Scalable (Re)design Frameworks for Optimal, Distributed Control in Power Networks

Xuan Zhang
Somerville College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2015

To my dearest family.

Acknowledgements

"Oxford is always a marvelous place" - this is how I feel when I am going to write this section. I am so lucky and honoured to spend about four years doing my D.Phil research here. It is not just a place to obtain a degree, but a place to experience life with all my dearest friends. Finally, I have the change to express my greatest gratitude to them.

First, I am deeply grateful to my supervisor, Professor Antonis Papachristodoulou, for his guidance and help during these years. He taught me how to conduct research, how to write papers and how to deliver presentations. Most importantly, he offered me freedom and support to pursue a variety of research topics. My gratitude also extends to Professor Dario Bauso at University of Palermo, and Professor Na (Lina) Li at Harvard University. Dario introduced me to the topic of mean-field games for routing problems in networks when he was visiting Oxford on sabbatical. I am extremely thankful for him teaching me about game theory. I visited Lina's group at Harvard from September to December in 2014, where we had developed a reverse- and forward-engineering framework for distributed control in network systems. Most of the material in Chapter 5 is the result of our joint discussions. Thanks a million, Lina, for hosting me at Harvard SEAS.

I would like to sincerely thank Professor Malcolm McCulloch and Ren Kang for sharing brilliant ideas in voltage control and detecting tap changer variations in power systems. Ren and I had many discussions on our joint project from which I had benefited a lot. Also, those attractive "stories" about his entrepreneurship were impressive.

I am truly grateful to my intelligent colleagues in Control Group. Special thanks go to some current and past group members for their help and encouragement: James Anderson, Mohamadreza Ahmadi, Bing Chu, Qifeng Cheng, Sandira Gayadeen, Andreas Harris, Edward Hancock, Richard Mason, Thomas Prescott, Dhruva Raman, Giorgio Valmorbida, Hao Wang, Shi Zhao, Xiaowei Zhao (in alphabetical order), etc.

I also wish to acknowledge an incomplete list of my friends at Oxford (it is really amazing that I have so many friends here): Biao Cai, Yiming Dong, Lu Feng, Cheng Ge, Jianan Hu, Mengchen Hu, Bo Lan, Qian Li, Bo Pang, Bohang Song, Tan Sui, Fan Xu, Siqi Ying, He Zhang (in alphabetical order), etc.

Last but not least, I would like to manifest my deepest gratitude to my dearest family, my mother Yunxiang Qin and my father Xiutian Zhang, for their unconditional love and support. I owe them much more than I would ever be able to express. I dedicate this thesis to them.

Abstract

In this thesis, we develop scalable frameworks to (re)design a class of large-scale network systems with built-in control mechanisms, including electric power systems and the Internet, in order to improve their economic efficiency and performance while guaranteeing their stability and robustness.

After a detailed introduction relating to power system control and optimization, as well as network congestion control, we turn our attention to merging primary and secondary frequency control for the power grid. We present modifications in the conventional generation control using a consensus design approach while considering the participation of controllable loads. The optimality, stability and delay robustness of the redesigned system are studied. Moreover, we extend the proposed control scheme to (i) networks with more complexity and (ii) the case where controllable loads are involved in the optimization. As a result, our controllers can balance power flow and drive the system to an economically optimal operating point in the steady state.

We then study a real-time control framework that merges primary, secondary and tertiary frequency control in power systems. In particular, we consider a transmission level network with tree topology. A distributed dynamic feedback controller is designed via a primal-dual decomposition approach and the stability of the overall system is studied. In addition, we introduce extra dynamics to improve system performance and emphasize the trade-off when choosing the gains of the extra dynamics. As a result, the proposed controller can balance supply and demand in the presence of disturbances, and achieve optimal power flow in the steady state. Furthermore, after introducing the extra dynamics, the transient performance of the system significantly improves.

A redesign framework for network congestion control is developed next. Motivated by the augmented Lagrangian method, we introduce extra terms to the Lagrangian, which is used to redesign the primal-dual, primal and dual algorithms. We investigate how the gains resulting from the extra dynamics influence the stability

and robustness of the system. Moreover, we show that the overall system can achieve added robustness to communication delays by appropriately tuning these gains. Also, the meaning of these extra dynamics is investigated and a distributed proportional-integral-derivative controller for solving network congestion control problems is further developed.

Finally, we concentrate on a reverse- and forward-engineering framework for distributed control of a class of linear network systems to achieve optimal steady-state performance. As a typical illustration, we use the proposed framework to solve the real-time economic dispatch problem in the power grid. On the other hand, we provide a general procedure to modify control schemes for a special class of dynamic systems. In order to investigate how general the reverse- and forward-engineering framework is, we develop necessary and sufficient conditions under which an linear time-invariant system can be reverse-engineered as a gradient algorithm to solve an optimization problem. These conditions are characterized using properties of system matrices and relevant linear matrix inequalities.

We conclude this thesis with an account for future research.

Contents

List of Figures

Chapter 1

Introduction

The 21st century society depends on a number of technological networks, such as the power grid, the Internet, transportation systems, etc. The efficient and robust operation of these network systems plays a key role in economic development and environmental sustainability.

Because these network systems are usually operated in an uncertain environment and with incomplete information, to ensure their economic efficiency, stability and robustness, the conventional operation is usually divided into two or more different time-scales. At a slow time-scale, efficient nominal operating points are determined using optimization methods with predictions of future uncertainties, e.g., disturbances. At a faster time-scale or in real time, the optimal points are tracked and stability of the system is ensured using control techniques. However, as *the uncertainties fluctuate faster and by a larger amount*, this time-scale separation framework could induce economic inefficiency, poor robustness, and even instability.

On the other hand, one key problem in the control and optimization of these network systems is *scale*. Given that the topology is many times changing or unknown, as the network becomes large, or of arbitrary size, the complexity of traditional tools and algorithms increases while the system (with existing control mechanisms) could be facing deteriorated performance, poor robustness, and even instability.

In this thesis, concentrating on a class of large-scale network systems including electric power systems and the Internet, we develop *scalable* frameworks for modifying/redesigning existing system dynamics and built-in control mechanisms to improve their economic efficiency and performance while guaranteeing their stability and robustness. As a result, in the presence of uncertainties, the redesigned systems can track the efficient operating points automatically in a distributed (i.e., communication with neighbours is allowed) or decentralized (i.e., no communication with neighbours allowed), and closed-loop (i.e., no information of uncertainty is needed) manner.

1.1 Electric Power Systems

1.1.1 Overview and Problem Descriptions

The electric power system is the largest, most advanced and most expensive engineered system in the twentieth century [3]. It consists of three main constituents: generation (i.e., energy supply), transmission and distribution, and consumption (i.e., energy demand) [4]. Each constituent has different components. Conventional generation contains synchronous generators, exciting systems, automatic voltage regulators, turbine-governor systems, etc., and produces electricity by converting mechanical energy into electrical energy. The electricity is transmitted over long distances to end-users via transmission level networks (which operate at high voltages) and distribution level networks (which operate at low voltages). These networks consist of overhead power lines, underground cables, transformers and substations. The demand side contains a variety of energy consumers, e.g., lighting, heating, electric motors, etc. An individual consumer may change its demand fast and frequently, however, after aggregating individual demands appropriately, the power consumption can become smooth and more or less predicable over hours, days, months and seasons [4]. All these components are interconnected and coupled to each other, forming a large-scale network system.

To date, various controls and actions are implemented in the current power grid in a layered/hierarchical way, as shown in Figure 1.1 based on [5] (detailed introduction

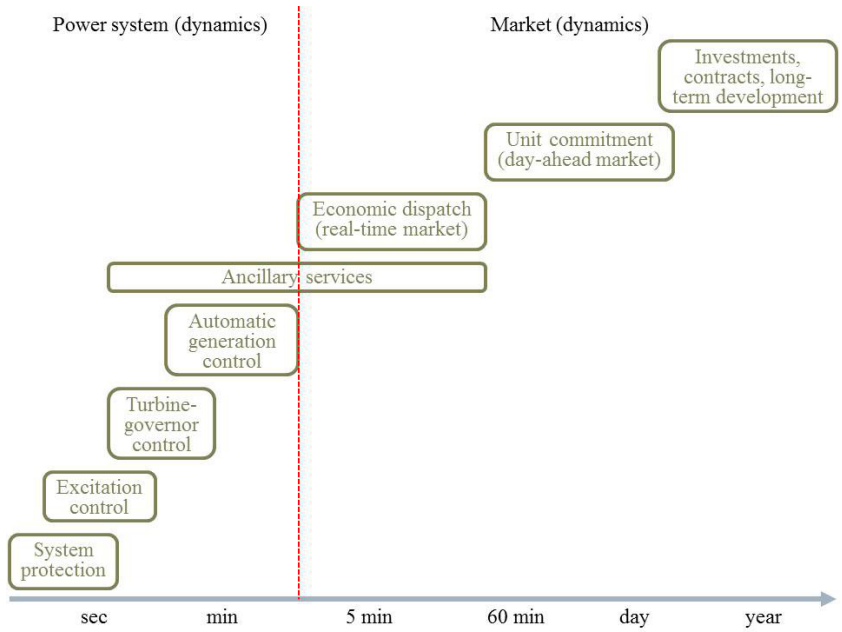

Figure 1.1: Time-scale decomposition of controls and actions in the power grid.

of each control and action can be found in [5, 6, 7]). Each layer is characterized by its operational/control objectives, corresponding to one time-scale. Generally speaking, there are mainly two types of controls and actions: the fast time-scale controls and actions for the power grid (i.e., a physical network system), and the slow time-scale controls and actions for the electricity market (i.e., a system that enables electricity to be bought, sold and traded as a commodity). The time-scale decomposition in system dynamics and market dynamics provides a premise to simplify problems when studying power systems, especially in control and optimization. For example, if considering long term contracts (slow time-scale) between a generation company and a factory, the detailed controls and actions in the transmission network (fast time-scale) are neglected. Such separation works well when the time-scales of the two layers are significantly different. Even if the difference is less obvious, e.g., between seconds and minutes, it is still acceptable because of high predictability (in future states) and low uncertainty (in disturbance injection) in the system [5].

However, the power grid is currently undergoing several fundamental changes such as increasing penetration of renewable and distributed energy resources, the proliferation of electric vehicles, the active participation of end-users, and the rapid deployment of sensing, communication and computing infrastructures [8]. The power network is in transition from a centralized structure to one that is more distributed, open, and autonomous. These trends provide tremendous opportunities for improvements in sustainability, efficiency, power quality and reliability. But they also present daunting technical challenges, particularly those imposed by non-dispatchable and volatile renewable generation, and the large number of active end-points in the future system. For instance, driven by the goal of sustainability, the penetration of renewable and distributed energy resources including solar power, wind power, tidal power and so on, is significantly increasing. Different from conventional generation, the availability of renewable energy is difficult to predict, and can fluctuate fast and by a large amount, as illustrated in Figures 1.2-1.3. Another big change is the growing use of electric vehicles. Electric vehicles are becoming more and more popular (see Table 1.1 [9]) due to their potential societal benefits, i.e., reducing petroleum consumption and greenhouse gas emissions, and improving air quality [10, 9]. They can be connected to the power grid for charging anytime and anywhere, however, these can cause unexpected disturbances to the system if not properly controlled [11, 12]. Because of the volatilities and uncertainties resulting from these new constituents, parts of the time-scale separation in the electricity market and power system dynamics are fading away, especially in the real-time market level (time-scale of usually 5 minutes

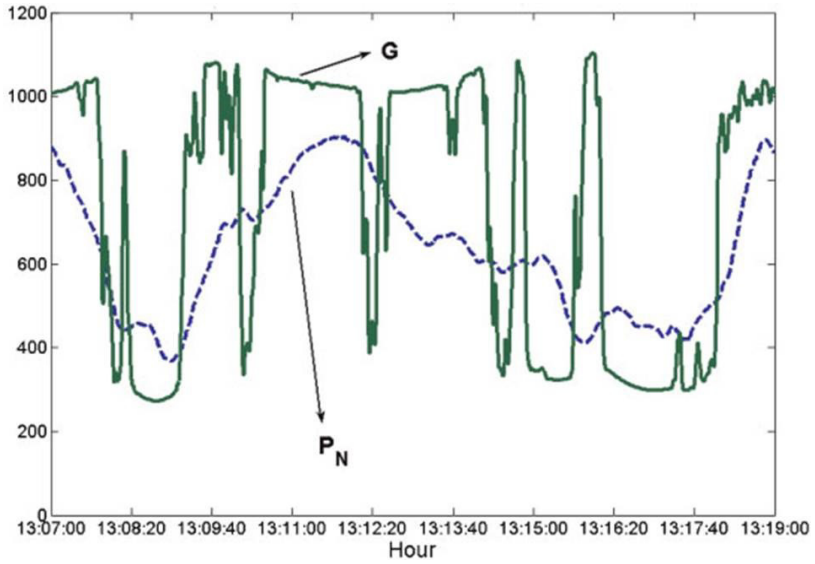

Figure 1.2: A typical solar power plant output profile from [1] (© 2011 John Wiley and Sons): Irradiance G (W/m^2) and output power P_N (normalized and scaled by a factor of 1000) recorded at Milagro site during a 15-min period on August 12th, 2008. The power curve is smoother than the irradiance due to the big size of the PV plant and the short sampling time.

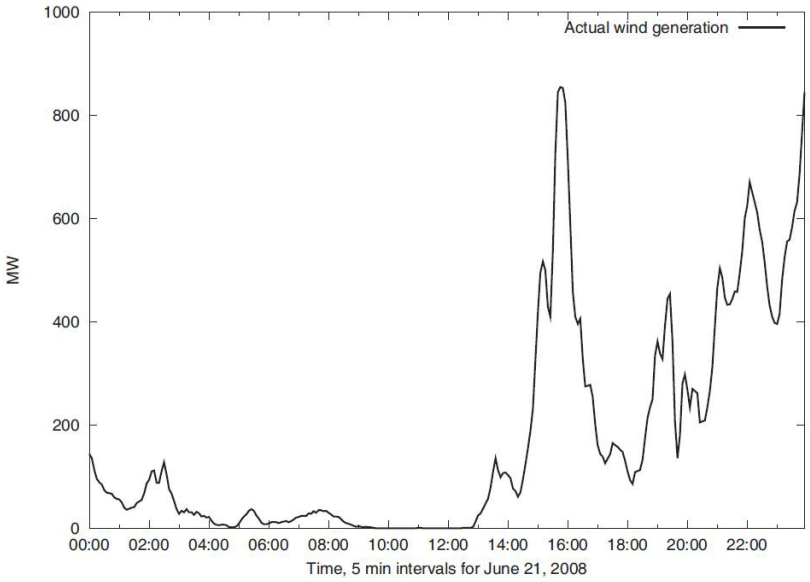

Figure 1.3: A typical wind farm output profile from [2] (© 2010 IEEE): The large positive and negative ramps in the hour just before and after 3:45pm on June 21, 2008 (data from wind farms in the Bonneville Power Administration region around the Columbia Basin).

and up). As a result, it would be more and more difficult for the current hierarchical control architecture to provide effectiveness and robustness to the real-time market,

Table 1.1: Number of cars in the UK stock by type under a stretch scenario: Electric Vehicles (EVs) and Internal Combustion Engine Vehicles (ICEVs). The stretch scenario is driven by a milestone for cars in line with the UK meeting a 42% CO_2 emission cut by 2020 [9] (© 2010 Element Energy Limited).

		2010	2015	2020	2030
Number of cars in UK car parc	EVs	11370	297400	4227300	26261200
	ICEVs	28980800	30173600	27798000	9114700
Percentage of UK car stock	EVs	0.04%	1.0%	13.2%	74.2%
	ICEVs	99.96%	99.0%	86.8%	25.8%

which separates power system dynamics from the real-time market mechanism, i.e., solving a static Optimal Power Flow (OPF)/Economic Dispatch (ED) problem only once within a certain period of time in a centralized way.

A main topic of this thesis is to redesign the conventional frequency control/active power control for power systems to adapt for those changes (this takes advantage of the weak physical coupling between active power flows and bus voltage magnitudes in transmission level networks to separate frequency control/active power control and voltage control/reactive power control [3, 7]). Conventional frequency control is designed and implemented in a hierarchical way, consisting of primary, secondary and tertiary control [6, 4]. When disturbances occur, primary frequency control operates at a time-scale of seconds to adjust the mechanical power input of generators based on local frequency deviations, which is implemented through turbine-governor systems equipped within the generation (there is a proportional controller in a turbine-governor system for frequency regulation). Meanwhile, there are also power demand deviations due to frequency deviations in the frequency-dependent loads. Then secondary frequency control, also known as Automatic Generation Control (AGC)/tie-line bias control, takes place at a time-scale of 10 seconds to minutes to restore nominal frequency and scheduled interchanges of tie-lines (there is an integral controller in AGC for frequency regulation): in an interconnected power system consisting of a number of control areas, AGC is implemented in each area in a centralized way, i.e., the active power regulation is allocated to the generators in each area via participation factors [13]. Lastly, tertiary control, also known as ED, operates at a time-scale of minutes to determine the economically-efficient nominal operating point of the mechanical power generation and power interchanges of tie-lines. ED is usually performed in the real-time market by solving an OPF problem, which aims

to minimize the operating cost subject to practical constraints. To conclude, the structure of the conventional frequency control is shown in Figure 1.4.

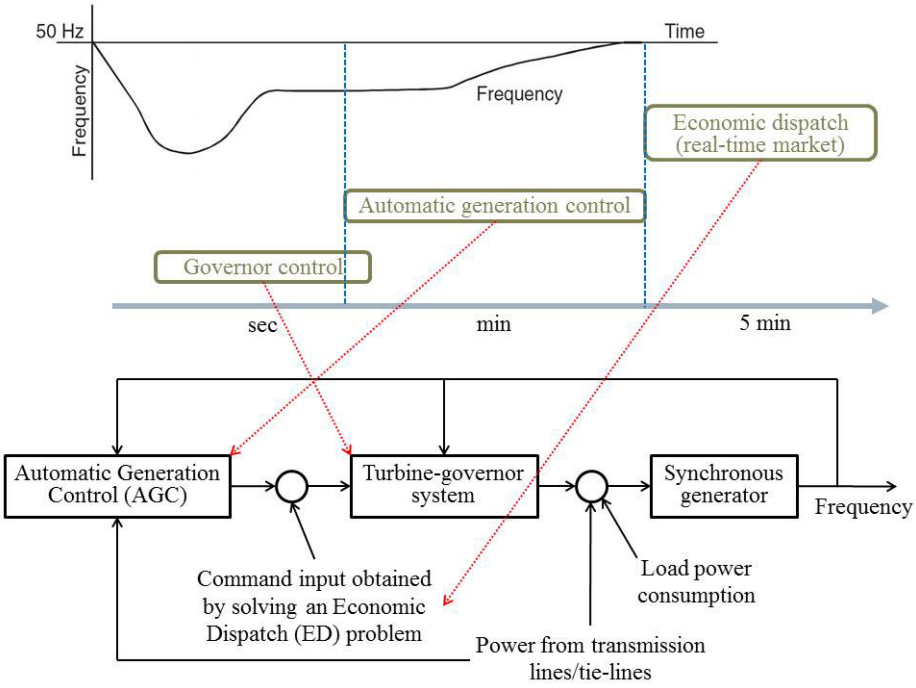

Figure 1.4: The structure of the conventional frequency control in power systems.

To redesign the conventional frequency control, two factors are important. The first one is the design of demand response schemes for (controllable) loads/(responsive) users, e.g., air conditioners, washers, electric vehicles, batteries, heaters, lights, etc. Demand response is viewed as a "price-based" mechanism that will encourage consumers to modify their power consumption when it is most difficult for the network to achieve a balance between supply and demand [14]. It can not only reduce peaks and shift loads for economic benefits, but also improve stability and reduce operating reserves by adapting elastic loads to intermittent and fluctuating renewable generation [15]. By introducing demand response, frequency control can be implemented in both the supply and demand side rather than only depend on generation control as in the traditional system. The second factor is to break the time-scale separation between the power system dynamics and the real-time market mechanism, i.e., to realize real-time (a time-scale of seconds to minutes) economic optimization under exogenous disturbances, e.g., renewable energy penetration. This scenario was first studied in [16]. It was shown that the stability region of the coupled system was very different from that of the market-only system, and therefore, the feedback signal (marginal prices derived in the real-time market) must be properly designed to

maintain the stability of the interconnected system. What needs to be emphasized is that after merging the different time-scales, demand response is not restricted to single price-based control but can include all intentional modifications in the power consumption of end-users. In the following subsection, we will review recent research relating to demand response and the redesign of conventional frequency control.

1.1.2 Literature Review

The idea of demand response was first suggested by Fred Schweppe and his co-workers in the 1980s [17, 18]. This concept was then extended and developed by researchers such as Hogan [19, 20] by introducing market and economic analysis, and Alvarado [21, 22, 16] by considering market stability. Nowadays, it is recognized as a key mechanism for ensuring reliability in the face of uncertainty in the future power network. However, demand response has not been widely deployed in practice. Once the reason was the lack of supporting information infrastructures in the power grid, while the situation is being gradually improved due to the increasing deployment of advanced sensing, communication and computing equipments. The current desiderata is to develop suitable control paradigms which can realize the full potential of demand response. On the other hand, power systems are large-scale and complicated. Because of the increase in uncertainty and variability, conventional centralized control for the power network becomes more costly and less efficient. As a consequence, distributed and decentralized control frameworks are necessary to achieve system-wide efficiency, reliability and robustness under those fundamental changes in the power network.

Driven by the goals of scalability and economic efficiency, recent work has focused on the redesign of the conventional frequency control that combines different layers/time-scales together in either distributed, decentralized or centralized manner, and some of them have included the design of demand response schemes. For example, a novel control scheme for achieving optimal power balancing and congestion management in electrical power systems via nodal prices was presented in [5], extending [16] to a more practical scenario. In [23, 24] a fast-acting decentralized load control scheme was presented for primary frequency regulation in power networks, and its implementation was studied in [25, 26, 27]. This load control scheme was then modified by considering more completed power flow constraints that preserved grid frequency [28] and interchanges of tie-lines [29], and was extended by adding decentralized generator-side primary frequency control [30]. Through combining primary and

secondary frequency control, two different distributed control schemes for both generators and controllable loads were explored in [31, 32] to achieve real-time economic efficiency. Another active research area is to redesign conventional AGC by incorporating ED. For instance, [33, 34, 35] considered distributed integral control and distributed Proportional-Integral (PI) control for generators to balance power systems after step disturbances of power loads and minimize generation cost. Moreover, sufficient conditions to guarantee the stability of the closed-loop system were proposed. In [13] the connections between AGC and ED were studied from an optimization viewpoint and a distributed control approach was presented which modified AGC by incorporating ED automatically and dynamically. In addition, the concept of "breaking" the hierarchical control architecture is applied to microgrids (i.e., low-voltage distribution level networks). Representative work includes [36, 37] in which a distributed-averaging proportional-integral controller was presented to ensure stability and power sharing in droop-controlled microgrids (where primary control and secondary control were combined), and [38] in which a centralized-averaging proportional-integral controller [39] was compared with the distributed-averaging proportional-integral controller in lossy and partially-controlled networks. Furthermore, [40, 41] explored different strategies for primary droop control, secondary frequency control and tertiary control (i.e., solving an OPF problem), and showed that the adoption of droop control was necessary and sufficient to achieve economic optimization, i.e., the time-scale separation of conventional hierarchical control strategies in microgrids could be broken.

1.1.3 Our Approach and Contributions

In the first two chapters of this thesis, we redesign the conventional frequency control by breaking the time-scale separation and considering the participation of loads. Based on [42, 43], Chapter 2 combines primary and secondary frequency control. We first present modifications in AGC where each bus in the power network represents a control area/an aggregated bus. The redesign methodology is via a consensus approach, and loads are used to improve the performance of the overall system under proportional control. We then extend the redesign to (i) arbitrary transmission level networks, i.e., each bus is not a control area anymore, and (ii) the case where loads are involved in ED. In all three scenarios, optimality, stability and delay robustness of the redesigned dynamics are studied. The performance of the proposed controllers is illustrated by several numerical examples. To summarize, this chapter shares some similarities with [34, 13] in that the economic efficiency of generator-side control is taken into account in the redesign and the resulting controllers have distributed PI

structures, while the difference is that both the load-side participation and the delay effect (we will show the delay-independent stability of the system) are considered.

Based on [44, 45, 46], Chapter 3 merges primary, secondary and tertiary frequency control through a primal-dual decomposition approach. We first present a distributed control architecture that can achieve real-time OPF in power networks under exogenous disturbances. Different from conventional market-based control, the feedback signals in our control architecture are not prices from a centralized market but information and feedback signals flowing between neighbouring components. To show how the design methodology and stability analysis can be performed, we consider a transmission level network with tree topology. Moreover, we introduce extra dynamics in the redesign to improve the transient behaviour of the closed-loop system. Numerical experiments illustrate that the proposed controller balances the power flow in the network quickly, and results in OPF in the steady state. Furthermore, after adding the extra dynamics, the transient performance of the system improves significantly. Similar to [31, 32], this chapter considers nonlinear active power flow equations which are more practical than linear equations used in conventional DC OPF problems. Also, the design framework can be immediately applied to droop-controlled microgrids [40, 41] for achieving ED with both equality and inequality constraints.

The contributions of the first two chapters are summarized as follows. Firstly, we present two systematic methods to redesign conventional generation control and design load behaviours for power networks to achieve real-time ED in the presence of uncertainties. Secondly, the resulting control schemes are implemented in either a distributed or decentralized and closed-loop manner, guaranteeing system stability and robustness, and leading to improved performance. Lastly, the proposed design frameworks can be scaled to networks of large size and more complexity.

1.2 Network Congestion Control

1.2.1 Literature Review and Problem Descriptions

As the second topic in this thesis, we propose a redesign framework for typical network congestion control algorithms to achieve improved performance and added robustness. The problem of network congestion control focuses in the design of the Transmission Control Protocol (TCP) and Active Queue Management (AQM) schemes for the Internet, whose aim is to make sure that users can get a share of the available bandwidth in a fair manner while ensuring that available bandwidth constraints are not violated. Violating capacity constraints was a problem in the mid-1980s in the early days of

the Internet: users were increasing their bandwidth until the capacity constraints were violated, congestion occurred and all packets on the network were dropped; this resulted in retransmissions of lost packets which led to more lost packets again. In [47], the first congestion control algorithm to alleviate this problem was proposed driven mainly by intuition, however, the efficiency of the protocols was not addressed. Ten years later, [48] presented two classes of congestion control algorithms that interpreted the resulting protocols as the solutions of distributed convex optimization problems which described the scenario of resource allocation in networks. This work allowed researchers to explain the profound scalability and robustness of congestion control protocols, and moreover enabled the design of algorithms in a rational fashion for better efficiency and fairness. As a result, network congestion control became an extremely active research topic in the past two decades [49, 50, 51].

Although many control schemes have been designed and applied to solve the network congestion control problem, there are still many open research issues [51, 52], e.g., global stability of more accurate models, trade-offs among system performance, robustness and complexity of controllers, etc. Recently, some work focused on redesigning dynamics to solve the network congestion control problem. In [53], the authors proposed a unified passivity framework as a tool to analyze existing congestion control algorithms and develop new control schemes. Also, [54] linked available congestion control algorithms with optimal control theory by introducing proper utility functionals so that the transient performance of the overall system could be taken into account when designing controllers. In [55], delay robustness of the available congestion control algorithms using Lyapunov-Krasovskii functionals was investigated and stability conditions for networks of arbitrary size with heterogeneous delays and nonlinear dynamics were derived. In [56], the primal-dual gradient dynamics were revisited with Krasovskii's method and LaSalle's invariance principle, and applications to cross-layer network optimization were presented. In [57], the authors proposed a framework from the control system viewpoint to analyze and design control dynamics for both centralized and distributed optimization problems. Recently in [58, 59], a distributed Newton-type second order algorithm for network utility maximization problems was proposed, with a superlinear convergence rate in terms of primal iterations. Other techniques such as event-triggered optimization [60] and Alternating Direction Method of Multipliers (ADMM) [61] were also applied to solve network congestion control problems.

1.2.2 Our Approach and Contributions

Based on [62, 63], Chapter 4 of this thesis proposes a method for redesigning existing congestion control algorithms at the level of fluid-flow models [64], in order to improve the transient behaviour, and provide robustness to uncertainties in the network structure and communication constraints. Our redesign process involves the introduction of simple terms in the Lagrangian (different from ADMM, these terms are irrelevant to the constraints of the optimization problem), leading to simple extra dynamics that however maintain the distributed structure for the overall system. We investigate the influence of the gains resulting from the extra dynamics on system stability and robustness to time delays. In addition, we study the meaning of the extra dynamics and further introduce distributed Proportional-Integral-Derivative (PID) control actions for solving network congestion control problems. Finally, we provide three illustrative examples to show the effectiveness of the proposed redesign framework.

This chapter highlights the idea of re-engineering existing systems and protocols based on optimization methods. Motivated by the case of modifying network congestion control algorithms, we propose a reverse- and forward-engineering framework for control redesign for a class of network systems to achieve optimal steady-state performance in Chapter 5 based on [65].

1.3 Reverse- and Forward-engineering for Control Redesign

The concept of reverse- and forward-engineering for system re-engineering was addressed in [51]: in general, "starting from a given protocol originally designed based on engineering heuristics, reverse-engineering discovers the underlying mathematical problems being solved by the protocols; forward-engineering based on the insights obtained from reverse-engineering then systematically improves the protocols". Following this concept, the objective of Chapter 5 is to modify/redesign the existing system dynamics and built-in control mechanisms in a network system so that the closed-loop system can track the efficient operating points automatically. We refer to this control as *optimal steady-state control*. The idea of using a dynamic system to track an implicitly defined optimal point originated in [66]. Also, this problem was recently studied in [67, 68, 69], in which the controllers were designed based on a dynamic extension of the Karush-Kuhn-Tucker (KKT) condition, a method of saddle point flows with backstepping, and a dual ϵ-subgradient method respectively.

Different from the recent work, we propose a reverse- and forward-engineering framework which consists of two steps: firstly, we reverse-engineer a dynamic system as a gradient algorithm to solve an optimization problem. Secondly, we use a forward-engineering approach to systematically design controllers or modify the existing control mechanisms. As a result, the redesigned system can automatically track the optimal solution of a predefined optimization problem. Under this framework, the resulting controller (i) has a distributed/decentralized structure and can be implemented in a closed-loop manner; (ii) respects the system operating constraints; (iii) ensures an efficient and reliable network operation. Recent research in frequency control of power grids and Internet congestion control has successfully demonstrated that these systems can be re-engineered using a reverse- and forward-engineering framework [13, 24, 70, 71, 51]. However, there is not much work investigating the general network systems – how general the reverse- and forward-engineering approach is. Specifically, (i) what kind of systems can be reverse-engineered; (ii) if a system can be reverse-engineered, how to use forward-engineering to do control (re)design. Chapter 5 will answer both of these questions, concentrating on Linear Time-Invariant (LTI) systems.

In the first half of Chapter 5, we provide a detailed example of using reverse- and forward-engineering to solve the real-time economic dispatch problem in power systems, which considers more realistic operating constraints than the scenario discussed in Chapter 2. The problem studied here actually shares some similarities with [29] while the design approaches are completely different. In the second half of this chapter, we generalize this approach to a class of large-scale network systems. In order to investigate how general this framework is, we establish necessary and sufficient conditions under which a linear dynamic system can be reverse-engineered as a gradient algorithm to solve an optimization problem. These conditions are characterized using properties of system matrices and relevant linear matrix inequalities. Lastly, we use the IEEE 14-bus network to demonstrate the effectiveness of our framework.

1.4 Connections between the Chapters

On the topic relating to power systems, according to the ideology of breaking the time-scale separation, Chapter 2 focuses on merging and redesigning primary and secondary frequency control in the power grid via a consensus approach. Chapter 3 goes one step further and considers merging and redesigning primary, secondary and tertiary frequency control via a primal-dual decomposition approach. In the first half

of Chapter 5, the problem considered in Chapter 2 is revisited (where more detailed operating constraints are considered) via reverse- and forward-engineering.

On the topic relating to network congestion control, Chapter 4 proposes a method to modify existing congestion control algorithms to improve their performance. This method can also be applied to improve the behaviour of the designed controllers in Chapters 2 and 3, as illustrated in Section 3.3. On the other hand, the second half of Chapter 5 generalizes this redesign method and links with the concept of reverse- and forward-engineering. The topic of optimal steady-state control using reverse- and forward-engineering is then studied in detail.

Lastly, the thesis is complemented with appendices on Lyapunov stability, synchronization in oscillator networks with non-homogeneous delays, saddle point dynamics and primal-dual gradient dynamics, classical loop shaping in Multiple-Input-Multiple-Output (MIMO) systems, and the proof of Vinnicombe's Lemma [72] which is frequently used in this thesis.

1.5 Notation

The notation throughout this thesis is as follows.

$|\mathcal{X}|$ is the cardinality of a set $\mathcal{X}$.

$x \in \mathbb{R}^m$ is a column vector in an m-dimensional Euclidean space $\mathbb{R}^m$, and $[x]_i$ denotes its i^{th} entry. x^T (x^H) is the transpose (conjugate transpose) of x. $x \succeq 0$ ($x \succ 0$) denotes that all components of a vector x are non-negative (positive). $\nabla_x f(x)$ or $\frac{\partial f}{\partial x}$ is the gradient (as a column vector) of a scalar function $f \in \mathcal{C}^1$ with respect to x, where $\mathcal{C}^n$ is the class of functions that are n times continuously differentiable. $\nabla_x^2 f(x)$ is the Hessian matrix of a scalar function $f \in \mathcal{C}^2$ with respect to x.

$\dot{x}$ denotes the derivative of a state variable/vector x with respect to time t, i.e., $\dot{x} = \frac{d}{dt}x(t)$. x^0 denotes the nominal value of a state variable/vector x. x^* denotes the equilibrium of a state variable/vector x.

$X \in \mathbb{R}^{m\times n}$ is an $m \times n$ real matrix. $\text{diag}\{\star\}$ denotes a diagonal matrix with corresponding entries $\star$ on the main diagonal, and $\text{diag}(x)$ is a diagonal matrix whose entries are the elements of a vector x. $\mathbf{1}$ ($\mathbf{0}$) denotes a matrix of ones (zeros) with an appropriate dimension determined by the context. I_m denotes an identity matrix of size $m \times m$. $\text{tr}(X)$ is the trace of a square matrix X. $X \succeq 0$ ($X \succ 0$) denotes that a symmetric matrix X is positive semi-definite (positive definite). $\overline{\sigma}(X)$ and $\underline{\sigma}(X)$ denote the maximum and minimum nonzero singular values of a matrix X respectively

(the singular values are the square roots of the eigenvalues of $X^H X$). $\mathrm{eig}(X)$ and $\rho(X)$ denote the spectrum and spectral radius of a square matrix X respectively.

$\mathcal{C}([a,b],\mathbb{R}^m)$ denotes the Banach space of continuous functions mapping the interval $[a,b]$ into $\mathbb{R}^m$ with the topology of uniform convergence. The norm on $\mathcal{C}([a,b],\mathbb{R}^m)$ is defined as $\|\phi\| = \sup_{a\leq\theta\leq b} |\phi(\theta)|$ where $|\cdot|$ is a norm in $\mathbb{R}^m$. For $\sigma \in \mathbb{R}$, $\tau \in [0,+\infty)$ and $a \geq 0$, $x \in \mathcal{C}([\sigma-\tau,\sigma+a],\mathbb{R}^m)$, and $t \in [\sigma,\sigma+a]$, define $x_t \in \mathcal{C}([\sigma-\tau,\sigma+a],\mathbb{R}^m)$ as $x_t(\theta) = x(t+\theta)$, $\theta \in [-\tau,0]$.

The positive projection of a continuous variable x is $\psi(x) = \max\{0,x\}$. $(h(y))_x^+$ denotes the positive projection of a function $h(y)$ on a variable x where

$$(h(y))_x^+ = \begin{cases} h(y) & \text{if } x > 0 \\ \max(0,h(y)) & \text{if } x = 0 \end{cases}.$$

Finally, in Chapter 4, j stands for the imaginary unit.

Chapter 2

Merging Primary and Secondary Frequency Control in Power Systems via a Consensus Approach

Due to increasing uncertainties resulting from renewable energy penetration and variability in both supply and demand, control and economic optimization for power networks will need to run on faster time-scales. Moreover, distributed (i.e., communication with neighbours allowed) and decentralized (i.e., no neighbour communication allowed) control architectures are necessary as power systems are large-scale networks with a lot of complexity, which makes centralized control costly, inefficient and hard to implement.

In this chapter, we merge conventional primary and secondary frequency control for power systems with (frequency-independent) controllable loads [73], i.e., redesign the generation control/tie-line bias control [6] and consider controllable loads participating in active power regulation, in order to improve the economic efficiency in the presence of uncertainties. We first obtain the state-space model of a transmission level network under exogenous disturbances, where each bus represents a control area/an aggregated bus. We then formulate an optimization problem relating to generator active power regulation. The redesign methodology is based on a consensus approach, and controllable loads are used to improve the performance of the overall system. The optimality, stability and delay robustness of the redesigned dynamics are studied. We next extend the redesign to cases where: (i) each bus in the network denotes either a generator bus (such a bus contains only one synchronous generator/generation unit) or a load bus (such a bus contains only one load which is an aggregation of a certain amount of loads/users at the bus it is connected to; alternatively speaking, each load corresponds to a distribution level network); (ii) controllable loads are involved in the

optimization problem related to active power regulation. Finally, the performance of the redesigned control schemes is illustrated by numerical examples.

2.1 Problem Setup

2.1.1 Network Model

Consider a transmission level network with arbitrary topology, described by a connected directed graph $(\mathcal{N}, \mathcal{E})$. Here $\mathcal{N}$ is the set of buses/areas and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of transmission lines connecting the buses/areas. Each bus locally contains both a synchronous generator/generation unit and a load which is an aggregation of a certain amount of loads/users at the bus it is connected to. We number the buses $1, \cdots, n$ ($\mathcal{N} = \{1, \cdots, n\}$) and the transmission lines $1, \cdots, p$ corresponding to a lexicographic ordering where $p = |\mathcal{E}|$. Define an orientation from bus i to bus j if $(i, j) \in \mathcal{E}, i < j$. View all buses as voltage sources and let $v_i \angle \delta_i, i \in \mathcal{N}$ be the voltage of each bus, where v_i is the voltage magnitude and δ_i is the voltage phase angle with respect to the rotating framework of nominal frequency ω_i^0, e.g., $2\pi \times 50$ rad/s in Europe. Assume that the network is working around a nominal operating point which is determined by an Economic Dispatch (ED) problem at a more slower time-scale. We then make a related assumption.

Assumption 2.1 *All bus voltage magnitudes are fixed, i.e., $v_i, i \in \mathcal{N}$ are constant. The resistance of transmission lines is negligible. Reactive power injections and flows are omitted.*

This assumption is reasonable since: (i) we consider the network working around a nominal operating point, which is determined by an ED problem at a more slower time-scale; (ii) realistic transmission networks are very close to being lossless [6, 7]. Similar assumptions are also used in [24, 13].

Let $\omega_i = \dot{\delta}_i, i \in \mathcal{N}$. Load power consumption may depend on either frequency or voltage magnitude, or both [3]. Since bus voltage magnitudes are fixed under Assumption 2.1, we can distinguish active power loads into three types: frequency-dependent loads whose power consumption is $d_i^{(1)} + D_i^{(1)}\omega_i$ where $D_i^{(1)} > 0$, frequency-independent controllable loads (heaters, dryers, lights, etc.) whose power consumption is P_{L_i}, and frequency-independent uncontrollable loads whose power consumption is $d_i^{(2)}$, $i \in \mathcal{N}$ [3, 24]. The dynamics at the buses/areas are given by the swing equations [74]

$$M_i\dot{\omega}_i + D_i^{(0)}\omega_i = P_{M_i} - d_i^{(1)} - D_i^{(1)}\omega_i - P_{L_i} - d_i^{(2)} - d_i^{(0)} - \sum_{(i,j)\in\mathcal{E}} v_i v_j B_{ij} \sin(\delta_i - \delta_j)$$
$$i \in \mathcal{N} \qquad (2.1)$$

where $M_i > 0$ is the generator inertia, $D_i^{(0)} > 0$ is the generator damping coefficient, P_{M_i} is the mechanical power input, $d_i^{(0)}$ is the disturbance injection, e.g., renewable energy injections and variations on both supply and demand, and $B_{ij} = B_{ji} > 0$ is the susceptance of the transmission line connecting buses i and j. Moreover, in realistic power systems, each generator contains a turbine-governor system given by [6, 13]

$$\dot{P}_{M_i} = \frac{1}{T_{TG_i}} \left(P_{C_i} - P_{M_i} - \frac{\omega_i}{R_i} \right), \ i \in \mathcal{N} \tag{2.2}$$

where P_{C_i} is the command/control input to the generator and $T_{TG_i}, R_i > 0$ are constant parameters. Note that we have simplified the dynamics of the turbine-governor system using a first-order model as in [13].

To simplify the notation, in the rest of the chapter, all the variables denote *deviations* from their nominal operating values. Let $\delta_{ij} = \delta_i - \delta_j, (i,j) \in \mathcal{E}$. Linearize system (2.1)-(2.2) around the nominal operating point to obtain

$$M_i \dot{\omega}_i + \underbrace{\left(D_i^{(0)} + D_i^{(1)} \right)}_{D_i} \omega_i = P_{M_i} - P_{L_i} - \underbrace{\left(d_i^{(0)} + d_i^{(1)} + d_i^{(2)} \right)}_{d_i} - \sum_{(i,j)\in\mathcal{E}} T_{ij} \delta_{ij}, \ i \in \mathcal{N} \tag{2.3a}$$

$$\dot{\delta}_{ij} = \omega_i - \omega_j, \ (i,j) \in \mathcal{E} \tag{2.3b}$$

$$\dot{P}_{M_i} = \frac{1}{T_{TG_i}} \left(P_{C_i} - P_{M_i} - \frac{\omega_i}{R_i} \right), \ i \in \mathcal{N} \tag{2.3c}$$

where $T_{ij} = v_i^0 v_j^0 B_{ij} > 0, (i,j) \in \mathcal{E}$, and each d_i still denotes the disturbance injection.

Conventionally, each area/bus in the network is equipped with a tie-line bias controller (i.e., an integral controller to drive the ACE to zero) given by [6]

$$\dot{P}_{C_i} = K_i \underbrace{\left(-B_i \omega_i - \sum_{(i,j)\in\mathcal{E}} T_{ij} \delta_{ij} \right)}_{\text{Area Control Error (ACE)}}, \ i \in \mathcal{N} \tag{2.4}$$

where $K_i, B_i > 0$ are constant (design) parameters. Under the tie-line bias control, when $P_{L_i} + d_i$ deviates from its nominal value, system (2.3) moves to a new state such that $P_{M_i} = P_{C_i} = P_{L_i} + d_i$, $\omega_i = 0$ and $\delta_{ij} = 0$ hold, i.e., each bus absorbs its own disturbances and scheduled interchanges of tie-line power are maintained. However, this could be economically inefficient as pointed out in [13]. On the other hand, the use of controllable loads can provide significant added stability and robustness to power systems, especially under real-time disturbances, contingency and renewable energy penetration [75, 73]. Due to these reasons, we will redesign the tie-line bias control and consider the participation of controllable loads in active power regulation.

We first present the state-space description of system (2.3), which is used to prove relevant results in this chapter. Let $\alpha_i = \delta_i - \delta_n, i \in \mathcal{N} \setminus \{n\}$, $\omega = [\omega_1, \cdots, \omega_n]^T$, $\alpha = [\alpha_1, \cdots, \alpha_{n-1}]^T$, and $\sigma = [\sigma_1, \cdots, \sigma_p]^T$ where $\sigma_k = \delta_{ij} = \delta_i - \delta_j, i < j$ if the transmission line $(i,j) \in \mathcal{E}$ is indexed by k. Let $T = \text{diag}\{T_k\} \in \mathbb{R}^{p \times p}$, where $T_k = T_{ij}$ if the transmission line $(i,j) \in \mathcal{E}$ is indexed by k. Moreover, we define three matrices to describe the interconnection structure of the network:

$$A_0 \in \mathbb{R}^{n \times p} = [a_{ik}] = \begin{cases} 1 & \text{if bus } i \text{ is connected to a bus } j > i \\ & \text{through transmission line indexed by } k, \\ -1 & \text{if bus } i \text{ is connected to a bus } j < i \\ & \text{through transmission line indexed by } k, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

$$A \in \mathbb{R}^{(n-1) \times p} = \begin{bmatrix} I_{n-1} & \mathbf{0} \end{bmatrix} \times A_0 \tag{2.6}$$

$$\Gamma \in \mathbb{R}^{n \times (n-1)} = \begin{bmatrix} I_{n-1} \\ -\mathbf{1} \end{bmatrix}. \tag{2.7}$$

Lemma 2.1 *The following equalities hold:*

$$A_0 = \Gamma A \tag{2.8a}$$

$$\sigma = A^T \alpha \tag{2.8b}$$

$$\dot{\sigma} = A_0^T \omega. \tag{2.8c}$$

Proof: Note that A_0 is the incidence matrix of the directed graph $(\mathcal{N}, \mathcal{E})$. So $\mathbf{1}A_0 = \mathbf{0}$ holds, which means that (2.8a) is true (actually A is the reduced incidence matrix resulting from A_0). Equations (2.8b) and (2.8c) can be verified from the definition above. A 3-bus (3-area) network in Figure 2.1 illustrates these relations. ■

The state-space version of the model is then given by

$$M\dot{\omega} + D\omega = P_M - P_L - d - A_0 T A^T \alpha \tag{2.9a}$$

$$\dot{\alpha} = \Gamma^T \omega \tag{2.9b}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega) \tag{2.9c}$$

where $M = \text{diag}\{M_i\} \in \mathbb{R}^{n \times n}$, $D = \text{diag}\{D_i\} \in \mathbb{R}^{n \times n}$, $P_M = [P_{M_1}, \cdots, P_{M_n}]^T$, $P_L = [P_{L_1}, \cdots, P_{L_n}]^T$, $d = [d_1, \cdots, d_n]$, $T_{TG} = \text{diag}\{T_{TG_i}\} \in \mathbb{R}^{n \times n}$, $P_C = [P_{C_1}, \cdots, P_{C_n}]^T$ and $R = \text{diag}\{R_i\} \in \mathbb{R}^{n \times n}$.

Remark 2.1 *Multiplying A^T on both sides of (2.9b) results in (2.8c) which is equivalent to (2.3b). Since the rank of A_0 is $n-1$ [76], system (2.8c) has a state-space of dimension $n-1$, although σ contains p state variables. So system (2.3) has a state-space of dimension $3n-1$ which is the same as system (2.9).*

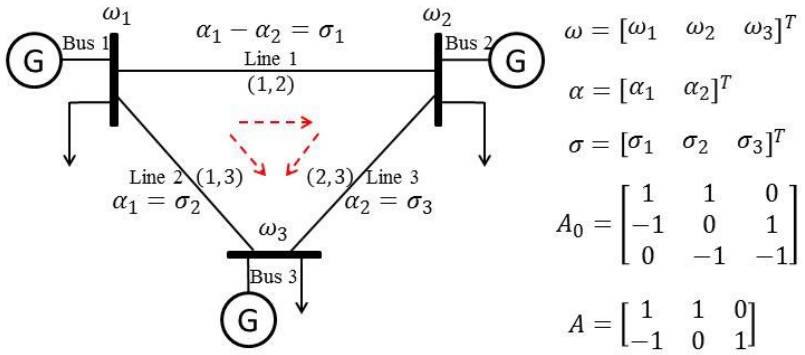

Figure 2.1: A 3-bus (3-area) network (the orientation is indicated by red dashed arrows).

Remark 2.2 *Given P_C, P_L, d, the equilibrium point of system (2.9) is uniquely determined as: $\omega^* = \mathbf{1}\nu$ where $\nu \in \mathbb{R} = \mathbf{1}(P_C - P_L - d)/(tr(D) + tr(R^{-1}))$, $P_M^* = P_C - R^{-1}\omega^*$ and*

$$\alpha^* = (ATA^T)^{-1} \begin{bmatrix} I_{n-1} \\ \mathbf{0} \end{bmatrix}^T (P_C - P_L - d - (D + R^{-1})\mathbf{1}\nu)$$

where ATA^T is invertible [76]. On the other hand, given P_C, P_L, d, the equilibria of system (2.3) are determined as: $\omega_i^* = \nu, P_{M_i}^* = P_{C_i} - R_i^{-1}\omega_i^*, i \in \mathcal{N}$ and

$$AT\sigma^* = \begin{bmatrix} I_{n-1} \\ \mathbf{0} \end{bmatrix}^T (P_C - P_L - d - (D + R^{-1})\mathbf{1}\nu).$$

Note that the above equation results in multiple solutions for σ^* when the graph $(\mathcal{N}, \mathcal{E})$ has mesh topology. However, all practical trajectories of system (2.3) converge to the unique equilibrium point $(\omega^*, A^T\alpha^*, P_M^*)$ [24] (Remark 5).

2.1.2 The Optimization Problem

Suppose that a *constant* disturbance d occurs in the network. As a result, the frequency deviates from its nominal value. It is expected that the generators can adjust the mechanical power P_M to not only balance the supply and demand at each bus but also to minimize the aggregate power generation cost. These two goals lead to a steady-state optimization problem given by

$$\min_{P_M, \alpha} \quad \sum_{i \in \mathcal{N}} C_i(P_{M_i}) \tag{2.10a}$$

$$\text{subject to} \quad P_M - d - A_0 TA^T\alpha = \mathbf{0} \tag{2.10b}$$

where $C_i(P_{M_i})$ is the cost function for the generator at bus i, and d is a given constant.

Remark 2.3 *The reason why P_L is not included in (2.10b) is the following: (i) the disturbance is expected to be cleared through only generation control, as in the standard tie-line bias control [6]; (ii) according to [24], controllable loads can participate in primary frequency control under proportional actions, which could result in better transient performance; (iii) the use of proportional control for controllable loads leads to a completely decentralized scheme; (iv) if using proportional control for controllable loads, $P_L = \mathbf{0}$ holds after the disturbance is cleared by generators. So we have excluded P_L in (2.10b), and moreover, we will use proportional actions for controllable loads.*

For problem (2.10), we have a related assumption.

Assumption 2.2 *Each $C_i(P_{M_i})$ is of quadratic form, i.e., $C_i(P_{M_i}) = \frac{1}{2}c_iP_{M_i}^2 + bP_{M_i}, b > 0, c_i > 0, i \in \mathcal{N}$. Problem (2.10) is feasible.*

This assumption is supported as follows. First, cost functions with quadratic form $\frac{1}{2}c_mP_m^2 + b_mP_m$, where $b_m > 0, c_m > 0$ and P_m is the mechanical power input, are widely used for generators [6, 16]. Let $C_i(P_{M_i}) = \frac{1}{2}c_iP_{M_i}^2 + b_iP_{M_i}$ where $b_i > 0, c_i > 0, i \in \mathcal{N}$. Note that we consider the network operating around a nominal state, i.e., for each generator, P_{M_i} denotes a deviation from its nominal operating value $P_{M_i}^0$. If there is no disturbance, i.e., $d = \mathbf{0}$, the system should stay at the nominal state so that $[P_M^{*T}, \alpha^{*T}]^T = \mathbf{0}$ is the optimal solution of problem (2.10). Since problem (2.10) is a convex problem with linear equality constraints, we can then obtain from its Karush-Kuhn-Tucker (KKT) conditions [77] that $b_i = b_j, (i, j) \in \mathcal{E}$. Finally, the feasibility of the problem is a necessary assumption.

Remark 2.4 *Adding (2.10b) results in $\mathbf{1}(P_M - d) = 0$, which indicates that $\omega = \mathbf{0}$ and $P_C = P_M$ hold in steady state for system (2.9). Based on Remark 2.2 (α can be uniquely derived given $P_C, P_L = \mathbf{0}, d$), problem (2.10) can be reformulated as*

$$\min_{P_C} \quad \sum_{i\in\mathcal{N}} C_i(P_{C_i}) \tag{2.11a}$$

$$\text{subject to} \quad \mathbf{1}(P_C - d) = 0 \tag{2.11b}$$

regardless of P_M, α. Under Assumption 2.2, the objective function becomes

$$\sum_{i\in\mathcal{N}} \frac{1}{2}c_iP_{C_i}^2 + bP_{C_i} = \sum_{i\in\mathcal{N}} \frac{1}{2}c_iP_{C_i}^2 + b\mathbf{1}d$$

which indicates that the choice of b does not affect the optimal solution. Moreover, c_i is chosen to be equal to c_i^0 in the global cost function of the generator at bus i, i.e., $C_i^0(P_{M_i}^0 + P_{M_i}) = \frac{1}{2}c_i^0(P_{M_i}^0 + P_{M_i})^2 + b_i^0(P_{M_i}^0 + P_{M_i})$, where $b_i^0 > 0, c_i^0 > 0, i \in \mathcal{N}$.

The goal is to redesign the tie-line bias control (i.e., redesign $\dot{P}_C$) so that it asymptotically stabilizes system (2.9) under disturbance injection d and the equilibria are the optimal solutions of the optimization problem (2.10)/(2.11). Note that solving (2.10)/(2.11) only leads to economic efficiency for secondary frequency control, which is different from solving an ED problem in tertiary control.

2.2 Redesign Methodology, Stability and Delay Robustness

2.2.1 Redesign Methodology

We first focus on the optimization problem (2.11). Since it is a strictly convex problem with a linear equality constraint under Assumption 2.2, strong duality holds and the KKT conditions are necessary and sufficient conditions for optimality [77], given by

$$c_i P_{C_i} = c_j P_{C_j},\ (i,j) \in \mathcal{E} \tag{2.12a}$$

$$\mathbf{1}(P_C - d) = 0 \tag{2.12b}$$

where (2.12a) indicates that the marginal costs of generators should be same at the optimal point. Recall the conventional tie-line bias control (2.4). It is actually an integral controller which aims to drive $-B_i\omega_i - \sum_{(i,j)\in\mathcal{E}} T_{ij}\delta_{ij}$ to zero. To keep the modification small, we replace this term by terms from the KKT conditions so that optimality is achieved at the equilibria:

$$\dot{P}_{C_i} = B_i(P_{M_i} - P_{C_i}) - K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i} - c_j P_{C_j}),\ i \in \mathcal{N} \tag{2.13}$$

where we have changed the positions of B_i, K_i since they are design parameters. Compared to the conventional tie-line bias control which is decentralized (among all control areas), the redesigned control scheme requires information exchange between neighbouring buses as illustrated in Figure 2.2, which is distributed. The communication can be carried out via the Wide-Area Network (WAN) [78].

In addition to redesigning $\dot{P}_{C_i}$, we consider the participation of controllable loads and use proportional control to regulate them as discussed in Remark 2.3:

$$P_{L_i} = P_{L_i}(k_{p_i}, \underline{\omega}_i, \overline{\omega}_i, \underline{P}_{L_i}, \overline{P}_{L_i}),\ i \in \mathcal{N} \tag{2.14}$$

where $k_{p_i} \geq 0$ is the proportional gain and $P_{L_i}(\cdot)$ is now a nonlinear function with dead zone (parameterized by $\underline{\omega}_i, \overline{\omega}_i, \underline{\omega}_i \leq 0 \leq \overline{\omega}_i$) and saturation (parameterized by

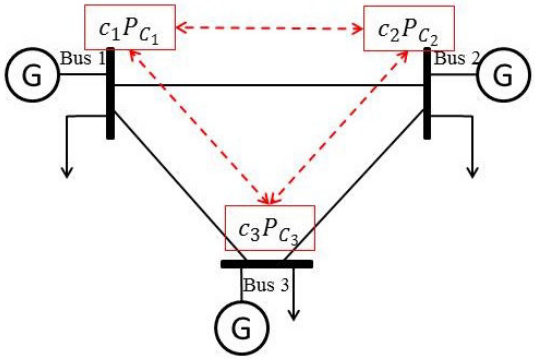

Figure 2.2: Information exchange illustration for a 3-bus network.

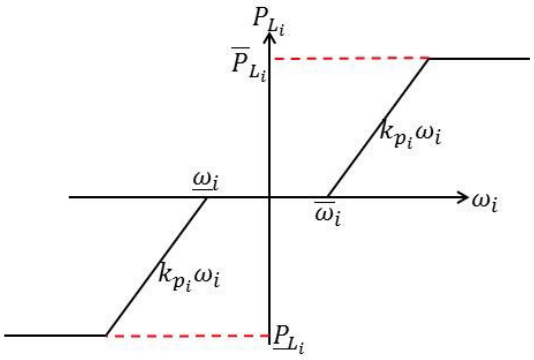

Figure 2.3: $P_{L_i}(\cdot), i \in \mathcal{N}$ with nonlinear characteristics.

$\underline{P}_{L_i}, \overline{P}_{L_i}, \underline{P}_{L_i} \leq 0 \leq \overline{P}_{L_i}$) characteristics, shown in Figure 2.3. These loads actually add damping to the system. Finally, the closed-loop system dynamics in state-space form are given by

$$M\dot{\omega} + D\omega = P_M - P_L - d - A_0 T A^T \alpha \tag{2.15a}$$

$$\dot{\alpha} = \Gamma^T \omega \tag{2.15b}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega) \tag{2.15c}$$

$$\dot{P}_C = B(P_M - P_C) - K A_0 A_0^T C P_C \tag{2.15d}$$

where $B = \text{diag}\{B_i\} \in \mathbb{R}^{n\times n}, K = \text{diag}\{K_i\} \in \mathbb{R}^{n\times n}, C = \text{diag}\{c_i\} \in \mathbb{R}^{n\times n}$, and P_{L_i}, i.e., the entry of P_L, is given by (2.14).

2.2.2 Main Results

We now present results relating to the optimality, stability and delay robustness of system (2.15).

Theorem 2.1 *Given a constant d, under Assumption 2.2, the equilibrium point of the overall system (2.15) is unique and (P_M^*, α^*) is the optimal solution of problem (2.10).*

Proof: At the equilibrium, we have

$$\begin{aligned}
\mathbf{1}(D\omega^* - P_M^* + P_L^* + d) &= 0 \\
\omega_i^* = \omega_j^*,\ (i,j) &\in \mathcal{E} \\
P_C^* - P_M^* - R^{-1}\omega^* &= \mathbf{0} \\
\mathbf{1}K^{-1}B(P_M^* - P_C^*) &= 0
\end{aligned}$$

which result in $\omega^* = \mathbf{0}, P_L^* = \mathbf{0}, \mathbf{1}(P_M^* - d) = 0, P_C^* = P_M^*$ and $A_0 A_0^T C P_C^* = \mathbf{0}$. Using Lemma 2.1, the last equation leads to $P_C^{*T} C\Gamma A A^T \Gamma^T C P_C^* = 0$. Since AA^T is positive definite [76], we then have $\Gamma^T C P_M^* = \Gamma^T C P_C^* = \mathbf{0}$. Therefore, (P_M^*, α^*) is the unique optimal solution of problem (2.10). According to Remark 2.2, the equilibrium point of the overall system (2.15) is unique. ■

Theorem 2.1 shows that the redesigned control scheme can ensure optimality for the network in the steady state. We then prove convergence, i.e. the stability of the overall system.

Theorem 2.2 *Given a constant d, under Assumption 2.2, starting from any feasible initial point, the trajectories generated by (2.15) converge to its unique equilibrium point if* $K = \kappa C R^{-1} B$ *where* $\kappa > 0$.

Proof: Define a candidate Lyapunov function for the overall system as

$$\begin{aligned}
V_{(2.15)} =& \frac{1}{2}\omega^T M \omega + \frac{1}{2}(\alpha - \alpha^*)^T A T A^T (\alpha - \alpha^*) + \frac{1}{2}(P_M - P_M^*)^T T_{TG} R (P_M - P_M^*) \\
&+ \frac{1}{2}(P_C - P_C^*)^T B^{-1} R (P_C - P_C^*).
\end{aligned} \tag{2.16}$$

If $K = \kappa C R^{-1} B$, we have

$$\begin{aligned}
\dot{V}_{(2.15)} =& -\omega^T D \omega - P_L^T \omega - (P_M - P_C)^T R (P_M - P_C) - \kappa (\Gamma^T C P_C)^T A A^T (\Gamma^T C P_C) \\
\leq& 0
\end{aligned} \tag{2.17}$$

where $P_L^T \omega \geq 0$ always holds. On the other hand, $\dot{V}_{(2.15)} = 0$ only when $\omega = P_L = \mathbf{0}, P_M = P_C, \Gamma^T C P_C = \mathbf{0}$, which only happens at the equilibrium point. Since $V_{(2.15)}$ is radially unbounded, using Krasovskii–LaSalle principle [79], the equilibrium point is globally asymptotically stable. ■

We next study delay robustness of the redesigned dynamics (2.15). As described previously, the implementation of the proposed generation control requires communication between neighbouring buses carried out via the WAN. So delays could occur in the network.

Consider the case in which the system is operating under finite communication delays, given by $\tau_{ji} > 0$ if a signal is passed from bus j to bus i. System (2.15) now becomes

$$M\dot{\omega}(t) = -D\omega(t) + P_M(t) - P_L(t) - d - A_0 T A^T \alpha(t) \tag{2.18a}$$

$$\dot{\alpha}(t) = \Gamma^T \omega(t) \tag{2.18b}$$

$$\dot{P}_M(t) = T_{TG}^{-1}(P_C(t) - P_M(t) - R^{-1}\omega(t)) \tag{2.18c}$$

$$\dot{P}_{C_i}(t) = B_i(P_{M_i}(t) - P_{C_i}(t)) - K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji})),\ i \in \mathcal{N} \tag{2.18d}$$

where $P_{C_i}(\theta) = \phi_i(\theta), \theta \in [-h, 0], \phi_i \in \mathcal{C}([-h, 0], \mathbb{R}), i \in \mathcal{N}$ and $h = \max_{(i,j)\in\mathcal{E}}\{\max\{\tau_{ij}, \tau_{ji}\}\}$. Note that we do not require $\tau_{ij} = \tau_{ji}$. For the optimality and stability of system (2.18), we have the following theorem.

Theorem 2.3 *Given a constant d, under Assumption 2.2, the equilibrium point of system (2.18) is unique and globally asymptotically stable if $K = \kappa C R^{-1} B$ where $\kappa > 0$. Moreover, (P_M^*, α^*) is the optimal solution of problem (2.10).*

Proof: It is sufficient to show the global asymptotic stability of the equilibrium. Define a candidate Lyapunov-Krasovskii functional as

$$\begin{aligned} V_{(2.18)}(t) =& \frac{1}{2}\omega(t)^T M\omega(t) + \frac{1}{2}(\alpha(t) - \alpha^*)^T ATA^T(\alpha(t) - \alpha^*) \\ &+ \frac{1}{2}(P_M(t) - P_M^*)^T T_{TG} R(P_M(t) - P_M^*) + \frac{1}{2}(P_C(t) - P_C^*)^T B^{-1} R(P_C(t) - P_C^*) \\ &+ \frac{\kappa}{2}\sum_{(i,j)\in\mathcal{E}} \left(\int_{t-\tau_{ji}}^{t} (c_j P_{C_j}(\beta) - c_j P_{C_j}^*)^2 d\beta + \int_{t-\tau_{ij}}^{t} (c_i P_{C_i}(\beta) - c_i P_{C_i}^*)^2 d\beta \right) \end{aligned} \tag{2.19}$$

where $\kappa > 0$. We move the equilibrium point to the origin and continue to use the same variable names, however, these are now deviations from the equilibrium point. Let $K = \kappa C R^{-1} B$. Differentiating $V_{(2.18)}$ with respect to time, we get

$$\begin{aligned} \dot{V}_{(2.18)}(t) =& -\omega(t)^T D\omega(t) - P_L(t)^T\omega(t) - (P_M(t) - P_C(t))^T R(P_M(t) - P_C(t)) \\ &- \frac{\kappa}{2}\sum_{(i,j)\in\mathcal{E}} \left((c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji}))^2 + (c_j P_{C_j}(t) - c_i P_{C_i}(t-\tau_{ij}))^2 \right) \\ \leq& 0. \end{aligned} \tag{2.20}$$

Moreover, $\dot{V}_{(2.18)}(t) = 0$ yields $\omega(t) = P_L(t) = \mathbf{0}, P_M(t) = P_C(t)$, $c_i P_{C_i}(t) = c_j P_{C_j}(t - \tau_{ji})$ and $c_i P_{C_i}(t - \tau_{ij}) = c_j P_{C_j}(t), (i, j) \in \mathcal{E}$. Substituting them into (2.18), we obtain that $P_M(t), P_C(t)$ are constant, which only happens at the equilibrium point.

Since $V_{(2.18)}$ is radially unbounded, using LaSalle's invariance principle for time-delay systems [80], the equilibrium point is globally asymptotically stable. ■

Theorem 2.3 indicates that the global asymptotic stability of the overall system is independent of (finite) delays by choosing $K = \kappa C R^{-1} B$ where $\kappa > 0$. This is also the nominal stability condition given in Theorem 2.2.

2.2.3 Discussion

It is worth noting that the designed controller (2.13) (Equation (2.18d) for the delayed case) can be interpreted as a modified version of the Kuramoto Model [81]:

$$\dot{x}_i = y_i - \sum_{i=1}^{N} A_{ij} \sin(x_i - x_j)$$

$$\dot{x}_i(t) = y_i(t) - \sum_{i=1}^{N} A_{ij} \sin(x_i(t) - x_j(t - \tau_{ji}))$$

where there are N coupled oscillators with phases $x_i \in [0, 2\pi]$ and natural frequencies y_i, A_{ij} is the coupling strength between neighbouring oscillators (i, j), and $\tau_{ji} > 0$ is the delay in the coupling from j to i. When system (2.15) (system (2.18) for the delayed case) achieves optimality, i.e., the optimization problem (2.10)/(2.11) is solved, the dynamics (2.13) (Equation (2.18d) for the delayed case) achieve consensus [82] given by (2.12a). Moreover, the proofs relating to stability are similar to those for the Kuramoto Model shown in [81] (see the Appendix).

2.3 Extensions

In this section, we show two direct extensions of the proposed redesign framework.

2.3.1 Topology

Firstly, consider a transmission level network with arbitrary topology, described by a connected directed graph $(\mathcal{G} \bigcup \mathcal{L}, \mathcal{E})$. Here $\mathcal{G}$ is the set of generator buses, $\mathcal{L}$ is the set of load buses, and $\mathcal{E} \subseteq (\mathcal{G} \bigcup \mathcal{L}) \times (\mathcal{G} \bigcup \mathcal{L})$ is the set of transmission lines connecting the buses. Each generator bus contains only one synchronous generator/generation unit, and each load bus contains only one load which is an aggregation of a certain amount of loads/users at the bus it is connected to. This can be realized through introducing fictitious buses which in fact represent the internal generation voltages [74]. An example is shown in Figure 2.4. The advantage of using the transformation is that

generators and loads can be partitioned into different groups (i.e., generator buses and load buses) so that the control design for them can be distinguished. *Note that the proposed design frameworks in this thesis can be naturally applied to the case where there exist buses containing both a synchronous generator and a load.*

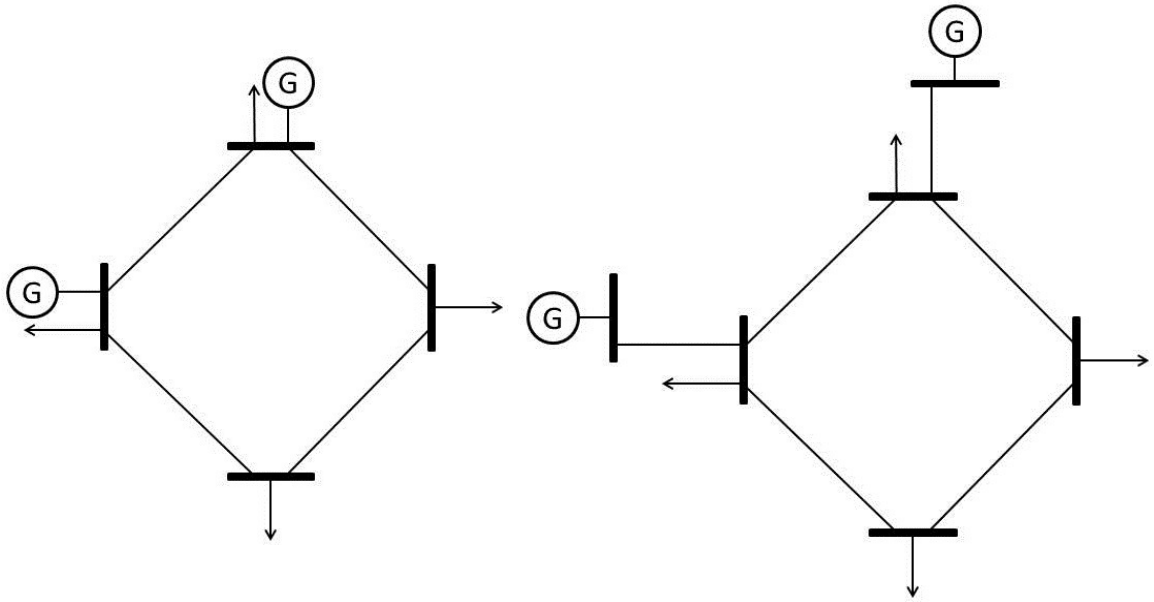

Figure 2.4: An example of topology transformation. Left: before. Right: after.

We number the generator buses $1, \cdots, m$ ($\mathcal{G} = \{1, \cdots, m\}$), the load buses $m+1, \cdots, m+n$ ($\mathcal{L} = \{m+1, \cdots, m+n\}$), and the transmission lines $1, \cdots, p$ corresponding to a lexicographic ordering where $p = |\mathcal{E}|$. Define an orientation from bus i to bus j if $(i, j) \in \mathcal{E}, i < j$. View all buses as voltage sources and let $v_i \angle \delta_i, i \in \mathcal{G} \bigcup \mathcal{L}$ be the voltage of each bus, where v_i is the voltage magnitude and δ_i is the voltage phase angle with respect to the rotating framework of nominal frequency ω_i^0, e.g., $2\pi \times 50$ rad/s in Europe. Assume that the network is working around a nominal operating point which is determined by an ED problem at a more slower time-scale. Assumption 2.1 is modified as follow.

Assumption 2.3 *All bus voltage magnitudes are fixed, i.e., $v_i, i \in \mathcal{G} \bigcup \mathcal{L}$ are constant. The resistance of transmission lines is negligible. Reactive power injections and flows are omitted.*

Similar to Equations (2.1)-(2.3), we use the linearized Structure Preserving Model (SPM) with frequency-dependent active power loads as the dynamics of the network, given by [74, 13]

$$M_i\dot{\omega}_i + D_i\omega_i = P_{M_i} - d_i - \sum_{(i,j)\in\mathcal{E}} T_{ij}\delta_{ij},\ i \in \mathcal{G} \tag{2.21a}$$

$$D_i\omega_i = -P_{L_i} - d_i - \sum_{(i,j)\in\mathcal{E}} T_{ij}\delta_{ij},\ i \in \mathcal{L} \tag{2.21b}$$

$$\dot{\delta}_{ij} = \omega_i - \omega_j,\ (i,j) \in \mathcal{E} \tag{2.21c}$$

$$\dot{P}_{M_i} = \frac{1}{T_{TG_i}}\left(P_{C_i} - P_{M_i} - \frac{\omega_i}{R_i}\right),\ i \in \mathcal{G} \tag{2.21d}$$

where all states and coefficients retain the same meaning as in Equation (2.3) (now $D_i > 0, i \in \mathcal{G}$ is the generator damping coefficient and $D_i > 0, i \in \mathcal{L}$ is the frequency-dependent load damping coefficient). To obtain the state-space description of system (2.21), let $\alpha_i = \delta_i - \delta_{m+n}, i \in \mathcal{G}\bigcup\mathcal{L} \setminus \{m+n\}$, $\omega_g = [\omega_1, \cdots, \omega_m]^T$, $\omega_l = [\omega_{m+1}, \cdots, \omega_{m+n}]^T$, and $\alpha = [\alpha_1, \cdots, \alpha_{m+n-1}]^T$. Moreover, $\sigma, T, A_0 \in \mathbb{R}^{(m+n)\times p}, A \in \mathbb{R}^{(m+n-1)\times p}$ are defined similarly as before. Define Γ_1, Γ_2 to be

$$\Gamma_1 \in \mathbb{R}^{m\times(m+n-1)} = \begin{bmatrix} I_m & \mathbf{0} \end{bmatrix} \tag{2.22}$$

$$\Gamma_2 \in \mathbb{R}^{n\times(m+n-1)} = \begin{bmatrix} \mathbf{0} & I_{n-1} \\ & -\mathbf{1} \end{bmatrix}. \tag{2.23}$$

Then the results in Lemma 2.1 still hold in this case with Equation (2.8a) changing to $A_0 = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} A$. An example is shown in Figure 2.5.

Finally, we obtain the state-space version of the model:

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \Gamma_1 A T A^T \alpha \tag{2.24a}$$

$$\dot{\alpha} = \Gamma_1^T\omega_g + \Gamma_2^T\omega_l \tag{2.24b}$$

$$\omega_l = D_l^{-1}(-P_L - d_l - \Gamma_2 A T A^T \alpha) \tag{2.24c}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega_g) \tag{2.24d}$$

where $M = \text{diag}\{M_i\} \in \mathbb{R}^{m\times m}$, $D_g = \text{diag}\{D_1, \cdots, D_m\}$, $P_M = [P_{M_1}, \cdots, P_{M_m}]^T$, $d_g = [d_1, \cdots, d_m]$, $D_l = \text{diag}\{D_{m+1}, \cdots, D_{m+n}\}$, $P_L = [P_{L_{m+1}}, \cdots, P_{L_{m+n}}]^T$, $d_l = [d_{m+1}, \cdots, d_{m+n}]$, $T_{TG} = \text{diag}\{T_{TG_i}\} \in \mathbb{R}^{m\times m}$, $P_C = [P_{C_1}, \cdots, P_{C_m}]^T$, $R = \text{diag}\{R_i\} \in \mathbb{R}^{m\times m}$.

Remark 2.5 *Multiplying A^T on both sides of (2.24b) results in $\dot{\sigma} = A_0^T\omega$ ($\omega = [\omega_1, \cdots, \omega_{m+n}]^T$) which is equivalent to (2.21c). Since the rank of A_0 is $m+n-1$ [76], system $\dot{\sigma} = A_0^T\omega$ has a state-space of dimension $m+n-1$, although σ contains p state variables. So system (2.21) has a state-space of dimension $3m+n-1$ (note that this system is described by differential algebraic equations which can be rearranged as only differential equations) which is the same as system (2.24).*

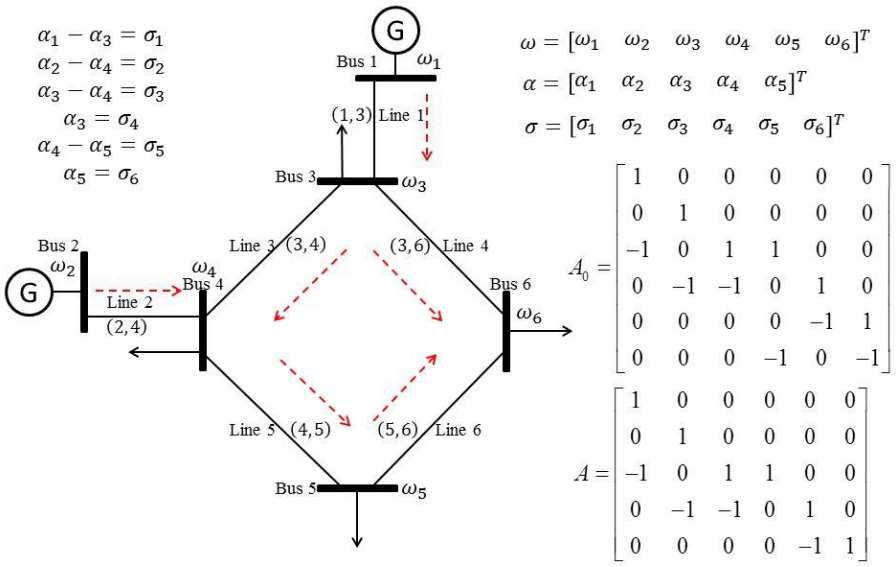

Figure 2.5: A 6-bus network (the orientation is indicated by red dashed arrows, and $\omega = [\omega_1, \cdots, \omega_{m+n}]^T$).

Remark 2.6 *Given P_C, P_L, d_g, d_l, the equilibrium point of system (2.24) is uniquely determined as: $\omega_g^* = \mathbf{1}\nu$ where $\nu \in \mathbb{R} = (\mathbf{1}(P_C - d_g) - \mathbf{1}(P_L + d_l))/(tr(D_g) + tr(R^{-1}) + tr(D_l))$, $P_M^* = P_C - R^{-1}\omega_g^*$ and*

$$\alpha^* = (ATA^T)^{-1} \begin{bmatrix} I_{m+n-1} \\ \mathbf{0} \end{bmatrix}^T \begin{bmatrix} P_C - d_g - (D_g + R^{-1})\mathbf{1}\nu \\ -P_L - d_l - D_l\mathbf{1}\nu \end{bmatrix}$$

where ATA^T is invertible [76]. On the other hand, given P_C, P_L, d_g, d_l, the equilibria of system (2.21) are determined as: $\omega_i^* = \nu, P_{M_i}^* = P_{C_i} - R_i^{-1}\omega_i^*, i \in \mathcal{G}$ and

$$AT\sigma^* = \begin{bmatrix} I_{m+n-1} \\ \mathbf{0} \end{bmatrix}^T \begin{bmatrix} P_C - d_g - (D_g + R^{-1})\mathbf{1}\nu \\ -P_L - d_l - D_l\mathbf{1}\nu \end{bmatrix}.$$

Note that the above equation results in multiple solutions for σ^* when the graph $(\mathcal{G} \bigcup \mathcal{L}, \mathcal{E})$ has mesh topology. However, all practical trajectories of system (2.21) converge to the unique equilibrium point $(\omega_g^*, A^T\alpha^*, P_M^*)$ [24] (Remark 5).

Similar to (2.10), we consider the steady-state optimization problem given by

$$\min_{P_M, \alpha} \quad \sum_{i \in \mathcal{G}} C_i(P_{M_i}) \tag{2.25a}$$

$$\text{subject to} \quad P_M - d_g - \Gamma_1 ATA^T\alpha = \mathbf{0} \tag{2.25b}$$

$$- d_l - \Gamma_2 ATA^T\alpha = \mathbf{0} \tag{2.25c}$$

where $C_i(P_{M_i})$ is the cost function for each generator, and d_g, d_l are given constants. As before, we design the following controller for generators and controllable loads:

$$\dot{P}_{C_i} = B_i(P_{M_i} - P_{C_i}) - K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i} - c_j P_{C_j}),\ i \in \mathcal{G} \tag{2.26a}$$

$$\dot{P}_{C_i} = -K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i} - c_j P_{C_j}),\ i \in \mathcal{L} \tag{2.26b}$$

$$P_{L_i} = P_{L_i}(k_{p_i}, \underline{\omega}_i, \overline{\omega}_i, \underline{P}_{L_i}, \overline{P}_{L_i}),\ i \in \mathcal{L} \tag{2.26c}$$

where $B_i > 0, i \in \mathcal{G}$, $K_i > 0, i \in \mathcal{G} \bigcup \mathcal{L}$, $P_{C_i}, i \in \mathcal{L}$ are auxiliary variables, $c_i = 1, i \in \mathcal{L}$, and $P_{L_i}(\cdot)$ is given by (2.14). The above control scheme requires information exchange between neighbouring buses as illustrated in Figure 2.6, which is distributed (while the control for loads is decentralized). It is clear to see that the benefit of introducing $c_i P_{C_i}, i \in \mathcal{L}$ is to keep communication only between neighbouring buses.

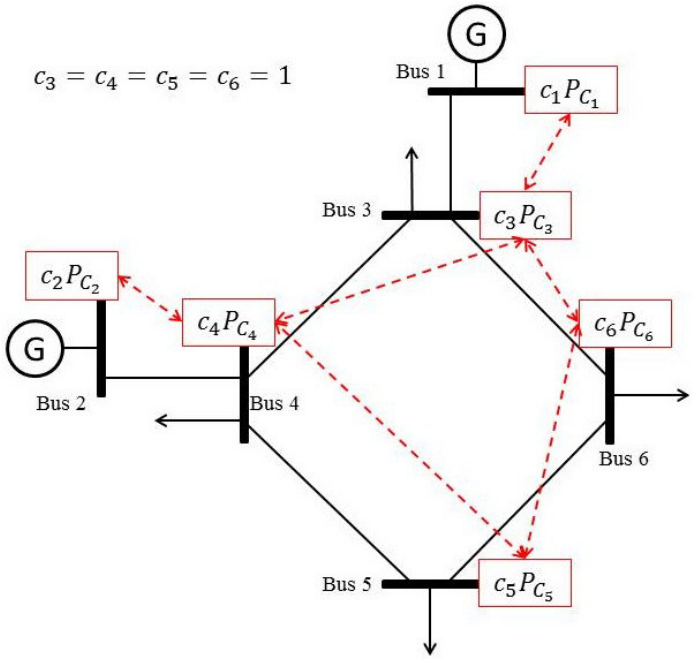

Figure 2.6: Information exchange illustration for a 6-bus network.

The closed-loop system in state-space form is then given by

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \Gamma_1 A T A^T \alpha \tag{2.27a}$$

$$\dot{\alpha} = \Gamma_1^T \omega_g + \Gamma_2^T D_l^{-1}(-P_L - d_l - \Gamma_2 A T A^T \alpha) \tag{2.27b}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega_g) \tag{2.27c}$$

$$\dot{P}_C = B(P_M - P_C) - \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix}^T K A_0 A_0^T \begin{bmatrix} CP_C \\ P_C' \end{bmatrix} \tag{2.27d}$$

$$\dot{P}_C' = -\begin{bmatrix} \mathbf{0} \\ I_n \end{bmatrix}^T K A_0 A_0^T \begin{bmatrix} CP_C \\ P_C' \end{bmatrix} \tag{2.27e}$$

where $B = \text{diag}\{B_i\} \in \mathbb{R}^{m\times m}$, $K = \text{diag}\{K_g, K_l\}, K_g = \text{diag}\{K_1, \cdots, K_m\}, K_l = \text{diag}\{K_{m+1}, \cdots, K_{m+n}\}$, $C = \text{diag}\{c_i\} \in \mathbb{R}^{m\times m}$, $P'_C = [P_{C_{m+1}}, \cdots, P_{C_{m+n}}]^T$, and P_{L_i}, i.e., the entry of P_L, is given by (2.14). Based on Theorems 2.1-2.2, the following proposition is immediate.

Proposition 2.1 *Given constant d_g, d_l, under Assumption 2.2 (problem (2.25) is feasible), the equilibrium point of system (2.27) is unique and globally asymptotically stable if $K_g = \kappa CR^{-1}B$ and $K_l = \kappa I_n$ hold where $\kappa > 0$. Moreover, (P^*_M, α^*) is the optimal solution of problem (2.25).*

Proof: The proof of uniqueness and optimality of the equilibrium point is similar to that of Theorem 2.1. The proof of stability of the equilibrium point is similar to that of Theorem 2.2, by replacing the candidate Lyapunov function $V_{(2.15)}$ with

$$\begin{aligned} V_{(2.27)} =& \frac{1}{2}\omega_g^T M\omega_g + \frac{1}{2}(\alpha-\alpha^*)^T ATA^T(\alpha-\alpha^*) + \frac{1}{2}(P_M - P^*_M)^T T_{TG}R(P_M - P^*_M) \\ &+ \frac{1}{2}(P_C - P^*_C)^T B^{-1}R(P_C - P^*_C) + \frac{1}{2}(P'_C - P'^*_C)^T(P'_C - P'^*_C) \end{aligned} \tag{2.28}$$

for which

$$\begin{aligned} \dot{V}_{(2.27)} =& -\omega_g^T D_g\omega_g - \omega_l^T D_l\omega_l - P_L^T\omega_l - (P_M - P_C)^T R(P_M - P_C) \\ &- \kappa \begin{bmatrix} CP_C \\ P'_C \end{bmatrix}^T \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} AA^T \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix}^T \begin{bmatrix} CP_C \\ P'_C \end{bmatrix} \le 0 \end{aligned} \tag{2.29}$$

when $K_g = \kappa CR^{-1}B$ and $K_l = \kappa I_n$ hold where $\kappa > 0$. ■

Now consider the case in which the system is operating under finite communication delays, given by $\tau_{ji} > 0$ if a signal is passed from bus j to bus i. System (2.27) then becomes

$$M\dot{\omega}_g(t) = -D_g\omega_g(t) + P_M(t) - d_g - \Gamma_1 ATA^T\alpha(t) \tag{2.30a}$$

$$\dot{\alpha}(t) = \Gamma_1^T\omega_g(t) + \Gamma_2^T D_l^{-1}(-P_L(t) - d_l - \Gamma_2 ATA^T\alpha(t)) \tag{2.30b}$$

$$\dot{P}_M(t) = T_{TG}^{-1}(P_C(t) - P_M(t) - R^{-1}\omega_g(t)) \tag{2.30c}$$

$$\dot{P}_{C_i}(t) = B_i(P_{M_i}(t) - P_{C_i}(t)) - K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji})),\ i \in \mathcal{G} \tag{2.30d}$$

$$\dot{P}_{C_i}(t) = -K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji})),\ i \in \mathcal{L} \tag{2.30e}$$

where $P_{C_i}(\theta) = \phi_i(\theta), \theta \in [-h, 0], \phi_i \in \mathcal{C}([-h, 0], \mathbb{R}), i \in \mathcal{G} \bigcup \mathcal{L}$ and $h = \max_{(i,j)\in\mathcal{E}}$ $\{\max\{\tau_{ij}, \tau_{ji}\}\}$. Note that we do not require $\tau_{ij} = \tau_{ji}$. The next proposition follows from Theorem 2.3.

Proposition 2.2 *Given constant d_g, d_l, under Assumption 2.2 (problem (2.25) is feasible), the equilibrium point of system (2.30) is unique and globally asymptotically stable if $K_g = \kappa C R^{-1} B$ and $K_l = \kappa I_n$ hold where $\kappa > 0$. Moreover, (P_M^*, α^*) is the optimal solution of problem (2.25).*

Proof: The proof of this proposition is similar to that of Theorem 2.3 by replacing the candidate Lyapunov-Krasovskii functional $V_{(2.18)}(t)$ with

$$\begin{aligned} V_{(2.30)}(t) =& \frac{1}{2}\omega_g(t)^T M \omega_g(t) + \frac{1}{2}(\alpha(t)-\alpha^*)^T A T A^T (\alpha(t)-\alpha^*) \\ &+ \frac{1}{2}(P_M(t)-P_M^*)^T T_{TG} R (P_M(t)-P_M^*) \\ &+ \frac{1}{2}(P_C(t)-P_C^*)^T B^{-1} R (P_C(t)-P_C^*) + \frac{1}{2}(P'_C(t)-P'^*_C)^T (P'_C(t)-P'^*_C) \\ &+ \frac{\kappa}{2}\sum_{(i,j)\in\mathcal{E}} \left(\int_{t-\tau_{ji}}^{t} (c_j P_{C_j}(\beta) - c_j P^*_{C_j})^2 d\beta + \int_{t-\tau_{ij}}^{t} (c_i P_{C_i}(\beta) - c_i P^*_{C_i})^2 d\beta \right) \end{aligned} \tag{2.31}$$

where $\kappa > 0$, for which

$$\begin{aligned} \dot{V}_{(2.30)}(t) =& -\omega_g(t)^T D_g \omega_g(t) - \omega_l(t)^T D_l \omega_l(t) - P_L(t)^T \omega_l(t) \\ &- (P_M(t) - P_C(t))^T R (P_M(t) - P_C(t)) \\ &- \frac{\kappa}{2}\sum_{(i,j)\in\mathcal{E}} \Big((c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji}))^2 + (c_j P_{C_j}(t) - c_i P_{C_i}(t-\tau_{ij}))^2 \Big) \\ \leq& 0 \end{aligned} \tag{2.32}$$

when $K_g = \kappa C R^{-1} B$ and $K_l = \kappa I_n$ hold. ■

Proposition 2.2 indicates that the global asymptotic stability of the overall system is independent of (finite) delays by choosing $K_g = \kappa C R^{-1} B$ and $K_l = \kappa I_n$ where $\kappa > 0$. This is also the nominal stability condition given in Proposition 2.1.

2.3.2 Controllable Load Participation

Secondly, we include controllable loads in the optimization problem related to active power regulation. Problem (2.25) is reformulated as

$$\min_{P_M, P_L, \alpha} \quad \sum_{i\in\mathcal{G}} C_i(P_{M_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) \tag{2.33a}$$

$$\text{subject to} \quad P_M - d_g - \Gamma_1 A T A^T \alpha = \mathbf{0} \tag{2.33b}$$

$$-P_L - d_l - \Gamma_2 A T A^T \alpha = \mathbf{0} \tag{2.33c}$$

where $U_i(P_{L_i})$ is the utility function for each controllable load, and other variables retain the same meaning as in problem (2.25). We do not add capacity constraints for controllable loads since it is convenient to extend the proposed redesign framework in this case. In Chapter 5, we will revisit the scenario in which capacity constraints are considered for generators, controllable loads and transmission lines.

As before, we have a related assumption for problem (2.33) (the following assumption will be relaxed in Chapter 5).

Assumption 2.4 *Each $C_i(P_{M_i})$ is of quadratic form, i.e., $C_i(P_{M_i}) = \frac{1}{2}c_iP_{M_i}^2 + bP_{M_i}, b > 0, c_i > 0, i \in \mathcal{G}$. Also, each $U_i(P_{L_i})$ is of quadratic form, i.e., $U_i(P_{L_i}) = -\frac{1}{2}c_iP_{L_i}^2 + bP_{L_i}, c_i > 0, i \in \mathcal{L}$. Problem (2.33) is feasible.*

This assumption is reasonable as: (i) cost and utility functions with quadratic form are widely used for generators and loads [16, 83]; (ii) if there is no disturbance, i.e., $d_g = \mathbf{0}, d_l = \mathbf{0}$, the system should stay at the nominal state so that $[P_M^{*T}, P_L^{*T}, \alpha^{*T}]^T = \mathbf{0}$ is the optimal solution of problem (2.33), which leads to the same first-order coefficient in the cost and utility functions; (iii) the feasibility of the problem is necessary. As before, we design the following controller for generators and controllable loads:

$$\dot{P}_{C_i} = B_i(P_{M_i} - P_{C_i}) - K_i \sum_{(i,j)\in\mathcal{E}} (c_iP_{C_i} - c_jP_{C_j}),\ i \in \mathcal{G} \tag{2.34a}$$

$$\dot{P}_{L_i} = B_i\omega_i + K_i \sum_{(i,j)\in\mathcal{E}} (c_iP_{C_i} - c_jP_{C_j}),\ i \in \mathcal{L} \tag{2.34b}$$

where $B_i > 0, K_i > 0, i \in \mathcal{G}\bigcup\mathcal{L}$, $P_{C_i}, i \in \mathcal{L}$ are auxiliary variables satisfying $P_{C_i} = -P_{L_i}$. In this control scheme, controllable loads measure bus frequency deviations and use information from neighbouring buses (as illustrated in Figure 2.6) to adjust active power consumption. The control scheme is completely distributed.

Remark 2.7 *Similar to Remark 2.4, problem (2.33) can be reformulated as*

$$\begin{aligned} \min_{P_C, P_L, \alpha}\ & \sum_{i\in\mathcal{G}} C_i(P_{C_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) \\ \textit{subject to}\ & \mathbf{1}(P_C - d_g) - \mathbf{1}(P_L + d_l) = 0 \end{aligned}$$

regardless of P_M, α. The corresponding KKT conditions are given by

$$\begin{aligned} c_iP_{C_i} &= c_jP_{C_j},\ i,j \in \mathcal{G}, (i,j) \in \mathcal{E} \\ c_iP_{C_i} &= -c_jP_{L_j},\ i \in \mathcal{G}, j \in \mathcal{L}, (i,j) \in \mathcal{E} \\ c_iP_{L_i} &= c_jP_{L_j},\ i,j \in \mathcal{L}, (i,j) \in \mathcal{E} \\ \mathbf{1}(P_C - d_g) &- \mathbf{1}(P_L + d_l) = 0. \end{aligned}$$

It is clear to see that introducing $P_{C_i}, i \in \mathcal{L}$ simplifies Equation (2.34).

The closed-loop system in state-space form is then given by

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \Gamma_1 ATA^T\alpha \tag{2.35a}$$

$$\dot{\alpha} = \Gamma_1^T\omega_g + \Gamma_2^T D_l^{-1}(-P_L - d_l - \Gamma_2 ATA^T\alpha) \tag{2.35b}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega_g) \tag{2.35c}$$

$$\dot{P}_C = B_g(P_M - P_C) - \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix}^T KA_0A_0^T \begin{bmatrix} C_gP_C \\ -C_lP_L \end{bmatrix} \tag{2.35d}$$

$$\dot{P}_L = B_lD_l^{-1}(-P_L - d_l - \Gamma_2 ATA^T\alpha) + \begin{bmatrix} \mathbf{0} \\ I_n \end{bmatrix}^T KA_0A_0^T \begin{bmatrix} C_gP_C \\ -C_lP_L \end{bmatrix}. \tag{2.35e}$$

In the above system, $B_g = \{B_1, \cdots, B_m\}$, $B_l = \{B_{m+1}, \cdots, B_{m+n}\}$, $K = \text{diag}\{K_g, K_l\}$, $K_g = \text{diag}\{K_1, \cdots, K_m\}$, $K_l = \text{diag}\{K_{m+1}, \cdots, K_{m+n}\}$, $C_g = \text{diag}\{c_1, \cdots, c_m\}$, $C_l = \text{diag}\{c_{m+1}, \cdots, c_{m+n}\}$, and other variables remain the same as in (2.27). Based on Theorems 2.1-2.2, the following proposition is immediate.

Proposition 2.3 *Given constant d_g, d_l, under Assumption 2.4, the equilibrium point of system (2.35) is unique and globally asymptotically stable if $K_g = \kappa C_g R^{-1} B_g$ and $K_l = \kappa C_l B_l$ hold where $\kappa > 0$. Moreover, (P_M^*, P_L^*, α^*) is the optimal solution of problem (2.33).*

Proof: The proof of uniqueness and optimality of the equilibrium point is similar to that of Theorem 2.1. The proof of stability of the equilibrium point is similar to that of Theorem 2.2, by replacing the candidate Lyapunov function $V_{(2.15)}$ with

$$\begin{aligned} V_{(2.35)} =& \frac{1}{2}\omega_g^T M\omega_g + \frac{1}{2}(\alpha - \alpha^*)^T ATA^T(\alpha - \alpha^*) + \frac{1}{2}(P_M - P_M^*)^T T_{TG}R(P_M - P_M^*) \\ &+ \frac{1}{2}(P_C - P_C^*)^T B_g^{-1}R(P_C - P_C^*) + \frac{1}{2}(P_L - P_L^*)^T B_l^{-1}(P_L - P_L^*) \end{aligned} \tag{2.36}$$

for which

$$\begin{aligned} \dot{V}_{(2.35)} =& -\omega_g^T D_g\omega_g - \omega_l^T D_l\omega_l - (P_M - P_C)^T R(P_M - P_C) \\ &- \kappa \begin{bmatrix} C_gP_C \\ -C_lP_L \end{bmatrix}^T \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} AA^T \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix}^T \begin{bmatrix} C_gP_C \\ -C_lP_L \end{bmatrix} \leq 0 \end{aligned} \tag{2.37}$$

when $K_g = \kappa C_g R^{-1} B_g$ and $K_l = \kappa C_l B_l$ hold where $\kappa > 0$. ■

We now consider the case in which the system is operating under finite communication delays, given by $\tau_{ji} > 0$ if a signal is passed from bus j to bus i. System (2.35) then becomes

$$M\dot{\omega}_g(t) = -D_g\omega_g(t) + P_M(t) - d_g - \Gamma_1 A T A^T \alpha(t) \tag{2.38a}$$

$$\dot{\alpha}(t) = \Gamma_1^T \omega_g(t) + \Gamma_2^T D_l^{-1}(-P_L(t) - d_l - \Gamma_2 A T A^T \alpha(t)) \tag{2.38b}$$

$$\dot{P}_M(t) = T_{TG}^{-1}(P_C(t) - P_M(t) - R^{-1}\omega_g(t)) \tag{2.38c}$$

$$\dot{P}_{C_i}(t) = B_i(P_{M_i}(t) - P_{C_i}(t)) - K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji})),\ i \in \mathcal{G} \tag{2.38d}$$

$$\dot{P}_{L_i}(t) = B_i\omega_i(t) + K_i \sum_{(i,j)\in\mathcal{E}} (c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji})),\ i \in \mathcal{L} \tag{2.38e}$$

where $P_{C_i}(\theta) = \phi_i(\theta), \theta \in [-h, 0], \phi_i \in \mathcal{C}([-h, 0], \mathbb{R}), i \in \mathcal{G} \bigcup \mathcal{L}$ and $h = \max_{(i,j)\in\mathcal{E}} \{\max\{\tau_{ij}, \tau_{ji}\}\}$. Note that we do not require $\tau_{ij} = \tau_{ji}$. The following proposition follows from Theorem 2.3 immediately.

Proposition 2.4 *Given constant d_g, d_l, under Assumption 2.4, the equilibrium point of system (2.38) is unique and globally asymptotically stable if $K_g = \kappa C_g R^{-1} B_g$ and $K_l = \kappa C_l B_l$ hold where $\kappa > 0$. Moreover, (P_M^*, P_L^*, α^*) is the optimal solution of problem (2.33).*

Proof: The proof of this proposition is similar to that of Theorem 2.3 by replacing the candidate Lyapunov-Krasovskii functional $V_{(2.18)}(t)$ with

$$\begin{aligned} V_{(2.38)}(t) =& \frac{1}{2}\omega_g(t)^T M \omega_g(t) + \frac{1}{2}(\alpha(t) - \alpha^*)^T A T A^T (\alpha(t) - \alpha^*) \\ &+ \frac{1}{2}(P_M(t) - P_M^*)^T T_{TG} R (P_M(t) - P_M^*) \\ &+ \frac{1}{2}(P_C(t) - P_C^*)^T B_g^{-1} R (P_C(t) - P_C^*) + \frac{1}{2}(P_L(t) - P_L^*)^T B_l^{-1} (P_L(t) - P_L^*) \\ &+ \frac{\kappa}{2} \sum_{(i,j)\in\mathcal{E}} \left(\int_{t-\tau_{ji}}^{t} (c_j P_{C_j}(\beta) - c_j P_{C_j}^*)^2 d\beta + \int_{t-\tau_{ij}}^{t} (c_i P_{C_i}(\beta) - c_i P_{C_i}^*)^2 d\beta \right) \end{aligned} \tag{2.39}$$

where $\kappa > 0$, for which

$$\begin{aligned} \dot{V}_{(2.38)}(t) =& -\omega_g(t)^T D_g \omega_g(t) - \omega_l(t)^T D_l \omega_l(t) - (P_M(t) - P_C(t))^T R (P_M(t) - P_C(t)) \\ &- \frac{\kappa}{2} \sum_{(i,j)\in\mathcal{E}} \left((c_i P_{C_i}(t) - c_j P_{C_j}(t-\tau_{ji}))^2 + (c_j P_{C_j}(t) - c_i P_{C_i}(t-\tau_{ij}))^2 \right) \\ \leq& 0 \end{aligned} \tag{2.40}$$

when $K_g = \kappa C_g R^{-1} B_g$ and $K_l = \kappa C_l B_l$ hold. ■

Proposition 2.4 indicates that the global asymptotic stability of the overall system is independent of (finite) delays by choosing $K_g = \kappa C_g R^{-1} B_g$ and $K_l = \kappa C_l B_l$ where $\kappa > 0$. This is also the nominal stability condition given in Proposition 2.3.

2.4 Numerical Investigations

2.4.1 Example 1

We first present a numerical example using the 3-bus (3-area) network illustrated in Figure 2.1. Table 2.1 shows the parameter values of the overall system and $K = \kappa C R^{-1} B$. We consider a scenario consisting of two disturbances, which is realized as

Table 2.1: Parameters of the overall system (p.u. means per unit).

Parameter	Value(p.u.)	Parameter	Value(p.u.)
M_i	$10, 8, 6$	D_i	$1.5, 3, 2$
T_k	$100, 120, 110$	T_{TG_i}	$5, 4, 6$
R_i	$0.05, 0.05, 0.05$	B_i	$2, 2, 2$
Function	Expression	Parameter	Value(p.u.)
$C_1(P_{M_1})$	$0.05P_{M_1}^2 + 5P_{M_1}$	$\underline{P}_{L_1}, \overline{P}_{L_1}$	$-20, 20$
$C_2(P_{M_2})$	$0.1P_{M_2}^2 + 5P_{M_2}$	$\underline{P}_{L_2}, \overline{P}_{L_2}$	$-10, 10$
$C_3(P_{M_3})$	$0.15P_{M_3}^2 + 5P_{M_3}$	$\underline{P}_{L_3}, \overline{P}_{L_3}$	$-15, 15$
		All $\underline{\omega}_i, \overline{\omega}_i$	0
Gain	Value(p.u.)	Parameter	Value(s)
κ	0.5	All τ_{ij}	10

follows: the system is stabilized at the nominal operating point at $t = 0$s; at $t = 10$s, there is a 10 per unit (p.u.) step change of load consumption at bus 1; after 50s, we make a -20p.u. step change of load consumption at bus 3. Both of these changes in practice represent variations on the demand side. The simulation results are shown in Figure 2.7.

From Figure 2.7 we see that under disturbances, the redesigned controller drives the network to a more economically efficient operating point compared with the conventional tie-line bias control. Here we have used the dynamics

$$\dot{P}_{C_i} = -B_i\omega_i - K_i \sum_{(i,j)\in\mathcal{E}} T_{ij}\delta_{ij},\ i \in \mathcal{N}$$

in the tie-line bias control. Moreover, the introduction of controllable loads can improve the transient behaviour of the overall system. As the proportional gain k_{p_i} (we have used a common proportional gain in this example) increases, the oscillations decrease and then vanish. Note that large k_{p_i} slows down the response speed of the overall system. When the effect of delays is considered, the redesigned dynamics still converge to the economically optimal operating point, as illustrated in Figure 2.8.

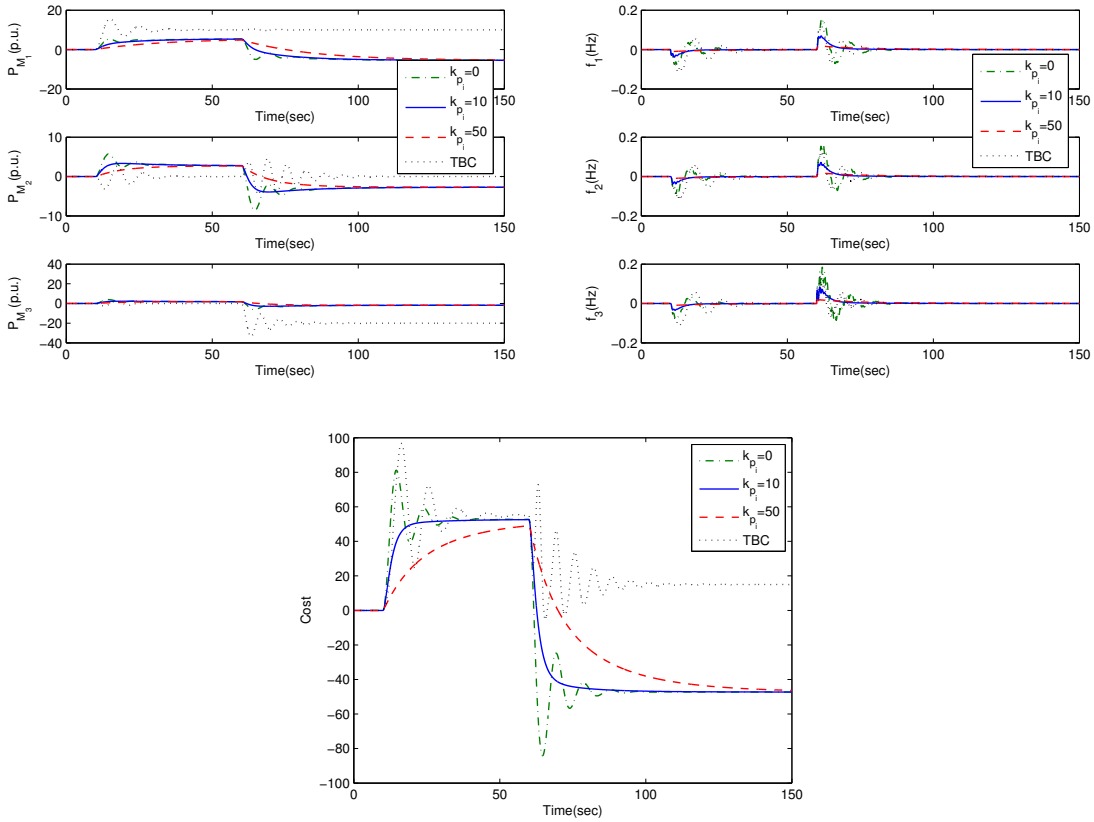

Figure 2.7: Responses of system (2.15) (dashdot lines for $k_{p_i} = 0$, solid lines for $k_{p_i} = 10$, and dashed lines for $k_{p_i} = 50$) and system (2.9) with the Tie-line Bias Control (dotted lines for TBC). Top left: mechanical power (deviations) at each bus. Top right: frequency (deviations) at each bus. Bottom: generation cost (deviations).

Furthermore, the participation of controllable loads can improve the transient performance of the overall network. We still see that there is a trade-off for the choice of k_{p_i}, relating to oscillations and response speed.

2.4.2 Example 2

In the second example, we consider the 6-bus network illustrated in Figure 2.5 and compare systems (2.27) and (2.35) (systems (2.30) and (2.38) in the delayed case) under a same scenario. Table 2.2 shows the parameter values of the overall system. The scenario consists of two disturbances, which is realized as follows: the system is stabilized at the nominal operating point at $t = 0$s; at $t = 10$s, there is a -3p.u. step change at bus 1; after 50s, we make a 1p.u. step change at bus 3. The simulation results of systems (2.27), (2.30), (2.35), and (2.38) are shown in Figures 2.9-2.12 respectively.

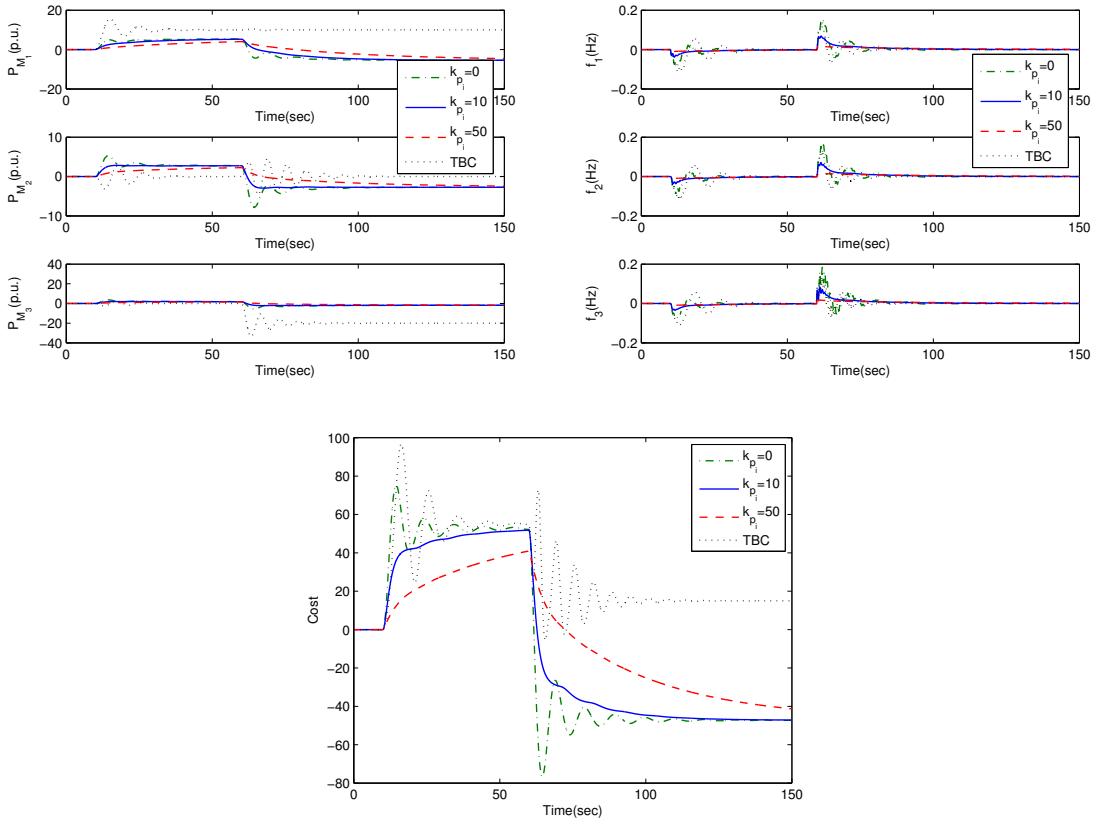

Figure 2.8: Responses of system (2.18) (dashdot lines for $k_{p_i} = 0$, solid lines for $k_{p_i} = 10$, and dashed lines for $k_{p_i} = 50$) under a 10s single delay, and system (2.9) with the Tie-line Bias Control (dotted lines for TBC). Top left: mechanical power (deviations) at each bus. Top right: frequency (deviations) at each bus. Bottom: generation cost (deviations).

From Figures 2.9 and 2.10, we can see that the introduction of controllable loads can improve the transient behaviour of the overall system in both undelayed and delayed cases. As the proportional gain k_{p_i} (we have used a common proportional gain in this example) increases, the oscillations decrease. As before, large k_{p_i} slows down the response speed of the overall system. On the other hand, when controllable loads are involved in the optimization problem, the operation cost decreases as shown in Figures 2.13 and 2.14. Note than in order to compare the operation cost under the same condition, it is necessary to set $\underline{P}_{L_i} = -\infty, \overline{P}_{L_i} = \infty, \underline{\omega}_i = \overline{\omega}_i = 0, i \in \mathcal{L}$ in systems (2.27) and (2.30).

Table 2.2: Parameters of the four systems.

Parameter	Value(p.u.)	Parameter	Value(p.u.)
M_i	$10, 8$	D_i	$2, 1.5, 1, 0.8, 0.3, 0.5$
T_k	$24, 25, 23, 26, 27, 22$	T_{TG_i}	$5, 6$
R_i	$0.05, 0.05$	B_i	$5, 5, 3, 3, 3, 3$
Function	Expression	Function	Expression
$C_1(P_{M_1})$	$0.05P_{M_1}^2 + 0.5P_{M_1}$	$U_4(P_{L_4})$	$-0.02P_{L_4}^2 + 0.5P_{L_4}$
$C_2(P_{M_2})$	$0.1P_{M_2}^2 + 0.5P_{M_2}$	$U_5(P_{L_5})$	$-0.025P_{L_5}^2 + 0.5P_{L_5}$
$U_3(P_{L_3})$	$-0.01P_{L_3}^2 + 0.5P_{L_3}$	$U_6(P_{L_6})$	$-0.015P_{L_6}^2 + 0.5P_{L_6}$
Parameter	Value(p.u.)	Parameter	Value(p.u.)
All $\underline{P}_{L_i}$	$-\infty$	All $\overline{P}_{L_i}$	∞
All $\underline{\omega}_i, \overline{\omega}_i$	0		
Gain	Value(p.u.)	Parameter	Value(s)
κ	2	All τ_{ij}	10

2.5 Conclusion

In this chapter, we have proposed modifications in the generation control in power systems to improve the economic efficiency, stability and robustness of generator regulation in real time. Moreover, the participation of controllable loads has been considered in the control redesign. We first obtained the state-space description of a conventional power network model which described system dynamics around a nominal operating state. We then formulated an optimization problem relating to system regulation under exogenous disturbances. The main results were presented in Section 2.2 where we proposed a new control scheme based on a consensus approach and studied its optimality, stability and delay robustness. In Section 2.3, we extended the designed control scheme to (i) networks with more complexity and (ii) the case where controllable loads were involved in the optimization problem. Finally, numerical examples showed that the proposed controllers could balance the power flow in the network quickly, and drive the system to an economically optimal operating point in the steady state.

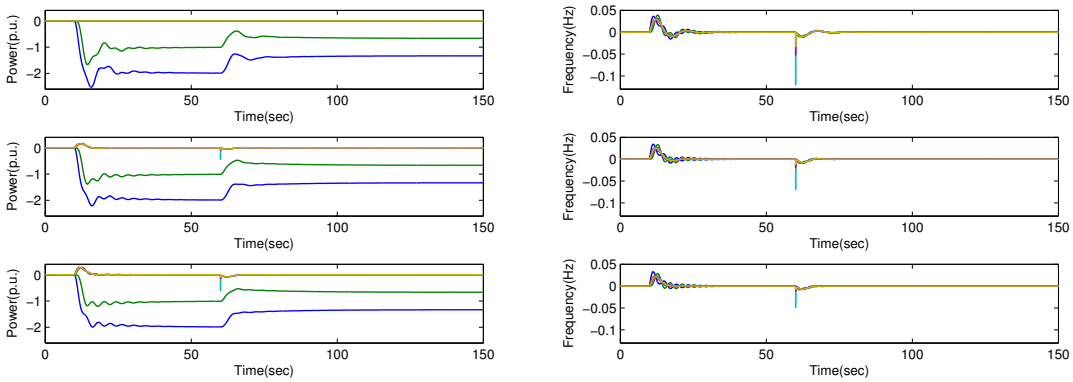

Figure 2.9: Response of system (2.27). Left: power generation and consumption (deviations). Right: bus frequency (deviations). (Top subfigure for $k_{p_i} = 0$; middle subfigure for $k_{p_i} = 1$; bottom subfigure for $k_{p_i} = 2$.)

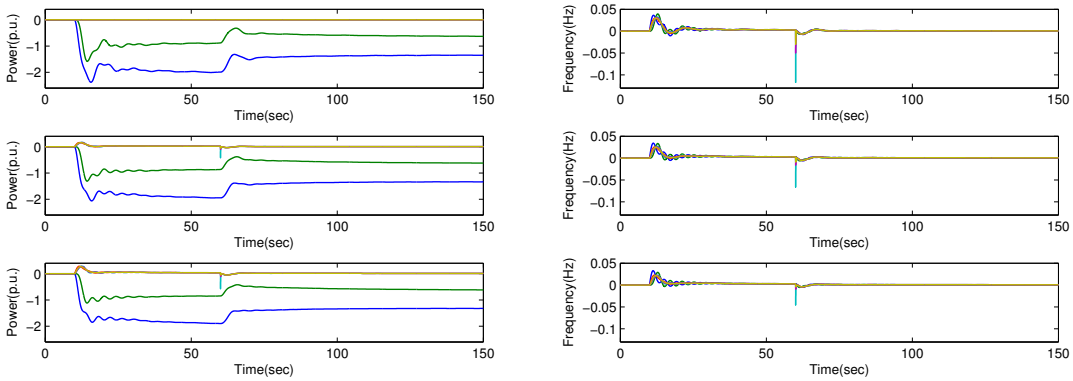

Figure 2.10: Response of system (2.30) under a 10s single delay. Left: power generation and consumption (deviations). Right: bus frequency (deviations). (Top subfigure for $k_{p_i} = 0$; middle subfigure for $k_{p_i} = 1$; bottom subfigure for $k_{p_i} = 2$.)

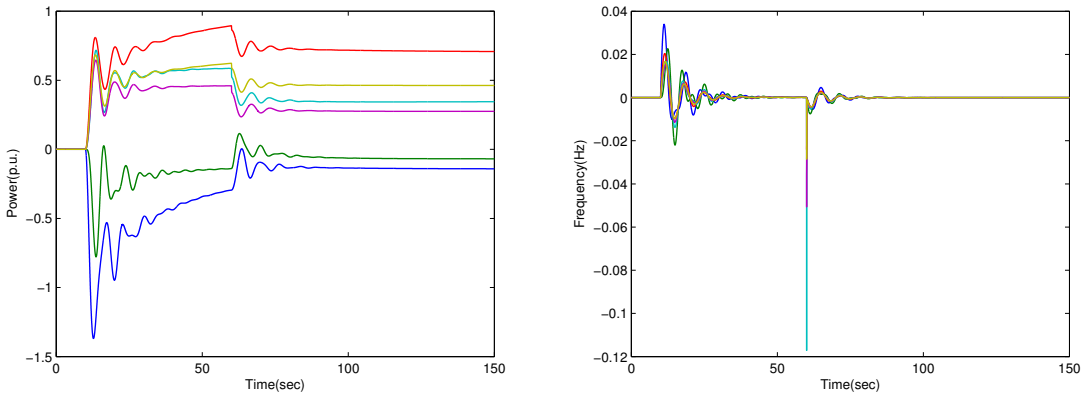

Figure 2.11: Response of system (2.35). Left: power generation and consumption (deviations). Right: bus frequency (deviations).

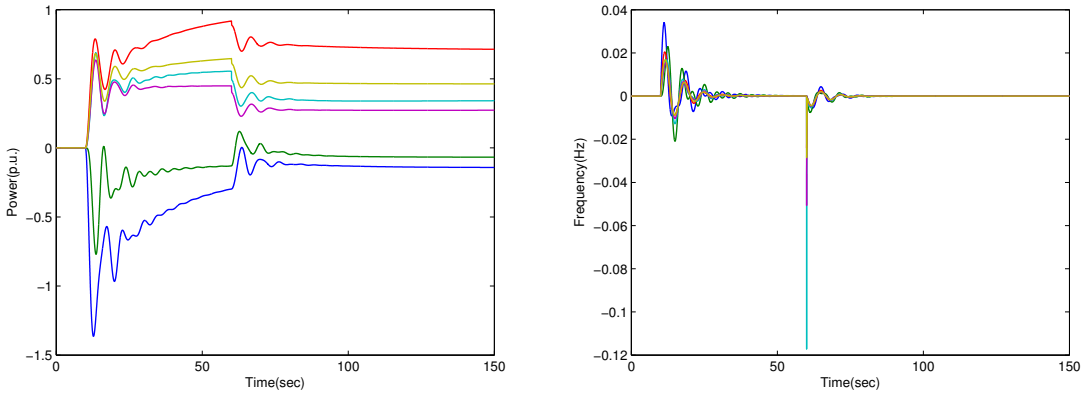

Figure 2.12: Response of system (2.38) under a 10s single delay. Left: power generation and consumption (deviations). Right: bus frequency (deviations).

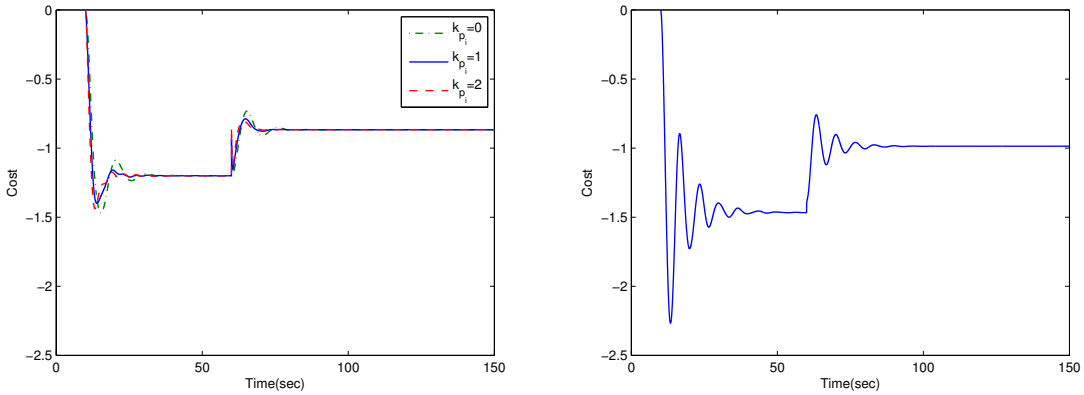

Figure 2.13: Operation cost comparison between systems (2.27) and (2.35). Left: cost (deviation) of system (2.27). Right: cost (deviation) of system (2.35).

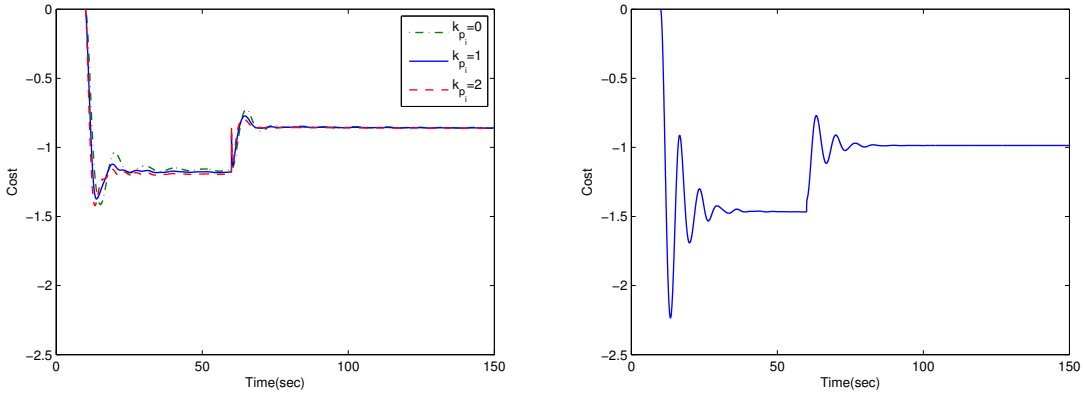

Figure 2.14: Operation cost comparison between systems (2.30) and (2.38) under a 10s single delay. Left: cost (deviation) of system (2.30). Right: cost (deviation) of system (2.38).

Chapter 3

Merging Primary, Secondary and Tertiary Frequency Control in Power Systems via a Primal-dual Decomposition Approach

In the previous chapter, we have proposed a redesign framework based on a consensus approach to merge conventional primary and secondary frequency control for power systems with (frequency-independent) controllable loads. This chapter goes one step further and considers the problem of merging primary, secondary and tertiary frequency control. As described in the introduction, in the traditional tertiary control, an Economic Dispatch (ED)/Optimal Power Flow (OPF) problem is solved in a centralized manner at a time-scale of usually 5 to 15 minutes, to determine the nominal operating values of the system states, e.g., generator command input, load power consumption and interchanges of tie-lines. However, due to fluctuations resulting from increasing distributed energy resources and variability in both supply and demand, the power grid is losing inertia and the economic optimization, i.e., solving the ED problem, will need to be run more and more faster to provide effectiveness and robustness. Thus, breaking the time-scale separation/hierarchical structure in frequency control becomes more and more necessary.

In this chapter, we propose a distributed control architecture to realize real-time economic optimization for the power network under exogenous disturbances. Unlike the traditional market-based control, the feedback signals in our control architecture are not prices from a centralized market but information and feedback signals flowing between neighbouring components in the system. Moreover, we consider nonlinear active power flow equations which are more practical than linear equations used in conventional DC OPF problems [7]. In particular, we focus on a network with tree

topology to demonstrate how the design methodology and stability analysis can be performed. In order to achieve better transient performance and added robustness for the closed-loop system, we also introduce extra dynamics and study the influence of the gains resulting from the extra dynamics on the system robustness. Finally, the performance of the proposed control schemes is illustrated by numerical examples.

3.1 Problem Setup

3.1.1 Control Architecture

The current power network consists of a number of regions, divided by areas, energy source types, etc. Each region corresponds to a transmission level network which contains generators, loads, transmission lines and buses. Each load corresponds to a distribution level network which is an aggregation of a certain amount of users at the bus it is connected to. A user could be an industrial company, a street with a certain amount of buildings, or a combination of several users (industrial companies, houses, etc). As demonstrated in Chapter 2 (Figure 2.4), by introducing fictitious buses, we can change the topology of a given transmission level network so that each bus is connected to either a generator or a load [74].

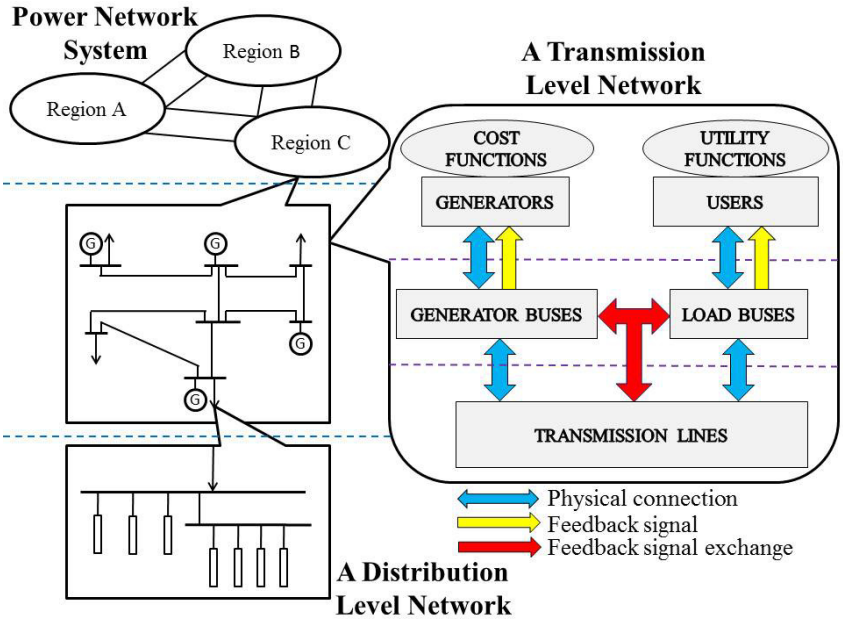

Figure 3.1: The architecture of the power network and the control scheme.

In order to deal with the problems of low predictability and high uncertainty that the current power network is faced with, we consider a real-time control architecture

with demand response, i.e., a mechanism that will encourage consumers to modify their demand when it is most difficult for the network to achieve a balance between supply and demand [14]. This means that loads are allowed to adjust their demand based on feedback signals (not prices from a centralized market), i.e., each user can respond to feedback from the power network. In this approach, every transmission level network can run in a distributed way: each generator receives a feedback signal from the generator bus it is connected to, and adjusts its power generated based on that signal and local information. So does each load to the load bus it is connected to. Each bus calculates feedback signals based on local information and signals from transmission lines and buses it is connected to. The local information contains cost functions, utility functions, capacity constraints, local frequency and power imbalance. All feedback signals can be transmitted via the Wide-Area Network (WAN), Local-Area Network (LAN) and Home-Area Network [78]. There is also information exchange between different regions. Figure 3.1 shows the architecture.

In the rest of the chapter, we study a transmission network with tree topology containing an arbitrary number of synchronous generators and loads (a 3-generator-3-load network with tree topology is illustrated in Figure 3.2). Tree structures are important as: (i) they are sufficiently complicated to offer promising approaches to handle more complicated cases; (ii) the (AC) OPF problem in a tree can be convexified [84], which is the technical reason for giving priority to tree networks in this chapter; (iii) if the use of phase shifters is allowed, the OPF problem for a mesh network can be simplified to one for a tree [85].

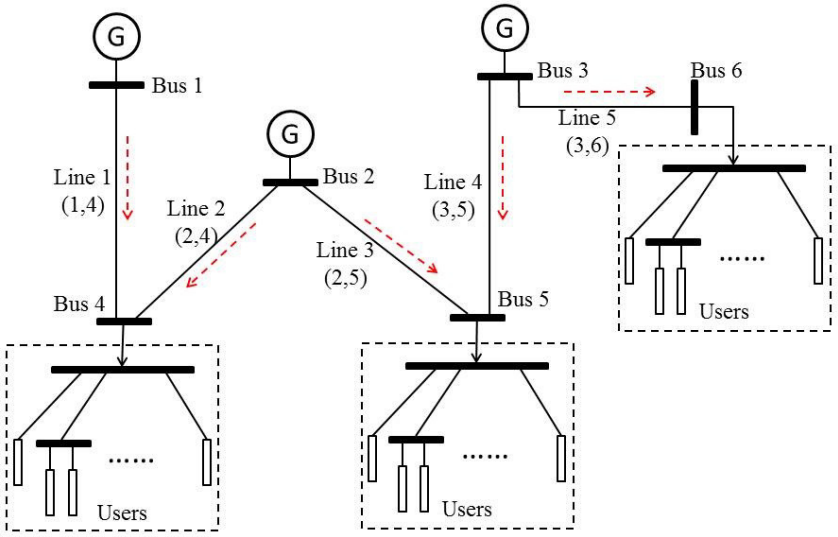

Figure 3.2: A 3-generator-3-load network with tree topology (the orientation is indicated by red dashed arrows).

3.1.2 System Model

Consider an m-generator-n-load transmission level network with tree topology (the number of transmission lines is $m+n-1$). We number the generator buses $1, \cdots, m$, the load buses $m+1, \cdots, m+n$, and the transmission lines $1, \cdots, m+n-1$ (corresponding to a lexicographic ordering, as illustrated in Figure 3.2). Let the sets of generator buses, load buses and transmission lines be $\mathcal{G} = \{1, \cdots, m\}$, $\mathcal{L} = \{m+1, \cdots, m+n\}$, and $\mathcal{E} \subseteq (\mathcal{G} \bigcup \mathcal{L}) \times (\mathcal{G} \bigcup \mathcal{L})$ respectively. View each bus as a voltage source and let $v_i \angle \delta_i, i \in \mathcal{G} \bigcup \mathcal{L}$ be its voltage, where v_i is the bus voltage magnitude, and δ_i is the bus voltage phase angle with respect to the rotating framework of nominal frequency, e.g., $2\pi \times 50$ rad/s in Europe. The main part of the model we use for the network is the Structure Preserving Model (SPM) with frequency-dependent active power loads, for which the following assumption is necessary [74, 6, 7].

Assumption 3.1 *All bus voltage magnitudes are fixed, i.e., $v_i, i \in \mathcal{G} \bigcup \mathcal{L}$ are constant. The resistance of transmission lines is negligible.*

The dynamics of the network is then given by [74, 13]

$$M_i \dot{\omega}_i + D_i \omega_i = P_{M_i} - d_i - \sum_{(i,j) \in \mathcal{E}} T_{ij} \sin \delta_{ij}, \ i \in \mathcal{G} \tag{3.1a}$$

$$D_i \omega_i = -P_{L_i} - d_i - \sum_{(i,j) \in \mathcal{E}} T_{ij} \sin \delta_{ij}, \ i \in \mathcal{L} \tag{3.1b}$$

$$\dot{\delta}_{ij} = \omega_i - \omega_j, \ (i,j) \in \mathcal{E} \tag{3.1c}$$

$$\dot{P}_{M_i} = \frac{1}{T_{TG_i}} \left(P_{C_i} - P_{M_i} - \frac{\omega_i}{R_i} \right), \ i \in \mathcal{G} \tag{3.1d}$$

where $M_i > 0$ is the generator inertia, $\omega_i = \dot{\delta}_i$ is the bus frequency deviation from the nominal frequency, $D_i > 0, i \in \mathcal{G}$ is the generator damping coefficient, $D_i > 0, i \in \mathcal{L}$ is the load bus damping coefficient, P_{M_i} is the mechanical power input, d_i is the disturbance injection, e.g., renewable energy injections and variations on both supply and demand, $T_{ij} = v_i v_j B_{ij} > 0, (i,j) \in \mathcal{E}$, $B_{ij} = B_{ji} > 0$ is the susceptance of the transmission line connecting buses i and j, $\delta_{ij} = \delta_i - \delta_j, (i,j) \in \mathcal{E}$, P_{L_i} is the load power consumption, P_{C_i} is the command/control input to the generator and $T_{TG_i}, R_i > 0$ are constant parameters. Note that we have simplified the dynamics of the turbine-governor system using a first-order model as in [13].

To obtain the state-space description of system (3.1), define $\omega_g = [\omega_1, \cdots, \omega_m]^T$, $\omega_l = [\omega_{m+1}, \cdots, \omega_{m+n}]^T$, and $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_{m+n-1}]^T$, where $\alpha_i = \delta_i - \delta_{m+n}$. Define transmission line angle differences to be $\sigma = [\sigma_1, \sigma_2, \cdots, \sigma_{m+n-1}]^T$, where $\sigma_k = \delta_i -$

$\delta_j, i < j$ for the k^{th} line (numbered using a lexicographic ordering) joining buses i and j. Let $T = \text{diag}\{T_k\} \in \mathbb{R}^{(m+n-1)\times(m+n-1)}$, where $T_k = T_{ij}$ if the transmission line $(i,j) \in \mathcal{E}$ is indexed by k. Moreover, we define four matrices to describe the interconnection structure of the network:

$$A_0 \in \mathbb{R}^{(m+n)\times(m+n-1)} = [a_{ik}] = \begin{cases} 1 & \text{if bus } i \text{ is connected to a bus } j > i \\ & \text{through transmission line indexed by } k, \\ -1 & \text{if bus } i \text{ is connected to a bus } j < i \\ & \text{through transmission line indexed by } k, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

$$A \in \mathbb{R}^{(m+n-1)\times(m+n-1)} = \left[\begin{array}{cc} I_{m+n-1} & \mathbf{0} \end{array}\right] \times A_0 \tag{3.3}$$

$$\Gamma_1 \in \mathbb{R}^{m\times(m+n-1)} = \left[\begin{array}{cc} I_m & \mathbf{0} \end{array}\right] \tag{3.4}$$

$$\Gamma_2 \in \mathbb{R}^{n\times(m+n-1)} = \left[\begin{array}{cc} \mathbf{0} & I_{n-1} \\ & -\mathbf{1} \end{array}\right]. \tag{3.5}$$

Define $f(\alpha) = [f_1(\alpha), f_2(\alpha), \cdots, f_{m+n-1}(\alpha)]^T$ where

$$f_i(\alpha) = \sum_{j=1, j\neq i}^{m+n-1} T_{ij} \sin(\alpha_i - \alpha_j) + T_{i(m+n)} \sin \alpha_i$$

is the power injection at the i^{th} bus. Define $g(\sigma) = [g_1(\sigma_1), g_2(\sigma_2), \cdots, g_{m+n-1}(\sigma_{m+n-1})]^T$ where

$$g_k(\sigma_k) = T_k \sin \sigma_k$$

is the power flow in the k^{th} transmission line. We then have the following lemma.

Lemma 3.1 *The following equalities hold:*

$$A_0 = \left[\begin{array}{c} \Gamma_1 \\ \Gamma_2 \end{array}\right] A \tag{3.6a}$$

$$\sigma = A^T \alpha \tag{3.6b}$$

$$f(\alpha) = Ag(\sigma) = Ag(A^T\alpha) \tag{3.6c}$$

$$\int_{\sigma^*}^{\sigma} (g(\beta) - g(\sigma^*))^T d\beta = \int_{\alpha^*}^{\alpha} (f(\beta) - f(\alpha^*))^T d\beta \tag{3.6d}$$

where $\omega = [\omega_1, \cdots, \omega_{m+n}]^T$, $f(\alpha^*) = Ag(\sigma^*)$ *and* $A^T\alpha^* = \sigma^*$.

Proof: Note that A_0 is the incidence matrix of the tree network (viewed as a directed graph with an orientation from i to j if buses i and j are connected where $i < j$). So $\mathbf{1}A_0 = \mathbf{0}$ holds, which means that (3.6a) is true (actually A is the reduced incidence

matrix obtained by eliminating the last row of A_0). Equations (3.6b)-(3.6c) can be verified from the definition above, while (3.6d) can be derived from (3.6c). ■

Finally, we obtain the state-space version of the model:

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \Gamma_1 f(\alpha) \tag{3.7a}$$

$$\dot{\alpha} = \Gamma_1^T\omega_g + \Gamma_2^T\omega_l \tag{3.7b}$$

$$\omega_l = D_l^{-1}(-P_L - d_l - \Gamma_2 f(\alpha)) \tag{3.7c}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega_g) \tag{3.7d}$$

where $M = \text{diag}\{M_i\} \in \mathbb{R}^{m\times m}$, $D_g = \text{diag}\{D_1, \cdots, D_m\}$, $P_M = [P_{M_1}, \cdots, P_{M_m}]^T$, $d_g = [d_1, \cdots, d_m]$, $D_l = \text{diag}\{D_{m+1}, \cdots, D_{m+n}\}$, $P_L = [P_{L_{m+1}}, \cdots, P_{L_{m+n}}]^T$, $d_l = [d_{m+1}, \cdots, d_{m+n}]$, $T_{TG} = \text{diag}\{T_{TG_i}\} \in \mathbb{R}^{m\times m}$, $P_C = [P_{C_1}, \cdots, P_{C_m}]^T$, $R = \text{diag}\{R_i\} \in \mathbb{R}^{m\times m}$. We have an additional assumption for the model.

Assumption 3.2 *The deviations between neighbouring bus voltage phase angles satisfy $\sigma_k \in [-\Phi, \Phi], k = 1, 2, \cdots, m+n-1$, where $\Phi \in (0, \frac{\pi}{2})$ is a constant.*

This assumption usually holds for normal operating conditions [74]. It ensures synchronization and guarantees uniqueness of the equilibrium α^*/σ^* [86].

3.1.3 The OPF Problem

The objective for the overall network in steady state is to achieve OPF:

$$\min_{P_{M_i}, P_{L_i}, \alpha} \quad \sum_{i\in\mathcal{G}} C_i(P_{M_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) \tag{3.8a}$$

$$\text{subject to} \quad P_M - d_g - \Gamma_1 f(\alpha) = \mathbf{0} \tag{3.8b}$$

$$-P_L - d_l - \Gamma_2 f(\alpha) = \mathbf{0} \tag{3.8c}$$

$$P_C^{\min} \preceq P_M \preceq P_C^{\max} \tag{3.8d}$$

$$P_L^{\min} \preceq P_L \preceq P_L^{\max} \tag{3.8e}$$

$$P_{TC}^{\min} \preceq g(A^T\alpha) \preceq P_{TC}^{\max}. \tag{3.8f}$$

Here $C_i(P_{M_i})$ is the cost function of generator i. $U_i(P_{L_i})$ is the utility function of load i. d_g and d_l are given *constant* vectors, representing fluctuations resulting from distributed energy resources and variability in both supply and demand. $P_C^{\max}, P_C^{\min} \in \mathbb{R}^m$ are the vectors of upper and lower bounds of generator capacity whose entries are $P_{C_i}^{\max}$ and $P_{C_i}^{\min}$ respectively. $P_L^{\max}, P_L^{\min} \in \mathbb{R}^n$ are the vectors of upper and lower bounds of load power consumption whose entries are $P_{L_i}^{\max}$ and $P_{L_i}^{\min}$

respectively. $P_{TC}^{\max}, P_{TC}^{\min} \in \mathbb{R}^{m+n-1}$ are the vectors of upper and lower bounds of transmission line capacity whose entries are $P_{TCi}^{\max}$ and $P_{TCi}^{\min}$ respectively. The objective of the OPF problem is to minimize generation cost and maximize load utility. Equations (3.8b)-(3.8c) ensure power balance at each bus. Equations (3.8d)-(3.8f) are capacity constraints for generators, loads, and transmission lines respectively. We then have the following assumption for the OPF problem.

Assumption 3.3 *The OPF problem is feasible. Moreover, the utility functions for loads, $U_i(P_{L_i})$, $\forall i \in \mathcal{L}$, are continuously differentiable, strictly concave, increasing functions satisfying $U_i''(P_{L_i}) \leq c_i < 0$. The generator cost functions take quadratic forms $C_i(P_{M_i}) = b_i P_{M_i} + \frac{c_i}{2} P_{M_i}^2$ where b_i and c_i are positive, $\forall i \in \mathcal{G}$. $|P_{TCi}^{max}|, |P_{TCi}^{min}| < T_i \sin\Phi$, $i = 1, 2, \cdots, m+n-1$.*

This assumption is supported as follows. First, the feasibility of the problem is a necessary assumption. Second, the more power the load is allocated, the larger the utility to the load. But the utility saturates as the power consumed increases. In practice, Equation (3.8e) decides the saturation point ensuring an upper bound for each U_i''. In addition, quadratic cost functions for generators are widely used as a good approximation in OPF problems [6, 16]. Finally, the constraints on $|P_{TCi}^{\max}|, |P_{TCi}^{\min}|$ are in accord with Assumption 3.2.

Remark 3.1 *Similar to Remark 2.4, problem (3.8) can be reformulated as*

$$\min_{P_{C_i}, P_{L_i}, \alpha} \quad \sum_{i \in \mathcal{G}} C_i(P_{C_i}) - \sum_{i \in \mathcal{L}} U_i(P_{L_i}) \tag{3.9a}$$

$$\text{subject to} \quad P_C - d_g - \Gamma_1 f(\alpha) = \mathbf{0} \tag{3.9b}$$

$$- P_L - d_l - \Gamma_2 f(\alpha) = \mathbf{0} \tag{3.9c}$$

$$P_C^{min} \preceq P_C \preceq P_C^{max} \tag{3.9d}$$

$$P_L^{min} \preceq P_L \preceq P_L^{max} \tag{3.9e}$$

$$P_{TC}^{min} \preceq g(A^T \alpha) \preceq P_{TC}^{max} \tag{3.9f}$$

regardless of P_M since $P_C = P_M$ holds in steady state for system (3.7).

Remark 3.2 *The choice of $P_{L_i}^{min}$ decides to what extent loads can respond to feedback signals. For example, for some i, if $0 = P_{L_i}^{min} < P_{L_i}^{max}$ holds, this load is fully responsive; if $P_{L_i}^{min} = P_{L_i}^{max}$ holds, this load has fixed consumption and is unresponsive.*

To conclude, the goal is to design a distributed control scheme (design P_C and P_L) to asymptotically stabilize the system (3.7); the equilibrium should be the optimal solution of the OPF problem (3.8)/(3.9).

3.2 Control Scheme Design

3.2.1 Design Methodology

Before designing the controller, we present a theorem relating to the OPF problem, according to Theorem 3 in [84].

Theorem 3.1 *Under Assumptions 3.2-3.3, the OPF problem (3.8)/(3.9) can be convexified and has a unique optimal solution. Moreover, strong duality holds.*

Proof: Based on Assumptions 3.2-3.3 and Theorem 3 in [84], since the network is a tree, the OPF problem can be convexified. In fact, because $f(\alpha) = Ag(A^T\alpha)$ is always true, Equations (3.8b)-(3.8f) can be rewritten as linear constraints regarding $P_C, P_L, g(A^T\alpha)$. Under Assumption 3.2, as A is the reduced incidence matrix of a tree, thus invertible, α can be uniquely determined given $g(A^T\alpha)$. So problem (3.8) can be convexified. Note that as A is invertible, $g(A^T\alpha)$ can be uniquely derived from (3.8b)-(3.8c) given P_M, P_L, d_g, d_l, i.e.,

$$g(A^T\alpha) = A^{-1} \begin{bmatrix} I_{m+n-1} \\ \mathbf{0} \end{bmatrix}^T \begin{bmatrix} P_M - d_g \\ -P_L - d_l \end{bmatrix}.$$

Under Assumption 3.3 (utility functions are strictly concave and cost functions are strictly convex), the optimal solution of the convexified problem is unique, indicating the uniqueness of the optimal solution of problem (3.8). Moreover, the convexified problem satisfies Slater's condition (it is feasible under linear constraints), therefore, strong duality holds [77]. This leads to strong duality for problem (3.8) under Assumption 3.2 since there is a bijection between $g(A^T\alpha)$ and α. ■

The Lagrangian of the OPF problem (3.9) is then given by

$$\begin{aligned}
&L_{(3.9)}(P_C, P_L, \alpha, \zeta, \lambda, \mu^+, \mu^-, \nu^+, \nu^-, l^+, l^-) = \\
&\sum_{i\in\mathcal{G}} C_i(P_{C_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) + \zeta^T(P_C - d_g - \Gamma_1 f(\alpha)) + \lambda^T(-P_L - d_l - \Gamma_2 f(\alpha)) \\
&+ \mu^{+T}(P_C - P_C^{\max}) + \mu^{-T}(P_C^{\min} - P_C) + \nu^{+T}(P_L - P_L^{\max}) + \nu^{-T}(P_L^{\min} - P_L) \\
&+ l^{+T}(g(A^T\alpha) - P_{TC}^{\max}) + l^{-T}(P_{TC}^{\min} - g(A^T\alpha))
\end{aligned} \tag{3.10}$$

where $\zeta, \mu^+, \mu^- \in \mathbb{R}^m$ with entries $\zeta_i, \mu_i^+, \mu_i^-$ respectively, $\lambda, \nu^+, \nu^- \in \mathbb{R}^n$ with entries $\lambda_i, \nu_i^+, \nu_i^-$ respectively, and $l^+, l^- \in \mathbb{R}^{m+n-1}$ with entries l_i^+, l_i^- respectively are Lagrange multipliers (dual variables) for the constraints (3.8b)-(3.8f). According to Theorem 3.1, the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient conditions for optimality [77], and are given by

$$\frac{\partial L_{(3.9)}}{\partial P_C} = C'(P_C) + \zeta + \mu^+ - \mu^- = \mathbf{0} \tag{3.11a}$$

$$\frac{\partial L_{(3.9)}}{\partial P_L} = -U'(P_L) - \lambda + \nu^+ - \nu^- = \mathbf{0} \tag{3.11b}$$

$$\frac{\partial L_{(3.9)}}{\partial \zeta} = P_C - d_g - \Gamma_1 f(\alpha) = \mathbf{0} \tag{3.11c}$$

$$\frac{\partial L_{(3.9)}}{\partial \lambda} = -P_L - d_l - \Gamma_2 f(\alpha) = \mathbf{0} \tag{3.11d}$$

$$\mathrm{diag}(\mu^+)(P_C - P_C^{\max}) = \mathbf{0},\ \mu^+, P_C^{\max} - P_C \succeq 0 \tag{3.11e}$$

$$\mathrm{diag}(\mu^-)(P_C^{\min} - P_C) = \mathbf{0},\ \mu^-, P_C - P_C^{\min} \succeq 0 \tag{3.11f}$$

$$\mathrm{diag}(\nu^+)(P_L - P_L^{\max}) = \mathbf{0},\ \nu^+, P_L^{\max} - P_L \succeq 0 \tag{3.11g}$$

$$\mathrm{diag}(\nu^-)(P_L^{\min} - P_L) = \mathbf{0},\ \nu^-, P_L - P_L^{\min} \succeq 0 \tag{3.11h}$$

$$\mathrm{diag}(l^+)(g(A^T\alpha) - P_{TC}^{\max}) = \mathbf{0},\ l^+, P_{TC}^{\max} - g(A^T\alpha) \succeq 0 \tag{3.11i}$$

$$\mathrm{diag}(l^-)(P_{TC}^{\min} - g(A^T\alpha)) = \mathbf{0},\ l^-, g(A^T\alpha) - P_{TC}^{\min} \succeq 0 \tag{3.11j}$$

$$\begin{aligned}\frac{\partial L_{(3.9)}}{\partial \alpha} &= -Ag'(\sigma)(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-) = \mathbf{0} \\ &\Leftrightarrow A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^- = \mathbf{0}\end{aligned} \tag{3.11k}$$

where $C(P_C) \in \mathbb{R}^m$ is the vector of cost functions of generators whose entries are $C_i(P_{C_i}(t))$, $U(P_L) \in \mathbb{R}^n$ is the vector of utility functions of loads whose entries are $U_i(P_{L_i}(t))$, and $g'(\sigma(t)) \in \mathbb{R}^{(m+n-1)\times(m+n-1)}$ is the Jacobian matrix of $g(\sigma)$. The last equation is derived based on the facts that $\sigma = A^T\alpha$ holds, A is invertible, and $g'(\sigma)$ is positive diagonal (since $\cos\sigma_k > 0$ and σ_k is bounded by Assumption 3.2).

The design methodology of the distributed controller is motivated by a primal-dual gradient algorithm [56]/saddle point algorithm [70] (see Chapter 5 and the Appendix). What should be underlined is that in the OPF problem, α is a decision vector. However, since we are designing real-time optimization, i.e., the inputs are only P_C and P_L, the dynamics of α can not be designed as in a standard primal-dual approach, but are described by (3.7b). The key idea is to insert (3.11k) into the dynamics of ζ, λ, l^+, l^- so that it could be automatically satisfied for the closed-loop system in steady state, which is necessary to achieve OPF for the system (this idea was also used in [5]). We then design the following controller

$$\dot{P}_C = K_{P_C}(-C'(P_C) - \zeta - \mu^+ + \mu^- + R(P_M - P_C)) \tag{3.12a}$$

$$\dot{P}_L = K_{P_L}(U'(P_L) + \lambda - \nu^+ + \nu^- + \omega_l) \tag{3.12b}$$

$$\dot{\zeta} = K_\zeta(P_C - d_g - \Gamma_1 f(\alpha) - \Gamma_1 AK(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-)) \tag{3.12c}$$

$$\dot{\lambda} = K_\lambda(-P_L - d_l - \Gamma_2 f(\alpha) - \Gamma_2 AK(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-)) \tag{3.12d}$$

$$\dot{\mu}^+ = K_{\mu^+}(P_C - P_C^{\max})^+_{\mu^+} \tag{3.12e}$$

$$\dot{\mu}^- = K_{\mu^-}(P_C^{\min} - P_C)^+_{\mu^-} \tag{3.12f}$$

$$\dot{\nu}^+ = K_{\nu^+}(P_L - P_L^{\max})^+_{\nu^+} \tag{3.12g}$$

$$\dot{\nu}^- = K_{\nu^-}(P_L^{\min} - P_L)^+_{\nu^-} \tag{3.12h}$$

$$\dot{l}^+ = K_{l^+}(g(A^T\alpha) - P_{TC}^{\max} + K(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-))^+_{l^+} \tag{3.12i}$$

$$\dot{l}^- = K_{l^-}(P_{TC}^{\min} - g(A^T\alpha) - K(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-))^+_{l^-} \tag{3.12j}$$

where $K_{P_C}, K_\zeta, K_{\mu^+}, K_{\mu^-} \in \mathbb{R}^{m\times m}$, $K_{P_L}, K_\lambda, K_{\nu^+}, K_{\nu^-} \in \mathbb{R}^{n\times n}$ and $K, K_{l^+}, K_{l^-} \in \mathbb{R}^{(m+n-1)\times(m+n-1)}$ are positive diagonal matrices, all representing the controller gains. Note that similar to the designed controllers in Chapter 2, the terms $R(P_M - P_C)$ and ω_l have been added to $\dot{P}_C$ and $\dot{P}_L$ respectively (the idea of introducing these terms follows from Equation (5.11) in Chapter 5). Also for simplicity, we have used vector forms of positive projection in (3.12e)-(3.12j), e.g., $(x)^+_{l^+} = [([x]_1)^+_{l_1^+}, ([x]_2)^+_{l_2^+}, \cdots]^T$.

In Equations (3.12c)-(3.12d), the information of d_g, d_l is needed. Since the disturbance injection is usually uncertain and/or hard to measure, we modify these two equations so that the implementation of the above controller is regardless of d_g, d_l:

$$\dot{\zeta} = K_\zeta(M\dot{\omega}_g + D_g\omega_g + P_C - P_M - \Gamma_1 AK(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-)) \tag{3.13a}$$

$$\dot{\lambda} = K_\lambda(D_l\omega_l - \Gamma_2 AK(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-)) \tag{3.13b}$$

where we have substituted system dynamics (3.7) into (3.12c)-(3.12d).

Remark 3.3 *The operation of the controller is completely distributed: each generator uses (3.12a), (3.12e) and (3.12f) to calculate the command input based on local information, i.e., cost function, mechanical power input, capacity constraints, and the feedback signal from the generator bus it is connected to, and so does each load using (3.12b), (3.12g) and (3.12h) based on local information and the feedback signal from the load bus it is connected to. Generator buses, load buses, and transmission line substations calculate feedback signals respectively with local information, i.e., power imbalance, frequency deviation, line capacity constraints and related signal*

aggregation, using (3.12c)/(3.13a), (3.12d)/(3.13b), (3.12i) and (3.12j). The signals exchanged between neighbours are ζ, λ, l^+, l^-, as illustrated in Figure 3.3. These are in accord with the architecture shown in Figure 3.1.

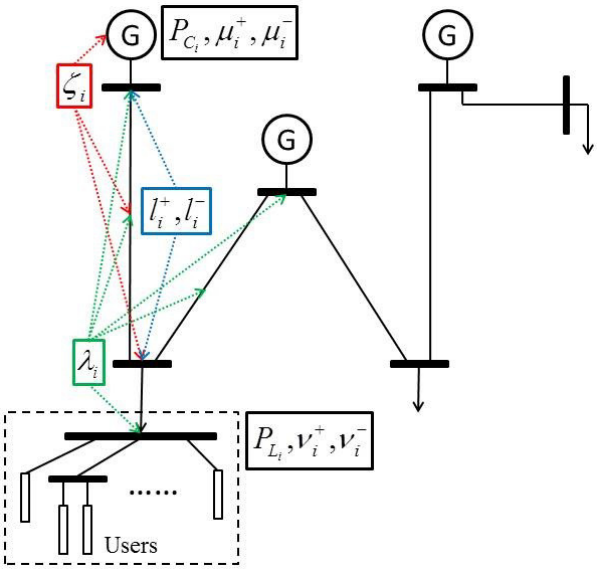

Figure 3.3: The operation of the control scheme.

3.2.2 Main Results

We now consider the overall system (3.7) and (3.12).

Theorem 3.2 *Under Assumptions 3.2-3.3, the equilibria of system (3.7) and (3.12) satisfy the KKT conditions (3.11).*

Proof: Comparing (3.7) and (3.12) with (3.11), the equilibria of the overall system satisfy (3.11) if and only if $\omega_i = 0, i \in \mathcal{G} \bigcup \mathcal{L}$ and $A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^- = \mathbf{0}$ hold. From (3.12c) and (3.12d) ($\dot{\zeta} = \mathbf{0}, \dot{\lambda} = \mathbf{0}$), we have $\mathbf{1}(P_C - d_g) - \mathbf{1}(P_L + d_l) = 0$. Substituting this equality into (3.7) ($\dot{\omega}_g = \mathbf{0}, \dot{\alpha} = \mathbf{0}, \dot{P}_M = \mathbf{0}$), we have $\omega_i = 0, i \in \mathcal{G} \bigcup \mathcal{L}, P_M - d_g - \Gamma_1 f(\alpha) = \mathbf{0}, -P_L - d_l - \Gamma_2 f(\alpha) = \mathbf{0}, P_C = P_M$ since $D_g, D_l, R \succ 0$ is true. Therefore, $A_0 K(A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-) = \mathbf{0}$ holds which leads to $A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^- = \mathbf{0}$. ■

Remark 3.4 *System (3.7) and (3.12) may have multiple equilibria since there may be dual degeneracy in the OPF problem. For example, if $P_{C_i}^{max} - P_{C_i}^* = P_{TC_k}^{max} - T_k \sin \sigma_k^* = 0$ holds and the i^{th} generator bus is only connected to the k^{th} transmission line, there are degenerate solutions/multiple solutions for $\zeta_i^*, \mu_i^{+*}, l_k^{+*}$ from (3.11). However, all equilibria share the same quantity in P_C^*, P_L^*, α^*, which is in accord with Theorem 3.1.*

Theorem 3.2 shows that the designed controller can ensure OPF for the network in the steady state. We next prove convergence, i.e. the stability of the equilibria.

Lemma 3.2 *Given $Z = diag\{z_i\} \in \mathbb{R}^{m\times m}$ where each z_i is positive, the following inequality holds:*

$$(P_M - \Gamma_1 f(\alpha))^T Z(Z+I_m)^{-1} D_g^{-1}(P_M - \Gamma_1 f(\alpha)) + (P_L + \Gamma_2 f(\alpha))^T D_l^{-1}(P_L + \Gamma_2 f(\alpha)) + P_M^T C_g P_M - P_L^T C_l P_L \geq g(A^T\alpha)^T A_0^T \Xi A_0 g(A^T\alpha) \tag{3.14}$$

where

$$\Xi = \begin{bmatrix} ZC_g(Z+(Z+I_m)D_gC_g)^{-1} & \mathbf{0} \\ \mathbf{0} & -C_l(I_n - D_lC_l)^{-1} \end{bmatrix}. \tag{3.15}$$

In the above equation, $C_g = diag\{\frac{c_iR_i}{c_i+R_i}, i \in \mathcal{G}\} \in \mathbb{R}^{m\times m}$ and $C_l = diag\{c_i, i \in \mathcal{L}\} \in \mathbb{R}^{n\times n}$ ($c_i, i \in \mathcal{G}\bigcup\mathcal{L}$ are defined in Assumption 3.3).

Proof: Denote $X = [P_M^T, P_L^T, f(\alpha)^T\Gamma_1^T, f(\alpha)^T\Gamma_2^T]^T$. The left hand side of the inequality equals

$$X^TQX = X^T \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} X$$

where

$$Q_{11} = \begin{bmatrix} Z(Z+I_m)^{-1}D_g^{-1} + C_g & \mathbf{0} \\ \mathbf{0} & D_l^{-1} - C_l \end{bmatrix}$$
$$Q_{12} = \begin{bmatrix} -Z(Z+I_m)^{-1}D_g^{-1} & \mathbf{0} \\ \mathbf{0} & D_l^{-1} \end{bmatrix}$$
$$Q_{22} = \begin{bmatrix} Z(Z+I_m)^{-1}D_g^{-1} & \mathbf{0} \\ \mathbf{0} & D_l^{-1} \end{bmatrix}.$$

Based on a Schur Complement [77], we have

$$\begin{aligned} X^TQX &\geq \begin{bmatrix} \Gamma_1 f(\alpha) \\ \Gamma_2 f(\alpha) \end{bmatrix}^T (Q_{22} - Q_{12}^TQ_{11}^{-1}Q_{12}) \begin{bmatrix} \Gamma_1 f(\alpha) \\ \Gamma_2 f(\alpha) \end{bmatrix} \\ &= g(A^T\alpha)^T A_0^T \Xi A_0 g(A^T\alpha) \end{aligned}$$

which completes the proof. ■

Theorem 3.3 *Under Assumptions 3.2-3.3, each equilibrium point of system (3.7) and (3.12) is locally asymptotically stable if*

$$4A_0^T\Xi A_0 - K^{-1} \succ 0 \tag{3.16}$$

where Ξ is defined in Lemma 3.2 and $Z = diag\{z_i\} \in \mathbb{R}^{m\times m} \succ 0$ satisfies

$$MZD_g^{-1}(\Gamma_1 ATA^T\Gamma_1^T + 2T_{TG}^{-1}R^{-1})D_g^{-1}ZM - 4ZM - 2M \prec 0. \tag{3.17}$$

Proof: Define a candidate Lyapunov function for the overall system as

$$
\begin{aligned}
&V_{(3.7)(3.12)} = \\
&\frac{1}{2}\omega_g^T M\omega_g - \int_{\mathbf{0}}^{\omega_g} \dot{\omega}_g(\beta)^T M^2 Z D_g^{-1} d\beta + \int_{\alpha^*}^{\alpha} (f(\beta) - f(\alpha^*))^T d\beta \\
&+ \frac{1}{2}(P_M - P_M^*)^T T_{TG} R (P_M - P_M^*) + \frac{1}{2}(P_C - P_C^*)^T K_{P_C}^{-1}(P_C - P_C^*) \\
&+ \frac{1}{2}(P_L - P_L^*)^T K_{P_L}^{-1}(P_L - P_L^*) + \frac{1}{2}(\zeta - \zeta^*)^T K_\zeta^{-1}(\zeta - \zeta^*) + \frac{1}{2}(\lambda - \lambda^*)^T K_\lambda^{-1}(\lambda - \lambda^*) \\
&+ \frac{1}{2}(\mu^+ - \mu^{+*})^T K_{\mu^+}^{-1}(\mu^+ - \mu^{+*}) + \frac{1}{2}(\mu^- - \mu^{-*})^T K_{\mu^-}^{-1}(\mu^- - \mu^{-*}) \\
&+ \frac{1}{2}(\nu^+ - \nu^{+*})^T K_{\nu^+}^{-1}(\nu^+ - \nu^{+*}) + \frac{1}{2}(\nu^- - \nu^{-*})^T K_{\nu^-}^{-1}(\nu^- - \nu^{-*}) \\
&+ \frac{1}{2}(l^+ - l^{+*})^T K_{l^+}^{-1}(l^+ - l^{+*}) + \frac{1}{2}(l^- - l^{-*})^T K_{l^-}^{-1}(l^- - l^{-*}) \qquad (3.18)
\end{aligned}
$$

where $Z = \text{diag}\{z_i\} \in \mathbb{R}^{m\times m} \succ 0$. We first show that by choosing Z appropriately, $V_{(3.7)(3.12)} \geq 0$ holds around the given equilibrium point. At the equilibrium point, we have $V_{(3.7)(3.12)} = 0$. Considering sufficiently small deviations around the equilibrium point that the dynamics of the overall system are linear, we have

$$
\begin{aligned}
&V^+_{(3.7)(3.12)} \\
&= \frac{1}{2}\omega_g^T M\omega_g - \dot{\omega}_g^T M^2 Z D_g^{-1}\omega_g + (\sigma - \sigma^*)^T T(\sigma - \sigma^*) + \frac{1}{2}(P_M - P_M^*)^T T_{TG} R (P_M - P_M^*) \\
&\quad + (\star) \\
&= \frac{1}{2}\omega_g^T M\omega_g - (-\omega_g^T MZ\omega_g + (P_M - P_M^*)^T MZD_g^{-1}\omega_g - (\sigma - \sigma^*)^T TA^T\Gamma_1^T MZD_g^{-1}\omega_g) \\
&\quad + (\sigma - \sigma^*)^T T(\sigma - \sigma^*) + \frac{1}{2}(P_M - P_M^*)^T T_{TG} R (P_M - P_M^*) + (\star) \qquad (3.19)
\end{aligned}
$$

where $(\star)$ represents non-negative quadratic items. $V^+_{(3.7)(3.12)} \geq 0$ requires

$$
\begin{bmatrix}
(Z + \frac{1}{2}I_m)M & \frac{1}{2}D_g^{-1}ZM\Gamma_1 AT & -\frac{1}{2}MZD_g^{-1} \\
\frac{1}{2}TA^T\Gamma_1^T MZD_g^{-1} & T & \mathbf{0} \\
-\frac{1}{2}MZD_g^{-1} & \mathbf{0} & \frac{1}{2}T_{TG}R
\end{bmatrix} \succ 0.
$$

Based on a Schur Complement [77], solving the inequality (3.17) yields the feasible Z making that $V_{(3.7)(3.12)} \geq 0$ holds around the equilibrium point and $V_{(3.7)(3.12)} = 0$ holds only at the equilibrium point, i.e., $0 \prec Z \prec Z_{\max}$.

Next we show $\dot{V}_{(3.7)(3.12)} \leq 0$ (although $\dot{V}_{(3.7)(3.12)}$ may be discontinuous, we only need $\dot{V}_{(3.7)(3.12)} \leq 0$ for asymptotic stability). Note that

$$
\begin{aligned}
(\mu^+ - \mu^{+*})^T K_{\mu^+}^{-1} \dot{\mu}^+ =& (\mu^+ - \mu^{+*})^T K_{\mu^+}^{-1} K_{\mu^+} (P_C - P_C^{\max})_{\mu^+}^+ \\
\overset{\text{(i)}}{\leq}& (\mu^+ - \mu^{+*})^T (P_C - P_C^{\max}) \\
=& (\mu^+ - \mu^{+*})^T (P_C - P_C^*) + (\mu^+ - \mu^{+*})^T (P_C^* - P_C^{\max}) \\
\overset{\text{(ii)}}{\leq}& (\mu^+ - \mu^{+*})^T (P_C - P_C^*)
\end{aligned}
$$

holds since: (i) when $(\star)_{\mu^+}^+$ is inactive, the item in the bracket is non-positive (otherwise, it is positive and then $(\star)_{\mu^+}^+$ would be active which indicates the equality case), and each entry of $(\mu^+ - \mu^{+*})$ is also non-positive so that $(\mu^+ - \mu^{+*})^T (\star)_{\mu^+}^+ \leq (\mu^+ - \mu^{+*})^T (\star)$ is always true; (ii) when $\mu^{+*} = \mathbf{0}$, $P_C^* \preceq P_C^{\max}$ is always true (otherwise, $P_C^* = P_C^{\max}$). Similarly, all positive projection can be removed in $\dot{V}_{(3.7)(3.12)}$. We move the equilibrium point to the origin and continue to use the same variable names, however, these are now deviations from the equilibrium point. It can be shown that

$$
\begin{aligned}
\dot{V}_{(3.7)(3.12)} \leq& \omega_g^T M \dot{\omega}_g - \dot{\omega}_g^T M^2 Z D_g^{-1} \dot{\omega}_g + f(\alpha)^T \dot{\alpha} + P_M^T T_{TG} R \dot{P}_M + P_C^T K_{P_C}^{-1} \dot{P}_C \\
&+ P_L^T K_{P_L}^{-1} \dot{P}_L + \zeta^T K_\zeta^{-1} \dot{\zeta} + \lambda^T K_\lambda^{-1} \dot{\lambda} + \mu^{+T} P_C - \mu^{-T} P_C + \nu^{+T} P_L - \nu^{+T} P_L \\
&+ l^{+T} (g(A^T \alpha) + K(A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)) \\
&+ l^{-T} (-g(A^T \alpha) - K(A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)) \\
=& - \omega_g^T D_g \omega_g - (D_g \omega_g - P_M + \Gamma_1 f(\alpha))^T Z D_g^{-1} (D_g \omega_g - P_M + \Gamma_1 f(\alpha)) \\
&- (P_L + \Gamma_2 f(\alpha))^T D_l^{-1} (P_L + \Gamma_2 f(\alpha)) - (P_M - P_C)^T R (P_M - P_C) \\
&- P_C^T C'(P_C) + P_L^T U'(P_L) - (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)^T g(A^T \alpha) \\
&- (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)^T K (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-) \\
\leq& - (D_g (Z + I_m) \omega_g - Z(P_M - \Gamma_1 f(\alpha)))^T (Z + I_m)^{-1} D_g^{-1} (D_g (Z + I_m) \omega_g \\
&- Z(P_M - \Gamma_1 f(\alpha))) - (P_M - \Gamma_1 f(\alpha))^T Z (Z + I_m)^{-1} D_g^{-1} (P_M - \Gamma_1 f(\alpha)) \\
&- (P_L + \Gamma_2 f(\alpha))^T D_l^{-1} (P_L + \Gamma_2 f(\alpha)) - \sum_{i \in \mathcal{G}} (c_i + R_i) \left(P_{C_i} - \frac{R_i}{c_i + R_i} P_{M_i} \right)^2 \\
&- P_M^T C_g P_M + P_L^T C_l P_L - (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)^T g(A^T \alpha) \\
&- (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)^T K (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-) \\
\leq& - (D_g (Z + I_m) \omega_g - Z(P_M - \Gamma_1 f(\alpha)))^T (Z + I_m)^{-1} D_g^{-1} (D_g (Z + I_m) \omega_g \\
&- Z(P_M - \Gamma_1 f(\alpha))) - \sum_{i \in \mathcal{G}} (c_i + R_i) \left(P_{C_i} - \frac{R_i}{c_i + R_i} P_{M_i} \right)^2 \\
&- g(A^T \alpha)^T A_0^T \Xi A_0 g(A^T \alpha) - (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)^T g(A^T \alpha) \\
&- (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)^T K (A_0^T [\zeta^T, \lambda^T]^T - l^+ + l^-)
\end{aligned} \tag{3.20}
$$

where C_g, C_l, Ξ are defined in Lemma 3.2, and the last step above is based on Lemma 3.2. Choosing K satisfying (3.16) makes $\dot{V}_{(3.7)(3.12)} \leq 0$. If $\dot{V}_{(3.7)(3.12)} = 0$, it is obvious that $A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^- = \mathbf{0}$, $g(A^T\alpha) = \mathbf{0}$, $P_{C_i} - \frac{R_i}{c_i+R_i}P_{M_i} = 0, \forall i \in \mathcal{G}$, then $\omega_g = P_M = P_C = \mathbf{0}$ and $P_L = \mathbf{0}$ hold, which only happens at the equilibrium point. Using LaSalle's invariance principle [79], each equilibrium point is locally asymptotically stable. ■

Remark 3.5 *Provided bounds of parameters of the system, i.e., bounds of the diagonal entries of $M, D_g, T, T_{TG}, R, D_l, C_g, C_l$, the range of K can be precisely derived using Theorem 3.3.*

Remark 3.6 *The condition (3.16) in Theorem 3.3 is sufficient but not necessary. It indicates that by choosing a large gain matrix K, i.e., the diagonal entries of K are large enough so that $4A_0^T\Xi A_0 - K^{-1} \succ 0$ holds, asymptotic stability of the equilibria is guaranteed. Alternatively, we can regard $A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^- = \mathbf{0}$ in the KKT conditions (3.11) as a "consensus condition", i.e., the k^{th} entry of $A_0^T[\zeta^T, \lambda^T]^T - l^+ + l^-$ corresponds to the k^{th} transmission line and needs to be zero in the steady state. By choosing a large gain matrix K, the overall system is forced to achieve the consensus and converges to an equilibrium point.*

3.2.3 A Special Case: Star Topology

In this subsection, we consider a special case in which the network has a star topology as illustrated in Figure 3.4. It will be shown that the result given by Theorem 3.3 can become much simpler for star networks.

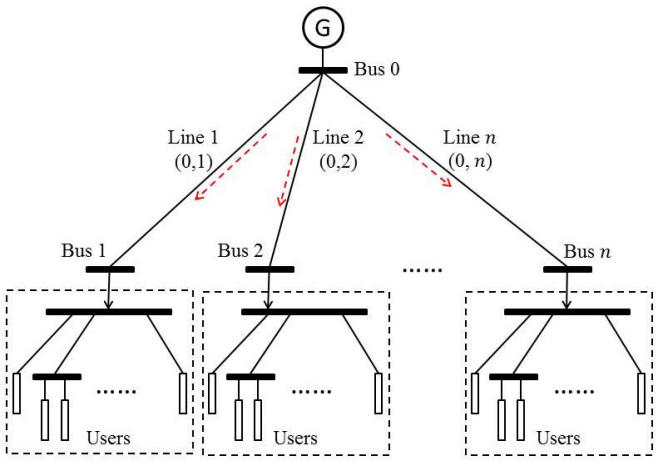

Figure 3.4: A one-generator-n-load network with star topology (the orientation is indicated by red dashed arrows).

For convenience, we number the generator bus 0, the load buses $1, \cdots, n$, and the transmission lines $1, \cdots, n$ (corresponding to a lexicographic ordering, as illustrated in Figure 3.4). Define $\mathcal{L} = \{1, \cdots, n\}$. The closed-loop system is then given by

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \sum_{i\in\mathcal{L}} T_i \sin\alpha_i \tag{3.21a}$$

$$\dot{\alpha}_i = \omega_g + \frac{1}{D_i}(P_{L_i} + d_i - T_i \sin\alpha_i),\ i \in \mathcal{L} \tag{3.21b}$$

$$\dot{P}_M = \frac{1}{T_{TG}}\left(P_C - P_M - \frac{\omega_g}{R}\right) \tag{3.21c}$$

$$\dot{P}_C = K_{P_C}(-C'(P_C) - \zeta - \mu^+ + \mu^- + R(P_M - P_C)) \tag{3.21d}$$

$$\dot{P}_{L_i} = K_{P_{L_i}}(U_i'(P_{L_i}) + \lambda_i - \nu_i^+ + \nu_i^- + \omega_i),\ i \in \mathcal{L} \tag{3.21e}$$

$$\dot{\zeta} = K_\zeta\left(P_C - d_g - \sum_{i\in\mathcal{L}} T_i \sin\alpha_i - \sum_{i\in\mathcal{L}} K_i(\zeta - \lambda_i - l_i^+ + l_i^-)\right) \tag{3.21f}$$

$$\dot{\lambda}_i = K_{\lambda_i}(-P_{L_i} - d_i + T_i \sin\alpha_i + K_i(\zeta - \lambda_i - l_i^+ + l_i^-)),\ i \in \mathcal{L} \tag{3.21g}$$

$$\dot{\mu}^+ = K_{\mu^+}(P_C - P_C^{\max})_{\mu^+}^+ \tag{3.21h}$$

$$\dot{\mu}^- = K_{\mu^-}(P_C^{\min} - P_C)_{\mu^-}^+ \tag{3.21i}$$

$$\dot{\nu}_i^+ = K_{\nu_i^+}(P_{L_i} - P_{L_i}^{\max})_{\nu_i^+}^+,\ i \in \mathcal{L} \tag{3.21j}$$

$$\dot{\nu}_i^- = K_{\nu_i^-}(P_{L_i}^{\min} - P_{L_i})_{\nu_i^-}^+,\ i \in \mathcal{L} \tag{3.21k}$$

$$\dot{l}_i^+ = K_{l_i^+}(T_i \sin\alpha_i - P_{TC_i}^{\max} + K_i(\zeta - \lambda_i - l_i^+ + l_i^-))_{l_i^+}^+,\ i \in \mathcal{L} \tag{3.21l}$$

$$\dot{l}_i^- = K_{l_i^-}(P_{TC_i}^{\min} - T_i \sin\alpha_i - K_i(\zeta - \lambda_i - l_i^+ + l_i^-))_{l_i^-}^+,\ i \in \mathcal{L} \tag{3.21m}$$

where $M > 0$ is the generator inertia, $\omega_g = \dot{\delta}_0$ and $\omega_i = \dot{\delta}_i, i \in \mathcal{L}$ are the bus frequency deviations from the nominal frequency (as before, $v_0\angle\delta_0, v_1\angle\delta_1, \cdots, v_n\angle\delta_n$ are the bus voltages, where v_i is the bus voltage magnitude, and δ_i is the bus voltage phase angle with respect to the rotating framework of nominal frequency), $D_g > 0$ is the generator damping coefficient, $D_i > 0$ is the load bus damping coefficient, P_M is the mechanical power input, d_g and d_i are the disturbance injections, e.g., renewable energy injections and variations on both supply and demand, $T_i = v_0 v_i B_i > 0, i \in \mathcal{L}$, $B_i > 0$ is the susceptance of the transmission line i, $\alpha_i = \delta_0 - \delta_i, i \in \mathcal{L}$, P_{L_i} is the load power consumption, P_C is the command/control input to the generator and $T_{TG}, R > 0$ are constant parameters. For the controller (3.21d)-(3.21m), all states and coefficients remain the same meaning as in (3.12) but now for the star network. The following proposition shows the stability of (3.21).

Proposition 3.1 *Under Assumptions 3.2-3.3, each equilibrium point of system (3.21) is locally asymptotically stable if*

$$K_i > \frac{D_i c_i - 1}{4c_i}, i \in \mathcal{L}. \tag{3.22}$$

Proof: Define a candidate Lyapunov function for the overall system as

$$\begin{aligned}
&V_{(3.21)} = \\
&\frac{1}{2}M\omega_g^2 + \sum_{i\in\mathcal{L}}\int_{\alpha_i^*}^{\alpha_i} T_i(\sin\beta - \sin\alpha_i^*)d\beta + \frac{T_{TG}R}{2}(P_M - P_M^*)^2 + \frac{1}{2K_{P_C}}(P_C - P_C^*)^2 \\
&+ \sum_{i\in\mathcal{L}}\frac{1}{2K_{P_{L_i}}}(P_{L_i} - P_{L_i}^*)^2 + \frac{1}{2K_\zeta}(\zeta - \zeta^*)^2 + \sum_{i\in\mathcal{L}}\frac{1}{2K_{\lambda_i}}(\lambda_i - \lambda_i^*)^2 + \frac{1}{2K_{\mu^+}}(\mu^+ - \mu^{+*})^2 \\
&+ \frac{1}{2K_{\mu^-}}(\mu^- - \mu^{-*})^2 + \sum_{i\in\mathcal{L}}\frac{1}{2K_{\nu_i^+}}(\nu_i^+ - \nu_i^{+*})^2 + \sum_{i\in\mathcal{L}}\frac{1}{2K_{\nu_i^-}}(\nu_i^- - \nu_i^{-*})^2 \\
&+ \sum_{i\in\mathcal{L}}\frac{1}{2K_{l_i^+}}(l_i^+ - l_i^{+*})^2 + \sum_{i\in\mathcal{L}}\frac{1}{2K_{l_i^-}}(l_i^- - l_i^{-*})^2.
\end{aligned} \tag{3.23}$$

Similar to the proof of Theorem 3.3, we remove all positive projection in $\dot{V}_{(3.21)}$, and move the equilibrium point to the origin. We continue to use the same variable names, however, these are now deviations from the equilibrium point. It can be shown that

$$\begin{aligned}
\dot{V}_{(3.21)} \leq & - D_g\omega_g^2 - R(P_M - P_C)^2 - P_C C'(P_C) + \sum_{i\in\mathcal{L}} c_i P_{L_i}^2 - \sum_{i\in\mathcal{L}} K_i(\zeta - \lambda_i - l_i^+ + l_i^-)^2 \\
& - \sum_{i\in\mathcal{L}}\frac{1}{D_i}(P_{L_i} - T_i\sin\alpha_i)^2 - \sum_{i\in\mathcal{L}} T_i\sin\alpha_i(\zeta - \lambda_i - l_i^+ + l_i^-) \\
= & - D_g\omega_g^2 - R(P_M - P_C)^2 + \sum_{i\in\mathcal{L}}\frac{D_i c_i - 1}{D_i}\left(P_{L_i} + \frac{1}{D_i c_i - 1}T_i\sin\alpha_i\right)^2 \\
& - P_C C'(P_C) - \sum_{i\in\mathcal{L}} K_i(\zeta - \lambda_i - l_i^+ + l_i^-)^2 - \sum_{i\in\mathcal{L}} T_i\sin\alpha_i(\zeta - \lambda_i - l_i^+ + l_i^-) \\
& + \sum_{i\in\mathcal{L}}\frac{-c_i}{D_i c_i - 1}(T_i\sin\alpha_i)^2
\end{aligned} \tag{3.24}$$

Choosing $K_i > \frac{D_i c_i - 1}{4c_i}$ makes $\dot{V}_{(3.21)} \leq 0$. If $\dot{V}_{(3.21)} = 0$, $\omega_g = P_M = P_C = P_{L_i} = \zeta - \lambda_i - l_i^+ + l_i^- = \sin\alpha_i = 0$, which only happens at the equilibrium point. Using LaSalle's invariance principle [79], each equilibrium point is locally asymptotically stable. ■

In Theorem 3.3, the stability condition regarding K is given in a centralized form, while it becomes decentralized as shown in Proposition 3.1. This indicates topology variations of a network system can influence its stability condition.

3.3 Performance Improvement

Although the primal-dual approach provides a framework to design control schemes, there are no guarantees on the transient behaviour of designed controller. In this section, under the same framework, we introduce extra dynamics [87, 88] that can be used to improve the performance of the previous controller.

3.3.1 Extra Dynamics

We rewrite the objective of the OPF problem (3.9) as

$$\min_{P_{C_i},P_{L_i},\alpha,\hat{P}_{C_i},\hat{P}_{L_i}} \quad \sum_{i\in\mathcal{G}} C_i(P_{C_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) + \sum_{i\in\mathcal{G}} \frac{K_{eP_{C_i}}}{2}(P_{C_i}-\hat{P}_{C_i})^2 + \sum_{i\in\mathcal{L}} \frac{K_{eP_{L_i}}}{2}(P_{L_i}-\hat{P}_{L_i})^2 \tag{3.25}$$

and keep the constraints unchanged. In the above objective, $\hat{P}_{C_i}, \hat{P}_{L_i}$ are extra decision variables and $K_{eP_{C_i}}, K_{eP_{L_i}} > 0$. It is clear that the new problem (3.25) and (3.9b)-(3.9f) is equivalent to the original one (3.9), with the optimal solution and $\hat{P}^*_{C_i} = P^*_{C_i}, \hat{P}^*_{L_i} = P^*_{L_i}$. Using the design methodology presented previously yields a new controller that consists of (3.12c)-(3.12j) and

$$\dot{P}_C = K_{P_C}(-C'(P_C) - \zeta - \mu^+ + \mu^- + R(P_M - P_C) - K_{eP_C}(P_C - \hat{P}_C)) \tag{3.26a}$$

$$\dot{\hat{P}}_C = \hat{K}_{eP_C}(P_C - \hat{P}_C) \tag{3.26b}$$

$$\dot{P}_L = K_{P_L}(U'(P_L) + \lambda - \nu^+ + \nu^- + \omega_l - K_{eP_L}(P_L - \hat{P}_L)) \tag{3.26c}$$

$$\dot{\hat{P}}_L = \hat{K}_{eP_L}(P_L - \hat{P}_L) \tag{3.26d}$$

where $K_{eP_C} = \text{diag}\{K_{eP_{C_i}}\} \in \mathbb{R}^{m\times m}, K_{eP_L} = \text{diag}\{K_{eP_{L_i}}\} \in \mathbb{R}^{n\times n}$, and $\hat{K}_{eP_C} \in \mathbb{R}^{m\times m}, \hat{K}_{eP_L} \in \mathbb{R}^{n\times n}$ are positive diagonal matrices, all representing the controller gains. Compared with (3.12), extra dynamics $\dot{\hat{P}}_C, \dot{\hat{P}}_L$ are introduced. We then have the following proposition.

Proposition 3.2 *Under Assumptions 3.2-3.3, each equilibrium point of system (3.7), (3.12c)-(3.12j), and (3.26) satisfies the KKT conditions (3.11) and is locally asymptotically stable if*

$$4A_0^T \Xi A_0 - K^{-1} \succ 0$$

where Ξ is defined in Lemma 3.2 and $Z = diag\{z_i\} \in \mathbb{R}^{m\times m} \succ 0$ satisfies

$$MZD_g^{-1}(\Gamma_1 ATA^T\Gamma_1^T + 2T_{TG}^{-1}R^{-1})D_g^{-1}ZM - 4ZM - 2M \prec 0.$$

Proof: The proof is similar to that of Theorems 3.2 and 3.3, using the Lyapunov function given by

$$V_{(3.7),(3.12c)\text{-}(3.12j),(3.26)} = V_{(3.7)(3.12)} + \frac{1}{2}(\hat{P}_C - \hat{P}_C^*)^T \hat{K}_{eP_C}^{-1} K_{eP_C}(\hat{P}_C - \hat{P}_C^*) + \frac{1}{2}(\hat{P}_L - \hat{P}_L^*)^T \hat{K}_{eP_L}^{-1} K_{eP_L}(\hat{P}_L - \hat{P}_L^*) \quad (3.27)$$

and the fact $\hat{P}_C^* = P_C^*, \hat{P}_L^* = P_L^*$. ■

Remark 3.7 *Similar modifications can also be applied to the control schemes proposed in Chapter 2, i.e., Equation (2.15d) can be redesigned as*

$$\dot{P}_C = B(P_M - P_C) - KA_0A_0^T CP_C - K_{eP_C}(P_C - \hat{P}_C)$$
$$\dot{\hat{P}}_C = \hat{K}_{eP_C}(P_C - \hat{P}_C),$$

Equations (2.27d)-(2.27e) can be redesigned as

$$\dot{P}_C = B(P_M - P_C) - \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix}^T KA_0A_0^T \begin{bmatrix} CP_C \\ P'_C \end{bmatrix} - K_{eP_C}(P_C - \hat{P}_C)$$
$$\dot{\hat{P}}_C = \hat{K}_{eP_C}(P_C - \hat{P}_C)$$
$$\dot{P}'_C = -\begin{bmatrix} \mathbf{0} \\ I_n \end{bmatrix}^T KA_0A_0^T \begin{bmatrix} CP_C \\ P'_C \end{bmatrix} - K_{eP'_C}(P'_C - \hat{P}'_C)$$
$$\dot{\hat{P}}'_C = \hat{K}_{eP'_C}(P'_C - \hat{P}'_C)$$

and Equations (2.35d)-(2.35e) can be redesigned as

$$\dot{P}_C = B_g(P_M - P_C) - \begin{bmatrix} I_m \\ \mathbf{0} \end{bmatrix}^T KA_0A_0^T \begin{bmatrix} C_gP_C \\ -C_lP_L \end{bmatrix} - K_{eP_C}(P_C - \hat{P}_C)$$
$$\dot{\hat{P}}_C = \hat{K}_{eP_C}(P_C - \hat{P}_C)$$
$$\dot{P}_L = B_lD_l^{-1}(-P_L - d_l - \Gamma_2 ATA^T\alpha) + \begin{bmatrix} \mathbf{0} \\ I_n \end{bmatrix}^T KA_0A_0^T \begin{bmatrix} C_gP_C \\ -C_lP_L \end{bmatrix} - K_{eP_L}(P_L - \hat{P}_L)$$
$$\dot{\hat{P}}_L = \hat{K}_{eP_L}(P_L - \hat{P}_L)$$

where $\hat{P}_C, \hat{P}'_C, \hat{P}_L$ are auxiliary state vectors and $K_{eP_C}, \hat{K}_{eP_C}, K_{eP'_C}, \hat{K}_{eP'_C}, K_{eP_L}, \hat{K}_{eP_L}$ are positive diagonal constant matrices with appropriate dimensions.

3.3.2 The Trade-off

We next investigate how the gains of the extra dynamics, i.e., $K_{eP_C}, \hat{K}_{eP_C}, K_{eP_L}, \hat{K}_{eP_L}$, affect system (linear) robustness. We linearize the overall system and rearrange it as a unity negative feedback system, as illustrated in Figure 3.5.

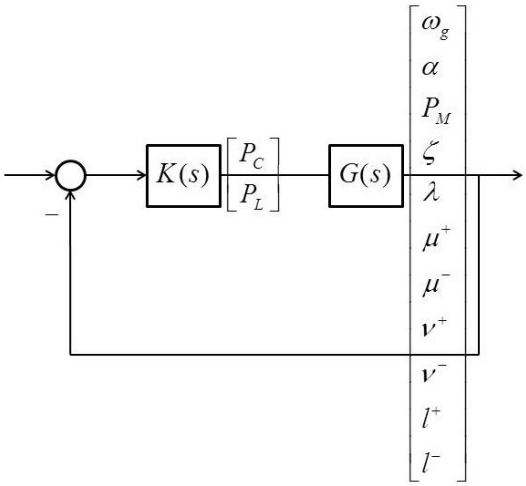

Figure 3.5: The overall feedback loop configuration.

Denote the plant by $G(s)$ whose dynamics are described by (3.7) and (3.12c)-(3.12j), and the controller by $K(s)$ whose dynamics are described by (3.26) (now $K(s)$ stands for a transfer function matrix but not a gain matrix). The outputs of $K(s)$, i.e., the inputs of $G(s)$, are P_C and P_L. Based on classical loop shaping methods (see the Appendix), $\underline{\sigma}(G(s)K(s))$ needs to be large at low frequencies for better performance and $\overline{\sigma}(G(s)K(s))$ to be small at high frequencies for better robust stability, noise attenuation and control energy reduction [89]. Due to

$$\underline{\sigma}(G(s))\underline{\sigma}(K(s)) \leq \underline{\sigma}(G(s)K(s))$$
$$\overline{\sigma}(G(s)K(s)) \leq \overline{\sigma}(G(s))\overline{\sigma}(K(s))$$

a sufficient condition is to make $\underline{\sigma}(K(s))$ large at low frequencies and $\overline{\sigma}(K(s))$ small at high frequencies. Note that $K(s)$ can be written as

$$K(s) = \text{diag}\,(K_1(s), \cdots, K_m(s), K_{m+1}(s), \cdots, K_{m+n}(s))\, B$$

where B is a matrix derived from (3.26) relating to the input of $K(s)$ and is independent of $K_{eP_C}, \hat{K}_{eP_C}, K_{eP_L}, \hat{K}_{eP_L}$, and $K_i(s)$ is given by

$$K_i(s) = \frac{K_{P_{C_i}}(s + \hat{K}_{eP_{C_i}})}{s^2 + (K_{P_{C_i}}\kappa_i + K_{P_{C_i}}K_{eP_{C_i}} + \hat{K}_{eP_{C_i}})s + K_{P_{C_i}}\kappa_i\hat{K}_{eP_{C_i}}},\ i \in \mathcal{G} \tag{3.28a}$$

$$K_i(s) = \frac{K_{P_{L_i}}(s + \hat{K}_{eP_{L_i}})}{s^2 + (K_{P_{L_i}}\kappa_i + K_{P_{L_i}}K_{eP_{L_i}} + \hat{K}_{eP_{L_i}})s + K_{P_{L_i}}\kappa_i\hat{K}_{eP_{L_i}}},\ i \in \mathcal{L} \tag{3.28b}$$

where $\kappa_i = c_i + R_i, i \in \mathcal{G}$ and $\kappa_i = -U_i''(P_{L_i}^*) + D_i^{-1}, i \in \mathcal{L}$. Due to

$$\underline{\sigma}(\text{diag}\,(K_1(s), \cdots, K_{m+n}(s)))\underline{\sigma}(B) \leq \underline{\sigma}(K(s))$$
$$\overline{\sigma}(K(s)) \leq \overline{\sigma}(\text{diag}\,(K_1(s), \cdots, K_{m+n}(s)))\overline{\sigma}(B)$$

we can obtain that $\underline{\sigma}(K_i(s))$ should be large at low frequencies and $\overline{\sigma}(K_i(s))$ should be small at high frequencies, $\forall i \in \mathcal{G} \bigcup \mathcal{L}$. We then have the following remark.

Remark 3.8 *Let j be the imaginary unit. Since*

$$\frac{\partial |K_i(j\omega)|^2}{\partial K_{eP_{C_i}}} < 0, \ i \in \mathcal{G}$$

$$\frac{\partial |K_i(j\omega)|^2}{\partial K_{eP_{L_i}}} < 0, \ i \in \mathcal{L}$$

$$\frac{\partial |K_i(j\omega)|^2}{\partial \hat{K}_{eP_{C_i}}} = \frac{2K_{P_{C_i}}^3 K_{eP_{C_i}} \omega^2 (\hat{K}_{eP_{C_i}} (\hat{K}_{eP_{C_i}} + 2K_{P_{C_i}} \kappa_i + K_{P_{C_i}} K_{eP_{C_i}}) - \omega^2)}{\left((\omega^2 - K_{P_{C_i}} \kappa_i \hat{K}_{eP_{C_i}})^2 + (K_{P_{C_i}} \kappa_i + K_{P_{C_i}} K_{eP_{C_i}} + \hat{K}_{eP_{C_i}})^2 \omega^2 \right)^2}, \ i \in \mathcal{G}$$

$$\frac{\partial |K_i(j\omega)|^2}{\partial \hat{K}_{eP_{L_i}}} = \frac{2K_{P_{L_i}}^3 K_{eP_{L_i}} \omega^2 (\hat{K}_{eP_{L_i}} (\hat{K}_{eP_{L_i}} + 2K_{P_{L_i}} \kappa_i + K_{P_{L_i}} K_{eP_{L_i}}) - \omega^2)}{\left((\omega^2 - K_{P_{L_i}} \kappa_i \hat{K}_{eP_{L_i}})^2 + (K_{P_{L_i}} \kappa_i + K_{P_{L_i}} K_{eP_{L_i}} + \hat{K}_{eP_{L_i}})^2 \omega^2 \right)^2}, \ i \in \mathcal{L}$$

where we have abused the notation ω, small $K_{eP_{C_i}}, K_{eP_{L_i}}$ increase the lower bound of $\underline{\sigma}(K(s))$ at low frequencies while also increasing the upper bound of $\overline{\sigma}(K(s))$ at high frequencies. Large $\hat{K}_{eP_{C_i}}, \hat{K}_{eP_{L_i}}$ increase the lower bound of $\underline{\sigma}(K(s))$ at low frequencies and decrease the upper bound of $\overline{\sigma}(K(s))$ at high frequencies.

This means that in order to improve the performance of the closed-loop system, there are trade-offs for $K_{eP_{C_i}}, K_{eP_{L_i}}$ (not very large or very small) and $\hat{K}_{eP_{C_i}}, \hat{K}_{eP_{L_i}}$ need to be large (on the other hand, as will be shown in Chapter 4, increasing $\hat{K}_{eP_{C_i}}, \hat{K}_{eP_{L_i}}$ too much leads to less influence of the extra dynamics on the closed-loop system).

3.4 Numerical Investigations

We now present a numerical example using the 6-bus network illustrated in Figure 3.2. Table 3.1 shows the parameter values of the overall system. We consider a scenario that consists of one of the generators suddenly losing power, two of the loads suddenly changing utility functions, one of the loads suddenly changing power consumption and the generator power recovering. This is realized as follows: at $t = 10$s (the system is stabilized at an equilibrium point at $t = 0$s), there is a step of magnitude -8 per unit (p.u.) (20 percent drop) in the capacity of generator 2; 20 seconds later, we make a sudden change in the load utility function at load bus 4 (from $-0.3P_{L_4}^2 + 70P_{L_4}$ to $-0.2P_{L_4}^2 + 70P_{L_4}$); after 20 seconds, there is a sudden change in the load utility function at load bus 5 (from $-0.4P_{L_5}^2 + 90P_{L_5}$ to $-0.35P_{L_5}^2 + 90P_{L_5}$); at $t = 70$s, we

Table 3.1: Parameters of the overall system (p.u. means per unit).

Parameter	Value(p.u.)	Parameter	Value(p.u.)
M_i	$9, 10, 11$	$P_{C_1}^{\max}, P_{C_1}^{\min}$	$60, 0$
D_i	$1.1, 1, 0.9, 0.4, 0.6, 0.5$	$P_{C_2}^{\max}, P_{C_2}^{\min}$	$40, 0$
T_{TG_i}	$5, 4, 6$	$P_{C_3}^{\max}, P_{C_3}^{\min}$	$50, 0$
R	$0.05I_3$	$P_{L_4}^{\max}, P_{L_4}^{\min}$	$40, 10$
T_k	$80, 100, 90, 120, 110$	$P_{L_5}^{\max}, P_{L_5}^{\min}$	$50, 0$
Function	Expression	Parameter	Value(p.u.)
$C_1(P_{C_1})$	$0.1P_{C_1}^2 + 25P_{C_1}$	$P_{L_6}^{\max}, P_{L_6}^{\min}$	$20, 15$
$C_2(P_{C_2})$	$0.3P_{C_2}^2 + 30P_{C_2}$	$P_{TC_1}^{\max}, P_{TC_1}^{\min}$	$40, -40$
$C_3(P_{C_3})$	$0.5P_{C_3}^2 + 35P_{C_3}$	$P_{TC_2}^{\max}, P_{TC_2}^{\min}$	$50, -50$
$U_4(P_{L_4})$	$-0.3P_{L_4}^2 + 70P_{L_4}$	$P_{TC_3}^{\max}, P_{TC_3}^{\min}$	$45, -45$
$U_5(P_{L_5})$	$-0.4P_{L_5}^2 + 90P_{L_5}$	$P_{TC_4}^{\max}, P_{TC_4}^{\min}$	$60, -60$
$U_6(P_{L_6})$	$-0.6P_{L_6}^2 + 100P_{L_6}$	$P_{TC_5}^{\max}, P_{TC_5}^{\min}$	$55, -55$
Gain	Value(p.u.)	Gain	Value(p.u.)
K_{P_C}	$3I_3$	K_{P_L}	$5I_3$
K_ζ	$4I_3$	K_λ	$6I_3$
K_{μ^+}, K_{μ^-}	$3.5I_3$	K_{ν^+}, K_{ν^-}	$4.5I_3$
K	$40I_5$	K_{l^+}, K_{l^-}	$5.5I_5$
K_{eP_C}	$3I_3$	$\hat{K}_{eP_C}$	I_3
K_{eP_L}	$3I_3$	$\hat{K}_{eP_L}$	I_3

make a sudden change in the upper bound of load power consumption at load bus 4 (from 40p.u. to 10p.u.): all of these changes in practice represent variations on the demand side; finally at $t = 90$s there is a step increase of magnitude +4p.u. (12.5 percent increase) in the capacity of generator 2 and this indicates the generator power recovery. We test the controller under this scenario without and with the exogenous input d_l at load buses ($d_g = \mathbf{0}$), representing uncontrollable active power injections from, e.g., uncontrollable loads/users and distributed energy resources. Note that here the assumption of constant d_l is relaxed, and instead, we consider time-varying exogenous input. The simulation results are shown in Figures 3.6 to 3.7. It is clear that the controller stabilizes the frequency and balances the supply and demand quickly during the whole scenario (the nominal frequency is 50Hz).

In addition, we introduce extra dynamics and test the system under the same scenario without and with the exogenous input as given on the bottom of Figure 3.7. The simulation results are shown in Figures 3.8 to 3.9. It is clear that with the extra dynamics, the controller behaves much better in that the transient time becomes

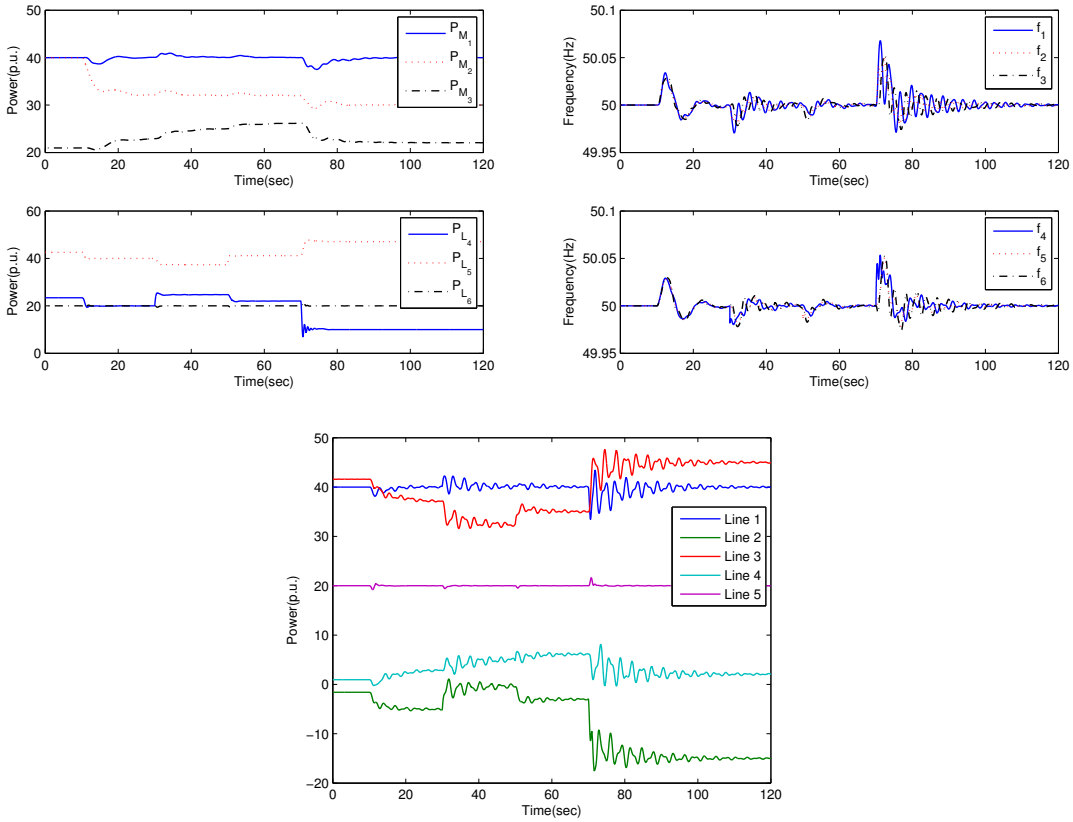

Figure 3.6: System response without the exogenous input. Top left: power generation and consumption. Top right: bus frequency. Bottom: power flow in transmission lines.

shorter and there are much less oscillations, since the extra dynamics hinder the power generated and consumed from changing too quickly.

3.5 Conclusion

In this chapter, we have studied a real-time control framework that merges conventional primary, secondary and tertiary frequency control in power systems. We considered a transmission level network with tree topology, and formulated an OPF problem with constraints containing the dynamics of the power network. We then used a primal-dual approach to design a distributed dynamic feedback controller for the system. The design process and stability proof were presented in Section 3.2. Moreover, we introduced the extra dynamics (3.26) to improve its behaviour, where we emphasized the trade-off when choosing the gains of the extra dynamics. Numerical experiments showed that the proposed controller could balance the power flow in

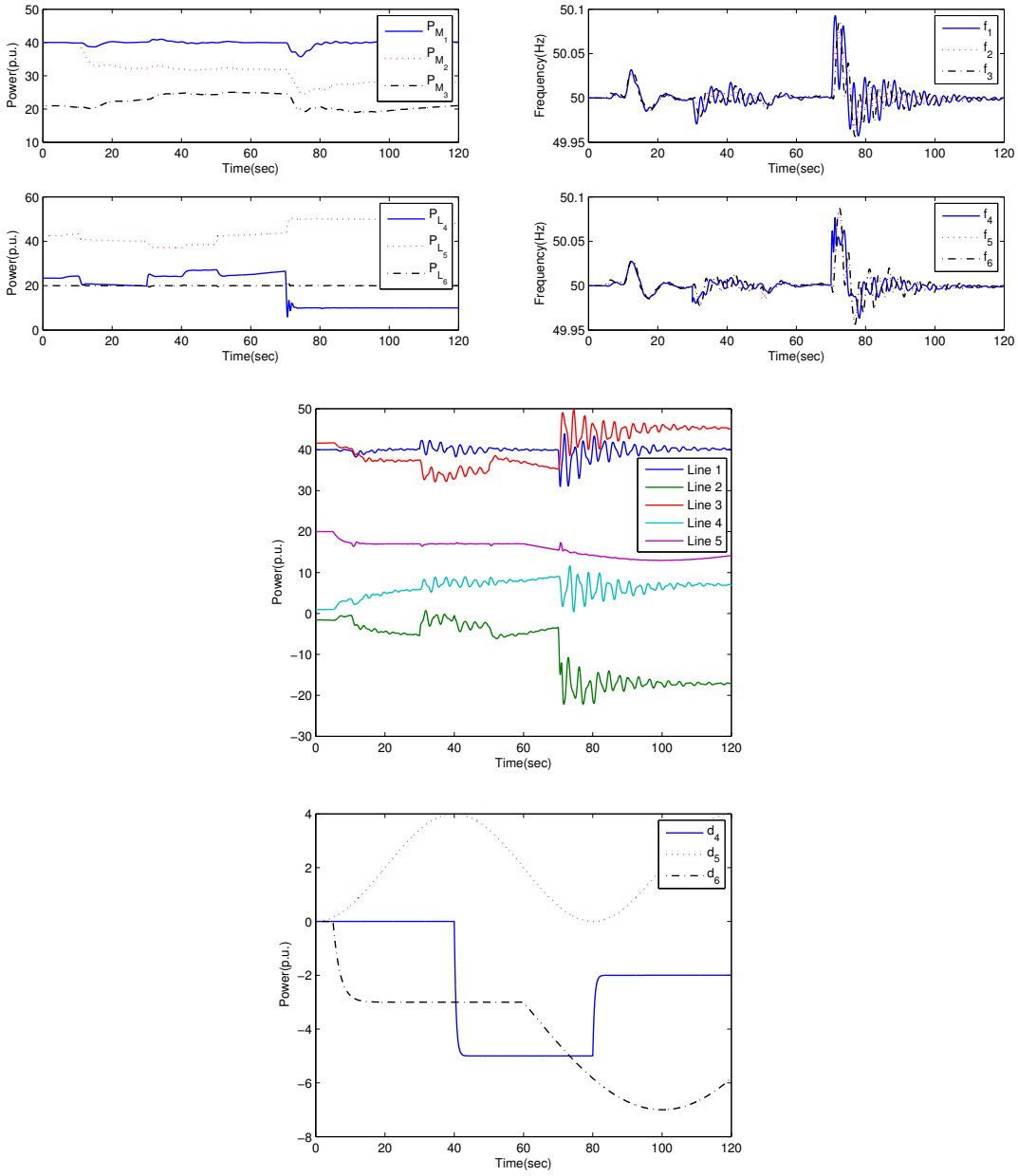

Figure 3.7: System response with the exogenous input. Top left: power generation and consumption. Top right: bus frequency. Middle: power flow in transmission lines. Bottom: exogenous input profile.

the network quickly, and achieve OPF in the steady state. Furthermore, after adding the extra dynamics, the transient performance of the system improved significantly.

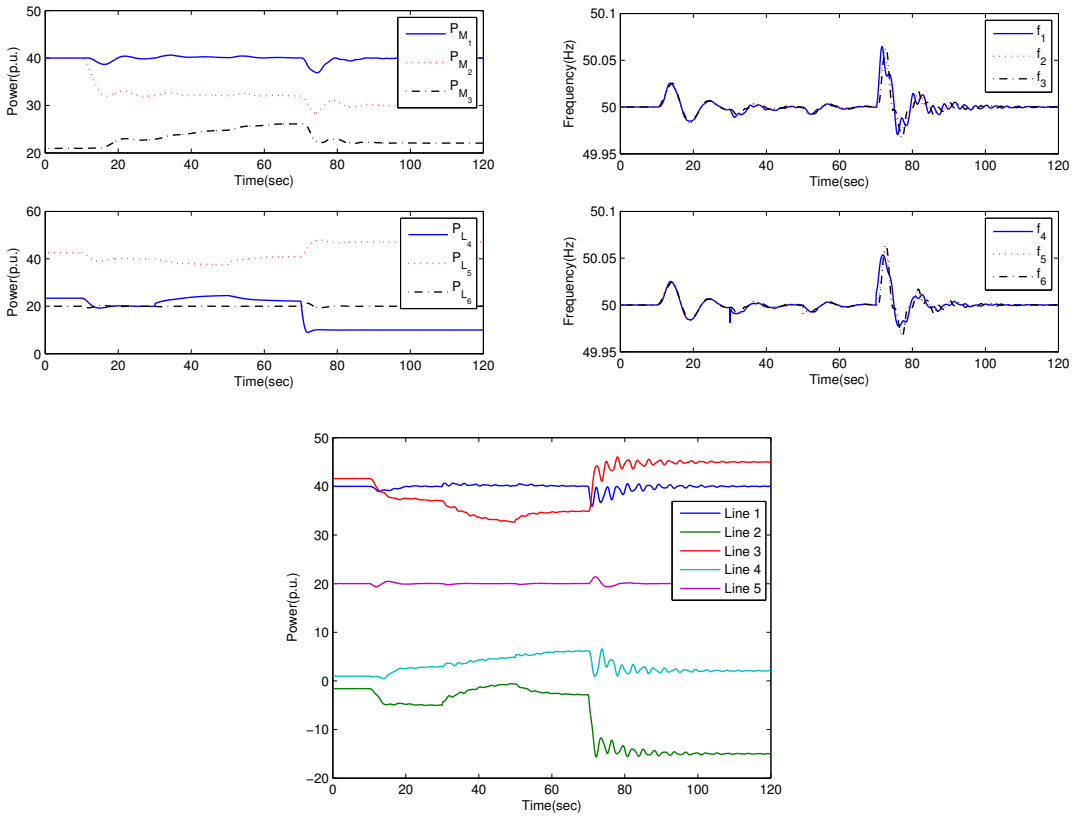

Figure 3.8: System response without the exogenous input, after adding the extra dynamics. Top left: power generation and consumption. Top right: bus frequency. Bottom: power flow in transmission lines.

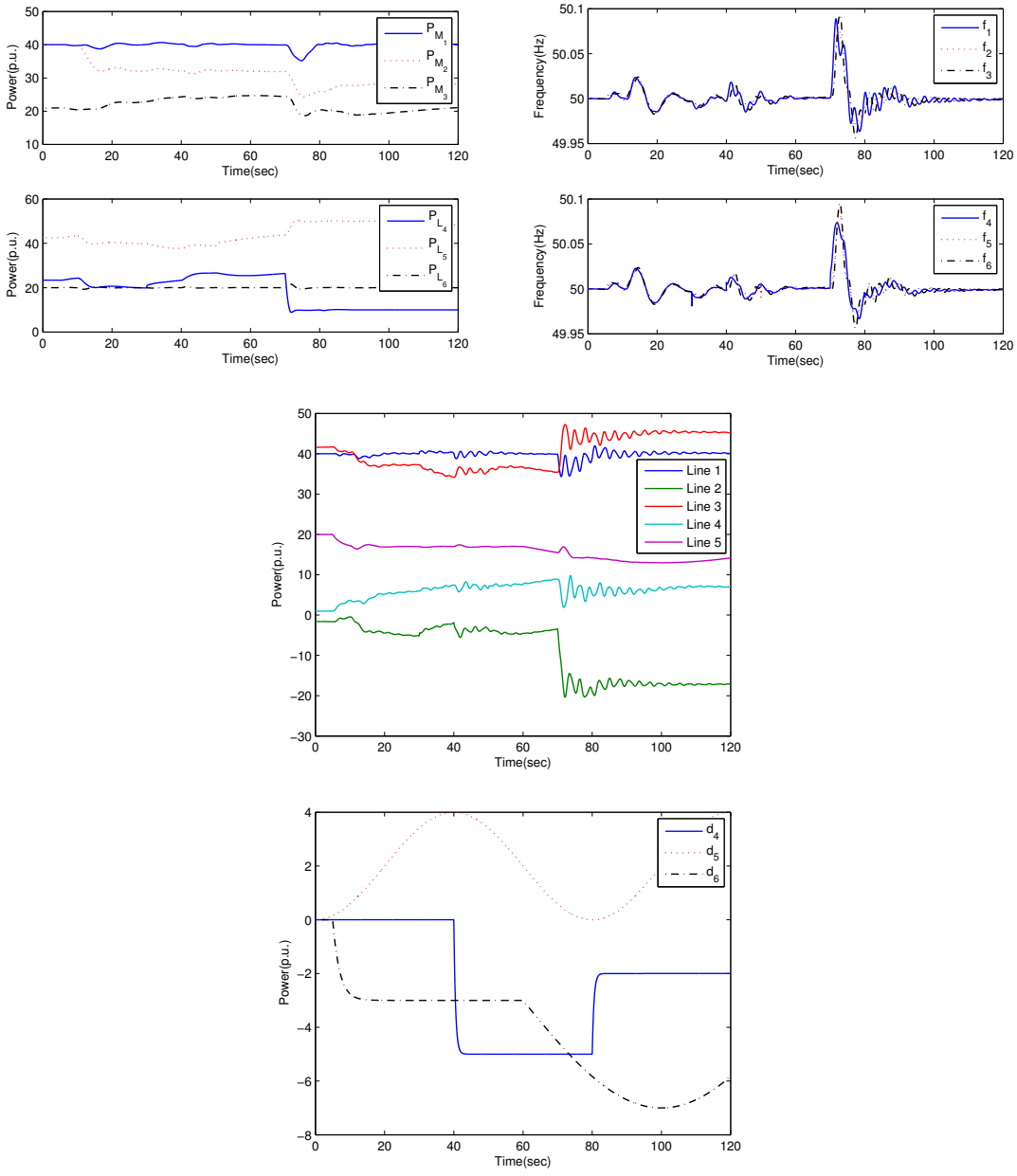

Figure 3.9: System response with the exogenous input, after adding the extra dynamics. Top left: power generation and consumption. Top right: bus frequency. Middle: power flow in transmission lines. Bottom: exogenous input profile.

Chapter 4

Improving the Performance of Network Congestion Control Algorithms

In the previous chapter, we have proposed a real-time distributed control framework to merge conventional primary, secondary and tertiary frequency control for power systems, and focused on networks with tree topology to demonstrate how the design methodology, stability analysis and robustness improvement can be performed. As shown in the numerical examples, introducing extra dynamics improved the transient performance of the feedback significantly. We now apply this redesign approach to the network congestion control problem, which is a typical problem in the study of networks. This problem has received increasing attention in the past three decades, especially following the work in [48]. Many control schemes have been designed and applied to solve it [49, 50, 51], nevertheless, there are still many open research issues [51, 52], e.g., global stability of more accurate models, trade-offs among system performance, robustness and complexity of controllers, etc.

In this chapter, motivated by the augmented Lagrangian method, we propose a method for redesigning existing congestion control algorithms at the level of fluid-flow models [64], in order to improve the transient behaviour, and provide robustness to uncertainties in the network structure and communication constraints. As pointed out in [90, 53, 51, 52], improving the performance and robustness of existing congestion control schemes is of practical importance.

This chapter is organized as follows. Firstly, we revisit the network congestion control problem and discuss the general redesign framework for the related control algorithms. We then apply the redesign framework to modify the primal-dual, primal and dual congestion control algorithms respectively by introducing quadratic penalty

terms. We mainly focus on scalable analysis relating to stability, linear robustness and delay robustness. In addition, we study the meaning of the extra dynamics and further introduce distributed Proportional-Integral-Derivative (PID) control actions to network congestion control algorithms. Finally, we present three illustrative examples, including an application to a jointly optimal congestion and contention control problem.

4.1 The Redesign Framework

4.1.1 Preliminaries

The standard network congestion control problem with N users using L links takes the form [55]

$$\max_{x_i \geq 0} \quad \sum_{i=1}^{N} U_i(x_i) \tag{4.1a}$$

$$\text{subject to} \quad \sum_{i=1}^{N} R_{li} x_i \leq c_l, \ l = 1, \cdots, L. \tag{4.1b}$$

Here $U_i(x_i)$ is the utility function of user i, which is assumed to be a continuously differentiable, monotonically increasing, strictly concave function of the transmission rate x_i. R is a routing matrix describing the interconnection, given by

$$R_{li} = \begin{cases} 1 & \text{if user } i \text{ uses link } l, \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

The constraints indicate that the aggregate rate of all users using link l is limited by c_l where $c_l > 0$. The Lagrangian is defined as

$$L(x_i, p_l) = \sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l \left(c_l - \sum_{i=1}^{N} R_{li} x_i \right) \tag{4.3}$$

where $p_l \geq 0$ is the Lagrange multiplier for the inequality constraint $\sum_{i=1}^{N} R_{li} x_i \leq c_l$. Thus, the dual optimization problem is given by

$$\min_{p_l} \sup_{x_i \geq 0} \quad \sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l \left(c_l - \sum_{i=1}^{N} R_{li} x_i \right) \tag{4.4a}$$

$$\text{subject to} \ \ p_l \geq 0, \ l = 1, \cdots, L. \tag{4.4b}$$

Let x_i^* and p_l^* denote the optimal solutions of the primal problem (4.1) and the dual problem (4.4) respectively. Since the primal problem has affine constraints and is

strictly convex, strong duality holds. x_i^* is unique because the primal problem is strictly convex and feasible. Moreover, x_i^* and p_l^* satisfy the Karush-Kuhn-Tucker (KKT) conditions given by

$$\frac{\partial L}{\partial x_i} = 0,\ i = 1, \ldots, N \tag{4.5a}$$

$$p_l \frac{\partial L}{\partial p_l} = 0, \quad p_l \geq 0, \quad \frac{\partial L}{\partial p_l} \geq 0,\ l = 1, \ldots, L. \tag{4.5b}$$

Generally speaking, p_l^* may not be unique because the capacity constraints may not be independent [91]. To simplify the problem and obtain a unique optimal solution to (4.4) (so that (x_i^*, p_l^*) is the unique saddle point of (4.3)), we make the following assumption [91, 50, 53, 51, 55]:

Assumption 4.1 *The routing matrix R has full row rank.*

For the case where Assumption 4.1 does not hold, the results in this chapter can be generalized in terms of equilibrium set/saddle point set.

For solving the centralized, large-scale, yet convex problem (4.1), primal-dual, primal and dual algorithms have been proposed in the literature [50], which are given as follows:

The primal-dual algorithm:

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - q_i) \tag{4.6a}$$

$$\dot{p}_l = k_{p_l}(y_l - c_l)_{p_l}^+ \tag{4.6b}$$

The primal algorithm:

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - q_i) \tag{4.7a}$$

$$p_l = f_l(y_l) \tag{4.7b}$$

The dual algorithm:

$$\dot{p}_l = k_{p_l}(y_l - c_l)_{p_l}^+ \tag{4.8a}$$

$$x_i = U_i'^{-1}(q_i) \tag{4.8b}$$

where $q_i = \sum_{l=1}^{L} R_{li} p_l$, $y_l = \sum_{i=1}^{N} R_{li} x_i$, $k_{x_i}, k_{p_l} > 0$ (note that when implementing these algorithms in a packet-level protocol, there are extra constraints for the gains to ensure convergence/stability in terms of stepsizes, as shown in [92]), and each $f_l(\beta)$ is a barrier/penalty function satisfying $f_l(\beta), f_l'(\beta) > 0$ [50, 55]. The characteristics

of these algorithms vary in efficiency, robustness, computational cost, communication, etc. Which algorithm should be chosen depends on the criteria set by users and operators in the network [51]. Figure 4.1 illustrates the feedback structure of these algorithms, where both users and links use the aggregate information available to them to find the optimal transmission rate x_i^* and price signal p_l^* in a distributed manner, which maximizes the total utility $\sum_{i=1}^{N} U_i(x_i)$. However, these algorithms only consider convergence to the equilibrium (the optimal solution to (4.1)), without paying attention to the transient behaviour of the overall system. This could result in, e.g., large overshoots, longer transient times and reduced robustness to communication constraints [51, 52], for the closed-loop system. In order to achieve better transient performance and added robustness, we next introduce our redesign framework.

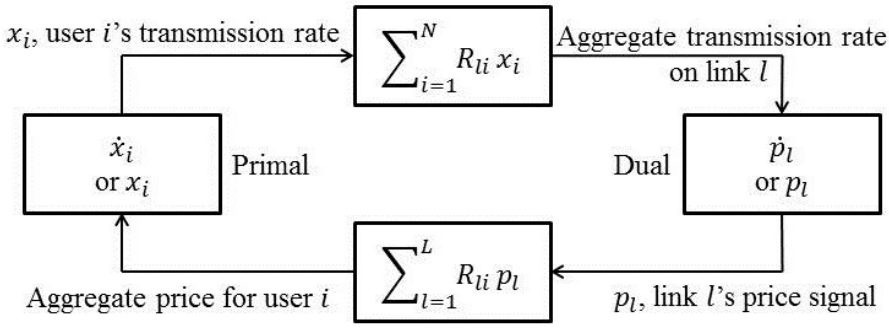

Figure 4.1: Feedback control structure of available congestion control problems.

4.1.2 The Redesign Framework

Motivated by the need for *scalability*, our redesign framework is based on an augmented Lagrangian method in that we add extra terms $h(z, \hat{z}, k)$ to (4.3). The Lagrangian then becomes

$$L_e(x_i, p_l, h(z, \hat{z}, k)) = \sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l \left(c_l - \sum_{i=1}^{N} R_{li} x_i \right) + h(z, \hat{z}, k) \tag{4.9}$$

where z stands for a group of existing states (decision variables, i.e., x_i, p_l), $\hat{z}$ stands for a group of unconstrained extra states, and k stands for a group of parameters/coefficients in the extra terms which can either be constant or time-varying. The structure of $h(z, \hat{z}, k)$, however, needs to satisfy certain requirements. The optimal solution should not change with $h(z, \hat{z}, k)$, i.e., if the unique saddle point of (4.9) is $(x_i^*, p_l^*, z^*, \hat{z}^*)$, then (x_i^*, p_l^*) should be the saddle point of (4.3). Moreover, to maintain the distributed structure, $h(z, \hat{z}, k)$ needs to be separable so that each

state can still be updated using local information. The idea behind this redesign is to increase the dimension of the problem by adding extra states – the control scheme, if designed properly, can result in improved performance: in fact the structure of $h(z, \hat{z}, k)$ depends on the problem considered. One approach is to introduce simple quadratic penalty terms, since there is little extra computational cost [93]. In this chapter, we consider $h(z, \hat{z}, k) = h(x_i, \hat{x}_i, k_{e_i}) = -\sum_{i=1}^{N} \frac{k_{e_i}}{2}(x_i - \hat{x}_i)^2$ and $h(z, \hat{z}, k) = h(p_l, \hat{p}_l, k_{e_l}) = \sum_{l=1}^{L} \frac{k_{e_l}}{2}(p_l - \hat{p}_l)^2$, and choose constant parameters. With the extra terms, we reformulate the Lagrangian as

$$L_e(x_i, p_l, \hat{x}_i) = \sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l \left(c_l - \sum_{i=1}^{N} R_{li} x_i \right) - \sum_{i=1}^{N} \frac{k_{e_i}}{2}(x_i - \hat{x}_i)^2 \tag{4.10a}$$

$$L_e(x_i, p_l, \hat{x}_i) = \sum_{i=1}^{N} U_i(x_i) - \sum_{l=1}^{L} \int_0^{y_l} f_l(\beta) d\beta - \sum_{i=1}^{N} \frac{k_{e_i}}{2}(x_i - \hat{x}_i)^2 \tag{4.10b}$$

$$L_e(x_i, p_l, \hat{p}_l) = \sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l \left(c_l - \sum_{i=1}^{N} R_{li} x_i \right) + \sum_{l=1}^{L} \frac{k_{e_l}}{2}(p_l - \hat{p}_l)^2. \tag{4.10c}$$

Equation (4.10a) will be used to redesign the primal-dual algorithm where $\{\hat{x}_1, \cdots, \hat{x}_N\}$ is a set of extra decision variables and $k_{e_i} > 0$. Equation (4.10b) will be used to redesign the primal algorithm where we have used the modified version of problem (4.1) (i.e., change (4.1) to an unconstrained optimization problem by introducing the penalty function $-\sum_{l=1}^{L} \int_0^{y_l} f_l(\beta) d\beta$) [50, 54], $\{\hat{x}_1, \cdots, \hat{x}_N\}$ is a set of extra decision variables and $k_{e_i} > 0$. Equation (4.10c) will be used to redesign the dual algorithm where $\{\hat{p}_1, \cdots, \hat{p}_L\}$ is a set of extra decision variables and $k_{e_l} > 0$. After introducing these extra states, the modified feedback control structure is shown in Figure 4.2 (with dashed blocks).

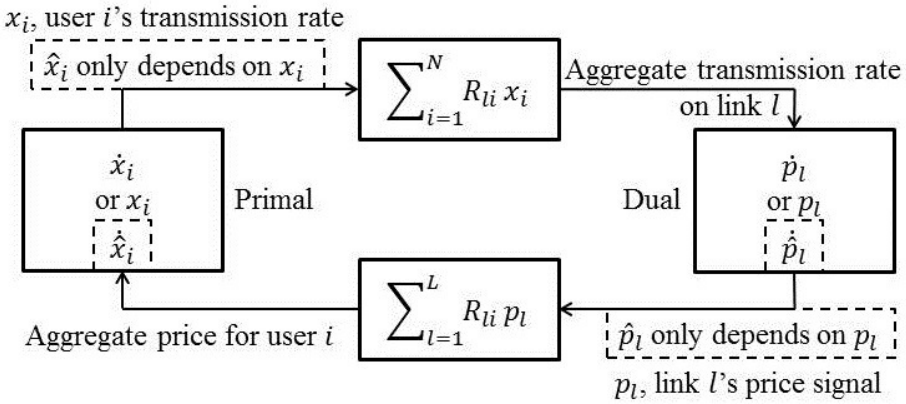

Figure 4.2: Modified feedback control structure of available congestion control problems.

Remark 4.1 *Adding the terms* $-\sum_{i=1}^{N} \frac{k_{e_i}}{2}(x_i - \hat{x}_i)^2$ *to (4.3) is equivalent to reformulating (4.1) as:*

$$\max_{x_i \geq 0, \hat{x}_i} \quad \sum_{i=1}^{N} U_i(x_i) - \sum_{i=1}^{N} \frac{k_{e_i}}{2}(x_i - \hat{x}_i)^2$$
$$\textit{subject to} \quad \sum_{i=1}^{N} R_{li} x_i \leq c_l, \ l = 1, \cdots, L.$$

The advantage of describing the redesign framework in terms of the augmented Lagrangian rather than reformulating the original problem is that both primal and dual variables can be merged in one formulation, as shown in (4.10c).

4.2 Redesign for the Primal-dual Algorithm

We now use (4.10a) to redesign the primal-dual algorithm, investigating how the gains resulting from the extra dynamics influence the stability and robustness of the closed-loop system. Moreover, we consider communication constraints and show that the modified system could achieve added delay robustness. In this section we will redesign the control algorithms at the level of fluid-flow models, which allows the use of continuous-time models for analysis and design. For implementing the redesigned algorithm in a practical packet-level context, we refer to [94] (Section IV) and [64] (Section V) for more details.

4.2.1 Modified Dynamics and Stability

Based on a primal-dual approach [50] (Section 3.4), we obtain the modified dynamics (in continuous-time) given by

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - q_i + k_{e_i}(\hat{x}_i - x_i)) \tag{4.11a}$$
$$\dot{\hat{x}}_i = \hat{k}_{e_i}(x_i - \hat{x}_i) \tag{4.11b}$$
$$\dot{p}_l = k_{p_l}(y_l - c_l)^+_{p_l} \tag{4.11c}$$

where $\hat{k}_{e_i} > 0$. Although the extra dynamics $\dot{\hat{x}}_i$ are introduced, the feedback control structure remains distributed.

Theorem 4.1 *Under Assumption 4.1, the equilibrium point of (4.11) is unique, satisfies the KKT conditions (4.5) and is globally asymptotically stable for all non-negative initial conditions.*

Proof: Noting that at the equilibrium point x_i and $\hat{x}_i$ are the same, the proof of uniqueness and optimality of the equilibrium point can be found in [50] (Section 3). Consider the following radially unbounded function as a candidate Lyapunov function

$$V_{(4.11)}(x_i, \hat{x}_i, p_l) = \sum_{i=1}^{N} \frac{1}{2k_{x_i}}(x_i - x_i^*)^2 + \sum_{l=1}^{L} \frac{1}{2k_{p_l}}(p_l - p_l^*)^2 + \sum_{i=1}^{N} \frac{k_{e_i}}{2\hat{k}_{e_i}}(\hat{x}_i - \hat{x}_i^*)^2 \tag{4.12}$$

where $x_i^* = \hat{x}_i^*$ holds. Differentiating $V_{(4.11)}$ with respect to time, we get

$$\begin{aligned}
\dot{V}_{(4.11)} =& \sum_{i=1}^{N}(U_i'(x_i) - q_i + k_{e_i}(\hat{x}_i - x_i))(x_i - x_i^*) + \sum_{l=1}^{L}(y_l - c_l)_{p_l}^{+}(p_l - p_l^*) \\
&+ \sum_{i=1}^{N} k_{e_i}(x_i - \hat{x}_i)(\hat{x}_i - \hat{x}_i^*) \\
\leq& \sum_{i=1}^{N}(U_i'(x_i) - q_i)(x_i - x_i^*) + \sum_{l=1}^{L}(y_l - c_l)(p_l - p_l^*) - \sum_{i=1}^{N} k_{e_i}(x_i - \hat{x}_i)^2 \\
\leq& \sum_{i=1}^{N}(U_i'(x_i) - U_i'(x_i^*))(x_i - x_i^*) + \sum_{l=1}^{L}(y_l^* - c_l)(p_l - p_l^*) - \sum_{i=1}^{N} k_{e_i}(x_i - \hat{x}_i)^2 \\
\leq& 0.
\end{aligned} \tag{4.13}$$

Now $\dot{V}_{(4.11)} = 0$ only when $x_i = \hat{x}_i = x_i^*$, and either $y_l^* = c_l$ or $p_l = p_l^*$, which can only happen at the equilibrium of interest. Using Krasovskii-LaSalle principle [79], the equilibrium point is globally asymptotically stable. ■

Theorem 4.1 indicates that the extra dynamics do not affect the global asymptotic stability of the overall system. Moreover, there is no restriction for the choice of k_{e_i} and $\hat{k}_{e_i}$, except positivity. We next investigate how to tune them in order to achieve better performance for the closed-loop system.

4.2.2 Linear Robust Analysis

Linearizing (4.11) around the equilibrium point and expressing it in the Laplace domain, we obtain

$$\delta x_i(s) = \frac{k_{x_i}}{s}\left(-\kappa_i \delta x_i(s) - \sum_{m=1}^{\bar{L}} \bar{R}_{mi}\delta\bar{p}_m(s) - \frac{k_{e_i}s}{s + \hat{k}_{e_i}}\delta x_i(s)\right) \tag{4.14a}$$

$$\delta\bar{p}_l(s) = \frac{k_{p_l}}{s}\sum_{n=1}^{N} \bar{R}_{ln}\delta x_n(s). \tag{4.14b}$$

In these transfer functions, $\delta x_i(s)$ denotes the linearized state and $\delta \bar{p}_l(s)$ denotes the linearized reduced state by eliminating non-bottleneck links (whose prices at the equilibrium point are zeros [64]). Also, $\bar{R}$ is obtained by eliminating non-bottleneck rows from R, $\bar{R}_{mi}$ and $\bar{R}_{ln}$ are the entries of $\bar{R}$ and $\bar{L}$ is the number of bottleneck links. Let $\kappa_i = -U_i''(x_i^*)$. Following Assumption 4.1, we will assume that $\bar{R}$ is of full row rank. We rearrange the transfer function as a standard unity negative feedback system so that the loop transfer function matrix is given by

$$L_{(4.14)}(s) = \text{diag}\left\{\frac{k_{x_i}(s+\hat{k}_{e_i})}{s(s^2+(k_{x_i}\kappa_i + k_{x_i}k_{e_i}+\hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i})}\right\}\bar{R}^T\text{diag}\{k_{p_l}\}\bar{R}. \quad (4.15)$$

Based on classical loop shaping methods (see the Appendix), in order to achieve better performance, $\underline{\sigma}(L_{(4.14)}(s))$ needs to be large at low frequencies and $\overline{\sigma}(L_{(4.14)}(s))$ needs to be small at high frequencies [89]. Since

$$\begin{aligned}
&\underline{\sigma}\left(\text{diag}\left\{\frac{k_{x_i}(s+\hat{k}_{e_i})}{s(s^2+(k_{x_i}\kappa_i + k_{x_i}k_{e_i}+\hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i})}\right\}\right)\underline{\sigma}\left(\bar{R}^T\text{diag}\{k_{p_l}\}\bar{R}\right)\\
\leq&\underline{\sigma}(L_{(4.14)}(s))\\
&\overline{\sigma}\left(\text{diag}\left\{\frac{k_{x_i}(s+\hat{k}_{e_i})}{s(s^2+(k_{x_i}\kappa_i + k_{x_i}k_{e_i}+\hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i})}\right\}\right)\overline{\sigma}\left(\bar{R}^T\text{diag}\{k_{p_l}\}\bar{R}\right)\\
\geq&\overline{\sigma}(L_{(4.14)}(s))
\end{aligned}$$

a sufficient condition is to make the lower bound of $\underline{\sigma}(L_{(4.14)}(s))$ large at low frequencies and the upper bound of $\overline{\sigma}(L_{(4.14)}(s))$ small at high frequencies. We then have the following remark.

Remark 4.2 *Let*

$$K_i(j\omega) = \frac{k_{x_i}(j\omega+\hat{k}_{e_i})}{j\omega(k_{x_i}\kappa_i\hat{k}_{e_i} - \omega^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i}+\hat{k}_{e_i})j\omega)}.$$

Since

$$\frac{\partial|K_i|^2}{\partial k_{e_i}} < 0$$

$$\frac{\partial|K_i|^2}{\partial \hat{k}_{e_i}} = \frac{2k_{x_i}^3k_{e_i}(\hat{k}_{e_i}(\hat{k}_{e_i}+2k_{x_i}\kappa_i+k_{x_i}k_{e_i})-\omega^2)}{((\omega^2-k_{x_i}\kappa_i\hat{k}_{e_i})^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})^2\omega^2)^2}$$

small k_{e_i} increases the lower bound of $\underline{\sigma}(L_{(4.14)}(s))$ at low frequencies while also increasing the upper bound of $\overline{\sigma}(L_{(4.14)}(s))$ at high frequencies. Large $\hat{k}_{e_i}$ increases the lower bound of $\underline{\sigma}(L_{(4.14)}(s))$ at low frequencies and decreases the upper bound of $\overline{\sigma}(L_{(4.14)}(s))$ at high frequencies.

This means that in order to improve the performance of the closed-loop system, there is a trade-off for k_{e_i} (not very large or very small) and $\hat{k}_{e_i}$ needs to be large.

4.2.3 Robust Stability to Delays

We now consider delay robustness of the overall system. In the delayed case, the closed-loop dynamics (4.11) change to

$$\dot{x}_i(t) = k_{x_i}\left(U_i'(x_i(t)) - \sum_{m=1}^{L} R_{mi}p_m(t-\tau_{i,m}^b) + k_{e_i}(\hat{x}_i(t) - x_i(t))\right) \tag{4.16a}$$

$$\dot{\hat{x}}_i(t) = \hat{k}_{e_i}(x_i(t) - \hat{x}_i(t)) \tag{4.16b}$$

$$\dot{p}_l(t) = k_{p_l}\left(\sum_{n=1}^{N} R_{ln}x_n(t-\tau_{n,l}^f) - c_l\right)_{p_l}^{+} \tag{4.16c}$$

where $\tau_{n,l}^f$ is the forward time delay of the transmission rate x_n at link l, and $\tau_{i,m}^b$ is the backward time delay of the price p_m for the user i. The forward and backward delays can be combined to yield the Round Trip Time (RTT) $\tau_i = \tau_{i,l}^f + \tau_{i,l}^b,\ \forall\ l = 1, \cdots, L$, for user i.

We linearize (4.16) around the equilibrium and express it in the Laplace domain

$$\delta x_i(s) = \frac{k_{x_i}}{s}\left(-\kappa_i\delta x_i(s) - \sum_{m=1}^{\bar{L}} \bar{R}_{mi}^b(s)\delta\bar{p}_m(s) - \frac{k_{e_i}s}{s+\hat{k}_{e_i}}\delta x_i(s)\right) \tag{4.17a}$$

$$\delta\bar{p}_l(s) = \frac{k_{p_l}}{s}\sum_{n=1}^{N} \bar{R}_{ln}^f(s)\delta x_n(s). \tag{4.17b}$$

In these transfer functions, $\delta x_i(s)$, $\delta\bar{p}_l(s)$, $\bar{L}$ and κ_i are defined as stated previously: $\bar{R}^f(s)$ and $\bar{R}^b(s)$ are obtained by eliminating non-bottleneck rows from R, and also replacing the "1" elements by the delay terms $e^{-\tau_{n,l}^f s}$ and $e^{-\tau_{i,m}^b s}$ respectively. $\bar{R}_{ln}^f(s)$ and $\bar{R}_{mi}^b(s)$ are the entries of $\bar{R}^f(s)$ and $\bar{R}^b(s)$ respectively. As before, we assume that $\bar{R}^f(0)$ and $\bar{R}^b(0)$ have full row rank. The loop transfer function matrix is given by

$$L_{(4.17)}(s) = \text{diag}\left\{\frac{k_{x_i}(s+\hat{k}_{e_i})}{s(s^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i})}\right\}\bar{R}^b(s)^T\text{diag}\{k_{p_l}\}\bar{R}^f(s).$$

Noting that $\bar{R}^b(s) = \bar{R}^f(-s)\text{diag}\,\{e^{-\tau_i s}\}$, we have

$$L_{(4.17)}(s) = \text{diag}\left\{\frac{k_{x_i}e^{-\tau_i s}(s+\hat{k}_{e_i})}{s(s^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i})}\right\}\bar{R}^f(-s)^T\text{diag}\{k_{p_l}\}\bar{R}^f(s). \tag{4.18}$$

Remark 4.3 *When $k_{e_i} = 0$, i.e., there is no extra dynamics added to the system, (4.18) becomes*

$$diag\left\{\frac{k_{x_i}e^{-\tau_i s}}{s(s+k_{x_i}\kappa_i)}\right\}\bar{R}^f(-s)^T diag\{k_{p_l}\}\bar{R}^f(s).$$

Comparing this transfer function matrix to (4.18), we can see that the redesigned dynamics add a band rejection filter $\frac{(s+k_{x_i}\kappa_i)(s+\hat{k}_{e_i})}{s^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})s+k_{x_i}\kappa_i\hat{k}_{e_i}}$ *to each* $\frac{k_{x_i}e^{-\tau_i s}}{s(s+k_{x_i}\kappa_i)}$ *in the modified system.*

Now we turn to developing stability conditions for (4.16). We first present a related lemma.

Lemma 4.1 *Given* $Q(j\omega)=\bar{R}^f(j\omega)^H diag\{k_{p_l}\}\bar{R}^f(j\omega)$*, where* $\omega\neq 0$ *and all parameters are defined in* $L_{(4.17)}(s)$*, the following inequality holds:*

$$\rho(Q(j\omega))\leq \max_i \sum_{l=1}^{\bar{L}}\sum_{n=1}^{N}\bar{R}_{li}\bar{R}_{ln}k_{p_l}.$$

Proof: The proof follows from the claim in the Appendix of [64]. Using the l_∞-induced norm, we have

$$\begin{aligned}\rho(Q(j\omega)) &= \rho\left(\bar{R}^f(j\omega)^H \mathrm{diag}\{k_{p_l}\}\bar{R}^f(j\omega)\right)\\ &\leq \left\|\bar{R}^f(j\omega)^H \mathrm{diag}\{k_{p_l}\}\bar{R}^f(j\omega)\right\|_{\infty-\mathrm{ind}}\\ &= \max_i \sum_{l=1}^{\bar{L}}\sum_{n=1}^{N}\bar{R}_{li}\bar{R}_{ln}k_{p_l}.\end{aligned}$$

■

Under Remark 4.3, Lemma 4.1 and following the proof of Theorem 1 in [64], we can derive a local stability condition for the system without extra dynamics, given by

$$\sum_{l=1}^{\bar{L}}\sum_{n=1}^{N}\bar{R}_{li}\bar{R}_{ln}k_{p_l} < \min_n \frac{\omega_n\sqrt{\omega_n^2+k_{x_n}^2\kappa_n^2}}{k_{x_n}},\ i=1,\cdots,N \tag{4.19}$$

where $\omega_n>0$ is the solution to $\tau_n\omega_n=\arctan\frac{k_{x_n}\kappa_n}{\omega_n}, n=1,\cdots,N$. With the extra dynamics, this stability condition can be much less conservative, as demonstrated in the following theorem.

Theorem 4.2 *For fixed full rank* $\bar{R}$ *(*$\bar{R}^f(0)=\bar{R}^b(0)=\bar{R}$*), the equilibrium point of (4.16) is unique, satisfies the KKT conditions (4.5) and is locally asymptotically stable provided*

$$\sum_{l=1}^{\bar{L}}\sum_{n=1}^{N}\bar{R}_{li}\bar{R}_{ln}k_{p_l} < \min_n\left(\frac{\omega_n\sqrt{\omega_n^2+k_{x_n}^2\kappa_n^2}}{k_{x_n}}\left(1+\frac{k_{x_n}^2\kappa_n k_{e_n}}{\omega_n^2+k_{x_n}^2\kappa_n^2}\right)\right), i=1,\cdots,N \tag{4.20}$$

where $\omega_n > 0$ *is the solution to* $\tau_n\omega_n = \arctan\frac{k_{x_n}\kappa_n}{\omega_n}$ *and* $\hat{k}_{e_n}$ *satisfies* $k_{x_n}\kappa_n\hat{k}_{e_n} = \omega_n^2, n = 1, \cdots, N$.

Proof: Noting that at the equilibria x_i and $\hat{x}_i$ are the same, the uniqueness and optimality of the equilibrium point follow from Theorem 4.1. For asymptotic stability, according to Proposition 2 in [64], it is sufficient to show that

$$-1 \notin \text{eig}(\mu L_{(4.17)}(j\omega)),\ \mu \in (0, 1],\ \omega \neq 0 \Leftrightarrow -1 \notin \text{eig}(Q(j\omega)\Lambda(j\omega))$$

where $Q(j\omega) = Q(j\omega)^H$ is positive semi-definite as defined in Lemma 4.1, and

$$\Lambda(j\omega) = \text{diag}\left\{\frac{k_{x_i}e^{-j\tau_i\omega}(j\omega + \hat{k}_{e_i})}{j\omega(k_{x_i}\kappa_i\hat{k}_{e_i} - \omega^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})j\omega)}\right\}.$$

Since the Nyquist plots of the entries of $\Lambda(j\omega)$ first cross the negative real axis at the frequency satisfying $\tau_i\omega = \arctan\frac{k_{x_i}\kappa_i}{\omega}$ when $k_{x_i}\kappa_i\hat{k}_{e_i} = \omega^2$ holds, based on Lemma 1 in [72], it is sufficient to show that

$$\rho(Q(j\omega))\max_n \frac{k_{x_n}\sqrt{\omega_n^2 + \hat{k}_{e_n}^2}}{\omega_n^2(k_{x_n}\kappa_n + k_{x_n}k_{e_n} + \hat{k}_{e_n})} < 1$$

where $\omega_n > 0$ is the solution to $\tau_n\omega_n = \arctan\frac{k_{x_n}\kappa_n}{\omega_n}, n = 1, \cdots, N$. According to Lemma 4.1, this is satisfied when (4.20) holds, which completes the proof. ■

The condition in Theorem 4.2 indicates that by increasing k_{e_i} and choosing $\hat{k}_{e_i}$ to satisfy $k_{x_i}\kappa_i\hat{k}_{e_i} = \omega^2$ where $\omega > 0$ solves $\tau_i\omega = \arctan\frac{k_{x_i}\kappa_i}{\omega}$, which is distributed, the overall system can achieve added delay robustness. Combining Remark 4.2 and Theorem 4.2, we can see that there is a trade-off between these gains, as expected.

Remark 4.4 *Another way to redesign the primal-dual algorithm is to use the augmented Lagrangian*

$$L_e(x_i, p_l, \hat{x}_i, \hat{p}_l) = \sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l\left(c_l - \sum_{i=1}^{N} R_{li}x_i\right) - \sum_{i=1}^{N}\frac{k_{ex_i}}{2}(x_i - \hat{x}_i)^2 + \sum_{l=1}^{L}\frac{k_{ep_l}}{2}(p_l - \hat{p}_l)^2$$

which introduces both extra dynamics $\dot{\hat{x}}_i$ *and* $\dot{\hat{p}}_l$*, and can result in more benefits. The analysis of linear robustness and delay robustness in this case is similar to that in this section but more complicated.*

4.3 Redesign for the Primal Algorithm

We now use (4.10b) to redesign the primal algorithm, investigating how the gains resulting from the extra dynamics influence the stability and robustness of the closed-loop system. Since the redesign process is similar to that in the previous section, some details are omitted for brevity.

4.3.1 Modified Dynamics and Stability

Based on a primal approach [50] (Section 3.1), we obtain the modified dynamics (in continuous-time) given by

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - q_i + k_{e_i}(\hat{x}_i - x_i)) \tag{4.21a}$$

$$\dot{\hat{x}}_i = \hat{k}_{e_i}(x_i - \hat{x}_i) \tag{4.21b}$$

$$p_l = f_l(y_l) \tag{4.21c}$$

where $\hat{k}_{e_i} > 0$. Although the extra dynamics $\dot{\hat{x}}_i$ are introduced, the feedback control structure remains distributed.

Theorem 4.3 *Under Assumption 4.1, starting from any non-negative initial condition, system (4.21) will converge to the unique optimal solution of:*

$$\max_{x_i \geq 0} \sum_{i=1}^{N} U_i(x_i) - \sum_{l=1}^{L} \int_0^{y_l} f_l(\beta) d\beta. \tag{4.22}$$

Proof: Noting that at the equilibrium point x_i and $\hat{x}_i$ are the same, the uniqueness and optimality of the equilibrium point can be found in [50] (Section 3). Consider the following radially unbounded function as a candidate Lyapunov function

$$V_{(4.21)}(x_i, \hat{x}_i) = -\sum_{l=1}^{L}(U_i(x_i) - U_i(x_i^*)) + \sum_{l=1}^{L}\int_{y_l^*}^{y_l} f_l(\beta)d\beta + \sum_{i=1}^{N}\frac{k_{e_i}}{2}(x_i - \hat{x}_i)^2 \tag{4.23}$$

where $x_i^* = \hat{x}_i^*$ holds. This function is non-negative with a minimum at the equilibrium. Differentiating $V_{(4.21)}$ with respect to time, we get

$$\begin{aligned}
\dot{V}_{(4.21)} =& -\sum_{i=1}^{N} U_i'(x_i)\dot{x}_i + \sum_{l=1}^{L} p_l \dot{y}_l + \sum_{i=1}^{N} k_{e_i}(x_i - \hat{x}_i)(\dot{x}_i - \dot{\hat{x}}_i) \\
=& -\sum_{i=1}^{N}(U_i'(x_i) - q_i + k_{e_i}(\hat{x}_i - x_i))\dot{x}_i - \sum_{i=1}^{N} k_{e_i}(x_i - \hat{x}_i)\dot{\hat{x}}_i \\
=& -\sum_{i=1}^{N} k_{x_i}(U_i'(x_i) - q_i + k_{e_i}(\hat{x}_i - x_i))^2 - \sum_{i=1}^{N} k_{e_i}\hat{k}_{e_i}(x_i - \hat{x}_i)^2 \\
\leq& 0.
\end{aligned} \tag{4.24}$$

Now $\dot{V}_{(4.21)} = 0$ only when $U_i'(x_i) - q_i + k_{e_i}(\hat{x}_i - x_i) = 0, x_i = \hat{x}_i$, which can only happen at the equilibrium of interest. ■

Theorem 4.3 indicates that the extra dynamics do not affect the global asymptotic stability of the overall system. Moreover, there is no restriction for the choice of k_{e_i} and $\hat{k}_{e_i}$, except positivity. We next investigate how to tune them in order to achieve better performance for the closed-loop system.

4.3.2 Linear Robust Analysis

Linearizing (4.21) around the equilibrium point and expressing it in the Laplace domain, we obtain

$$\delta x_i(s) = \frac{k_{x_i}}{s}\left(-\kappa_i \delta x_i(s) - \sum_{l=1}^{L} R_{li} f_l'(y_l^*) \sum_{n=1}^{N} R_{ln}\delta x_n(s) - \frac{k_{e_i} s}{s + \hat{k}_{e_i}}\delta x_i(s)\right) \tag{4.25}$$

where $\delta x_i(s)$ denotes the linearized state, $\kappa_i = -U_i''(x_i^*)$ and $y_l^* = \sum_{i=1}^{N} R_{li} x_i^*$. We rearrange the transfer functions as a standard unity negative feedback system so that the loop transfer function matrix is given by

$$L_{(4.25)}(s) = \text{diag}\left\{\frac{k_{x_i}(s + \hat{k}_{e_i})}{s^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i}}\right\} R^T \text{diag}\{f_l'(y_l^*)\}R. \tag{4.26}$$

Based on classical loop shaping methods, in order to achieve better performance, $\underline{\sigma}(L_{(4.25)}(s))$ needs to be large at low frequencies and $\overline{\sigma}(L_{(4.25)}(s))$ needs to be small at high frequencies [89]. Since

$$\begin{aligned}
&\underline{\sigma}\left(\text{diag}\left\{\frac{k_{x_i}(s + \hat{k}_{e_i})}{s^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i}}\right\}\right)\underline{\sigma}\left(R^T \text{diag}\{f_l'(y_l^*)\}R\right)\\
\leq &\underline{\sigma}(L_{(4.25)}(s))\\
&\overline{\sigma}\left(\text{diag}\left\{\frac{k_{x_i}(s + \hat{k}_{e_i})}{s^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})s + k_{x_i}\kappa_i\hat{k}_{e_i}}\right\}\right)\overline{\sigma}\left(R^T \text{diag}\{f_l'(y_l^*)\}R\right)\\
\geq &\overline{\sigma}(L_{(4.25)}(s))
\end{aligned}$$

a sufficient condition is to make the lower bound of $\underline{\sigma}(L_{(4.25)}(s))$ large at low frequencies and the upper bound of $\overline{\sigma}(L_{(4.25)}(s))$ small at high frequencies. We then have the following remark.

Remark 4.5 *Let*

$$K_i(j\omega) = \frac{k_{x_i}(j\omega + \hat{k}_{e_i})}{k_{x_i}\kappa_i\hat{k}_{e_i} - \omega^2 + (k_{x_i}\kappa_i + k_{x_i}k_{e_i} + \hat{k}_{e_i})j\omega}.$$

Since

$$\frac{\partial|K_i|^2}{\partial k_{e_i}} < 0$$

$$\frac{\partial|K_i|^2}{\partial \hat{k}_{e_i}} = \frac{2k_{x_i}^3 k_{e_i}\omega^2(\hat{k}_{e_i}(\hat{k}_{e_i}+2k_{x_i}\kappa_i+k_{x_i}k_{e_i})-\omega^2)}{((\omega^2-k_{x_i}\kappa_i\hat{k}_{e_i})^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})^2\omega^2)^2}$$

small k_{e_i} increases the lower bound of $\underline{\sigma}(L_{(4.25)}(s))$ at low frequencies while also increasing the upper bound of $\overline{\sigma}(L_{(4.25)}(s))$ at high frequencies. Large $\hat{k}_{e_i}$ increases the lower bound of $\underline{\sigma}(L_{(4.25)}(s))$ at low frequencies and decreases the upper bound of $\overline{\sigma}(L_{(4.25)}(s))$ at high frequencies.

This means that in order to improve the performance of the closed-loop system, there is a trade-off for k_{e_i} (not very large or very small) and $\hat{k}_{e_i}$ needs to be large.

4.3.3 Robust Stability to Delays

We now consider delay robustness of the overall system. In the delayed case, the closed-loop dynamics (4.21) change to

$$\dot{x}_i(t) = k_{x_i}\left(U_i'(x_i(t)) - \sum_{l=1}^{L} R_{li} f_l\left(\sum_{n=1}^{N} R_{ln}x_n(t-\tau_{i,l}^b-\tau_{n,l}^f)\right) + k_{e_i}(\hat{x}_i(t)-x_i(t))\right) \tag{4.27a}$$

$$\dot{\hat{x}}_i(t) = \hat{k}_{e_i}(x_i(t)-\hat{x}_i(t)) \tag{4.27b}$$

where $\tau_{i,l}^f$ and $\tau_{i,l}^b$ are the same as previously defined, and $\tau_i = \tau_{i,l}^f + \tau_{i,l}^b,\ \forall\ l = 1,\cdots,L$.

We linearize (4.27) around the equilibrium and express it in the Laplace domain

$$\delta x_i(s) = \frac{k_{x_i}}{s}\left(-\kappa_i\delta x_i(s) - \sum_{l=1}^{L} R_{li}^b(s) f_l'(y_l^*)\sum_{n=1}^{N} R_{ln}^f(s)\delta x_n(s) - \frac{k_{e_i}s}{s+\hat{k}_{e_i}}\delta x_i(s)\right) \tag{4.28}$$

where all states and parameters are defined as stated previously. We rearrange the transfer functions as a standard unity negative feedback system so that the loop transfer function matrix is given by

$$L_{(4.28)}(s) = \text{diag}\left\{\frac{k_{x_i}(s+\hat{k}_{e_i})}{s^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})s+k_{x_i}\kappa_i\hat{k}_{e_i}}\right\} R^b(s)^T \text{diag}\{f_l'(y_l^*)\}R^f(s).$$

Noting that $R^b(s) = R^f(-s)\text{diag}\{e^{-\tau_i s}\}$, we have

$$L_{(4.28)}(s) = \text{diag}\left\{\frac{k_{x_i}e^{-\tau_i s}(s+\hat{k}_{e_i})}{s^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})s+k_{x_i}\kappa_i\hat{k}_{e_i}}\right\} R^f(-s)^T \text{diag}\{f_l'(y_l^*)\}R^f(s). \tag{4.29}$$

Remark 4.6 *When $k_{e_i} = 0$, i.e., there is no extra dynamics added to the system, (4.29) becomes*

$$diag\left\{\frac{k_{x_i}e^{-\tau_i s}}{s+k_{x_i}\kappa_i}\right\}R^f(-s)^T\,diag\{f_l'(y_l^*)\}R^f(s).$$

Comparing this transfer function matrix to (4.29), we can see that the redesigned dynamics add a band rejection filter $\frac{(s+k_{x_i}\kappa_i)(s+\hat{k}_{e_i})}{s^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})s+k_{x_i}\kappa_i\hat{k}_{e_i}}$ to each $\frac{k_{x_i}e^{-\tau_i s}}{s(s+k_{x_i}\kappa_i)}$ in the modified system.

Under Remark 4.6, Lemma 4.1 and following the proof of Theorem 1 in [95], we can derive a local stability condition for the system without extra dynamics, given by

$$\sum_{l=1}^{L}\sum_{n=1}^{N}R_{li}R_{ln}f_l'(y_l^*) < \min_n \frac{\sqrt{\omega_n^2+k_{x_n}^2\kappa_n^2}}{k_{x_n}},\ i=1,\cdots,N \tag{4.30}$$

where $\omega_n > 0$ is the solution to $\tau_n\omega_n - \frac{\pi}{2} = \arctan\frac{k_{x_n}\kappa_n}{\omega_n}, n = 1,\cdots,N$. With the extra dynamics, this stability condition can be much less conservative, as demonstrated in the following theorem.

Theorem 4.4 *For fixed full rank R, the equilibrium point of (4.27) is unique, optimal with respect to problem (4.22), and is locally asymptotically stable provided*

$$\sum_{l=1}^{L}\sum_{n=1}^{N}R_{li}R_{ln}f_l'(y_l^*) < \min_n\left(\frac{\sqrt{\omega_n^2+k_{x_n}^2\kappa_n^2}}{k_{x_n}}\left(1+\frac{k_{x_n}^2\kappa_n k_{e_n}}{\omega_n^2+k_{x_n}^2\kappa_n^2}\right)\right), i=1,\cdots,N \tag{4.31}$$

where $\omega_n > 0$ is the solution to $\tau_n\omega_n - \frac{\pi}{2} = \arctan\frac{k_{x_n}\kappa_n}{\omega_n}$ and $\hat{k}_{e_n}$ satisfies $k_{x_n}\kappa_n\hat{k}_{e_n} = \omega_n^2, n = 1,\cdots,N$.

Proof: Noting that at the equilibria x_i and $\hat{x}_i$ are the same, the uniqueness and optimality of the equilibrium point follow from Theorem 4.3. For asymptotic stability, according to the proof of Theorem 1 in [95], it is sufficient to show that

$$-1 \notin \text{eig}(Q(j\omega)\Lambda(j\omega))$$

where

$$Q(j\omega) = Q(j\omega)^H = R^f(j\omega)^H\text{diag}\{f_l'(y_l^*)\}R^f(j\omega)$$
$$\Lambda(j\omega) = \text{diag}\left\{\frac{k_{x_i}e^{-j\tau_i\omega}(j\omega+\hat{k}_{e_i})}{k_{x_i}\kappa_i\hat{k}_{e_i}-\omega^2+(k_{x_i}\kappa_i+k_{x_i}k_{e_i}+\hat{k}_{e_i})j\omega}\right\}.$$

Since the Nyquist plots of the entries of $\Lambda(j\omega)$ first cross the negative real axis at the frequency satisfying $\tau_i\omega - \frac{\pi}{2} = \arctan\frac{k_{x_i}\kappa_i}{\omega}$ when $k_{x_i}\kappa_i\hat{k}_{e_i} = \omega^2$ holds, based on Lemma 1 in [72], it is sufficient to show that

$$\rho(Q(j\omega))\max_n \frac{k_{x_n}\sqrt{\omega_n^2 + \hat{k}_{e_n}^2}}{\omega_n(k_{x_n}\kappa_n + k_{x_n}k_{e_n} + \hat{k}_{e_n})} < 1$$

where $\omega_n > 0$ is the solution to $\tau_n\omega_n - \frac{\pi}{2} = \arctan\frac{k_{x_n}\kappa_n}{\omega_n}, n = 1, \cdots, N$. According to Lemma 4.1, this is satisfied when (4.31) holds, which completes the proof. ■

The condition in Theorem 4.4 indicates that by increasing k_{e_i} and choosing $\hat{k}_{e_i}$ to satisfy $k_{x_i}\kappa_i\hat{k}_{e_i} = \omega^2$ where $\omega > 0$ solves $\tau_i\omega - \frac{\pi}{2} = \arctan\frac{k_{x_i}\kappa_i}{\omega}$, which is distributed, the overall system can achieve added delay robustness. Combining Remark 4.5 and Theorem 4.4, we can see that there is a trade-off between these gains, as expected.

4.4 Redesign for the Dual Algorithm

We now use (4.10c) to redesign the dual algorithm, investigating how the gains resulting from the extra dynamics influence the stability and robustness of the closed-loop system. Since the redesign process is similar to that in the previous section, some details are omitted for brevity.

4.4.1 Modified Dynamics and Stability

Based on a dual approach [50] (Section 3.2), we obtain the modified dynamics (in continuous-time) given by

$$\dot{p}_l = k_{p_l}(y_l - c_l + k_{e_l}(\hat{p}_l - p_l))^+_{p_l} \tag{4.32a}$$

$$\dot{\hat{p}}_l = \hat{k}_{e_l}(p_l - \hat{p}_l) \tag{4.32b}$$

$$x_i = U_i'^{-1}(q_i) \tag{4.32c}$$

where $\hat{k}_{e_l} > 0$. Although the extra dynamics $\dot{\hat{p}}_l$ are introduced, the feedback control structure remains distributed.

Theorem 4.5 *Under Assumption 4.1, the equilibrium point of (4.32) is unique, satisfies the KKT conditions (4.5) and is globally asymptotically stable for all non-negative initial conditions.*

Proof: Noting that at the equilibrium point p_l and $\hat{p}_l$ are the same, the proof of uniqueness and optimality of the equilibrium point can be found in [50] (Section 3). Consider the following radially unbounded function as a candidate Lyapunov function

$$V_{(4.32)}(p_l, \hat{p}_l) = \sum_{l=1}^{L}(c_l - y_l^*)p_l + \sum_{i=1}^{N}\int_{q_i^*}^{q_i}(x_i^* - U_i'^{-1}(\beta))d\beta + \sum_{l=1}^{L}\frac{k_{e_l}}{2}(p_l - \hat{p}_l)^2 \quad (4.33)$$

where $p_l^* = \hat{p}_l^*$ holds. This function is non-negative for $p_l \geq 0$ with a minimum at the equilibrium. Differentiating $V_{(4.32)}$ with respect to time, we get

$$\begin{aligned}
\dot{V}_{(4.32)} =& \sum_{l=1}^{L}(c_l - y_l^* - k_{e_l}(\hat{p}_l - p_l) + k_{e_l}(\hat{p}_l - p_l))\dot{p}_l + \sum_{i=1}^{N}(x_i^* - U_i'^{-1}(q_i))\dot{q}_i \\
&+ \sum_{l=1}^{L} k_{e_l}(p_l - \hat{p}_l)(\dot{p}_l - \dot{\hat{p}}_l) \\
=& \sum_{l=1}^{L}(c_l - y_l^* - k_{e_l}(\hat{p}_l - p_l))\dot{p}_l + \sum_{i=1}^{N}(x_i^* - U_i'^{-1}(q_i))\dot{q}_i - \sum_{l=1}^{L} k_{e_l}\hat{k}_{e_l}(p_l - \hat{p}_l)^2 \\
=& -\sum_{l=1}^{L} k_{p_l}(y_l - c_l + k_{e_l}(\hat{p}_l - p_l))(y_l - c_l + k_{e_l}(\hat{p}_l - p_l))_{p_l}^{+} - \sum_{l=1}^{L} k_{e_l}\hat{k}_{e_l}(p_l - \hat{p}_l)^2 \\
\leq& 0. \qquad (4.34)
\end{aligned}$$

Now $\dot{V}_{(4.32)} = 0$ only when $y_l - c_l + k_{e_l}(\hat{p}_l - p_l) = 0, p_l = \hat{p}_l$ or $y_l - c_l + k_{e_l}(\hat{p}_l - p_l) < 0, p_l = \hat{p}_l = 0$, which can only happen at the equilibrium of interest. ■

Theorem 4.5 indicates that the extra dynamics do not affect the global asymptotic stability of the overall system. Moreover, there is no restriction for the choice of k_{e_l} and $\hat{k}_{e_l}$, except positivity. We next investigate how to tune them in order to achieve better performance for the closed-loop system.

4.4.2 Linear Robust Analysis

Linearizing (4.32) around the equilibrium point and expressing it in the Laplace domain, we obtain

$$\delta\bar{p}_l(s) = \frac{k_{p_l}}{s}\left(-\sum_{i=1}^{N}\bar{R}_{li}\frac{1}{\kappa_i}\sum_{m=1}^{\bar{L}}\bar{R}_{mi}\delta\bar{p}_m(s) - \frac{k_{e_l}s}{s + \hat{k}_{e_l}}\delta\bar{p}_l(s)\right). \quad (4.35)$$

In this transfer function, $\delta\bar{p}_l(s)$ denotes the linearized reduced state by eliminating non-bottleneck links. $\bar{R}$ is obtained by eliminating non-bottleneck rows from R and $\bar{R}_{li}$ is the entries of $\bar{R}$. $\bar{L}$ is the number of bottleneck links. Following Assumption 4.1, we will assume that $\bar{R}$ is of full row rank. $\kappa_i = -U_i''(x_i^*)$. We rearrange the transfer

functions as a standard unity negative feedback system so that the loop transfer function matrix is given by

$$L_{(4.35)}(s) = \text{diag}\left\{\frac{k_{p_l}(s+\hat{k}_{e_l})}{s(s+\hat{k}_{e_l}+k_{p_l}k_{e_l})}\right\}\bar{R}\text{diag}\left\{\frac{1}{\kappa_i}\right\}\bar{R}^T. \tag{4.36}$$

Based on classical loop shaping methods, in order to achieve better performance, $\underline{\sigma}(L_{(4.35)}(s))$ needs to be large at low frequencies and $\overline{\sigma}(L_{(4.35)}(s))$ needs to be small at high frequencies [89]. Since

$$\underline{\sigma}\left(\text{diag}\left\{\frac{k_{p_l}(s+\hat{k}_{e_l})}{s(s+\hat{k}_{e_l}+k_{p_l}k_{e_l})}\right\}\right)\underline{\sigma}\left(\bar{R}\text{diag}\left\{\frac{1}{\kappa_i}\right\}\bar{R}^T\right) \leq \underline{\sigma}(L_{(4.35)}(s))$$

$$\overline{\sigma}(L_{(4.35)}(s)) \leq \overline{\sigma}\left(\text{diag}\left\{\frac{k_{p_l}(s+\hat{k}_{e_l})}{s(s+\hat{k}_{e_l}+k_{p_l}k_{e_l})}\right\}\right)\overline{\sigma}\left(\bar{R}\text{diag}\left\{\frac{1}{\kappa_i}\right\}\bar{R}^T\right)$$

a sufficient condition is to make the lower bound of $\underline{\sigma}(L_{(4.35)}(s))$ large at low frequencies and the upper bound of $\overline{\sigma}(L_{(4.35)}(s))$ small at high frequencies. We then have the following remark.

Remark 4.7 *Let*

$$K_l(j\omega) = \frac{k_{p_l}(j\omega+\hat{k}_{e_l})}{j\omega(j\omega+\hat{k}_{e_l}+k_{p_l}k_{e_l})}.$$

Since

$$\frac{\partial|K_l|^2}{\partial k_{e_l}} < 0$$

$$\frac{\partial|K_l|^2}{\partial \hat{k}_{e_l}} = \frac{2k_{p_l}^3 k_{e_l}(\hat{k}_{e_l}(\hat{k}_{e_l}+k_{p_l}k_{e_l})-\omega^2)}{\omega^2(\omega^2+(\hat{k}_{e_l}+k_{p_l}k_{e_l})^2)^2}$$

small k_{e_l} increases the lower bound of $\underline{\sigma}(L_{(4.35)}(s))$ at low frequencies while also increasing the upper bound of $\overline{\sigma}(L_{(4.35)}(s))$ at high frequencies. Large $\hat{k}_{e_l}$ increases the lower bound of $\underline{\sigma}(L_{(4.35)}(s))$ at low frequencies and decreases the upper bound of $\overline{\sigma}(L_{(4.35)}(s))$ at high frequencies.

This means that in order to improve the performance of the closed-loop system, there is a trade-off for k_{e_l} (not very large or very small) and $\hat{k}_{e_l}$ needs to be large.

4.4.3 Robust Stability to Delays

We now consider delay robustness of the overall system. In the delayed case, the closed-loop dynamics (4.32) change to

$$\dot{p}_l(t) = k_{p_l}\left(\sum_{i=1}^{N} R_{li}U_i'^{-1}\left(\sum_{m=1}^{L} R_{mi}p_m(t-\tau_{i,l}^f-\tau_{i,m}^b)\right) - c_l + k_{e_l}(\hat{p}_l(t)-p_l(t))\right)_{p_l(t)}^{+} \tag{4.37a}$$

$$\dot{\hat{p}}_l(t) = \hat{k}_{e_l}(p_l(t)-\hat{p}_l(t)) \tag{4.37b}$$

where $\tau_{i,l}^f$ and $\tau_{i,l}^b$ are the same as previously defined, and $\tau_i = \tau_{i,l}^f + \tau_{i,l}^b,\ \forall\ l = 1,\cdots,L$.

We linearize (4.37) around the equilibrium and express it in the Laplace domain

$$\delta\bar{p}_l(s) = \frac{k_{p_l}}{s}\left(-\sum_{i=1}^{N}\bar{R}_{li}^f(s)\frac{1}{\kappa_i}\sum_{m=1}^{\bar{L}}\bar{R}_{mi}^b(s)\delta\bar{p}_m(s) - \frac{k_{e_l}s}{s+\hat{k}_{e_l}}\delta\bar{p}_l(s)\right) \tag{4.38}$$

where all states and parameters are defined as stated previously. We rearrange the transfer functions as a standard unity negative feedback system so that the loop transfer function matrix is given by

$$L_{(4.38)}(s) = \text{diag}\left\{\frac{k_{p_l}(s+\hat{k}_{e_l})}{s(s+\hat{k}_{e_l}+k_{p_l}k_{e_l})}\right\}\bar{R}^f(s)\text{diag}\left\{\frac{1}{\kappa_i}\right\}\bar{R}^b(s)^T.$$

Noting that $\bar{R}^b(s) = \bar{R}^f(-s)\text{diag}\{e^{-\tau_i s}\}$, we have

$$L_{(4.38)}(s) = \text{diag}\left\{\frac{k_{p_l}(s+\hat{k}_{e_l})}{s(s+\hat{k}_{e_l}+k_{p_l}k_{e_l})}\right\}\bar{R}^f(s)\text{diag}\left\{\frac{e^{-\tau_i s}}{\kappa_i}\right\}\bar{R}^f(-s)^T. \tag{4.39}$$

Remark 4.8 *When $k_{e_l} = 0$, i.e., there is no extra dynamics added to the system, (4.39) becomes*

$$diag\left\{\frac{k_{p_l}}{s}\right\}\bar{R}^f(s)\,diag\left\{\frac{e^{-\tau_i s}}{\kappa_i}\right\}\bar{R}^f(-s)^T.$$

Comparing this transfer function matrix to (4.39), we can see that the redesigned dynamics add a filter $\frac{s+\hat{k}_{e_l}}{s+\hat{k}_{e_l}+k_{p_l}k_{e_l}}$ to each $\frac{k_{p_l}}{s}$ in the modified system.

Under Remark 4.8, Lemma 4.1 and following the proof of Theorem 1 in [64], we can derive a local stability condition for the system without extra dynamics, given by

$$k_{p_l}\sum_{i=1}^{N}\sum_{n=1}^{\bar{L}}\frac{\bar{R}_{li}\bar{R}_{ni}\tau_i}{\kappa_i} < \frac{\pi}{2},\ l = 1,\cdots,\bar{L}. \tag{4.40}$$

With the extra dynamics, this stability condition can be much less conservative, as demonstrated in the following theorem.

Theorem 4.6 *For fixed full rank $\bar{R}$ ($\bar{R}^f(0) = \bar{R}^b(0) = \bar{R}$), the equilibrium point of (4.37) is unique, satisfies the KKT conditions (4.5) and is locally asymptotically stable provided*

$$k_{p_l} \sum_{i=1}^{N} \sum_{n=1}^{\bar{L}} \frac{\bar{R}_{li}\bar{R}_{ni}\tau_i}{\kappa_i} < \min_i \tau_i\omega_i \sqrt{1 + \frac{2\hat{k}k + k^2}{\omega_i^2 + \hat{k}^2}}, l = 1, \cdots, \bar{L} \tag{4.41}$$

where $\omega_i > 0$ is the solution to $\tau_i\omega_i = \frac{\pi}{2} + \arctan\frac{\omega_i}{\hat{k}} - \arctan\frac{\omega_i}{\hat{k}+k}, i = 1, \cdots, N$, and $\hat{k}_{e_l} = \hat{k} > 0, k_{p_l}k_{e_l} = k > 0, l = 1, \cdots, L$.

Proof: Noting that at the equilibria p_l and $\hat{p}_l$ are the same, the uniqueness and optimality of the equilibrium point follow from Theorem 4.5. Under the condition $\hat{k}_{e_l} = \hat{k} > 0, k_{p_l}k_{e_l} = k > 0, l = 1, \cdots, L$, (4.39) becomes

$$L_{(4.38)}(s) = \text{diag}\{k_{p_l}\}\bar{R}^f(s)\text{diag}\left\{\frac{\tau_i}{\kappa_i}\right\}\text{diag}\left\{\frac{e^{-\tau_i s}(s+\hat{k})}{\tau_i s(s+\hat{k}+k)}\right\}\bar{R}^f(-s)^T.$$

For asymptotic stability, according to Proposition 2 in [64], it is sufficient to show that

$$-1 \notin \text{eig}(\mu L_{(4.38)}(j\omega)),\ \mu \in (0, 1],\ \omega \neq 0 \Leftrightarrow -1 \notin \text{eig}(Q(j\omega)\Lambda(j\omega))$$

where

$$Q(j\omega) = Q(j\omega)^H = \text{diag}\left\{\sqrt{\frac{\tau_i}{\kappa_i}}\right\} R^f(j\omega)^H \text{diag}\{k_{p_l}\}\bar{R}^f(j\omega)\text{diag}\left\{\sqrt{\frac{\tau_i}{\kappa_i}}\right\}$$

$$\Lambda(j\omega) = \text{diag}\left\{\frac{e^{-j\tau_i\omega}(j\omega+\hat{k})}{j\tau_i\omega(j\omega+\hat{k}+k)}\right\}.$$

Since the Nyquist plots of the entries of $\Lambda(j\omega)$ first cross the negative real axis at the frequency satisfying $\tau_i\omega = \frac{\pi}{2} + \arctan\frac{\omega}{\hat{k}} - \arctan\frac{\omega}{\hat{k}+k}$, based on Lemma 1 in [72], it is sufficient to show that

$$\rho(Q(j\omega)) \max_i \frac{\sqrt{\omega_i^2 + \hat{k}^2}}{\tau_i\omega_i\sqrt{\omega_i^2 + (\hat{k}+k)^2}} < 1$$

where $\omega_i > 0$ is the solution to $\tau_i\omega_i = \frac{\pi}{2} + \arctan\frac{\omega_i}{\hat{k}} - \arctan\frac{\omega_i}{\hat{k}+k}, i = 1, \cdots, N$. According to Lemma 4.1, this is satisfied when (4.41) holds, which completes the proof. ■

Since

$$\left(\frac{\pi}{2} + \arctan\frac{\omega_i}{\hat{k}} - \arctan\frac{\omega_i}{\hat{k}+k}\right)\sqrt{1 + \frac{2\hat{k}k + k^2}{\omega_i^2 + \hat{k}^2}} > \frac{\pi}{2},\ \forall\ i$$

the stability condition can be much less conservative after adding the extra dynamics, indicating improved robustness to time delays.

Remark 4.9 *The condition in Theorem 4.6 requires* $\hat{k}_{e_l} = \hat{k} > 0, k_{p_l} k_{e_l} = k > 0, l = 1, \cdots, L$*, i.e., homogenous gains in the extra dynamics. In fact, there are alternative stability conditions which can be derived following from Theorem 4.6. For example, the equilibrium point of (4.37) is locally asymptotically stable provided*

$$k_{p_l} \sum_{i=1}^{N} \sum_{n=1}^{\bar{L}} \frac{\bar{R}_{li} \bar{R}_{ni} \tau_i}{\kappa_i} < \min_i \tau_i \omega_i \sqrt{1 + \frac{2\hat{k} k_{p_m} k_{e_m} + k_{p_m}^2 k_{e_m}^2}{\omega_i^2 + \hat{k}^2}}, l = 1, \cdots, \bar{L}$$

where $\omega_i > 0$ *is the solution to* $\tau_i \omega_i = \frac{\pi}{2} + \arctan \frac{\omega_i}{\hat{k}} - \arctan \frac{\omega_i}{\hat{k} + k_{p_m} k_{e_m}}, i = 1, \cdots, N$*,* $\hat{k}_{e_l} = \hat{k} > 0, l = 1, \cdots, L$*, and* (k_{p_m}, k_{e_m}) *are derived from the set of* (k_{p_l}, k_{e_l}) *such that* $k_{p_m} k_{e_m} = \min_l \{k_{p_l} k_{e_l}\}$*.*

4.5 A Distributed PID Controller

In the previous sections, we have redesigned the primal-dual, primal and dual algorithms by introducing extra dynamics. We now turn to study the meaning of these extra dynamics and further introduce a design framework of distributed PID control for network congestion control problems.

4.5.1 Derivative Action with Filtering

We first focus on the standard primal-dual algorithm given by

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - q_i)$$
$$\dot{p}_l = k_{p_l}(y_l - c_l)_{p_l}^+$$

where $q_i = \sum_{l=1}^{L} R_{li} p_l$, $y_l = \sum_{i=1}^{N} R_{li} x_i$. From a control perspective, we can regard the scheme as one in which each user adopts integral control action to adjust his source rate in a distributed manner. This is because the above scheme can be written as

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - u_i) \tag{4.42a}$$
$$u_i = q_i = \sum_{l=1}^{L} R_{li} p_l \tag{4.42b}$$
$$p_l = \int_0^t k_{p_l} \left(\sum_{i=1}^{N} R_{li} x_i(\beta) - c_l \right)_{p_l}^+ d\beta, \ \sum_{i=1}^{N} R_{li} x_i(0) - c_l \leq 0. \tag{4.42c}$$

This is in accord with the paradigm proposed in [57]. On the other hand, consider the redesign primal-dual algorithm given by

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - \underbrace{(q_i + k_{e_i}(x_i - \hat{x}_i))}_{u_i}) \tag{4.43a}$$

$$\dot{\hat{x}}_i = \hat{k}_{e_i}(x_i - \hat{x}_i) \tag{4.43b}$$

$$\dot{p}_l = k_{p_l}(y_l - c_l)^+_{p_l}. \tag{4.43c}$$

Express (4.43b) in the Laplace domain as

$$\hat{x}_i(s) = \frac{\hat{k}_{e_i}}{s + \hat{k}_{e_i}} x_i(s).$$

The controller, from (4.43a), is therefore

$$\begin{aligned} u_i(s) =& q_i(s) + \frac{k_{e_i}}{s + \hat{k}_{e_i}} s x_i(s) \\ =& q_i(s) + \frac{k_{e_i}}{\hat{k}_{e_i}} s\hat{x}_i(s) \end{aligned} \tag{4.44}$$

which contains both integral action and derivative action with filtering.

We can compare (4.44) with another potential derivative action given by

$$u_i = q_i + k_{d_i}\dot{x}_i \tag{4.45}$$

where $k_{d_i} > 0$. This is equivalent to changing the gain k_{x_i} to $\frac{k_{x_i}}{1+k_{x_i}k_{d_i}}$ in (4.42). On the contrary, the term $\frac{k_{e_i}}{\hat{k}_{e_i}} s\hat{x}_i(s)$ in (4.44) is actually the high frequency component of x_i which can hinder x_i from changing too quickly. By tuning the gains appropriately, it can improve the behaviour of the overall system at high frequencies while not affecting the system greatly at low frequencies. Figure 4.3 shows the system responses with (4.45) and (4.44) respectively (We use a 3-source-2-link network as illustrated in Figure 4.4. All utility functions of the sources are $U_i(x_i) = \log(x_i)$. $R = [1\ 1\ 0;\ 1\ 0\ 1]$ and $c_l = (2, 1)$. The controller gains are: all $k_{x_i} = 3$; all $k_{p_l} = 1$; all $k_{d_i} = 1$ in (4.45); all $k_{e_i} = 2$ and all $\hat{k}_{e_i} = 2$ (i.e., $\frac{k_{e_i}}{\hat{k}_{e_i}} = 1$) in (4.44). The scenario consists of one of the sources suddenly changing its utility function, from $\log(x_i)$ to $3\log(x_i)$). It can be seen that using derivative action with filtering results in less overshoot and oscillations than using (4.45) under the same conditions.

Similarly, let us compare the standard primal algorithm given by

$$\begin{aligned} \dot{x}_i &= k_{x_i}(U_i'(x_i) - q_i) \\ p_l &= f_l(y_l) \end{aligned}$$

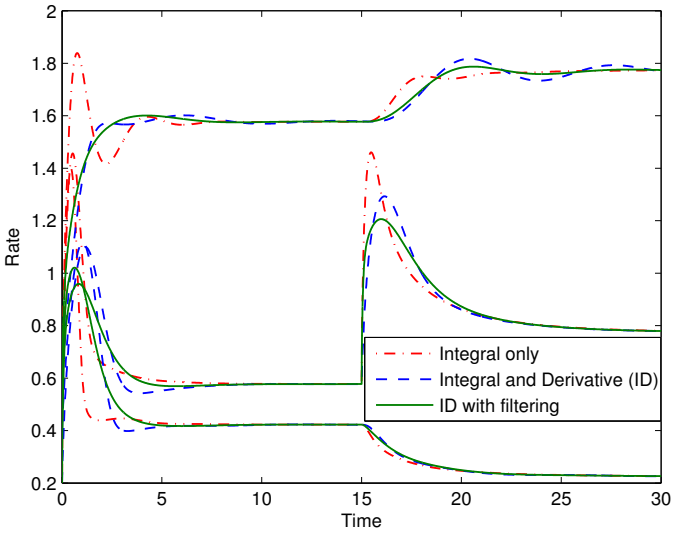

Figure 4.3: Simulation results comparing (4.45) (ID) and (4.44) (ID with filtering): responses of source rate.

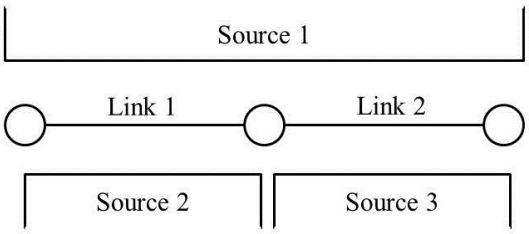

Figure 4.4: A 3-source-2-link network.

and redesign primal algorithm given by

$$\dot{x}_i = k_{x_i}(U_i'(x_i) - \underbrace{(q_i + k_{e_i}(x_i - \hat{x}_i))}_{u_i})$$
$$\dot{\hat{x}}_i = \hat{k}_{e_i}(x_i - \hat{x}_i)$$
$$p_l = f_l(y_l).$$

In the standard algorithm, the input actually corresponds to proportional action; in the redesigned algorithm, derivative active with filtering is introduced. Therefore, the redesigned primal algorithm contains a distributed Proportional-Derivative (PD) controller.

Finally, we compare the standard dual algorithm given by

$$\dot{p}_l = k_{p_l}(\underbrace{y_l - c_l}_{-u_l})^+_{p_l}$$
$$x_i = U_i'^{-1}(q_i)$$

and the redesigned dual algorithm given by

$$\dot{p}_l = k_{p_l}(\underbrace{y_l - c_l + k_{e_l}(\hat{p}_l - p_l)}_{-u_l})^+_{p_l}$$
$$\dot{\hat{p}}_l = \hat{k}_{e_l}(p_l - \hat{p}_l)$$
$$x_i = U_i'^{-1}(q_i).$$

Likewise, we can see that the redesigned dual algorithm contains a distributed PD controller.

4.5.2 A Distributed PID Controller

In this subsection, we extend the redesign of the primal-dual algorithm and propose a distributed PID controller for problem (4.1). Consider the following augmented Lagrangian

$$\begin{aligned} L_e(x_i, p_l, \hat{x}_i) = &\sum_{i=1}^{N} U_i(x_i) + \sum_{l=1}^{L} p_l \left(c_l - \sum_{i=1}^{N} R_{li} x_i \right) - \sum_{i=1}^{N} \frac{k_{e_i}}{2} (x_i - \hat{x}_i)^2 \\ &- k \sum_{l=1}^{L} \frac{k_{p_l}}{2} \psi \left(\sum_{i=1}^{N} R_{li} x_i - c_l \right)^2 \end{aligned} \tag{4.46}$$

where $k > 0$, k_{p_l} equals to the gain of the dynamics of p_l and $\psi(x) = \max\{0, x\}$. This modified Lagrangian then leads to the dynamics with a distributed PID controller given by

$$\dot{x}_i = k_{x_i} \left(U_i'(x_i) - \underbrace{\left(q_i + k \sum_{l=1}^{L} R_{li} \psi(\dot{p}_l) + k_{e_i}(x_i - \hat{x}_i) \right)}_{u_i} \right) \tag{4.47a}$$
$$\dot{\hat{x}}_i = \hat{k}_{e_i}(x_i - \hat{x}_i) \tag{4.47b}$$
$$\dot{p}_l = k_{p_l}(y_l - c_l)^+_{p_l}. \tag{4.47c}$$

Since $\dot{p}_l$ is proportional to x_i, when $\dot{p}_l > 0$ ($\psi(\dot{p}_l) = \dot{p}_l > 0$), there is proportional action activated in each $\dot{x}_i$, and k is therefore the common proportional gain. We then have the following theorem.

Theorem 4.7 *Under Assumption 4.1, the equilibrium point of (4.47) is unique, satisfies the KKT conditions (4.5), and is globally asymptotically stable for all non-negative initial conditions.*

Proof: Noting that at the equilibrium point x_i and $\hat{x}_i$ are the same, the proof of uniqueness and optimality of the equilibrium point can be found in [50] (Section 3). For the global asymptotic stability, see the proof of Theorem 4.8 in the sequel. ■

To achieve more flexibility for the choice of controller gains, we next consider the case that each user can tune the proportional gain autonomously. Then (4.47a) changes to

$$\dot{x}_i = k_{x_i}\left(U_i'(x_i) - \underbrace{\left(q_i + k_i\sum_{l=1}^{L} R_{li}\psi(\dot{p}_l) + k_{e_i}(x_i - \hat{x}_i)\right)}_{u_i}\right) \tag{4.48}$$

where $k_i > 0$ stands for each proportional gain. Note that the dynamics (4.48) and (4.47b)-(4.47c) do not result from L_e.

Theorem 4.8 *Under Assumption 4.1, the equilibrium point of (4.48) and (4.47b)-(4.47c) is unique, satisfies the KKT conditions (4.5), and is globally asymptotically stable for all non-negative initial conditions.*

Proof: We only need to show that the equilibrium is globally asymptotically stable, where we use techniques from [56]. Consider the following radially unbounded function as a candidate Lyapunov function

$$V_{(4.48),(4.47b)-(4.47c)}(x_i, p_l, \hat{x}_i) = \sum_{i=1}^{N}\frac{1}{2k_{x_i}}\dot{x}_i^2 + \sum_{l=1}^{\bar{L}}\frac{1}{2k_{\bar{p}_l}}\dot{\bar{p}}_l^2 + \sum_{i=1}^{N}\frac{k_{e_i}}{2\hat{k}_{e_i}}\dot{\hat{x}}_i^2 \tag{4.49}$$

where $\bar{p}_l$ denotes the reduced state from the set $\{p_l\}$ by eliminating non-bottleneck links, $k_{\bar{p}_l}$ is the controller gain in the dynamics of $\bar{p}_l$ and $\bar{L}$ is the number of bottleneck links. $V_{(4.48),(4.47b)-(4.47c)}$ may be discontinuous during the transient period since $\{\bar{p}_l\}$ may change.

For a time interval with fixed $\{\bar{p}_l\}$, differentiating $V_{(4.48),(4.47b)-(4.47c)}$ with respect to time, we get

$$\begin{aligned}\dot{V}_{(4.48),(4.47b)-(4.47c)} \leq& \sum_{i=1}^{N}\dot{x}_i U_i''(x_i)\dot{x}_i - \dot{x}^T\mathrm{diag}\{k_i\}\tilde{R}^T\mathrm{diag}\{k_{\tilde{p}_l}\}\tilde{R}\dot{x} - \sum_{i=1}^{N}k_{e_i}(\dot{x}_i - \dot{\hat{x}}_i)^2\\ \leq& 0\end{aligned}$$

where x stands for the vector with entries x_i, $\tilde{R}$ is obtained by first eliminating non-bottleneck rows from R and then eliminating rows whose corresponding $\dot{p}_l$ is

non-positive, and $k_{\tilde{p}_l}$ is the controller gain in the dynamics of $\tilde{p}_l$ which is obtained through the same way as $\tilde{R}$. Note that $\tilde{R}$ may change during the transient period, however, $\dot{V}_{(4.48),(4.47b)-(4.47c)}$ is non-positive all the time. Now $\dot{V}_{(4.48),(4.47b)-(4.47c)} = 0$ only when $\dot{x}_i = \dot{\hat{x}}_i = 0$, i.e., x_i and $\hat{x}_i$ are constant. Substituting them into (4.47b)-(4.47c), we have $x_i = \hat{x}_i$ and each $\dot{\bar{p}}_l$ is constant. Noting that R has full row rank, i.e., RR^T is nonsingular, using (4.48) we can obtain that each $\bar{p}_l$ is constant, i.e., $\dot{\bar{p}}_l = 0$. So $\dot{V}_{(4.48),(4.47b)-(4.47c)} = 0$ only happens at the equilibrium point.

Consider the case when the cardinality of $\{\bar{p}_l\}$ increases at time t. Then an additional term $\dot{\bar{p}}_m$ is added to $V_{(4.48),(4.47b)-(4.47c)}$ that must satisfy $\bar{p}_m(t^-) = 0$, $\dot{\bar{p}}_m(t^-) \leq 0$, $\bar{p}_m(t^+) > 0$ and $\dot{\bar{p}}_m(t^+) > 0$. Since $\dot{\bar{p}}_m$ has zero initial value at time t, $V_{(4.48),(4.47b)-(4.47c)}$ is still continuous in this case. When the cardinality of $\{\bar{p}_l\}$ decreases at time t, there must be a term $\dot{\bar{p}}_m$ satisfying $\bar{p}_m(t^-) > 0$, $\dot{\bar{p}}_m(t^-) \leq 0$, $\bar{p}_m(t^+) = 0$ and $\dot{\bar{p}}_m(t^+) \leq 0$. So $V_{(4.48),(4.47b)-(4.47c)}$ loses a non-negative term and is non-increasing at time t, although now it can be discontinuous. Therefore, $V_{(4.48),(4.47b)-(4.47c)}$ is either non-increasing when there is discontinuity, or decreasing, until the system reaches the equilibrium point. Thus, the global asymptotic stability is ensured. ■

With the PID controllers (4.47) as well as (4.48) and (4.47b)-(4.47c), the block diagram of the system is illustrated in Figure 4.5, where there is extra but limited information relating to $\psi(\dot{p}_l)$ passed from links to sources, while its operation remains distributed. Compared with the standard primal-dual algorithm, extra gains are introduced, i.e., $k_i, k_{e_i}, \hat{k}_{e_i}$. The analysis of linear robustness and delay robustness is similar to that in the previous sections but more complicated. In the next section, we will present an example to illustrate how the extra gains affect the performance and robustness of the closed-loop system.

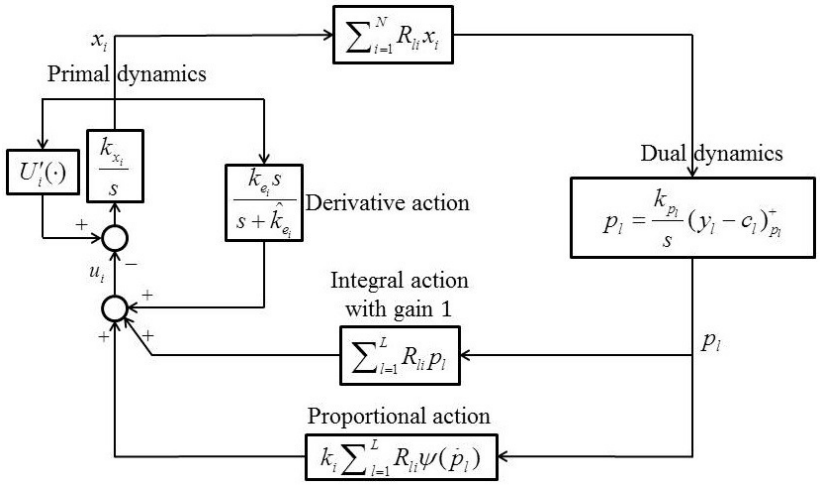

Figure 4.5: The block diagram of the system with the distributed PID controller.

4.6 Numerical Investigations

4.6.1 An Example that Illustrates the Trade-off

Firstly, we present an example to illustrate how the gains of the extra dynamics affect the performance and robustness of the closed-loop system. Here we focus on the redesigned prima-dual algorithm with PID control actions (4.48) and (4.47b)-(4.47c). We continue to use the same network, scenario and parameters corresponding to Figure 4.3. Since the PID controller is completely distributed, we only investigate the gains of the controller for user 2 (in Figure 4.3, this corresponds to the curves on top). As for users 1 and 3, the proportional gains are $k_1 = k_3 = 0.5$.

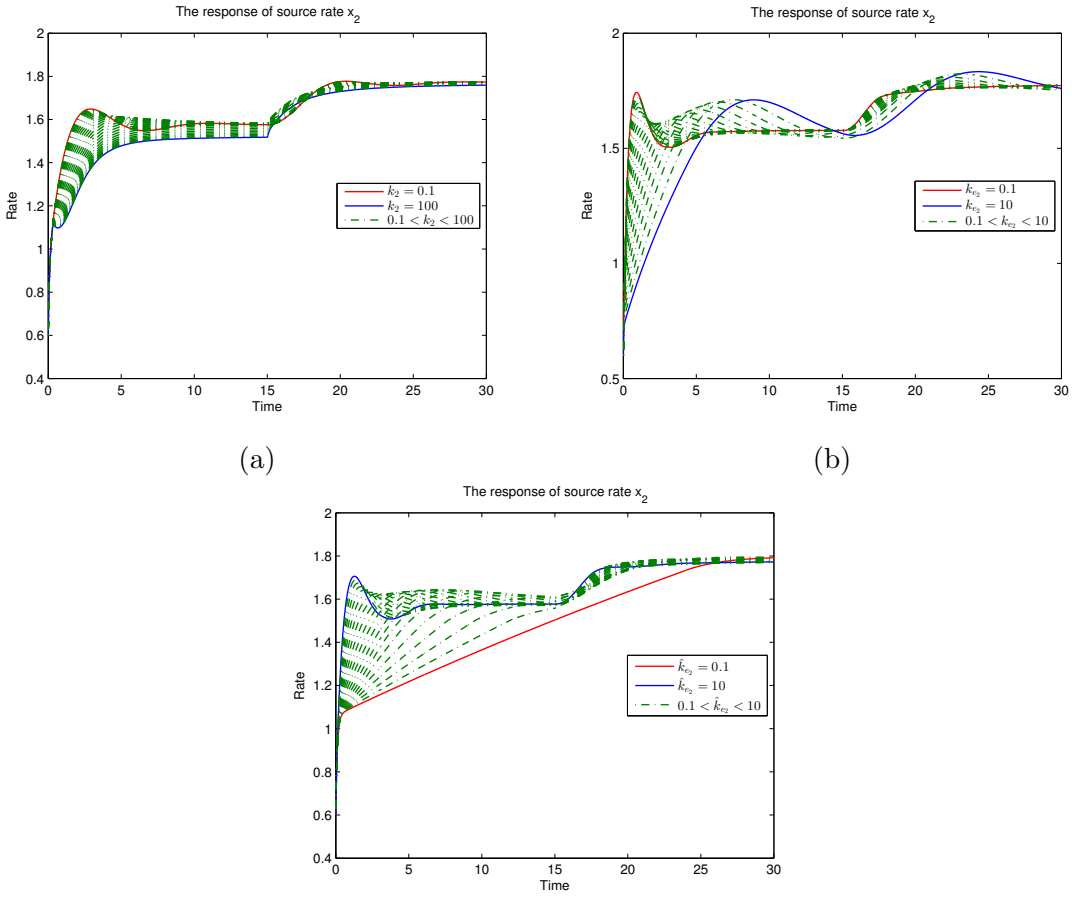

Figure 4.6: The response of x_2 when tuning k_2 ((a)), k_{e_2} ((b)) and $\hat{k}_{e_2}$ ((c)).

We first fix $k_{e_2} = 2$ and $\hat{k}_{e_2} = 2$ and tune the proportional gain k_2 (from 0.1 to 100). Figure 4.6(a) illustrates the change of the response of x_2. As k_2 increases, the response slows down and the system achieves better performance. However, for large

k_2, the system converges too slowly. We then consider the case when $k_2 = 0.5$, $\hat{k}_{e_2} = 2$ and k_{e_2} changes (from 0.1 to 10). The results are exhibited in Figure 4.6(b). We can see that the larger the derivative gain is, the slower the system behaves. When k_{e_2} is not very large or very small, the performance is better balanced out. Finally, let $k_2 = 0.5$, $k_{e_2} = 2$, and tune $\hat{k}_{e_2}$ (from 0.1 to 10). As demonstrated in Figure 4.6(c), small $\hat{k}_{e_2}$ decreases the response speed of x_2 since it results in small cutoff frequency for the filtering part in the derivative control. Note that increasing $\hat{k}_{e_2}$ too much leads to weaker filtering function. To conclude, in order to improve the performance of the closed-loop system, there are trade-offs for these gains.

4.6.2 An Example of Jointly Optimal Congestion and Contention Control

In the second example, we focus on the jointly optimal congestion and contention control problem, as a further application of the proposed redesign framework. Consider the following problem [96]

$$\max_{x_i \geq 0, p_l \geq 0} \quad \sum_{i=1}^{N} U_i(x_i) \tag{4.50a}$$

$$\text{subject to} \quad \sum_{i=1}^{N} R_{li} x_i \leq c_l p_l \prod_{m \in N_{to}^I(l)} \left(1 - \sum_{r \in L_{out}(m)} p_r\right), \ l = 1, \cdots, L \tag{4.50b}$$

$$\sum_{r \in L_{out}(n)} p_r \leq 1, \ n = 1, \cdots, N \tag{4.50c}$$

where p_l is the persistence probability of link l controlled by contention-based Medium Access Control (MAC) (note that p_l is not the link price any more), c_l is the physical layer transmission rate of link l which is a constant, $N_{to}^I(l)$ is the set of nodes whose transmissions cause interference to the receiver node of link l, excluding the transmitter node of link l denoted by t_l, $L_{out}(n)$ is the set of outgoing links of node n, and the rest of notation remains the same as in (4.1). In addition, let $L_{from}^I(n)$ denote the set of links whose transmissions get interfered from the transmission of node n, excluding the outgoing link l. For simplicity, we have removed the constraints of maximum and minimum source rate for each x_i. Note that here R is a routing matrix of a given logical topology, since this problem is mainly concerned with ad-hoc networks. Compared with (4.1), each link capacity is now a nonlinear function of relevant persistence probabilities.

Generally speaking, the above problem is non-convex and non-separable. Taking the logarithm of each variables x_i and every link congestion constraint we can convexify it, provided that each utility function satisfies $U_i''(x_i)x_i + U_i'(x_i) \leq 0$ [96]. Define $y_i = \log x_i$, then the convexified version of the problem (4.50) is formulated as

$$\max_{p_l>0} \quad \sum_{i=1}^{N} U_i(e^{y_i}) \tag{4.51a}$$

$$\text{subject to} \quad \log\left(\sum_{i=1}^{N} R_{li}e^{y_i}\right) \leq \log c_l + \log p_l + \sum_{m\in N_{to}^I(l)} \log\left(1 - \sum_{r\in L_{out}(m)} p_r\right) \quad l = 1, \cdots, L \tag{4.51b}$$

$$\sum_{r\in L_{out}(n)} p_r \leq 1, \ n = 1, \cdots, N. \tag{4.51c}$$

We can now apply the proposed redesign framework to solve the problem (4.51), through which (4.50) can be solved. The modified dynamics are given by

$$\dot{y}_i = k_{y_i}\left(U_i'(e^{y_i})e^{y_i} - e^{y_i}\sum_{l=1}^{N} R_{li}\frac{\lambda_l + k_l\psi(\dot{\lambda}_l)}{\sum_{n=1}^{N} R_{ln}e^{y_n}} - k_{ey_i}(y_i - \hat{y}_i)\right) \tag{4.52a}$$

$$\dot{\hat{y}}_i = \hat{k}_{ey_i}(y_i - \hat{y}_i) \tag{4.52b}$$

$$\dot{p}_l = k_{p_l}\left(\frac{\lambda_l + k_l\psi(\dot{\lambda}_l)}{p_l} - \frac{\sum_{m\in L_{from}^I(t_l)}(\lambda_m + k_m\psi(\dot{\lambda}_m))}{1 - \sum_{r\in L_{out}(t_l)} p_r} - \mu_{t_l} - k_{ep_l}(p_l - \hat{p}_l)\right) \tag{4.52c}$$

$$\dot{\hat{p}}_l = \hat{k}_{ep_l}(p_l - \hat{p}_l) \tag{4.52d}$$

$$\dot{\lambda}_l = k_{\lambda_l}\left(\log\left(\sum_{i=1}^{N} R_{li}e^{y_i}\right) - \log c_l - \log p_l - \sum_{m\in N_{to}^I(l)} \log\left(1 - \sum_{r\in L_{out}(m)} p_r\right)\right)_{\lambda_l}^{+} \tag{4.52e}$$

$$\dot{\mu}_n = k_{\mu_n}\left(\sum_{r\in L_{out}(n)} p_r - 1\right)_{\mu_n}^{+} \tag{4.52f}$$

where $k_{y_i}, k_l, k_{ey_i}, \hat{k}_{ey_i}, k_{p_l}, k_{ep_l}, \hat{k}_{ep_l}, k_{\lambda_l}, k_{\mu_n} > 0$ are the controller gains, λ_l is the price relating to congestion for link l, μ_n is the price relating to the sum of persistence probability for node n. It is clear that the implementation of (4.52) is completely distributed, and moreover, with PID control actions.

We next use a 1-source-2-link-3-node network give in Figure 4.7 to illustrate the performance of the proposed controller. Here the transmission of node 2 will cause interference to link 1. The parameters of the overall system are: $U_1(x_1) = \log(x_1), k_{y_1} =$

$k_{p_1} = k_{p_2} = 3, k_1 = k_2 = 1, k_{ey_1} = k_{ep_2} = 1, k_{ep_1} = 2, \hat{k}_{ey_1} = \hat{k}_{ep_2} = 2, \hat{k}_{ep_1} = 1, k_{\lambda_1} = k_{\lambda_2} = 4, k_{\mu_1} = k_{\mu_2} = 2$. We consider a scenario involving a sudden drop of transmission rate in both links, from 11 Mbps to 8 Mbps. The simulation results are shown in Figure 4.8, using only integral control, Integral and Derivative (ID) control with filtering, and PID control respectively. It is clear that with the distributed PID controller, the transient performance of the closed-loop system is much better than that under only integral control and that under ID control with filtering.

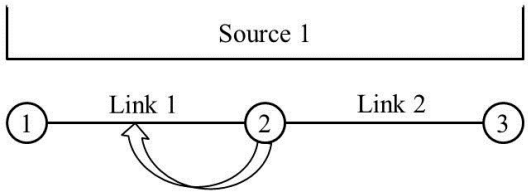

Figure 4.7: A 1-source-2-link-3-node network.

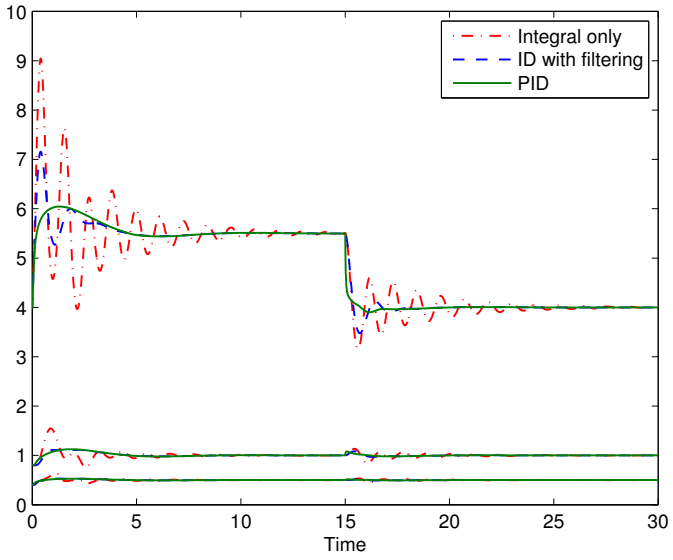

Figure 4.8: Simulation results of the 1-source-2-link-3-node network. The top curves are the responses of the source rate. The curves in the middle are the responses of the persistence probability of link 1. The curves in the bottom are the responses of the persistence probability of link 2.

4.6.3 An Example that Illustrates the Improved Delay Robustness

Finally, to illustrate the improved delay robustness of the redesigned schemes, we consider the redesigned primal-dual algorithm (4.11) and use a network with 3 links

and 4 sources as given in Figure 4.9. The parameters of the overall system are: $U_i(x_i) = \log(x_i), k_{x_i} = 0.1, k_{e_i} = 100, \hat{k}_{e_i} = 1.7, \forall i, k_{p_l} = 100, \forall l$. The scenario here consists of source 3 suddenly changing its utility function at $t = 150$s, from $\log(x_3)$ to $3\log(x_3)$, and there is a 0.25s-delay in the price feedback of link 3. The simulation results are shown in Figure 4.10. In addition, the results of using the standard primal-dual dynamics and the modified primal-dual dynamics with Strictly Positive Real (SPR) filter [53], i.e.,

$$\begin{aligned}\dot{\xi}_i &= A_i\xi_i + B_i(U_i'(x_i) - q_i)\\ \dot{x}_i &= C_i\xi_i + D_i(U_i'(x_i) - q_i)\\ \dot{p}_l &= k_{p_l}(y_l - c_l)^+_{p_l}\end{aligned}$$

are provided under the same condition ($A_i = -20, B_i = 1, C_i = -1.9, D_i = 0.1, \forall i, k_{p_l} = 100, \forall l$). It is clear that the controller behaves much better after redesign in the presence of time delay. Moreover, the redesigned dynamics converge faster than the modified primal-dual dynamics with SPR filter in this case.

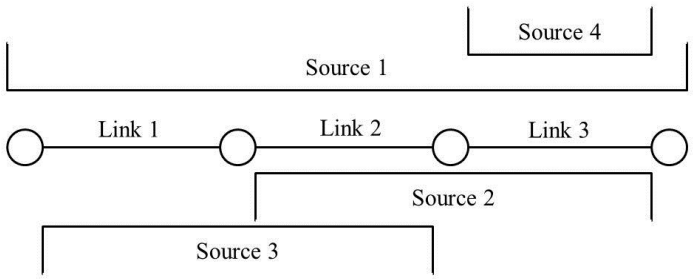

Figure 4.9: A 4-source-3-link network.

4.7 Conclusion

In this chapter, we have presented a redesign framework for network congestion control problems at the level of fluid-flow models. Motivated by the augmented Lagrangian method, we introduced extra terms to the Lagrangian, which was then used to redesign the primal-dual, primal and dual algorithms. We investigated how the gains resulting from the extra dynamics influenced the stability and robustness of the system. Moreover, we showed that the overall system could achieve added robustness to communication delays by appropriately tuning these gains. In Section 4.5, we studied the meaning of the extra dynamics in the redesign and further developed

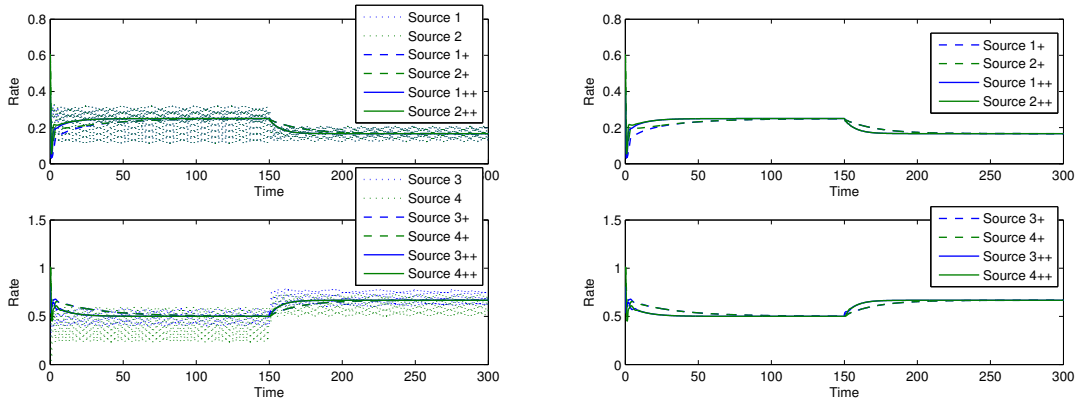

Figure 4.10: Simulation results of the 3-link-4-source network under 0.25s delay in the price feedback of link 3, using the standard primal-dual dynamics, the modified primal-dual dynamics with SPR filter (marked with "+"), and the redesigned dynamics (marked with "++").

a distributed PID controller for solving the network congestion control problem. Finally, we illustrated that compared to standard algorithms, the modified algorithms resulted in better transient performance and improved robustness for the overall system, without changing the distributed structure of the feedback.

Chapter 5

Distributed Optimal Steady-state Control Using Reverse- and Forward-engineering

In the previous chapter, we have presented a redesign framework for network congestion control problems at the level of fluid-flow models. The framework is based on an augmented Lagrangian method in which we introduce extra dynamics to algorithms resulting from traditional optimization decomposition methods. Such an approach actually follows a "forward-engineering" procedure [51]. In this chapter, we go back to the problem of merging primary and secondary frequency control in power systems. We will use a reverse- and forward-engineering approach to design distributed control schemes for generators and controllable loads to track the optimal solution of a predefined optimization problem automatically, i.e., to achieve real-time economic dispatch, in the presence of disturbances. Unlike Chapter 2, the optimization problem considered here is more general and includes capacity constraints for each component in the network.

In addition to power systems, we propose a reverse- and forward-engineering design framework for linear network systems to achieve optimal steady-state performance, which consists of two steps. Firstly, we reverse-engineer a dynamic system as a gradient algorithm to solve an optimization problem. Secondly, we use a forward-engineering approach to systematically design distributed control or modify the existing control. As a result, the system can automatically track the optimal solution of a predefined optimization problem and the control scheme can be implemented in a distributed (i.e., communicating with neighbours) and closed-loop (i.e., no information of disturbance is needed) manner. In order to investigate how general this framework is, we establish necessary and sufficient conditions under which a linear

dynamic system can be reverse-engineered as a gradient algorithm to solve an optimization problem. These conditions are characterized using properties of system matrices and relevant linear matrix inequalities. Finally, a practical example using the IEEE 14-bus network is given to demonstrate the effectiveness of the framework.

5.1 Achieving Real-time Economic Dispatch in Power Networks

In this section, we focus on the problem of designing distributed control schemes for power networks to achieve real-time economic dispatch.

5.1.1 Problem Setup

The network configuration was previously described in Section 2.3. For conciseness, we directly provide the system model (i.e., system (2.24)):

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \Gamma_1 A T A^T \alpha \tag{5.1a}$$

$$\dot{\alpha} = \Gamma_1^T \omega_g + \Gamma_2^T \omega_l \tag{5.1b}$$

$$\omega_l = D_l^{-1}(-P_L - d_l - \Gamma_2 A T A^T \alpha) \tag{5.1c}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega_g). \tag{5.1d}$$

Suppose that the network is operating around a nominal state determined by an Economic Dispatch (ED) problem at a more slower time-scale, and that *constant* disturbances d_g, d_l occur. As a result, the frequency deviates from the nominal value. It is expected that generators and controllable loads can adjust their power generation P_C and consumption P_L respectively in real time using *local* information, not only to release the overall system from stress, i.e., to restore the grid frequency, but also to effect a better supply-demand balance, e.g., minimizing generation cost, maximizing load utility, and satisfying transmission line congestion constraints. These goals lead to a steady-state optimization/DC Optimal Power Flow (OPF) problem:

$$\min_{P_M, P_L, \alpha} \quad \sum_{i\in\mathcal{G}} C_i(P_{M_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) \tag{5.2a}$$

$$\text{subject to} \quad P_M - d_g - \Gamma_1 A T A^T \alpha = \mathbf{0} \tag{5.2b}$$

$$- P_L - d_l - \Gamma_2 A T A^T \alpha = \mathbf{0} \tag{5.2c}$$

$$P_C^{\min} \preceq P_M \preceq P_C^{\max} \tag{5.2d}$$

$$P_L^{\min} \preceq P_L \preceq P_L^{\max} \tag{5.2e}$$

$$P_{TC}^{\min} \preceq T A^T \alpha \preceq P_{TC}^{\max} \tag{5.2f}$$

where $C_i(P_{M_i})$ is the cost function for each generator, $U_i(P_{L_i})$ is the utility function for each controllable load, Equations (5.2b)-(5.2c) represent power flow balance at each bus, Equations (5.2d)-(5.2f) are capacity constraints for generators, loads, and transmission lines respectively, and $P_C^{\min}, P_C^{\max} \in \mathbb{R}^m$, $P_L^{\min}, P_L^{\max} \in \mathbb{R}^n$, $P_{TC}^{\min}, P_{TC}^{\max} \in \mathbb{R}^p$ are corresponding capacity vectors. Note that by setting $P_{TC_{ij}}^{\max} = P_{TC_{ij}}^{\min}$, the scheduled power flow in line (i, j) can be maintained. For problem (5.2), we have a related assumption, which is extensively used in the literature [13, 24].

Assumption 5.1 *Problem (5.2) is feasible. Moreover, each C_i is a strictly convex function in P_{M_i} and each U_i is a strictly concave function in P_{L_i}.*

Compared with similar Assumptions 2.2 and 2.4 in Chapter 2, and Assumption 3.3 in Chapter 3, this assumption is more general in which we do not require particular forms of cost and utility functions. Since the dynamics of P_M and α are constrained by (5.1) during the transient, they therefore cannot be instantaneously set to the solution of problem (5.2). Note that Equations (5.2b)-(5.2c) require the frequency to be restored when system (5.1) reaches steady state, i.e., $\omega_i^* = 0, i \in \mathcal{G} \bigcup \mathcal{L}$. According to Remark 2.6, this results in $P_M = P_C$ in steady state. So we can reformulate problem (5.2) as one that can be used to design P_C and P_L, given by

$$\min_{P_C, P_L, \theta} \quad \sum_{i \in \mathcal{G}} C_i(P_{C_i}) - \sum_{i \in \mathcal{L}} U_i(P_{L_i}) \tag{5.3a}$$

$$\text{subject to} \quad P_C - d_g - \Gamma_1 A T A_0^T \theta = \mathbf{0} \tag{5.3b}$$

$$- P_L - d_l - \Gamma_2 A T A_0^T \theta = \mathbf{0} \tag{5.3c}$$

$$P_C^{\min} \preceq P_C \preceq P_C^{\max} \tag{5.3d}$$

$$P_L^{\min} \preceq P_L \preceq P_L^{\max} \tag{5.3e}$$

$$P_{TC}^{\min} \preceq T A_0^T \theta \preceq P_{TC}^{\max} \tag{5.3f}$$

where $\theta = [\theta_1, \cdots, \theta_{m+n}]^T$ is the vector of ancillary decision variables. In fact, we have replaced P_M, α in (5.2) by P_C, θ to derive (5.3).

Proposition 5.1 *Given the control input P_C^* and P_L^* to system (5.1), where (P_C^*, P_L^*, θ^*) is an optimal solution of (5.3), the equilibrium point of system (5.1) satisfies $P_M^* = P_C^*$ and $\alpha^* = [\Gamma_1^T, \Gamma_2^T]\theta^*$. As a result, (P_M^*, P_L^*, α^*) is an optimal solution of problem (5.2).*

Proof: The proof follows Remark 2.6 immediately. ■

According to the above proposition, the optimality of problem (5.2) is preserved after the reformulation. Finally, the real-time economic dispatch problem is described as: design P_C and P_L so that for given d_g, d_l, system (5.1) is driven to an equilibrium point where the steady-state optimization problem (5.2)/(5.3) is solved.

5.1.2 A Reverse- and Forward-engineering Design Framework

As the first step, we reverse-engineer system (5.1) as one with primal-dual gradient dynamics [56, 70] (see the Appendix) to solve a saddle point problem. The resulting saddle point problem will allow us to embed the steady-state optimization problem for control design. Two lemmas relating to properties of a saddle point of a function and primal-dual gradient dynamics are now presented.

Lemma 5.1 *Let $f \in \mathcal{C}^2$: $\mathbb{R}^a \times \mathbb{R}^b \to \mathbb{R}$ satisfy: for all y, z, $\nabla_y^2 f \succeq 0, \nabla_z^2 f \preceq 0$. Then $(\tilde{y}, \tilde{z})$ is a saddle point of f, i.e., $f(\tilde{y}, z) \leq f(\tilde{y}, \tilde{z}) \leq f(y, \tilde{z})$, if and only if $\nabla_{y,z} f|_{y=\tilde{y}, z=\tilde{z}} = \mathbf{0}$.*

Proof: This Lemma is a corollary of the Karush-Kuhn-Tucker (KKT) conditions, the proof of which can be found in the Appendix. ■

Lemma 5.2 *Let $f \in \mathcal{C}^2$: $\mathbb{R}^a \times \mathbb{R}^b \to \mathbb{R}$ satisfy: for all y, z, $\nabla_y^2 f \succeq 0, \nabla_z^2 f \preceq 0$, and the set $\{(y, z) | \nabla_{y,z} f = \mathbf{0}\}$ is nonempty. Then the trajectories of the primal-dual gradient dynamics [56]/saddle point dynamics [70] given by*

$$\dot{y} = -K_y \frac{\partial f}{\partial y} \tag{5.4a}$$

$$\dot{z} = K_z \frac{\partial f}{\partial z} \tag{5.4b}$$

are bounded, where $K_y \in \mathbb{R}^{a \times a}, K_z \in \mathbb{R}^{b \times b}$ are positive definite constant matrices. Furthermore, if f is either strictly convex in y or strictly concave in z, then the trajectories of (5.4) asymptotically converge to a saddle point of f.

Proof: The proof follows the content of Section 2 in [56] and Proposition 13 in [70], which can be found in the Appendix. ■

Now consider a candidate Lagrangian $L_{(5.1)}(\alpha, P_M, \omega_g | P_C, P_L, d_g, d_l) \in \mathcal{C}^2 : \mathbb{R}^{m+n-1} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ given by

$$\begin{aligned} L_{(5.1)} = &-\frac{1}{2}(D_g\omega_g - P_M + d_g + \Gamma_1 A T A^T \alpha)^T D_g^{-1}(D_g\omega_g - P_M + d_g + \Gamma_1 A T A^T \alpha) \\ &+ \frac{1}{2}(P_M - d_g - \Gamma_1 A T A^T \alpha)^T D_g^{-1}(P_M - d_g - \Gamma_1 A T A^T \alpha) \\ &+ \frac{1}{2}(P_L + d_l + \Gamma_2 A T A^T \alpha)^T D_l^{-1}(P_L + d_l + \Gamma_2 A T A^T \alpha) \\ &+ \frac{1}{2}(P_M - P_C)^T R (P_M - P_C) \end{aligned} \tag{5.5}$$

where $\alpha \in \mathbb{R}^{m+n-1}, P_M \in \mathbb{R}^m, \omega_g \in \mathbb{R}^m$ are decision variables, and $P_C, d_g \in \mathbb{R}^m, P_L, d_l \in \mathbb{R}^n$ are constant. Note that for simplicity, we have used the same notation as in the dynamics (5.1). Since $\nabla^2_{\alpha,P_M} L_{(5.1)} = \text{diag}\{ATA^T\Gamma_2^T D_l^{-1}\Gamma_2 ATA^T, R\} \succeq 0, \nabla^2_{\omega_g} L_{(5.1)} = -D_g \prec 0$ are true, $L_{(5.1)}$ is convex in α, P_M and strictly concave in ω_g. We then formulate a saddle point problem as

$$\min_{\alpha,P_M} \max_{\omega_g} L_{(5.1)}. \tag{5.6}$$

Through straightforward derivation, we can show that system dynamics (5.1) is actually the saddle point dynamics (5.4) of $L_{(5.1)}$, due to the following equalities:

$$\begin{aligned}
\frac{\partial L_{(5.1)}}{\partial \omega_g} &= M\dot{\omega}_g \ (\dot{\omega}_g \text{ is given in (5.1a)}) \\
\frac{\partial L_{(5.1)}}{\partial \alpha} &= -ATA^T\dot{\alpha} \ (\ \dot{\alpha} \text{ is given in (5.1b)}) \\
\frac{\partial L_{(5.1)}}{\partial P_M} &= -T_{TG}R\dot{P}_M \ (\dot{P}_M \text{ is given in (5.1d)}).
\end{aligned}$$

We then have the following theorem.

Theorem 5.1 *The trajectories of system (5.1) asymptotically converge to a saddle point of problem (5.6).*

Proof: The proof follows Lemmas 5.1-5.2 immediately. ■

As for problem (5.3), under Assumption 5.1, strong duality holds. Formulate the Lagrangian of (5.3) as

$$\begin{aligned}
&L_{(5.3)}(P_C, P_L, \theta, \zeta, \lambda, \mu^+, \mu^-, \nu^+, \nu^-, l^+, l^-) = \\
&\sum_{i\in\mathcal{G}} C_i(P_{C_i}) - \sum_{i\in\mathcal{L}} U_i(P_{L_i}) + \zeta^T(P_C - d_g - \Gamma_1 ATA_0^T\theta) + \lambda^T(-P_L - d_l - \Gamma_2 ATA_0^T\theta) \\
&+ \mu^{+T}(P_C - P_C^{\max}) + \mu^{-T}(P_C^{\min} - P_C) + \nu^{+T}(P_L - P_L^{\max}) + \nu^{-T}(P_L^{\min} - P_L) \\
&+ l^{+T}(TA_0^T\theta - P_{TC}^{\max}) + l^{-T}(P_{TC}^{\min} - TA_0^T\theta)
\end{aligned} \tag{5.7}$$

where $\zeta, \mu^+, \mu^- \in \mathbb{R}^m$, $\lambda, \nu^+, \nu^- \in \mathbb{R}^n$ and $l^+, l^- \in \mathbb{R}^p$ are Lagrange multipliers (dual vectors) for the constraints in (5.3). Then we obtain the following saddle point problem:

$$\min_{P_C,P_L,\theta\ \mu^+,\mu^-,\nu^+,\nu^-,l^+,l^-\succeq 0,\zeta,\lambda} \max L_{(5.3)}. \tag{5.8}$$

As a result, solving problem (5.3) is equivalent to solving problem (5.8). Consider the augmented saddle point problem given by

$$\min_{\alpha,P_M,P_C,P_L,\theta\ \mu^+,\mu^-,\nu^+,\nu^-,l^+,l^-\succeq 0,\zeta,\lambda,\omega_g} \max L_{au} = L_{(5.1)} + \gamma L_{(5.3)} \tag{5.9}$$

where $\gamma > 0$ is constant. We have a related proposition regarding the saddle point problems (5.6), (5.8) and (5.9).

Proposition 5.2 *If* $(\omega_g^*, \alpha^*, P_M^*, P_C^*, P_L^*, \theta^*, \zeta^*, \lambda^*, \mu^{+*}, \mu^{-*}, \nu^{+*}, \nu^{-*}, l^{+*}, l^{-*})$ *is a saddle point of* L_{au}*, then* $(\omega_g^*, \alpha^*, P_M^*)$ *is a saddle point of* $L_{(5.1)}$ *and* $(P_C^*, P_L^*, \theta^*, \zeta^*, \lambda^*, \mu^{+*}, \mu^{-*}, \nu^{+*}, \nu^{-*}, l^{+*}, l^{-*})$ *is a saddle point of* $L_{(5.3)}$.

Proof: According to Lemma 5.1, any saddle point $(\omega_g^*, \alpha^*, P_M^*, P_C^*, P_L^*, \theta^*, \zeta^*, \lambda^*, \mu^{+*}, \mu^{-*}, \nu^{+*}, \nu^{-*}, l^{+*}, l^{-*})$ of L_{au} satisfies (here $C(P_C) \in \mathbb{R}^m$ is the vector of cost functions of generators whose entries are $C_i(P_{C_i}(t))$, $U(P_L) \in \mathbb{R}^n$ is the vector of utility functions of loads whose entries are $U_i(P_{L_i}(t))$ in the following equations):

$$-D_g\omega_g^* + P_M^* - d_g - \Gamma_1 A T A^T \alpha^* = \mathbf{0} \tag{5.10a}$$

$$\Gamma_1^T \omega_g^* + \Gamma_2^T D_l^{-1}(-P_L^* - d_l - \Gamma_2 A T A^T \alpha^*) = \mathbf{0} \tag{5.10b}$$

$$P_C^* - P_M^* - R^{-1}\omega_g^* = \mathbf{0} \tag{5.10c}$$

$$R(P_M^* - P_C^*) - \gamma(C'(P_C^*) + \zeta^* + \mu^{+*} - \mu^{-*}) = \mathbf{0} \tag{5.10d}$$

$$D_l^{-1}(-P_L^* - d_l - \Gamma_2 A T A^T \alpha^*) + \gamma(U'(P_L^*) + \lambda^* - \nu^{+*} + \nu^{-*}) = \mathbf{0} \tag{5.10e}$$

$$A_0 T(A^T \Gamma_1^T \zeta^* + A^T \Gamma_2^T \lambda^* - l^{+*} + l^{-*}) = \mathbf{0} \tag{5.10f}$$

$$P_C^* - d_g - \Gamma_1 A T A_0^T \theta^* = \mathbf{0} \tag{5.10g}$$

$$-P_L^* - d_l - \Gamma_2 A T A_0^T \theta^* = \mathbf{0} \tag{5.10h}$$

$$\mathrm{diag}(\mu^{+*})(P_C^* - P_C^{\max}) = \mathbf{0},\ \mu^{+*}, P_C^{\max} - P_C^* \succeq 0 \tag{5.10i}$$

$$\mathrm{diag}(\mu^{-*})(P_C^{\min} - P_C^*) = \mathbf{0},\ \mu^{-*}, P_C^* - P_C^{\min} \succeq 0 \tag{5.10j}$$

$$\mathrm{diag}(\nu^{+*})(P_L^* - P_L^{\max}) = \mathbf{0},\ \nu^{+*}, P_L^{\max} - P_L^* \succeq 0 \tag{5.10k}$$

$$\mathrm{diag}(\nu^{-*})(P_L^{\min} - P_L^*) = \mathbf{0},\ \nu^{-*}, P_L^* - P_L^{\min} \succeq 0 \tag{5.10l}$$

$$\mathrm{diag}(l^{+*})(T A_0^T \theta^* - P_{TC}^{\max}) = \mathbf{0},\ l^{+*}, P_{TC}^{\max} - T A_0^T \theta^* \succeq 0 \tag{5.10m}$$

$$\mathrm{diag}(l^{-*})(P_{TC}^{\min} - T A_0^T \theta^*) = \mathbf{0},\ l^{-*}, T A_0^T \theta^* - P_{TC}^{\min} \succeq 0. \tag{5.10n}$$

Equations (5.10g)-(5.10h) result in $\mathbf{1}(P_C^* - d_g) - \mathbf{1}(P_L^* + d_l) = 0$. Based on Remark 2.6, we have $\omega_i^* = 0, i \in \mathcal{G} \bigcup \mathcal{L}$ and $P_C^* = P_M^*$. Substituting these into Equations (5.10d)-(5.10e) leads to

$$C'(P_C^*) + \zeta^* + \mu^{+*} - \mu^{-*} = \mathbf{0}$$
$$U'(P_L^*) + \lambda^* - \nu^{+*} + \nu^{-*} = \mathbf{0}.$$

According to Lemma 5.1, $(\omega_g^*, \alpha^*, P_M^*)$ is a saddle point of $L_{(5.1)}$ and $(P_C^*, P_L^*, \theta^*, \zeta^*, \lambda^*, \mu^{+*}, \mu^{-*}, \nu^{+*}, \nu^{-*}, l^{+*}, l^{-*})$ is a saddle point of $L_{(5.3)}$. ■

Since $\nabla^2_{\alpha,P_M,P_C,P_L,\theta} L_{au} \succeq 0$ and $\nabla^2_{\mu^+,\mu^-,\nu^+,\nu^-,l^+,l^-,\zeta,\lambda,\omega_g} L_{au} \preceq 0$ hold, under Lemma 5.2, the trajectories of the primal-dual gradient dynamics given by Equation (5.11) in the sequel are bounded. Note that the projection in (5.11i)-(5.11n) does not affect the result in Lemma 5.2 [56]. In system (5.11), $K_{P_C}, K_\zeta, K_{\mu^+}, K_{\mu^-} \in \mathbb{R}^{m\times m}$, $K_{P_L}, K_\lambda, K_{\nu^+}, K_{\nu^-} \in \mathbb{R}^{n\times n}$, $K_\theta \in \mathbb{R}^{(m+n)\times(m+n)}$, and $K_{l^+}, K_{l^-} \in \mathbb{R}^{p\times p}$ are positive diagonal matrices, all representing the controller gains. Also for simplicity, we have used vector forms of positive projection in (5.11i)-(5.11n).

Power network dynamics:

$$M\dot{\omega}_g + D_g\omega_g = P_M - d_g - \Gamma_1 A T A^T \alpha \tag{5.11a}$$

$$\dot{\alpha} = \Gamma_1^T \omega_g + \Gamma_2^T D_l^{-1}(-P_L - d_l - \Gamma_2 A T A^T \alpha) \tag{5.11b}$$

$$\dot{P}_M = T_{TG}^{-1}(P_C - P_M - R^{-1}\omega_g) \tag{5.11c}$$

Control input dynamics:

$$\dot{P}_C = K_{P_C}(R(P_M - P_C) - \gamma(C'(P_C) + \zeta + \mu^+ - \mu^-)) \tag{5.11d}$$

$$\dot{P}_L = K_{P_L}(D_l^{-1}(-P_L - d_l - \Gamma_2 A T A^T \alpha) + \gamma(U'(P_L) + \lambda - \nu^+ + \nu^-)) \tag{5.11e}$$

Ancillary variable dynamics:

$$\dot{\theta} = K_\theta A_0 T(A^T \Gamma_1^T \zeta + A^T \Gamma_2^T \lambda - l^+ + l^-) \tag{5.11f}$$

$$\dot{\zeta} = K_\zeta(P_C - d_g - \Gamma_1 A T A_0^T \theta) \tag{5.11g}$$

$$\dot{\lambda} = K_\lambda(-P_L - d_l - \Gamma_2 A T A_0^T \theta) \tag{5.11h}$$

$$\dot{\mu}^+ = K_{\mu^+}(P_C - P_C^{\max})^+_{\mu^+} \tag{5.11i}$$

$$\dot{\mu}^- = K_{\mu^-}(P_C^{\min} - P_C)^+_{\mu^-} \tag{5.11j}$$

$$\dot{\nu}^+ = K_{\nu^+}(P_L - P_L^{\max})^+_{\nu^+} \tag{5.11k}$$

$$\dot{\nu}^- = K_{\nu^-}(P_L^{\min} - P_L)^+_{\nu^-} \tag{5.11l}$$

$$\dot{l}^+ = K_{l^+}(T A_0^T \theta - P_{TC}^{\max})^+_{l^+} \tag{5.11m}$$

$$\dot{l}^- = K_{l^-}(P_{TC}^{\min} - T A_0^T \theta)^+_{l^-}. \tag{5.11n}$$

In Equations (5.11g)-(5.11h), the information of d_g, d_l is needed. Since the disturbance injection is usually uncertain and/or hard to measure, we modify these two equations so that the implementation of the above controller is independent of d_g, d_l:

$$\dot{\zeta} = K_\zeta(M\dot{\omega}_g + D_g\omega_g + P_C - P_M + \Gamma_1 A T A^T \alpha - \Gamma_1 A T A_0^T \theta) \tag{5.12a}$$

$$\dot{\lambda} = K_\lambda(D_l\omega_l + \Gamma_2 A T A^T \alpha - \Gamma_2 A T A_0^T \theta) \tag{5.12b}$$

where we have substituted system dynamics (5.1) into (5.11g)-(5.11h).

It is important to note that controller (5.11d)-(5.11n) is completely distributed, i.e., states are updated using only local information and signals from their neighborhood.

Theorem 5.2 *Under Assumption 5.1, the trajectories of system (5.11) asymptotically converge to a saddle point of L_{au}, denoted by $(\omega_g^*, \alpha^*, P_M^*, P_C^*, P_L^*, \theta^*, \zeta^*, \lambda^*, \mu^{+*}, \mu^{-*}, \nu^{+*}, \nu^{-*}, l^{+*}, l^{-*})$. Moreover, (P_M^*, P_L^*, α^*) is an optimal solution of problem (5.2).*

Proof: Define a candidate Lyapunov function for the overall system as

$$
\begin{aligned}
&V_{(5.11)} = \\
&\frac{1}{2}\omega_g^T M\omega_g + \frac{1}{2}(\alpha-\alpha^*)^T ATA^T(\alpha-\alpha^*) + \frac{1}{2}(P_M - P_M^*)^T T_{TG} R(P_M - P_M^*) \\
&+ \frac{1}{2}(P_C - P_C^*)^T K_{P_C}^{-1}(P_C - P_C^*) + \frac{1}{2}(P_L - P_L^*)^T K_{P_L}^{-1}(P_L - P_L^*) \\
&+ \frac{\gamma}{2}(\theta-\theta^*)^T K_\theta^{-1}(\theta-\theta^*) + \frac{\gamma}{2}(\zeta-\zeta^*)^T K_\zeta^{-1}(\zeta-\zeta^*) + \frac{\gamma}{2}(\lambda-\lambda^*)^T K_\lambda^{-1}(\lambda-\lambda^*) \\
&+ \frac{\gamma}{2}(\mu^+-\mu^{+*})^T K_{\mu^+}^{-1}(\mu^+-\mu^{+*}) + \frac{\gamma}{2}(\mu^--\mu^{-*})^T K_{\mu^-}^{-1}(\mu^--\mu^{-*}) \\
&+ \frac{\gamma}{2}(\nu^+-\nu^{+*})^T K_{\nu^+}^{-1}(\nu^+-\nu^{+*}) + \frac{\gamma}{2}(\nu^--\nu^{-*})^T K_{\nu^-}^{-1}(\nu^--\nu^{-*}) \\
&+ \frac{\gamma}{2}(l^+-l^{+*})^T K_{l^+}^{-1}(l^+-l^{+*}) + \frac{\gamma}{2}(l^--l^{-*})^T K_{l^-}^{-1}(l^--l^{-*}).
\end{aligned} \tag{5.13}
$$

According to the proof of Theorem 3.3, we can remove all positive projection in $\dot{V}_{(5.11)}$ to get

$$
\begin{aligned}
\dot{V}_{(5.11)} \leq & -(P_L - P_L^* + \Gamma_2 ATA^T(\alpha-\alpha^*))^T D_l^{-1}(P_L - P_L^* + \Gamma_2 ATA^T(\alpha-\alpha^*)) \\
& - \omega_g^T D_g \omega_g - (P_M - P_C)^T R(P_M - P_C) - \gamma(P_C - P_C^*)^T(C'(P_C) - C'(P_C^*)) \\
& + \gamma(P_L - P_L^*)^T(U'(P_L) - U'(P_L^*)) \\
\leq & 0.
\end{aligned} \tag{5.14}
$$

Moreover, $\dot{V}_{(5.11)} = 0$ leads to $\omega_i = 0, i \in \mathcal{G} \bigcup \mathcal{L}$, $P_M = P_C = P_C^* = P_M^*$, $P_L = P_L^*$, and $\mu^+, \mu^-, \nu^+, \nu^-$ are all constant. We then obtain $\dot{\omega}_g = \dot{P}_M = \dot{P}_C = \mathbf{0}$, $\dot{\alpha} = \mathbf{0}$ and $\dot{P}_L = \mathbf{0}$, which result in $\alpha = \alpha^*$ based on Remark 2.6, and ζ, λ are constant. Therefore, we have

$$
\begin{aligned}
P_C^* - d_g - \Gamma_1 ATA_0^T\theta = \mathbf{0} \\
-P_L^* - d_l - \Gamma_2 ATA_0^T\theta = \mathbf{0}.
\end{aligned}
$$

Since $\mathbf{1}K_\theta^{-1}\dot{\theta} = 0$ is always true, $\mathbf{1}K_\theta^{-1}\theta$ is constant and is determined by the initial condition. Together with the above two equations, θ can be uniquely determined as

$\theta = \theta^*$. According to Remark 2.6, we obtain $\alpha^* = [\Gamma_1^T, \Gamma_2^T]\theta^*$, which leads to that l^+, l^- are constant. Thus, we have

$\{(\omega_g, \alpha, P_M, P_C, P_L, \theta, \zeta, \lambda, \mu^+, \mu^-, \nu^+, \nu^-, l^+, l^-) | \dot{V}_{(5.11)} = 0\} \subseteq$
$\{(\omega_g, \alpha, P_M, P_C, P_L, \theta, \zeta, \lambda, \mu^+, \mu^-, \nu^+, \nu^-, l^+, l^-) | \dot{\omega}_g = \dot{P}_M = \dot{P}_C = \dot{\zeta} = \dot{\mu}^+ = \dot{\mu}^- = \mathbf{0},$
$\dot{\alpha} = \mathbf{0}, \dot{P}_L = \dot{\lambda} = \dot{\nu}^+ = \dot{\nu}^- = \mathbf{0}, \dot{\theta} = \mathbf{0}, \dot{l}^+ = \dot{l}^- = \mathbf{0}\} \subseteq$
$\{(\omega_g, \alpha, P_M, P_C, P_L, \theta, \zeta, \lambda, \mu^+, \mu^-, \nu^+, \nu^-, l^+, l^-) | (\omega_g, \alpha, P_M, P_C, P_L, \theta, \zeta, \lambda, \mu^+, \mu^-, \nu^+,$
$\nu^-, l^+, l^-)$ is a saddle point of $L_{au}\}$.

According to the proof of Proposition 13 in [70], we obtain that the trajectories of the overall system asymptotically converge to a saddle point of L_{au}, at which the steady-state optimization problem (5.2) is solved. ■

Theorem 5.2 immediately leads to the optimality and stability of system (5.11). The benefit of the saddle point design approach is summarized as follows. First, the approach allows us to embed different kinds of steady-state convex optimization problems, where problem (5.3) is a typical example. Second, as long as the optimization problem is distributed, e.g., with separable objective functions and local constraints, the resulting controller is completely distributed. Third, the trajectories of the closed-loop system are bounded, and can be further proven to converge to an equilibrium point where the optimization problem is solved, as illustrated in Theorem 5.2.

5.2 Distributed Optimal Steady-state Control in Linear Network Systems

In the previous section, we have presented a reverse- and forward-engineering design framework to solve the real-time economic dispatch problem for power systems. We now generalize this framework for a class of Linear Time-Invariant (LTI) network systems to achieve optimal steady-state performance.

5.2.1 Problem Setup and Preliminaries

Consider a linear network system consisting of N subsystems:

$$\dot{x}_i = \sum_{j \in \mathcal{N}(i)} A_{ij} x_j + B_i u_i + C_i w_i, \ i = 1, \cdots, N \tag{5.15}$$

where $x_i(t) \in \mathbb{R}^{n_i}$ is the state vector of subsystem i, $\mathcal{N}(i)$ is the set of neighbouring subsystems of subsystem i, $A_{ij} \in \mathbb{R}^{n_i \times n_j}$, $B_i \in \mathbb{R}^{n_i \times m_i}$, $u_i(t) \in \mathbb{R}^{m_i}$ is the control

input vector to subsystem i, $C_i \in \mathbb{R}^{n_i \times p_i}$, and $w_i(t) \in \mathbb{R}^{p_i}$ is the exogenous input vector to subsystem i, e.g., disturbance injection.

To make our discussion practical and general, in addition to system dynamics (5.15), we also consider the scenario in which some or all subsystems are equipped with existing built-in controllers which can be described as

$$\dot{u}_i = \sum_{j \in \mathcal{N}(i)} D_{ij} x_j + \sum_{j \in \mathcal{N}(i)} E_{ij} u_j + F_i w_i, i = 1, \cdots, N \tag{5.16}$$

where $D_{ij} \in \mathbb{R}^{m_i \times n_j}$, $E_{ij} \in \mathbb{R}^{m_i \times m_j}$, $F_i \in \mathbb{R}^{m_i \times p_i}$. Here the dynamic feedback controller is assumed to have a distributed structure. For convenience, let $x = \{x_1, \cdots, x_N\}$, $u = \{u_1, \cdots, u_N\}$, $w = \{w_1, \cdots, w_N\}$, $n = \sum_{i=1}^N n_i$, $m = \sum_{i=1}^N m_i$ and $p = \sum_{i=1}^N p_i$.

Remark 5.1 *System (5.15)-(5.16) describes a general class of network systems with built-in control mechanisms. Equation (5.16) represents any distributed dynamic feedback controller with order one or more. Moreover, $D_{ij} = \mathbf{0}, E_{ij} = \mathbf{0}, F_i = \mathbf{0}, \forall j \in \mathcal{N}(i)$ means that there is no such controller equipped with subsystem i. Note that a zero-order distributed state feedback controller, i.e., $u_i = \sum_{j \in \mathcal{N}(i)} K_{ij} x_j$ where $K_{ij} \in \mathbb{R}^{m_i \times n_j}$, can be included in the system dynamics (5.15) through A_{ij}. Lastly, $F_i = \mathbf{0}$ indicates that the controller of subsystem i does not use any information of disturbance w_i.*

Suppose that the exogenous disturbance w is *constant*, and that the closed-loop system (5.15)-(5.16) is stabilized and each trajectory converges to one point in the equilibrium set $\mathcal{X} := \{x | \sum_{j \in \mathcal{N}(i)} A_{ij} x_j + B_i u_i + C_i w_i = \mathbf{0}, \sum_{j \in \mathcal{N}(i)} D_{ij} x_j + \sum_{j \in \mathcal{N}(i)} E_{ij} u_j + F_i w_i = \mathbf{0}, i = 1, \cdots, N\}$. Now consider the following optimization problem associated with the system under a constant disturbance w:

$$\min_{x \in \mathbb{R}^n, u \in \mathbb{R}^m} \quad \sum_{i=1}^N f_i(x_i) + \sum_{i=1}^N g_i(u_i) \tag{5.17a}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{N}(i)} A_{ij} x_j + B_i u_i + C_i w_i = \mathbf{0} \tag{5.17b}$$

$$\sum_{j \in \mathcal{N}(i)} D_{ij} x_j + \sum_{j \in \mathcal{N}(i)} E_{ij} u_j + F_i w_i = \mathbf{0} \tag{5.17c}$$

$$h_i(x_i, u_i) \leq 0 \tag{5.17d}$$

where $i = 1, \cdots, N$, $f_i \in \mathcal{C}^2$: $\mathbb{R}^{n_i} \to \mathbb{R}$, $g_i \in \mathcal{C}^2$: $\mathbb{R}^{m_i} \to \mathbb{R}$, $h_i \in \mathcal{C}^2$: $\mathbb{R}^{n_i} \times \mathbb{R}^{m_i} \to \mathbb{R}$. We assume that problem (5.17) is convex (i.e., $\nabla^2_{x_i} f_i \succeq 0$, $\nabla^2_{u_i} g_i \succeq 0$

and $\nabla^2_{x_i,u_i} h_i \succeq 0$ hold), feasible and satisfies Slater's constraint qualification [77]. This optimization problem usually defines a desired operating point for the network system. For instance, in power systems, (5.17) may be an OPF problem in which generation cost is minimized and consumption utility is maximized subject to power flow balance constraints and network operating constraints [3]. Another example is the space satellite formation control in which (5.17) is usually formulated as a consensus problem with respect to the positions and velocities of space satellites [97].

Remark 5.2 *In problem (5.17), there can be more inequality constraints for each subsystem, i.e., Equation (5.17d) can be reformulated in a vector form. For convenience, in this section, we consider the case in which there is only one inequality constraint for each subsystem, while the result here can be immediately extended.*

The problem we focus on is to modify the existing controller (5.16) in a distributed and closed-loop manner so that the system can track the optimal solution of (5.17) automatically. In the existing literature and practice, e.g., frequency control in power systems, controller (5.16) is designed to drive the system to a predefined nominal operating point which is derived by solving (5.17) using a prediction on the future w. Moreover, the optimization problem is usually solved (i) at a much slower time-scale compared with system dynamics (5.15)-(5.16) and (ii) by using centralized optimization or distributed optimization methods, both of which require a certain amount of communication and computation. If controller (5.16) can be modified to track the optimal solution of (5.17) automatically and be implemented in a distributed and closed-loop manner, then the system itself can adapt to fast change of uncertain disturbance w. Furthermore measurement, communication and computation can be saved.

Formally speaking, we focus on a network system (5.15)-(5.16) which belongs to following class:

Class-$\mathcal{S}'$: System (5.15)-(5.16) belongs to Class-$\mathcal{S}'$ means that there exists a function $L_{\text{sys}}(x^{(1)}, x^{(2)}, u) : \mathbb{R}^{m+n} \to \mathbb{R}$ and positive definite matrices $P_{x^{(1)}}$, $P_{x^{(2)}}$ and $P_{u_i}, i = 1, \cdots, N$ such that $\nabla^2_{x^{(1)}} L_{\text{sys}} \preceq 0$, $\nabla^2_{x^{(2)},u} L_{\text{sys}} \succeq 0$, the saddle point set $\{(x^{(1)}, x^{(2)}, u) | \nabla_{x^{(1)},x^{(2)},u} L_{\text{sys}} = \mathbf{0}\}$ is nonempty, and (5.15)-(5.16) is a primal-dual gradient algorithm to solve $\max_{x^{(1)}} \min_{x^{(2)},u} L_{\text{sys}}$, i.e.,

$$\dot{x} = \text{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}\} \frac{\partial L_{\text{sys}}}{\partial x}$$
$$\dot{u}_i = -P_{u_i} \frac{\partial L_{\text{sys}}}{\partial u_i}, \ i = 1, \cdots, N.$$

Remark 5.3 *In the above definition, (i) state x is partitioned into $x^{(1)}$ and $x^{(2)}$; (ii) u is required as a minimizer; (iii) the gain matrix for $\dot{u}$ is block diagonal consisting of $P_{u_i}, i = 1, \cdots, N$, i.e., $\dot{u}$ has a distributed structure. The definition can be extended to the case where u is required as a maximizer (by adding a minus sign before L_{sys}). For ease of exposition, we only focus on the class in the above definition.*

The motivation to consider this special class includes: (i) recent research has demonstrated that there are many cyber-physical systems, such as power systems and Internet congestion control protocols, belonging to this class [71, 51]; (ii) if a system belongs to Class-$\mathcal{S}'$, then we have a "forward-engineering" procedure to modify the controller (5.16) for achieving the optimal steady-state performance described in (5.17), which is demonstrated in Section 5.2.2.

Although some recent work has shown that power system frequency control and Internet congestion control belong to Class-$\mathcal{S}'$, fundamental questions remain. For example, how general this class is? Do we have a rigorous and general characterization of this class in terms of the property of system matrices? These questions lead to Section 5.2.3 in which we study reverse-engineering for general LTI systems.

5.2.2 Forward-engineering for Control Modification

In this subsection we provide a procedure to modify the controller of system (5.15)-(5.16) so that the closed-loop system can track the optimal solution of (5.17) automatically, under the premise that system (5.15)-(5.16) belongs to Class-$\mathcal{S}'$.

Step 1): Introduce auxiliary decision variables
Modify problem (5.17) by replacing x (x_i) with an auxiliary decision vector $y \in \mathbb{R}^n$ ($y_i \in \mathbb{R}^{n_i}$):

$$\min_{y\in\mathbb{R}^n, u\in\mathbb{R}^m} \quad \sum_{i=1}^{N} f_i(y_i) + \sum_{i=1}^{N} g_i(u_i) \tag{5.18a}$$

$$\text{subject to} \quad \sum_{j\in\mathcal{N}(i)} A_{ij}y_j + B_i u_i + C_i w_i = \mathbf{0} \tag{5.18b}$$

$$\sum_{j\in\mathcal{N}(i)} D_{ij}y_j + \sum_{j\in\mathcal{N}(i)} E_{ij}u_j + F_i w_i = \mathbf{0} \tag{5.18c}$$

$$h_i(y_i, u_i) \le 0, \text{ where } i = 1, \cdots, N. \tag{5.18d}$$

Step 2): Merge objective functions
Since problem (5.18) is convex and strong duality holds, derive a saddle point problem corresponding to (5.18), given by

$$\max_{\zeta_i\in\mathbb{R}^{n_i},\lambda_i\in\mathbb{R}^{m_i},\mu_i\geq 0}\ \min_{y\in\mathbb{R}^n,u\in\mathbb{R}^m} L_{\rm op}=\sum_{i=1}^N f_i(y_i)+\sum_{i=1}^N g_i(u_i)+\sum_{i=1}^N \mu_i h_i(y_i,u_i)$$
$$-\sum_{i=1}^N \zeta_i^T\left(\sum_{j\in\mathcal{N}(i)} A_{ij}y_j+B_iu_i+C_iw_i\right)-\sum_{i=1}^N\lambda_i^T\left(\sum_{j\in\mathcal{N}(i)} D_{ij}y_j+\sum_{j\in\mathcal{N}(i)} E_{ij}u_j+F_iw_i\right) \tag{5.19}$$

where $\zeta_i\in\mathbb{R}^{n_i},\lambda_i\in\mathbb{R}^{m_i},\mu_i\geq 0$ are Lagrangian multipliers (dual variables) for the constraints in (5.18). By adding $L_{\rm sys}$, we obtain an augmented saddle point problem:

$$\max_{\zeta_i\in\mathbb{R}^{n_i},\lambda_i\in\mathbb{R}^{m_i},\mu_i\geq 0,x^{(1)}\in\mathbb{R}^{n^{(1)}}}\ \min_{y\in\mathbb{R}^n,u\in\mathbb{R}^m,x^{(2)}\in\mathbb{R}^{n^{(2)}}} L_{\rm au}=L_{\rm sys}+\gamma L_{\rm op} \tag{5.20}$$

where $\gamma>0$ is a constant. The next lemma shows the properties of $L_{\rm au}$.

Lemma 5.3 *$\nabla^2_{y,u,x^{(2)}}L_{au}\succeq 0$ and $\nabla^2_{\zeta_i,\lambda_i,\mu_i,x^{(1)}}L_{au}\preceq 0$ hold. Moreover, if A is invertible in (5.15) (here we rewrite (5.15) as $\dot{x}=Ax+Bu+Cw$, where $A\in\mathbb{R}^{n\times n},B\in\mathbb{R}^{n\times m},C\in\mathbb{R}^{n\times p}$), then $(y,u,x,\zeta_i,\lambda_i,\mu_i)$ is a saddle point of L_{au} if and only if $(y,u,\zeta_i,\lambda_i,\mu_i)$ is a saddle point of L_{op} and (x,u) is a saddle point of L_{sys}.*

Proof: Due to $\nabla^2_{y,u,x^{(2)}}L_{\rm au}=\nabla^2_{y,u,x^{(2)}}L_{\rm sys}+\gamma\nabla^2_{y,u,x^{(2)}}L_{\rm op}$, $\nabla^2_{y,u,x^{(2)}}L_{\rm sys}\succeq 0$ and $\nabla^2_{y,u,x^{(2)}}L_{\rm op}\succeq 0$, we have $\nabla^2_{y,u,x^{(2)}}L_{\rm au}\succeq 0$. Similarly, $\nabla^2_{\zeta_i,\lambda_i,\mu_i,x^{(1)}}L_{\rm au}\preceq 0$ holds. Based on the KKT conditions [77] (saddle points of a function satisfy the KKT conditions), any saddle point of $L_{\rm au}$ satisfies

$$\sum_{j\in\mathcal{N}(i)} A_{ij}x_j+B_iu_i+C_iw_i=\mathbf{0} \tag{5.21a}$$
$$\sum_{j\in\mathcal{N}(i)} D_{ij}x_j+\sum_{j\in\mathcal{N}(i)} E_{ij}u_j+F_iw_i-\gamma P_{u_i}\left(\frac{\partial g_i}{\partial u_i}-B_i^T\zeta_i-\sum_{j\in\mathcal{N}(i)} E_{ji}^T\lambda_j+\mu_i\frac{\partial h_i}{\partial u_i}\right)=\mathbf{0} \tag{5.21b}$$
$$\frac{\partial f_i}{\partial y_i}-\sum_{j\in\mathcal{N}(i)} A_{ji}^T\zeta_j-\sum_{j\in\mathcal{N}(i)} D_{ji}^T\lambda_j+\mu_i\frac{\partial h_i}{\partial y_i}=\mathbf{0} \tag{5.21c}$$
$$\sum_{j\in\mathcal{N}(i)} A_{ij}y_j+B_iu_i+C_iw_i=\mathbf{0} \tag{5.21d}$$
$$\sum_{j\in\mathcal{N}(i)} D_{ij}y_j+\sum_{j\in\mathcal{N}(i)} E_{ij}u_j+F_iw_i=\mathbf{0} \tag{5.21e}$$
$$\mu_ih_i(y_i,u_i)=0,\quad \mu_i\geq 0,\quad h_i(y_i,u_i)\leq 0 \tag{5.21f}$$

where $i = 1, \cdots, N$. If A is invertible in (5.15), from (5.21a), (5.21d)-(5.21e), $y = x$ and $\sum_{j\in\mathcal{N}(i)} D_{ij}x_j + \sum_{j\in\mathcal{N}(i)} E_{ij}u_j + F_i w_i = \mathbf{0}$ hold. On the other hand, any saddle point of L_{sys} satisfies

$$\sum_{j\in\mathcal{N}(i)} A_{ij}x_j + B_i u_i + C_i w_i = \mathbf{0}$$
$$\sum_{j\in\mathcal{N}(i)} D_{ij}x_j + \sum_{j\in\mathcal{N}(i)} E_{ij}u_j + F_i w_i = \mathbf{0}$$

where $i = 1, \cdots, N$, and any saddle point of L_{op} satisfies

$$\frac{\partial g_i}{\partial u_i} - B_i^T\zeta_i - \sum_{j\in\mathcal{N}(i)} E_{ji}^T\lambda_j + \mu_i\frac{\partial h_i}{\partial u_i} = \mathbf{0}$$
$$\frac{\partial f_i}{\partial y_i} - \sum_{j\in\mathcal{N}(i)} A_{ji}^T\zeta_j - \sum_{j\in\mathcal{N}(i)} D_{ji}^T\lambda_j + \mu_i\frac{\partial h_i}{\partial y_i} = \mathbf{0}$$
$$\sum_{j\in\mathcal{N}(i)} A_{ij}y_j + B_i u_i + C_i w_i = \mathbf{0}$$
$$\sum_{j\in\mathcal{N}(i)} D_{ij}y_j + \sum_{j\in\mathcal{N}(i)} E_{ij}u_j + F_i w_i = \mathbf{0}$$
$$\mu_i h_i(y_i, u_i) = 0, \quad \mu_i \geq 0, \quad h_i(y_i, u_i) \leq 0$$

where $i = 1, \cdots, N$. Therefore, the following two sets are equivalent:

$$\{(y, u, x, \zeta_i, \lambda_i, \mu_i) | (y, u, x, \zeta_i, \lambda_i, \mu_i) \text{ is a saddle point of } L_{\text{au}}\} \Leftrightarrow$$
$$\{(y, u, x, \zeta_i, \lambda_i, \mu_i) | (x, u) \text{ is a saddle point of } L_{\text{sys}},$$
$$(y, u, \zeta_i, \lambda_i, \mu_i) \text{ is a saddle point of } L_{\text{op}}\}$$

and furthermore, $y = x$ holds. So $(y, u, x, \zeta_i, \lambda_i, \mu_i)$ is a saddle point of L_{au} if and only if $(y, u, \zeta_i, \lambda_i, \mu_i)$ is a saddle point of L_{op} and (x, u) is a saddle point of L_{sys}. ■

Step 3): Forward-engineering

With the new saddle point function L_{au}, derive the following saddle point dynamics:

$$\dot{x}_i = \sum_{j\in\mathcal{N}(i)} A_{ij}x_j + B_i u_i + C_i w_i \tag{5.22a}$$
$$\dot{u}_i = \sum_{j\in\mathcal{N}(i)} D_{ij}x_j + \sum_{j\in\mathcal{N}(i)} E_{ij}u_j + F_i w_i$$
$$- \gamma P_{u_i}\left(\frac{\partial g_i}{\partial u_i} - B_i^T\zeta_i - \sum_{j\in\mathcal{N}(i)} E_{ji}^T\lambda_j + \mu_i\frac{\partial h_i}{\partial u_i} + K_{eu_i}(u_i - \hat{u}_i)\right) \tag{5.22b}$$
$$\dot{\hat{u}}_i = \hat{K}_{eu_i}(u_i - \hat{u}_i) \tag{5.22c}$$

$$\dot{y}_i = - K_{y_i} \left(\frac{\partial f_i}{\partial y_i} - \sum_{j \in \mathcal{N}(i)} A_{ji}^T \zeta_j - \sum_{j \in \mathcal{N}(i)} D_{ji}^T \lambda_j + \mu_i \frac{\partial h_i}{\partial y_i} + K_{ey_i}(y_i - \hat{y}_i) \right) \tag{5.22d}$$

$$\dot{\hat{y}}_i = \hat{K}_{ey_i}(y_i - \hat{y}_i) \tag{5.22e}$$

$$\dot{\zeta}_i = - K_{\zeta_i} \left(\sum_{j \in \mathcal{N}(i)} A_{ij} y_j + B_i u_i + C_i w_i + K_{e\zeta_i}(\zeta_i - \hat{\zeta}_i) \right) \tag{5.22f}$$

$$\dot{\hat{\zeta}}_i = \hat{K}_{e\zeta_i}(\zeta_i - \hat{\zeta}_i) \tag{5.22g}$$

$$\dot{\lambda}_i = - K_{\lambda_i} \left(\sum_{j \in \mathcal{N}(i)} D_{ij} y_j + \sum_{j \in \mathcal{N}(i)} E_{ij} u_j + F_i w_i + K_{e\lambda_i}(\lambda_i - \hat{\lambda}_i) \right) \tag{5.22h}$$

$$\dot{\hat{\lambda}}_i = \hat{K}_{e\lambda_i}(\lambda_i - \hat{\lambda}_i) \tag{5.22i}$$

$$\dot{\mu}_i = k_{\mu_i}(h_i(y_i, u_i))_{\mu_i}^+ \tag{5.22j}$$

where $K_{eu_i}, \hat{K}_{eu_i}, K_{\lambda_i}, K_{e\lambda_i}, \hat{K}_{e\lambda_i} \in \mathbb{R}^{m_i \times m_i}$, K_{y_i}, K_{ey_i}, $\hat{K}_{ey_i}, K_{\zeta_i}, K_{e\zeta_i}, \hat{K}_{e\zeta_i} \in \mathbb{R}^{n_i \times n_i}$ are positive definite constant diagonal matrices, $k_{\mu_i} > 0$, and $i = 1, \cdots, N$.

The above saddle point algorithm is not exactly the primal-dual gradient algorithm (5.4). It is a modified saddle point algorithm given by the following lemma.

Lemma 5.4 *Let $f \in \mathcal{C}^2$: $\mathbb{R}^a \times \mathbb{R}^b \to \mathbb{R}$ satisfy: for all y, z, $\nabla_y^2 f \succeq 0, \nabla_z^2 f \preceq 0$, and the set $\{(y,z) | \nabla_{y,z} f = \mathbf{0}\}$ is nonempty. Then each trajectory of the modified saddle point dynamics given by*

$$\dot{y} = -K_y \left(\frac{\partial f}{\partial y} + K_{ey}(y - \hat{y}) \right) \tag{5.23a}$$

$$\dot{\hat{y}} = \hat{K}_{ey}(y - \hat{y}) \tag{5.23b}$$

$$\dot{z} = K_z \left(\frac{\partial f}{\partial z} - K_{ez}(z - \hat{z}) \right) \tag{5.23c}$$

$$\dot{\hat{z}} = \hat{K}_{ez}(z - \hat{z}) \tag{5.23d}$$

asymptotically converges to an equilibrium point at which (y, z) is a saddle point of f. Here $\hat{y}(t) \in \mathbb{R}^a, \hat{z}(t) \in \mathbb{R}^b$ are auxiliary decision variables, and $K_{ey}, \hat{K}_{ey} \in \mathbb{R}^{a \times a}, K_{ez}, \hat{K}_{ez} \in \mathbb{R}^{b \times b}$ are positive definite constant diagonal matrices.

Proof: Let (y^*, z^*) be a saddle point of f. Define a candidate Lyapunov function for (5.23) as

$$\begin{aligned} V_{(5.23)} =& \frac{1}{2}(y - y^*)^T K_y^{-1}(y - y^*) + \frac{1}{2}(z - z^*)^T K_z^{-1}(z - z^*) + \frac{1}{2}(\hat{y} - y^*)^T K_{ey} \hat{K}_{ey}^{-1}(\hat{y} - y^*) \\ & + (\hat{z} - z^*)^T K_{ez} \hat{K}_{ez}^{-1}(\hat{z} - z^*) \end{aligned}$$

which is radially unbounded and positive definite with respect to (y^*, z^*, y^*, z^*) (note that at any equilibrium of (5.23), $y^* = \hat{y}^*, z^* = \hat{z}^*$ hold). The derivative of $V_{(5.23)}$ with respect to time along the trajectory of system (5.23) is given by

$$\begin{aligned}
\dot{V}_{(5.23)} =& -\left(\frac{\partial f}{\partial y}\right)^T (y - y^*) + \left(\frac{\partial f}{\partial z}\right)^T (z - z^*) - (y - \hat{y})^T K_{ey}(y - \hat{y}) \\
& - (z - \hat{z})^T K_{ez}(z - \hat{z}) \\
\leq& - f(y, z) + f(y^*, z) + f(y, z) - f(y, z^*) - (y - \hat{y})^T K_{ey}(y - \hat{y}) \\
& - (z - \hat{z})^T K_{ez}(z - \hat{z}) \\
=& f(y^*, z) - f(y^*, z^*) + f(y^*, z^*) - f(y, z^*) - (y - \hat{y})^T K_{ey}(y - \hat{y}) \\
& - (z - \hat{z})^T K_{ez}(z - \hat{z}) \\
\leq& 0
\end{aligned}$$

where the first inequality comes from the fact that f is convex in y and concave in z, and the last inequality follows that (y^*, z^*) is a saddle point of f. When $\dot{V}_{(5.23)} = 0$, we have $y = \hat{y}, z = \hat{z}$, leading to $\dot{\hat{y}} = \dot{y} = \mathbf{0}, \dot{\hat{z}} = \dot{z} = \mathbf{0}$. So $\frac{\partial f}{\partial y} = \mathbf{0}$ and $\frac{\partial f}{\partial z} = \mathbf{0}$ hold, indicating that (y, z) is a saddle point of f. From LaSalle's invariance principle [79], we conclude that each trajectory of system (5.23) asymptotically converges to an equilibrium point at which (y, z) is a saddle point of f. ■

Step 4): Closed-loop implementation

Let $\tilde{\zeta}_i = x_i + K_{\zeta_i}^{-1}\zeta_i$ and $\tilde{\lambda}_i = u_i + K_{\lambda_i}^{-1}\lambda_i$. Rewrite Equations (5.22f)-(5.22i) as (also substitute $\zeta_i = K_{\zeta_i}(\tilde{\zeta}_i - x_i), \lambda_i = K_{\lambda_i}(\tilde{\lambda}_i - u_i)$ to Equations (5.22b) and (5.22d))

$$\dot{\tilde{\zeta}}_i = \sum_{j \in \mathcal{N}(i)} A_{ij}(x_j - y_j) - K_{e\zeta_i}(K_{\zeta_i}(\tilde{\zeta}_i - x_i) - \hat{\zeta}_i) \tag{5.24a}$$

$$\dot{\hat{\zeta}}_i = \hat{K}_{e\zeta_i}(K_{\zeta_i}(\tilde{\zeta}_i - x_i) - \hat{\zeta}_i) \tag{5.24b}$$

$$\begin{aligned}
\dot{\tilde{\lambda}}_i =& \sum_{j \in \mathcal{N}(i)} D_{ij}(x_j - y_j) - K_{e\lambda_i}(K_{\lambda_i}(\tilde{\lambda}_i - u_i) - \hat{\lambda}_i) \\
& - \gamma P_{u_i}\left(\frac{\partial g_i}{\partial u_i} - B_i^T K_{\zeta_i}(\tilde{\zeta}_i - x_i) - \sum_{j \in \mathcal{N}(i)} E_{ji}^T K_{\lambda_j}(\tilde{\lambda}_j - u_j) + \mu_i \frac{\partial h_i}{\partial u_i} + K_{eu_i}(u_i - \hat{u}_i)\right)
\end{aligned} \tag{5.24c}$$

$$\dot{\hat{\lambda}}_i = \hat{K}_{e\lambda_i}(K_{\lambda_i}(\tilde{\lambda}_i - u_i) - \hat{\lambda}_i) \tag{5.24d}$$

so that the extra states $\tilde{\zeta}_i, \tilde{\lambda}_i$ are independent of w.

For the optimality and stability of system (5.22), we have the following theorem.

Theorem 5.3 *If (5.15)-(5.16) belongs to Class-$\mathcal{S}'$ and A is Hurwitz (recall Lemma 5.3 for the definition of A), each trajectory of (5.22) asymptotically converges to an equilibrium point at which (x, u) is an optimal solution of (5.17).*

Proof: Since A is Hurwitz in (5.15), under Lemma 5.3, at any equilibrium point of (5.22), (x, u) is an optimal solution of problem (5.17). The proof of convergence is similar to that in Lemma 5.4, by constructing a quadratic Lyapunov function and showing that its derivative with respect to time is non-increasing along the trajectory of system (5.22). When this derivative is 0, $u_i = \hat{u}_i$, $y_i = \hat{y}_i$, $\zeta_i = \hat{\zeta}_i$, $\lambda_i = \hat{\lambda}_i$ hold, which lead to $\dot{u}_i = \mathbf{0}, \dot{y}_i = \mathbf{0}, \dot{\zeta}_i = \mathbf{0}, \dot{\lambda}_i = \mathbf{0}$. Given constant u and w, system (5.15) eventually converges to an equilibrium point at which $x = y$. Then $\dot{\mu}_i = 0, i = 1, \cdots, N$ are true. From LaSalle's invariance principle [79], each trajectory of (5.22) asymptotically converges to an equilibrium point at which (x, u) is an optimal solution of problem (5.17). ■

Remark 5.4 *The condition in the above theorem requires A to be Hurwitz. This can be realized by introducing a static feedback controller $u = Kx, K \in \mathbb{R}^{n \times n}$ to pre-stabilize (5.15) if A is not Hurwitz. A similar assumption is also used in [5], Chapter 5.2. Note that system (5.22) is slightly different from the formulation (5.23), in which only $u_i, y_i, \zeta_i, \lambda_i$ are equipped with auxiliary decision variables $\hat{u}_i, \hat{y}_i, \hat{\zeta}_i, \hat{\lambda}_i$ respectively. Similar examples have been shown in Chapter 4.*

Going from the original closed-loop system (5.15)-(5.16) to the modified one (5.22), we have introduced extra dynamics while the structure of the original dynamic feedback controller is preserved as shown in Equation (5.22b). The benefit of this forward-engineering modification is summarized as follows. First, the modification allows us to embed different types of steady-state convex optimization problems. Second, as long as the steady-state optimization problem is distributed, e.g., with separable objective functions and local constraints as shown in problem (5.17), the resulting extra dynamics are completely distributed as given in system (5.22). Third, the modification can ensure that each trajectory of the redesigned closed-loop system converges to an equilibrium point where the steady-state optimization problem is solved.

5.2.3 LTI Systems as Gradient Algorithms for Quadratic Optimization

In the previous subsection, a forward-engineering framework is presented to modify an existing control scheme for solving optimal steady-state control problems. Class-

$\mathcal{S}'$ is considered as a prerequisite, which implies that the original closed-loop system (5.15)-(5.16) can be reverse-engineered as a primal-dual gradient algorithm to solve an unconstrained quadratic saddle point problem. In this subsection, we investigate Class-$\mathcal{S}'$ by studying what kind of systems belongs to it. Generally speaking, a system should be at least marginally stable [98] if it can be interpreted as a primal-dual gradient algorithm for an unconstrained quadratic saddle point problem, since system trajectories are bounded in this case (recall Lemma 5.2). Also, the equilibrium set of the system should be equivalent to the saddle point set of the resulting problem. To limit the scope of discussion, we only focus on reverse-engineering continuous-time LTI systems that are in accord with (5.15)-(5.16).

1): LTI autonomous systems

We begin our investigation by studying a simple class of systems, i.e., LTI autonomous systems. This allows us to understand the insight behind the mathematical derivation better. Consider an LTI autonomous system:

$$\dot{x} = Ax \tag{5.25}$$

where $x(t) \in \mathbb{R}^n$ is the state vector and $A \in \mathbb{R}^{n\times n}$. First, let us study the following class of autonomous systems which can be reverse-engineered as gradient algorithms for solving an unconstrained convex quadratic programming problem.

Class-$\mathcal{O}$: System (5.25) belongs to Class-$\mathcal{O}$ means that there exists a function $L(x) : \mathbb{R}^n \to \mathbb{R}$ and a positive definite matrix $P \in \mathbb{R}^{n\times n}$ such that $\nabla_x^2 L \preceq 0$, $\{x|\nabla_x L = \mathbf{0}\}$ is nonempty, and (5.25) is a gradient algorithm to solve $\max_x L$, i.e.,

$$\dot{x} = P\frac{\partial L}{\partial x}.$$

Since (5.25) is linear, if (5.25) belongs to Class-$\mathcal{O}$, then the associated L must be a convex quadratic function, i.e., $L = \frac{1}{2}x^T Qx$ for some $Q \preceq 0$. Therefore, system (5.25) belongs to Class-$\mathcal{O}$ if and only if there exist matrices $P \succeq 0$ and $Q \preceq 0$ such that $A = PQ$ holds. This leads to our first result regarding reverse-engineering.

Lemma 5.5 *Given $A \in \mathbb{R}^{n\times n}$, the following statements are equivalent:*

(i) There exists a matrix $P = P^T \in \mathbb{R}^{n\times n}$ and $P \succ 0$, so that $PA = A^T P$ holds;

(ii) $eig(A) \in \mathbb{R}$ and A is diagonalizable (here we use $eig(A) \in \mathbb{R}$ to indicate that all eigenvalues of A are real numbers, although $eig(A)$ is a set).

Proof: (i)⇒(ii). Write A in its Jordan canonical form $A = J\Lambda J^{-1}$, where $\Lambda = \text{diag}\{\Lambda_1, \cdots, \Lambda_a\}$ and each Λ_i is a Jordan block. Without loss of generality, let

$$\Lambda_1 = \begin{bmatrix} \lambda_1 & 1 & \mathbf{0} \\ 0 & \lambda_1 & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots \end{bmatrix}$$

and the dimension of Λ_1 is at least 2×2. Then (i) leads to $PJ\Lambda J^{-1} = (J^{-1})^H \Lambda^H J^H P$ which is equivalent to

$$J^H P J \Lambda = \Lambda^H J^H P J. \tag{5.26}$$

Note that $J^H P J$ is a Hermitian matrix, and $J^H P J \succ 0$ since $P \succ 0$ and J is invertible. So we obtain that all the diagonal entries of $J^H P J$ are positive real. Let

$$J^H P J = \begin{bmatrix} s_{11} & s_{12} & \cdots \\ s_{12}^H & s_{22} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

where $s_{11}, s_{22} > 0$. From (5.26), we have $\lambda_1 s_{11} = \lambda_1^H s_{11}$ and $s_{11} + s_{12}\lambda_1 = \lambda_1^H s_{12}$ which result in $\lambda_1 = \lambda_1^H$ and $s_{11} = 0$, i.e., a contradiction. Therefore, each Jordan block should have the dimension 1×1 and a real diagonal entry, i.e., A is diagonalizable and $\text{eig}(A) \in \mathbb{R}$ are true.

(ii)⇒(i). Based on (ii), write A as its diagonal canonical form $A = J\Lambda J^{-1}$, where Λ is a diagonal matrix with real entries on the diagonal and $J \in \mathbb{R}^{n\times n}$. Define a matrix $V \in \mathbb{R}^{n\times n}$ so that $V = V^T \succ 0$ and $V\Lambda = \Lambda V$ hold. Such a matrix V always exists and the simplest choice is $V = I_n$. We can then find a matrix $P = P^T = (J^{-1})^T V J^{-1} \in \mathbb{R}^{n\times n}$ so that $P \succ 0$ and $PA = A^T P$ hold, i.e., (i) is true.■

Theorem 5.4 *System (5.25) belongs to Class-$\mathcal{O}$ if and only if (5.25) is marginally or asymptotically stable, eig$(A) \in \mathbb{R}$ and A is diagonalizable.*

Proof: Necessity. Suppose that system (5.25) is a gradient algorithm to solve an unconstrained convex quadratic programming problem given by

$$\max_{x\in\mathbb{R}^n} L(x) = \frac{1}{2} x^T Q x \tag{5.27}$$

where $Q \in \mathbb{R}^{n\times n}$ satisfies $Q = Q^T \preceq 0$. Then the trajectories of (5.25) are bounded under Lemma 5.2, and there exists a matrix $P = P^T \in \mathbb{R}^{n\times n}$ and $P \succ 0$, so that $\dot{x} = P\frac{\partial L}{\partial x} = PQx = Ax$ is always true. i.e., $A = PQ$ holds. This leads to $P^{-1}A = A^T P^{-1}$ which is equivalent to $\text{eig}(A) \in \mathbb{R}$ and A is diagonalizable, based on Lemma 5.5. Also, system (5.25) is marginally or asymptotically stable.

Sufficiency. Because system (5.25) is marginally or asymptotically stable, $\text{eig}(A) \in \mathbb{R}$ and A is diagonalizable, A can be written as a diagonal canonical form $A = J\Lambda J^{-1}$ where $J \in \mathbb{R}^{n\times n}$ and $\Lambda \preceq 0$. Under Lemma 5.5, there exists a positive definite matrix $V \in \mathbb{R}^{n\times n}$ so that $V\Lambda = \Lambda V$ holds. Based on Lemma 1 in [72], we have $V\Lambda \preceq 0$. Define a matrix $P = JV^{-1}J^T$. Then $P^{-1}A = A^T P^{-1} \preceq 0$ holds. Considering the unconstrained convex quadratic programming problem (5.27) where $Q = P^{-1}A$, we conclude that system (5.25) belongs to Class-$\mathcal{O}$. ■

Theorem 5.4 proposes a necessary and sufficient condition to reverse-engineer system (5.25) as a gradient algorithm to solve a related unconstrained convex quadratic programming problem. We now study the case where we can interpret system (5.25) as a primal-dual gradient algorithm to solve an unconstrained quadratic saddle point problem. Consider the following class of systems.

Class-$\mathcal{S}$: System (5.25) belongs to Class-$\mathcal{S}$ means that there exists a function $L(x^{(1)}, x^{(2)}) : \mathbb{R}^n \to \mathbb{R}$ and positive definite matrices $P_{x^{(1)}}$, $P_{x^{(2)}}$ such that $\nabla^2_{x^{(1)}} L \preceq 0$, $\nabla^2_{x^{(2)}} L \succeq 0$, the saddle point set $\{x | \nabla_x L = \mathbf{0}\}$ is nonempty, and (5.25) is a primal-dual gradient algorithm to solve $\max_{x^{(1)}} \min_{x^{(2)}} L$, i.e.,

$$\dot{x} = \text{diag}\left\{P_{x^{(1)}}, -P_{x^{(2)}}\right\} \frac{\partial L}{\partial x}.$$

In the above definition, state x is partitioned into $x^{(1)}$ and $x^{(2)}$. Similarly, (5.25) is rearranged in the following form (in this subsection, we abuse notation e.g., A_{ij}, n_i, λ_i, B_i, C_i, whose meaning should be clear from the context)

$$\underbrace{\begin{bmatrix} \dot{x}^{(1)} \\ \dot{x}^{(2)} \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_{A} x \tag{5.28}$$

where $x^{(1)}(t) \in \mathbb{R}^{n_1}$, $x^{(2)}(t) \in \mathbb{R}^{n_2}$, and $n_1 + n_2 = n$.

Using similar arguments as before, we only need to focus on a function L in a quadratic form, i.e.,

$$L = \frac{1}{2} x^T \underbrace{\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}}_{Q} x \tag{5.29}$$

where $Q_{11} \in \mathbb{R}^{n_1\times n_1}$ satisfies $Q_{11} = Q_{11}^T \preceq 0$ (i.e., L is concave in $x^{(1)}$), $Q_{22} \in \mathbb{R}^{n_2\times n_2}$ satisfies $Q_{22} = Q_{22}^T \succeq 0$ (i.e., L is convex in $x^{(2)}$), and $Q_{12} \in \mathbb{R}^{n_1\times n_2}$. Based on the definition of Class-$\mathcal{S}$, we have the following theorem.

Theorem 5.5 *System (5.28) belongs to Class-$\mathcal{S}$ if and only if the following conditions are satisfied: (i) (5.28) is marginally or asymptotically stable; (ii) the eigenvalues of both A_{11} and A_{22} are non-positive real; (iii) both A_{11} and A_{22} are diagonalizable with the diagonal canonical forms given by $A_{11} = J_1\Lambda_1 J_1^{-1}, J_1 \in \mathbb{R}^{n_1\times n_1}, A_{22} = J_2\Lambda_2 J_2^{-1}, J_2 \in \mathbb{R}^{n_2\times n_2}$, and there exist V_1 and V_2 such that*

$$(J_1^{-1})^T V_1 J_1^{-1} A_{12} + A_{21}^T (J_2^{-1})^T V_2 J_2^{-1} = \mathbf{0} \tag{5.30a}$$

$$V_1\Lambda_1 = \Lambda_1 V_1,\ V_2\Lambda_2 = \Lambda_2 V_2 \tag{5.30b}$$

$$V_1 \succ 0,\ V_2 \succ 0. \tag{5.30c}$$

Proof: Necessity. Suppose that system (5.28) is a primal-dual gradient algorithm to solve $\max_{x^{(1)}} \min_{x^{(2)}} L$ where L is given by (5.29). Then the trajectories of (5.28) are bounded under Lemma 5.2, and there exist matrices $P_{x^{(1)}} \in \mathbb{R}^{n_1\times n_1}$ and $P_{x^{(2)}} \in \mathbb{R}^{n_2\times n_2}$ satisfying $P_{x^{(1)}} = P_{x^{(1)}}^T \succ 0$ and $P_{x^{(2)}} = P_{x^{(2)}}^T \succ 0$, so that $Q = \text{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}\}^{-1}A = A^T\text{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}\}^{-1}$ holds. This equation leads to

$$P_{x^{(1)}}^{-1} A_{11} = A_{11}^T P_{x^{(1)}}^{-1} = Q_{11} \preceq 0 \tag{5.31a}$$

$$P_{x^{(2)}}^{-1} A_{22} = A_{22}^T P_{x^{(2)}}^{-1} = -Q_{22} \preceq 0 \tag{5.31b}$$

$$P_{x^{(1)}}^{-1} A_{12} + A_{21}^T P_{x^{(2)}}^{-1} = \mathbf{0}. \tag{5.31c}$$

Based on Lemma 5.5, (5.31a)-(5.31b) are equivalent to $\text{eig}(A_{11}) \in \mathbb{R}$, A_{11} is diagonalizable, $\text{eig}(A_{22}) \in \mathbb{R}$, A_{22} is diagonalizable, and moreover, the eigenvalues of both A_{11} and A_{22} are non-positive (otherwise, for example, there would exist a positive eigenvalue of A_{11} and a corresponding non-zero eigenvector, denoted by λ_1 and x_{λ_1}, for which $x_{\lambda_1}^T P_{x^{(1)}}^{-1} A_{11} x_{\lambda_1} = \lambda_1 x_{\lambda_1}^T P_{x^{(1)}}^{-1} x_{\lambda_1} > 0$, i.e., a contradiction to (5.31a)). By defining $V_1 = J_1^T P_{x^{(1)}}^{-1} J_1$, $V_2 = J_2^T P_{x^{(2)}}^{-1} J_2$, condition (iii) holds, which completes the proof of necessity.

Sufficiency. Let conditions (i)-(iii) be true. Consider the following unconstrained quadratic saddle point problem:

$$\max_{x^{(1)}\in\mathbb{R}^{n_1}} \min_{x^{(2)}\in\mathbb{R}^{n_2}} L = \frac{1}{2} x^T P^{-1} A x \tag{5.32}$$

where $P^{-1} = \text{diag}\{(J_1^{-1})^T V_1 J_1^{-1}, -(J_2^{-1})^T V_2 J_2^{-1}\}$. Due to $V_1\Lambda_1 \preceq 0, V_2\Lambda_2 \preceq 0$, L is concave in $x^{(1)}$ and convex in $x^{(2)}$. Define matrices $P_{x^{(1)}} = J_1 V_1^{-1} J_1^T \succ 0$, $P_{x^{(2)}} = J_2 V_2^{-1} J_2^T \succ 0$. Under Lemma 5.2, the trajectories of the primal-dual gradient algorithm given by $\dot{x} = \text{diag}\left\{P_{x^{(1)}}, -P_{x^{(2)}}\right\} \frac{\partial L}{\partial x}$ are bounded, which is the same as (5.28). So we conclude that (5.28) belongs to Class-$\mathcal{S}$. ■

Theorem 5.5 establishes a necessary and sufficient condition to reverse-engineer (5.28) as a primal-dual gradient algorithm to solve an unconstrained quadratic saddle point problem. One of the conditions is to check the feasibility of (5.30). This problem is actually a Semi-Definite Programming (SDP) problem which can be solved using SDP solvers, e.g., SeDuMi [99]. Similar to Remark 5.3, adding a minus sign before L leads to an alternative formulation of (5.32) in which $x^{(1)}$ appears a minimizer and $x^{(2)}$ appears a maximizer.

Remark 5.5 *Table 5.1 provides examples on how to reverse-engineer LTI autonomous systems using Theorems 5.4 and 5.5. Note that there can be multiple ways to reverse-engineer a given system (compare line 1 and 3 in the table).*

Table 5.1: Examples of reverse-engineering LTI autonomous systems.

A	P^{-1} or $\text{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}\}^{-1}$	$Q = P^{-1}A$	Theorem
$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	5.4
$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$	5.4
$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	5.5
$\begin{bmatrix} -2 & -1 \\ 1 & -2 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -2 & -1 \\ -1 & 2 \end{bmatrix}$	5.5
$\begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -2 & -1 \\ -1 & 0 \end{bmatrix}$	5.5
$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	5.5
$\begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	5.5
$\begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$	5.5

2): LTI systems with inputs

In this subsection, we extend the results in Theorems 5.4-5.5 to LTI systems with inputs. Consider the following system

$$\dot{x} = Ax + Bu + Cw \tag{5.33}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $u(t) \in \mathbb{R}^m$ is the control input vector, $C \in \mathbb{R}^{n\times p}$, and $w(t) \in \mathbb{R}^p$ is the exogenous input vector, e.g., disturbance injection.

Remark 5.6 *A given LTI closed-loop system with either static feedback or dynamic feedback can be rearranged to fit (5.33). More specifically, system (5.15)-(5.16) can be rearranged as:*

$$\underbrace{\begin{bmatrix} \dot{x} \\ \dot{u} \end{bmatrix}}_{\dot{\tilde{x}}} = \underbrace{\begin{bmatrix} A & B \\ D & E \end{bmatrix}}_{\tilde{A}} \tilde{x} + \underbrace{\begin{bmatrix} C \\ F \end{bmatrix}}_{\tilde{C}} w$$

where u is included in the augmented state vector $\tilde{x}$.

The following result follows from Theorem 5.4.

Theorem 5.6 *Let u, w be constant in system (5.33) and the set $\{x|Ax+Bu+Cw = \mathbf{0}\}$ be nonempty. System (5.33) belongs to Class-$\mathcal{O}$ if and only if (5.33) is marginally or asymptotically stable, eig$(A) \in \mathbb{R}$ and A is diagonalizable.*

Proof: The proof is similar to that of Theorem 5.4, by replacing problem (5.27) with the following unconstrained convex quadratic programming problem:

$$\max_{x\in\mathbb{R}^n} L(x) = \frac{1}{2}x^T P^{-1}Ax + x^T P^{-1}(Bu + Cw)$$

where P is positive definite satisfying $P^{-1}A = A^T P^{-1}$. ■

In the above theorem, we have extended the definition of Class-$\mathcal{O}$ to include linear systems with inputs, which is straightforward. Similar to (5.28), let us partition system (5.33) in the form

$$\underbrace{\begin{bmatrix} \dot{x}^{(1)} \\ \dot{x}^{(2)} \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_{A} x + Bu + Cw \tag{5.34}$$

where $x^{(1)}(t) \in \mathbb{R}^{n_1}$, $x^{(2)}(t) \in \mathbb{R}^{n_2}$ and $n_1 + n_2 = n$. We then obtain the following theorem.

Theorem 5.7 *Let u, w be constant in system (5.34) and the set $\{x|Ax+Bu+Cw = \mathbf{0}\}$ be nonempty. System (5.34) belongs to Class-$\mathcal{S}$ if and only if the following conditions are satisfied: (i) (5.34) is marginally or asymptotically stable; (ii) the eigenvalues of both A_{11} and A_{22} are non-positive real; (iii) both A_{11} and A_{22} are diagonalizable with the diagonal canonical forms given by $A_{11} = J_1\Lambda_1 J_1^{-1}, J_1 \in \mathbb{R}^{n_1\times n_1}, A_{22} = J_2\Lambda_2 J_2^{-1}, J_2 \in \mathbb{R}^{n_2\times n_2}$, and there exist V_1 and V_2 such that (5.30) holds.*

Proof: The proof is similar to that of Theorem 5.5, by replacing problem (5.32) with the following unconstrained quadratic saddle point problem:

$$\max_{x^{(1)}\in\mathbb{R}^{n_1}} \min_{x^{(2)}\in\mathbb{R}^{n_2}} L(x) = \frac{1}{2}x^T P^{-1}Ax + x^T P^{-1}(Bu + Cw)$$

where $P^{-1} = \text{diag}\{(J_1^{-1})^T V_1 J_1^{-1}, -(J_2^{-1})^T V_2 J_2^{-1}\}$. ■

In the above theorem, we have extended the definition of Class-$\mathcal{S}$ to include linear systems with inputs. Because Class-$\mathcal{S}'$ is a subset of Class-$\mathcal{S}$, we can apply the results to system (5.15)-(5.16) and Class-$\mathcal{S}'$. Partition (5.15) in the form (5.34) and rearrange (5.15)-(5.16) as

$$\underbrace{\begin{bmatrix} \dot{x}^{(1)} \\ \dot{x}^{(2)} \\ \dot{u} \end{bmatrix}}_{\dot{\tilde{x}}} = \underbrace{\begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ D_1 & D_2 & E \end{bmatrix}}_{\tilde{A}} \tilde{x} + \underbrace{\begin{bmatrix} C_1 \\ C_2 \\ F \end{bmatrix}}_{\tilde{C}} w \tag{5.35}$$

where $x^{(1)}(t) \in \mathbb{R}^{n_1}$, $x^{(2)}(t) \in \mathbb{R}^{n_2}$, $n_1 + n_2 = n$. Then the following proposition is immediate.

Proposition 5.3 *Let w be constant in system (5.35) and the set $\{\tilde{x}|\tilde{A}\tilde{x}+\tilde{C}w = \mathbf{0}\}$ be nonempty. System (5.35) belongs to Class-$\mathcal{S}'$ if and only if the following conditions are satisfied: (i) (5.35) is marginally or asymptotically stable; (ii) the eigenvalues of A_{11}, A_{22}, E, $E_{ii}, i = 1, \cdots, N$ and $\begin{bmatrix} A_{22} & B_2 \\ D_2 & E \end{bmatrix}$ are non-positive real, and these matrices are diagonalizable; (iii) let the diagonal canonical forms of A_{11}, A_{22}, $E_{ii}, i = 1, \cdots, N$ be $A_{11} = J_1\Lambda_1 J_1^{-1}, J_1 \in \mathbb{R}^{n_1\times n_1}$, $A_{22} = J_2\Lambda_2 J_2^{-1}, J_2 \in \mathbb{R}^{n_2\times n_2}$, $E_{ii} = J_{E_i}\Lambda_{E_i} J_{E_i}^{-1}, J_{E_i} \in \mathbb{R}^{m_i\times m_i}$, there exist V_1, V_2, V_{E_i} such that*

$$(J_1^{-1})^T V_1 J_1^{-1} A_{12} + A_{21}^T (J_2^{-1})^T V_2 J_2^{-1} = \mathbf{0} \tag{5.36a}$$

$$(J_1^{-1})^T V_1 J_1^{-1} B_1 + D_1^T \, diag\{(J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1}\} = \mathbf{0} \tag{5.36b}$$

$$(J_2^{-1})^T V_2 J_2^{-1} B_2 - D_2^T \, diag\{(J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1}\} = \mathbf{0} \tag{5.36c}$$

$$(J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1} E_{ij} - E_{ji}^T (J_{E_j}^{-1})^T V_{E_j} J_{E_j}^{-1} = \mathbf{0}, \ j \in \mathcal{N}(i), i = 1, \cdots, N \tag{5.36d}$$

$$V_1\Lambda_1 = \Lambda_1 V_1, \ V_2\Lambda_2 = \Lambda_2 V_2, \ V_{E_i}\Lambda_{E_i} = \Lambda_{E_i} V_{E_i}, \ i = 1, \cdots, N \tag{5.36e}$$

$$V_1 \succ 0, \ V_2 \succ 0, \ V_{E_i} \succ 0, \ i = 1, \cdots, N. \tag{5.36f}$$

Proof: Necessity. Suppose that system (5.35) belongs to Class-$\mathcal{S}'$ and is a primal-dual gradient algorithm to solve $\max_{x^{(1)}} \min_{x^{(2)},u} L_{\rm sys}$ and

$$L_{\rm sys} = \frac{1}{2}\tilde{x}^T \underbrace{\left[\begin{array}{cc} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{array} \right]}_{Q} \tilde{x} + \tilde{x}^T P^{-1} \tilde{C} w$$

where $Q_{11} \in \mathbb{R}^{n_1 \times n_1}$ satisfies $Q_{11} = Q_{11}^T \preceq 0$ (i.e., $L_{\rm sys}$ is concave in $x^{(1)}$), $Q_{22} \in \mathbb{R}^{(n_2+m)\times(n_2+m)}$ satisfies $Q_{22} = Q_{22}^T \succeq 0$ (i.e., $L_{\rm sys}$ is convex in $(x^{(2)}, u)$), $Q_{12} \in \mathbb{R}^{n_1 \times (n_2+m)}$ and $P \in \mathbb{R}^{(n_1+n_2+m)\times(n_1+n_2+m)}$. Then the trajectories of (5.35) are bounded under Lemma 5.2, and there exist matrices $P_{x^{(1)}} \in \mathbb{R}^{n_1 \times n_1}$, $P_{x^{(2)}} \in \mathbb{R}^{n_2 \times n_2}$ and $P_{u_i} \in \mathbb{R}^{m_i \times m_i}$ satisfying $P_{x^{(1)}} = P_{x^{(1)}}^T \succ 0$, $P_{x^{(2)}} = P_{x^{(2)}}^T \succ 0$ and $P_{u_i} = P_{u_i}^T \succ 0, i = 1, \cdots, N$, so that the equality $Q = \mathrm{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}, -\mathrm{diag}\{P_{u_i}\}\}^{-1} \tilde{A} = \tilde{A}^T \mathrm{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}, -\mathrm{diag}\{P_{u_i}\}\}^{-1}$ holds (let $P = \mathrm{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}, -\mathrm{diag}\{P_{u_i}\}\}$). This leads to

$$\begin{gathered} P_{x^{(1)}}^{-1} A_{11} = A_{11}^T P_{x^{(1)}}^{-1} = Q_{11} \preceq 0 \\ \left[\begin{array}{cc} P_{x^{(2)}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}\{P_{u_i}^{-1}\} \end{array} \right] \left[\begin{array}{cc} A_{22} & B_2 \\ D_2 & E \end{array} \right] = \left[\begin{array}{cc} A_{22}^T & D_2^T \\ B_2^T & E^T \end{array} \right] \left[\begin{array}{cc} P_{x^{(2)}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}\{P_{u_i}^{-1}\} \end{array} \right] = -Q_{22} \preceq 0 \\ P_{x^{(1)}}^{-1} \left[\begin{array}{cc} A_{12} & B_1 \end{array} \right] = - \left[\begin{array}{cc} A_{21}^T & D_1^T \end{array} \right] \left[\begin{array}{cc} P_{x^{(2)}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}\{P_{u_i}^{-1}\} \end{array} \right] \end{gathered}$$

which can be further rearranged as

$$\begin{gathered} P_{x^{(1)}}^{-1} A_{11} = A_{11}^T P_{x^{(1)}}^{-1} \preceq 0 \\ P_{x^{(2)}}^{-1} A_{22} = A_{22}^T P_{x^{(2)}}^{-1} \preceq 0 \\ P_{x^{(2)}}^{-1} B_2 = D_2^T \mathrm{diag}\{P_{u_i}^{-1}\} \\ \mathrm{diag}\{P_{u_i}^{-1}\} \left[\begin{array}{ccc} E_{11} & \cdots & E_{1N} \\ \cdots & \cdots & \cdots \\ E_{N1} & \cdots & E_{NN} \end{array} \right] = \left[\begin{array}{ccc} E_{11}^T & \cdots & E_{N1}^T \\ \cdots & \cdots & \cdots \\ E_{1N}^T & \cdots & E_{NN}^T \end{array} \right] \mathrm{diag}\{P_{u_i}^{-1}\} \preceq 0 \\ P_{x^{(1)}}^{-1} A_{12} + A_{21}^T P_{x^{(2)}}^{-1} = \mathbf{0} \\ P_{x^{(1)}}^{-1} B_1 + D_1^T \mathrm{diag}\{P_{u_i}^{-1}\} = \mathbf{0}. \end{gathered}$$

Based on Lemma 5.5, the above equations are equivalent to conditions (ii) and (iii) by defining $V_1 = J_1^T P_{x^{(1)}}^{-1} J_1, V_2 = J_2^T P_{x^{(2)}}^{-1} J_2, V_{E_i} = J_{E_i}^T P_{u_i}^{-1} J_{E_i}, i = 1, \cdots, N$.

Sufficiency. Let conditions (i)-(iii) be true. Consider the following unconstrained quadratic saddle point problem:

$$\max_{x^{(1)} \in \mathbb{R}^{n_1}} \min_{x^{(2)} \in \mathbb{R}^{n_2}, u \in \mathbb{R}^m} L_{\rm sys} = \frac{1}{2}\tilde{x}^T P^{-1} \tilde{A} \tilde{x} + \tilde{x}^T P^{-1} \tilde{C} w$$

where $P^{-1} = \text{diag}\{(J_1^{-1})^T V_1 J_1^{-1}, -(J_2^{-1})^T V_2 J_2^{-1}, -\text{diag}\{(J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1}\}\}$. Due to $V_1\Lambda_1 \preceq 0$ and

$$\begin{aligned}&\begin{bmatrix} (J_2^{-1})^T V_2 J_2^{-1} & \mathbf{0} \\ \mathbf{0} & \text{diag}\{J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1}\} \end{bmatrix} \begin{bmatrix} A_{22} & B_2 \\ D_2 & E \end{bmatrix} \\ =&\begin{bmatrix} A_{22}^T & D_2^T \\ B_2^T & E^T \end{bmatrix} \begin{bmatrix} (J_2^{-1})^T V_2 J_2^{-1} & \mathbf{0} \\ \mathbf{0} & \text{diag}\{J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1}\} \end{bmatrix} \preceq 0 \end{aligned} \tag{5.37}$$

L_{sys} is concave in $x^{(1)}$ and convex in $(x^{(2)}, u)$. Equation (5.37) is derived as follow. Let $\tilde{A}_{22} = \begin{bmatrix} A_{22} & B_2 \\ D_2 & E \end{bmatrix}$ and $\tilde{P}_2^{-1} = \begin{bmatrix} (J_2^{-1})^T V_2 J_2^{-1} & \mathbf{0} \\ \mathbf{0} & \text{diag}\{J_{E_i}^{-1})^T V_{E_i} J_{E_i}^{-1}\} \end{bmatrix}$. Since the eigenvalues of $\tilde{A}_{22}$ are non-positive real and $\tilde{A}_{22}$ is diagonalizable, $\tilde{A}_{22}$ can be written as $\tilde{A}_{22} = \tilde{J}\tilde{\Lambda}\tilde{J}^{-1}$ where $\tilde{\Lambda} \preceq 0$ is a diagonal matrix. Due to $\tilde{P}_2^{-1}\tilde{J}\tilde{\Lambda}\tilde{J}^{-1} = (\tilde{J}^{-1})^T\tilde{\Lambda}\tilde{J}^T\tilde{P}_2^{-1}$, based on Lemma 1 in [72], $\tilde{J}^T\tilde{P}_2^{-1}\tilde{J}\tilde{\Lambda} = \tilde{\Lambda}\tilde{J}^T\tilde{P}_2^{-1}\tilde{J} \preceq 0$ holds. For any given vector $v \in \mathbb{R}^{n_2+m}$, $(\tilde{J}v)^T\tilde{P}_2^{-1}\tilde{J}\tilde{\Lambda}\tilde{J}^{-1}(\tilde{J}v) = v^T\tilde{J}^T\tilde{P}_2^{-1}\tilde{J}\tilde{\Lambda}v \leq 0$, which leads to $\tilde{P}_2^{-1}\tilde{J}\tilde{\Lambda}\tilde{J}^{-1} \preceq 0$, i.e., Equation (5.37) holds.

Now define matrices $P_{x^{(1)}} = J_1 V_1^{-1} J_1^T \succ 0$, $P_{x^{(2)}} = J_2 V_2^{-1} J_2^T \succ 0$ and $P_{u_i} = J_{E_i} V_{E_i}^{-1} J_{E_i}^T \succ 0, i = 1, \cdots, N$. Under Lemma 1, the trajectories of the primal-dual gradient algorithm given by $\dot{x} = \text{diag}\{P_{x^{(1)}}, -P_{x^{(2)}}\}\frac{\partial L_{\text{sys}}}{\partial x}$ and $\dot{u}_i = -P_{u_i}\frac{\partial L_{\text{sys}}}{\partial u_i}, i = 1, \cdots, N$ are bounded, which is the same as (5.35). So we conclude that system (5.35) belongs to Class-$\mathcal{S}'$. ■

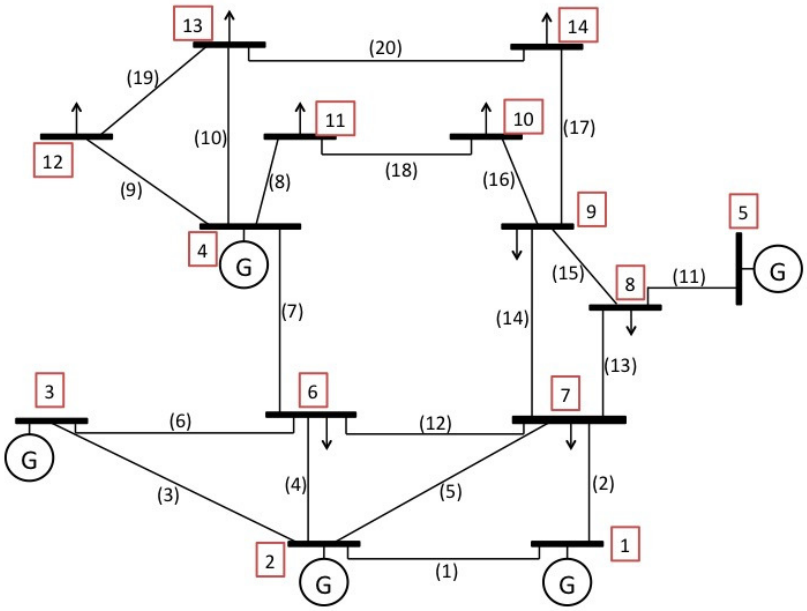

Figure 5.1: IEEE 14-bus network, with 5 generator buses, 9 load buses and 20 transmission lines. The buses are numbered with red blocks. The transmission lines are numbered with brackets.

5.3 Numerical Investigations

We now present numerical investigations using the IEEE 14-bus network illustrated in Figure 5.1. The parameter of the overall system (5.11) are shown in Table 5.2.

Table 5.2: Parameters of the overall system (p.u. means per unit).

Parameter	Value(p.u.)	Parameter	Value(p.u.)
M_i	$10, 11, 9, 8, 12$	$D_i, i \in \mathcal{G}$	$2, 1.2, 1.5, 0.8, 1$
$D_i, i \in \mathcal{L}$	$0.8, 0.9, 1, 1.1, 1.2, 0.7, 0.3, 1.4, 0.6$	T_{TG_i}	$5, 6, 3, 4, 7$
R	$0.05I_5$	γ	2

Parameter	Value(p.u.)
T_k	$24, 25, 23, 26, 27, 22, 24, 25, 23, 26, 27, 22, 24, 25, 23, 26, 27, 22, 20, 30$
$P_{C_i}^{\max}, -P_{C_i}^{\min}$	$P_{C_i}^{\max} = -P_{C_i}^{\min} : 0.3, 0.4, 0.2, 0.1, 0.5$
$P_{L_i}^{\max}, -P_{L_i}^{\min}$	$P_{L_i}^{\max} = -P_{L_i}^{\min} : 0.2, 0.2, 0.3, 0.3, 0.5, 0.4, 0.4, 0.1, 0.1$
$P_{TC_k}^{\max}, -P_{TC_k}^{\min}$	All $P_{TC_k}^{\max} = -P_{TC_k}^{\min} = 1$ except $P_{TC_7}^{\max} = -P_{TC_7}^{\min} = 0.5$, $P_{TC_{13}}^{\max} = -P_{TC_{13}}^{\min} = 0.3$ and $P_{TC_{14}}^{\max} = -P_{TC_{14}}^{\min} = 0.8$

Function	Expression
$C_i(P_{C_i})$	$C_i(P_{C_i}) = \frac{c_i}{2} P_{C_i}^2$, $c_i = 0.01, 0.02, 0.03, 0.04, 0.05$
$U_i(P_{L_i})$	$U_i(P_{L_i}) = -\frac{c_i}{2} P_{L_i}^2$, $c_i = 0.1, 0.2, 0.05, 0.08, 0.06, 0.09, 0.12, 0.15, 0.07$

Gain	Value(p.u.)	Gain	Value(p.u.)
K_{P_C}	$15I_5$	K_{P_L}	$15I_9$
K_θ	$10I_{14}$	K_ζ	$10I_5$
K_λ	$10I_9$	K_{μ^+}, K_{μ^-}	$10I_5$
K_{ν^+}, K_{ν^-}	$10I_9$	K_{l^+}, K_{l^-}	$10I_{20}$

We consider a scenario consisting of two disturbances which is realized as follows: the system is stabilized at the nominal operating point at $t = 0$s; at $t = 5$s, there is a step change $d_i = +2$p.u. at bus 6; 40s later, there is a step change $d_i = -2.5$p.u. at bus 9. The simulation results are shown in Figures 5.2-5.6. Compared to the uncontrolled case (i.e., without any control, shown in figures on the left), we can see that the proposed control scheme restores the grid frequency quickly under disturbances. Moreover, it can be verified that the optimization problem is solved when system reaches steady state, i.e., generation cost is minimized within capacity limit, load utility is maximized within capacity limit, and transmission line congestion constraints are satisfied.

On the other hand, to fit the framework proposed in Section 5.2, we can regard a decentralized frequency integral control scheme given by

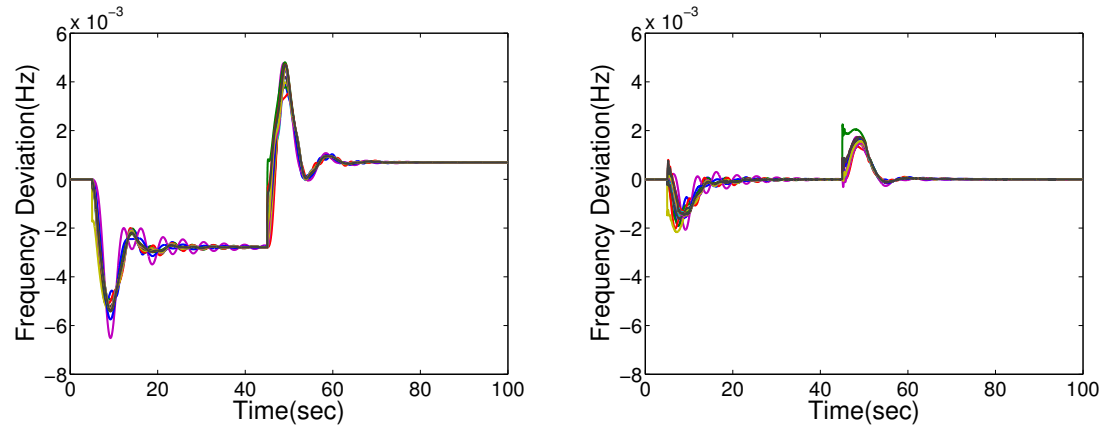

Figure 5.2: Bus frequency (deviations). Left: uncontrolled case. Right: controlled case.

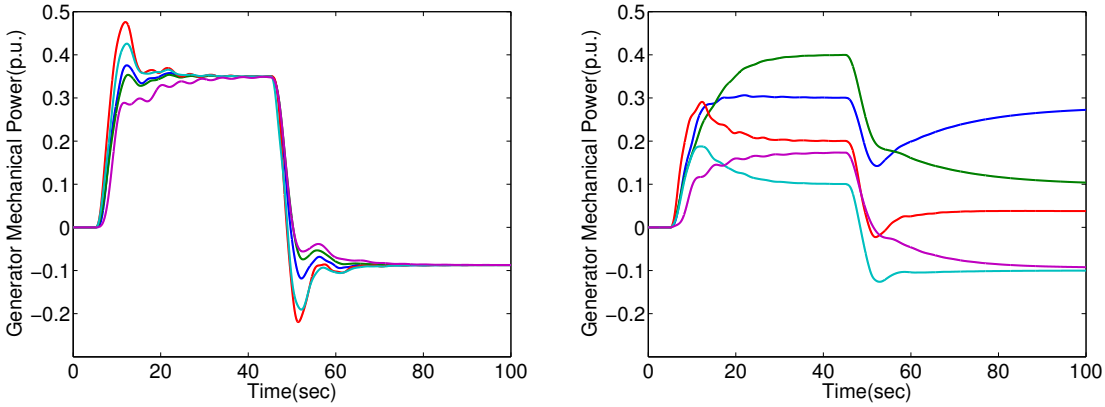

Figure 5.3: Generator mechanical power (deviations). Left: uncontrolled case. Right: controlled case.

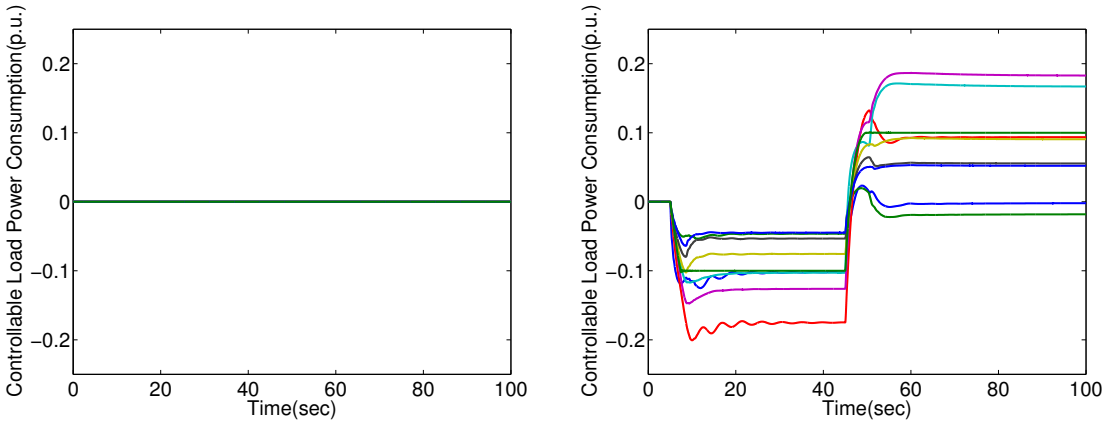

Figure 5.4: Controllable load power consumption (deviations). Left: uncontrolled case. Right: controlled case.

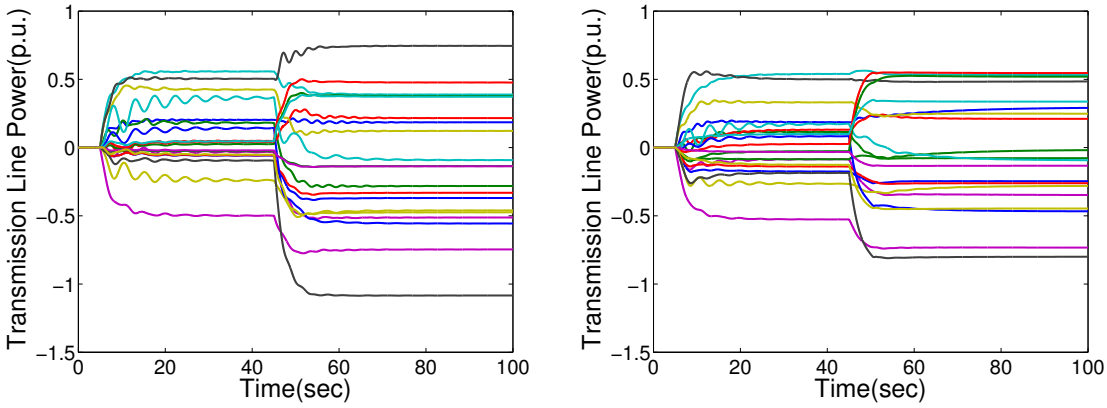

Figure 5.5: Transmission line power flow (deviations). Left: uncontrolled case. Right: controlled case.

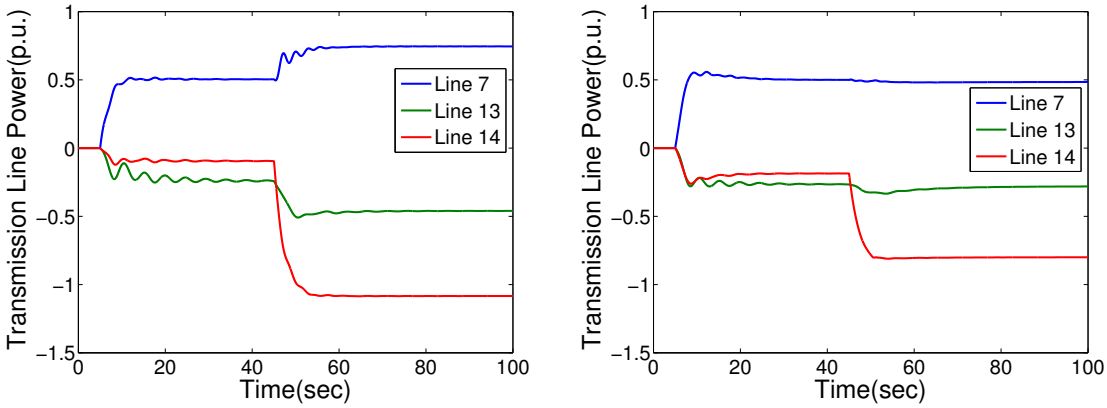

Figure 5.6: Power flow (deviations) through line 7 (blue), 13 (green), 14 (red). Left: uncontrolled case. Right: controlled case.

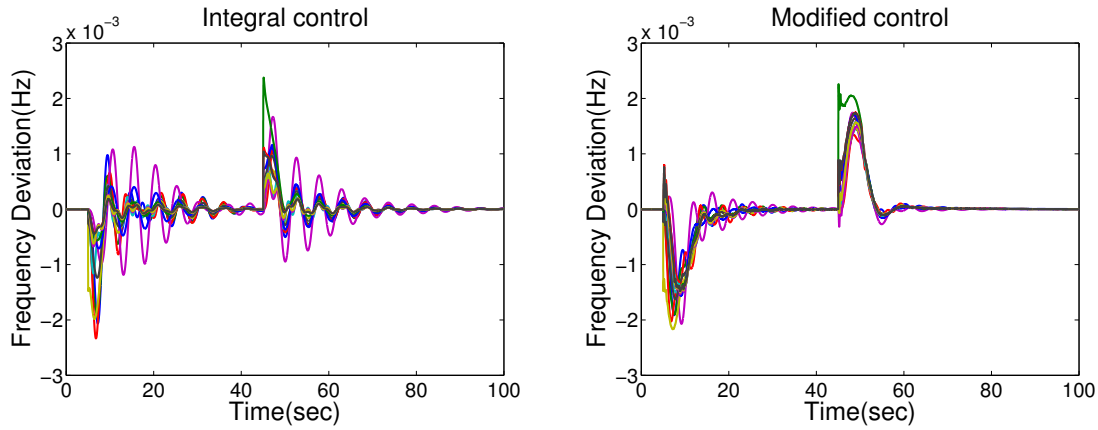

Figure 5.7: Bus frequency (deviations). Left: using integral control. Right: using modified control.

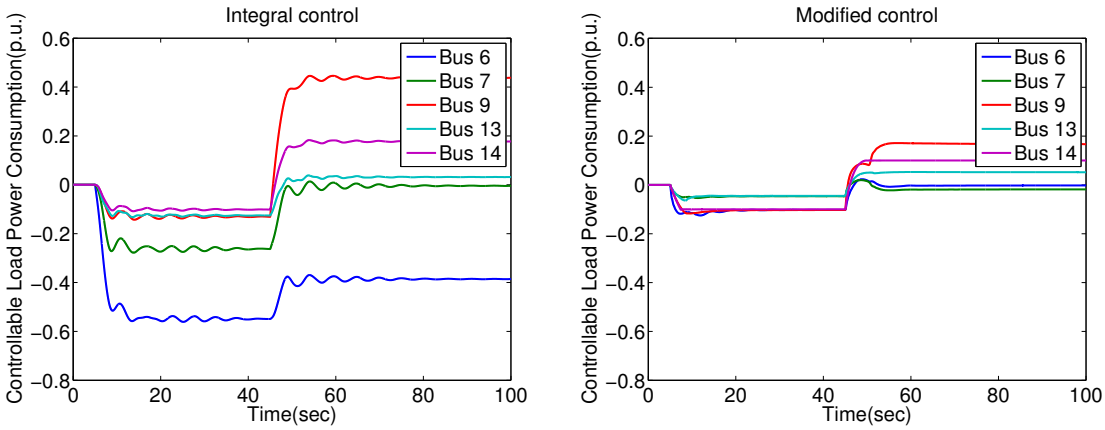

Figure 5.8: Controllable load power consumption (deviations) at selected buses. Left: using integral control. Right: using modified control.

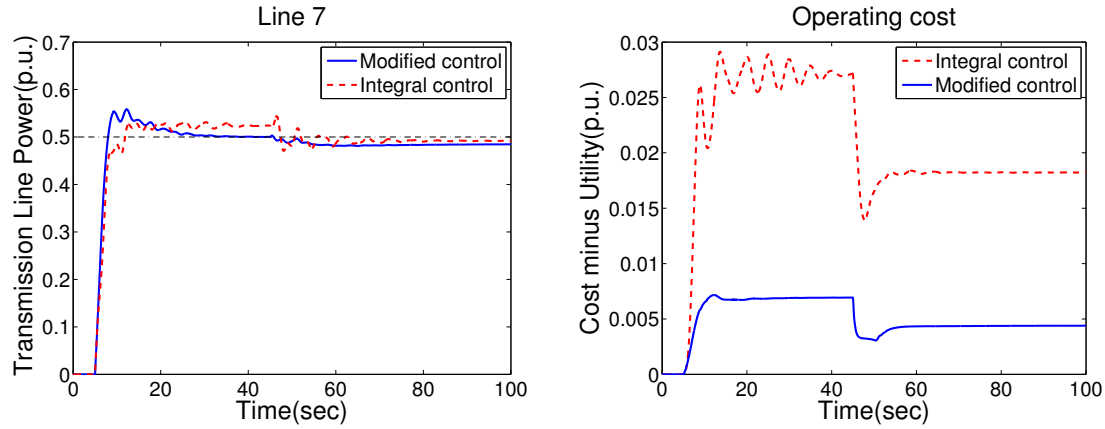

Figure 5.9: Left: power flow (deviation) in line 7 (the black dashdot line is the capacity upper bound of line 7 in the OPF problem). Right: operating cost comparison.

$$\dot{P}_C = K_{P_C}(R(P_M - P_C))$$
$$\dot{P}_L = K_{P_L}(D_l^{-1}(-P_L - d_l - \Gamma_2 A T A^T \alpha))$$

as a built-in control mechanism, and apply the proposed procedure to redesign the closed-loop system that belongs to Class-$\mathcal{S}'$. Under the same scenario and conditions, we compare the decentralized frequency integral controller with the redesigned controller (i.e., Equations (5.11d)-(5.11n)). The simulation results are shown in Figures 5.7-5.9. Compared with the case of using only integral control, the modified control scheme leads to less oscillations and less cost, and moreover, satisfies operating constraints (in the case of using integral control, load power consumption at bus 6, 7, 9, 13, 14 does not meet the capacity constraints, and the power flow in line 7

does not meet its capacity constraint).

5.4 Conclusion

In this chapter, we have studied a reverse- and forward-engineering framework for distributed control of a class of linear network systems to achieve optimal steady-state performance. This framework consists of two steps: firstly, seek an appropriate optimization problem that the system dynamics implicitly solve (*reverse-engineering*); secondly, modify the resulting optimization problem by incorporating a predefined optimization problem and derive control mechanisms to solve the augmented optimization problem (*forward-engineering*). In the first half the chapter, we used the proposed framework to solve the real-time economic dispatch problem in power systems. Under exogenous disturbances, the distributed controller derived through this framework asymptotically stabilized the power network at an equilibrium point which solved the steady-state optimization problem (5.2)/(5.3). In the second half of the chapter, we provided a general procedure to modify control schemes for a special class of network systems. In order to investigate how general the reverse- and forward-engineering framework was, we developed necessary and sufficient conditions under which an LTI system could be reverse-engineered as a gradient algorithm to solve either (i) an unconstrained convex quadratic programming problem or (ii) an unconstrained quadratic saddle point problem. These conditions were characterized using properties of system matrices and relevant linear matrix inequalities. Finally, a numerical example using the IEEE 14-bus network was presented to demonstrate the effectiveness of the proposed framework.

Chapter 6

Conclusions

In this thesis, we have developed several scalable frameworks to modify/redesign a class of large-scale network systems with built-in control mechanisms to improve their economic efficiency and performance while guaranteeing their stability and robustness. These frameworks were mainly applied to electric power systems and network congestion control. As a result, in the presence of uncertainties, the redesigned systems can track the efficient operating points automatically in a distributed or decentralized, and closed-loop manner. Before closing, we provide a summary of the main results and briefly discuss future research directions.

6.1 Summary

In Chapter 2, we have proposed modifications in the generation control in the power grid to improve its economic efficiency, stability and robustness. Moreover, the participation of controllable loads has been considered. We first obtained the state-space description of a conventional power network model which described system dynamics around a nominal operating state, and formulated an optimization problem relating to active power regulation under exogenous disturbances. We then proposed a new control scheme via a consensus approach and studied its optimality, stability and delay robustness. Moreover, we extended the designed control scheme to (i) networks with more complexity and (ii) the case where controllable loads were involved in the optimization problem. Finally, numerical examples showed that the proposed controllers could balance the power flow in the network quickly, and drive the system to an economically optimal operating point in the steady state.

In Chapter 3, a real-time control framework that merges conventional primary, secondary and tertiary frequency control in power systems has been studied. In particular, we considered a transmission level network with tree topology, and formulated

an Optimal Power Flow (OPF) problem with constraints containing the dynamics of the power network. We then used a primal-dual decomposition approach to design a distributed dynamic feedback controller and showed the stability of the overall system. In addition, we introduced extra dynamics to improve system behaviour, where we emphasized the trade-off when choosing the gains of the extra dynamics. Numerical experiments showed that the proposed controller could balance the power flow in the network quickly, and achieve OPF in the steady state. Furthermore, after adding the extra dynamics, the transient performance of the system improved significantly.

In Chapter 4, we presented a redesign framework for network congestion control problems. Motivated by the augmented Lagrangian method, we introduced extra terms to the Lagrangian, which was then used to redesign the primal-dual, primal and dual algorithms. We investigated how the gains resulting from the extra dynamics influenced the stability and robustness of the system. Moreover, we showed that the overall system could achieve added robustness to communication delays by appropriately tuning these gains. We next studied the meaning of the extra dynamics in the redesign and developed a distributed proportional-integral-derivative controller for solving network congestion control problems. Finally, we illustrated that compared to standard algorithms, the modified algorithms resulted in better transient performance and improved robustness, without changing the distributed structure of the feedback.

In Chapter 5, a reverse- and forward-engineering framework for distributed control of a class of linear network systems to achieve optimal steady-state performance was developed. In the first half of this chapter, we used the proposed framework to solve the real-time economic dispatch problem in power systems. In the second half, we provided a general procedure to modify control schemes for a special class of network systems. In order to investigate how general the reverse- and forward-engineering framework was, we developed necessary and sufficient conditions under which an Linear Time-Invariant (LTI) system could be reverse-engineered as a gradient algorithm to solve either (i) an unconstrained convex quadratic programming problem or (ii) an unconstrained quadratic saddle point problem. Finally, a numerical example was presented to demonstrate the effectiveness of the proposed framework.

6.2 Future Research Directions

We would like to mention a few future research directions based on the work presented in this thesis.

Firstly, an immediate research direction is to extend the redesign framework presented in Chapter 3 to power networks with mesh topology. Mesh networks are harder since the corresponding OPF problem is non-convex and nonlinear. But recent work by Low [100, 101] has suggested that tunable phase shifters can be introduced to convexify the (AC) OPF problem, which will assist in extending our redesign framework. Note that after introducing phase shifters, their dynamics should be considered. Also, a further step is to include voltage dynamics and (re)design real-time voltage control/reaction power regulation schemes.

Secondly, Chapters 2 and 5 have considered frequency control/active power regulation in the power grid through merging primary and secondary control. It would be interesting to investigate the problem of voltage control under the same scenario. Since the physical coupling between active power flows and bus voltage magnitudes, and between reactive power flows and bus angle differences is weak in transmission level networks, a standard decoupling approximation can be adopted [3, 7] when linearizing system dynamics around a nominal operating state. This will simplify the design of real-time voltage control schemes.

The effect of time delay in the redesigned frequency control schemes is only considered in Chapter 2. Future work will focus on the control schemes proposed in Chapters 3 and 5 with time delays. According to Chapter 4, introducing extra dynamics can lead to added delay robustness. We are therefore interested in investigating the choice of controller gains to achieve improved delay robustness for the overall system.

We are also interested in more accurate models of exogenous input in power system control and optimization, i.e., fluctuations resulting from distributed energy resources and variability in both supply and demand. The exogenous input considered in this thesis is only constant or slow varying. In the future, we will consider the stochastic effects.

Also, note that if gradient-based methods are applied on an optimization problem to design control schemes, the proposed redesign framework in Chapter 4 can then be used to improve their properties. Moreover, a different choice of the extra terms (e.g., $h(z,\hat{z},k) = h(x_i,\hat{x}_i,\tilde{x}_i,k_{1e_i},k_{2e_i}) = -\sum_{i=1}^{N}\left(\frac{k_{1e_i}}{2}(x_i-\hat{x}_i)^2+\frac{k_{2e_i}}{2}(x_i-\hat{x}_i-\tilde{x}_i)^2\right)$, where $\hat{x}_i,\tilde{x}_i$ are extra states and $k_{1e_i},k_{2e_i}>0$ are parameters) can lead to different types of extra dynamics. In the future, we will investigate different types of extra terms and the implementation of the resulting dynamics.

Last but not least, Chapter 5 serves as our initial step towards developing a reverse- and forward-engineering framework for system control (re)design. In the

future, we will focus on more general systems with more complexity. We will extend our result to discrete-time LTI systems, and then focus on linear time-varying systems. We will consider developing necessary and sufficient conditions to reverse-engineer nonlinear dynamic systems. We will also develop matrix partition methods for related theorems in Section 5.2.3. Finally, besides power networks, we are interested in applying our result to do control (re)design for other practical systems, e.g., Hamiltonian systems.

Appendix A

Lyapunov Stability

In this appendix, we review some results relating to Lyapunov stability theorems based on [79, 80], which are frequently used in this thesis.

Let $x(t) \in \mathbb{R}^n$ and $f : D \to \mathbb{R}^n$ be locally Lipschitz (i.e., $\|f(x)-f(y)\| \leq L\|x-y\|$ where $\|\cdot\|$ denotes any p-norm and $L > 0$ is constant) in a domain $D \subset \mathbb{R}^n$. Consider the following autonomous system

$$\dot{x}(t) = f(x(t)) \tag{A.1}$$

and suppose x^* is an equilibrium point of (A.1), i.e., $f(x^*) = \mathbf{0}$. For convenience, we assume $x^* = \mathbf{0}$ which can be achieved using a simple change of coordinates. Here, we focus on the Lyapunov stability of the equilibrium $x^* = \mathbf{0}$, which is summarized as follows [79] ($\|\cdot\|$ denotes a norm in $\mathbb{R}^n$).

Definition A.1 *The equilibrium $x^* = \mathbf{0}$ of system (A.1) is:*

(i) Stable, if for any $\epsilon > 0$ there is a $\delta = \delta(\epsilon) > 0$ such that

$$\|x(0)\| < \delta \Rightarrow \|x(t)\| < \epsilon, \forall t \geq 0;$$

(ii) Unstable, if it is not stable;

(iii) Asymptotically stable, if it is stable and δ can be chosen such that

$$\|x(0)\| < \delta \Rightarrow \lim_{t\to\infty} \|x(t)\| = 0.$$

In many cases, stability can be shown by constructing an energy-like function, called a Lyapunov function. The Lyapunov stability theorem is provided below [79].

Theorem A.1 *Let $D \subseteq \mathbb{R}^n$ be a neighborhood of the origin. If there is a continuously differentiable function $V : D \to \mathbb{R}$ such that the following two conditions are satisfied:*

(i) $V(x) > 0, \forall x \in D \setminus \{\mathbf{0}\}$ and $V(\mathbf{0}) = 0$, i.e., $V(x)$ is positive definite in D,

(ii) $\dot{V}(x) = \frac{\partial V}{\partial x} f(x) \leq 0, \forall x \in D$, *i.e.,* $\dot{V}(x)$ *is negative semi-definite in* D, *then the origin is a stable equilibrium of (A.1). If in condition (ii),* $\dot{V}(x) < 0, \forall x \in D \setminus \{\mathbf{0}\}$, *i.e.,* $\dot{V}(x)$ *is negative definite in* D, *then the origin is asymptotically stable. Furthermore, if* $D = \mathbb{R}^n$ *and* $V(x)$ *is radially unbounded, i.e.,* $V(x) \to \infty$ *as* $\|x\| \to \infty$, *then the origin is globally asymptotically stable.*

To show the asymptotic stability of the equilibrium of (A.1), an alternative way is to use the LaSalle's invariance principle, given by the following theorem [79].

Theorem A.2 *Let* $\Omega \subset D$ *be a compact set that is positively invariant with respect to system (A.1). Let* $V : D \to \mathbb{R}$ *be a continuously differentiable function such that* $\dot{V}(x) \leq 0$ *in* Ω. *Let* E *be the set of all points in* Ω *where* $\dot{V}(x) = 0$. *Let* M *be the largest invariant set in* E. *Then every solution starting in* Ω *approaches* M *as* $t \to \infty$.

Likewise, for the global asymptotic stability of the equilibrium of (A.1), an alternative way is to use the Krasovskii-LaSalle principle as follow [79].

Theorem A.3 *Let* $x^* = \mathbf{0}$ *be an equilibrium point of (A.1). Let* $V : \mathbb{R}^n \to \mathbb{R}$ *be a continuously differentiable, radially unbounded, positive definite function such that* $\dot{V}(x) \leq 0, \forall x(t) \in \mathbb{R}^n$. *Let* $S = \{x(t) \in \mathbb{R}^n | \dot{V}(x) = 0\}$ *and suppose that no solution can stay identically in* S, *other than the trivial solution* $x^* = \mathbf{0}$. *Then the origin is globally asymptotically stable.*

We next consider the case in which a system is modeled by an autonomous delayed differential equation, i.e.,

$$\dot{x}(t) = f(x(t), x(t - \tau_1), \cdots, x(t - \tau_r)) \tag{A.2}$$

where each τ_i is a discrete delay in the system. Without loss of generality, we assume that $\mathbf{0}$ is a steady state for the system. The following definition extends Definition A.1, relating to the stability of the solution $x = \mathbf{0}$ of system (A.2) [80].

Definition A.2 *The solution* $x = \mathbf{0}$ *of system (A.2) is:*

(i) Stable, if for any $\epsilon > 0$ *there is a* $\delta = \delta(\epsilon) > 0$ *such that* $\|x(t)|_\phi\| \leq \epsilon$ *for any initial condition* $\phi \in \Omega^\delta$, *where* $x(t)|_\phi$ *denotes the solution of (A.2) given the initial condition* ϕ, $\Omega^\delta = \{\phi \in \mathcal{C}([-\tau, 0], \mathbb{R}^n) | \|\phi\| \leq \delta\}$, *and* $\tau = \max_i\{\tau_i\}$;

(ii) Unstable, if it is not stable;

(iii) Asymptotically stable, if it is stable and there is a δ *such that* $\lim_{t\to\infty} x(t)|_\phi = \mathbf{0}$ *for any initial condition* $\phi \in \Omega^\delta$.

Just as in the case of systems described by (A.1), a Lyapunov argument can be formulated to establish the stability of the steady state of (A.2) when many, incommensurate delays appear in the system.

Consider a continuous functional $V : \mathcal{C}([-\tau, \infty), \mathbb{R}^n) \to \mathbb{R}$ and define

$$\dot{V}(x(t)|_\phi) = \lim_{h \to 0^+} \frac{1}{h}(V(x(t+h)|_\phi) - V(x(t)|_\phi))$$

to be the derivative of V along a solution of (A.2). We then have the following Lyapunov-Krasovskii theorem [80].

Theorem A.4 *Let f be completely continuous and $f(\mathbf{0}) = \mathbf{0}$ in (A.2). Assume that there exist $a(s)$ and $b(s)$ that are non-negative continuous, $a(0) = b(0) = 0$, $\lim_{s\to\infty} a(s) = +\infty$, and $V : \Omega \to \mathbb{R}$ that is continuous and satisfies*

$$V(x(t)|_\phi) \geq a(\|\phi(0)\|) \text{ in } \Omega$$
$$\dot{V}(x(t)|_\phi) \leq -b(\|\phi(0)\|) \text{ in } \Omega$$

where $\Omega \subset \mathcal{C}([-\tau, \infty), \mathbb{R}^n)$. Then the solution $x = \mathbf{0}$ of (A.2) is stable, and every solution is bounded. Furthermore, if $b(s) > 0$ for $s > 0$, then $x = \mathbf{0}$ is asymptotically stable.

Finally, we present a theorem which extends the LaSalle's invariance principle to the case of time-delay systems (A.2) [80].

Theorem A.5 *Let $V : \Omega \to \mathbb{R}$ be a Lyapunov functional with respect to system (A.2), i.e., V is continuous on the closure of Ω, and $\dot{V} \leq 0$ on Ω, where $\Omega \subset \mathcal{C}([-\tau, \infty), \mathbb{R}^n)$. Let E be the set of ϕ on the closure of Ω where $\dot{V}(x(t)|_\phi) = 0$. Let M be the largest invariant set in E with respect to system (A.2). Then for a bounded solution $x(t)|_\phi$ of (A.2) that remains in Ω, $x(t)|_\phi$ tends to M as $t \to \infty$.*

Appendix B

Synchronization in Oscillator Networks with Non-homogeneous Delays

In this appendix, we review some results relating to synchronization in oscillator networks with non-homogeneous delays according to [81, 82].

Consider N coupled oscillators with natural frequencies ω_i and phases $\theta_i \in [0, 2\pi]$, $i = 1, \cdots, N$. The dynamics of each oscillator are described by the modified Kuramoto Model with non-homogeneous delays:

$$\dot{\theta}_i(t) = \omega_i(t) + \frac{K}{n_i} \sum_{j=1}^{N} A_{ij} \sin(\theta_j(t - \tau_{ji}) - \theta_i(t)) \tag{B.1}$$

where K is the coupling strength between the oscillators which is assumed to be the same, n_i is the number of neighbouring oscillators to oscillator i, and τ_{ji} is the delay in the coupling from oscillator j to oscillator i. Note that we require $\tau_{ji} = \tau_{ij}$. In addition, A_{ij} is the entry of the adjacency matrix A of the underlying undirected graph consisting of these oscillators. Here the adjacency matrix is an $N \times N$ matrix with elements $A_{ij} = 1$ only if oscillators i and j are neighbours to each other (i.e., they are connected by a link), otherwise, $A_{ij} = 0$.

Formally speaking, we consider the case where all the oscillators synchronize, i.e., $\theta_i(t) = \theta(t) = \Omega t + \alpha, i = 1 \cdots, N$. Then the self-consistency equation for ω_i and K needs to be satisfied given by

$$\Omega = \omega_i - \frac{K}{n_i} \sum_{j=1}^{N} A_{ij} \sin(\Omega \tau_{ji}). \tag{B.2}$$

Let $\theta_i(t) = \Omega t + \phi_i(t)$. Linearize (B.1) around the steady state $\phi_i(t) = \alpha$ to get

$$\dot{\phi}_i(t) = \frac{K}{n_i} \sum_{j=1}^{N} G_{ij}(\phi_j(t - \tau_{ji}) - \phi_i(t)) \tag{B.3}$$

where $G_{ij} = G_{ji} = A_{ij}\cos(\Omega\tau_{ji})$ since $\tau_{ij} = \tau_{ji}$ and $A_{ij} = A_{ji}$. Note that the equilibria of this system belong to the consensus set $\mathcal{CS} = \{(\phi_1, \cdots, \phi_N)|\phi_1 = \phi_2 = \cdots = \phi_N = \alpha\}$ which is one-dimensional. The stability of (B.3) is shown below.

Theorem B.1 *If* $\cos(\Omega\tau_{ji}) > 0$ *for all* i, j *for which* $A_{ij} = 1$ *and* Ω *solves (B.2), then the equilibrium set of system (B.3), i.e.,* $\mathcal{CS} = \{(\phi_1, \cdots, \phi_N)|\phi_1 = \phi_2 = \cdots = \phi_N = \alpha\}$*, is asymptotically attracting.*

Proof: Consider the following candidate Lyapunov-Krasovskii functional

$$V_{(B.3)}(t) = \frac{1}{2}\sum_{i=1}^{N} n_i\phi_i^2(t) + \frac{K}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}\int_{t-\tau_{ji}}^{t} \phi_j^2(\beta)d\beta.$$

Differentiating $V_{(B.3)}(t)$ with respect to time, we get

$$\begin{aligned}\dot{V}_{(B.3)}(t) =& -K\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}\phi_i^2(t) + K\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}\phi_i(t)\phi_j(t-\tau_{ji}) \\ &+ \frac{K}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}\phi_j^2(t) - \frac{K}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}\phi_j^2(t-\tau_{ji}) \\ =& -\frac{K}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}(\phi_i(t) - \phi_j(t-\tau_{ji}))^2.\end{aligned}$$

If $\cos(\Omega\tau_{ji}) > 0$ for all i, j for which $A_{ij} = 1$, $\dot{V}_{(B.3)}(t) \leq 0$. Moreover, $\dot{V}_{(B.3)}(t) = 0$ only happens when $\phi_i(t) - \phi_j(t - \tau_{ji}) = 0$ holds for all i, j for which $A_{ij} = 1$. This results in $\dot{\phi}_i(t) = 0, i = 1, \cdots, N$. Using LaSalle's invariance principle for time-delay systems [80], the equilibrium set is asymptotically attracting. ■

We can see that the proofs of Theorem 2.3 and Propositions 2.2 and 2.4 are similar to the above proof, using the same type of Lyapunov-Krasovskii functionals.

Appendix C

Saddle Point Dynamics and Primal-dual Gradient Dynamics

In this appendix, we review some mathematical preliminaries about the saddle point dynamics and the primal-dual gradient dynamics (for solving convex optimization problems) based on [77, 56, 70].

Consider a convex-concave function $f \in \mathcal{C}^2$: $\mathbb{R}^a \times \mathbb{R}^b \to \mathbb{R}$ satisfying that for all y, z, $\nabla_y^2 f \succeq 0, \nabla_z^2 f \preceq 0$ hold. We say $(\tilde{y}, \tilde{z})$ is a saddle point of f if

$$f(\tilde{y}, z) \leq f(\tilde{y}, \tilde{z}) \leq f(y, \tilde{z}) \tag{C.1}$$

holds. A property of a saddle point of a function is shown in Lemma 5.1, the proof of which is provided below.

Proof of Lemma 5.1: Necessity. Since the saddle point condition (C.1) holds, $\tilde{y}$ minimizes $f(y, \tilde{z})$ over all y and $\tilde{z}$ maximizes $f(\tilde{y}, z)$ over all z. Thus, $\nabla_y f|_{y=\tilde{y}, z=\tilde{z}} = \mathbf{0}$ and $\nabla_z f|_{y=\tilde{y}, z=\tilde{z}} = \mathbf{0}$ hold, i.e., $\nabla_{y,z} f|_{y=\tilde{y}, z=\tilde{z}} = \mathbf{0}$.

Sufficiency. Since $\nabla_y f|_{y=\tilde{y}, z=\tilde{z}} = \mathbf{0}$ and $f(y, \tilde{z})$ is convex in y, we conclude that $\tilde{y}$ minimizes $f(y, \tilde{z})$ over y, i.e., for all y, we have $f(\tilde{y}, \tilde{z}) \leq f(y, \tilde{z})$. This is one of the inequalities in (C.1). We can argue in the same way about z. ■

Denote the set of saddle points of f by $\mathcal{S}$ and assume that $\mathcal{S}$ is nonempty. Consider the saddle point dynamics given by

$$\dot{y} = -K_y \frac{\partial f}{\partial y} \tag{C.2a}$$

$$\dot{z} = K_z \frac{\partial f}{\partial z} \tag{C.2b}$$

where $K_y \in \mathbb{R}^{a\times a}, K_z \in \mathbb{R}^{b\times b}$ are positive definite constant matrices. A property of the saddle point dynamics is shown in Lemma 5.2, the proof of which is provided below.

Proof of Lemma 5.2: Let (y^*, z^*) be a saddle point of f. Define a candidate Lyapunov function as

$$V_{(C.2)} = \frac{1}{2}(y-y^*)^T K_y^{-1}(y-y^*) + \frac{1}{2}(z-z^*)^T K_z^{-1}(z-z^*)$$

which is radially unbounded and positive definite with respect to (y^*, z^*). The derivative of $V_{(C.2)}$ with respect to time along the trajectory of (C.2) is given by

$$\begin{aligned}\dot{V}_{(C.2)} =& -\left(\frac{\partial f}{\partial y}\right)^T (y-y^*) + \left(\frac{\partial f}{\partial z}\right)^T (z-z^*)\\ \leq& -f(y,z) + f(y^*,z) + f(y,z) - f(y,z^*)\\ =& f(y^*,z) - f(y^*,z^*) + f(y^*,z^*) - f(y,z^*)\\ \leq& 0\end{aligned}$$

where the first inequality comes from the fact that f is convex in y and concave in z, and the last inequality follows that (y^*, z^*) is a saddle point of f. When $\dot{V}_{(C.2)} = 0$, we have $f(y^*, z) = f(y^*, z^*)$ and $f(y^*, z^*) = f(y, z^*)$. Based on LaSalle's invariance principle [79], we conclude that the trajectories of (C.2) converge to a compact subset of the invariant set given by

$$\mathcal{I} = \{(y,z) | \dot{V}_{(C.2)}(y,z,y^*,z^*) = 0\}$$

which indicates that the trajectories of (C.2) are bounded.

We next show that if f is either strictly convex in y or strictly concave in z, then the trajectories of (C.2) asymptotically converge to a saddle point of f. Firstly, if $(\tilde{y}, \tilde{z}) \in \mathcal{I}$, then $f(y^*, \tilde{z}) = f(y^*, z^*)$ and $f(y^*, z^*) = f(\tilde{y}, z^*)$. Since f is either strictly convex in y or strictly concave in z, either $\tilde{y} = y^*$ or $\tilde{z} = z^*$ holds. Let us assume $\tilde{y} = y^*$, which means $\dot{y}|_{y=\tilde{y}} = \mathbf{0}$. Then $\frac{\partial f}{\partial y}|_{y=\tilde{y}} = \mathbf{0}$, i.e., $\tilde{y}$ satisfies the first order optimality condition for $f(y, z)$. On the other hand, since $f(\tilde{y}, z^*) = f(y^*, z^*) = f(y^*, \tilde{z})$, $\tilde{z}$ is a maximizer of $f(y^*, z)$, i.e., $\frac{\partial f}{\partial z}|_{y=y^*=\tilde{y}, z=\tilde{z}} = \mathbf{0}$. Thus, we have $\nabla_{y,z} f|_{y=\tilde{y}, z=\tilde{z}} = \mathbf{0}$, i.e., $(\tilde{y}, \tilde{z}) \in \mathcal{S}$. Similarly, if $\tilde{z} = z^*$, we can still derive $(\tilde{y}, \tilde{z}) \in \mathcal{S}$. To conclude, under the condition that f is either strictly convex in y or strictly concave in z, $\mathcal{I} \subseteq \mathcal{S}$ is true.

To show the pointwise convergence, since $(y(t), z(t))$ converges to a compact subset of $\mathcal{I} \subseteq \mathcal{S}$, there exists a subsequence $\{(y_k, z_k)\}$ where $y_k = y(t_k), z_k = z(t_k)$ that

converges to a point (y^∞, z^∞). This means that $v_k = V_{(C.2)}(y_k, z_k, y^\infty, z^\infty)$ asymptotically converges to 0. On the other hand, we have

$$\lim_{t\to\infty} V_{(C.2)}(y(t), z(t), y^\infty, z^\infty) = \lim_{t\to\infty} v(t) = v^\infty$$

where v^∞ is constant. Since v_k is a subsequence of $v(t)$ and converges to 0, $v^\infty = 0$ holds. Therefore, we conclude that $(y(t), z(t))$ converges to (y^∞, z^∞). ■

Let us now consider a convex optimization problem given by

$$\min_y C(y) \tag{C.3a}$$

$$\text{subject to} \quad g(y) \preceq 0 \tag{C.3b}$$

where $C \in \mathcal{C}^2 : \mathbb{R}^a \to \mathbb{R}$ satisfying $\nabla_y^2 C \succeq 0$, and $g(y) : \mathbb{R}^a \to \mathbb{R}^b$ is the vector of $g_i \in \mathcal{C}^2 : \mathbb{R}^a \to \mathbb{R}$ satisfying $\nabla_y^2 g_i \succeq 0$. Assume that this problem is feasible and satisfies Slater's constraint qualification [77], i.e., strong duality holds for this optimization problem. The Lagrangian of the problem is given by

$$f(y, z) = C(y) + z^T g(y)$$

where $z \succeq 0$ is the vector of Lagrangian multipliers (dual variables) z_i for the constraint $g(y) \preceq 0$. Thus, the dual optimization problem is given by

$$\max_z \inf_y f(y, z) \tag{C.4a}$$

$$\text{subject to} \quad z \succeq 0. \tag{C.4b}$$

Let y^* and $z^* \succeq 0$ denote the optimal solution of the primal problem (C.3) and the dual problem (C.4) respectively. Due to strong duality, (y^*, z^*) is a saddle point of f. Moreover, (y^*, z^*) satisfies the Karush-Kuhn-Tucker (KKT) conditions given by

$$\frac{\partial f}{\partial y} = \mathbf{0} \tag{C.5a}$$

$$z_i \frac{\partial f}{\partial z_i} = 0, \quad z_i \geq 0, \quad \frac{\partial f}{\partial z_i} \leq 0. \tag{C.5b}$$

The corresponding saddle point dynamics, i.e. the primal-dual gradient dynamics, are given by

$$\dot{y} = -K_y \frac{\partial f}{\partial y} \tag{C.6a}$$

$$\dot{z} = K_z \left(\frac{\partial f}{\partial z}\right)_z^+ \tag{C.6b}$$

where $K_y \in \mathbb{R}^{a\times a}, K_z \in \mathbb{R}^{b\times b}$ are positive definite constant matrices. Note that we have used a vector form of positive projection in (C.6b).

Similar to Lemma 5.2, we have the following results relating the stability of (C.6).

Proposition C.1 *The trajectories of the primal-dual gradient dynamics (C.6) are bounded. Furthermore, if C is strictly convex in y, the trajectories of (C.6) asymptotically converge to a saddle point of f.*

Proof: Consider the following candidate Lyapunov function

$$V_{(C.6)} = V_{(C.2)} = \frac{1}{2}(y-y^*)^T K_y^{-1}(y-y^*) + \frac{1}{2}(z-z^*)^T K_z^{-1}(z-z^*).$$

According to the proof of Theorem 3.3, we can remove all positive projection in $\dot{V}_{(C.6)}$ to get

$$\dot{V}_{(C.6)} \leq -\left(\frac{\partial f}{\partial y}\right)^T (y-y^*) + \left(\frac{\partial f}{\partial z}\right)^T (z-z^*).$$

Following the proof of Lemma 5.2, we can obtain that the trajectories of (C.6) are bounded, and moreover, if C is strictly convex in y, the trajectories of (C.6) asymptotically converge to a saddle point of f. ■

We see that from the saddle point dynamics (C.2) to the primal-dual gradient dynamics (C.6), although the positive projection is added, the results in Lemmas 5.1 and 5.2 still hold (with $\nabla_z f|_{y=\tilde{y},z=\tilde{z}} = \mathbf{0}$ changing to the complementary slackness condition (C.5b)).

Appendix D

Classical Loop Shaping in Multiple-Input-Multiple-Output (MIMO) Systems

In this appendix, we review the classical loop shaping method for Multiple-Input-Multiple-Output (MIMO) systems, according to [102, 89].

Consider the one degree-of-freedom configuration shown in Figure D.1 [89]. Define the (output) sensitivity function to be $S = (I + GK)^{-1}$ and the (output) complementary sensitivity function to be $T = GK(I + GK)^{-1} = I - S$. Then we have

$$y(s) = T(s)(r(s) - n(s)) + S(s)d(s)$$
$$u(s) = K(s)S(s)(r(s) - n(s) - d(s)).$$

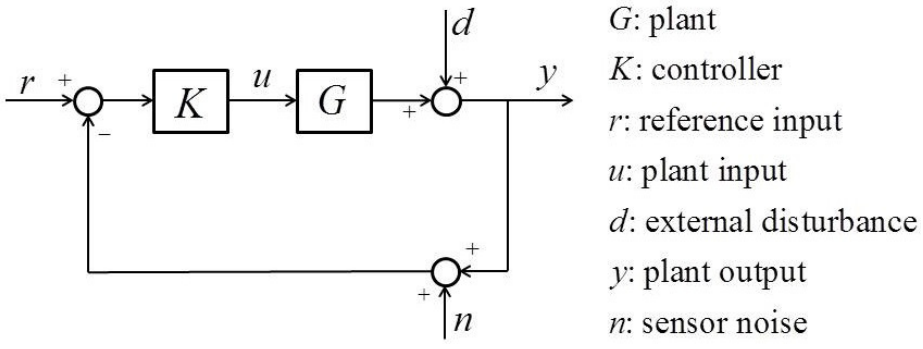

Figure D.1: One degree-of-freedom feedback configuration.

(i) For disturbance rejection, $\overline{\sigma}(S)$ needs to be small. When $\underline{\sigma}(GK) \gg 1$, $S = (I + GK)^{-1} \approx (GK)^{-1}$. Then making $\overline{\sigma}(S)$ small can be approximated by making $\underline{\sigma}(GK)$ large when $\underline{\sigma}(GK) \gg 1$. Here $\overline{\sigma}(X)$ and $\underline{\sigma}(X)$ are the maximum and minimum

nonzero singular values of a matrix X respectively (the singular values are the square roots of the eigenvalues of $X^H X$ where X^H is the conjugate transpose of X).

(ii) For noise attenuation, $\overline{\sigma}(T)$ needs to be small. When $\overline{\sigma}(GK) \ll 1$, $T = GK(I + GK)^{-1} \approx GK$. Then making $\overline{\sigma}(T)$ small can be approximated by making $\overline{\sigma}(GK)$ small when $\overline{\sigma}(GK) \ll 1$.

(iii) For reference tracking, both $\overline{\sigma}(T)$ and $\underline{\sigma}(T)$ need to be close to 1. When $\underline{\sigma}(GK) \gg 1$, $T = GK(I + GK)^{-1} \approx I$. So making $\underline{\sigma}(GK)$ large leads to $\overline{\sigma}(T) \approx \underline{\sigma}(T) \approx 1$ when $\underline{\sigma}(GK) \gg 1$.

(iv) For control energy reduction, $\overline{\sigma}(KS)$ needs to be small. When $\overline{\sigma}(GK) \ll 1$, $KS = K(I + GK)^{-1} \approx K$. Then making $\overline{\sigma}(KS)$ small can be approximated by making $\overline{\sigma}(K)$ small when $\overline{\sigma}(GK) \ll 1$.

(v) For robust stability to an additive perturbation in plant G, $\overline{\sigma}(KS)$ needs to be small [103]. Same as (iv) above, we can make $\overline{\sigma}(K)$ small when $\overline{\sigma}(GK) \ll 1$.

(vi) For robust stability to a multiplicative perturbation in plant G, $\overline{\sigma}(T)$ needs to be small [103]. When $\overline{\sigma}(GK) \ll 1$, $T = GK(I + GK)^{-1} \approx GK$. Then making $\overline{\sigma}(T)$ small can be approximated by making $\overline{\sigma}(GK)$ small when $\overline{\sigma}(GK) \ll 1$.

Typically, (i) and (iii) are valid at low frequencies and (ii), (iv), (v), (vi) are valid at high frequencies. Thus, in terms of the loop transfer function matrix GK, $\underline{\sigma}(GK)$ needs to be large at low frequencies for good performance of the closed-loop system, and $\overline{\sigma}(GK)$ needs to be small at high frequencies for robust stability , noise attenuation and control energy reduction of the closed-loop system.

Appendix E

Vinnicombe's Lemma

Lemma E.1 *Let $Q = Q^H \succeq 0$ be an $n \times n$ positive semi-definite Hermitian matrix, and $\Lambda = diag\{\lambda_i\}$ be an $n \times n$ diagonal matrix where $\lambda_i \in \mathbb{C}, i = 1, \cdots, n$, i.e., each λ_i is a complex number. Then the eigenvalues of $Q\Lambda$ belong to the convex hull of $\{0, \lambda_1, \cdots, \lambda_n\}$, scaled by the spectral radius $\rho(Q)$.*

Proof: If 0 is an eigenvalue of $Q\Lambda$, the result is obvious. Due to $Q = Q^H \succeq 0$, Q can be rewritten as $Q = Q^{\frac{1}{2}}Q^{\frac{1}{2}}$ where $Q^{\frac{1}{2}} = Q^{\frac{1}{2}H} \succeq 0$. Since nonzero eigenvalues are invariant under commutation, the nonzero eigenvalues of $Q\Lambda$ are identical to the nonzero eigenvalues of $Q^{\frac{1}{2}}\Lambda Q^{\frac{1}{2}}$. Let v be a normalized eigenvector of $Q\Lambda$, i.e., $\|v\| = 1$, corresponding to a nonzero eigenvalue of $Q\Lambda$. Then this nonzero eigenvalue is

$$v^H Q^{\frac{1}{2}}\Lambda Q^{\frac{1}{2}} v = \rho(Q) \sum_{i=1}^{n} \frac{|w_i|^2}{\rho(Q)} \lambda_i$$

where w_i is the entry of the vector $Q^{\frac{1}{2}}v$. Since $\|Q^{\frac{1}{2}}v\| \leq \sqrt{\rho(Q)}$, we have

$$\sum_{i=1}^{n} \frac{|w_i|^2}{\rho(Q)} \leq 1.$$

This demonstrates that the eigenvalues of $Q\Lambda$ belong to the convex hull of $\{0, \lambda_1, \cdots, \lambda_n\}$, scaled by the spectral radius $\rho(Q)$. ■

Bibliography

[1] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira, and E. Izco. Power output fluctuations in large scale PV plants: One year observations with one second resolution and a derived analytic model. *Progress in Photovoltaics: Research and Applications*, 19(2):218–227, 2011.

[2] C. Kamath. Understanding wind ramp events through analysis of historical data. In *Proc. of IEEE PES Transmission and Distribution Conference and Exposition*, 2010.

[3] P. Kundur. *Power System Stability and Control*. New York: McGraw-Hill, 1994.

[4] J. Machowski, J. W. Bialek, and J. R. Bumby. *Power Systems Dynamics Stability and Control*. John Wiley & Sons, Ltd, second edition, 2008.

[5] A. Jokic. Price-based optimal control of electrical power systems. *PhD dissertation, Department of Electrical Engineering, Eindhoven University of Technology*, 2007.

[6] A. R. Bergen and V. Vittal. *Power Systems Analysis*. Pearson Education Inc. and Dorling Kindersley Publishing Inc., second edition, 2000.

[7] D. Kirschen and G. Strbac. *Fundamentals of Power System Economics*. John Wiley & Sons, Ltd, first edition, 2004.

[8] A. Ipakchi and F. Albuyeh. Grid of the future. *IEEE Power Energy Magzine*, 8(4):52–62, 2009.

[9] Element Energy Limited. Electric vehicles in the UK and Republic of Ireland: Greenhouse gas emission reductions & infrastructure needs. *Final report for WWF UK*, 2010.

[10] M. Duvall and E. Knipping. Environmental assessment of plug-in hybrid electric vehicles, Volume 1: Nationwide greenhouse gas emissions. *EPRI Technical Report 1015325*, 2007.

[11] J. Taylor, A. Maitra, M. Alexander, D. Brooks, and M. Duvall. Evaluations of plug-in electric vehicle distribution system impacts. In *Proc. of IEEE Power and Energy Society General Meeting*, 2010.

[12] J. A. P. Lopes, F. J. Soares, and P. M. R. Almeida. Integration of electric vehicles in the electric power system. *Proc. of IEEE*, 99(1):168–183, 2011.

[13] N. Li, L. Chen, C. Zhao, and S. H. Low. Connecting automatic generation control and economic dispatch from an optimization view. In *Proc. of 2014 American Control Conference*, pages 735–740, 2014.

[14] S. Borenstein, M. Jaske, and A. Rosenfeld. Dynamic pricing, advanced metering, and demand response in electricity markets. *Technical Report, Center for the Study of Energy Markets, CSEMWP105*, 2002.

[15] T. J. Lui, W. Stirling, and H. O. Marcy. Get smart. *IEEE Power and Energy Magazine*, 8(3):66–78, 2010.

[16] F. L. Alvarado, J. Meng, C. DeMarco, and W. Mota. Stability analysis of interconnected power systems coupled with market dynamics. *IEEE Transactions on Power Systems*, 16(4):695–701, 2001.

[17] F. C. Schweppe, R. D. Tabors, J. L. Kirtley, H. R. Outhred, F. H. Pickel, and A. J. Cox. Homeostatic utility control. *IEEE Transactions on Power Apparatus and Systems*, 99(3):1151–1163, 1980.

[18] M. C. Caramanis, R. E. Bohn, and F. C. Schweppe. Optimal spot pricing: Practice and theory. *IEEE Transactions on Power Apparatus and Systems*, 109(9):3234–3245, 1982.

[19] W. W. Hogan. Contract networks for electric power transmission. *Regulatory Economics*, 4:211–242, 1992.

[20] J. Cardell, C. C. Hitt, and W. W. Hogan. Market power and strategic interaction in electricity networks. *Resource and Energy Economics*, 19(1–2):109–137, 1997.

[21] H. Glavitsch and F. L. Alvarado. Management of multiple congested conditions in unbundled operation of a power system. *IEEE Transactions on Power Systems*, 13(3):1013–1019, 1998.

[22] F. L. Alvarado. The stability of power system markets. *IEEE Transactions on Power Systems*, 14(3):509–511, 1999.

[23] C. Zhao, U. Topcu, and S. Low. Swing dynamics as primal-dual algorithm for optimal load control. In *Proc. of 3rd IEEE International Conference on Smart Grid Communications*, pages 570–575, 2012.

[24] C. Zhao, U. Topcu, N. Li, and S. Low. Design and stability of load-side primary frequency control in power systems. *IEEE Transactions on Automatic Control*, 59(5):1177–1189, 2014.

[25] C. Zhao, U. Topcu, and S. Low. Frequency-based load control in power systems. In *Proc. of 2012 American Control Conference*, pages 4423–4430, 2012.

[26] C. Zhao, U. Topcu, and S. Low. Fast load control with stochastic frequency measurement. In *Proc. of IEEE Power and Energy Society General Meeting*, 2012.

[27] C. Zhao, U. Topcu, and S. Low. Optimal load control via frequency measurement and neighborhood area communication. *IEEE Transactions on Power Systems*, 28(4):3576–3587, 2013.

[28] E. Mallada and S. H. Low. Distributed frequency-preserving optimal load control. In *Proc. of 19th IFAC World Congress*, pages 5411–5418, 2014.

[29] E. Mallada, C. Zhao, and S. Low. Optimal load-side control for frequency regulation in smart grids. In *Proc. of Allerton Conference on Communication, Control, and Computing*, pages 731–738, 2014.

[30] C. Zhao and S. Low. Optimal decentralized primary frequency control in power networks. In *Proc. of 53rd IEEE Conference on Decision and Control*, pages 2467–2473, 2014.

[31] E. Mallada, C. Zhao, and S. Low. Distributed generator and load-side secondary frequency control in power networks. In *Proc. of Conference on Information Sciences and Systems (CISS)*, 2015.

[32] C. Zhao, E. Mallada, and F. Dörfler. Distributed frequency control for stability and economic dispatch in power networks. In *Proc. of 2015 American Control Conference*, pages 4423–4430, 2015.

[33] M. Andreasson, H. Sandberg, D. V. Dimarogonas, and K. H. Johansson. Distributed integral action: Stability analysis and frequency control of power systems. In *Proc. of 51st IEEE Conference on Decision and Control*, pages 2077–2083, 2012.

[34] M. Andreasson, D. V. Dimarogonas, K. H. Johansson, and H. Sandberg. Distributed vs. centralized power systems frequency control. In *Proc. of 2013 European Control Conference*, pages 3524–3529, 2013.

[35] M. Andreasson, D. V. Dimarogonas, H. Sandberg, and K. H. Johansson. Distributed PI control with applications to power systems frequency control. In *Proc. of 2014 American Control Conference*, pages 3183–3188, 2014.

[36] F. Bullo, J. W. Simpson-Porco, and F. Dörfler. Synchronization and power sharing for droop-controlled inverters in islanded microgrids. *Automatica*, 49(9):2603–2611, 2013.

[37] J. W. Simpson-Porco, F. Dörfler, Q. Shafiee, J. M. Guerrero, and F. Bullo. Stability, power sharing and distributed secondary control in droop-controlled microgrids. In *Proc. of 4th IEEE International Conference on Smart Grid Communications*, pages 672–677, 2013.

[38] H. Bouattour, J. W. Simpson-Porco, F. Dörfler, and F. Bullo. Further results on distributed secondary control in microgrids. In *Proc. of 52nd IEEE Conference on Decision and Control*, pages 1514–1519, 2013.

[39] Q. Shafiee, J. C. Vasquez, and J. M. Guerrero. Distributed secondary control for islanded microgrids - A networked control systems approach. In *Proc. of Annual Conference on IEEE Industrial Electronics Society*, pages 5637–5642, 2012.

[40] F. Dörfler, J. W. Simpson-Porco, and F. Bullo. Plug-and-play control and optimization in microgrids. In *Proc. of 53rd IEEE Conference on Decision and Control*, pages 211–216, 2014.

[41] F. Dörfler, J. W. Simpson-Porco, and F. Bullo. Breaking the hierarchy: Distributed control and economic optimality in microgrids. *IEEE Transactions on Control of Network Systems*, 2015, to appear.

[42] X. Zhang and A. Papachristodoulou. Redesigning generation control in power systems: Methodology, stability and delay robustness. In *Proc. of 53rd IEEE Conference on Decision and Control*, pages 953–958, 2014.

[43] X. Zhang, R. Kang, M. McCulloch, and A. Papachristodoulou. Real-time active and reactive power regulation in power systems with tap-changing transformers and controllable loads. *Sustainable Energy, Grids and Networks*, 5:27–38, 2016.

[44] X. Zhang and A. Papachristodoulou. A real-time control framework for smart power networks with star topology. In *Proc. of 2013 American Control Conference*, pages 5062–5067, 2013.

[45] X. Zhang and A. Papachristodoulou. Distributed dynamic feedback control for smart power networks with tree topology. In *Proc. of 2014 American Control Conference*, pages 1156–1161, 2014.

[46] X. Zhang and A. Papachristodoulou. A real-time control framework for smart power networks: Design methodology and stability. *Automatica*, 58:43–50, 2015.

[47] V. Jacobson. Congestion avoidance and control. *ACM SIGCOMM Computer Communication Review*, 18(4):314–329, 1988.

[48] F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: Shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.

[49] S. H. Low, J. C. Doyle, and F. Paganini. Internet congestion control. *IEEE Control System Magazine*, 21(1):28–43, 2002.

[50] R. Srikant. *The Mathematics of Internet Congestion Control*. Cambridge, MA: Birkhäuser, 2004.

[51] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proc. of IEEE*, 95(1):255–312, 2007.

[52] D. Papadimitriou, M. Welzl, M. Scharf, and B. Briscoe. *Open research issues in Internet congestion control*. RFC 6077, Internet Research Task Force (IRTF), 2011.

[53] J. T. Wen and M. Arcak. A unifying passivity framework for network flow control. *IEEE Transactions on Automatic Control*, 49(2):162–174, 2004.

[54] J. Lavaei, J. C. Doyle, and S. H. Low. Utility functionals associated with available congestion control algorithms. In *Proc. of IEEE INFOCOM*, 2010.

[55] A. Papachristodoulou and A. Jadbabaie. Delay robustness of nonlinear internet congestion control schemes. *IEEE Transactions on Automatic Control*, 55(6):1421–1428, 2010.

[56] D. Feijer and F. Paganini. Stability of primal-dual gradient dynamics and applications to network optimization. *Automatica*, 46(12):1974–1981, 2010.

[57] J. Wang and N. Elia. A control perspective for centralized and distributed convex optimization. In *Proc. of 50th IEEE Conference on Decision and Control*, pages 3800–3805, 2011.

[58] E. Wei, A. Ozdaglar, and A. Jadbabaie. A distributed Newton method for network utility maximization, Part I: Algorithm. *IEEE Transactions on Automatic Control*, 58(9):2162–2175, 2013.

[59] E. Wei, A. Ozdaglar, and A. Jadbabaie. A distributed Newton method for network utility maximization, Part II: Convergence. *IEEE Transactions on Automatic Control*, 58(9):2176–2188, 2013.

[60] P. Wan and M. D. Lemmon. Distributed network utility maximization using event-triggered augmented lagrangian methods. In *Proc. of 2009 American Control Conference*, pages 3298–3303, 2009.

[61] J. Mota, J. Xavier, P. Aguiar, and M. Puschel. Distributed ADMM for model predictive control and congestion control. In *Proc. of 51st IEEE Conference on Decision and Control*, pages 5110–5115, 2012.

[62] X. Zhang and A. Papachristodoulou. A distributed PID controller for network congestion control problems. In *Proc. of 2014 American Control Conference*, pages 5453–5458, 2014.

[63] X. Zhang and A. Papachristodoulou. Improving the performance of network congestion control algorithms. *IEEE Transactions on Automatic Control*, 60(2):522–527, 2015.

[64] F. Paganini, Z. Wang, J. C. Doyle, and S. H. Low. Congestion control for high performance, stability and fairness in general networks. *IEEE/ACM Transactions on Networking*, 13(1):43–56, 2005.

[65] X. Zhang, N. Li, and A. Papachristodoulou. Distributed optimal steady-state control using reverse- and forward-engineering. In *Proc. of 54th IEEE Conference on Decision and Control*, 2015, to appear.

[66] K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in Linear and Nonlinear Programming*. Stanford University Press, 1958.

[67] A. Jokic, M. Lazar, and P. Van Den Bosch. On constrained steady-state regulation: Dynamic KKT controllers. *IEEE Transactions on Automatic Control*, 54(9):2250–2254, 2009.

[68] F. D. Brunner, H-B. Dürr, and C. Ebenbauer. Feedback design for multi-agent systems: A saddle point approach. In *Proc. of 51st IEEE Conference on Decision and Control*, pages 3783–3789, 2012.

[69] E. Dall'Anese, S. Dhople, and G. B. Giannakis. Regulation of dynamical systems to optimal solutions of semidefinite programs: Algorithms and applications to AC optimal power flow. In *Proc. of 2015 American Control Conference*, pages 2087–2092, 2015.

[70] S. You and L. Chen. Reverse and forward engineering of frequency control in power networks. In *Proc. of 53rd IEEE Conference on Decision and Control*, pages 191–198, 2014.

[71] X. Zhang, N. Li, and A. Papachristodoulou. Achieving real-time economic dispatch in power networks via a saddle point design approach. In *Proc. of IEEE Power and Energy Society General Meeting*, 2015.

[72] G. Vinnicombe. On the stability of end-to-end congestion control for the internet. In *Technical Report, University of Cambridge, CUED/F-INFENG/TR.398*, 2000.

[73] D. Callaway and I. Hiskens. Achieving controllability of electric loads. *Proc. of IEEE*, 99(1):184–199, 2011.

[74] A. R. Bergen and D. J. Hill. A structure preserving model for power system stability analysis. *IEEE Transactions on Power Apparatus and Systems*, 100(1):25–35, 1981.

[75] J. A. Short, D. G. Infield, and L. L. Ferris. Stabilization of grid frequency through dynamic demand control. *IEEE Transactions on Power Systems*, 23(3):1284–1293, 2007.

[76] P. V. Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, 2011.

[77] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[78] H. Farhangi. The path of the smart grid. *IEEE Power and Energy Magazine*, 8(1):18–28, 2010.

[79] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, third edition, 2002.

[80] J. K. Hale and S. M. V. Lunel. *Introduction to Functional Differential Equations*, volume 99. Springer-Verlag New York, Inc, 1993.

[81] A. Papachristodoulou and A. Jadbabaie. Synchronization in oscillator networks: Switching topologies and non-homogeneous delays. In *Proc. of 44th IEEE Conference on Decision and Control*, pages 5692–5697, 2005.

[82] A. Papachristodoulou, A. Jadbabaie, and U. Münz. Effects of delay in multi-agent consensus and oscillator synchronization. *IEEE Transactions on Automatic Control*, 55(6):1471–1477, 2010.

[83] A. Kiani and A. Annaswamy. Wholesale energy market in a smart grid: Dynamic modeling, and stability. In *Proc. of 50th IEEE Conference on Decision and Control*, pages 2202–2207, 2011.

[84] J. Lavaei, D. Tse, and B. Zhang. Geometry of power flows in tree networks. In *Proc. of IEEE Power and Energy Society General Meeting*, 2012.

[85] S. Sojoudi and J. Lavaei. Physics of power networks makes hard optimization problems easy to solve. In *Proc. of IEEE Power and Energy Society General Meeting*, 2012.

[86] F. Dörfler, F. Chertkov, and F. Bullo. Synchronization in complex oscillator networks and smart grids. *Proc. of the National Academy of Sciences*, 110(6):2005–2010, 2013.

[87] A. M. Cervantes, A. Wächter, R. H. Tütüncü, and L. T. Biegler. A reduced space interior point strategy for optimization of differential algebraic systems. *Computers and Chemical Engineering*, 24(1):39–51, 2000.

[88] R. Tonita. Applications of control theory to economics. *PhD dissertation, Department of Engineering, University of Cambridge*, 2010.

[89] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control: Analysis and Design*. John Wiley & Sons, Ltd, 2005.

[90] X. Fan, M. Arcak, and J. T. Wen. Robustness of network flow control against disturbances and time-delay. *System and Control Letters*, 53(1):13–29, 2004.

[91] F. Paganini. A global stability result in network flow control. *System and Control Letters*, 46(3):165–172, 2002.

[92] S. H. Low and D. E. Lapsley. Optimization flow control, I: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, 1999.

[93] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer-Verlag New York, Inc, 1999.

[94] F. Paganini, Z. Wang, S. H. Low, and J. C. Doyle. A new TCP/AQM for stable operation in fast networks. In *Proc. of IEEE INFOCOM*, pages 96–105, 2003.

[95] G. Vinnicombe. On the stability of networks operating TCP-like congestion control. In *Proc. of 15th IFAC World Congress*, 2002.

[96] J. Lee, M. Chiang, and R. Calderbank. Jointly optimal congestion and contention control in wireless ad hoc networks. *IEEE Communications Letters*, 10(3):216–218, 2006.

[97] M. Andreasson, D. V. Dimarogonas, H. Sandberg, and K. H. Johansson. Distributed control of networked dynamical systems: Static feedback, integral action and consensus. *IEEE Transactions on Automatic Control*, 59(7):1750–1764, 2014.

[98] C-T. Chen. *Linear System Theory and Design*. Oxford University Press, Inc. New York, third edition, 1999.

[99] J. F. Sturm. Using SEDUMI 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(12):625–653, 1999.

[100] S. H. Low. Convex relaxation of optimal power flow – Part I: Formulations and equivalence. *IEEE Transactions on Control of Network Systems*, 1(1):15–27, 2014.

[101] S. H. Low. Convex relaxation of optimal power flow – Part II: Exactness. *IEEE Transactions on Control of Network Systems*, 1(2):177–189, 2014.

[102] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, Englewood Cliffs, New Jersey, 1996.

[103] J. C. Doyle, B. Francis, and A. Tannenbaum. *Feedback Control Theory*. Macmillan Publishing Company, New York, 1992.