often be solved by greedy algorithms effectively with provable optimality guarantees. The concept of submodularity has been studied in a wide range of areas including sensor placement, feature selection, etc. The submodularity for power system stability problems, however, has not been exploited prior to this work.

The goal of this dissertation is to provide novel approaches towards addressing the power system control and stability challenges by exploiting the submodularity. This dissertation covers the voltage control problem, input and output selection problem in wide-area damping control for small signal stability, and controlled islanding problem in cascading failure.

# TABLE OF CONTENTS

# LIST OF FIGURES

Figure Number Page

# ACKNOWLEDGMENTS

My PhD study has been a precious experience to me thanks to many people. I would like to thank my advisors, Professor Radha Poovendran and Professor Linda Bushnell, for their guidance and mentoring over the years. I also thank them for providing me the opportunity to study in electrical engineering when I was a student without engineering background. This dissertation would not have been possible without their support. I thank Professor Daniel Kirschen, from whom I excelled my knowledge in power systems. I would also like to thank the rest of my PhD supervisory committee members, Professor Baosen Zhang and Professor Arvind Krishnamurthy for their time and helpful discussion.

I thank the former members of the Network Security Lab (NSL), Andrew Clark and Phillip Lee, who were my mentors, research collaborators, and also friends. I extend my thanks to all other current or former NSL members, including Xuhang (Shaun) Ying, Hossein Hosseini, Baicen Xiao, Sang Sagong, Joanna Mazer, Dinuka Sahabandu, Yize Chen, Shana Moothedath, Kalikinkar Mandal, Chou-Chang (Jack) Yang, Elisabeth Senmarti Robla, Truc Pham, Laila Abudahi, and Qiong Lin for their friendship.

At last, I send the deepest thanks to my parents and sisters for their unconditional love and support.

# DEDICATION

to my parents and my sisters

# Chapter 1

# INTRODUCTION

Power systems are large-scale interconnected cyber-physical systems. Due to the increasing demand for electricity and deregulation for new economic objectives, power systems are being operated close to their stability limits. Maintaining stable power generation and transmission in the presence of disturbances will be a significant engineering challenge in coming decades. The enhanced monitoring and communication technologies, as well as future smart grids, are envisioned to help ensure the stability of power system through real-time sensing and control, fundamentally shifting the paradigm of power system operation from open-loop to closed-loop control.

Power system instability often arises following disturbances, such as unexpected load shifting and transmission line tripping. Depending on both the type of the disturbance and its impact, the major stability problems have been classified into different forms including voltage stability and generator rotor angle stability, where the rotor angle stability is further classified into two categories: small signal stability and transient stability.

The available control actions to address these stability problems are often discrete, e.g., switching capacitor banks, cutting transmission lines, etc. As a result, many control design problems for power system stability are inherently combinatorial optimization problems, which are NP-hard to solve.

*Submodularity* is a diminishing-return property of set functions and is analogous to the convexity of continuous functions. A combinatorial problem that possesses submodularity, e.g., maximizing a submodular function subject to a cardinality/matroid constraint, can be approximately solved by computationally efficient algorithms, where solutions are proved to be within a constant factor the optimal solutions. Unfortunately, the submodularity does not always hold in many problems. In that case, one may exploit a weakened form of submodularity, namely *submodularity ratio*, which is sufficient to guarantee the optimality of a greedy algorithm with a relaxed approximation factor.

The concept of submodularity has been studied in a wide range of areas and has led to several effective greedy solutions to problems such as sensor placement, leader selection, feature selection, etc. The submodularity for power system stability problems, however, has not been exploited yet in existing literature. The goal of this dissertation is to provide novel approaches towards addressing the challenges in power system control and stability, by exploiting the submodularity/weak submodularity.

## *1.1 Power System Stability*

In this section, we briefly introduce the major stability problems and control challenges in power systems.

### *1.1.1 Voltage Stability*

*Voltage stability* refers to the ability of a power system to maintain acceptable voltages under normal conditions and after disturbances [36]. The inability to prevent system voltages from decreasing uncontrollably (voltage collapse) has resulted in extensive blackouts, including the 2003 Northeast US blackout, 1983 Sweden power outage, and 1987 Tokyo blackout [24]. Voltage instability occurs when a power system is unable to meet the reactive power demand (Fig. 1.1), and is typically corrected by injecting reactive power at some buses using capacitor/reactor bank switching or transformer tap changing [36]. In order to limit the complexity associated with coordinating actions taken over large networks by multiple system operators, these corrective actions are traditionally performed as local controls [79]. A purely local approach that does not take the interactions between neighboring buses into account, however, may be insufficient to meet reactive power demands when voltage deviations occur simultaneously at multiple buses.

Enhanced monitoring and communication capabilities could help improve voltage stability by enabling centralized or hierarchical controls. One such architecture was proposed in [71], and is currently being implemented in southern California. Under this approach, voltage magnitudes and

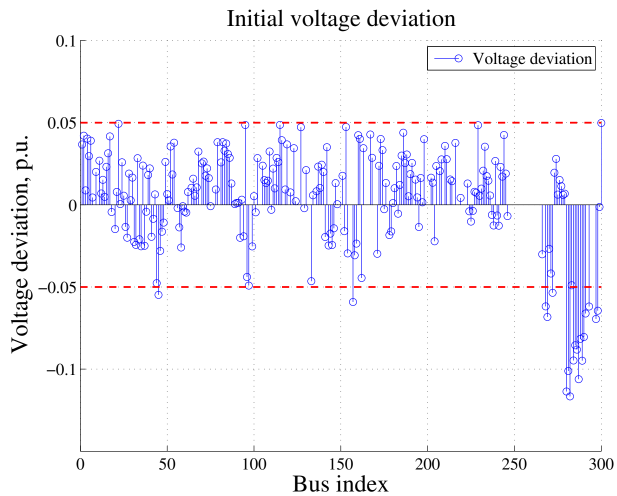



Figure 1.1: An example of unstable voltage profile. Voltages at some buses deviate from the acceptable region.

phase angles are measured using Phasor Measurement Units (PMUs). These measurements are used to evaluate the effectiveness of distributed voltage controls and, when necessary, devise a coordinated, centralized response.

A crucial step in a coordinated response is selecting a set of control actions based on the state of the system. If the set of actions at each bus is discrete (e.g., switch a capacitor bank on or off), designing a centralized control response to resolve multiple voltage deviations is a discrete subset selection problem. At present, such actions are heuristically chosen via enumeration and evaluation of combinations of control actions at multiple buses, after pruning the search space of all possible combinations using human expert domain knowledge [71]. Since the computation time of this approach is worst-case exponential in the number of buses, it does not scale to large power systems. Furthermore, it may result in a suboptimal response or fail to timely identify a set of control actions to prevent an impending voltage collapse. A computationally efficient optimization approach that exploits the underlying structure of the voltage regulation problem to select provably optimal control actions is thus needed to enable a timely and effective coordinated response to voltage instability.

### *1.1.2 Small Signal Stability*

Growing energy demand and integration of renewable energy sources have led to a steady increase in power transfers across wide geographical areas [83]. Such power transfers can create low-frequency phase oscillations between generators in different regions (Fig. 1.2), which, when poorly damped, cause the phase angles between generators to increase beyond the stability limit [36]. The ability of the power system to damp these oscillations before angle separation and outages occur is defined as the *small-signal stability* of the power system. Small-signal instability has resulted in generator tripping and catastrophic power outages, such as the 1996 Western United States blackout [72]. Ensuring small-signal stability has been identified as one of the critical power system stability challenges by the North American Synchrophasor Initiative (NASPI) [1].

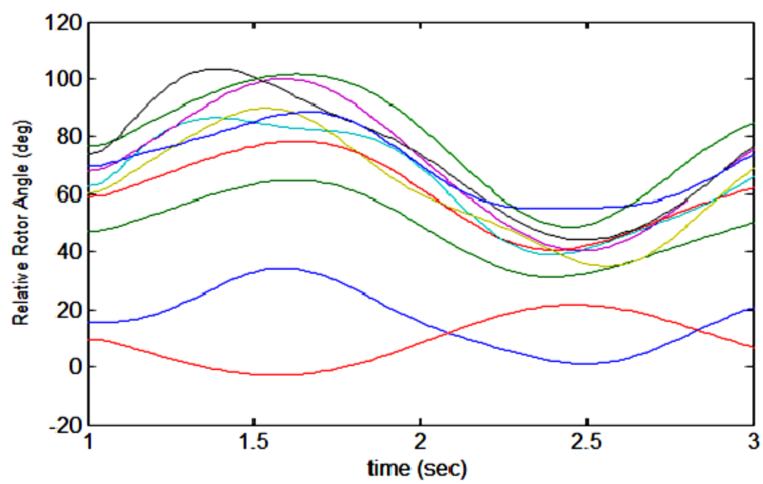



Figure 1.2: Generator rotor angle oscillations

The current approach to ensuring small-signal stability of the power system is through local control at generators using power system stabilizers (PSS) [36]. More recently, as increases in power transfers have made the stability of inter-area modes increasingly unpredictable, hierarchical, wide-area control mechanisms have been proposed in the power system literature [83, 82, 32]. Under the hierarchical approach, local PSSs are supplemented by a wide-area controller that receives state information from buses in multiple areas, evaluates the stability of the overall system, and sends control inputs to a subset of generators in order to ensure stability (Fig. 1.3). This control methodology is enabled by the availability of real-time state information from Phasor Measurement Units (PMUs) and the integration of reliable, low-latency communications [33].

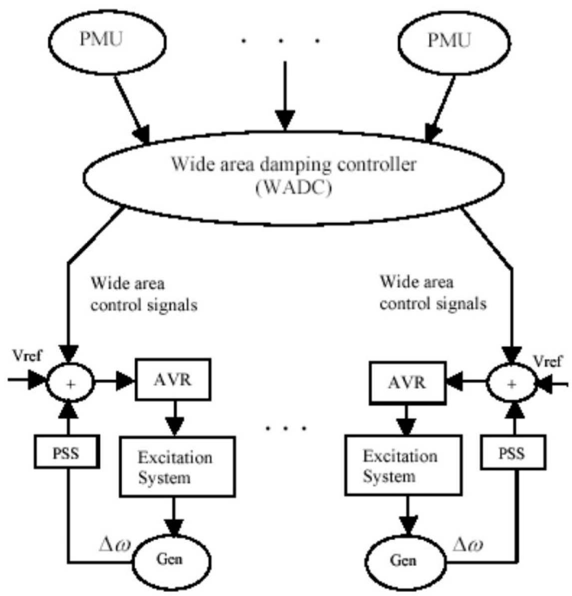



Figure 1.3: Wide-area damping control [83]

State of the art design techniques for wide-area control are based on linearizing the power system dynamics and identifying inter-area oscillations as unstable modes of the linear dynamics [83, 82, 32]. The goal of the wide-area control is to drive the system state to the equilibrium. A key step is selecting a subset of inputs (PSSs) and outputs (PMUs) to participate in the wide-area control design [83, 82]. Selecting such a subset, however, is inherently a discrete (combinatorial) optimization problem, and hence is computationally intractable for large-scale power systems unless additional problem structure can be identified.

Current approaches for selecting inputs and outputs for wide-area damping control use heuristic techniques that do not guarantee stability, and require the additional post-processing step of synthesizing a controller using the selected inputs/outputs, evaluating the stability properties, and re-selecting the inputs/outputs if stability is not satisfied.

On the other hand, most existing approaches to the minimal input/output selection problem assume that no uncertainty or time delay is present in the system, and consider one or more of the following system models: linear systems with known, fixed parameters [50], systems with a discrete set of parameter values [10], or systems with arbitrary real-valued parameters [54]. At present, there has been little study of input or output selections in the presence of continuous system uncertainties, such as small perturbations in system matrices. It has been shown that inputs and outputs that are selected based on a nominal system model may fail to guarantee the system stability with these uncertainties present [15]. Current works addressing the minimal selection problem for systems with uncertainties are based on relaxing the inherently combinatorial problem to continuous convex optimization form. This approach does not guarantee the optimality of solution.

### *1.1.3 Transient Stability and Cascading Failure*

The ability of a power system to return to stable condition and maintain its synchronism following a relatively large disturbance is referred to as the *transient stability*. When a large disturbance such as a transmission line outage occurs in one geographic area, its nearby lines have to compensate for the failed component by carrying more power flows [36]. This shift of power flows may cause other transmission lines to exceed their capacities, which leads to further outages. This process is referred to as *cascading failure*, in which local failures propagate and trigger successive failures throughout the entire power system (Fig. 1.4) [17]. Cascading failures can cause significant economic damage. For instance, in the 2003 North American blackout, two transmission line outages in the Ohio state created a cascading failure that left 55 million people out of power [4].

*Controlled islanding* is an effective corrective control to mitigate the impending cascading failure [2, 77]. Controlled islanding partitions the unstable power system into a set of smaller, disjoint subsystems, or islands, by tripping a selected set of transmission lines. Each of these islands is an internally stable and self-contained system, which can be reconnected to restore the entire system when the failure is cleared.

It is inevitable to shed a certain amount of load in some islands in order to match the limited

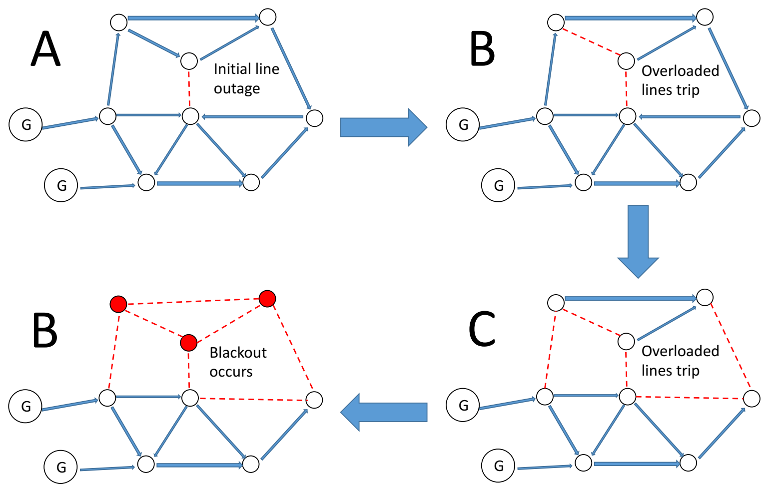



Figure 1.4: Cascading failure in power systems

generation within the island. A large amount of power imbalance between load and generation within an island may cause the power frequency deviating from the acceptable region and hence lead to an eventual collapse of the island [77]. Thus, it is important to minimize the overall load-generation imbalance in each island. On the other hand, generators tend to oscillate at different frequencies during cascading failure. If the generators in an island do not oscillate at similar frequencies, the island may become unstable as well [9]. Thus, it is also critical to limit the generator frequencies within each island when choosing an islanding strategy.

Algorithms that select islanding (partition) strategies based on these considerations have been proposed in [77, 16], which address the problem in two steps. These methods first identify clusters of generators with similar rotor angle frequencies, namely coherent generator groups. After forming such groups, a set of transmission lines is selected to trip, such that the coherent groups of generators are separated into different islands while the load-generation imbalance within each island is minimized. Current two-step approaches to controlled islanding rely on graph-based heuristics such as min-cut and spectral clustering. Such approaches are unable to guarantee that the resulting load-generation imbalance is within a constant factor of the optimal solution. On the other hand, these methods lack the flexibility in making trade-off between generator coherency and load-generation imbalance, defined as the ability to sacrifice a small amount of generator coherency within each island in order to gain a large reduction in overall power imbalance.

### *1.2 Contributions of this Thesis*

- *For voltage stability:* This thesis proposed a submodular optimization approach for designing a control strategy that prevents voltage instability. Our key insight is that the voltage deviation from the desired level is a supermodular function of the set of reactive power injections that are employed, leading to computationally efficient control algorithms for stabilization with provable optimality guarantees. This submodular control framework is tested on the IEEE 300-bus transmission system with comparison to the state-of-the-art hierarchical method. We found that our proposed submodular voltage control approach achieves voltage profiles with less voltage deviation in the same computation time.

- *For small signal stability:* This thesis studies the problem of identifying the minimum-size sets of input and output nodes in wide-area damping control to guarantee small signal stability. Our approach is generalized to any networked systems with uncertainties and time delays. We derive sufficient conditions to guarantee existence of a stabilizing controller for an uncertain linear system, based on a subset of system modes lying within the controllability and observability subspaces induced by the selected inputs and outputs. We then formulate the problems of selecting minimum-size sets of input and output nodes to satisfy the derived conditions, and prove that they are equivalent to discrete optimization problems with bounded submodularity ratios. We develop polynomial-time selection algorithms with provable guarantees on the minimum number of inputs and outputs required. Our approach is applicable to various types of uncertainties, including additive uncertainty, multiplicative uncertainty, uncertain output delay, and structured uncertainty. In a numerical study, we test our approach for the wide-area damping control in power systems to ensure small signal stability. Our results are validated on the IEEE 39-bus test power system.

- *For transient stability:* This thesis proposed a novel approach to controlled islanding based on weak submodularity. Our formulation jointly captures the minimal generator non-coherency and minimal load-generation imbalance in one objective function. We relax the problem to a formulation with bounded submodularity ratio and a matroid constraint, and propose an approximation algorithm which achieves a provable optimality bound on minimal non-coherency and load-generation imbalance. The proposed framework is tested on IEEE 39-bus and 118-bus power systems with comparison to a state-of-the-art spectral clustering islanding method. We observed that the proposed islanding approach can significantly reduce the power imbalance in resulting islands, compared to the state of the art, with the same or slightly higher generator non-coherency.

## *1.3 Organization of this Thesis*

This thesis is organized as follows. Chapter 2 gives backgrounds on submodularity and matroids. Chapter 3 presents our submodular optimization approach for voltage control. Chapter 4 studies the input and output selection problem for small signal stability and proposes a weak submodularity based approach which is generalized to any networked systems with uncertainties. Chapter 5 presents a weak submodularity based controlled islanding scheme to mitigate cascading failure.

# Chapter 2

# SUBMODULARITY AND MATROID

In this chapter, we give backgrounds on the submodularity, weak submodularity and matroids.

## *2.1 Submodularity and Weak Submodularity*

Submodularity is a diminishing-returns property of set functions, wherein the incremental benefit of adding an element to a set $S$ decreases as more elements are added to $S$. Let $2^\Omega$ denote the power set of $\Omega$. A set function $f : 2^\Omega \to R$ is *submodular* if, for any sets $S \subseteq T \subseteq \Omega$ and any element $v \in \Omega \setminus T$, $f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T)$. An equivalent definition is stated as follows.

**Definition 1** ([21]). *Let $V$ be a finite set, and let $2^V$ denote the set of all subsets of $V$ (power set). A function $f : 2^V \to \mathbb{R}$ is* submodular *if, for any subsets $S \subseteq V$ and $T \subseteq V$,*

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T). \tag{2.1}$$

*In addition, $f$ is* supermodular *if $-f$ is submodular. $f$ is modular if the equality holds in (2.1). A nonnegative weighted sum of submodular functions is also submodular.*

Submodularity is analogous to concavity of continuous functions. The problem of maximizing a monotone submodular function subject to a cardinality constraint can be approximately solved by polynomial-time greedy algorithms, where solutions are proved to be within a factor of $(1 - 1/e)$ of the optimal solutions [21].

The submodularity ratio measures to what extent a function $f$ has submodular properties, defined as follows.

**Definition 2** (Submodular Ratio [13], Weak Submodularity). *Let $f : 2^\Omega \to \mathbb{R}$ be a non-negative set function. The submodularity ratio of $f$ with respect to a set $U \subseteq \Omega$ and a parameter $k \geq 1$ is*

$$\gamma_{U,k} = \min_{\substack{L,S: L \cap S = \emptyset \\ L \subseteq U, |S| \leq k}} \frac{\sum_{z \in S} (f(L \cup \{z\}) - f(L))}{f(L \cup S) - f(L)}$$

*$f$ is called weakly submodular at $U$ and $k$ if $\gamma_{U,k} > 0$.*

A function is submodular if and only if $\gamma_{U,k} \geq 1$ for all $U$ and $k$. Generalizing to the functions with $0 < \gamma_{U,k} < 1$ provides a notion of *weak submodularity* [34]. For weakly submodular functions, even though the function may not be submodular, it still provides approximation guarantees for greedy algorithms.

An important concept related to submodularity ratio is coefficient of determination, which is defined as follows.

**Definition 3** (Coefficient of determination [13]). *For a linear regression problem $\min_z \|Hz - v\|_2^2$, assuming the columns of $H \in \mathbb{R}^{n \times m}$ and $v \in \mathbb{R}^n$ are normalized to have norm 1, the coefficient of determination (or $R^2$ statistic) is*

$$R^2 = n\, \bar{v}^T \bar{C}^{-1} \bar{v}$$

*where $\bar{C} = H^T H / n$ and $\bar{v} = H^T v / n$.*

### *2.2 Matroids*

A matroid is defined as follows.

**Definition 4.** *Let $V$ denote a finite set, and let $\overline{\mathcal{I}}$ denote a collection of subsets of $V$. The tuple $\mathcal{M} = (V, \overline{\mathcal{I}})$ is a* matroid *if (i) $\emptyset \in \overline{\mathcal{I}}$, (ii) for any $B \in \overline{\mathcal{I}}$ and $A \subseteq B$, $A \in \overline{\mathcal{I}}$, and (iii) if $A, B \in \overline{\mathcal{I}}$ and $|A| < |B|$, then there exists $v \in B \setminus A$ such that $(A \cup \{v\}) \in \overline{\mathcal{I}}$.*

The collection $\overline{\mathcal{I}}$ is referred to as the set of *independent sets* of the matroid $\mathcal{M}$. A maximal independent set of a matroid is a *basis*, denoted by $\mathcal{I}$. It can be shown that all bases of a matroid have the same cardinality. An important property of matroid bases is shown below.

**Lemma 1.** *If $\mathcal{I}_1$ and $\mathcal{I}_2$ are two bases of a matroid $\mathcal{M}(V, \overline{\mathcal{I}})$, then there exists a bijection $\pi : \mathcal{I}_1 \rightarrow \mathcal{I}_2$ such that*

$$\mathcal{I}_1 - x + \pi(x) \in \overline{\mathcal{I}} \text{ for all } x \in \mathcal{I}_1$$

An important class of matroids, denoted *graphic matroids*, is defined as follows.

**Lemma 2.** *Let $G = (N, E)$ be a graph with node set $N$ and edge set $E$. Define $\overline{\mathcal{I}}$ to be a collection of subsets of $E$ such that $A \in \overline{\mathcal{I}}$ iff $A$ does not contain any cycle (i.e., $A$ is a forest). Then $\mathcal{M} = (E, \overline{\mathcal{I}})$ is a matroid.*

For connected graphs, the set of bases of a graph matroid is the set of all spanning trees.

Another sub-class of matroids, namely *partition matroid*, is defined as follows.

**Lemma 3.** *Let $V$ denote a finite set, and let $V_1, \ldots, V_m$ be a partition of $V$, i.e., a collection of sets such that $V_1 \cup \cdots \cup V_m = V$ and $V_i \cap V_j = \emptyset$ for $i \neq j$. Let $d_1, \ldots, d_m$ be a collection of nonnegative integers. Define a set $\mathcal{I}$ by $A \in \mathcal{I}$ iff $|A \cap V_i| \leq d_i$ for all $i = 1, \ldots, m$. Then $\mathcal{M} = (V, \mathcal{I})$ is a matroid.*

The following lemma gives a property of submodular functions over matroid bases.

**Lemma 4** ([39]). *Let $\mathcal{M} = (V, \mathcal{I})$ be a matroid, and let $f : 2^V \rightarrow \mathbb{R}$ be a submodular function defined on $V$. Define $S$ to be a matroid basis such that, for all $u$ and $v$ with $u \in S$, $v \notin S$, and $(S \setminus \{u\} \cup \{v\}) \in \mathcal{I}$), $(1 + \epsilon) f(S) \geq f(S \setminus \{u\} \cup \{v\})$. Then for any basis $C$ of $\mathcal{M}$,*

$$2(1 + \epsilon) f(S) \geq f(S \cup C) + f(S \cap C). \tag{2.2}$$

# Chapter 3

# SUBMODULAR OPTIMIZATION FOR VOLTAGE CONTROL

## *3.1 Introduction*

The study of voltage instability problems dates back to the 1980s. Dobson and Chiang [18] introduced a model of voltage collapse that identifies the point of bifurcation as the operating point leading to voltage collapse. Traditional approaches to mitigating voltage instability are mainly preventive actions based on contingency ranking [23].

In the early 1990s, a three-level hierarchical voltage control scheme was proposed and later implemented in several European countries [37, 12, 73, 57]. The idea is analogous to the hierarchical levels of frequency control. In these schemes, a "pilot bus" that represents the voltage profile of a given region is selected and all control actions aim at regulating the voltage at this bus. Ongoing research address the pilot bus selection problem with improvements mostly in selection theory and decentralization of the controller [3].

Selecting a pilot point or a control region that is loosely coupled to the rest of the system is difficult. Furthermore, in North America, most voltage control actions are achieved by switching in or out capacitor or reactor banks. An online voltage regulation approach was proposed in [78] based on capacitor switching and power generation resetting. This approach, however, is limited in practice because operators hesitate to switch capacitor banks frequently. A slow centralized online voltage control scheme was proposed in [7] and further improved in [62]. In [7], the controller utilizes SCADA measurements and state estimation to predict the incremental effects of a switching action. The goal is to resolve any voltage violation with minimum switching actions and circular reactive flows. In [62], the state estimator is replaced by an integer programming method for deciding switching actions directly from the SCADA data and hence the computational overhead is significantly reduced. These approaches, however, are based on the assumption that the problematic areas containing voltage violations are loosely coupled and can be considered as separate control regions. When buses with voltage violations are closely coupled, the computational burden of these proposed methods either increases exponentially or the optimality deteriorates because the size of the control area is reduced.

Recently, a hierarchical, centralized strategy that addresses the voltage regulation problem for large power systems was proposed in [71]. In this approach, reactive power control is exerted through switching of reactive power injection devices to regulate voltage while minimizing the operating costs. This approach requires an enumeration of all possible configurations of reactive devices. To keep this enumeration computationally tractable, the system is partitioned into small areas and the enumeration is done within each area. This partitioning does not take into account the interactions between the voltages in the various areas and thus makes the solutions suboptimal.

Distributed or decentralized approaches for voltage regulation based on control theory are presented in [79, 56]. Zhang et al. [79] identify a sufficient condition for regulating voltage in a distributed manner with limited communication between buses. Robbins et al. [56] derive a distributed control law that guarantees voltage stability by varying the reactive power injection at each bus using distributed energy resources (DER). These works assume that the reactive power injection at every bus can be dynamically adjusted and do not take into account the associated cost.

Other voltage control approaches based on wide-area measurements are described in [25, 26, 22]. The control schemes in [25] and [26] rely on a constrained optimization, which is solved using linear or stochastic programming depending on the model used. In [22], a real-time control scheme is proposed to assess the power system stress and adjust voltage control actions for different stressed situations. These approaches, however, mainly rely on the control of generator voltage setpoints and load shedding, which are not suited to the deregulated power systems of the North American power grid where operators do not have direct control over generators.

Submodular optimization techniques for control of networked systems have been proposed in [67, 11]. These techniques focus on optimizing performance parameters, such as robustness to noise, smooth convergence, and controllability of linear systems. Prior to our work, however, submodularity has not been explored in the context of power system stability and control.

This chapter presents our submodular optimization approach to voltage control in power systems. Our fundamental insight is that the metrics typically used to evaluate the effectiveness of a voltage control strategy, such as the deviation from the desired voltage and the capacitor/reactor switching cost, have an inherent submodular structure. We make the following specific contributions:

- We formulate the problem of selecting a set of reactive power injection devices to minimize the deviation of the voltages from their desired values and the total operating cost. We prove that this problem is equivalent to submodular maximization with a matroid basis constraint.

- We propose polynomial-time algorithms for computing the optimal voltage control strategy. The performance of a control strategy is measured by its total operating costs plus remaining voltage deviations after control. We prove that the control strategies returned by our algorithms have performance no worse than a factor $1/3$ of the optimal control strategy.

- We present both a non-adaptive algorithm, which assumes that switching each reactive device results in a fixed change in system voltages based on the initial system operating point, as well as an adaptive algorithm that updates the operating point after each selection.

- We evaluate our submodular approach through a numerical study on the IEEE 300-bus test system. We find that, within the same computational time, our approach results in voltage profiles with less deviation compared to a state-of-the-art algorithm that searches over a subset of the possible control actions based on voltage sensitivities to reactive power injections.

The rest of this chapter is organized as follows. Section 3.2 presents the power system model used for voltage control and provides background information on submodularity. Section 3.3 presents the problem formulation and the approximation algorithms based on submodular optimization. Section 3.4 presents the simulation results. Section 3.5 concludes this chapter.

### *3.2 Power System Model*

We consider a power system consisting of $n$ PQ buses, $m$ PV buses and a slack bus. We index PQ buses by $i = 1, \ldots, n$, PV buses by $i = n+1, \ldots, n+m$ and the slack bus by $i = 0$. At each bus $i$, the voltage magnitude $V_i$ and angle $\theta_i$ can be measured by PMUs and are available to the centralized controller. A set of capacitor and reactor banks, denoted as $\Omega$, is located at a subset of PQ buses. One PQ bus $i$ may have multiple capacitor/reactor banks, denoted as $\Omega_i$, available to switch in or out. Switching a capacitor/reactor bank $j \in \Omega_i$ located at bus $i$ increases the reactive power injection at bus $i$ by $\Delta Q_j$. The magnitude of the change is taken to be a function of pre-switching voltage at the bus where the capacitor/reactor bank is located, e.g., $|\Delta Q_j| = C_j V_i^2$, where $C_j$ is the capacitance of the capacitor/reactor bank $j$. If a subset of capacitor/reactor banks, $S_i \subseteq \Omega_i$, are switched at bus $i$, then the total change in reactive power injection at bus $i$ is the sum of changes due to individual switches, denoted as $(\Delta Q)_i = \sum_{j \in S_i} \Delta Q_j$.

The impact of reactive power injection on the system voltages is derived as follows using a linearization of the power flow equations. We denote the admittance of the transmission line between bus $i$ and $j$ as $y_{ij} = g_{ij} - jb_{ij}$ where $g_{ij}$ and $b_{ij}$ denote the conductance and susceptance, respectively. Note that $g_{ii} = -\sum_{j \neq i}^{n+m} g_{ij}$ and $b_{ii} = -\sum_{j \neq i}^{n+m} b_{ij}$. At each bus $i$, the active power injection $P_i$ and reactive power injection $Q_i$ satisfy the power flow equations

$$P_i = V_i \sum_{j=0}^{n+m} V_j \left[ g_{ij} \cos(\theta_i - \theta_j) + b_{ij} \sin(\theta_i - \theta_j) \right], \tag{3.1}$$

$$Q_i = V_i \sum_{j=0}^{n+m} V_j \left[ g_{ij} \sin(\theta_i - \theta_j) - b_{ij} \cos(\theta_i - \theta_j) \right]. \tag{3.2}$$

Define $P = [P_1, P_2, \ldots, P_{n+m}]^T$ as the vector of active power injections at all PQ and PV buses and $Q = [Q_1, Q_2, \ldots, Q_n]^T$ as the vector of reactive power injections at PQ buses. By linearizing these equations around the current operating point, we obtain the following relation that contains a Jacobian matrix, $J$, of power flow equations, $[\Delta P^T \ \Delta Q^T]^T = J[\Delta\theta^T \Delta V^T]^T$, where $\Delta P = [(\Delta P)_1, (\Delta P)_2, \ldots, (\Delta P)_{n+m}]^T$ is the vector of changes in active power injections at PQ and PV buses, while $\Delta Q = [(\Delta Q)_1, (\Delta Q)_2, \ldots, (\Delta Q)_n]^T$ is the vector of changes in reactive power injections at PQ buses. The vector $\Delta\theta = [\Delta\theta_1, \Delta\theta_2, \ldots, \Delta\theta_{n+m}]^T$ refers to changes in bus angles at PQ and PV buses, while $\Delta V = [\Delta V_1, \Delta V_2, \ldots, \Delta V_n]^T$ refers to changes in voltage magnitudes at PQ buses.

The Jacobian matrix $J$ is assumed to be invertible under normal operating condition. This assumption is consistent with existing literature [18, 58], where the operating point at which the Jacobain matrix $J$ becomes singular was identified as the point of voltage collapse since no change in power injection could stabilize the change in voltage magnitudes.

Inverting the matrix $J$ and using the fact that the changes in real power injections are zero ($\Delta P = 0$) implies that voltage changes at all PQ buses due to capacitor/reactor bank switches can be approximated by

$$\Delta V = D\Delta Q, \tag{3.3}$$

where $D$ is the $n \times n$ sub matrix of $J^{-1}$ describing the effect of changes in reactive power injections on changes in voltage magnitudes.

## 3.3 Voltage Control Framework

This section formulates the problem of minimizing the deviation from the desired voltage and the operating cost by selecting a set of reactive power injection devices. We first formulate the problem and prove that it satisfies submodularity. Based on the submodularity property, we present a polynomial-time voltage control algorithm with provable optimality guarantees. We then describe an adaptive selection procedure added to the submodular control algorithm that mitigates the error in state prediction due to linearization.

### 3.3.1 Problem formulation

A centralized voltage controller can monitor and control the voltages continuously for all PQ buses. In a well-maintained power system, the voltage magnitude typically varies within $5\%$ of the nominal value, e.g., $0.95 \sim 1.05$ p.u. (per unit).

Switching a capacitor/reactor bank incurs an operating cost. We denote $c_i$ as the operating cost of switching in a capacitor bank $i$ that is inactive or switching out a reactor bank $i$ that is active, and denote $b_i$ as the cost of switching out a capacitor bank $i$ that is active or switching in a reactor bank $i$ that is inactive. Let $O \subseteq \Omega$ be the set of capacitor banks that are currently active and reactor banks that are inactive, and $F \subseteq \Omega$ be the set of capacitor banks that are inactive and reactor banks that are active.

The centralized controller exerts control by selecting a set $O' \subseteq \Omega$ consisting of capacitors that should be active and reactors that should be inactive. This set $O'$ can be interpreted as a set of capacitor/reactor banks that are selected to inject reactive power to the buses where they are located. Based on the current status set $O$ and decision set $O'$, the actual control action is to switch a set of capacitor/reactor banks, denoted as $S \subseteq \Omega$, from on to off or from off to on with $S = (O' \setminus O) \cup (O \setminus O')$.

The goal of the centralized controller is to choose the injecting set $O'$ that minimizes the resulting voltage deviations at PQ buses and the operating cost. The cost of voltage deviation and operation can be captured by the metric $f(O')$, defined as

$$f(O') = \sum_{j \in O' \setminus O} c_j + \sum_{j \in O \setminus O'} b_j + \lambda \sum_{i=1}^{n} h(V_i + \Delta V_i - V_{ref,i}), \tag{3.4}$$

where $h$ is a penalty function on voltage deviation at PQ buses. The function $h$ is chosen to be convex as heavier penalty should be applied if the voltage deviates further from the reference. The first two terms in Equation (3.4) give the total operating cost of switching capacitor/reactor banks in $S$, while the third term is a sum of penalties on voltage deviations at each PQ bus after switching capacitor/reactor banks in $S$. $\lambda$ is a nonnegative trade-off factor between the operating cost and the penalty on voltage deviations.

By using the approximation of Eq. (3.3), the last term of Eq. (3.4) can be re-written as

$$\lambda \sum_{i=1}^{n} h\left(\sum_{j \in S} \omega_{ij}(\Delta Q_j) - V_{ref,i} + V_i\right),$$

where $\omega_{ij} = D_{ik}$ for $j \in \Omega_k$, i.e., the capacitor/reactor bank $j \in S$ located at bus $k$. (Here, $D_{ik}$ denotes the $i$th row and $k$th column entry of the matrix $D$ in Section 3.2.) The parameter, $\omega_{ij}$, can be interpreted as the approximated incremental increase in voltage at bus $i$ from a per unit reactive power injection by capacitor/reactor bank $j$.

By substituting $S$ with $(O' \setminus O) \cup (O \setminus O')$ and rearranging terms, we have

$$f(O') = \sum_{j \in O' \setminus O} c_j + \sum_{j \in O \setminus O'} b_j + \lambda \sum_{i=1}^{n} h\left(\sum_{j \in O'} \omega_{ij} |\Delta Q_j| - \bar{V}_i\right), \tag{3.5}$$

where $\bar{V}_i = V_{ref,i} - V_i + \sum_{j \in O} \omega_{ij} |\Delta Q_j|$. The problem of selecting an optimal set $O'$ of capacitor/reactor banks to inject reactive power can then be formulated as a discrete optimization problem

$$\min \{f(O') : O' \subseteq \Omega\}. \tag{3.6}$$

This is a combinatorial optimization problem, and hence cannot be efficiently solved or approximated unless the objective function possesses additional structure, such as submodularity. While the objective function $f(O')$ is not supermodular, an equivalent submodular objective function can still be derived, as shown in the following subsection.

### *3.3.2 Submodular optimization approach*

We first give a new problem formulation with a matroid constraint, followed by the proof of its equivalence to the previous formulation. We then demonstrate that this problem has the structure of matroid basis-constrained submodular maximization. The submodular formulation is established on an expanded ground set $\overline{\Omega}$, defined as

$$\overline{\Omega} = \{v_{jl} : j \in \Omega,\ l = 0, 1\},$$

where $v_{j0}$ denotes the event that the capacitor bank $j$ is inactive (or active if $j$ is reactor bank), while $v_{j1}$ is the event that the capacitor bank $j$ is active (or inactive for reactor bank).

For each PQ bus $i$, let $P_i = \{j \in \Omega : \omega_{ij} > 0\}$ and $R_i = \{j \in \Omega : \omega_{ij} < 0\}$, i.e., the set of capacitor/reactor banks that, by injecting reactive power at its located bus, causes an increase ($P_i$) or decrease ($R_i$) in voltage at bus $i$. For any set $A \subseteq \overline{\Omega}$, we define sets $A_0$ and $A_1$ by

$$A_0 = \{j \in \Omega : v_{j0} \in A\}, \quad A_1 = \{j \in \Omega : v_{j1} \in A\}.$$

The submodular formulation consists of an objective function $\hat{f} : 2^{\overline{\Omega}} \to \mathbb{R}$ that is supermodular and equivalent to Eq. (3.6), and a matroid constraint with the matroid bases collection denoted as $\mathcal{B}$. For each PQ bus $i$, we define a cost function $\hat{f}_i : 2^{\overline{\Omega}} \to \mathbb{R}$ by

$$\hat{f}_i(A) = h\left(\sum_{j \in P_i \cap A_1} \omega_{ij}|\Delta Q_j| + \sum_{j \in R_i \cap A_0} |\omega_{ij}||\Delta Q_j| - \hat{V}_i\right),$$

where $\hat{V}_i = \bar{V}_i - \sum_{j \in R_i} \omega_{ij}|\Delta Q|_j$. Considering $n$ PQ buses overall, then the system-wide cost function $\hat{f}$ is defined as

$$\hat{f}(A) = \sum_{j \in A_1 \cap F} c_j + \sum_{j \in A_0 \cap O} b_j + \lambda \sum_{i=1}^{n} \hat{f}_i(A). \tag{3.7}$$

The equivalence between $\hat{f}(A)$ and the objective function in the previous subsection is established by the following lemma.

**Lemma 5.** *Suppose that the set $A \subseteq \overline{\Omega}$ satisfies $|A \cap \{v_{j0}, v_{j1}\}| = 1$ for all $j \in \Omega$. Then $\hat{f}(A) = f(A_1)$.*

The proof follows from the fact that, under the assumption on $A$, $\Omega = A_0 \cup A_1$ is a partition of the set $\Omega$, and hence $R_i = (R_i \cap A_0) \cup (R_i \cap A_1)$ and $R_i \cap A_1 = R_i \setminus (R_i \cap A_0)$.

*Proof.* The proof contains two parts. First, the condition $|A \cap \{v_{j0}, v_{j1}\}| = 1$ for all $j \in \Omega$ implies $\Omega = A_0 \cup A_1$ and hence we have

$$\sum_{j \in A_1 \setminus O} c_j + \sum_{j \in O \setminus A_1} b_j = \sum_{j \in A_1 \setminus O} c_j + \sum_{j \in A_0 \cap O} b_j.$$

Then it suffices to show that, for each PQ bus $i$,

$$\hat{f}_i(A) = h\left(\sum_{j \in A_1} \omega_{ij}|\Delta Q_j| - \bar{V}_i\right)$$

when the condition $|A \cap \{v_{j0}, v_{j1}\}| = 1$ holds for all $j \in \Omega$. We have

$$\begin{aligned}\sum_{j \in A_1} \omega_{ij}|\Delta Q_j| &= \sum_{j \in A_1 \cap P_i} \omega_{ij}|\Delta Q_j| + \sum_{j \in A_1 \cap R_i} \omega_{ij}|\Delta Q_j| \\ &= \sum_{j \in A_1 \cap P_i} \omega_{ij}|\Delta Q_j| + \sum_{j \in R_i} \omega_{ij}|\Delta Q_j| - \sum_{j \in R_i \cap A_0} \omega_{ij}|\Delta Q_j|, \end{aligned} \tag{3.8}$$

where Eq. (3.8) follows from the fact that, under the assumption on $A$, $\Omega = A_0 \cup A_1$ is a partition of the set $\Omega$, and hence $R_i = (R_i \cap A_0) \cup (R_i \cap A_1)$ and $R_i \cap A_1 = R_i \setminus (R_i \cap A_0)$. Then

$$\begin{aligned} h\left(\sum_{j\in A_1} \omega_{ij}|\Delta Q_j| - \bar{V}_i\right) &= h\left(\sum_{j\in A_1\cap P_i} \omega_{ij}|\Delta Q_j| + \sum_{j\in A_0\cap R_i} |\omega_{ij}||\Delta Q_j| - \left[\bar{V}_i - \sum_{j\in R_i}\omega_{ij}|\Delta Q_j|\right]\right) \\ &= \hat{f}_i(A), \end{aligned}$$

as desired. □

By Lemma 5, the problem (3.6) of selecting a set of buses to inject reactive power is equivalent to minimizing $\hat{f}(A)$ subject to the constraint that $|A \cap \{v_{j0}, v_{j1}\}| = 1$ for all $j \in \Omega$. We define a collection $\mathcal{B}$ of subsets of $\overline{\Omega}$ by $A \in \mathcal{B}$ iff $|A \cap \{v_{j0}, v_{j1}\}| = 1$ for all $j \in \Omega$. Then the equivalent discrete optimization problem can be formulated as

$$\min\{\hat{f}(A) : A \in \mathcal{B}\}. \tag{3.9}$$

Next, we show $\mathcal{B}$ is the basis of a matroid and hence the condition that $A \in \mathcal{B}$ in Eq. (3.9) is a matroid basis constraint on the set $A \subseteq \overline{\Omega}$. The intuition comes from the fact that each device can be either on or off, which corresponds to $|A \cap \{v_{j0}, v_{j1}\}| = 1$ for all $j \in \Omega$, i.e., $A \in \mathcal{B}$. As a result, for any optimal selection $A$, the resulting $A_0$ and $A_1$ should always form a partition of $\Omega$.

We define $\overline{\Omega}_j = \{v_{j0}, v_{j1}\}$ for $j \in \Omega$. Then the sets $\overline{\Omega}_{j\in\Omega}$ form a partition of $\overline{\Omega}$, i.e., $\overline{\Omega} = \cup_{j\in\Omega}\overline{\Omega}_j$ and $\overline{\Omega}_i \cap \overline{\Omega}_j = \emptyset$ for $i \neq j$. By Lemma 3, the constraint $A \in \mathcal{I}$ iff $|A \cap \overline{\Omega}_j| \leq 1$ for all $j \in \Omega$ defines a partition matroid $\mathcal{M} = (\overline{\Omega}, \mathcal{I})$ . The bases of $\mathcal{M}$ are the sets $A$ satisfying $|A \cap \overline{\Omega}_j| = 1$ for all $j \in \Omega$, which are exactly the sets in $\mathcal{B}$.

It remains to show $\hat{f}(A)$ is supermodular, which is established by the following theorem.

**Theorem 1.** *The function $\hat{f}(A)$ is supermodular in $A \subseteq \overline{\Omega}$.*

*Proof.* The approach of the proof is to show a more general result, namely, that any function $g(A)$ defined by $g(A) = h\left(\sum_{j\in A}\alpha_j - \beta\right)$, where $\alpha_j \geq 0$ and $\beta \in \mathbb{R}$, is supermodular as a function of $A$. Define $\rho = \frac{\sum_{j\in A\setminus B}\alpha_j}{\sum_{j\in A\Delta B}\alpha_j}$, where $\Delta$ is the symmetric difference operator. Since all $\alpha_j$'s are nonnegative, $\rho \in [0, 1]$. We have

$$\sum_{j \in A} \alpha_j = \rho \sum_{j \in A \cup B} \alpha_j + (1-\rho) \sum_{j \in A \cap B} \alpha_j,$$
$$\sum_{j \in B} \alpha_j = (1-\rho) \sum_{j \in A \cup B} \alpha_j + \rho \sum_{j \in A \cap B} \alpha_j.$$

By substitution, we have

$$\begin{aligned}
g(A) + g(B) =& h\left(\sum_{j \in A} \alpha_j - \beta\right) + h\left(\sum_{j \in B} \alpha_j\right) \\
=& h\left(\rho\left(\sum_{j \in A \cup B} \alpha_j - \beta\right) + (1-\rho)\left(\sum_{j \in A \cap B} \alpha_j - \beta\right)\right) \\
&+ h\left((1-\rho)\left(\sum_{j \in A \cup B} \alpha_j - \beta\right) + \rho\left(\sum_{j \in A \cap B} \alpha_j - \beta\right)\right) \\
\leq& \rho h\left(\sum_{j \in A \cup B} \alpha_j - \beta\right) + (1-\rho) h\left(\sum_{j \in A \cap B} \alpha_j - \beta\right) \\
&+ (1-\rho) h\left(\sum_{j \in A \cup B} \alpha_j - \beta\right) + \rho h\left(\sum_{j \in A \cap B} \alpha_j - \beta\right) \\
=& g(A \cup B) + g(A \cap B),
\end{aligned}$$

which establishes the supermodularity of $g$.

Since the function $\hat{f}_i(A)$ can be obtained from $g(A)$ by setting $\alpha_j = |\omega_{ij}||\Delta Q_j|$ and $\beta = \hat{V}_i$, each $\hat{f}_i(A)$ is supermodular. The terms $\sum_{i \in A_1 \setminus O} c_i + \sum_{i \in A_0 \cap O} b_i$ can be shown to be modular. The function $\hat{f}(A)$ is thus a sum of supermodular and modular functions, and hence is supermodular. $\square$

Combining Theorem 1 with the above discussion of $\mathcal{B}$, we have that Eq. (3.9) is a supermodular minimization problem, which can be transformed to an equivalent submodular maximization problem, subject to a matroid basis constraint. The submodularity of this problem allows constructing approximation algorithms with provable optimality bound, as presented in the following subsection.

### *3.3.3 Voltage control algorithms*

We now present a voltage control algorithm that is derived from techniques for maximizing a submodular function subject to a matroid basis constraint.

As the constraint $A \in \mathcal{B}$ in Eq. (3.9) implies that each capacitor/reactor bank $j \in \Omega$ can be either active $v_{1j}$ or inactive $v_{0j}$, the algorithm proceeds as follows. We initialize a set $A \subseteq \overline{\Omega}$ by the current configuration of capacitor/reactor banks such that $A = \{v_{j1} : j \in O\} \cup \{v_{j0} : j \in F\}$. For a chosen constant parameter $\epsilon > 0$, the algorithm iteratively selects a capacitor/reactor bank $j$ from $A_0$ or $A_1$ such that

$$\hat{f}(A \setminus \{v_{j0}\} \cup \{v_{j1}\}) < (1-\epsilon)\hat{f}(A) \text{ if } v_{j0} \in A \text{ (i.e., } j \in A_0),$$

or

$$\hat{f}(A \setminus \{v_{j1}\} \cup \{v_{j0}\}) < (1-\epsilon)\hat{f}(A) \text{ if } v_{j1} \in A \text{ (i.e., } j \in A_1).$$

The algorithm terminates when no such device $j$ can be found. A pseudocode description is given as Algorithm 1.

Intuitively, at each iteration, the algorithm identifies a capacitor/reactor bank $j \in \Omega$ such that switching it will reduce the system-wide cost $\hat{f}(A)$.

The algorithm converges to a $\epsilon$-local minimum, defined as a set $A \subseteq \overline{\Omega}$ that satisfies $(1-\epsilon)\hat{f}(A) < \hat{f}(A \cup \{v\} \setminus \{w\})$ for any $v, w \in \overline{\Omega}$ such that $(A \cup \{v\} \setminus \{w\}) \in \mathcal{B}$. The local minimum is known to be within a factor of $1/6$ for the nonmonotone submodular maximization problem with matroid basis constraint [39]. This optimality bound can be improved by $1/3$ by exploiting the partition matroid structure of (3.9), as shown in the following theorem.

**Theorem 2.** *Let $M$ satisfy $\hat{f}(A) \leq M$ for all $A \subseteq \overline{\Omega}$. Define $O'$ to be the set chosen by Algorithm 1, and let $O^*$ be the optimal solution to* $\min\{f(O') : O' \subseteq \Omega\}$. *Then*

$$M - f(O') \geq \left(\frac{1}{3+\epsilon}\right)(M - f(O^*)). \tag{3.10}$$

**Algorithm 1:** Algorithm for selecting a set of capacitor/reactor banks to inject reactive power.

**procedure** SUBMODULARVC($\mathbf{v}$, $\mathbf{v_{ref}}$, $\boldsymbol{\omega}$, $\mathbf{q}$, $\mathbf{b}$, $\mathbf{c}$, $O$, $F$)

**Input**: Initial voltages at each bus $\mathbf{v}$, reference voltages for PQ buses $\mathbf{v_{ref}}$, weights $\omega_{ij}$ from inverse Jacobian $J^{-1}$, possible reactive power injections by each capacitor/reactor bank $\mathbf{q} = \{|\Delta Q_j|, j \in \Omega\}$, switching costs $\mathbf{b}$ and $\mathbf{c}$, set of active capacitor banks and inactive reactor banks $O$ that initially injecting reactive power, $F = \Omega \setminus O$.

**Output:** Set of capacitor/reactor banks $O'$ to inject reactive power.

**Initialization:** $A \leftarrow \{v_{j0} : j \in F\} \cup \{v_{j1} : j \in O\}$, $A_0 \leftarrow F$, $A_1 \leftarrow O$, $flag = 1$

**while** $flag == 1$ **do**

$flag = 0$, $\Delta\hat{f}_0 = 0$, $\Delta\hat{f}_1 = 0$

**if** *there exists* $j_0 \in A_0$ *with* $\hat{f}(A \setminus \{v_{j_0 0}\} \cup \{v_{j_0 1}\}) < (1-\epsilon)\hat{f}(A)$ **then**

$j_0 \leftarrow \arg\min_{j \in A_0} \hat{f}(A \setminus \{v_{j0}\} \cup \{v_{j1}\})$

$flag = 1$, $\Delta\hat{f}_0 = \hat{f}(A \setminus \{v_{j_0 0}\} \cup \{v_{j_0 1}\})$

**if** *there exists* $j_1 \in A_1$ *with* $\hat{f}(A \setminus \{v_{j_1 1}\} \cup \{v_{j_1 0}\}) < (1-\epsilon)\hat{f}(A)$ **then**

$j_1 \leftarrow \arg\min_{j \in A_1} \hat{f}(A \setminus \{v_{j1}\} \cup \{v_{j0}\})$

$flag = 1$, $\Delta\hat{f}_1 = \hat{f}(A \setminus \{v_{j_1 1}\} \cup \{v_{j_1 0}\})$

**if** $\Delta\hat{f}_0 < \Delta\hat{f}_1$ **then**

$A \leftarrow A \setminus \{v_{j_0 0}\} \cup \{v_{j_0 1}\}$

$A_1 \leftarrow A_1 \cup \{j_0\}$, $A_0 \leftarrow A_0 \setminus \{j_0\}$

**else if** $\Delta\hat{f}_0 > \Delta\hat{f}_1$ **then**

$A \leftarrow A \setminus \{v_{j_1 1}\} \cup \{v_{j_1 0}\}$

$A_0 \leftarrow A_0 \cup \{j_1\}$, $A_1 \leftarrow A_1 \setminus \{j_1\}$

**if** $\hat{f}(A) \leq \hat{f}(\overline{\Omega} \setminus A)$ **then**

$O' \leftarrow A_1$, **return** $O'$

**else**

$O' \leftarrow \Omega \setminus A_1$, **return** $O'$

**end procedure**

*Proof.* Let $A^* = \{v_{i1} : i \in O^*\} \cup \{v_{i0} : i \in \Omega \setminus O^*\}$ and $A = \{v_{i1} : i \in O'\} \cup \{v_{i0} : i \in \Omega \setminus O'\}$. Define $\hat{g}(A) = M - \hat{f}(A)$, so that $\hat{g}$ is a nonnegative submodular function. By Lemma 4,

$$(2+\epsilon)\hat{g}(A) \geq \hat{g}(A \cup A^*) + \hat{g}(A \cap A^*). \tag{3.11}$$

From Algorithm 1, Line $22-26$ implies that $\hat{f}(A) \leq \hat{f}(\overline{\Omega} \setminus A)$. Applying it to (3.11) yields

$$(3+\epsilon)\hat{g}(A) \geq \hat{g}(A \cup A^*) + \hat{g}(\overline{\Omega} \setminus A) + \hat{g}(A \cap A^*). \tag{3.12}$$

By submodularity and nonnegativity of $\hat{g}$, we have

$$\begin{aligned} \hat{g}(A \cup A^*) + \hat{g}(\overline{\Omega} \setminus A) &\geq \hat{g}(A^c \cap (A \cup A^*)) + \hat{g}(A^c \cup (A \cup A^*)) \\ &= \hat{g}(\overline{\Omega}) + \hat{g}(A^* \setminus A) \geq \hat{g}(A^* \setminus A). \end{aligned}$$

Applying this inequality to (3.12) yields

$$(3+\epsilon)\hat{g}(A) \geq \hat{g}(A^* \setminus A) + \hat{g}(A \cap A^*) \geq \hat{g}(A^*) + \hat{g}(\emptyset) \geq \hat{g}(A^*)$$

by submodularity of $\hat{g}$. By Lemma 5, $\hat{g}(A) = M - f(O')$ and $\hat{g}(A^*) = M - f(O^*)$. Substitution of $\hat{g}$ then gives (3.10). $\square$

Denote the cardinality $|\Omega|$ by $N$. The complexity of Algorithm 1 is described below.

**Proposition 1.** *Let $f_{min} = \min\{f(O') : O' \subseteq \Omega\}$ and $f_{max} = \max\{f(O') : O' \subseteq \Omega\}$. The run-time of the algorithm is bounded above by $O(NT_0)$, where $T_0 = \lceil \frac{\log\left[\frac{f_{min}}{f_{max}}\right]}{\log(1-\epsilon)} \rceil$.*

*Proof.* Let $O_i'$ denote the set $O'$ after $i$ iterations of the algorithm. By construction, $f(O_i') < (1-\epsilon)f(O_{i-1}')$, and hence $f(O_i') < (1-\epsilon)^i f(O_0')$. Let $T$ denote the number of iterations before the algorithm terminates. By definition,

$$f_{min} \leq f(O_T') < (1-\epsilon)^T f(O_0') \leq (1-\epsilon)^T f_{max}.$$

Rearranging terms yields $T \leq T_0$. Each iteration requires at most $N$ evaluations of the objective function $f(O')$ in order to identify a capacitor/reactor bank to activate or deactivate, implying that the computation is $O(NT_0)$ in the worst case. $\square$

### 3.3.4 Adaptive selection algorithm

The approach of Algorithm 1 introduces additional errors due to the approximation resulting from the linearization of the nonlinear power flow equations (3.1)–(3.2). In what follows, we present an adaptive approach that reduces these errors. Instead of using fixed values of $\Delta Q$ and the Jacobian matrix $J$, the adaptive algorithm solves the power flow equations at each iteration after selecting a capacitor/reactor bank to switch. The values of $J$ and $\Delta Q$ are updated based on the new operating point. Algorithm 2 embodies this adaptive approach.

---

**Algorithm 2:** Algorithm for adaptively selecting a set of capacitor/reactor banks to inject reactive power.

---

**procedure** ADAPTIVE_VC($\mathbf{v}$, $\mathbf{v_{ref}}$, $\boldsymbol{\omega}$, $\mathbf{q}$, $\mathbf{b}$, $\mathbf{c}$, $O$, $F$)

  Line $2 - 14$ in Algorithm 1

  Compare $\Delta\hat{f}_0$ and $\Delta\hat{f}_1$

  **if** $\Delta\hat{f}_0 < \Delta\hat{f}_1$ **then**

    $A \leftarrow A \setminus \{v_{j_0 0}\} \cup \{v_{j_0 1}\}$

    $A_1 \leftarrow A_1 \cup \{j_0\}$, $A_0 \leftarrow A_0 \setminus \{j_0\}$

  [Switch $j_0$ and update $\hat{f}$: voltages $\mathbf{v}$; $O = A_1$ and $F = A_0$; Jacobian inverse matrix $J^{-1}$ and weights $\boldsymbol{\omega}$]

  **else if** $\Delta\hat{f}_0 > \Delta\hat{f}_1$ **then**

    $A \leftarrow A \setminus \{v_{j_1 1}\} \cup \{v_{j_1 0}\}$

    $A_0 \leftarrow A_0 \cup \{j_1\}$, $A_1 \leftarrow A_1 \setminus \{j_1\}$

  [Switch $j_1$ and Update $\hat{f}$: voltages $\mathbf{v}$; $O = A_1$ and $F = A_0$; Jacobian inverse matrix $J^{-1}$ and weights $\boldsymbol{\omega}$]

  Line $23 - 28$ in Algorithm 1

  **return** $O'$

**end procedure**

---

Algorithm 2 proceeds with a logic similar to Algorithm 1. At each iteration, the algorithm selects a device $j \in \Omega$, such that switching device $j$ will lower the cost function $\hat{f}(A)$. Before entering the next iteration, the centralized controller implements this switching action in simulation, and updates the cost function $\hat{f}$ by requesting the post-switching system state from the nonlinear power flow solver.

Compared to Algorithm 1, Algorithm 2 requires one extra step at each iteration for solving the power flow equations, denoted by $F(x) = 0$. If the nonlinear power flow solver uses Newton-Raphson iterative method, provided that a good initial guess is known, then it adds a complexity of $O((\log k)H(k))$ to each iteration of Algorithm 2, where $H(k)$ is the cost of calculating $F(x)/F'(x)$ with $k$-digit precision. Thus, the overall computational complexity of Algorithm 2 is $O((N + O((\log k)H(k)))T_0)$.

## *3.4 Case Study: 300-bus Transmission System*

This section presents test results of our proposed voltage controller on the IEEE 300-bus test system. A state-of-the-art algorithm [71] based on voltage sensitivities to reactive power changes is also implemented for comparison.

### *3.4.1 Test setup*

Power system state information (including voltage magnitudes $V$ and angles $\theta$) is obtained by solving power flows using Matpower [88]. Initial power flow data for the IEEE 300-bus test system is provided by the program default function `case300.m` and all test cases considered in this section are generated from this base case with variations in power demands. A power flow solution of the original `case300.m` produces an unrealistic voltage profile with a number of voltages both above and below the acceptable $[0.95, 1.05]$ per unit voltage range.

Because an optimal power flow (OPF) is not able to find a feasible solution for the original loading conditions, the following procedure was used to create realistic low voltage initial conditions. We obtain an initial operating point by scaling the active and reactive loads in the original data by a factor $\alpha < 1$. We then solve the OPF to get all bus voltages within the range $V_{\min}^{opf} = 0.95$ to $V_{\max}^{opf} = 1.05$. Finally we scale the active and reactive loads by a factor $\beta > 1$ at each PQ bus and solve the power flow equations. A larger $\beta$ value corresponds to an increase in loading and hence additional violations of the lower voltage limit. The following parameter values were used:

1) Case 1: $\alpha = 0.89, \beta = 1.07$.

2) Case 2: $\alpha = 0.89, \beta = 1.08$.

Once an initial operating point was created, we initialized the rest of the simulation as follows. The reference voltages at PQ buses were set to $\mathbf{v_{ref}} = \mathbf{1}$ p.u.. We assume that a capacitor bank is available at each PQ bus. The IEEE 300-bus system consists of a main transmission network with high nominal voltages (66-345 kV) and a small part (buses labeled 9001-9533 in `case300.m`) with low nominal voltages (0.6-13.8 kV). Thus, two types of capacitor banks are assumed to be installed at different PQ buses. Capacitors in the main transmission network are assumed to have capacitance $C = 0.003\ Mvar/kV^2$ (e.g., 155-Mvar capacitor bank at 230 kV, [71]), while capacitors at the low-voltage branch have capacitance $C = 0.03 Mvar/kV^2$ (e.g., 6-Mvar capacitor bank at 13.8 kV, [14]). We do not consider reactors in this simulation. We assume that no capacitor banks are energized at the initial operating point, i.e., $O = \emptyset$.

All algorithms are implemented in Matlab™. Given necessary information including the initial voltages, corresponding Jacobian matrix and capacitances of capacitor banks, each algorithm returns a set $S$ of buses where reactive power will be injected.

The total cost $f(S)$ of any switching action $S$ is calculated using Eq. (3.4). We choose a trade-off factor $\lambda = 1$. For switching costs at bus $i$, we set $b_i = 1$ and $c_i = 1$ for switching capacitor banks off and on, respectively. The parameters used in cost function $f$ are chosen to make the switching costs relatively small compared to penalties on the voltage deviation, such that switching solutions are differentiated mainly by their impact on the voltage.

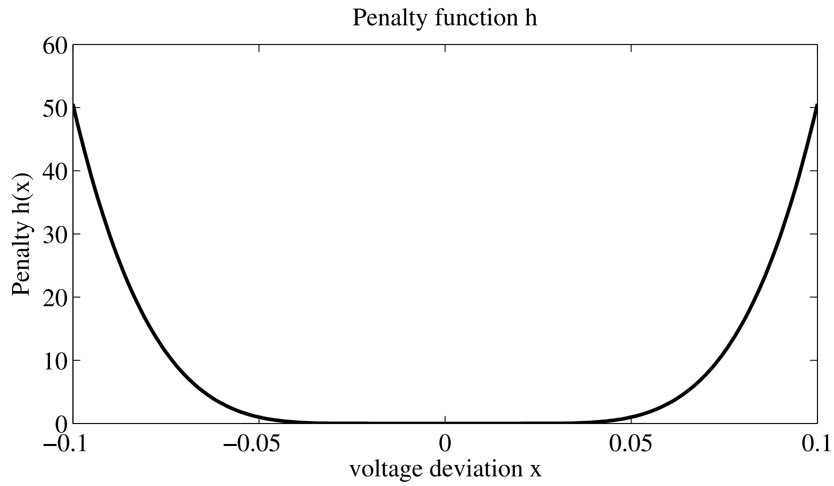



Figure 3.1: Convex penalty function $h(x)$ for voltage deviation.

The convex penalty function, $h$, on voltage deviations from Eq. (3.4) is defined by

$$h(x) = \begin{cases} \left(\frac{x - V_{high}}{V_{max} - V_{high}}\right)^4, & x > V_{high} \\ 0, & V_{low} < x \leq V_{high} \\ \left(\frac{x - V_{low}}{V_{min} - V_{low}}\right)^4, & x \leq V_{low} \end{cases}$$

where $V_{low} = -0.02$ and $V_{high} = 0.02$ are the lower and upper bounds on the desired voltage region, while $V_{min} = -0.05$ and $V_{max} = 0.05$ are limits defining voltage violations (Figure 3.1).

For comparison, we implement the sensitivity-based control algorithm of [71], which proceeds as follows. The controller first calculates voltage sensitivities of each bus $i$ due to a reactive power change at bus $k$ using the Jacobian matrix $J$ by

$$\Lambda_i = \min_j \left| \frac{\partial V_i}{\partial Q_k} / \frac{\partial V_j}{\partial Q_k} \right|, \ 0 \leq \Lambda_i \leq 1.$$

It then divides the system topology into $M$ control areas for $M$ voltage violating nodes. Given any threshold $\Lambda^*$, each area $k$ has nodes defined by the set $A_k = \{\forall i : \Lambda_i > \Lambda^* \text{ for any injection at bus } k\}$. Areas are combined if they are not disjoint, $A_x \cap A_y \neq \emptyset$. In each area, the controller enumerates all possible control actions and selects the set of actions $S$ that minimizes $f(S)$.

In the following subsections, we present the performance of submodular control Algorithm 1 and adaptive control Algorithm 2, as well as the sensitivity algorithm, for the two test cases we created. Buses are labeled by integers from $1$ to $300$ following the order in `case300.m`. We choose the parameter $\epsilon = 0$ for both Algorithm 1 and Algorithm 2.

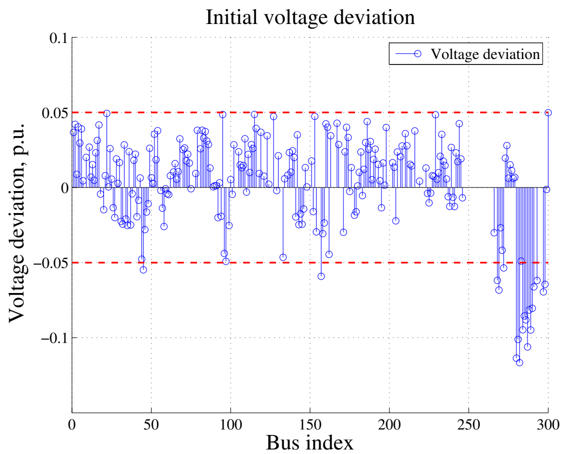
### *3.4.2 Comparison of submodular and sensitivity-based methods*

Figure 3.2(a) shows the initial voltages prior to any control actions for test case 1. 19 PQ buses have voltages below $0.95$ with a minimum voltage of $0.8834$ p.u. at Bus $282$. The maximum voltage is $1.0498$ p.u. at Bus $300$, which is within the limits. The initial operating point has a total cost $f(\emptyset) = 537.6943$.

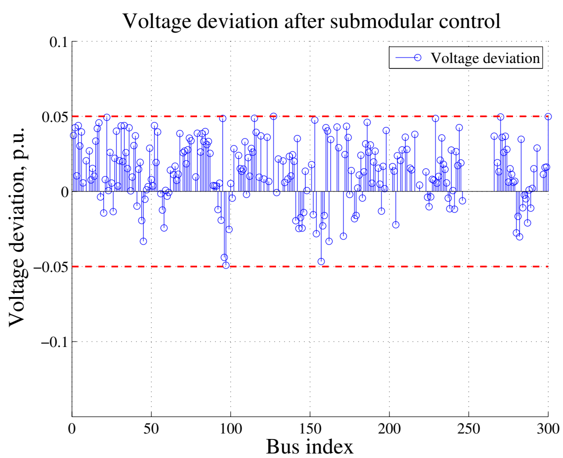
Since the error in the linear prediction of voltage deviation is negligible, we observe that Algorithm 1 and Algorithm 2 give the same switching set $S_{sub} = \{133, 157, 270\}$, which lowers the total cost to $f(S_{sub}) = 20.4634$. After control is exerted, no bus voltage falls below $0.95$ p.u., while the voltage of Bus $127$ is at the upper limit (Figure 3.2(b)). We observe that, while $19$ buses experienced low voltages, only one of them (157) was selected to inject reactive power.

For the sensitivity approach in test case 1, a threshold $\Lambda^* = 0.2$ is used as suggested in [71]. The sensitivity-based algorithm took five times as long to complete as the submodular control algorithm on a Macbook with $2.4$GHz dual-core Intel Core i5 processor and $8$GB RAM. The sensitivity approach identifies four areas ($A_1, A_2, A_3, A_4$) containing $1$, $3$, $5$ and $18$ buses, respectively. Four sets of buses, $S_1 = \{293\}$, $S_2 = \{44, 45, 48\}$, $S_3 = \{124, 157, 158, 159\}$, and $S_4 = \{268, 269, 271, 272, 285, 290, 291\}$ are selected to inject reactive power in areas $A_1, A_2, A_3$ and $A_4$, respectively (Figure 3.2(c)). This control solution ($S_{sens} = S_1 \cup S_2 \cup S_3 \cup S_4$) results in 3 PQ buses with voltage below $0.95$ p.u. while 1 bus (127) has a voltage of $1.0599$ p.u., which is above the upper limit. The total cost has been lowered to $f(S_{sens}) = 40.9636$.

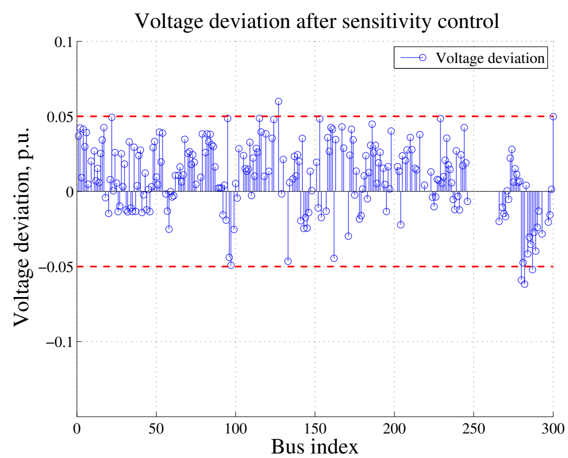
An intuitive explanation for the different outcomes of the submodular and sensitivity-based algorithms is as follows. In order to reduce computation time, the sensitivity algorithm partitions the power system into regions around buses with voltage violations, and attempts to find a set of reactive power injections within each region to resolve any voltage deviations. In this case,




(a)




(b)




(c)

Figure 3.2: Comparing submodular voltage control with sensitivity-based voltage control in test case 1. (a) Initial voltage deviation. Initially, 19 buses are below the desired operating voltage. (b) Effect of the submodular control algorithm on system voltages. After exerting control, one bus is at $1.05$ p.u. and no bus is below $0.95$ p.u.. (c) Change in voltage due to the sensitivity voltage control. After exerting the control, 3 PQ buses have voltages below the desired operating range of $[0.95, 1.05]$ p.u..

however, any additional switching actions beyond $S_4$ within region $A_4$ results in a minor reduction in voltage deviation at a high switching cost. The sensitivity algorithm therefore selects $S_4$ as the local optimal control actions within this region. On the other hand, our submodular control algorithm restores the voltages to the desired range by selecting devices outside the region $A_4$ to inject reactive power. In summary, this test case suggests that the submodular algorithm provides a better performance than sensitivity-based algorithm when different areas of the grid are highly interdependent. This is because an optimal strategy has to take into account local voltage deviations as well as the coupling between areas.

### *3.4.3 Comparison of non-adaptive and adaptive algorithms*

We now compare the performance of the adaptive control algorithm (Algorithm 2) with the non-adaptive algorithm (Algorithm 1) in test case 2. The initial operating point has $31$ buses with low voltages, including a minimum voltage of $0.8338$ p.u.. Voltages at the other buses are within range, with a maximum voltage of $1.0498$ p.u.. Before taking any control actions, the initial total cost is $f(\emptyset) = 3853.8877$.

This is a much more heavily loaded operating point compared to test case 1. Buses $280$, $282$ and $287$ have voltages that drop below $0.85$ p.u., which must be mitigated through reactive power injection. We set $\epsilon = 0$ for both the non-adaptive Algorithm 1 and the adaptive Algorithm 2. A large threshold $\Lambda^* = 0.92$ is used for the sensitivity algorithm, which divides the system into 10 areas with a maximum area size of 17 buses. This value of the threshold is chosen to limit the number of buses in the largest control area to a reasonable amount. Table 3.1 shows the control actions and the resulting costs. While all algorithms significantly reduce the total voltage deviations, Table 3.2 shows that the submodular algorithms (Algorithm 1 and 2) achieve better voltage profiles than the sensitivity-based controller. Figure 3.3 illustrates the performance of adaptive Algorithm 2 compared to non-adaptive Algorithm 1. We observe that the adaptive algorithm provides better performance due to a more accurate prediction of voltage deviations in the heavy loading case.

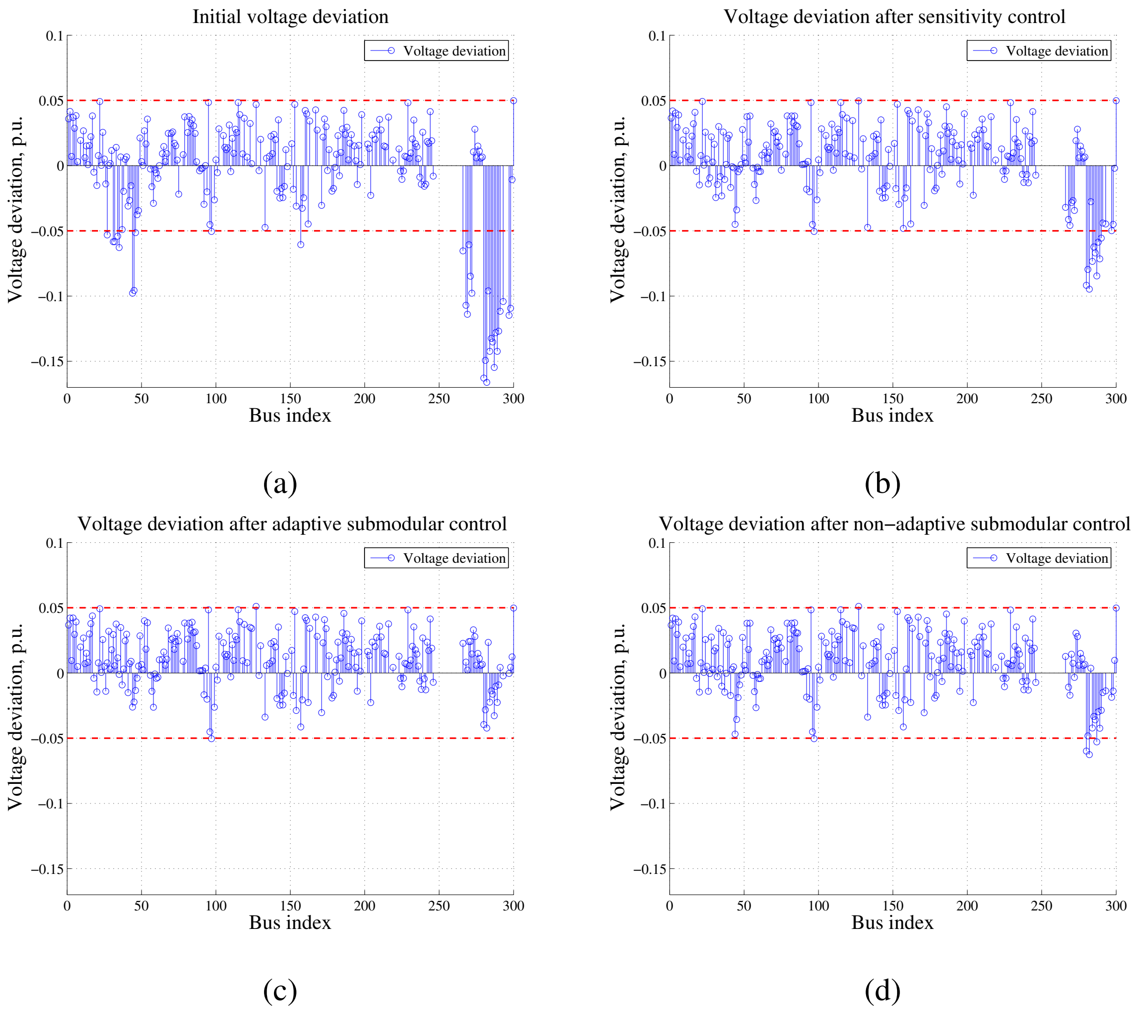



Figure 3.3: Comparing adaptive submodular, non-adaptive submodular and sensitivity-based voltage control in test case 2. (a) Initial voltage deviation. Initially, 31 buses are below the desired operating voltage. (b) Change in voltage due to the sensitivity voltage control. After exerting the control, no bus has voltage above $1.05$ p.u. but 11 buses have voltages below $0.95$ p.u.. (c) Effect of the adaptive submodular control algorithm on system voltages. After exerting control, 1 bus is above $1.05$ p.u. and 1 bus is below $0.95$ p.u.. (d) Effect of the non-adaptive submodular control algorithm on system voltages. After exerting control, 1 bus is above $1.05$ p.u. and 4 buses are below $0.95$ p.u..

Table 3.1: Comparison of control actions selected by adaptive submodular, non-adaptive submodular, and sensitivity-based algorithms in test case 2

| Algorithms | Buses selected to inject reactive power, $S$ | Total cost, $f(S)$ (initial 3853.8877) |
|---|---|---|
| Adaptive | {31,44,45,124,162,266,268} | 21.0875 |
| Non-adaptive | {45,124,162,266} | 27.8693 |
| Sensitivity | {27,32,34,35,44,45,46,157,269,272,280,281,282, 283,284,285,286,287,288,289,290,291,293,297,298} | 184.2068 |

Table 3.2: Voltage profile after exerting control under adaptive submodular, non-adaptive submodular, and sensitivity-based algorithms in test case 2

| Algorithms | Number of buses with low voltage (below $0.95$ p.u.) | Number of buses with high voltage (above $1.05$ p.u.) | Lowest voltage | Highest voltage |
|---|---|---|---|---|
| Adaptive | $1$ | $1$ | $0.9496$ | $1.0510$ |
| Non-adaptive | $4$ | $1$ | $0.9374$ | $1.0510$ |
| Sensitivity | $11$ | $0$ | $0.9053$ | $1.0498$ |

### *3.4.4 Voltage control following generator tripping*

Any change in system topology, such as a generator or transmission line tripping, could also cause low voltages ($V < 0.95$ p.u.). This is because after rebalancing the generation and load following the topology change, the total reactive power loss may increase, for example, due to more power flows on high inductance transmission lines . This subsection presents a test of the proposed submodular voltage control algorithm on an initial operating point that has dozens of buses experiencing low voltage following a generator tripping, with a comparison to the sensitivity-based control algorithm. The goal of this test is to demonstrate the ability of our proposed submodular voltage control algorithms to correct voltage deviations arising from a system topology change.

As in the previous two test cases, the initial operating point for this test case is obtained from the IEEE 300-bus system by scaling the active and reactive loads in the original data `case300.m` down by the factor $\alpha = 0.89$. The OPF solution shows that all bus voltages are then within the range $V_{\min}^{opf} = 0.95$ to $V_{\max}^{opf} = 1.05$ p.u.. All other parameters and assumptions are the same as in the previous test cases.

The generator at Bus 165 is chosen to be tripped. Without load shedding, the resulting loss in active power is fully compensated by the slack bus. After this generator tripping, a power flow solution produces a voltage profile with 25 buses having voltages below $0.95$, the lowest voltage being $0.8614$ p.u. These low voltages occur due to an increase in reactive power losses from $3972.5$ Mvar to $5641.6$ Mvar throughout the grid. Before exerting any controls, the voltage deviations resulting from the generator tripping incur a total cost $f(\emptyset) = 1564.8701$.

We select the same parameters for algorithms as in test case 2, in order to limit the computational times of the different algorithms to a comparable amount. For both the non-adaptive Algorithm 1 and the adaptive Algorithm 2, we set $\epsilon = 0$. For the sensitivity algorithm, the threshold $\Lambda^* = 0.92$ is used, which divides the power system into 5 areas with a maximum area size of 17 buses. In this test case, the sensitivity-based algorithm requires at least three times the computing time required by the submodular control algorithms.

Table 3.3 lists the buses selected for reactive power injection and the resulting cost for each algorithm. The detailed voltage profile after exerting control with each algorithm is given in Table 3.4. We observe that the submodular voltage control algorithms select fewer buses to inject reactive power but results in a solution with more buses returning to the normal operating region compared with the sensitivity algorithm. In particular, both the adaptive submodular algorithm and non-adaptive submodular algorithm are able to find a “key” bus, Bus 270, where injecting reactive power can mitigate voltage deviations at most buses. The voltage at this bus, however, is less sensitive to its own reactive power injection compared to its effect on some other buses. Because a large threshold $\Lambda^* = 0.92$ is used to ensure that the sensitivity algorithm is computationally feasible, Bus 270 is excluded from the control area in this case. Although a large number of buses are selected, the sensitivity-based algorithm cannot effectively reduce the voltage deviations at all buses compared to the submodular algorithms.

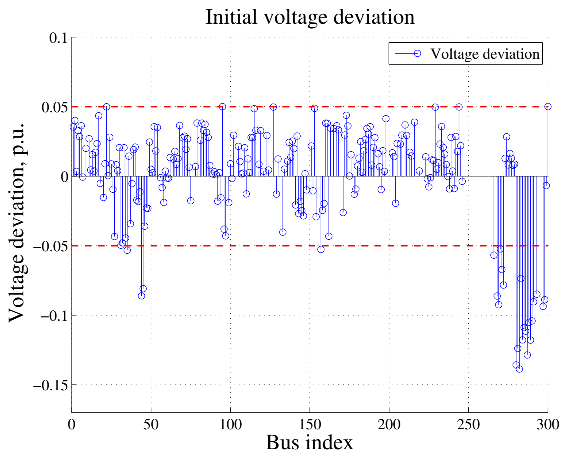
Figure 3.4 demonstrates the shift of voltages at all PQ buses with the non-adaptive submodular and the sensitivity-based algorithms. We observe that in the case of tripping a generator at Bus 165, the proposed submodular voltage control can achieve better voltage profiles compared to the sensitivity approach.

Table 3.3: Comparison of control actions selected by the adaptive submodular, non-adaptive submodular, and sensitivity-based algorithms after generator tripping.

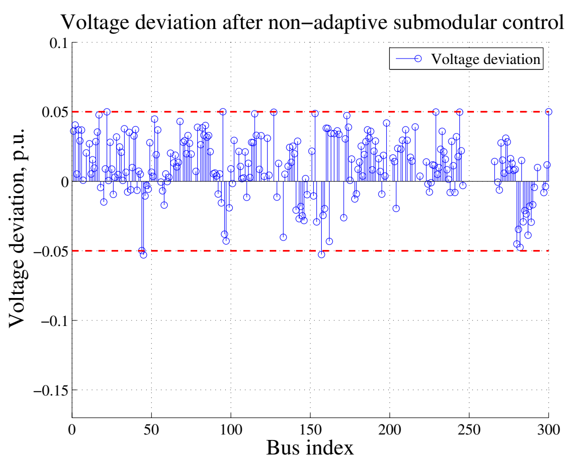
| Algorithms | Buses selected to inject reactive power, $S$ | Total cost, $f(S)$ (initial 3853.8877) |
|---|---|---|
| Adaptive | $\{$44,270$\}$ | 19.1012 |
| Non-adaptive | $\{$270$\}$ | 20.5082 |
| Sensitivity | $\{$35,44,45,157,269,272,280,281,282,283, 284,285,286,287,288,289,290,291,293$\}$ | 165.4577 |

Table 3.4: Voltage profile after exerting control with the adaptive submodular, non-adaptive submodular, and sensitivity-based algorithms after generator tripping.

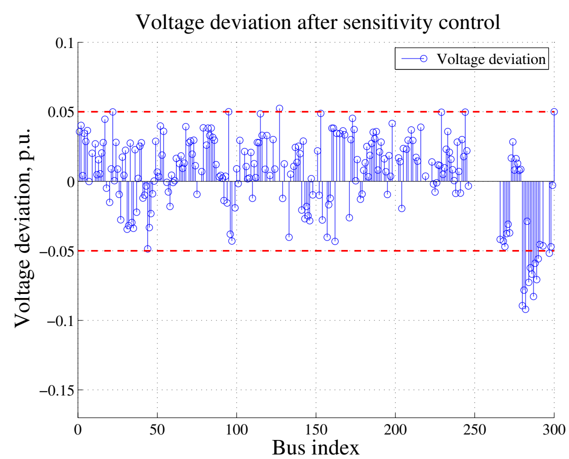
| Algorithms | Number of buses with low voltage (below $0.95$ p.u.) | Number of buses with high voltage (above $1.05$ p.u.) | Lowest voltage | Highest voltage |
|---|---|---|---|---|
| Adaptive | $1$ | $0$ | $0.9473$ | $1.0500$ |
| Non-adaptive | $2$ | $0$ | $0.9471$ | $1.0510$ |
| Sensitivity | $11$ | $1$ | $0.9080$ | $1.0524$ |




(a)




(b)




(c)

Figure 3.4: Comparing submodular voltage control and sensitivity-based voltage control in generator tripping case. (a) Initial voltage deviation. In total, 25 buses are below the desired voltage. (b) Effect of the submodular control algorithm on system voltages. After exerting control, no bus is above $1.05$ p.u. and 2 buses are below $0.95$ p.u.. (c) Change in voltage due to the sensitivity voltage control. After exerting control, 1 bus is above $1.05$ p.u. and 11 buses are below $0.95$ p.u..

### 3.5 Conclusion

In this chapter, we proposed computationally efficient centralized voltage control algorithms with provable optimality bounds that are scalable to large power systems. It presented a discrete optimization approach for selecting a subset of reactive power injection devices to minimize switching costs and voltage deviations from desired levels. We proved that this formulation is equivalent to submodular maximization with a matroid basis constraint, leading to efficient approximation algorithms with provable bounds. The algorithms were evaluated numerically on the IEEE 300-bus transmission system and compared to an existing voltage control algorithm. We found that our proposed submodular control approach achieves voltage profiles with less voltage deviation in the same computation time.

# Chapter 4

# INPUT AND OUTPUT SELECTION FOR SMALL SIGNAL STABILITY

## *4.1 Introduction*

Small-signal stability is a critical property of power systems, and hence has received extensive research attention over the past several decades [36, 82, 32, 33, 28]. Recently, wide-area control methodologies that use real-time state information from PMUs have been proposed to mitigate inter-area oscillations [83, 82]. These related works have identified the selection of generators to implement wide-area control as a key component of the design, and have proposed selecting generators based on geometric controllability and observability indices [83]. These current approaches, however, do not guarantee that the number of generators and measurements selected is within a provable bound of the minimum size.

Input and output selection ensuring stability-related properties is a critical step in controlling networked systems and hence has received extensive research attention over decades. Early approaches were summarized in [70], where various selection criteria were discussed, including controllability/observability, frequency-dependent minimum singular value, and condition number. These methods, however, cannot scale to large-scale systems that have a nearly unlimited number of interacting parts, such as biological systems and the smart grid. Initial studies of identifying a minimum-size set of input nodes to control a large-scale complex network were presented in [40].

In recent years, a series of works have been proposed to address the problem of selecting the minimum-size set of inputs for linear control systems with a diagonal input matrix while satisfying certain controllability requirements, known as the minimal controllability problem. This problem was first presented in [50], where greedy algorithms were proposed to approximate the optimal solution with provable guarantees on sparsity. The exact solution to the minimal controllability problem was studied in [55], where authors show that this problem can be reduced to the minimum set covering problem.

Most existing works on input and output selection focus on selecting a minimum-size set of nodes to satisfy one or more performance criteria [54, 10, 69, 63, 53, 52]. In [69], the authors studied the problem of selecting a minimum-size set of inputs to ensure controllability while the average minimum control effort for stabilizing the system is within a bound. In [54], the constraint is to ensure structural controllability, defined by the existence of a controllable numerical realization of the linear system with the same structure (i.e., zero/nonzero patterns) in system matrices. This problem was extended in [53] by considering output selection and an additional constraint on the communication cost associated with feeding outputs to inputs. The submodular structures of the minimal input selection problems with different constraints were exploited in [10] and [63], and efficient selection algorithms have been proposed to approximate each problem with provable optimality bounds. The joint problem of selecting control nodes and designing controller actions is addressed in [52]. All these approaches select minimal sets of input nodes to ensure the controllability of a nominal system, while our approach focuses on the stabilizability of uncertain systems, defined by the ability of a system to reach an equilibrium.

Controllability is a sufficient but not necessary condition to ensure system stability in many control systems. Controllability implies that all system modes are controllable, while stabilizability only requires that a subset of unstable modes be controllable [6]. The problem of selecting a minimal set of inputs to achieve stabilizability was studied in [68], where approximation algorithms were proposed to greedily select input nodes until the desired modes are contained in the controllability matrix subspace. These works assume known, fixed system parameters and have not taken into account plant uncertainties or time delays.

Uncertainties and time delays are common problems in linear systems. In [15], the authors have proposed a heuristic approach towards the input and output selection problem guaranteeing system robustness to uncertainties. More recent approaches to minimal input and output selection for systems with uncertainties rely on relaxing the discrete combinatorial problem to continuous convex optimization forms. For example, the sparsity-promoting method using $l_1$-norm [19], the $H_2/H_\infty$

optimization approach for centralized control systems [49], and the minimal input selection for decentralized control systems [60]. These heuristic methods, however, do not provide guarantees on the optimality of solutions of the initial problem.

In this chapter, we propose a bounded submodularity ratio approach with provable optimality guarantees to the minimal input and output selection problem that ensures the existence of a robust control and achieves system reachability from a given state to the origin. Our key insights are the following. First, we show that it is sufficient to guarantee the existence of such a stabilizing controller if a subset of system modes, determined by the $H_\infty$ norm of the uncertain perturbation, have eigenvectors lying in both the controllability and observability matrix subspaces induced by the selected input and output nodes, respectively. Second, we show that the problem of selecting the minimum-size sets of input and output nodes that satisfy the above controllability and observability conditions is equivalent to a column-subset selection problem with bounded submodularity ratio [13]. We make the following specific contributions:

- We derive a sufficient condition for a set of input nodes and a set of output nodes to guarantee the existence of a feedback controller that can stabilize a given uncertain linear system. This condition is based on the fact that a set of system modes is required to lie in the controllability and observability matrix subspace resulting from the input and output nodes. We extend our results to systems with uncertain output delays, as well as systems with structured uncertainties.

- We formulate the problem of selecting minimum-cardinality sets of input and output nodes to satisfy the derived conditions. We prove that the input and output selection problems are equivalent to minimizing the Euclidean distance between eigenvectors of the identified system modes and the controllability/observability subspace. We also prove that these distances are monotone decreasing functions of the selected input/output sets with bounded submodularity ratio.

- We propose polynomial-time approximation algorithms for selecting minimum-cardinality sets of input and output nodes guaranteeing system stabilizability. We characterize the optimality gap/bound of our approach via the submodularity ratio, which we derive as a function of the system parameters. We analyze different types of uncertainties under our framework, including uncertain delays and structured uncertainties.

- We evaluate our approach through a numerical study on the IEEE 39-bus power system [48]. From the results, we conclude that our proposed input/output selection approach better ensures the system robustness to uncertainties, compared to selections that do not take uncertainties into account. With the same perturbation, the proposed algorithm with robustness consideration results in a system that is stable in a higher percentage of cases compared to the selection without considering uncertainties. In comparison to a geometric indices based input/output selection approach, our method requires fewer input and output nodes to satisfy the stability conditions.

The difficulty of approximating minimal reachability problems is presented in [30], which proves that there is no polynomial-time algorithm to achieve a set of inputs within a constant factor of true optimal selections. In this chapter, we show that the bounded submodularity ratio alone is sufficient to achieve certain approximations, even without submodularity or supermodularity. The optimality bound, however, can be arbitrarily small for certain system parameters and hence is consistent with [30].

The rest of this chapter is organized as follows. Section 4.2 introduces the system model. Section 4.3 presents the proposed approach to input and output selections. Section 4.4 contains analysis for different types of uncertainties. Section 4.5 presents numerical results. Section 4.6 concludes this chapter.

## 4.2 System Model and Preliminaries

### 4.2.1 Power System Model

We consider a multi-machine power system with $p$ generators, where each generator's dynamics are represented by its swing equation and locally closed-loop controls including excitation system, Automatic Voltage Regulator (AVR) and PSS. Generators are coupled through power flows. Using a Kron reduction [36], all $PQ$ buses that are not directly connected to a generator can be eliminated and the system topology can be reduced to a fully connected circuit graph where nodes represent generators. The resistance of the transmission lines is neglected. The Kron-reduced admittance matrix $Y$ has pure imaginary entries $-\mathbf{j}Y_{ij}$, where $\mathbf{j} = \sqrt{-1}$, that describe the interactions between generators $i$ and $j$. Therefore, the power flow from generator $i$ to $j$ is given as

$$P_{ij} = \Re\left[E_i e^{\theta_i}\left(-\mathbf{j}Y_{ij}(E_i e^{\mathbf{j}\theta_i} - E_j e^{\mathbf{j}\theta_j})\right)^*\right] = Y_{ij}E_iE_j\sin(\theta_i - \theta_j) \tag{4.1}$$

where $E_i$ is the magnitude of the internal electromotive force (e.m.f) of generator $i$, which is assumed to remain constant at the pre-disturbance value; $\theta_i$ is the rotor angle of generator $i$. The generator internal e.m.f magnitudes $E_i, E_j$ are controlled by the exciter, AVR, and PSS of each generator.

At each generator $i$, linearizing the swing dynamics around an operating point, denoted by $\{\theta_i^0, i = 1, \ldots, p\}$, gives

$$\Delta\dot{\theta}_i = \Delta\omega_i \tag{4.2}$$

$$\Delta\dot{\omega}_i = \frac{1}{M_i}\left(-D_i\Delta\omega_i - \sum_{j=1;j\neq i}^{p} Y_{ij}E_iE_j\cos(\theta_i^0 - \theta_j^0)(\Delta\theta_i - \Delta\theta_j)\right) \tag{4.3}$$

where $\omega_i$ represents the rotor frequency of generator $i$; $M_i$ and $D_i$ are the inertia and damping coefficients. The mechanical power injections $P_m$ are assumed to be constants.

Considering $p$ generators at the same time, the swing dynamics are given by

$$\begin{bmatrix}\Delta\dot{\theta}\\ \Delta\dot{\omega}\end{bmatrix} = \begin{bmatrix}0_{p\times p} & I_{p\times p}\\ M^{-1}L & -M^{-1}D\end{bmatrix}\begin{bmatrix}\Delta\theta\\ \Delta\omega\end{bmatrix} \tag{4.4}$$

where $\Delta\theta = [\Delta\theta_1, \Delta\theta_2, \ldots, \Delta\theta_p]^T$ and $\Delta\omega = [\Delta\omega_1, \Delta\omega_2, \ldots, \Delta\omega_p]^T$; $M$ and $D$ are diagonal matrices with diagonal elements $M_i$ and $D_i$ respectively; $L$ is a dense matrix with off-diagonal elements $L_{ij} = Y_{ij}E_iE_j\cos(\theta_i^0 - \theta_j^0)$ and diagonal elements $L_{ii} = -\sum_{j=1;j\neq i}^{p} L_{ij}$.

Considering other local dynamics such as AVR and PSS, we specify the internal e.m.f magnitude changes $\Delta E_i$ at each generator as a function of local state variables $\Delta E_{fd_i}, \Delta E_{f_i}$ and the central control signal $u_i$. As a result, the change in rotor frequency $\Delta\omega_i$ is also affected by these local variables in the following way: A new term $-\frac{K_2}{M_i}\Delta E_{f_i}$ is added to $\Delta\omega_i$'s dynamics (4.3), while $\Delta E_{fd_i}$ and $\Delta E_{f_i}$ have dynamics

$$\begin{aligned}
\Delta\dot{E}_{fd_i} &= -\frac{1}{\tau_{Ai}}\left[K_{Ai}K_{5i}\Delta\theta_i + \Delta E_{fd_i} + K_{Ai}K_{6i}\Delta E_{f_i}\right] + \frac{K_{Ai}}{\tau_{Ai}}(u_{pss_i} + u_i) \\
\Delta\dot{E}_{f_i} &= \frac{1}{\tau_{doi}}\left[-K_{4i}\Delta\theta_i + \Delta E_{fd_i} - \frac{1}{K_{3i}}\Delta E_{f_i}\right]
\end{aligned}$$

where $\tau_{Ai}, \tau_{doi}, K_{Ai}, K_{2i}, K_{3i}, K_{4i}, K_{5i}, K_{6i}$ are local AVR parameters; $u_{pss_i} = K_{pss_i}\Delta\omega_i$ is the local PSS closed-loop control input with feedback parameter $K_{pss_i}$; $u_i$ is the central control signal sent to local PSS.

Define local state vectors $\Delta E_{fd} = [\Delta E_{fd_1}, \ldots, \Delta E_{fd_p}]^T$ and $\Delta E_f = [\Delta E_{f_1}, \ldots, \Delta E_{f_p}]^T$. The system model in (4.4) can then be extended to include the local AVR and PSS dynamics. Hence we have

$$\begin{bmatrix} \Delta\dot{E}_{fd} \\ \Delta\dot{E}_f \\ \Delta\dot{\theta} \\ \Delta\dot{\omega} \end{bmatrix} = A \begin{bmatrix} \Delta E_{fd} \\ \Delta E_f \\ \Delta\theta \\ \Delta\omega \end{bmatrix} + \begin{bmatrix} T_A^{-1}K_A \\ 0_{p\times p} \\ 0_{p\times p} \\ 0_{p\times p} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_p \end{bmatrix} \tag{4.5}$$

where

$$A = \begin{bmatrix} -T_A^{-1} & T_A^{-1}K_AK_6 & T_A^{-1}K_AK_5 & T_A^{-1}K_AK_{pss} \\ T_{do}^{-1} & -T_{do}^{-1}K_3^{-1} & -T_{do}^{-1}K_4 & 0_{p\times p} \\ 0_{p\times p} & 0_{p\times p} & 0_{p\times p} & I_{p\times p} \\ 0_{p\times p} & -M^{-1}K_2 & M^{-1}L & -M^{-1}D \end{bmatrix}$$

and $T_A$, $T_{do}$, $K_A$, $K_2$, $K_3$, $K_4$, $K_5$, $K_6$, $K_{pss}$ are diagonal matrices with diagonal elements $\tau_{Ai}$, $\tau_{doi}$, $K_{Ai}$, $K_{2i}$, $K_{3i}$, $K_{4i}$, $K_{5i}$, $K_{6i}$, $K_{pss_i}$ respectively.

The full system model (4.5) can be generalized into the following state-space form:

$$\dot{x} = Ax + Bu \tag{4.6}$$

$$y = Cx \tag{4.7}$$

where $x = [\Delta E_{fd}^T, \Delta E_f^T, \Delta\theta^T, \Delta\omega^T]^T \in R^{4p\times 1}$ is the state vector; $y \in R^{m\times 1}$ is a vector of those states in $x$ that are observable; $u = [u_1, \ldots, u_p]^T \in R^{p\times 1}$ is the input vector whose entries $u_i$ are control signals sent by the central controller to generator $i$; $A \in R^{4p\times 4p}$ has a set $\Gamma$ of modes with corresponding orthonormal eigenvectors $v_i, i \in \Gamma$; $B \in R^{4p\times p}$ and $C \in R^{m\times 4p}$ are rectangular diagonal matrices. We denote $B = [b_1, \ldots, b_p]$ where $b_i$ is the $i$th column with nonzero entry $b_{ii}$.

The power system is stable without the help of control input $u$ when all eigenvalues of matrix $A$ have only non-positive real parts. Small signal instability arises when $A$ has eigenvalues with positive real parts under some operating conditions. This occurs when generators in different areas oscillate against each other. In such cases, the central control signal $u$ attempts to stabilize the system by driving current state $x_0$ to the origin.

### *4.2.2 Generalized System Model with Uncertainties*

We consider a more general linear system with a perturbation $\delta(t)$ in system dynamics, described by

$$\dot{x}(t) = Ax(t) + Bu(t) + \delta(t) \tag{4.8}$$

$$y(t) = Cx(t), \tag{4.9}$$

where $x \in \mathbb{R}^n$ is the vector of $n$ states, $y \in \mathbb{R}^m$ is the output vector, and $u \in \mathbb{R}^p$ is the input vector. The $i$th entries of $u$ and $y$ are denoted by $u_i$ and $y_i$, representing an input control signal and an output observation, respectively. Matrices $A$, $B$ and $C$ are constant and given, where $A \in \mathbb{R}^{n\times n}$ has eigenvalues $\{\lambda_i\}$ and corresponding eigenvectors $\{v_i\}$.

The vector $\delta \in \mathbb{R}^n$ represents an unknown perturbation affecting the system dynamics and the resulting system (4.8)-(4.9) is an uncertain system. For many types of perturbations, the uncertain term $\delta$ can be expressed as a linear function of the state $x$ and an unknown matrix $\Delta$ whose singular values are bounded by $\|\Delta\|_\infty \leq \sigma$. For example, when an additive perturbation occurs in matrix $A$ and results in a system $\dot{x}(t) = (A + \Delta)x(t) + Bu(t)$, the uncertain term is represented by $\delta(t) = \Delta x(t)$.

Eq. (4.8)-(4.9) generalize both uncertain linear systems and certain linear systems. A system with no uncertainty will have $\delta(t) = 0$ for all $t$, or equivalently, $\|\Delta\|_\infty = 0$ assuming $\delta(t) = \Delta x(t)$.

Given a linear system (4.8)-(4.9), the input matrix $B \in \mathbb{R}^{n\times p}$ has columns $\{b_1, \ldots, b_p\}$ where each column $b_i$ corresponds to the influence of a control input $u_i$ to the system, while the output matrix $C \in \mathbb{R}^{m\times n}$ has rows $\{c_1, \ldots, c_m\}$ where each row $c_i$ represents the connection between an output $y_i$ and the states $x$.

Let $\Omega_{in} = \{1, \ldots, p\}$ be the set of indices of all inputs and let $\Omega_{out} = \{1, \ldots, m\}$ be the set of indices of all outputs. Selecting a subset $S \subseteq \Omega_{in}$ of inputs and a subset $T \subseteq \Omega_{out}$ of outputs changes the system dynamics to

$$\dot{x}(t) = Ax(t) + B_{S}u_{S}(t) + \delta(t) \tag{4.10}$$

$$y_{T}(t) = C_{T}x(t), \tag{4.11}$$

where the input vector $u_S$ has entries $\{u_i : i \in S\}$; the matrix $B_S$ has columns $\{b_i : i \in S\}$; the output vector $y_T$ has entries $\{y_i : i \in T\}$; and the matrix $C_T$ has rows $\{c_i : i \in T\}$. The input and output selections therefore impact the system dynamics by determining the columns of matrix $B$ and the rows of matrix $C$, respectively.

The goal of the input and output selection is to ensure the existence of a feedback control $u_S$ based on output $y_T$ such that the system (4.10)-(4.11) can be driven from the initial state $x(t_0)$ at time $t_0$ to the origin at some time $t_1$ in the presence of uncertainty $\delta(t)$. Such control $u_S$ is called a stabilizing controller. In our approach, we consider the following design of such a controller.

Let $\hat{x}(t)$ be an estimate of the system state $x(t)$ with dynamics

$$\dot{\hat{x}}(t) = A\hat{x}(t) + B_{S}u_{S}(t) + L(y(t) - C_{T}\hat{x}(t)),$$

where $L$ is a weighting matrix on the output error $(y(t) - C_T\hat{x}(t))$ that helps correct the estimation dynamics. Defining the estimation error $e(t) = x(t) - \hat{x}(t)$ induces the following dynamics

$$\dot{e}(t) = (A - LC_T)e(t) + \delta(t). \tag{4.12}$$

We consider a feedback control based on the state estimation, defined by

$$u_S(t) = -K\hat{x}(t), \tag{4.13}$$

where $K$ is the control matrix to be designed. Then by involving the feedback control (4.13) and the state estimation dynamics (4.12), the closed-loop dynamics of system (4.10)-(4.11) is given by

$$\begin{bmatrix} \dot{x}(t) \\ \dot{e}(t) \end{bmatrix} = \begin{bmatrix} A - B_S K & B_S K \\ 0 & A - LC_T \end{bmatrix} \begin{bmatrix} x(t) \\ e(t) \end{bmatrix} + \begin{bmatrix} I \\ I \end{bmatrix} \delta(t). \tag{4.14}$$

The system (4.14) is asymptotically stable if there exists a feedback matrix $K$ and a state-estimate matrix $L$ such that the initial state $[x(t_0), e(t_0)]$ at time $t_0$ can be driven to zero at some time $t_1$ in the presence of uncertainty $\delta(t)$. Such a pair of matrices $(K, L)$ gives a stabilizing controller.

#### *4.2.3 $M - \Delta$ Loop Representation*

A system in $M - \Delta$ loop has structures as shown in Fig. 4.1, where the block $\Delta$ is an uncertain matrix with singular values bounded by $\|\Delta\|_\infty \leq \sigma$, while the block $M$ is the nominal dynamics of the system. The vector $\overline{y}$ represents the parameters in system $M$ that incurs perturbation and $\overline{u}$ represents the impact of the uncertainty $\Delta$ into the system $M$. They are related by $\overline{u} = \Delta\overline{y}$.

The block $M$ has a state-space realization described by the following equations

$$\dot{\overline{x}} = A_{cl}\overline{x} + B_{cl}\overline{u} \tag{4.15}$$

$$\overline{y} = C_{cl}\overline{x}, \tag{4.16}$$

where $\overline{x}$ is the system state and $A_{cl}, B_{cl}, C_{cl}$ are known constant matrices.

The closed-loop system (4.14) can be transformed into an $M - \Delta$ loop system for stability analysis [86]. The parameters in (4.15)-(4.16) relate to those in Eq. (4.14) by the following equa-

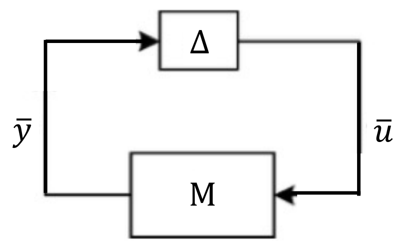



Figure 4.1: M-Δ loop for stability analysis.

tions

$$\overline{x} = \begin{bmatrix} x \\ e \end{bmatrix}, \ A_{cl} = \begin{bmatrix} A - B_S K & B_S K \\ 0 & A - LC_T \end{bmatrix}, \ B_{cl}\Delta C_{cl}\overline{x} = \begin{bmatrix} I \\ I \end{bmatrix} \delta.$$

The values of $B_{cl}$ and $C_{cl}$ depend on how the uncertainty $\delta$ is represented in terms of $\Delta$. For example, when $\delta = \Delta x$, we have $B_{cl} = \begin{bmatrix} I & I \end{bmatrix}^T$ and $C_{cl} = \begin{bmatrix} I & 0 \end{bmatrix}$. Denote $b$ and $c$ as the smallest numbers satisfying $B_{cl}B_{cl}^T \preceq bI$ and ${C_{cl}}^T C_{cl} \preceq cI$. The transformations to $M - \Delta$ systems for different types of uncertainties will be studied in Section 4.4.

A sufficient and necessary condition guaranteeing the stability of the $M - \Delta$ system is given by the following lemma.

**Lemma 6** ([86], Small Gain Theorem)**.** *The interconnected system in Fig. 4.1 is internally stable for all $\Delta$ with $\|\Delta\|_\infty \leq \sigma$ if and only if $\|M\|_\infty < 1/\sigma$.*

The following lemma shows a condition equivalent to the stability requirement given by Lemma 6, on time-domain matrices $(A_{cl}, B_{cl}, C_{cl})$.

**Lemma 7** ([20], Lemma 7.4)**.** *Suppose $M = \left[\begin{array}{c|c} A_{cl} & B_{cl} \\ \hline C_{cl} & 0 \end{array}\right]$. Then $\|M\|_\infty < 1$ if and only if there exists a positive definite matrix $X \succ 0$ such that*

$$\begin{bmatrix} C_{cl}^T \\ 0 \end{bmatrix} \begin{bmatrix} C_{cl} & 0 \end{bmatrix} + \begin{bmatrix} \left(A_{cl}^T X + XA_{cl}\right) & XB_{cl} \\ B_{cl}^T X & -I \end{bmatrix} \prec 0. \tag{4.17}$$

### 4.3 Proposed Input and Output Selection Framework

In this section, we propose a framework to solve the problem of selecting a minimal set of inputs and a minimal set of outputs that guarantee the existence of a stabilizing controller. In the first subsection, we give the exact problem formulation, followed by a sufficient condition in order to achieve computational feasibility. In the remaining subsections, we prove that the minimal input selection and minimal output selection to satisfy the sufficient condition are equivalent to two column-subset selection problems with bounded submodularity ratios. By exploiting the submodular structure, we present a polynomial-time greedy algorithm for input and output selections with provable optimality guarantees on the minimum number of inputs and outputs required.

#### 4.3.1 Problem Formulation

The problem studied in this chapter is to select a minimum-size set $S \subseteq \Omega_{in}$ of inputs and a minimum-size set $T \subseteq \Omega_{out}$ of outputs such that there exists a controller $u_s = -K\hat{x}$ that stabilizes the system in the presence of any uncertainty $\delta$. In what follows, we formulate the problem mathematically and derive a sufficient condition on selections to satisfy the stability requirement.

In the equivalent $M - \Delta$ representation (4.15)-(4.16) of the system (4.10)-(4.11), the input selection $S$ and output selection $T$ have impact on matrix $A_{cl}$ by determining matrices $B_S$ and $C_T$ respectively,

$$A_{cl} = \begin{bmatrix} A - B_S K & B_S K \\ 0 & A - LC_T \end{bmatrix}. \tag{4.18}$$

Assume the uncertainty block $\Delta$ has bounded singular values, i.e., $\|\Delta\|_\infty \le \sigma$. By Lemma 6, a stable $M - \Delta$ system requires the largest singular value of the transfer function $M$ less than $1/\sigma$. In the time domain, $M$ has a realization in state-space system $(A_{cl}, B_{cl}, C_{cl})$ as defined by (4.15)-(4.16).

For any transfer function $M = C_{cl}\left(sI - A_{cl}\right)^{-1} B_{cl}$, an equivalent condition to $\|M\|_\infty < 1/\sigma$ is that $\|M'\|_\infty < 1$ where $M' = \sqrt{\sigma} C_{cl}\left(sI - A_{cl}\right)^{-1} \sqrt{\sigma} B_{cl}$.

By substituting $M'$ for $M$ in Lemma 7, we have the following result.

**Corollary 1.** $\|M\|_\infty < 1/\sigma$ *if and only if*

$$\begin{bmatrix} \sqrt{\sigma}C_{cl}^T \\ 0 \end{bmatrix} \begin{bmatrix} \sqrt{\sigma}C_{cl} & 0 \end{bmatrix} + \begin{bmatrix} \left(A_{cl}^T X + X A_{cl}\right) & \sqrt{\sigma} X B_{cl} \\ \sqrt{\sigma} B_{cl}^T X & -I \end{bmatrix} \prec 0 \tag{4.19}$$

*for some* $X \succ 0$.

Applying a Schur complement operation, the inequality (4.19) is equivalent to

$$\sigma {C_{cl}}^T C_{cl} + A_{cl}^T X + X A_{cl} + \sigma X B_{cl} B_{cl}^T X \prec 0. \tag{4.20}$$

By Lemma 6 and 7, the problem of selecting minimal sets of inputs and outputs to ensure existence of a stabilizing control robust to uncertainties, $\|\Delta\|_\infty \leq \sigma$, can be formulated as follows.

For input selection, the objective is

$$\min_{S \subseteq \Omega_{in}} \; |S| \quad s.t. \; (4.20) \text{ holds for some } X \succ 0. \tag{4.21}$$

For output selection, the objective is

$$\min_{T \subseteq \Omega_{out}} \; |T| \quad s.t. \; (4.20) \text{ holds for some } X \succ 0. \tag{4.22}$$

Both problems are subject to the system stability constraint that for some $X \succ 0$,

$$\sigma {C_{cl}}^T C_{cl} + A_{cl}^T X + X A_{cl} + \sigma X B_{cl} B_{cl}^T X \prec 0. \tag{4.23}$$

Note that (4.23) contains the closed-loop system matrix $A_{cl}$ defined by Eq. (4.18) and hence is a function of $S$ and $T$.

Without additional structure, problems (4.21) and (4.22) require exhaustive search over all selections of $S$ and $T$.

In what follows, we present a sufficient condition on $S$ and $T$ to satisfy the stability constraint (4.23). The sufficient condition is established by constructing a solution $X$ to inequality (4.23) and mapping the requirement on $X$ to a constraint on the eigenvalues of $A_{cl}$. We show later that the sufficient constraints have bounded submodularity ratio and hence lead to computationally efficient input and output selection algorithms.

The following lemma constructs a solution $X$ to (4.23).

**Lemma 8.** *Suppose that ${C_{cl}}^T C_{cl} \preceq cI$ and $B_{cl}B_{cl}^T \preceq bI$. For any Hurwitz matrix $A_{cl}$, suppose that there exists $\epsilon > 0$ such that $X \prec \epsilon I$ where $X$ is defined by*

$$X = \int_0^\infty e^{A_{cl}^T t}(\sigma cI + \sigma b\epsilon^2 I)e^{A_{cl}t}\, dt. \tag{4.24}$$

*Then X is a solution to (4.23).*

*Proof.* Suppose ${C_{cl}}^T C_{cl} \preceq cI, \;\; B_{cl}B_{cl}^T \preceq bI$. Then the expression in (4.23) is bounded by

$$\sigma {C_{cl}}^T C_{cl} + A_{cl}^T X + XA_{cl} + \sigma XB_{cl}B_{cl}^T X \preceq \sigma cI + A_{cl}^T X + XA_{cl} + \sigma bX^2.$$

Thus, any positive definite matrix $X$ that satisfies

$$\sigma cI + A_{cl}^T X + XA_{cl} + \sigma bX^2 \prec 0 \tag{4.25}$$

is a solution to inequality (4.23).

In what follows, we prove that the $X$ constructed in (4.24) is a solution to (4.25).

First, we show that when $X \prec \epsilon I$, we have $X^2 \prec \epsilon^2 I$. For any $X \succ 0$, we have $\epsilon^2 I - X^2 = (\epsilon I - X)(\epsilon I + X)$, where $\epsilon I - X$ is positive definite by assumptio and $\epsilon I + X \succ X \succ 0$ is also positive definite. Both $(\epsilon I - X)$ and $(\epsilon I + X)$ are symmetric positive definite and commute. Hence, $(\epsilon I - X)(\epsilon I + X) = \epsilon^2 I - X^2$ is also positive definite, which implies $X^2 \prec \epsilon^2 I$.

Substituting (4.24) into (4.25) gives

$$\begin{aligned}
&\sigma cI + A_{cl}^T X + XA_{cl} + \sigma bX^2\\
=&\sigma cI + \sigma bX^2 + \int_0^\infty A_{cl}^T e^{A_{cl}^T t}(\sigma cI + \sigma b\epsilon^2 I)e^{A_{cl}t}\, dt + \int_0^\infty e^{A_{cl}^T t}(\sigma cI + \sigma b\epsilon^2 I)e^{A_{cl}t}A_{cl}\, dt\\
=&\sigma cI + \sigma bX^2 + \int_0^\infty \frac{d}{dt}\left(e^{A_{cl}^T t}(\sigma cI + \sigma b\epsilon^2 I)e^{A_{cl}t}\right)\, dt\\
=&\ \sigma cI + \sigma bX^2 + e^{A_{cl}^T t}(\sigma cI + \sigma b\epsilon^2 I)e^{A_{cl}t}\Big|_{t=0}^{\infty}\\
=&\sigma cI + \sigma bX^2 - (\sigma cI + \sigma b\epsilon^2 I)\\
=&\sigma b(X^2 - \epsilon^2 I) \prec 0,
\end{aligned}$$

which implies that the $X$ constructed in (4.24) is a solution to inequality (4.25) and hence is a solution to inequality (4.23). $\square$

By Lemma 8, in order to meet the system stability constraint (4.23), it is sufficient to have the matrix $A_{cl}$ satisfy the following: for some $\epsilon > 0$,

$$\int_0^\infty e^{A_{cl}^T t}(\sigma c I + \sigma b \epsilon^2 I) e^{A_{cl} t}\, dt \prec \epsilon I. \tag{4.26}$$

The following lemma simplifies the inequality (4.26), which results in a sufficient condition to constraint (4.23).

**Lemma 9.** *Let $\lambda_0$ denote the eigenvalue of $A_{cl}$ with the largest real part. Define $\alpha = \lambda_1(A_{cl} + A_{cl}^T)/(2Re(\lambda_0))$, where the notation $\lambda_1(\cdot)$ refers to the largest eigenvalue. Assume $A_{cl}$ is asymptotically stable, i.e., $Re(\lambda_0) < 0$. Then the inequality (4.26) holds if*

$$Re(\lambda_0) < -\sigma\sqrt{bc}/\alpha. \tag{4.27}$$

*Proof.* Given the fact that a symmetric matrix $(\epsilon I - X)$ is positive definite if and only if all its eigenvalues are positive, the inequality $X \prec \epsilon I$ in (4.26) holds if the largest eigenvalue of $X$ is smaller than $\epsilon$, i.e., $\lambda_1(X) < \epsilon$. The constructed $X$ in (4.26) is a solution of the Lyapunov equation

$$A_{cl}^T X + X A_{cl} + Q = 0, \text{ where } Q = (\sigma c + \sigma b \epsilon^2) I.$$

The eigenvalues of any solution $X$ are bounded by $\lambda_1(X) \leq -\lambda_1(Q)/\lambda_1(A_{cl} + A_{cl}^T)$ when $A_{cl} + A_{cl}^T \prec 0$ [38]. Substituting $\lambda_1(A_{cl} + A_{cl}^T)$ with $2\alpha\text{Re}(\lambda_0)$, we have

$$\lambda_1(X) \leq \frac{\sigma c + \sigma b \epsilon^2}{-2\alpha \text{Re}(\lambda_0)}.$$

In order to guarantee $\lambda_1(X) < \epsilon$, it suffices to require

$$\text{Re}(\lambda_0) < -\frac{\sigma}{2\alpha}(\frac{c}{\epsilon} + b\epsilon). \tag{4.28}$$

After finding the maximum value of the right-hand-side function in (4.28) respect to $\epsilon$, we have the inequality (4.28) holds for some $\epsilon > 0$ if $\text{Re}(\lambda_0) < -\sigma\sqrt{bc}/\alpha$. □

By Lemma 8 and 9, the constraint (4.23) for input and output selections is relaxed to a sufficient condition (4.27) on eigenvalues of $A_{cl}$, which still guarantees the design of a feedback control $K$ that can stabilize the system in the presence of uncertainty $\Delta$.

Given that $A_{cl}$ in (4.18) is a block upper triangular matrix with matrices $A - B_{_S}K$ and $A - LC_{_T}$ on the diagonal, the eigenvalue condition (4.27) is satisfied if both $A - B_{_S}K$ and $A - LC_{_T}$ have eigenvalues less than $-\sigma\sqrt{bc}/\alpha$. Thus, the sufficient condition guaranteeing system stability becomes

$$\text{Re}\left(\lambda(A - B_{_S}K)\right) < -\sigma\sqrt{bc}/\alpha \tag{4.29}$$

for input selection $S$, and

$$\text{Re}\left(\lambda(A - LC_{_T})\right) < -\sigma\sqrt{bc}/\alpha \tag{4.30}$$

for output selection $T$, where the notation $\lambda(\cdot)$ refers to all eigenvalues of a given matrix.

Hence, the input selection problem can be formulated as

$$\begin{aligned} \min_{S \subseteq \Omega_{in}} \quad & |S| \\ s.t. \quad & \text{Re}\left(\lambda(A - B_{_S}K)\right) < -\sigma\sqrt{bc}/\alpha \end{aligned} \tag{4.31}$$

and the output selection problem can be formulated as

$$\begin{aligned} \min_{T \subseteq \Omega_{out}} \quad & |T| \\ s.t. \quad & \text{Re}\left(\lambda(A - LC_{_T})\right) < -\sigma\sqrt{bc}/\alpha \end{aligned} \tag{4.32}$$

where ${C_{cl}}^T C_{cl} \preceq cI, \;\; B_{cl}B_{cl}^T \preceq bI$.

The input selection is to achieve the stabilizability of the system while the output selection is to achieve the system detectability, both in the presence of uncertainties. For a nominal system, i.e., $\Delta = 0$, our stability threshold $(-\sigma\sqrt{bc}/\alpha = 0)$ is consistent with the existing literature [6].

By the fact that $2\text{Re}(\lambda_0) \leq \lambda_1(A_{cl} + A_{cl}^T) < 0$ [80], we have that $\alpha$ is in the range $(0, 1]$. We take the following approach to approximating the solution to (4.31) and (4.32). We first set $\alpha = 1$, yielding the constraints

$$\text{Re}\left(\lambda(A - B_{_S}K)\right) < -\sigma\sqrt{bc}, \tag{4.33}$$

$$\text{Re}\left(\lambda(A - LC_{_T})\right) < -\sigma\sqrt{bc}, \tag{4.34}$$

and select a set of inputs and outputs using the procedure enumerating in the following sections. After synthesizing a controller and computing the matrices $A_{cl}$, $B_{cl}$, and $C_{cl}$, if the condition of Lemma 7 does not hold, then we reduce the value of $\alpha$ and reuse the procedure of the following sections, but with the constraints Re $(\lambda(A - B_S K)) < -\sigma\sqrt{bc}/\alpha$ and Re $(\lambda(A - LC_T)) < -\sigma\sqrt{bc}/\alpha$. We continue this procedure until an input and an output sets satisfying Lemma 7 are selected. In practice, during the simulations described in Section 4.5, the conditions (4.33) and (4.34) were sufficient to ensure system stability during all simulation trials. This is because the small gain theorem (Lemma 6) is known to be conservative and hence by picking $\alpha = 1$, the conditions (4.33) and (4.34) are in general sufficient to ensure stability.

### *4.3.2 Submodularity Ratio of Input Selection Constraint*

In this section, we establish a metric for input selection (4.31) and prove the metric has a bounded submodularity ratio.

First, we present a lemma, as follows, which shows a sufficient condition that guarantees the existence of a feedback control $K$ satisfying the input selection constraint (4.31), i.e., Re $(\lambda(A - B_S K)) < -\sigma\sqrt{bc}$.

**Lemma 10.** *Let*

$$\dot{x}(t) = Ax(t) + B_S u(t) \tag{4.35}$$

$$y(t) = Cx(t) \tag{4.36}$$

*be a fully observable control system. If all of the eigenvectors $v_i$ of $A$ with eigenvalues $\lambda_i$ satisfying Re$(\lambda_i) \geq \lambda_1$ lie in the span of the controllability matrix of the system, then there exists a feedback control $K$ such that all eigenvalues of the closed-loop system $A - B_S K$ satisfy Re$(\lambda(A - B_S K)) < \lambda_1$.*

*Proof.* The proof is a straightforward generalization of Theorem 8.3 of [6] and is included for completeness.

Let $\{q_1, \ldots, q_{n_1}\}$ be a maximal set of linearly independent columns of the controllability matrix $\mathcal{C}$, and define a matrix $P$ such that $P^{-1} = (q_1 \cdots q_{n_1} \cdots q_n)$, where $q_{n_1+1}, \ldots, q_n$ can be arbitrarily chosen so long as $P^{-1}$ is nonsingular. Under the transformation $z = Px$, the system can be written as $\dot{z} = \overline{A}z + \overline{B}u$ or more specifically

$$\begin{pmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} B_1 \\ 0 \end{pmatrix} u(t),$$

where the $n_1$-dimensional subsystem $\dot{z}_1 = A_{11}z_1 + B_1u$ is controllable (Chen [6, Theorem 6.6]).

Suppose first that $\lambda$ is an eigenvalue where the corresponding eigenvector $v$ lies in the span of the controllability matrix, i.e., $v \in \text{span}(\mathcal{C}) = \text{span}([B_S \; AB_S \; A^2B_S \; \ldots \; A^{n-1}B_S]$. We show that $\lambda$ is an eigenvalue of $A_{11}$, and hence all eigenvalues with Re$(\lambda) \geq \lambda_1$ are eigenvalues of $A_{11}$ by assumption.

Indeed, letting $w = Pv$, we have that

$$\overline{A}w = PAP^{-1}w = PAP^{-1}Pv = PAv = \lambda Pv = \lambda w,$$

establishing that $w$ is an eigenvector of $\overline{A}$ corresponding to $\lambda$. Furthermore, we have that $v = P^{-1}w$, and hence $w$ is the representation of $v$ with respect to the columns of $P^{-1}$. Since $v$ is in the span of $\mathcal{C}$, it also lies in the span of $\{q_1, \ldots, q_{n_1}\}$, and thus $w_i = 0$ for $i > n_1$. We then have $w = (w_1 \; 0)^T$, where $w_1$ is an eigenvector of $A_{11}$ corresponding to eigenvalue $\lambda$.

A similar argument to the above establishes that all eigenvalues corresponding to "undesired modes" (Re$(\lambda) \geq \lambda_1$) are eigenvalues of $A_{11}$, and that the system $(A_{11}, B_1)$ is controllable. Hence, we can design a feedback controller $K$ such that the system $A_{11} - B_1K$ has eigenvalues in the desired region (Re$(\lambda) < \lambda_1$) [6, Theorem 8.3], while the remaining eigenvalues are unchanged (Re$(\lambda) < \lambda_1$). Applying the transformation $P$ and using the fact that the eigenvalues of the system are invariant under a similarity transform yields the desired result. $\square$

By Lemma 10, it suffices to ensure the existence of a feedback control $K$ that satisfies constraint (4.31), if all eigenvectors $v_i$ of $A$ corresponding to "undesired modes" (Re$(\lambda_i) \geq -\sigma\sqrt{bc}$) lie in the span of the controllability matrix of the system $(A, B_S)$, i.e.,

$$\{v_i : \text{Re}(\lambda_i) \geq -\sigma\sqrt{bc}\} \in \text{span}\left(\mathcal{C}(S)\right), \tag{4.37}$$

where $\mathcal{C}(S) = [B_S \; AB_S \; A^2B_S \; \dots \; A^{n-1}B_S]$ denotes the controllability matrix of system $(A, B_S)$.

These "undesired modes" with eigenvalues satisfying $\text{Re}(\lambda_i) \geq -\sigma\sqrt{bc}$ refer to dynamics in the system (4.10)-(4.11) that are unstable in the presence of uncertainty $\Delta$ with $\|\Delta\|_\infty \leq \sigma$.

It is shown that for a linear system as defined in (4.35), the span of its controllability matrix equals to the span of its controllability Gramian [59], i.e.,

$$\text{span}(\mathcal{C}(S)) = \text{span}(W(S)), \tag{4.38}$$

where $W(S)$ is the controllability Gramian defined by

$$W(S) = \int_{t_0}^{t_1} e^{A(t-t_0)} B_S B_S^T e^{A^T(t-t_0)} dt \tag{4.39}$$

for some $t_1 > t_0$.

By (4.37) and (4.38), we construct a metric for the input selection in the presence of bounded uncertainty $\|\Delta\|_\infty \leq \sigma$, given by

$$f(S) \triangleq \sum_{i:\ Re(\lambda_i) \geq -\sigma\sqrt{bc}} dist^2(v_i, \text{span}(W(S))), \tag{4.40}$$

where $dist(\cdot)$ denotes the Euclidean distance; $\lambda_i$ and $v_i$ are eigenvalues and corresponding eigenvectors of $A$, respectively; and $b, c$ are constants of the assumption ${C_{cl}}^T C_{cl} \preceq cI, \;\; B_{cl} B_{cl}^T \preceq bI$.

Given any input selection $S$, the function $f(S)$ measures the distance between the span of the controllability Gramian and the eigenvectors of "undesired modes" where eigenvalues $\text{Re}(\lambda_i) \geq -\sigma\sqrt{bc}$. Intuitively, this metric is a measure of how close the unstable modes of the system (4.10)-(4.11) are to being controllable.

To ensure the existence of a feedback control $K$ that satisfies the input selection constraint (4.31), it suffices to require

$$f(S) = 0. \tag{4.41}$$

In the following, we prove that the function $f(S)$ has bounded submodularity ratio. Before giving the bound, we first define some notations as follows.

Let $\mathcal{C} = [B\ AB\ A^2B\ \dots\ A^{n-1}B]$ be the controllability matrix and let $P \in \mathbb{R}^{np\times np}$ be a nonsingular matrix that normalizes $\mathcal{C}$ to $\mathcal{C}' = \mathcal{C}P$ where each column of $\mathcal{C}'$ has norm 1. For simplicity, we use $\mathcal{C}_s$ to denote the function $\mathcal{C}(S)$ in the following and denote $\mathcal{C}'_s = \mathcal{C}_s P$.

Define $\bar{C} \in \mathbb{R}^{np\times np} = \mathcal{C}'^T\mathcal{C}'/n$ and $\bar{C}_s = \mathcal{C}_s'^T\mathcal{C}'_s/n$. We note that $\bar{C}_s$ is a submatrix[1] of $\bar{C}$ with rows and columns selected by set $S$. For any matrix $\bar{C}_s$, we denote its smallest eigenvalue as $\lambda_{\min}(\bar{C}_s)$ and let $\lambda_{\min}(\bar{C}, k) = \min_{S:|S|=k} \lambda_{\min}(\bar{C}_s)$.

The submodularity ratio of the input selection constraint (4.41) is described by the following theorem.

**Theorem 3.** *Let $\gamma_{U,k}$ be the submodularity ratio of the set function $f(S)$ defined in (4.40). Then for any set $U \subseteq \Omega_{in}$ and $k \geq 1$, the submodularity ratio $\gamma_{U,k}$ is bounded by*

$$\gamma_{U,k} \geq \lambda_{\min}(\bar{C}, k + |U|) \geq \lambda_{\min}(\bar{C}).$$

Before giving the proof, we first present some preliminary results define some notations as follows.

Given any vector $v$ with $\|v\|_2 = 1$, consider the function $f_v(S) = dist^2\left(v, \text{span}(\mathcal{C}(S))\right) = \min_z \|\mathcal{C}(S)z - v\|_2^2$.

**Lemma 11.** *For any nonsingular matrix $P$, the following equality holds:*

$$\min_z \|\mathcal{C}_s z - v\|_2^2 = \min_z \|\mathcal{C}_s P z - v\|_2^2.$$

*Proof.* Let $z'^* = \arg\min_z \|\mathcal{C}_s P z - v\|_2^2$. Then

$$z'^* = P^{-1}(\mathcal{C}_s^T\mathcal{C}_s)^{-1}P^{-T}P^T\mathcal{C}_s^T v = P^{-1}(\mathcal{C}_s^T\mathcal{C}_s)^{-1}\mathcal{C}_s^T v = P^{-1}z^*,$$

where $z^* = \arg\min_z \|\mathcal{C}_s z - v\|_2^2$. Thus,

$$\min_z \|\mathcal{C}_s P z - v\|_2^2 = \|\mathcal{C}_s P z'^* - v\|_2^2 = \|\mathcal{C}_s P P^{-1} z^* - v\|_2^2 = \|\mathcal{C}_s z^* - v\|_2^2 = \min_z \|\mathcal{C}_s z - v\|_2^2,$$

completing the proof. □

[1] $\bar{C}_s$ consists of $n \times |S|$ rows and columns selected from $\bar{C}$.

By Lemma 11, the function $f(S)$ is equivalent to

$$f_v(S) = \min_z \ \|\mathcal{C}'_s z - v\|_2^2. \tag{4.42}$$

Denote $z_s^*$ as the optimal vector that solves $\min_z \|\mathcal{C}'_s z - v\|_2^2$. The value of $z_s^*$ is given by $z_s^* = (\mathcal{C}'^T_s \mathcal{C}'_s)^{-1} \mathcal{C}'^T_s v$ [31]. Knowing that $v - \mathcal{C}'_s z_s^*$ and $\mathcal{C}'_s z_s^*$ are orthogonal, we have

$$f_v(S) = \|\mathcal{C}'_s z_s^* - v\|_2^2 = \|v\|_2^2 - \|\mathcal{C}'_s z_s^*\|_2^2 = 1 - \|\mathcal{C}'_s z_s^*\|_2^2.$$

Define a new set function

$$g_v(S) = \|\mathcal{C}'_s z_s^*\|_2^2 = \|P_s v\|_2^2, \tag{4.43}$$

where $P_s = \mathcal{C}'_s (\mathcal{C}'^T_s \mathcal{C}'_s)^{-1} \mathcal{C}'^T_s$ is the projection matrix for orthogonal projection onto the span of columns of $\mathcal{C}'_s$, and hence

$$g_v(S) = 1 - f_v(S). \tag{4.44}$$

**Lemma 12** ([34]). *The function $g_v(S)$ is the $R^2$ statistic (Definition 3) for the linear regression problem* $\min_z \|\mathcal{C}_s z - v\|_2^2$.

*Proof.* By definition,

$$g_v(S) = \|\mathcal{C}'_s (\mathcal{C}'^T_s \mathcal{C}'_s)^{-1} \mathcal{C}'^T_s v\|_2^2 = (\mathcal{C}'^T_s v)^T (\mathcal{C}'^T_s \mathcal{C}'_s)^{-1} (\mathcal{C}'^T_s v) = n \, \bar{v}_s^T \bar{C}_s^{-1} \bar{v}_s,$$

where $\bar{C}_s = \mathcal{C}'^T_s \mathcal{C}'_s / n$ and $\bar{v}_s = \mathcal{C}'^T_s v / n$, satisfying the definition of $R^2$ statistic for the linear regression (4.42) [13]. $\square$

**Lemma 13** ([13]). *Let $\gamma'_{U,k}$ be the submodularity ratio of the $R^2$ statistic $g_v(S)$. Then*

$$\gamma'_{U,k} \geq \lambda_{\min}(\bar{C}, k + |U|) \geq \lambda_{\min}(\bar{C}). \tag{4.45}$$

By Lemma 12 and 13, the function $g_v(S)$ has submodularity ratio $\gamma'_{U,k}$ bounded by (4.45). Using this result, now we are ready to give the proof of Theorem 3.

*Proof of Theorem 3:* We first show that $f_v(S)$ and $g_v(S)$ have the same submodularity ratio $\gamma'_{U,k}$. From (4.44), we have $g_v(S) = 1 - f_v(S)$. By definition,

$$\begin{aligned}
\gamma'_{U,k} &= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_{z\in S}\left(g_v(L\cup\{z\}) - g_v(L)\right)}{g_v(L\cup S) - g_v(L)} \\
&= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_{z\in S}\left((1 - f_v(L\cup\{z\})) - (1 - f_v(L))\right)}{(1 - f_v(L\cup S)) - (1 - f_v(L))} \\
&= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_{z\in S}\left(f_v(L\cup\{z\}) - f_v(L)\right)}{f_v(L\cup S) - f_v(L)}
\end{aligned}$$

Then we show that $f(S)$ has submodularity ratio $\gamma_{U,k}$ bounded by $\gamma'_{U,k}$. By Eq. (4.38) and Lemma 11, we have $f(S) = \sum_i f_{v_i}(S)$ where $v_i$ is the eigenvector of the $i$th "undesired mode" with eigenvalue (Re$(\lambda) \geq -\sigma\sqrt{bc}$). Then

$$\begin{aligned}
\gamma_{U,k} &= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_{z\in S}\left(f(L\cup\{z\}) - f(L)\right)}{f(L\cup S) - f(L)} \\
&= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_i \sum_{z\in S}\left(f_{v_i}(L\cup\{z\}) - f_{v_i}(L)\right)}{\sum_i \left(f_{v_i}(L\cup S) - f_{v_i}(L)\right)} \\
&\geq \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \left(\min_i \frac{\sum_{z\in S}\left(f_{v_i}(L\cup\{z\}) - f_{v_i}(L)\right)}{\left(f_{v_i}(L\cup S) - f_{v_i}(L)\right)}\right) \geq \gamma'_{U,k}
\end{aligned}$$

By Lemma 13, we have $\gamma_{U,k} \geq \gamma'_{U,k} \geq \lambda_{\min}(\bar{C}, k + |U|) \geq \lambda_{\min}(\bar{C})$, completing the proof. □

### *4.3.3 Submodularity Ratio of Output Selection Constraint*

In this section, we derive a metric for output selection (4.32) that has bounded submodularity ratio by analogy with input selection.

By condition (4.32), to achieve a detectable system in the presence of uncertainty $\Delta$ with $\|\Delta\|_\infty \leq \sigma$ requires that

$$\text{Re}\left(\lambda(A - LC_T)\right) < -\sigma\sqrt{bc},$$

or equivalently,

$$\text{Re}\left(\lambda(A^T - C_T^T L^T)\right) < -\sigma\sqrt{bc}, \tag{4.46}$$

where $A^T, C_{_T}^T, L^T$ are the transpose of matrices $A, C_{_T}, L$ respectively.

By duality theory, Lemma 10 implies that for the system $(A^T, C_{_T}^T)$, if all of the eigenvectors $v_i'$ of $A^T$ with eigenvalues $\lambda_i$ satisfying $\lambda_i \geq \lambda_0$ lie in the span of the matrix

$$M(T) = \int_{t_0}^{t_1} e^{A^T(t-t_0)} C_{_T}^T C_{_T} e^{A(t-t_0)} \, dt, \tag{4.47}$$

there exists a matrix $L$ such that all eigenvalues of the closed-loop system $A^T - C_{_T}^T L^T$ satisfy $\text{Re}(\lambda(A^T - C_{_T}^T L^T)) < \lambda_0$.

It is shown that the matrix $M(T)$ defined by (4.47) is the observability Gramian of the system $(A, C)$. Thus, a sufficient condition to ensure system detectability in the presence of uncertainty $\|\Delta\|_\infty \leq \sigma$ is that all eigenvectors $v_i'$ of $A^T$ with eigenvalues $\lambda_i \geq -\sigma\sqrt{bc}$ lie in the span of the observability Gramian $M(S)$ of the system $(A, C)$, i.e.,

$$\{v_i' : \text{Re}(\lambda_i) \geq -\sigma\sqrt{bc}\} \in \text{span}\left(M(T)\right). \tag{4.48}$$

We present a metric for output selection as follows, which measures how far the system is from being detectable in the presence of bounded uncertainty $\|\Delta\|_\infty \leq \sigma$,

$$g(T) \triangleq \sum_{i:\ Re(\lambda_i) \geq -\sigma\sqrt{bc}} dist^2(v_i', \text{span}(M(T))), \tag{4.49}$$

where $v_i'$ are eigenvectors of the matrix $A^T$.

For output selection, we have the main results, as follows, similar to the input selection. To ensure the existence of a matrix $L$ that satisfies the output selection constraint (4.32), it suffices to require

$$g(T) = 0. \tag{4.50}$$

The submodularity ratio of the constraint (4.50) is described by the following theorem.

**Theorem 4.** *Let $\eta_{U,k}$ be the submodularity ratio of the set function $g(T)$ defined in (4.49). Then for any set $U \subseteq \Omega$ and $k \geq 1$, the submodularity ratio $\eta_{U,k}$ is bounded by*

$$\eta_{U,k} \geq \lambda_{\min}(\bar{O}, k + |U|) \geq \lambda_{\min}(\bar{O}).$$

*where $\bar{O}$ and $\lambda_{\min}(\cdot)$ are defined as follows. Let*

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}^T \quad \textit{and} \quad \mathcal{O}(T) = \begin{bmatrix} C_T \\ C_T A \\ C_T A^2 \\ \vdots \\ C_T A^{n-1} \end{bmatrix}^T$$

*Let $D \in \mathbb{R}^{nm \times nm}$ be a nonsingular matrix that normalizes the observability matrix $\mathcal{O}$ to $\mathcal{O}' = \mathcal{O}D$ where each column of $\mathcal{O}'$ has norm 1, and hence $\mathcal{O}'(T) = \mathcal{O}(T)D$. We define $\bar{O} \in \mathbb{R}^{nm \times nm} = \mathcal{O}'^T\mathcal{O}'/n$ and $\bar{O}(T) = \mathcal{O}'(T)^T\mathcal{O}'(T)/n$. $\lambda_{\min}(\bar{O}(T))$ refers to the smallest eigenvalue of $\bar{O}(T)$ and $\lambda_{\min}(\bar{O}, k) = \min_{T:|T|=k} \lambda_{\min}(\bar{O}(T))$.*

*Proof.* The proof can be established directly by following the same steps as the proof of Theorem 3, while the controllability Gramian $W$ and the controllability matrix $\mathcal{C}$ are replaced by the observability Gramian $M$ and the observability matrix (transpose) $\mathcal{O}$. □

### *4.3.4 Minimal Input and Output Selection Algorithm*

Using the metrics we established in previous sections, the problem of selecting the minimal set of input nodes to ensure system stabilizability in the presence of uncertainties can be formulated as

$$\min_{S \subseteq \Omega_{in}} |S|, \quad s.t. \;\; f(S) = 0 \tag{4.51}$$

and the output selection for system detectability is given as

$$\min_{T \subseteq \Omega_{out}} |T|, \quad s.t. \;\; g(T) = 0 \tag{4.52}$$

For some control systems, the input selection and output selection are dependent to each other. For example, the wide-area damping system in a power grid selects generators as input to participate in the centralized control while observing rotor speeds at these selected generators as system outputs [83]. For such systems that input and output selections share the same ground set denoted by $\Omega_{joint}$, we propose a joint selection formulation

$$\min_{S\subseteq\Omega_{joint}} |S|, \quad s.t.\ f(S)+g(S)=0 \tag{4.53}$$

Note that $f+g$ in constraint (4.53) is the sum of two monotone decreasing set functions with submodularity ratios $\gamma_{U,k}$ and $\eta_{U,k}$, respectively. By definition, the submodularity ratio of $f+g$ is bounded by $\min(\gamma_{U,k}, \eta_{U,k})$. Since $\gamma_{U,k} \geq \lambda_{\min}(\bar{C}, k+|U|)$ and $\eta_{U,k} \geq \lambda_{\min}(\bar{O}, k+|U|)$, thus a lower bound of the submodularity ratio of $f+g$ is $\min(\lambda_{\min}(\bar{C}, k+|U|), \lambda_{\min}(\bar{O}, k+|U|))$. We also note that, since $f(S) \geq 0$ and $g(S) \geq 0$, the constraint (4.53) is satisfied if and only if $f(S)=0$ and $g(S)=0$.

The sufficient conditions (4.51)-(4.52) for selecting input and output nodes guarantee the existence of a stabilizing controller that can drive the system state to the origin in the presence of uncertainties. With selected inputs and outputs satisfying these conditions, one can find such a stabilizing controller using robust control design with pole placement constraints, such as the LMI-based $H_\infty$ synthesis proposed in [8].

We now present a polynomial-time algorithm, shown as Algorithm 3, that approximately solves the above input selection, output selection, and joint selection problems, with an optimality bound on the solution set cardinality.

For the input selection problem (4.51), we set $\Omega = \Omega_{in}$ and $F = f$. For the output selection problem (4.52), we set $\Omega = \Omega_{out}$ and $F = g$. For the joint selection problem (4.53), we set $\Omega = \Omega_{joint}$ and $F = f+g$.

The algorithm proceeds as follows. Given $\Omega$ and $F$, a set of selected inputs/outputs, $S$, is initialized to be empty. At each iteration, the algorithm searches an element $v$ from $\Omega/S$ that maximizes $F(S) - F(S\cup\{v\})$ and adds it to the selection set $S$. The algorithm terminates when $F(S)$ is reduced to 0 or $S=\Omega$.

The input and output sets selected by Algorithm 3 have cardinalities within a factor of the true optimal solution of the original problem (4.51), (4.52), or (4.53). The optimality bound of Algorithm 3 is defined by the following lemma.

**Algorithm 3:** Algorithm for selecting minimum-size set of inputs/outputs for a stabilizing controller.

**procedure** MINSET($\Omega$, $F$)

  **Input:** Ground set $\Omega$ of inputs/outputs, i.e., $\Omega = \Omega_{in}$ for input selection, $\Omega = \Omega_{out}$ for output selection, or $\Omega = \Omega_{joint}$ for joint selection;

  Supermodular metric $F$ in constraint, i.e., $F = f$ for input selection, $F = g$ for output selection, or $F = f + g$ for joint selection.

  **Output:** Set of inputs/outputs $S$ to participate in the system control.

  **Initialization:** $S \leftarrow \emptyset$

  **while** $F(S) > 0$ *and* $S \neq \Omega$ **do**

  **for** $v \in \Omega \setminus S$ **do**

  $\delta_v \leftarrow F(S) - F(S \cup \{v\})$

  $v^* \leftarrow \arg\max_v \delta_v$

  $S \leftarrow S \cup \{v^*\}$

  **return** $S$

**end procedure**

**Proposition 2.** *Let $S^*$ denote the true optimal solution to the input/output selection problem (4.51), (4.52), or (4.53) that has the form $\{\min_{S \in \Omega} |S| : F(S) = 0\}$ and let $S$ denote the solution returned by Algorithm 3. Then we have*

$$\frac{|S| - 1}{|S^*|} \leq \frac{1}{\gamma_0} \log \frac{F(\emptyset)}{F(S_{t-1})},$$

*where $S_{t-1}$ denotes the selected input set $S$ at the second-to-last iteration of Algorithm 3 and $\gamma_0$ is the lower bound of the submodularity ratio of $F$.*

*For input selection (4.51), $\gamma_0 = \lambda_{\min}(\bar{C}, 2|S|)$; For output selection (4.52), $\gamma_0 = \lambda_{\min}(\bar{O}, 2|S|)$; For joint selection (4.53), $\gamma_0 = \min(\lambda_{\min}(\bar{C}, 2|S|), \lambda_{\min}(\bar{O}, 2|S|))$.*

*Proof.* Suppose $F$ has submodularity ratio $\gamma_{U,k}$ for any given set $U$ and integer $k$. Let $\{s_1, \ldots, s_k\}$ denote the elements selected by the Algorithm 3 in the first $k$ iterations. By the definition of submodularity ratio, we have

$$\frac{\sum_{z\in S^*}\left(F(\emptyset)-F(\{z\})\right)}{F(\emptyset)-F(S^*)}\geq\gamma_{S,|S|}\geq\gamma_0 \tag{4.54}$$

The lower bound $\gamma_0$ depends on function $F$. When $F=f$, by Theorem 3, we have $\gamma_{S,|S|}\geq\lambda_{\min}(\bar{C},2|S|)$. When $F=g$, by Theorem 4, we have $\gamma_{S,|S|}\geq\lambda_{\min}(\bar{O},2|S|)$. When $F=f+g$, we have $\gamma_{S,|S|}\geq\min(\lambda_{\min}(\bar{C},2|S|),\lambda_{\min}(\bar{O},2|S|))$.

Inequality (4.54) implies that

$$|S^*|\left(F(\emptyset)-F(\{s_1\})\right)\geq\gamma_0\left(F(\emptyset)-F(S^*)\right),$$

or equivalently,

$$F(\{s_1\})-F(S^*)\leq\left(1-\frac{\gamma_0}{|S^*|}\right)\left(F(\emptyset)-F(S^*)\right).$$

Similarly, we have that

$$\frac{\sum_{z\in S^*}\left(F(s_1)-F(\{z,s_1\})\right)}{F(\{s_1\})-F(S^*\cup\{s_1\})}\geq\gamma_0$$

implying that

$$\begin{aligned}F(\{s_1\})-F(\{s_1,s_2\})&\geq\frac{\gamma_0}{|S^*|}(F(\{s_1\})-F(S^*\cup\{s_1\}))\\&\geq\frac{\gamma_0}{|S^*|}(F(\{s_1\})-F(S^*)),\end{aligned}$$

which is further equivalent to

$$\begin{aligned}F(\{s_1,s_2\})-F(S^*)&\leq\left(1-\frac{\gamma_0}{|S^*|}\right)\left(F(\{s_1\})-F(S^*)\right)\\&\leq\left(1-\frac{\gamma_0}{|S^*|}\right)^2\left(F(\{\emptyset\})-F(S^*)\right).\end{aligned}$$

By mathematical induction, we have that

$$F(\{s_1,\ldots,s_k\})-F(S^*)\leq\left(1-\frac{\gamma_0}{|S^*|}\right)^k(F(\{\emptyset\})-F(S^*)).$$

Note that $F(S^*)=0$ and hence

$$F(\{s_1,\ldots,s_k\})\leq\left(1-\frac{\gamma_0}{|S^*|}\right)^kF(\{\emptyset\}).$$

Use the fact that $1 - 1/z \leq \log(z), \forall z \geq 1$. Then when

$$k = \frac{|S^*|}{\gamma_0} \log \frac{F(\emptyset)}{F(S_{t-1})},$$

we have

$$F(\{s_1, \ldots, s_k\}) \leq F(S_{t-1}),$$

implying that one additional element is sufficient. Hence

$$\frac{|S| - 1}{|S^*|} \leq \frac{1}{\gamma_0} \log \frac{F(\emptyset)}{F(S_{t-1})}.$$

□

For a set $\Omega = \{1, \ldots, |\Omega|\}$, Algorithm 3 terminates in at most $|\Omega|$ iterations. For a system with $n$ states, each iteration solves at most $n$ least-squares problems where each problem requires computation of $O(n^3)$. Thus, the overall computational complexity of Algorithm 3 is $O(|\Omega| n^4)$.

We note that the optimality bound in Proposition 2 can be arbitrarily small for certain system matrices. This is consistent with the claim that minimal reachability is hard to approximate as presented in [30].

### *4.4 Analysis for Different Uncertainties*

In this section, we analyze our proposed input and output selection approach for different uncertainty cases. For additive and multiplicative uncertainties in system matrix $A$, we show a direct transformation to the $M - \Delta$ loop model and give the specific values of the thresholds in metrics $f$ and $g$. For uncertain time delay in output $y$, we first incorporate the delayed dynamics into the original system and then transform the modified linear system to the $M - \Delta$ model.

We also study the structured uncertainty in matrix $A$. From a known robustness criterion, we derive a sufficient condition on eigenvalues of the closed-loop system matrix $A_{cl}$, such that metrics similar to (4.40) and (4.49) that have bounded submodularity ratios can be established for input and output selection problems.

### *4.4.1 Additive Uncertainty in Matrix $A$*

A linear system with additive perturbation $\Delta$ in matrix $A$ has dynamics

$$\dot{x}(t) = (A + \Delta)x(t) + Bu(t)$$
$$y(t) = Cx(t).$$

When matching this system to our model (4.8)-(4.9), the uncertain term $\delta(t) = \Delta x(t)$.

In the equivalent $M - \Delta$ loop representation (4.15)-(4.16), we have $B_{cl} = \begin{bmatrix} I & I \end{bmatrix}^T$ and $C_{cl} = \begin{bmatrix} I & 0 \end{bmatrix}$.

By (4.27), we have the stability condition Re$\left(\lambda(A_{cl})\right) < -\sigma\sqrt{bc}$, assuming $B_{cl}B_{cl}^T \preceq bI$, $C_{cl}^T C_{cl} \preceq cI$ and $\|\Delta\|_\infty \leq \sigma$. In this case, we get $b = 2$ and $c = 1$ by the fact that

$$B_{cl}B_{cl}^T = \begin{bmatrix} I & I \\ I & I \end{bmatrix} \preceq 2 \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \; C_{cl}^T C_{cl} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

Therefore, for additive uncertainties in matrix $A$, the metric $f(S)$ defined by (4.40) for input selection guaranteeing system stabilizability becomes

$$f(S) = \sum_{i:\ Re(\lambda_i) \geq -\sigma\sqrt{2}} dist^2(v_i, \text{span}(W(S))), \tag{4.55}$$

and the metric $g(T)$ for output selection guaranteeing system detectability is given by

$$g(T) = \sum_{i:\ Re(\lambda_i) \geq -\sigma\sqrt{2}} dist^2(v_i', \text{span}(M(T))). \tag{4.56}$$

These changes in metrics $f$ and $g$ therefore impact the Algorithm 3 by specifying the parameters $b$ and $c$ in the eigenvalue threshold $(-\sigma\sqrt{bc})$.

### *4.4.2 Multiplicative Uncertainty in Matrix $A$*

For a multiplicative perturbation in matrix $A$, it changes the system dynamics to

$$\dot{x}(t) = (I + \Delta)Ax(t) + Bu(t)$$
$$y(t) = Cx(t),$$

where $\Delta$ is assumed to be bounded by $\|\Delta\|_\infty \leq \sigma$. In terms of our model (4.8)-(4.9), the uncertain term $\delta(t) = \Delta A x(t)$.

In this case, the equivalent $M - \Delta$ system has matrices $B_{cl} = \begin{bmatrix} I & I \end{bmatrix}^T$ and $C_{cl} = \begin{bmatrix} A & 0 \end{bmatrix}$. By the same fact that $B_{cl} B_{cl}^T \preceq 2I$ as shown in previous case, we have $b = 2$.

Let $\rho$ denote the largest singular value of $A$, i.e., $\|A\|_\infty = \rho$. Then the following inequality holds,

$$C_{cl}^T C_{cl} = \begin{bmatrix} A^T A & 0 \\ 0 & 0 \end{bmatrix} \preceq \rho^2 \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

which implies that $c = \rho^2$.

Hence, the metric $f(S)$ for input selection in the presence of a multiplicative uncertainty in matrix $A$ is given by

$$f(S) = \sum_{i:\ Re(\lambda_i) \geq -\sigma\rho\sqrt{2}} dist^2(v_i, \text{span}(W(S))), \tag{4.57}$$

and the output selection metric $g(T)$ becomes

$$g(T) = \sum_{i:\ Re(\lambda_i) \geq -\sigma\rho\sqrt{2}} dist^2(v_i', \text{span}(M(T))). \tag{4.58}$$

In this case, the Algorithm 3 selects inputs/outputs based on a different threshold $(-\sigma\rho\sqrt{2})$ that relays on the largest singular value of $A$ compared to the additive uncertainty case.

### *4.4.3 Uncertain Time Delay in Output $y$*

For a control system with an uncertain time delay $\tau$ in output $y$, the delayed output is denoted by $y_d(t) = y(t - \tau)$. In frequency domain, the output delay can be represented by a first-order Padé approximation [81],

$$y_d = e^{-s\tau} y \approx \frac{-0.5s\tau + 1}{0.5s\tau + 1} y. \tag{4.59}$$

A state-space realization of the transfer function (4.59) is given by

$$\dot{x}_d(t) = \frac{2}{\tau}(-x_d(t) + 2y(t)) \tag{4.60}$$

$$y_d(t) = x_d(t) - y(t). \tag{4.61}$$

Consider a system with $m$ outputs $\{y_1, \ldots, y_m\}$. When each output $y_i$ has a different delay $\tau_i$, we define a diagonal matrix of delays by $\Gamma = diag(\tau_1, \ldots, \tau_m)$. By involving $m$ such delayed dynamics (4.60)-(4.61) into the original system ($\dot{x} = Ax + Bu,\ y = Cx$), the overall system dynamics are given by

$$\begin{bmatrix} \dot{x}(t) \\ \dot{x}_d(t) \end{bmatrix} = \begin{bmatrix} A & 0 \\ 4\Gamma^{-1}C & -2\Gamma^{-1} \end{bmatrix} \begin{bmatrix} x(t) \\ x_d(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) \tag{4.62}$$

$$y_d(t) = \begin{bmatrix} -C & I \end{bmatrix} \begin{bmatrix} x(t) \\ x_d(t) \end{bmatrix}, \tag{4.63}$$

where $x_d \in \mathbb{R}^m$ is the state vector of delayed dynamics; $y_d \in \mathbb{R}^m$ is the vector of delayed outputs.

Suppose that each time delay $\tau_i$ is varying within a certain known interval such that $\Gamma^{-1} = \Gamma_0^{-1} + \Delta$, where $\Gamma_0^{-1}$ is the nominal values of delays and $\Delta$ is the uncertain part of delays bounded by $\|\Delta\|_\infty \leq \sigma$. By defining a new state vector $\tilde{x} = [x, x_d]^T$, then the system (4.62)-(4.63) can be mapped to our system model (4.8)-(4.9),

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t) + \delta(t) \tag{4.64}$$

$$y_d(t) = \tilde{C}\tilde{x}(t), \tag{4.65}$$

where $\tilde{A} = \begin{bmatrix} A & 0 \\ 4\Gamma_0^{-1}C & -2\Gamma_0^{-1} \end{bmatrix}$, $\tilde{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}$, $\tilde{C} = \begin{bmatrix} -C & I \end{bmatrix}$, and $\delta(t) = \begin{bmatrix} 0 & 0 \\ 4\Delta C & -2\Delta \end{bmatrix} \tilde{x}(t)$.

In the $M - \Delta$ loop representation of the system (4.64)-(4.65), the system matrices are given by

$$A_{cl} = \begin{bmatrix} \tilde{A} - \tilde{B}K & \tilde{B}K \\ 0 & \tilde{A} - L\tilde{C} \end{bmatrix},\ B_{cl} = \begin{bmatrix} 0 & 4I & 0 & 4I \end{bmatrix}^T,\ C_{cl} = \begin{bmatrix} I & -\frac{1}{2}I & 0 & 0 \end{bmatrix}.$$

By exploring the largest eigenvalues of $B_{cl}$ and $C_{cl}$, we have tight bounds on $B_{cl}B_{cl}^T \leq 32I$ and $C_{cl}^T C_{cl} \leq \frac{5}{4}I$. Therefore, the parameters of the input and output selection metrics are given by $b = 32$ and $c = 5/4$. Hence, the input selection metric for system with uncertain time delays is

$$f(S) = \sum_{i:\ Re(\lambda_i) \geq -2\sigma\sqrt{10}} dist^2(v_i, \text{span}(W(S))), \tag{4.66}$$

and the output selection metric is

$$g(T) = \sum_{i:\ Re(\lambda_i) \geq -2\sigma\sqrt{10}} dist^2(v_i', \text{span}(M(T))). \tag{4.67}$$

For the uncertain time delay, Algorithm 3 uses a much higher eigenvalue threshold ($-2\sigma\sqrt{10}$) compared to the additive uncertainty in $A$ ($-\sigma\sqrt{2}$).

### *4.4.4 Structured Uncertainty in Matrix $A$*

In previous analysis, it has been assumed that the each entry of the system matrix is perturbed independently. In some problems (e.g., constant output feedback with uncertainty in the gain matrix), there are only a small number of uncertain parameters, but each uncertain parameter affects multiple entries of the system matrix [87]. In what follows, we assume that the uncertainty matrix $\Delta$ has the following structure:

$$\Delta = \sum_{i=1}^{N} \delta_i E_i,$$

where $E_i$ are constant matrices and $\delta_i$ are uncertain parameters varying in the interval $[-\epsilon_i, \epsilon_i]$ for each $i$. When such a structured uncertainty $\Delta$ is added to the matrix $A$, the system has the dynamics

$$\begin{aligned} \dot{x}(t) &= (A + \Delta)x(t) + Bu(t) \\ y(t) &= Cx(t). \end{aligned}$$

By involving the state estimation $\hat{x}$ and feedback control $u = -K\hat{x}$, the closed-loop dynamics (4.14) of the above system have the form

$$\dot{\overline{x}} = A_{cl}\overline{x} + \left(\sum_{i=1}^{N} \delta_i \overline{E}_i\right) \overline{x}(t), \quad \text{where } \overline{E}_i = \begin{bmatrix} E_i & 0 \\ E_i & 0 \end{bmatrix}. \tag{4.68}$$

A sufficient condition guaranteeing the stability of the system (4.68) is given by the following lemma.

**Lemma 14** ([87])**.** *Define* $P_i = (\overline{E}_i^T P + P\overline{E}_i)/2$, *where* $P$ *is the unique solution of the Lyapunov equation*

$$PA_{cl} + A_{cl}^T P + 2I = 0. \tag{4.69}$$

*The closed-loop system (4.68) is asymptotically stable if the following condition holds:*

$$|k_j| < 1/\|\sum_{i=1}^{N} P_i\|_\infty \textit{ for all } j. \tag{4.70}$$

The following lemma presents a sufficient condition on the solution $P$ of the Lyapunov equation (4.69) to satisfy the stability condition (4.70).

**Lemma 15.** *Let* $\sigma$ *be the largest singular value of* $\left(\sum_{i=1}^{N} \overline{E}_i\right)$. *Assume* $k_i \in [-\epsilon_i, \epsilon_i]$ *for all* $i$. *Then the Lyapunov equation (4.69) has solution* $P$ *satisfying the condition (4.70) if*

$$P \prec \frac{1}{\sigma\epsilon} I \tag{4.71}$$

*where* $\epsilon = \max_i\{\epsilon_i\}$.

*Proof.* Given $k_i \in [-\epsilon_i, \epsilon_i]$, we have $|k_i| \leq \epsilon_i$. Define $\epsilon = \max_i\{\epsilon_i\}$. Then the condition (4.70) is satisfied if

$$\|\sum_{i=1}^{N} P_i\|_\infty < \frac{1}{\epsilon}. \tag{4.72}$$

Since the matrix $P$ constructing $P_i$ is the solution of the Lyapunov equation (4.69), $P$ is symmetric positive definite and hence

$$\|\sum_{i=1}^{N} P_i\|_\infty = \|\sum_{i=1}^{N}(\overline{E}_i^T P + P\overline{E}_i)/2\|_\infty \leq \frac{1}{2}\|\sum_{i=1}^{N} \overline{E}_i^T P\|_\infty + \frac{1}{2}\|\sum_{i=1}^{N} P\overline{E}_i\|_\infty \leq \|\sum_{i=1}^{N} \overline{E}_i\|_\infty \|P\|_\infty.$$

When the largest singular value of $\left(\sum_i^N \overline{E}_i\right)$ is denoted by $\sigma$, the inequality (4.72) holds if

$$\|P\|_\infty < \frac{1}{\sigma\epsilon}.$$

For symmetric positive definite matrix $P$, it is equivalent to requiring

$$P \prec \frac{1}{\sigma\epsilon} I,$$

which completes the proof. $\square$

By Lemma 14 and 15, if $P \prec (1/\sigma\epsilon)I$, then the system (4.68) is stable. Assuming $A_{cl}$ is Hurwitz, a positive definite solution $P$ to the Lyapunov equation (4.69) is given by [6],

$$P = \int_0^\infty e^{A_{cl}^T t}(2I)e^{A_{cl}t}\, dt.$$

Let $\lambda(A_{cl})$ be the largest eigenvalue of $A_{cl}$. Then the sufficient condition to guarantee system stability in the presence of structured uncertainty $\Delta$ becomes

$$\int_0^\infty e^{\lambda(A_{cl})t}(2)e^{\lambda(A_{cl})t}\, dt < \frac{1}{\epsilon\sigma}.$$

Simplifying the above expression, we obtain

$$\lambda(A_{cl}) < -\epsilon\sigma \tag{4.73}$$

where $\epsilon = \max_i\{\epsilon_i\}$ and $\sigma = \| \sum_{i=1}^N \overline{E}_i\|_\infty$.

The condition (4.73) on eigenvalues of $A_{cl}$ is similar to the condition (4.27). Following the rest steps of our approach, we have the metrics

$$f(S) = \sum_{i:\ Re(\lambda_i)\geq -\epsilon\sigma} dist^2(v_i, \text{span}(W(S))) \tag{4.74}$$

for the input selection and

$$g(T) = \sum_{i:\ Re(\lambda_i)\geq -\epsilon\sigma} dist^2(v_i', \text{span}(M(T))) \tag{4.75}$$

for the output selection.

Since the analysis for structured uncertainty does not depend on the $M - \Delta$ transformation, the eigenvalue threshold $(-\epsilon\sigma)$ that is used in the Algorithm 3 has a different structure compared to previous cases $(-\sigma\sqrt{bc})$.

### *4.5 Numerical Study*

This section presents numerical results of our proposed input and output selection approach on the IEEE 39-bus power system. We adopt the Linear Quadratic Regulator (LQR) for controller design and study the system stability in the presence of uncertainties of different types and magnitudes. We illustrate the optimality of our method by comparing with the existing geometric indices based selection strategy [28].

### *4.5.1 System Setup*

We consider the wide-area damping control in power systems to ensure small signal stability [83]. Given an unstable initial operating point, the goal of the controller is to damp inter-area oscillations between generators and stabilize rotor speeds at each generator. Each generator's Power System Stabilizers (PSS) is an input node and the control signals are to adjust excitation voltages of each generator, which eventually impact the power output. The control outputs consist of measurable parameters throughout the system, including power outputs of generators and power flows over transmission lines. The test system in this numerical study has 10 input nodes and 56 output nodes.

The initial power system, including system topology and generator configurations, is obtained from the IEEE 39-bus power system data [48]. We create initial operating points with different unstable modes by varying the overall load level and gain parameters of PSS at each generator.

Linearizing the system dynamics at any initial operating point gives a linear approximation of the control system with matrices $A, B, C$ in the state-space form (4.8)-(4.9). In this control system, uncertainties may arise due to load changing, unexpected transmission line tripping, and time delays.

The problem we studied is to select a minimal set of generators (inputs) and observations (outputs) to be involved in the wide-area damping control, which ensures the existence of a stabilizing controller that can damp all generator rotor oscillations. We test our proposed selection algorithm and present simulation results in two subsections. In the first, we demonstrate the minimal input and output selections for nominal systems and compare our proposed selection approach with a method based on geometric indices. In the second subsection, we show how different types and magnitudes of uncertainties affect the minimal selections. We also show that the proposed selection approach guarantees the system stability when uncertainties are within the assumed bound, i.e., $\|\Delta\|_\infty \leq \sigma$.

The geometric indices based input/output selection approach is implemented as follows. The algorithm first computes the geometric measures of controllability $m_{ci}(k)$ and observability $m_{oi}(k)$

for each unstable mode $k$ and each input/output $i$ as follows [28],

$$m_{ci}(k) = \frac{b_i^T \psi_k}{\|\psi_k\| \|b_i\|}, \quad m_{oi}(k) = \frac{c_i \phi_k}{\|\phi_k\| \|c_i\|},$$

where $b_i$ is the $i$th column of matrix $B$ and $c_i$ is the $j$th row of matrix $C$; $\psi_k$ and $\phi_k$ are left and right eigenvectors of matrix $A$ respectively corresponding to the eigenvalue $\lambda_k$. Then the algorithm iteratively selects inputs and outputs in a greedy manner. At each iteration, the input and output with the highest controllability and observability indices are selected and added to the sets $S$ and $T$, respectively. The algorithm terminates when both of the stabilizability and detectability constraints are satisfied, i.e., $f(S) = 0$ and $g(T) = 0$.

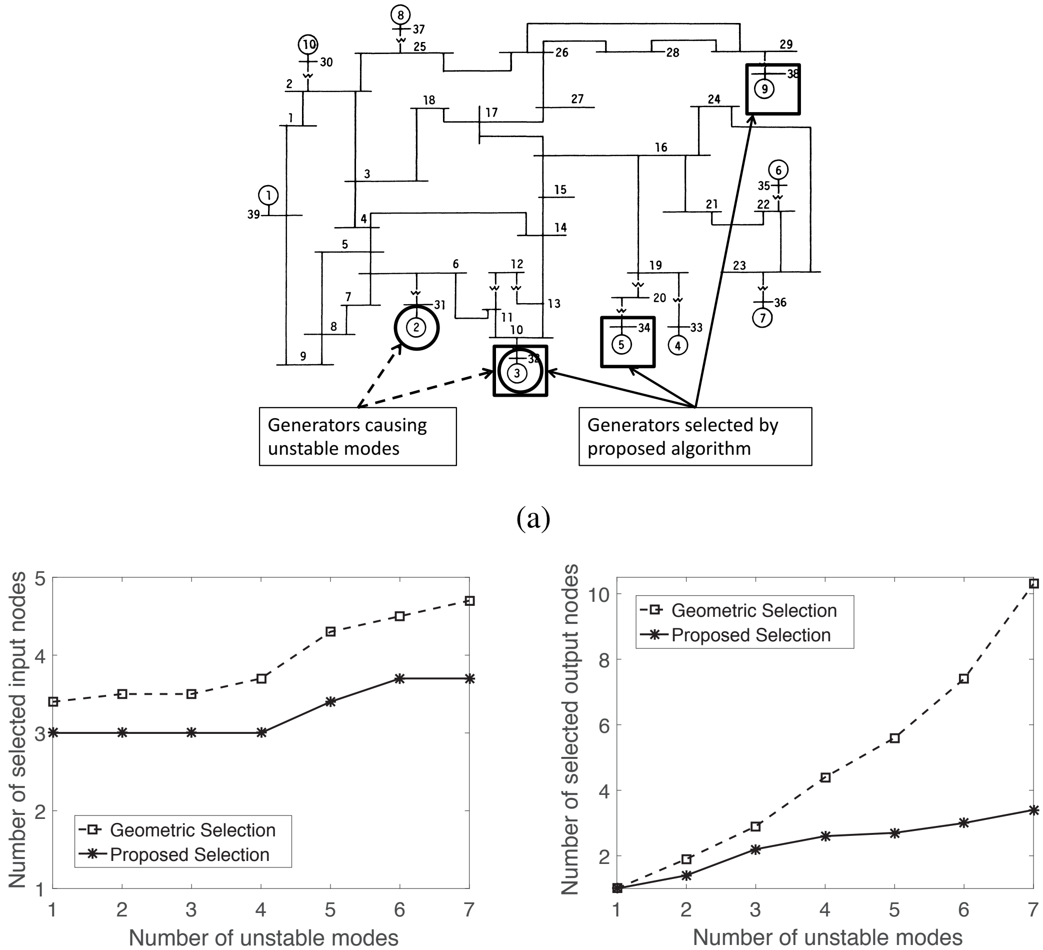
### *4.5.2 Input and Output Selection for Nominal Systems*

In this test case, we assume that there is no perturbation in the power system, $\sigma = 0$. We test our proposed selection algorithm for scenarios when there are different numbers of unstable modes. Unstable modes of a power system are typically caused by load shifting and improper functioning of power system stabilizers. We generate operating points with up to seven unstable modes by varying the load level randomly between $80\% - 150\%$ of the initial value, as well as scaling the gain parameters of all power system stabilizers to negative 116.2%.

Fig. 4.2(a) shows the system topology considered in this study, where one unstable mode occurs due to uncoordinated behaviors at generator 2 and 3. Our proposed algorithm selects generators $\{3, 5, 9\}$ as control input nodes to ensure system stabilizability.

A comparison between the proposed input/output selection and the geometric indices based selection is shown in Fig. 4.2(b) and 4.2(c). Each data point is the average number of selections at ten random operating points with the same number of unstable modes.

We observed that the average number of input and output nodes required by wide-area damping control increases as the number of unstable modes increases. We also found that the proposed selection algorithm requires fewer inputs and outputs for designing a stabilizing controller.




(a)

(b) (c)

Figure 4.2: Simulation study of our proposed approach to minimal input and output selection for power system stability. (a) IEEE 10-generator 39-bus power system topology considered in this study. Circled generators have initial rotor angles differing from the rest, causing unstable system modes. Generators selected by our proposed algorithm are indicated by squares. (b) Comparison of the average numbers of controlled generators (input nodes) required to meet the stabilizability condition ($f(S) = 0$) using the proposed approach and the method based on geometric indices. (c) Comparison of average numbers of observations (output nodes) for the proposed approach and the geometric indices based method to achieve the detectability condition ($g(T) = 0$).

### *4.5.3 Input and Output Selection for Uncertain Systems*

In this test case, we study how uncertainties impact our selection decisions at a fixed operating point, and evaluate the system robustness within our proposed selection framework. The initial operating point is obtained by using the default load level but scaling up the gains of all power system stabilizers by 16.2% while two gains are changed to negative.

Tables 4.1 and 4.2 demonstrate the input and output selections for different types of uncertainties at the same operating point. The metrics established in Section 4.4 are used in our proposed algorithm for different uncertainty cases. For structured uncertainty, we assume that the uncertain parameters $k_i$ are bounded by $k_i \in [-1, 1]$ and hence $\epsilon = \max_i\{|k_i|\} = 1$. We compare selections for different uncertainties as the largest singular value $\sigma$ increases from 0 to 0.35. We observed that the number of selections required to satisfy the stabilizability/detectability conditions increases as the uncertainty magnitude increases. It has also been observed that the number of selections for multiplicative uncertainty increases faster than other uncertainty cases, which indicates that the proposed selection approach is more sensitive to the multiplicative uncertainty. This is because that, comparing the thresholds $\text{Re}(\lambda) \leq -\sigma\sqrt{bc}$ for different uncertainties, the multiplicative uncertainty requires a much larger number of "desired modes" (proportional to $\sigma$) to lie in the controllability/observability subspaces.

In what follows, results on evaluating the system robustness to uncertainties are presented. For any set of input and output nodes, we consider the following LQR design of a feedback control $u = -Kx$, which is to minimize the cost function defined by

$$J(u) = \int_0^\infty (2\, x^T x + u^T u)\, dt.$$

To evaluate if a set of selections can guarantee system stability, we randomly generate 1000 uncertain matrices $\Delta$, with the same largest singular value $\sigma$, and add to matrix $A$. For each case, we apply the same LQR controller designed from the nominal system parameters. We measure the robustness of this control system by the percentage of cases that is stable out of these 1000 random trials.

Selections by Algorithm 3 with threshold $\sigma = 0$ are used as a baseline in comparison to the

Table 4.1: Comparing inputs selected by the proposed algorithm under different uncertainties.

| Uncertainty bound, $\sigma$ | 0 | 0.02 | 0.04 | 0.06 | 0.35 |
|---|---|---|---|---|---|
| Outputs for additive uncertainty | {4, 5, 10} | {3, 5, 9} | {3, 5, 9} | {3, 5, 9} | {2, 3, 5, 9} |
| Outputs for multiplicative uncertainty | {4, 5, 10} | {2, 3, 5, 9} | {3, 7, 8, 10} | {3, 7, 8, 10} | {3, 7, 8, 10} |
| Outputs for uncertain delays | {4, 5, 10} | {3, 5, 9} | {3, 5, 9} | {2, 3, 5, 9} | {3, 7, 8, 10} |
| Outputs for structured uncertainty | {4, 5, 10} | {3, 5, 9} | {3, 5, 9} | {3, 5, 9} | {2, 3, 5, 9} |

Table 4.2: Comparing outputs selected by the proposed algorithm under different uncertainties.

| Uncertainty bound, $\sigma$ | 0 | 0.02 | 0.04 | 0.06 | 0.35 |
|---|---|---|---|---|---|
| Outputs for additive uncertainty | {26, 35} | {25, 38} | {22, 28, 45} | {22, 28, 45} | {2, 19, 43} |
| Outputs for multiplicative uncertainty | {26, 35} | {2, 19, 43} | {13, 21, 46} | {13, 21, 46} | {13, 21, 46} |
| Outputs for uncertain delays | {26, 35} | {19, 20, 54} | {19, 20, 54} | {2, 19, 43} | {13, 21, 46} |
| Outputs for structured uncertainty | {26, 35} | {25, 38} | {25, 38} | {22, 28, 45} | {2, 19, 43} |

proposed selection approach that guarantees stability. When $\sigma = 0$, the selected inputs ensure stability of the nominal system but do not guarantee robustness to any uncertainties, while $\sigma \neq 0$ results in a selection for stabilizing control assuming $\|\Delta\|_\infty \leq \sigma$. At this operating point, the Algorithm 3 with threshold $\sigma = 0$ selects generators $\{4, 5, 10\}$ as inputs and observations $\{26, 35\}$ as outputs. For additive uncertainties in matrix $A$ assuming largest singular values less than $\sigma = 40$, our proposed algorithm selects generators $\{3, 7, 8, 10\}$ as input nodes and observations $\{13, 21, 46\}$ as outputs.

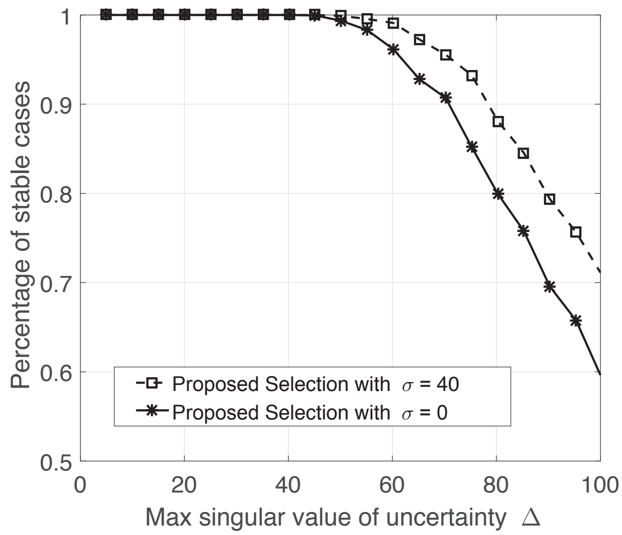



Figure 4.3: Comparison of percentages of cases that the controlled system is stable in the presence of additive uncertainty $\Delta$ using the proposed selection with $\sigma = 0$ and the selection with $\sigma = 40$. Each data point has 1000 random trials in $\Delta$. The proposed algorithm with robustness consideration ($\sigma = 40$) is stable in a higher percentage of cases compared to the selection without considering uncertainties ($\sigma = 0$).

Fig. 4.3 shows a comparison of system robustness between the two systems with input/output nodes selected by $\sigma = 0$ and $\sigma = 40$. Each data point is the percentage of stable cases in 1000 trials of random additive uncertainty $\Delta$ with the same largest singular value. We observed that when uncertainties have magnitudes within the assumption, $\|\Delta\|_\infty \leq 40$, then the controller based on the selections for $\sigma = 40$ ensures system stability in all cases. When uncertainties have largest singular values over the bound, for example $\|\Delta\|_\infty = 53 > 40$, both selections fail to guarantee stability. Our approach with robust consideration $\sigma = 40$, however, achieves a system with more stable cases (93.25%) compared to the selection based on $\sigma = 0$ (85.20%).

### 4.6 Conclusion

In this chapter, we studied the problem of selecting minimal inputs and outputs to ensure small signal stability. The problem was generalized to minimal input and output selection to ensure the existence of a stabilizing controller for any uncertain linear system. We derived sufficient conditions for input selection and output selection that guarantee system stabilizability and detectability respectively in the presence of uncertainties, requiring that a subset of system modes lie in the controllability/observability subspace. We formulated the problem of selecting minimal inputs/outputs to ensure that the distance from these desired modes to the span of the controllability/observability Gramian is zero. We showed that both input and output selection problems have bounded submodularity ratios. We developed computationally efficient algorithms with provable optimality bounds for minimal input selection, minimal output selection, and joint input and output selections, all ensuring the system stability in the presence of uncertainties. We also analyzed four types of uncertainties under our framework, including additive system uncertainty, multiplicative uncertainty, output time delay, and structured uncertainty. Our approach was validated through numerical study on the IEEE 39-bus power system with comparison to existing input/output selection methods.

# Chapter 5

# CONTROLLED ISLANDING TO MITIGATE CASCADING FAILURE

## *5.1 Introduction*

Controlled islanding has been studied extensively especially since the 2003 North American blackout that was caused by cascading failure [77, 16, 75, 61, 74, 64]. One important branch of study is slow coherency-based islanding, which was first introduced in [77]. This approach first classifies generators into groups with minimum non-coherency, and then finds a best cutset of transmission lines that separate generator groups into different islands while minimizing the overall load-generation imbalance through exhaustive search. This method, however, is not scalable to large systems.

Minimizing the load-generation imbalance is typically NP-hard and hence is challenging [77]. Using the minimal power-flow disruption as the objective function, instead of the minimal power imbalance, can reduce the time complexity of the problem and allow computationally efficient algorithms to be developed. Such an islanding scheme is presented in [74], in which the step of searching for minimum-load-generation islands is approximated by finding minimal power-flow disruption between islands, and hence the problem is mapped to a classic graph min-cut problem. This allows the use of existing graph theory based tools, but the solution optimality is not guaranteed. In [75], the slow coherency-based generator grouping and min-cut approach to islanding have been implemented in a simulation study of the 2003 blackout scenario.

Other approaches to controlled islanding include *Ordered Binary Decision Diagram (OBDD)* methods [64] and machine learning based methods [61]. In [64], a large power system is first simplified into a topology with less than 30 nodes and then a fast brute-force search is used to find feasible cutsets. The topology simplification is heavily based on system operation heuristics, such as integrating important transmission lines that should not be cut. As a consequence, the

OBDD approach cannot provide any guarantee on the optimality of solutions. Off-line machine learning based approaches have been investigated in [61]. Given a large set of off-line simulation data, a probability model can be established to assist evaluation of online islanding strategies. The optimality of such data-driven methods, however, is not guaranteed.

A two-step spectral clustering based control islanding is proposed in [16]. In the first step, the generators that have strong dynamic coupling are grouped, by applying the normalized spectral clustering to the system eigenbasis. In the second step, other buses are grouped into islands by constrained spectral clustering with the goal of minimizing power-flow disruption on cutting edges. Although this method is computationally efficient, the gap between the solution and the optimal strategy is not guaranteed. On the other hand, such two-step based methods lack the flexibility in making trade-off between generator coherency and load-generation imbalance, defined as the ability to sacrifice a small amount of generator coherency within each island in order to gain a large reduction in overall power imbalance.

In this chapter, we propose a one-step approach to controlled islanding that jointly minimizes the generator non-coherency and power imbalance when selecting an islanding strategy, with provable optimality guarantees. Our key insight is that selecting a subset of transmission lines to trip in order to form desired partition is inherently a combinatorial problem with weak submodularity property. Weak submodularity is a weakened form of diminishing property of set functions, which is sufficient to guarantee the optimality of a greedy selection algorithm with a relaxed approximation factor [34]. We make the following specific contributions:

- We formulate the problem of selecting a set of transmission lines to trip in order to minimize the weighted sum of generator non-coherency and load-generation imbalance, subject to constraints that ensure that the reference generators are in disjoint islands.

- We relax the islanding problem by removing the limits on load and generation and then prove that the relaxed formulation has bounded submodularity ratio (weak submodularity). We also prove a sufficient condition for the constraints using graph matroid.

- A polynomial-time greedy algorithm is proposed based on the weak submodularity to approximately solve the relaxed islanding problem and We prove that the algorithm has optimality guarantees on the minimal non-coherency and load-generation imbalance.

- A submodular maximization approach to selecting reference generators is proposed based on slow-coherency theory, which provides a optimality guarantee of $(1 - 1/e)$ on reference generator independence.

- The proposed weak-submodular approach to controlled islanding is tested on the IEEE 39-bus and 118-bus systems. The optimality of the proposed islanding method is validated by comparing to a two-step spectral clustering based approach [16].

The rest of this chapter is organized as follows. Section 5.2 presents the power system model and background on submodularity and matroid. Section 5.3 presents the problem formulation and the islanding algorithm. Section 5.4 discusses the slow-coherency based reference generator selection and proposes a submodular maximization approach to selecting reference generators. Section 5.5 presents the simulation results. Section 5.6 concludes this chapter.

## 5.2 System Model and Preliminaries

### 5.2.1 Power System Dynamics and Generator Coherency

Consider a power system consisting of $n$ generators, $m$ buses and $l$ transmission lines that connect buses. Let $\delta_i$ denote the rotor angle of generator $i$ and $\Delta\delta_i$ denote the rotor angle deviation from a steady state operating point. Let $M_i$ denote the inertia of generator $i$. Given a base frequency $\omega_0$, the generator dynamics based on the classical linearized generator swing equation is given as [9]

$$\Delta\ddot{\delta} = M^{-1}K\,\Delta\delta, \tag{5.1}$$

where $\Delta\delta = [\Delta\delta_1, \ldots, \Delta\delta_n]^T$ is the state vector and $M = diag(2M_1/\omega_0, \ldots, 2M_n/\omega_0)$ is the inertia matrix. The matrix $K$ has entries

$$K_{ij} = \begin{cases} -V_i V_j B_{ij} \cos(\delta_i - \delta_j), & \text{if } i \neq j, \\ -\displaystyle\sum_{k=1, k\neq i}^{n} K_{ik}, & \text{if } i = j, \end{cases}$$

where $V_i$ is the per unit voltage behind transient reactance at generator $i$ and $B_{ij}$ is the imaginary part of the $(i, j)$-entry of the admittance matrix reduced to the internal generator nodes.

Following a large disturbance, generators often oscillate at different frequencies. Two synchronous generators are called $\epsilon$-coherent if the maximum change in their rotor angle difference falls within a specified tolerance, $\delta_{ij}(t) - \delta_{ij}(0) \leq \epsilon$, at any time $t$ [27]. The $\epsilon$ is omitted for simplicity of notation. In the design of a strategy to mitigate impending cascading failures, it is critical to have coherent generators remain connected and separate any generators that are non-coherent into disjoint subsystems. Each subsystem is an island and the set of generators in the same island is called a generator group. Let $r$ be the desired number of islands.

Given a set of eigenvalues $\sigma_r = \{\lambda_1, \ldots, \lambda_r\}$ of the system matrix $M^{-1}K$, let $U$ be the eigenbasis of the $\sigma_r$ eigenspace. Two generators (states) are coherent with respect to $\sigma_r$ if their corresponding row vectors in $U$ are linearly dependent. With selection of the $r$ "slowest modes" of the system matrix $M^{-1}K$, representing the $r$ eigenvalues with smallest magnitudes, the generators dominant in each mode will have eigenbasis rows that are most linearly independent to each other [9]. Such a set of $r$ generators are called reference generators and can be used to initiate the $r$ islands in controlled islanding. The classic heuristics as well as our proposed submodular approach to reference generator selection are presented in Section 5.4.

Given a set of reference generators, an $n \times r$ coherency matrix $L$ can be calculated, whose entries represent how coherent each generator is respect to a reference generator. The algorithm to obtain $L$ has the following steps [9]:

1) Choose the set of $r$ slowest modes, $\sigma_r$.

2) Compute a basis matrix $U_{n\times r}$ of the $\sigma_r$-eigenspace.

3) Let $U_1$ denote the matrix composed of the rows of $U$ corresponding to the reference generators.

4) Compute $L = UU_1^{-1}$.

Any generator grouping strategy can be represented by a partition matrix $L_g$ that has exactly one 1 per row and all other entries zero. Matrix $L_g$ has the same size as $L$. Each column of $L_g$ represents a generator group dominated by a reference generator. The $(i, j)$th entry of $L_g$ being 1 can be interpreted as that the generator $i$ is in the group of the $j$th reference generator. Given any grouping strategy, we define a metric that quantifies how coherent generators are within each group, given by

$$\overline{H}(L_g) = \|L - L_g\|_F^2, \tag{5.2}$$

where $||\cdot||_F$ corresponds to the Frobenius norm. Note that $\overline{H}(L_g)$ measures the overall non-coherency of generators in each group and hence $\overline{H}(L_g) = 0$ when all generators within the same group are perfectly coherent ($\delta_{ij}(t) - \delta_{ij}(0) = 0, \forall t$).

This subsection has established the dynamical model of the power system. The algebraic properties of the system will be modeled in the following subsection.

### 5.2.2 *Power flow network*

Given a power system with $m$ buses and $l$ transmission lines, its topology can be depicted by a graph $G(N, E)$ where $N = \{1, \ldots, m\}$ is the index set of buses and $E \subseteq N \times N$ is the set of edges. Denote $e_k$ as the $k$th element in $E$. An edge $e_k = (i, j)$ is in $E$ if and only if there is a transmission line connecting bus $i$ and bus $j$. For any bus $i$, the set of its neighboring buses is defined by $\Delta(i) = \{j : (i, j) \in E\}$. A bus with a generator connected to it is called a generator bus while others are called load buses. Let $g_i$ be the active power generation and $d_i$ be the active power demand at bus $i$. Any distributed generation has been integrated as a part of the load at each bus [47]. The amount of active power flowing over a transmission line $(i, j) \in E$ from bus $i$ to $j$ is denoted by $p_{ij}$. At the initial operating point, set $g_i = g_i^0$ and $d_i = d_i^0$ for all buses $i$ and $p_{ij} = p_{ij}^0$ for all transmission lines.

With transmission line resistance neglected, the net active power flowing into any bus equals the net active power demand at this bus, which is known as the flow conservation property:

$$\sum_{j \in \Delta(i)} p_{ji} = d_i - g_i, \ \forall i \in N. \tag{5.3}$$

Using the incidence matrix $A$ of an orientation of the graph $G(N, E)$, the linear equations in (5.3) are equivalent to the matrix expression

$$Ax = b, \tag{5.4}$$

where $A$ is an $m \times l$ matrix with each row corresponding to a bus and each column corresponding to a transmission line. The entry of $A$ in the $i$th row and $k$th column is denoted by $(A)_{ik}$, where $(A)_{ik} = -1$ if the line $e_k \in E$ connects bus $i$ to some bus $j$ with $i > j$, $(A)_{ik} = 1$ if $e_k \in E$ connects $i$ to $j$ with $i < j$, and $(A)_{ik} = 0$ otherwise. The vector $x \in \mathbb{R}^l$ consists of entries $x_k$, where the $k$th entry $x_k = p_{ij}$ corresponding to the power flow on the edge $e_k = (i, j) \in E$. The vector $b \in \mathbb{R}^m$ has entries $(d_i - g_i)$ corresponding to the net active power demand at each bus $i$.

Let $a_i$ be the $i$th column of the incidence matrix $A$ and for any subset $S \subseteq E$, define $A(S)$ to be a matrix consisting of columns $\{a_i, \forall e_i \in S\}$. Tripping any transmission line $e_k \in E$ will change the incidence matrix $A$ in Eq. (5.4) to $A(E \setminus \{e_k\}) = [a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_l]$.

### *5.2.3 Spectral Clustering Based Controlled Islanding*

Spectral clustering is a technique making use of the eigenvalues/eigenvectors of the similarity matrix to perform dimensionality reduction before clustering data. The spectral clustering based approach to controlled islanding is implemented in two steps. Denote $\Omega_G \subseteq N$ as the set of generator buses. The first step is to classify generators into two coherent groups $T_1, T_2 \subseteq \Omega_G$ by minimizing the dynamic coupling between groups:

$$\min_{T_1, T_2 \subseteq \Omega_G} \sum_{j \in T_2} \sum_{i \in T_1} \left( \frac{\partial p_{ij}}{\partial \delta_{ij}} \left( \frac{1}{M_i} + \frac{1}{M_j} \right) \right),$$

where $\partial p_{ij}/\partial \delta_{ij} = |V_i V_j B_{ij} \cos(\delta_i - \delta_j)|$. In the second step, the algorithm finds a set of edges to cut such that the system is split into two subsystems $S_1, S_2 \subseteq N$ where $T_1 \subseteq S_1$ and $T_2 \subseteq S_2$, while minimizing the power-flow disruption on cutting edges:

$$\min_{S_1, S_2 \subseteq N} \sum_{i \in S_1, j \in S_2} |p_{ij}|, \;\; s.t. \; T_1 \subseteq S_1, T_2 \subseteq S2.$$

For $r > 2$, repeat the above process for each subsystem until the desired number of islands $r$ is achieved.

## 5.3 Proposed Controlled Islanding framework

This section formulates the problem of partitioning a power system into internally stable islands and presents a greedy islanding algorithm with provable optimality guarantees by exploiting the weak submodularity of the problem.

### 5.3.1 Problem Formulation

Suppose that we are given a power system with graph $G(N, E)$ and a set of reference generators at buses $\{s_1, \ldots, s_r\} \subseteq N$. We are also given a maximum generation limit $g^i_{\max}$ for each generator and a desired load $d^i_{\max}$ for each load at bus $i$. Note that $g^i_{\max} = 0$ for load buses and $d^i_{\max} = 0$ for buses without any load.

The problem studied in this chapter is to select a set of edges $S' \subseteq E$ to trip such that the system is partitioned into $r$ disjoint internally stable islands where each island contains a unique reference generator. Let $S$ denote the complementary set of $S'$ such that $S = E \setminus S'$. Such a set $S$ is called a partitioning set. The partition of the system induced by $S$ is called an islanding strategy.

An island being internally stable requires that the generators within the island must be coherent and the imbalance between load and generation at all buses must be minimized [77]. The load-generation imbalance is an estimate of how much load must be shed during controlled islanding. In the existing literature [16, 74], the load-generation imbalance is often defined as sum of mismatch between load and generation within each island or sum of power flows on lines to be tripped, based on a power flow solution at a steady state prior to the disturbance. For islands having power imbalance, the load that must be shed depends on the new steady state established in the island.

A metric that captures the minimum load that must be shed, given any partitioning set $S$, is defined as follows

$$F(S) \triangleq \begin{array}{ll} \min\limits_{x,d,g} & ||A(S)x - (d_0 - g_0)||_2^2 \\ \text{s.t.} & 0 \leq d \leq d_{\max},\ 0 \leq g \leq g_{\max}, \\ & A(S)x = d - g, \end{array} \tag{5.5}$$

where $A(S)$ is the incidence matrix induced by the set of edges $S$. The vectors $d_{\max} = [d_{\max}^1, \ldots, d_{\max}^m]^T$ and $g_{\max} = [g_{\max}^1, \ldots, g_{\max}^m]^T$ are the maximum available load and generation at each bus, respectively. The vector $(d_0 - g_0) = [d_1^0 - g_1^0, \ldots, d_m^0 - g_m^0]^T$ is the net load at initial operating point. The metric $F(S)$ is equal to the minimum load-generation imbalance for any power flow $x$ that satisfies the flow conservation constraint (5.4) at each bus. Note that given any partitioning $S$, the optimal value of $x$ in (5.5) can infer the post-islanding load $d$ and generation $g$ at each bus. This can approximately indicate the amount and location of load to be shed. For more accurate calculations, other methods including the rate of frequency decline-based load shedding can be used [76].

Generator coherency is another essential property to achieve the internal stability of islands. In the following, a metric is defined based on (5.2) that evaluates the generator non-coherency given any partitioning set $S$. By linearizing the system dynamics at a steady-state operating point, a model can be established as described by (5.1). Following the algorithm in Section 5.2.1, one can obtain a coherency matrix $L$ with respect to the given reference generators. Given any set $S$, a partition matrix $L_g$ can be calculated and the overall generator non-coherency is measured by $\overline{H}(L_g)$ in Eq. (5.2).

Let $u_i$ denote the bus index of the $i$th generator. For each non-reference generator at bus $u_i \in N \setminus \{s_1, \ldots, s_r\}$, define a function $H_i$ to be the optimal value of a quadratic program:

$$H_i(S) \triangleq \begin{array}{ll} \min\limits_{x} & ||A(S)x - c^i||_2^2 \\ \text{s.t.} & (A(S)x)_j = 0,\ \forall j \notin \{u_i, s_1, \ldots, s_r\}, \end{array} \tag{5.6}$$

where $A(S)$ is the incidence matrix induced by $S$ and $c^i$ is a vector defined by

$$(c^i)_j = \begin{cases} 1, & j = u_i \\ -L_{ik}, & j = s_k \in \{s_1, \dots, s_r\} \\ 0, & \text{else.} \end{cases}$$

The following lemma establishes the relation between the metrics $H_i$ and $\overline{H}(L_g)$.

**Lemma 16.** *If the set of edges $S$ induces a partition such that each reference generator $s_k \in \{s_1, \dots, s_r\}$ is in a unique island and the generator at bus $u_i \in N \setminus \{s_1, \dots, s_r\}$ is connected to exactly one reference generator at bus $s_j$, then*

$$H_i(S) = \frac{1}{2}(1 - L_{ij})^2 + \frac{1}{2}\sum_{k=1, k\neq j}^{r} L_{ik}^2 = \frac{1}{2}\|(L - L_g)_i\|_2^2,$$

*where $(L - L_g)_i$ is the $i$th row of the matrix $(L - L_g)$.*

*Proof.* Suppose that the set of edges $S$ induces a partition of the system $G(N, E)$, where each reference generator $s_k \in \{s_1, \dots, s_r\}$ is in a unique island and the generator at bus $u_i \in N \setminus \{s_1, \dots, s_r\}$ is connected to a reference generator $s_j$. With an orientation of the graph $G$, the matrix $A(S)$ has the following block pattern

$$A(S) = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{bmatrix},$$

where each block $A_k$ corresponds to the incidence matrix of an island containing reference generator $s_k$. Similarly, the vectors in (5.6) become $x = [\vec{x}_1, \dots, \vec{x}_r]^T$ and $c^i = [c_1^i, \dots, c_r^i]^T$. Thus,

$$H_i(S) = \sum_{k=1}^{r} \min_{\vec{x}_k} \|A_k\vec{x}_k - c_k^i\|_2^2, \tag{5.7}$$

subject to $(A(S)x)_k = 0,\ \forall k \notin \{u_i, s_1, \dots, s_r\}$.

Each term of $H_i(S)$ in (5.7) is a minimization problem. The solution to each problem is equivalent to the load-generation imbalance defined by (5.5) of a unique island. For the island $k = j$

which contains the bus $u_i$, given a "generation" of $1$ at $u_i$ and a "load" of $L_{ij}$ at $s_j$, it has a "load-generation imbalance" of $(1-L_{ij})^2/2$ resulting from the corresponding minimization problem. For each of the rest islands where $k \neq j$, there is only a "load" of $L_k$ at bus $s_k$ without any generation in the island, which results in a "load-generation imbalance" of $L_{ik}^2$. With the analogy above, we get

$$H_i(S) = \frac{1}{2}(1 - L_{ij})^2 + \frac{1}{2}\sum_{k=1,k\neq j}^{r} L_{ik}^2.$$

$L_g$ is the generator partitioning matrix induced by $S$. If the generator at bus $u_i$ is in the same island of the reference generator $s_j$, then the $i$th row of $L_g$ has a $1$ at $j$th column and $0$ anywhere else. Thus, we have

$$\|(L - L_g)_i\|_2^2 = (1 - L_{ij})^2 + \sum_{k=1,k\neq j}^{r} L_{ik}^2,$$

and hence $H_i(S) = \|(L - L_g)_i\|_2^2/2$, completing the proof. □

The proof of Lemma 16 is established using the analogy of the formulation in (5.5), in which we have a "generation' of $1$ at bus $u_i$ and a "load demand"' of $L_{ik}$ at bus $s_k$, $\forall k \in \{1, \ldots, r\}$. We restrict $d_{\max}$ and $g_{\max}$ to be $0$ for buses not in the set $\{u_i, s_1, \ldots, s_r\}$.

The function $H_i(S)$ can be interpreted as the sum of non-coherency in each group caused by the $i$th generator, given an islanding strategy $S$. We note that $H_i(S) = 0$ if the generator $i$ is a reference generator.

By Lemma 16 and the fact that $\overline{H}(L_g) = \|L - L_g\|_F^2 = \sum_{i=1}^n \|(L - L_g)_i\|_2^2$, we find the equivalence of the metrics $\overline{H}(L_g)$ and $H_i(S)$, given by $\overline{H}(L_g) = 2\sum_{i=1}^n H_i(S)$.

The controlled islanding problem can be formulated as

$$\begin{aligned} \min_{S} \quad & \xi F(S) + \textstyle\sum_{i=1}^n H_i(S) \\ \text{s.t.} \quad & S = S_1 \cup \cdots \cup S_r,\ N = N_1 \cup \cdots \cup N_r, \\ & s_k \in N_k,\ \forall k \in 1, \ldots, r \end{aligned} \tag{5.8}$$

where the objective function is to minimize the sum of load-generation imbalance and generator non-coherency with a trade-off parameter $\xi \geq 0$ balancing the two objectives. The constraint requires that each reference generator must be in a unique island.

In classic two-step formulations, when given the groups of coherent generators, the problem of partitioning a system with minimum load-generation imbalance is typically NP-hard [85]. For the one-step problem formulation (5.8), the objective function is the sum of a finite number of least square metrics, where each metric measures the distance between a vector and the subspace spanned by the selected column vectors from $A$. As is shown later in Section 5.3.2, the constraint of (5.8) can be formulated as a graph matroid constraint and hence the problem in (5.8) is equivalent to a subspace selection optimization problem with matroid constraint which is known to be NP-hard [84].

In the following section, the weak submodularity of problem (5.8) is exploited, which leads to polynomial-time approximation algorithms with provable bound on the optimality of solutions.

### 5.3.2 Weak Submodularity Approach

In order to define a relaxation of the problem (5.8) which has weak submodularity, the following metrics are introduced:

$$h_i(S) = \min_x ||A(S)x - c^i||_2^2,$$

for all $i \in \{1, \ldots, n\}$ and

$$f(S) = \min_x ||A(S)x - b_0||_2^2,$$

where $b_0 = d_0 - g_0$ and the restrictions on variables $d$, $g$ and $x$ are removed. The following lemma establishes the connection between metrics $f(S)$ and $F(S)$, as well as $h_i(S)$ and $H_i(S)$.

**Lemma 17.** *Given any set $S \subseteq E$, we have $f(S) = 0$ if $F(S) = 0$, and $f(S) \leq F(S)$ if $F(S) > 0$; we have $h_i(S) = 0$ if $H_i(S) = 0$, and $h_i(S) \leq H_i(S)$ if $H_i(S) > 0$, for all $i$.*

*Proof.* If $F(S) = 0$, then there exists a power flow $\bar{x}$ such that $A(S)\bar{x} = b_0$. Thus, we have

$$f(S) = \min_x \ \|A(S)x - b_0\|_2^2 = \|A(S)\bar{x} - b_0\|_2^2 = 0.$$

If $F(S) > 0$, let $\hat{x}$ be the optimal solution of $F(S)$, i.e., $\hat{x} = \arg\min_x \ \|A(S)x - (d_0 - g_0)\|_2^2$ $s.t.\ 0 \leq d \leq d_0, 0 \leq g \leq g_0$ and $A(S)x = d - g$. Then we have

$$f(S) = \min_x \ \|A(S)x - b_0\|_2^2 \leq \|A(S)\hat{x} - b_0\|_2^2 = F(S).$$

Similar proofs for $h_i(S)$ and $H_i(S)$ can be established following the same steps. $\square$

By Lemma 17, the function $f(S)$ is a lower bound to the load-generation imbalance $F(S)$ by relaxing the restrictions on generation and load at each bus, while $h_i(S)$ is a lower bound to the generator non-coherency $H_i(S)$ for each generator $i$.

Thus, the problem (5.8) has a relaxed formulation, given as follows, that provides a lower bound on the optimal solution:

$$\begin{aligned}
\min_S \quad & \xi f(S) + \textstyle\sum_{i=1}^n h_i(S) \\
\text{s.t.} \quad & S = S_1 \cup \cdots \cup S_r,\ N = N_1 \cup \cdots \cup N_r, \\
& s_k \in N_k,\ \forall k \in 1, \ldots, r
\end{aligned} \tag{5.9}$$

The following theorem establishes the weak submodularity structure of the objective function in (5.9).

**Theorem 5.** *Let $J(S) = \xi f(S) + \sum_{i=1}^n h_i(S)$. Denote $\gamma_{U,k}$ as the submodularity ratio of $J(S)$. Then for any set $U \subseteq E$ and $k \geq 1$, the submodularity ratio $\gamma_{U,k}$ is bounded by*

$$\gamma_{U,k} \geq \lambda_{\min}(C, k + |U|) \geq \lambda_{\min}(C),$$

*where $C = A^T A/(2n)$; $\lambda_{\min}(C)$ is the smallest eigenvalue of $C$ and $\lambda_{\min}(C,k)$ is the smallest $k$-sparse eigenvalue of $C$, defined by $\lambda_{\min}(C,k) = \min_{S:|S|=k} \lambda_{\min}(C(S))$ where $C(S) = A(S)^T A(S)/(2n)$.*

*Proof.* For any vector $v$, we define a metric

$$g_v(S) = \min_x \|A(S)x - v\|_2^2.$$

Equivalently, $g_v(S) = \min_y \| \left(A(S)/\sqrt{2}\right) y - v\|_2^2$ where each column of the matrix $\left(A(S)/\sqrt{2}\right)$ has norm 1 and sum of entries 0. The metric $g_v(S)$ has been shown to have bounded submodularity ratio $\gamma'_{U,k} \geq \lambda_{\min}(C, k + |U|) \geq \lambda_{\min}(C)$ [34, 13, 46]. Hence, it suffices to show that $J(S)$ has submodularity ratio $\gamma_{U,k}$ bounded by $\gamma'_{U,k}$, i.e., $\gamma_{U,k} \geq \gamma'_{U,k}$.

Let $\{v_i\}$ be a set of vectors where $v_0 = b_0$ and $v_i = c^i$ for $i \in \{1, \ldots, n\}$. Let $\{a_i\}$ be a set of weights where $a_0 = \xi$ and $a_i = 1$ for $i \in \{1, \ldots, n\}$. Then $J(S) = \sum_{i=0}^{n} a_i g_{v_i}(S)$.

$$\begin{aligned}
\gamma_{U,k} &= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_{z\in S}\left(J(L\cup\{z\}) - J(L)\right)}{J(L\cup S) - J(L)} \\
&= \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \frac{\sum_i \sum_{z\in S}\left(a_i g_{v_i}(L\cup\{z\}) - a_i g_{v_i}(L)\right)}{\sum_i \left(a_i g_{v_i}(L\cup S) - a_i g_{v_i}(L)\right)} \\
&\geq \min_{\substack{L,S:L\cap S=\emptyset \\ L\subseteq U,|S|\leq k}} \left(\min_i \frac{\sum_{z\in S}\left(g_{v_i}(L\cup\{z\}) - g_{v_i}(L)\right)}{\left(g_{v_i}(L\cup S) - g_{v_i}(L)\right)}\right) \\
&\geq \gamma'_{U,k}
\end{aligned}$$

completing the proof. □

In what follows, we define a graphic matroid constraint that is sufficient to satisfy the constraint on $S$ of problem (5.9). Given the graph $G(N, E)$ and reference generators at buses $\{s_1, \ldots, s_r\} \in N$, we construct a new graph $G_0(N_0, E_0)$ as follows. The vertex set $N_0 = N \cup \{0\}$ and the edge set $E_0 = E \cup \{(0, s_i), i = 1, \ldots, r\}$. Intuitively, the graph $G_0$ is obtained by adding a new node $0$ to the graph $G$ and creating new edges to connect all reference generators to node $0$.

Let $\mathcal{M} = (E_0, \mathcal{I}_0)$ denote the graphic matroid defined on the graph $G_0$, with basis $\mathcal{I}$ being the collection of all spanning trees of $G_0$. The following lemma relates the basis $\mathcal{I}$ of $\mathcal{M}$ to the constraint in (5.9).

**Lemma 18.** *Given a graph $G(N, E)$ and a set of reference generators $\{s_1, \ldots, s_r\} \in N$, a set of edges $S \in E$ induces a partition of the graph: $S = \{\cup_{k=1}^{r} S_k\}$, $N = \{\cup_{k=1}^{r} N_k\}$ and $S_i \cap S_j = \emptyset, N_i \cap N_j = \emptyset, \forall i \neq j$, where each reference generator satisfies $s_k \in N_k, \forall k$, if*

$$S \cup \{(0, s_i), i = 1, \ldots, r\} \in \mathcal{I}. \tag{5.10}$$

*Proof.* The proof is established by contradiction. If $S \cup \{(0, s_i), i = 1, \ldots, r\} \in \mathcal{I}$, then the set of edges $S \cup \{(0, s_i), \forall i\}$ forms a spanning tree of $G_0$. The node 0 is the root of the tree and $S = \{\cup_{k=1}^{r} S_k\}$ and $N = \{\cup_{k=1}^{r} N_k\}$ form $r$ branches of the tree. Given any two reference generators at nodes $s_i$ and $s_j$, then they are connected by a path $(s_i, 0), (0, s_j)$ in graph $G_0$. Suppose that two reference generators are in the same branch, i.e., $s_i \in N_i$ and $s_j \in N_i$. Then $S$ contains edges that connect $s_i \in N_i$ and $s_j \in N_i$ by a path $(s_i, v_1), (v_1, v_2), \dots, (v_k, s_j) \subseteq S$ in graph $G$. Since node $0 \notin N$, we have $0 \neq v_1 \neq \cdots \neq v_k$. Thus, the path $(0, s_i), (s_i, v_1), (v_1, v_2), \dots, (v_k, s_j), (s_j, 0) \subseteq S \cup \{(0, s_i), \forall i\}$ forms a cycle in graph $G_0$, which contradicts with the definition of a spanning tree. Therefore, by contradiction, we have that $s_k \in N_k, \forall k$. □

By Lemma 18, the problem of selecting a set of edges $S \subseteq E$ to form disjoint islands where each island contains a unique reference generator, while minimizing load-generation imbalance and generator non-coherency can be formulated as

$$\begin{aligned} \min_S \quad & J(S) \\ \text{s.t.} \quad & S \cup \{(0, s_i), i = 1, \dots, r\} \in \mathcal{I}. \end{aligned} \tag{5.11}$$

where $J(S) = \xi f(S) + \sum_{i=1}^{n} h_i(S)$.

### *5.3.3 Proposed Islanding Algorithm*

In this subsection, we present a greedy algorithm that approximately solves (5.11) with provable optimality guarantees. The algorithm consists of two stages, namely greedy selection and local search. Each stage proceeds as follows. In greedy selection, the set $S$ is initialized to be empty and the groundset $\Omega$ is initialized to be the set of edges $E$. At each iteration, the algorithm selects an edge $e \in \Omega \setminus S$ that maximizes $J(S) - J(S \cup \{e\})$ and deletes $e$ from $\Omega$. The edge $e$ is added to the set $S$ if it does not generate any cycle in the graph consisting of edges $S \cup \{e\} \cup \{(0, s_i), \forall i\}$. The greedy selection terminates when the groundset $\Omega$ is empty.

In local search, let $\epsilon > 0$ be a constant parameter. The algorithm selects a pair of edges $(v \in S, e \in E \setminus S)$ that by swapping them can lower the function $J$ value, i.e., $J(S \setminus \{v\} \cup \{e\}) < (1 - \epsilon)J(S)$, and the graph consisting of edges $S \setminus \{v\} \cup \{e\} \cup \{(0, s_i), \forall i\}$ contains no cycle. A pseudocode description is given as Algorithm 4.

Intuitively, the algorithm first selects a spanning tree from graph $G_0$ while minimizing the function $J$. Then in local search, the algorithm swaps edges from the tree to converge to a local

**Algorithm 4:** Algorithm for selecting a set of edges that partitions the system into internally stable islands

**procedure** CONTROL_ISLAND($G(N, E)$, $\{s_i\}$, $b_0$, $L$, $\xi$)

**Input**: Power system topology $G(N, E)$, set of reference generators $\{s_i\} \subseteq N$, net load at each bus $b_0$, generator coherency matrix $L$, trade-off parameter $\xi$.

**Output:** Set of edges $S \in E$ that form a partition of $G$, where each partition contains a group of coherent generators while minimizing load-generation imbalance.

**Initialization:** Construct metric $J = \xi f + \sum_{i=1}^{n} h_i$; Construct new graph $G_0(N_0, E_0)$ by adding new node 0 and edges $(0, s_i), \forall i$ to $G$; Set $\Omega = E$, $S = \emptyset$.

*1) Greedy Selection:*

**while** $\Omega \neq \emptyset$ **do**

$e \leftarrow \arg\max_{v \in \Omega \setminus S} \ (J(S) - J(S \cup \{v\}))$

**if** $S \cup \{e\} \cup \{(0, s_i), \forall i\}$ *contains no cycle* **then**

$S \leftarrow S \cup \{e\}$

$\Omega \leftarrow \Omega \setminus \{e\}$

*2) Local Search:*

**while** $\exists v \in S$, $\exists e \in E \setminus S : J(S \setminus \{v\} \cup \{e\}) < (1 - \epsilon)J(S)$ *and* $S \setminus \{v\} \cup \{e\} \cup \{(0, s_i), \forall i\}$ *has no cycle* **do**

$S \leftarrow S \setminus \{v\} \cup \{e\}$

**return** $S$

**end procedure**

optimum which further minimizes $J$. The following proposition describes the optimality guarantee of Algorithm 4.

**Proposition 3.** *Let $S^*$ denote the optimal solution to (5.11) and let $S$ be the solution returned by*

*Algorithm 4. Then, we have*

$$J(S) \leq (m - r - \gamma_0)\, J(S_{t-1}) + \gamma_0 J(S^*),$$

*where $\gamma_0 = \lambda_{\min}(C, 2|S|)$ and $S_{t-1}$ denotes the selected set at the second-to-last iteration of Algorithm 4.*

*Proof.* The matroid $\mathcal{M}$ has rank $m$ and hence $|S^*| = |S| = m - r$. Let $e_i$ be the $i$th element selected by the Algorithm 4 and let $S_i = \{e_1, \ldots, e_i\}$ be the set containing the first $i$ elements picked by the algorithm. Let $\pi : S^* \to S$ be a bijection as in Lemma 1 so that the elements $e_i^*$ of $S^*$ are sorted to satisfy $S \setminus \{e_i\} \cup \{e_i^*\} \in \overline{\mathcal{I}}$ for all $i$.

By the definition of submodularity ratio, we have

$$\frac{\sum_{z \in S^*} \left(J(S_{i-1}) - J(S_{i-1} \cup \{z\})\right)}{J(S_{i-1}) - J(S_{i-1} \cup S^*)} \geq \gamma_{S,|S|} \geq \gamma_0,\ \forall i,$$

where $\gamma_0 = \lambda_{\min}(C, 2|S|)$. By the greediness of Algorithm 4 and that $S_{i-1} \cup \{e_i^*\} \in \overline{\mathcal{I}}$ for all $i$, we have

$$J(S_{i-1}) - J(S_i) + \sum_{z \in S^* \setminus e_i^*} \left(J(S_{i-1}) - J(S_{i-1} \cup \{z\})\right) \geq \gamma_0 \left(J(S_{i-1}) - J(S_{i-1} \cup S^*)\right).$$

Since $J$ is monotone decreasing, we have

$$J(S_{i-1}) - J(S_i) + \sum_{z \in S^* \setminus e_i^*} J(S_{i-1}) \geq \gamma_0 \left(J(S_{i-1}) - J(S^*)\right),$$

or equivalently,

$$|S^*| J(S_{i-1}) - J(S_i) \geq \gamma_0 \left(J(S_{i-1}) - J(S^*)\right),$$

and hence

$$J(S_i) \leq (|S^*| - \gamma_0)\, J(S_{i-1}) + \gamma_0 J(S^*).$$

At the last iteration, we have

$$J(S) \leq (|S^*| - \gamma_0)\, J(S_{t-1}) + \gamma_0 J(S^*),$$

where $S_{t-1}$ denotes the selected set at the second-to-last iteration of Algorithm 4.

□

The complexity of Algorithm 4 is described by the following proposition.

**Proposition 4.** *The runtime of Algorithm 4 is bounded above by $O(l^5 n + l^2 C)$, where the greedy selection stage requires a computation of at most $O(l^5 n)$ and the local search requires $O(l^2 C)$, where*

$$C = \frac{\log J(E) - \log J(\emptyset)}{\log(1-\epsilon)}.$$

*Proof.* The runtime of Algorithm 4 is determined by the complexity of evaluating the objective function $J(S)$ and checking the graphical matroid constraint. For greedy selection stage, the algorithm requires $O(l^2)$ evaluations of $J(S)$, each of which involves solving $(n+1)$ least-squares problems with complexity of $O(l^3)$, and $O(l)$ checks of the matroid constraints. Checking the matroid constraint is to detect cycle in an undirected graph which requires computation of $O(l)$. Thus, the total complexity for greedy selection is $O(l^5 n)$.

In local search, let $S_i$ denote the set $S$ after $i$ iterations of the algorithm. By construction, $J(S_i) < (1-\epsilon)J(S_{i-1})$ and hence $J(S_i) < (1-\epsilon)^i J(S_0)$. Let $k$ be the number of iterations before the algorithm terminates. By the monotonicity of $J$, we have $J(E) = \min_S J(S)$ and $J(\emptyset) = \max_S J(S)$. Therefore,

$$J(E) \leq J(S_k) < (1-\epsilon)^k J(S_0) \leq (1-\epsilon)^k J(\emptyset).$$

Rearranging terms yields $k \leq C$. Each iteration requires at most $l^2$ evaluations of the objective function, implying that the worst-case computation is $O(l^2 C)$. □

### 5.4 Reference Generator Selection

In this section, we discuss the slow-coherency based reference generator selection and present a submodular optimization approach to selecting reference generators.

From the discussion in Section 5.2.1, given the set of $r$ slowest modes $\sigma_r$ and its corresponding eigenbasis matrix $U$, two generators (states) are coherent with respect to $\sigma_r$ if their corresponding row vectors in $U$ are linearly dependent. Then selecting $r$ reference generators is equivalent to find $r$ rows from $U$ that are most linearly independent.

Define a matrix $U(T)$ consisting of the selected set $T$ of rows from $U$. A measure of the linear independence of the row vectors of $U(T)$ is given by the Gramian

$$G_U(T) = \det(U(T)U(T)'),$$

where $U(T)'$ is the transpose of $U(T)$.

A physical interpretation of the Gramian $G_U(T)$ is that the row vectors of $U(T)$ forms a parallelepiped in the case of $|S| = 3$ and $G_U(T)$ is the volume of the parallelepiped. For $|T| > 3$, the parallelepiped becomes a $|T|-$dimensional polytope formed by the row vectors of $U(T)$.

The linear independence of the rows of $U(T)$ reaches maximum when the volume of the parallelepiped/polytope formed by $U(T)$ is largest. Therefore, the reference generator selection can be formulated as selecting a set $T$ of rows from $U$ such that

$$\max_{T \subseteq \{1,\ldots,n\}} G_U(T), \quad s.t. \ |T| = m. \tag{5.12}$$

A classic heuristic method to approximately solve (5.12) is by applying complete-pivoting Gaussian elimination to matrix $U$ [9]. Following the elimination, the generator (state) dominant in each column is chosen as a reference generator. This heuristic, however, does not have any guarantee on the optimality of solutions.

In what follows, we present a submodular optimization based approach to reference generator selection, which provides a provable guarantee on the maximum linear independence achieved. At each iteration, the algorithm selects a reference generator $v \in \Omega \setminus T$ that maximizes $\phi(T \cup \{v\}) - \phi(T)$ and adds to the selection $T$. The algorithm terminates when the groundset $\Omega$ is empty.

Since $U(T)U(T)'$ is positive semi-definite, its determinant $G_U(T)$ is nonnegative. By the fact that $\log(\cdot)$ is a monotone increasing function, the optimal point of $\max_T G_U(T)$ is the same as $\max_T \log G_U(T)$. Thus, the problem (5.12) is equivalent to

$$\max_{T} \ \log\det(U(T)U(T)'), \ \ s.t. \ |T| = m. \tag{5.13}$$

**Theorem 6.** *The function $\phi(T) = \log\det(U(T)U(T)')$ is submodular in $T$.*

**Algorithm 5:** Algorithm for selecting reference generators.

**procedure** REFERENCE_GEN($\hat{A}, r$)

  **Input:** Matrix $\hat{A} = M^{-1}K$, number of islands $r$

  **Output:** Set of reference generators $T$

  **Initialization:** Calculate the set of $r$ slowest modes $\sigma_r$ and basis matrix $U$; construct the set function $\phi$; $T \leftarrow \emptyset$

  **while** $|T| < r$ **do**

   **for** $v \in \Omega \setminus T$ **do**

   $\delta_v \leftarrow \phi(T \cup \{v\}) - \phi(T)$

  $v^* \leftarrow \arg\max_v \delta_v$

  $T \leftarrow T \cup \{v^*\}$

  **return** $T$

**end procedure**

*Proof.* Let $\Sigma = UU'$. Suppose $X \in \mathbb{R}^r$ is multivariate Gaussian with probability density function

$$p(x) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp\left(-\frac{1}{2}(x-u)'\Sigma^{-1}(x-u)\right)$$

with some mean vector $u$. The (differential) entropy of the random variable $X$ is given by

$$\Phi(X) = \log\sqrt{|2\pi\Sigma|} = \log\sqrt{(2\pi)^r|\Sigma|}.$$

For a subset $T$ of variables, $X_T$, we denote $\Sigma_T = V(T)V(T)'$ and we have

$$\Phi(X_T) = \log\sqrt{(2\pi)^{|T|}|\Sigma_T|} = \frac{1}{2}|T|\log 2\pi + \frac{1}{2}\log|\Sigma_T|,$$

where the term $(|T|\log 2\pi)$ is modular in $T$.

Since entropy $\Phi(X_T)$ is submodular in $T$, the term $\log|\Sigma_T| = \log\det(V(T)V(T)') = \phi(T)$ is submodular. □

By Theorem 6, the reference generator selection (5.13) is a submodular maximization problem subject to a cardinality constraint. A simple greedy algorithm can approximately solve (5.13) in polynomial time $O(nr)$ up to a worst-case optimality bound of $(1-1/e)$ [5]. We present such an algorithm, described by Algorithm 5.

## 5.5 Numerical Study

This section presents the simulation results of our proposed approach to controlled islanding on IEEE 39-bus system and IEEE 118-bus system, with comparison to the two-step spectral clustering based islanding method which is the state of the art [16].

### 5.5.1 System Setup

For both test cases, the initial system data including topology, transmission line impedance, load and generator configuration are specified by the test system. We obtain an initial operating point including knowledge of generator voltage magnitudes $V_i$ and angles $\delta_i$, by solving an optimal power flow using *Matpower* [88]. The generator inertia $M_i$ for the 39-bus and the 118-bus systems are given by [29] and [66], respectively. The desired number of islands is set to $r = 3$.

The performance of an islanding strategy $S$ is evaluated by calculating the resulting load-generation imbalance and generator non-coherency. The metric $\sqrt{f(S)} = dist\left(b_0, span(A(S))\right)$ is used for estimating the load-generation imbalance with unit MW and $\overline{H}(L_g) = \|L - L_g\|_F^2$ is used for generator non-coherency where $L_g$ is the generator partition matrix induced by $S$. The trade-off parameter $\xi$ is chosen from the interval [0,1], in order to scale the load-generation imbalance down to have the same order of magnitude as the generator non-coherency.

### 5.5.2 Test Case 1 : IEEE 39-bus Test System

The IEEE 39-bus test system contains 10 generators and 46 transmission lines. The system topology is shown in Fig. 5.1. Algorithm 5 identifies generators {G1, G5, G9} as reference generators, which is consistent with the result of the classic Gaussian elimination based reference generator selection.

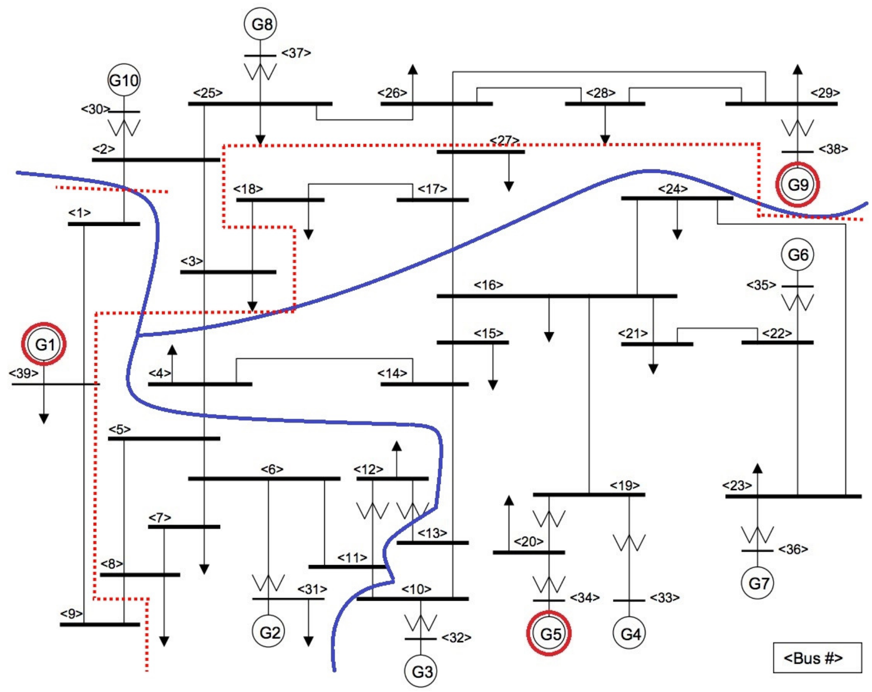



Figure 5.1: IEEE 39-bus system topology. The circled generators {G1,G5,G9} are reference generators identified by Algorithm 5. The solid curves represent the cutset selected by our Algorithm 4 with parameter $\xi = 10^{-7}$. The dotted lines represent the cutset selected by the spectral clustering method.

As a measure of the order of magnitude difference between the load-generation imbalance and generator non-coherency, when $S = \emptyset$, the ratio of $\sum_{i=1}^{n} h_i(S)/f(S)$ is $2 \times 10^{-6}$. For arbitrary $S$, the ratio of $\sum_{i=1}^{n} h_i(S)/f(S)$ generally varies within $[10^{-7}, 10^{-6}]$. By choosing $\xi = 10^{-7}$, Algorithm 4 finds the optimal islanding strategy based on metric $J(S)$ by tripping lines {1-2, 3-4, 4-5, 10-11, 12-13, 16-17}. The coherent generator groups found are {G1,G2}, { G3,G4,G5,G6,G7}, {G8,G9,G10}.

The two-step spectral clustering method finds generator groups {G1}, {G2,G3,G4,G5,G6,G7}, {G8,G9,G10} in the first step and a min-cut set of edges {1-2, 8-9, 3-4, 3-18, 17-27} in the second step.

Table 5.1: Comparison of the weak-submodular islanding ($\xi = 10^{-7}$) and spectral clustering islanding results on the 39-bus system.

| Method | $J(S)$ | Load-generation Imbalance $\sqrt{f(S)}$ | Generator Non-coherency $\overline{H}$ |
|---|---|---|---|
| Weak-Submodular | 0.1696 | 241.1 MW | 1.4237 |
| Spectral Clustering | 0.2149 | 330.3 MW | 1.3104 |

Table 5.2: Comparing results of the weak-submodular islanding with different $\xi$ values on the 39-bus system.

| Trade-off Parameter | Islands[1] | Imbalance $\sqrt{f(S)}$ | Non-coherency $\overline{H}$ |
|---|---|---|---|
| $\xi = 0$ | {1,4,5,7,8,9,39}<br>{2,3,17,18,25,26,27,28,29,30,37,38}<br>{6,10:16,19:24,31:36} | 758.3 MW | 1.3104 |
| $\xi > 10^{-5}$ | {2,3,4,5,25,26,27,28,29,30,37,38}<br>{1,6,7,8,9,10,11,31,32,39}<br>{12:24,33,34,35,36} | 28.5 MW | 1.5900 |

Fig. 5.1 illustrates the difference between islanding strategies of the proposed weak-submodular approach and the spectral clustering method on topology. A comparison of performance between the two islanding strategies is shown in Table 5.1. It is observed that the weak-submodular strategy reduces the load-generation imbalance by 27% compared to the spectral clustering islanding, although the generator non-coherency is 8.6% higher.

Table 5.2 shows islanding results from Algorithm 4 with different $\xi$ values. When choosing $\xi = 0$, only generator non-coherency is considered in the problem (5.11) and hence the objec-

[1] {10:16} represents {10,11,12,13,14,15,16} and so on.

tive function $J(S) = \sum_{i=1}^{n} h_i(S)$. The resulting submodular islanding strategy is able to find the same coherent generator groups as the spectral clustering method. When $\xi > 10^{-5}$, we have $J(S) \approx \sum_{i=1}^{n} h_i(S)$. The objective function $J(S)$ is dominated by the load-generation imbalance, while the effect of generator non-coherency is negligible. The Algorithm 4 finds a near-balanced islanding strategy by sacrificing 21.3% on generator non-coherency compared to the spectral clustering method.

For the IEEE 39-bus system, the proposed Algorithm 4 on average returns an islanding solution in 7 seconds, which consists of $4.5$ seconds for greedy selection and $2.5$ seconds for local swap (on a Macbook with 2.4GHz dual-core Intel Core i5 processor and 8GB RAM). The spectral clustering algorithm costs less computational time but does not provide provable optimality guarantees.

### *5.5.3 Test Case 2 : IEEE 118-bus Test System*

For this test case, the system topology is given by Fig. 5.2. Generators at buses {10, 54, 87} are selected by Algorithm 5 as reference generators, which is the same as classic reference generator selection result.

For the IEEE 118-bus system, the computational time of Algorithm 4 on average increases to $9$ minutes, which consists of $1$ min for greedy selection and $8$ min for local swap.

When $\xi = 10^{-6}$, Algorithm 4 returns a solution that has the same generator non-coherency but a lower load-generation imbalance compared to the spectral clustering solution, as shown in Table 5.3. A comparison between the two solutions on system topology is shown in Fig. 5.2.

In conclusion, the weak-submodular islanding approach can identify solutions that achieve lower load-generation imbalance compared to the two-step spectral clustering method, with the same or slightly higher generator non-coherency. This is due to the trade-off mechanism of the proposed approach that is able to reduce the load-generation imbalance by changing coherent generator groups.

Table 5.3: Comparison of the weak-submodular islanding ($\xi = 10^{-6}$) and spectral clustering islanding results on the 118-bus system.

| Method | $J(S)$ | Load-generation Imbalance $\sqrt{f(S)}$ | Generator Non-coherency $\overline{H}$ |
|---|---|---|---|
| Weak-Submodular | 0.0524 | 9.7 MW | 1.4515 |
| Spectral Clustering | 0.0527 | 22.1 MW | 1.4515 |

## 5.6 Conclusion

In this chapter, we studied the problem of selecting a set of transmission lines to trip in order to partition an unstable power system into internally stable islands. The proposed formulation captures the generator coherency and load-generation imbalance jointly in one objective function. The weak submodularity of the formulation was exploited. Based on the weak submodularity, a greedy islanding algorithm with provable optimality guarantees was proposed, which is currently not available in the existing literature. In this chapter, we also studied the reference generator selection and proposed a submodular maximization approach. The proposed islanding algorithms were validated on the IEEE 39-bus and 118-bus power systems, with a comparison to the spectral clustering-based islanding method. It was observed that the proposed islanding approach can significantly reduce the power imbalance compared to the state of the art, with the same or slightly higher generator non-coherency.

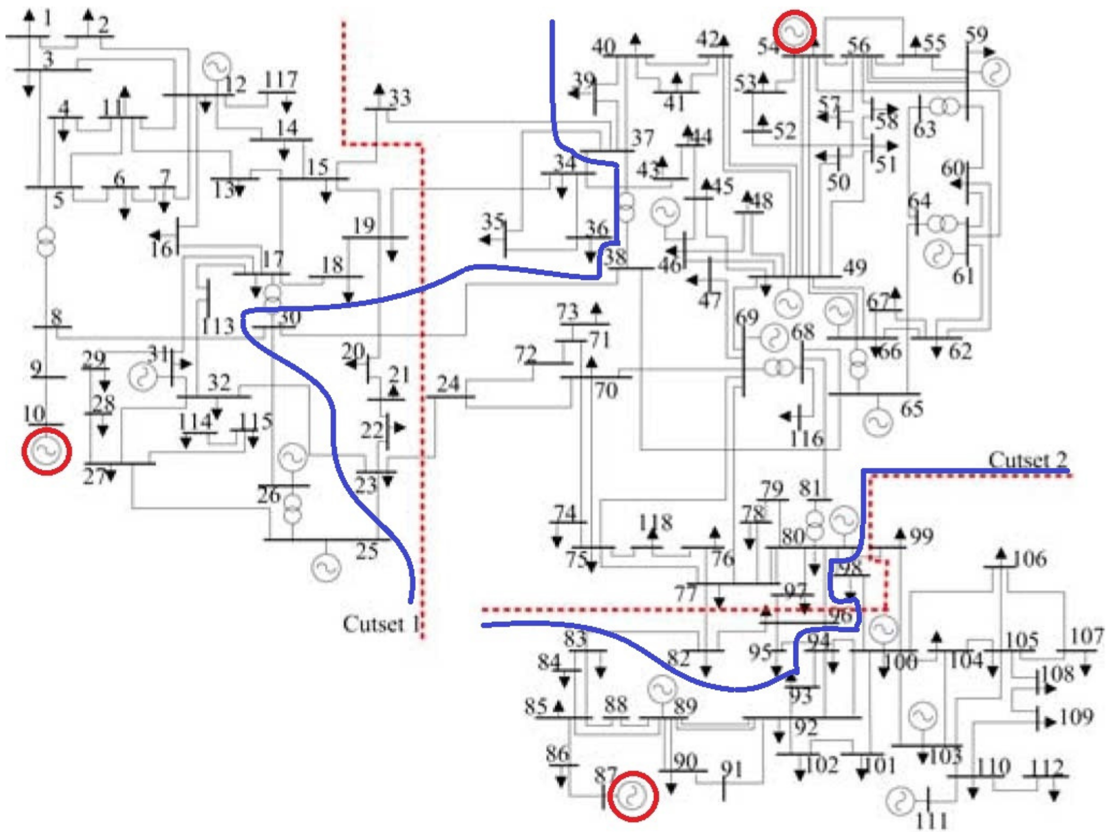



Figure 5.2: IEEE 118-bus system topology. The circled generators at buses {10,54,87} are reference generators identified by Algorithm 5. The solid curves represent the cutset selected by Algorithm 4 with parameter $\xi = 10^{-6}$. The dotted lines represent the cutset selected by the spectral clustering method.

# BIBLIOGRAPHY

[1] North american synchrophasor initiative 2016 meeting agenda. https://www.naspi.org/meetings.

[2] S. S. Ahmed, N. C. Sarker, et al. A scheme for controlled islanding to prevent subsequent blackout. *IEEE Transactions on Power Systems*, 18(1):136–143, 2003.

[3] T. Amraee, A. Soroudi, and A. M. Ranjbar. Probabilistic determination of pilot points for zonal voltage control. *IET Generation, Transmission Distribution*, 6(1):1–10, 2012.

[4] G Andersson, P Donalek, R Farmer, et al. Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance. *IEEE Transactions on Power Systems*, 20(4):1922–1928, 2005.

[5] Niv Buchbinder, Moran Feldman, Joseph Seffi Naor, and Roy Schwartz. Submodular maximization with cardinality constraints. In *25th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1433–1452. SIAM, 2014.

[6] Chi-Tsong Chen. *Linear System Theory and Design*. Oxford University Press, Inc., 1995.

[7] Yonghong Chen. *Development of automatic slow voltage control for large power systems*. PhD thesis, Washington State University, Pullman, WA, 2001.

[8] M. Chilali and P. Gahinet. $H_\infty$ design with pole placement constraints: an LMI approach. *IEEE Transactions on Automatic Control*, 41(3):358–367, 1996.

[9] Joe H. Chow. *Time-Scale Modeling of Dynamic Networks with Applications to Power Systems*, volume 46. New York: Springer-Verlag, 1982.

[10] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran. Minimizing convergence error in multi-agent systems via leader selection: A supermodular optimization approach. *IEEE Transactions on Automatic Control*, 59(6):1480–1494, 2014.

[11] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran. *Submodularity in Dynamics and Control of Networked Systems*. Springer, 2016.

[12] S. Corsi, P. Marannino, N. Losignore, G. Moreschini, and G. Piccini. Coordination between the reactive power scheduling function and the hierarchical voltage control of the EHV ENEL system. *IEEE Transactions on Power Systems*, 10(2):686–694, 1995.

[13] Abhimanyu Das and David Kempe. Submodular meets spectral: greedy algorithms for subset selection, sparse approximation and dictionary selection. In *28th International Conference on International Conference on Machine Learning*, pages 1057–1064. Omnipress, 2011.

[14] J. C. Das. Analysis and control of large-shunt-capacitor-bank switching transients. *IEEE Transactions on Industry Applications*, 41(6):1444–1451, 2005.

[15] B. de Jager and M. van de Wal. Efficient selection of inputs and outputs for robust control. In *38th IEEE Conference on Decision and Control (CDC)*, pages 4517–4522, 1999.

[16] Lei Ding, Francisco M Gonzalez-Longatt, Peter Wall, and Vladimir Terzija. Two-step spectral clustering controlled islanding algorithm. *IEEE Transactions on Power Systems*, 28(1):75–84, 2013.

[17] Ian Dobson, Benjamin A Carreras, Vickie E Lynch, and David E Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103, 2007.

[18] Ian Dobson and Hsiao-Dong Chiang. Towards a theory of voltage collapse in electric power systems. *Systems & Control Letters*, 13(3):253–262, 1989.

[19] Florian Dörfler, Mihailo R Jovanovic, Michael Chertkov, and Francesco Bullo. Sparsity-promoting optimal wide-area control of power networks. *IEEE Transactions on Power Systems*, 29(5):2281–2291, 2014.

[20] Geir E Dullerud and Fernando Paganini. *A course in robust control theory: a convex approach*, volume 36. Springer Science & Business Media, 2013.

[21] Satoru Fujishige. *Submodular Functions and Optimization*, volume 58. Elsevier, 2005.

[22] M. Glavic, D. Novosel, E. Heredia, D. Kosterev, A. Salazar, F. Habibi-Ashrafi, and M. Donnelly. See it fast to keep calm: Real-time voltage control under stressed conditions. *IEEE Power and Energy Magazine*, 10(4):43–55, 2012.

[23] Scott Greene, Ian Dobson, and Fernando L Alvarado. Contingency ranking for voltage collapse via sensitivities from a single nose curve. *IEEE Transactions on Power Systems*, 14(1):232–240, 1999.

[24] Leonard L Grigsby. *Power System Stability and Control*, volume 5. CRC press, 2012.

[25] M. Hajian, M. Glavic, W. D. Rosehart, and H. Zareipour. A chance-constrained optimization approach for control of transmission voltages. *IEEE Transactions on Power Systems*, 27(3):1568–1576, 2012.

[26] M. Hajian, W. Rosehart, M. Glavic, H. Zareipour, and T. Van Cutsem. Linearized power flow equations based predictive control of transmission voltages. In *46th Hawaii International Conference on System Sciences*, pages 2298–2304, 2013.

[27] MH Haque and AHMA Rahim. Identification of coherent generators using energy function. In *IEE Proceedings C (Generation, Transmission and Distribution)*, volume 137, pages 255–260. IET, 1990.

[28] Annissa Heniche and I Kamwa. Control loops selection to damp inter-area oscillations of electrical networks. *IEEE Transactions on Power Systems*, 17(2):378–384, 2002.

[29] Ian Hiskens. IEEE PES task force on benchmark systems for stability controls. *Technical Report*, 2013.

[30] Ali Jadbabaie, Alexander Olshevsky, George J Pappas, and Vasileios Tzoumas. Minimal reachability is hard to approximate. *arXiv preprint arXiv:1710.10244*, 2017.

[31] Richard Arnold Johnson, Dean W Wichern, et al. *Applied Multivariate Statistical Analysis*, volume 4. Prentice-Hall New Jersey, 2014.

[32] Innocent Kamwa, Robert Grondin, and Yves Hébert. Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach. *IEEE Transactions on Power Systems*, 16(1):136–153, 2001.

[33] M. Kezunovic, S. Meliopoulos, V. Venkatasubramanian, and V. Vittal. *Application of Time-Synchronized Measurements in Power System Transmission Networks*. Springer, 2014.

[34] Rajiv Khanna, Ethan Elenberg, Alex Dimakis, Sahand Negahban, and Joydeep Ghosh. Scalable greedy feature selection via weak submodularity. In *Artificial Intelligence and Statistics*, pages 1560–1568, 2017.

[35] Andreas Krause and Daniel Golovin. Submodular function maximization. *Tractability: Practical Approaches to Hard Problems*, 3:19, 2012.

[36] Prabha Kundur. *Power System Stability and Control*. McGraw-Hill New York, 1994.

[37] P. Lagonotte, J. C. Sabonnadiere, J. Y. Leost, and J. P. Paul. Structural analysis of the electrical system: application to secondary voltage control in France. *IEEE Transactions on Power Systems*, 4(2):479–486, 1989.

[38] CH Lee. Solution bounds of the continuous and discrete Lyapunov matrix equations. *Journal of Optimization Theory and Applications*, 120(3):559–578, 2004.

[39] Jon Lee, Vahab S Mirrokni, Viswanath Nagarajan, and Maxim Sviridenko. Maximizing non-monotone submodular functions under matroid or knapsack constraints. *SIAM Journal on Discrete Mathematics*, 23(4):2053–2078, 2010.

[40] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[41] Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. Mingen: Minimal generator set selection for small signal stability in power systems: A submodular framework. In *IEEE 55th Conference on Decision and Control (CDC)*, pages 4122–4129, 2016.

[42] Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. Towards scalable voltage control in smart grid: A submodular optimization approach. In *ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)*, pages 1–10, 2016.

[43] Z. Liu, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. Submodular optimization for voltage control. *IEEE Transactions on Power Systems*, 33(1):502–513, 2018.

[44] Z. Liu, L. Yao, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. Minimal input selection for robust control. In *IEEE 56th Conference on Decision and Control (CDC)*, 2017.

[45] Zhipeng Liu, Andrew Clark, Phillip Lee, Linda Bushnell, Daniel Kirschen, and Radha Poovendran. A submodular optimization approach to controlled islanding under cascading failure. In *8th International Conference on Cyber-Physical Systems*, pages 187–196. ACM, 2017.

[46] Zhipeng Liu, Yao Long, Andrew Clark, Phillip Lee, Linda Bushnell, Daniel Kirschen, and Radha Poovendran. Minimal input selection for robust control. *arXiv preprint arXiv:1712.01232*, 2017.

[47] JA Pecas Lopes, N Hatziargyriou, J Mutale, P Djapic, and N Jenkins. Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities. *Electric power systems research*, 77(9):1189–1203, 2007.

[48] A. Moeini, I. Kamwa, P. Brunelle, and G. Sybille. Open data IEEE test systems implemented in SimPowerSystems for education and research in power grid dynamics and control. In *50th International Universities Power Engineering Conference (UPEC)*, pages 1–6, 2015.

[49] U. Münz, M. Pfister, and P. Wolfrum. Sensor and actuator placement for linear systems based on $H_2$ and $H_\infty$ optimization. *IEEE Transactions on Automatic Control*, 59(11):2984–2989, 2014.

[50] A. Olshevsky. Minimal controllability problems. *IEEE Transactions on Control of Network Systems*, 1(3):249–258, 2014.

[51] Anantha Pai. *Energy function analysis for power system stability*. Springer Science & Business Media, 2012.

[52] F. Pasqualetti, S. Zampieri, and F. Bullo. Controllability metrics, limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems*, 1(1):40–52, 2014.

[53] S. Pequito, S. Kar, and G. J. Pappas. Minimum cost constrained input-output and control configuration co-design problem: A structural systems approach. In *2015 American Control Conference (ACC)*, pages 4099–4105.

[54] Sérgio Pequito, Soummya Kar, and A Pedro Aguiar. On the complexity of the constrained input selection problem for structural linear systems. *Automatica*, 62:193–199, 2015.

[55] Sergio Pequito, Guilherme Ramos, Soummya Kar, A Pedro Aguiar, and Jaime Ramos. On the exact solution of the minimal controllability problem. *arXiv preprint arXiv:1401.4209*, 2014.

[56] Brett Robbins, Christoforos N Hadjicostis, Alejandro D Domínguez-García, et al. A two-stage distributed architecture for voltage control in power distribution systems. *IEEE Transactions on Power Systems*, 28(2):1470–1482, 2013.

[57] J. L. Sancha, J. L. Fernandez, A. Cortes, and J. T. Abarca. Secondary voltage control: analysis, solutions and simulation results for the spanish transmission system. *IEEE Transactions on Power Systems*, 11(2):630–638, 1996.

[58] PW Sauer and MA Pai. Power system steady-state stability and the load-flow jacobian. *IEEE Transactions on Power Systems*, 5(4):1374–1383, 1990.

[59] Carsten Scherer. Controllability and stabilizability. University of Stuttgart, 2013.

[60] Simone Schuler, Ulrich Münz, and Frank Allgöwer. Decentralized state feedback control for interconnected systems with application to power systems. *Journal of Process Control*, 24(2):379–388, 2014.

[61] Nilanjan Senroy, Gerald T Heydt, and Vijay Vittal. Decision tree assisted controlled islanding. *IEEE Transactions on Power Systems*, 21(4):1790–1797, 2006.

[62] Jingdong Su. *A heuristic slow voltage control scheme for large power systems*. PhD thesis, Washington State University, Pullman, WA, 2006.

[63] T. H. Summers, F. L. Cortesi, and J. Lygeros. On submodularity and controllability in complex dynamical networks. *IEEE Transactions on Control of Network Systems*, 3(1):91–101, 2016.

[64] Kai Sun, Da-Zhong Zheng, and Qiang Lu. Splitting strategies for islanding operation of large-scale power systems using OBDD-based methods. *IEEE Transactions on Power Systems*, 18(2):912–923, 2003.

[65] Maxim Sviridenko, Jan Vondrák, and Justin Ward. Optimal approximation for submodular and supermodular optimization with bounded curvature. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1134–1148. SIAM, 2014.

[66] J. Q. Torts and V. Terzija. Controlled islanding strategy considering power system restoration constraints. In *IEEE Power and Energy Society General Meeting*, pages 1–8, 2012.

[67] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie. Minimal actuator placement with optimal control constraints. In *American Control Conference (ACC)*, pages 2081–2086, 2015.

[68] Vasileios Tzoumas, Ali Jadbabaie, and George J Pappas. Minimal reachability problems. In *54th IEEE Conference on Decision and Control (CDC)*, pages 4220–4225, 2015.

[69] Vasileios Tzoumas, Mohammad Amin Rahimian, George J Pappas, and Ali Jadbabaie. Minimal actuator placement with bounds on control effort. *IEEE Transactions on Control of Network Systems*, 3(1):67–78, 2016.

[70] Marc Van De Wal and Bram De Jager. A review of methods for input/output selection. *Automatica*, 37(4):487–510, 2001.

[71] Vaithianathan Venkatasubramanian, Javier Guerrero, Jingdong Su, Hong Chun, Xun Zhang, Farrokh Habibi-Ashrafi, Armando Salazar, and Backer Abu-Jaradeh. Hierarchical two-level voltage controller for large power systems. *IEEE Transactions on Power Systems*, 31(1):397–411, 2016.

[72] Vaithianathan Venkatasubramanian and Yuan Li. Analysis of 1996 western american electric blackouts. *Bulk Power System Dynamics and Control-VI, Cortina dAmpezzo, Italy*, pages 22–27, 2004.

[73] H. Vu, P. Pruvot, C. Launay, and Y. Harmand. An improved voltage control on large-scale power system. *IEEE Transactions on Power Systems*, 11(3):1295–1303, 1996.

[74] X. Wang and V. Vittal. System islanding using minimal cutsets with minimum net flow. In *Power Systems Conference and Exposition. IEEE PES*, pages 379–384 vol.1, 2004.

[75] Bo Yang, Vijay Vittal, and Gerald T Heydt. Slow-coherency-based controlled islanding - a demonstration of the approach on the August 14, 2003 blackout scenario. *IEEE Transactions on Power Systems*, 21(4):1840–1847, 2006.

[76] Haibo You, V. Vittal, and Zhong Yang. Self-healing in power systems: an approach using islanding and rate of frequency decline-based load shedding. *IEEE Transactions on Power Systems*, 18(1):174–181, 2003.

[77] Haibo You, Vijay Vittal, and Xiaoming Wang. Slow coherency-based islanding. *IEEE Transactions on Power Systems*, 19(1):483–491, 2004.

[78] Chien-Ning Yu, Yong T. Yoon, M. D. Ilic, and A. Catelli. On-line voltage regulation: the case of New England. *IEEE Transactions on Power Systems*, 14(4):1477–1484, 1999.

[79] Baosen Zhang, Alejandro D Dominguez-Garcia, and David Tse. A local control approach to voltage regulation in distribution networks. *North American Power Symposium (NAPS)*, pages 1–6, 2013.

[80] Fuzhen Zhang. *Matrix Theory: Basic Results and Techniques*. Springer Science & Business Media, 2011.

[81] S. Zhang and V. Vittal. Design of wide-area damping control robust to transmission delay using $\mu$-synthesis approach. In *IEEE PES General Meeting — Conference Exposition*, pages 1–5, 2014.

[82] Song Zhang and Vijay Vittal. Wide-area control resiliency using redundant communication paths. *IEEE Transactions on Power Systems*, 29(5):2189–2199, 2014.

[83] Yang Zhang and Anjan Bose. Design of wide-area damping controllers for interarea oscillations. *IEEE Transactions on Power Systems*, 23(3):1136–1143, 2008.

[84] Z. Zhang, Y. Wang, E. K. P. Chong, and A. Pezeshki. Subspace selection for projection maximization with matroid constraints. *IEEE Transactions on Signal Processing*, 65(5):1339–1351, 2017.

[85] Qianchuan Zhao, Kai Sun, Da-Zhong Zheng, Jin Ma, and Qiang Lu. A study of system splitting strategies for island operation of power system: a two-phase method based on obdds. *IEEE Transactions on Power Systems*, 18(4):1556–1565, 2003.

[86] Kemin Zhou and John Comstock Doyle. *Essentials of Robust Control*, volume 104. Prentice Hall Upper Saddle River, NJ, 1998.

[87] Kemin Zhou and P. Khargonekar. Stability robustness bounds for linear state-space models with structured uncertainty. *IEEE Transactions on Automatic Control*, 32(7):621–623, 1987.

[88] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1), 2011.

# Appendix A

## LIST OF PUBLICATIONS

1. **Z. Liu**, A. Clark, L. Bushnell, D. Kirschen, and R. Poovendran. "Controlled Islanding via Weak Submodularity", *IEEE Transactions on Power Systems*, Nov 2018.

2. **Z. Liu**, Y. Long, A. Clark, L. Bushnell, D. Kirschen, and R. Poovendran. "Minimal Input and Output Selection for Stability of Systems with Uncertainties", *IEEE Transactions on Automatic Control*, Jun 2018.

3. **Z. Liu**, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. "Submodular Optimization for Voltage Control", *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 502-513, Jan 2018.

4. **Z. Liu**, Y. Long, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. "Minimal Input Selection for Robust Control", in *IEEE 56th Annual Conference on Decision and Control (CDC)*, Melbourne, Australia, Dec 2017.

5. **Z. Liu**, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. "A Submodular Optimization Approach to Controlled Islanding under Cascading Failure", in *ACM 8th International Conference on Cyber-Physical Systems (ICCPS)*, Pittsburgh, USA, Apr 2017.

6. **Z. Liu**, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. "MinGen: Minimal Generator Set Selection for Small Signal Stability in Power Systems: A Submodular Framework", in *IEEE 55th Conference on Decision and Control (CDC)*, Las Vegas, USA, Dec 2016.

7. **Z. Liu**, A. Clark, P. Lee, L. Bushnell, D. Kirschen, and R. Poovendran. "Towards Scalable Voltage Control in Smart Grid: A Submodular Optimization Approach", in *ACM 7th International Conference on Cyber-Physical Systems (ICCPS)*, Vienna, Austria, Apr 2016.

# VITA

Zhipeng Liu was born in Binzhou, Shandong Province, People's Republic of China, 1989. He received the B.S. degree in Mathematics from the City University of Hong Kong in 2013 and the M.S. degree in Applied Mathematics from the University of Washington in 2014. He is a Graduate Fellow of Clean Energy Institute at University of Washington. His research interests include submodular optimization, power system stability, and stabilizing control, .