Fundamental Limitations of Distributed Feedback Control in Large-Scale Networks

EMMA TEGLING

Doctoral Thesis
Stockholm, Sweden 2018

KTH
TRITA-EECS-AVL-2019:5 Skolan för elektroteknik och datavetenskap
ISBN 978-91-7873-059-9 100 44 Stockholm

Akademisk avhandling som med tillstånd av Kungliga Tekniska högskolan framläggs till offentlig granskning för avläggande av teknologie doktorsexamen i elektro- och systemteknik fredagen den 18 januari 2019 klockan 14.00 i sal F3, Kungliga Tekniska högskolan, Lindstedtsvägen 26, Stockholm.

Tryck: Universitetsservice US AB

Abstract

Networked systems accomplish global behaviors through local feedback interactions. The purpose of a distributed control design is to select interaction rules and control protocols that achieve desired global control objectives. In this thesis, we address the question of fundamental limitations to such control designs, in terms of the global performance that is achievable in large-scale networks.

We consider networked dynamical systems with single- and double-integrator dynamics controlled with linear consensus-like protocols. Such systems can be used to model, for example, vehicular formation dynamics and synchronization in electric power networks. We assume that the systems are subject to distributed disturbances and study performance in terms of $\mathcal{H}_2$-norm metrics that capture the notion of network coherence. In the context of power networks, we also show how such metrics can be used to quantify resistive losses caused by non-equilibrium, or transient, power flows due to a lack of synchrony.

Distributed *static* feedback control based on localized, relative state measurements is subject to known limitations that, for example, cause coherence metrics to scale unfavorably with network size in lattices of low spatial dimensions. This causes an inevitable lack of rigidity in one-dimensional formations, such as strings of vehicles. We show here that the same limitations in general apply also to *dynamic* feedback controllers that are locally of first order. The proof relies partly on a fundamental limitation of localized relative feedback in networks of integrators of order three or higher, which we show to cause instability if the network grows beyond a certain finite size.

This result holds unless the controller can access measurements of its local state with respect to an absolute reference frame, in which case dynamic feedback in the form of distributed derivative or integral control can fundamentally improve performance. This case applies, for example, to frequency control in power networks. However, if the absolute state measurements are subject to noise, the advantage of the distributed integral controller in terms of its performance scaling is lost. We show that scalable integral control of networks in principle requires centralization or all-to-all communication.

For electric power networks, we show that performance in terms of transient power losses scales with the number of generator nodes in a network. However, in sharp contrast to the previous results, an increased connectivity does not in general improve performance. We discuss possible implications of these results in terms of the design of future power grids with increasingly distributed electricity generation.

Sammanfattning

Distribuerad reglering av nätverk går ut på att definiera lokala styrregler och kommunikationsprotokoll, som leder till att nätverket i stort uppvisar önskat beteende. Denna avhandling undersöker fundamentala begränsningar hos sådana distribuerade regulatorer, utifrån den prestanda som kan uppnås globalt i storskaliga nätverk.

Vi modellerar nätverk där den lokala dynamiken är en enkel- eller dubbelintegrator, och som regleras genom linjär återkoppling från ett begränsat antal grannsystem. En sådan modell kan till exempel representera formationskörning med autonoma fordon eller frekvenssynkronisering i elnät. Vi antar att systemet utsätts för störningar och mäter dess prestanda utifrån systemets $\mathcal{H}_2$-norm med avseende på en lämplig utsignal. På det sättet kan vi mäta ett nätverks koherens, som är ett mått på hur stor varians noderna har i förhållande till nätverksgenomsnittet. Vi visar också att den här metoden kan uppskatta de elektricitetsförluster som uppstår på grund av cirkulerande strömmar i synkroniseringstransienten hos elnät.

Det är känt att det finns begränsningar hos distribuerad reglering som är proportionell mot relativa tillståndsmätningar mellan grannsystem. Dessa innebär en dålig skalning av koherensmåttet i nätverk med lågdimensionell gitterstruktur och leder till att en endimensionell fordonsformation inte kan uppvisa en stelkroppsliknande rörelse. Avhandlingen visar att dynamiska regulatorer med ett ytterligare regulatortillstånd generellt sett omfattas av samma begränsningar. Beviset bygger delvis på en fundamental begränsning hos relativ återkoppling i nätverk med tre eller fler lokala integratorer. Vi visar att sådana system alltid blir instabila om nätverket tillåts växa utöver en viss (ändlig) storlek.

Ett undantag till resultatet ovan inträffar om regulatorn har tillgång till sitt eget lokala tillstånd med avseende på en absolut referensram. Det är till exempel fallet med frekvensreglering i elnät, där den lokala frekvensavvikelsen kan mätas. I dessa fall kan en dynamisk regulator innehålla en deriverande eller integrerande del som leder till en fundamental prestandaförbättring och ökar skalbarheten. Om den absoluta tillståndsmätningen är brusig tappar dock den integrerande regulatorn sina fördelar och försämrar istället skalbarheten. Vi visar att integralverkan i stora nätverk därför i princip måste vara centraliserad eller tillåta kommunikation mellan samtliga noder.

När det gäller synkronisering i elnät visar vi att de transientförluster som uppstår på grund av till exempel laststörningar ökar linjärt med antalet generatornoder. Prestandan kan dock inte, som tidigare, förbättras av att öka konnektiviteten hos nätverket. Vi diskuterar vad detta kan innebära för framtidens elnät med distribuerad kraftproduktion.

For Martin

Acknowledgements

I am lucky to have my own large-scale network to rely on in my work.

First of all my main advisor Henrik Sandberg, whose intelligent, supportive and friendly advice has been invaluable. Thank you for always finding room in your calendar, for your attention both to detail and to the big picture, and especially for your many late evenings spent reading near deadlines. I could not have wished for better.

I am beyond grateful to Dennice Gayme, without whom I would never have started doing research. Thank you for your encouragement and friendship, your time and your patience, and, of course, for all your smart ideas. My warmest thanks also to Bassam Bamieh, for welcoming me to UCSB, and for sharing so much of your knowledge and insight with me. Your curiosity is contagious!

My co-advisor Karl Henrik Johansson is always a source of inspiration. Thank you for your relentless optimism and all the good advice! I am also lucky to have other excellent collaborators in John W. Simpson-Porco, Hendrik Flamme and Partha Mitra. I wish to thank you, as well as Florian Dörfler, Bart Besselink, Richard Pates and Anders Rantzer for numerous interesting and fruitful discussions. I look forward to picking them up again!

I was privileged to have a great friend and collaborator in Martin Andreasson. I am saddened beyond words that he is no longer with us. I keep being inspired by his enthusiasm, for work as well as life.

My colleagues at the Automatic Control department, thank you for always making my day! I cannot mention all of you here, but please know that you and all our fun chats over coffee breaks are important to me. I would, however, like to give a special thanks to those of you who helped me proof read parts of this thesis. David, Ehsan, Jezdimir, Matin, Michelle, and Mohammad, thanks for your time and your sharp eyes!

My meetings with the intelligent, strong and kind women in our field inspire me more than you may think. In no particular order, thank you Jana, Matin, Michelle, Alessandra, Line, Stacy, Dennice, Necmiye, Maria, Sonja, Linnea, Sofie, Mariette, Joana, Mina, Pian, Danica, Lina, Jacquelien, Luz, Weiwei, Kasia, Emily, Yuke, Elisabet, Laila, Floria and Sahar. The future is female.

This thesis would not have been possible without financial support, in my case from the Swedish Research Council through grants 2013-5523 and 2016-0861, and and the Swedish Foundation for Strategic Research through the project ICT-Psi. I also thank the Knut and Alice Wallenberg foundation and the ACCESS Linnaeus Centre for travel grants.

Finally, there are those who had less to do with the work that led up to this thesis, but who have my deepest gratitude all the same. You know who you are. Måns, thank you for being my greatest supporter. I love you.

Emma Tegling
Stockholm, December 2018

Contents

Chapter 1

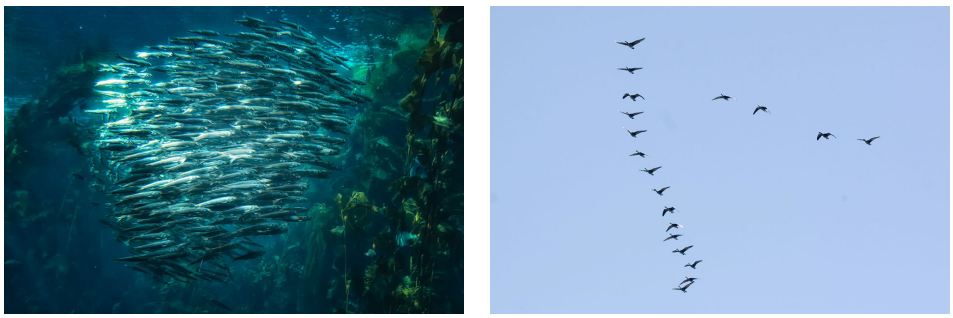
Introduction

Understanding networked systems is essential for understanding today's world. Modern sensing, computation and communication technologies have made networked engineering systems ubiquitous, and concepts such as the Internet of Things, self-driving cars, or smart power grids no longer appear fictional. Along with the increased relevance of networks in engineering, a networking perspective has permeated both the natural and social sciences. Phenomena in areas ranging from signaling in nervous systems and epidemic spread to financial systems and social networks are increasingly described in terms of interactions of individual subsystems in large-scale networks. This thesis takes a control-theoretical approach to understanding such interactions.

The desired behavior of a network can typically be described in terms of *global* control objectives. Such an objective may be synchrony in an electric power network, an agreement over measurements in a sensor network, or a common velocity in a string of self-driving cars. In large and complex networks, centralized control is often infeasible or undesirable, and the communication burden of all-to-all connectivity becomes overwhelming. Objectives are instead fulfilled through control actions that are *distributed* across the individual subsystems, and that rely on feedback interactions from *local* neighborhoods. Here, we address the question of what this implies in terms of the overall *performance* of the network. In particular, can performance be maintained as networks grow large?

We explore the limitations of local, distributed control in terms of achieving global control and performance objectives in large-scale networks. We study how such limitations depend on factors such as network topology, plant and controller dynamics, and the quality and type of measurements. Since interactions of even simple subsystems in large-scale networks can give rise to complex phenomena, these dependencies are oftentimes counterintuitive. Understanding them is therefore essential for an efficient analysis and design of the networked systems of the future.

In this introductory chapter, we present the problem formulation and point to its relevance in applications. We also summarize our technical contributions and provide an outline for the remainder of the thesis.



(a) School of sardines (b) Geese in formation flight

Figure 1.1: The study of distributed coordination of multi-agent networks often takes inspiration from flocking behaviors in the animal kingdom. In the fish school and the bird formation above, the indivuduals can be seen to align themselves into formation based on observations of their neighbors' headings [77]. The limitations to this type of feedback, which are treated in this thesis, can partly explain why the 3-dimensional fish school can resemble a rigid body, while the 1-dimensional bird formation has a meandering motion, see also [14]. We discuss whether alternative feedback laws in engineered multi-agent networks can alleviate these limitations, and increase the rigidity of 1-dimensional formations. Photos by (a) Erik Kilby/CC BY-SA 2.0 (b) Kelly Martin/CC BY-SA 3.0.

1.1 A performance perspective on networked systems

A central issue in the control of networked systems is to understand and quantify how the limited sensing, actuation and connectivity of a distributed controller structure affect global performance. A prototypical problem is that of distributed consensus, where the objective is to drive a network of agents to a common state, but where each agent only has access to limited and localized measurements. Natural questions arise as to how well a state of consensus can be upheld, for example, under external disturbances, and how this depends on the size of sensing neighborhoods and the topological connectivity of the controller. In other words, how does the best-achievable performance of a system depend on the architectural controller constraints that are imposed by a network structure? An understanding of these issues is key in achieving efficient and robust control performance in a wide range of applications, including vehicle platooning and formation control problems, wireless sensor networks and electric power systems.

We will address this problem from several angles in this thesis. Before detailing the problem formulation and the objectives, we will use this section to highlight a few motivational problems from applications.

1.1.1 Examples

The problems we consider in this thesis are all examples of distributed coordination and control problems, which can be modeled through consensus-type dynamics. Since the seminal works by Fax and Murray [51], Olfati-Saber and Murray [118], and Jadbabaie *et al.* [77] at the beginning of the 2000's, these problems have become some of the most active research topics in the field. While problems of stability and convergence subject to various properties of the network are well-studied, several questions related to emerging behaviors and performance limitations in large-scale networks have remained open. Many of these questions are highly relevant in applications, and problems range from animal flocking behaviors (Figure 1.1) to electric power transmission (Figure 1.7). We next describe a few examples.

Example 1.1 (Vehicle platooning) The problem of controlling strings of vehicles – the platooning problem – has raised much interest since some pioneering work in the 1960's [89, 104]. The main control objective in a vehicle platoon is to maintain a desired cruising velocity with a pre-determined constant spacing between the vehicles. By keeping this spacing small, the aerodynamic drag can be reduced, saving fuel costs and emissions. This has motivated significant platooning efforts in road freight transportation [1, 22]. With the advent of automated driving, large-scale platoons may emerge on highways. This would create long-envisioned automated highway systems [70, 74, 173, 182] that have the potential to increase road capacity and throughput.

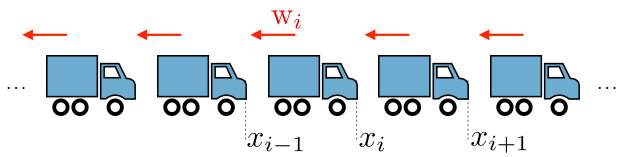


Figure 1.2: A truck platoon subject to disturbances.

The principle for a vehicle platoon is illustrated in Figure 1.2. An example of a simple linear control law is the combined look-ahead and look-behind control, see for example [68]. This gives the simplified closed-loop dynamics of the i^{th} vehicle of the platoon as:

$$\ddot{x}_i(t) = \dot{v}_i(t) = f_+(x_{i+1}(t) - x_i(t) - \Delta_{(i+1,i)}) + f_-(x_{i-1}(t) - x_i(t) - \Delta_{(i-1,i)}) + g_+(v_{i+1}(t) - v_i(t)) + g_-(v_{i-1(t)} - v_i(t)) + \mathrm{w}_i(t), \quad (1.1)$$

where $x_i(t)$ is the vehicle's position and $\dot{x}_i(t) = v_i(t)$ its velocity at the time t, f_+, f_-, g_+, g_- are constant gains and $\Delta_{i,j}$ denotes the desired spacing between vehicles i and j. An external disturbance that can be due to, for example, wind gusts, varying road conditions, and measurement errors is modeled through the signal $\mathrm{w}_i(t)$.

Apart from a multitude of practical challenges of which [182] gives an early overview, a number of theoretical questions arise. Some examples are:

- How can it be ensured that disturbances are not amplified along the string of vehicles, in particular as the string grows large? This is an issue of so-called *string stability* [17, 106, 149, 169, 172].
- What control architecture performs better: a predecessor-follower architecture, or a bidirectional architecture as in (1.1)? [68]
- What are optimal controller gains? [91]
- What types of measurements are needed to achieve reasonable performance? Do relative measurements with respect to neighbors (using, for example, radar) suffice, or is each vehicle required to access its absolute position and velocity with respect to the lead vehicle or a global frame (based on, for example, speedometers or GPS)? [14, 67]

We return to some of these problems in the literature review of Section 2.1.3. In this thesis, we address the limitations of feedback of the type (1.1) in terms of its ability to maintain *coherence* in large vehicle platoons. Coherence, in this case, describes the rigidity of the formation and is a concept that is distinct from string stability mentioned above. Lack of coherence results in undesired motions in the platoon, examples of which can be previewed in Figure 6.4 on page 115.

Example 1.2 (Frequency synchronization in power networks) The electric power system, possibly rivaled by telecommunications, is the most significant networked system ever engineered, as well as one of the largest and most complex. At present, this system is undergoing important changes. An integration of increased levels of renewable energy leads to an increasingly distributed system as conventional high-capacity generation is replaced by local, small-scale resources [49, 75, 108]. The future grid is also expected to have higher levels of uncertainty [180] and to feature new load patterns [145].

On the one hand, these changes can affect the synchronous stability of the power network, that is, its ability to recover a synchronous frequency after a disturbance. On the other hand, the emerging distributed generation paradigm both enables and demands novel distributed, scalable control schemes. These topics have therefore seen an increased research interest over the last decade.

Synchronization in power networks is typically studied through a system of coupled swing equations, which describe the electromechanical oscillations in synchronous generators. The linearized swing equation can, under some simplifying assumptions, be written as:

$$m_i\ddot{\theta}_i(t) + d_i\dot{\theta}_i(t) = \sum_{j\in\mathcal{N}_i} b_{ij}(\theta_j(t) - \theta_i(t)) + P_{m,i}. \tag{1.2}$$

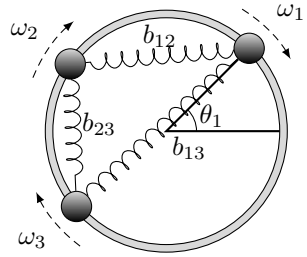


Figure 1.3: Mechanical analogy of the coupled swing equation, due to Dörfler *et al.* [40]. Any deviation in phase angle θ_i or frequency $\dot{\theta}_i = \omega_i$ is propagated over the springs (power lines) across the network.

where $\theta_i(t)$ is the phase angle and $\dot{\theta}_i(t) = \omega_i(t)$ is the frequency at the i^{th} generator, and m_i and d_i are, respectively, its inertia and damping coefficients. The constants b_{ij} denotes the susceptances of the power lines connecting the i^{th} generator to its neighbors in the set $\mathcal{N}_i$, and $P_{m,i}$ can be regarded as the net power injection at the generator. This swing equation can be illustrated by the mechanical analogy in Figure 1.3.

Even if the network is synchronously stable, a lack of synchrony arises whenever the system is subject to disturbances. To secure safe and efficient operation of the power grid, it is important that the frequency is kept close to its nominal operating point, typically 50 Hz or 60 Hz. It is therefore desirable to keep the transient fluctuations "small". This is a question of the power system's *performance*, which is what we will address in this thesis. In particular, we analyze how the system's transient performance can potentially be affected by a distributed generation paradigm where the number of generator nodes may increase by several orders of magnitude.

Example 1.3 (Dynamic load balancing in distributed computing) In distributed computing, a network of computers share a computational task in order to increase total processing capabilities and thereby efficiency. An important issue in distributed computing is *load balancing*, that is, ensuring that each processor is assigned a "fair share" of the total work load. If the computational problem is complex, it is difficult to distribute sub-tasks a priori and the work load must be balanced during the execution. Such a strategy is referred to as *dynamic* load balancing [34].

A simple algorithm for dynamic load balancing is through the standard average consensus algorithm, which can be written as follows [34, 128, 190]:

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i} a_{ij}(x_j(t) - x_i(t)) + \mathrm{w}_i(t+1), \quad i = 1, \ldots, N \tag{1.3}$$

where the state $x_i(t)$ is the amount of work assigned to computer i at time instance t, $\mathcal{N}_i$ is the set of neighbors that the i^{th} computer communicates with, and a_{ij} are

nonnegative constants. The parameter $\mathrm{w}_i(t)$ accounts for the new work that is generated at time t, less the amount of work that can be executed in one time step. If the computer is precisely able to keep up with the expected amount of new work, then $\mathrm{w}_i(t)$ can be seen as a zero-mean random variable, and can be thought of as a distributed disturbance in the process.

The objective of this algorithm is to converge to an equilibrium where all work loads are equal, that is $x_i = x_j$ for all $i, j \in \{1, \ldots, N\}$. Such a state of consensus can only hold approximately when the system is subject to persistent disturbances. Still, it is desirable to keep deviations from this state small and the system's performance can therefore be measured through

$$\mathbb{E}\left\{\left(x_i(t) - \frac{1}{N}\sum_{j=1}^{N} x_j(t)\right)^2\right\}, \tag{1.4}$$

that is, the variance of the state x_i's deviation from the network average.

Today's computer networks are several orders of magnitude larger than in 1989 when [34] was published. An interesting question is therefore whether a load balancing algorithm like (1.3) would be able to keep measures like (1.4) small in a network of millions, or billions, of nodes. And if so, what would the requirements be on the network topology? In this thesis, we address such questions for continuous-time equivalents of the system (1.3).

Example 1.4 (Voltage control in multi-terminal HVDC networks) Transmitting power over long distances while maintaining low losses is one of the greatest challenges in power transmission systems. For example, this is one of the issues that arise in the deployment of large off-shore wind farms, which require long-distance undersea power transmission. This motivates the use of high-voltage direct current (HVDC) technology. Its higher investment costs compared to AC transmission lines are compensated by lower resistive losses for sufficiently long distances. These are typically 500-800 km for overhead lines [121], but less than 100 km for undersea cable connections [24]. As more energy sources and consumers are connected by HVDC lines, the individual lines eventually form a grid with multiple terminals, resulting in *multi-terminal HVDC* (MTDC) systems [181]. Figure 1.4 shows an example of such a network.

MTDC networks require the ability to control the DC voltages at the terminals; first, in order to govern the network's current flows, and second, in order to avoid damage to power electronic equipment [86, 181]. The voltage dynamics can be assumed to be given by:

$$c_i \dot{v}_i(t) = -\sum_{j \in \mathcal{N}_i} \left(\frac{1}{r_{ij}}(v_i(t) - v_j(t)) - i_{ij}^{\mathrm{ref}}\right) + u_i(t), \tag{1.5}$$

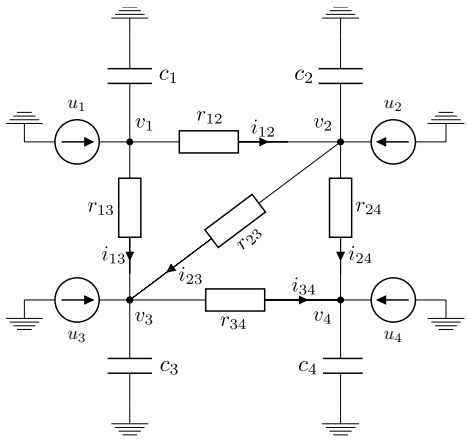


Figure 1.4: Example of an MTDC network consisting of 4 terminals (buses) and 5 lines.

where $c_i > 0$ is the capacitance of the i^{th} terminal, r_{ij} is the line resistance and i_{ij}^{ref} is the nominal current on the $(i,j)^{\text{th}}$ line. The signal $u_i(t)$ is the controlled injected current. Different control schemes have been proposed in the literature. Examples are:

i) *Droop control,* which is a simple proportional controller:

$$u_i(t) = k_{P,i}(v_i^{\text{ref}} - v_i(t)),$$

where $k_{P,i} > 0$ is the droop gain [66, 198].

ii) *Distributed-averaging proportional-integral (DAPI) control,* which appends integral action to the droop controller and thus eliminates stationary errors:

$$\begin{aligned} u_i(t) &= k_{P,i}(v_i^{\text{ref}} - v_i(t)) - z_i(t), \\ k_{I,i}\dot{z}_i(t) &= v_i^{\text{ref}} - v_i(t) - \sum_{j\in\mathcal{N}_i} a_{ij}(z_i(t) - z_j(t)), \end{aligned}$$

where $z_i(t)$ is an auxiliary controller state and $k_{I,i}, a_{ij} > 0$ are fixed gains. The DAPI controller has been proposed for frequency control in AC power networks in [7, 159] and for DC microgrids in [198]. We will revisit it in Chapter 4.

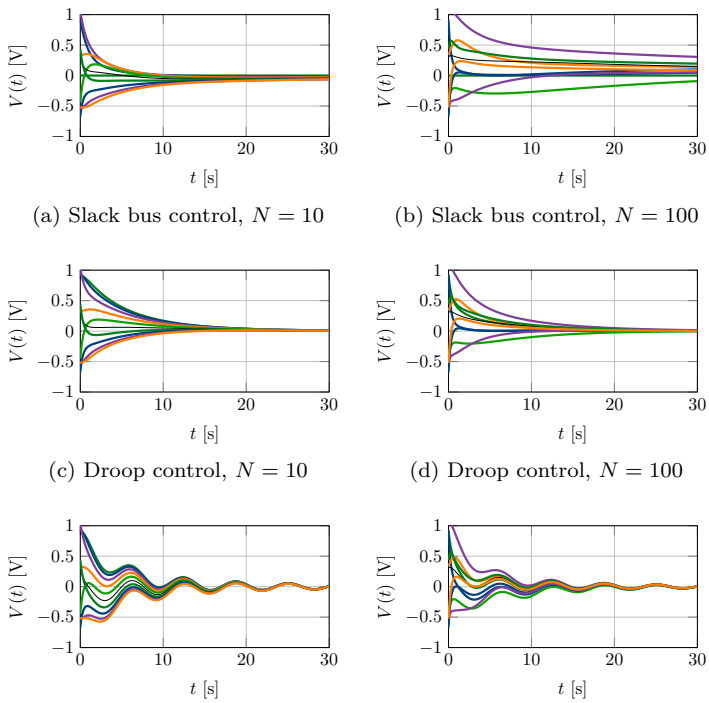
iii) *Slack bus control,* where the voltage at a single terminal is controlled and allowed to regulate the network's voltage drift. Somewhat idealized, we can then set

$$v_1(t) = v_1^{\text{ref}}, \quad \forall t \geq 0.$$

Remaining buses' voltages then evolve according to (1.5) with $u_i(t) = 0$ for $i = 2, \ldots, N$.

While all of these controllers succeed in stabilizing the voltage, their performance differ. In particular, the slack bus controller does in many cases have worse *scalability* to large networks. That is, its transient performance can be shown to deteriorate as the network size grows. This is showcased through the simulation in Figure 1.5, where the size of the network increases from $N = 10$ to $N = 100$.

The system (1.5) describes consensus dynamics (the slack bus control case describes *leader-follower* consensus). In this thesis, we will discuss reasons for lack of scalability of the consensus algorithm. We will show that *absolute feedback*, which in the MTDC case is introduced through voltage droop control with the term $-k_{P,i}v_i(t)$, is important for scalability, and that DAPI control has the potential to further improve performance. We will not revisit the particular MTDC control problem in this thesis, but refer the interested reader to [9].



(a) Slack bus control, $N = 10$ (b) Slack bus control, $N = 100$

(c) Droop control, $N = 10$ (d) Droop control, $N = 100$

(e) DAPI control, $N = 10$ (f) DAPI control, $N = 100$

Figure 1.5: Simulation of voltage droop controllers in multi-terminal HVDC networks with $N = 10$ and $N = 100$ terminals. The network topology is a path graph (shown in Figure 3.1). In the slack bus case, performance deteriorates substantially for the larger network (b) compared to the smaller (a), which is not the case with droop or DAPI control. The oscillatory behavior in the DAPI case is characteristic of integral control. In this thesis, we formalize and explain the limited scalability of the slack bus controller.

1.2 Problem formulation

1.2.1 Prototypical model

This thesis treats a number of distributed control problems with analogous mathematical settings. We will consider first-order systems, where there is one local state x_i at each node i of the network, as well as second-order systems, in which there are two such states: x_i and $\dot{x}_i$. The state x_i may, for example, represent the phase angle of the i^{th} generator in an electric power network, while $\dot{x}_i$ represents its angular frequency.

We will assume that the systems are subject to distributed disturbances w_i, which in most cases will be modeled as white noise[1] that is uncorrelated across the network. The systems can then be written as

$$\dot{x}_i(t) = u_i(t) + \mathrm{w}_i(t), \tag{1.6}$$

in the first-order case, and

$$\ddot{x}_i(t) = u_i(t) + \mathrm{w}_i(t), \tag{1.7}$$

in the second-order case. Here, $u_i(t)$ represents a control input.

The control objective is *global*; to drive the states at all nodes to a common equilibrium, such as synchrony in a power network or a common cruising speed in a vehicle platoon. However, in the problems we consider, the control is *localized* and *distributed*, meaning that the control signal at any given node will only depend on feedback from that node itself and from neighboring nodes in a bounded neighborhood. As we shall see, this fact often leads to limitations to the system performance.

We consider linear feedback throughout. With static state feedback, we can therefore write the control input for the second-order system (1.7) as

$$u_i(t) = -\sum_{j\in\mathcal{N}_i} f_{ij}(x_i(t)-x_j(t)) - \sum_{j\in\mathcal{N}_i} g_{ij}(\dot{x}_i(t)-\dot{x}_j(t)) - f_{0,i}x_i(t) - g_{0,i}\dot{x}_i(t), \tag{1.8}$$

where $\mathcal{N}_i$ denotes the neighbor set of the i^{th} node and $f_{ij}, g_{ij}, f_{0,i}, g_{0,i}$ are constant gains. In the first-order case (1.6), the control input can be written as in (1.8), but without the terms containing $\dot{x}_i(t)$. We refer to the control law (1.8) as *static*, or *proportional*, feedback control since the feedback is proportional to state measurements at time t.

Throughout the thesis, we will distinguish between and compare *relative* and *absolute* state feedback. In (1.8) the terms $-f_{0,i}x_i(t)$ and $-g_{0,i}\dot{x}_i(t)$ represent absolute feedback from the states x_k and $\dot{x}_k$ respectively. If $f_{0,i} = g_{0,i} = 0$, the control

[1] We refer to "white noise" in continuous time as a stationary zero-mean stochastic process with autocorrelation $\mathbb{E}\{\mathrm{w}(\tau)\mathrm{w}^*(t)\} = \delta(t-\tau)I$, where $\delta(t)$ denotes the Dirac delta distribution. This idealized process can be thought of as the time derivative of a Brownian motion, dB/dt, although such a derivative does not formally exist, see [10, Theorem 4.1].

relies only on relative differences between neighboring nodes. The availability of relative and absolute feedback measurements depends on the application. For example, the swing equation (1.2) describes relative feedback from phase angles, but absolute feedback from angular frequency. Absolute feedback can also be regarded as a type of self-damping.

We will also consider systems where the local dynamics are of order three or higher, and where the control dynamics are analogous to (1.8). In this case, the problem in focus will be one of stability, not of performance.

1.2.2 Performance measures

The systems (1.6)–(1.7) are subject to persistent stochastic disturbances w_i, which will cause the states at each network site to fluctuate around the equilibrium. Throughout this thesis, we will characterize system performance through the variance of such fluctuations, which in a sense captures the amount of "disorder" in the network. More precisely, we consider the quantities:

Global error (coherence)

$$V_i^{\mathrm{dav}}(t) := \mathbb{E}\left\{\left(x_i(t) - \frac{1}{N}\sum_{j=1}^{N} x_j(t)\right)^2\right\}, \tag{1.9}$$

that is, the variance of the i^{th} node's deviation from the average over all N nodes in the network. It captures the amount of global disorder and therefore a notion of network *coherence*. In the example of a vehicle platoon, V_i^{dav} characterizes how closely the platoon resembles a rigid body.

Local error (lack of synchrony)

$$V_i^{\mathrm{loc}}(t) := \mathbb{E}\left\{\left(\sum_{j\in\mathcal{N}_i} a_{ij}(x_i(t) - x_j(t))\right)^2\right\}, \tag{1.10}$$

that is, the variance of the i^{th} node's deviation from a local, weighted average over its neighboring nodes (the a_{ij} being appropriate weights). The measure (1.10) captures local disorder in the network, and in a sense quantifies the amount of interaction between neighboring nodes.

In power networks, V_i^{loc} signifies a *lack of synchrony*, which incurs non-equilibrium power flows. In Chapter 8, we will discuss how the losses associated with such flows are a meaningful measure of performance that we term the "Price of Synchrony." That measure can be written as a weighted sum of local errors on the form (1.10).

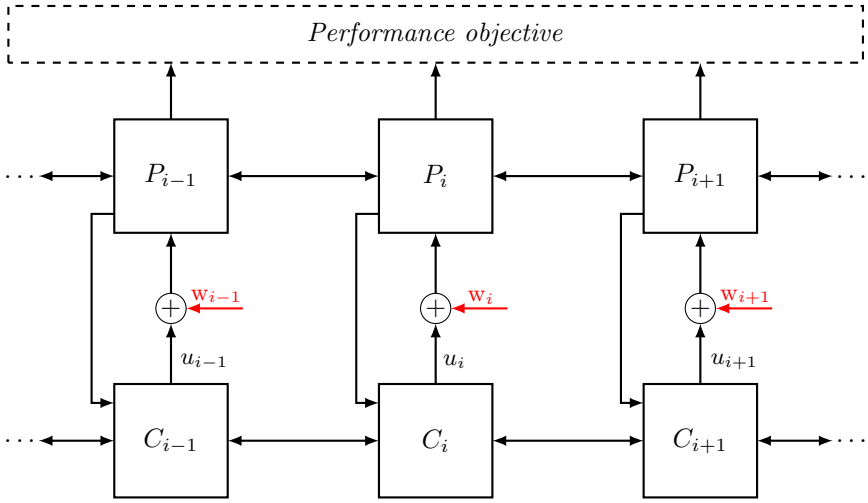


Figure 1.6: Illustration of the problem formulation. The plants P_i are, in our case, integrator chains controlled by the distributed controllers C_i. The arrows between plants and controllers represent relative feedback interactions between neighboring nodes. The performance objective is evaluated from a global perspective.

While the performance measures (1.9) and (1.10) may appear very similar, they capture two distinct phenomena in networked systems. For example, we will show in Chapter 8 that it is possible for two networks with very different global error properties to have the same performance in terms of local errors. Conversely, a given system may appear well-regulated locally, although the network, when viewed as a whole, has unregulated modes. Such situations are explored in Chapters 6–7. The problem setup is illustrated in Figure 1.6.

1.2.3 Objectives

The objective of this thesis is to characterize the performance and scalability of distributed control laws on the form (1.8) in terms of the their ability to keep the variance metrics (1.9)–(1.10) small under persistent stochastic disturbances. In particular, we are interested in scaling bounds on these metrics in network size.

We show that this performance characterization, at steady state and typically aggregated over the network, amounts to an evaluation of $\mathcal{H}_2$ performance with suitably defined performance outputs. It will therefore also be relevant for describing transient performance under other input scenarios, such as impulse responses.

We mainly focus on two problem settings. First, the characterization of coherence properties in networks of first- and second-order integrators. Here, the main question of interest is whether *dynamic* feedback laws perform better than the static state feedback (1.8) in large-scale networks. Second, we consider electric

power networks, where we evaluate performance in terms of resistive power losses incurred in maintaining network synchrony. Here, the objective is to characterize the losses through their dependence on physical properties of the network. In both of these settings, we address how performance depends on

- the size of the network,
- the topology of the network and the connectivity of local controllers,
- the plant and controller dynamics, and
- the availability and quality of absolute state measurements.

We also consider a third problem setting with networks of third- and higher-order integrators. Here, we study how system stability depends on the above properties.

Our overall objective is to identify and characterize *fundamental limitations* to the performance achievable through distributed feedback control. In our case, such limitations would imply that a given performance specification cannot be satisfied with finite control effort in a network of any size. Awareness of fundamental limitations is at the very core of control theory, and is used not only to discard impossible specifications but also to prove optimality of control designs. Unacceptable performance limitations indicate a need to change the system design or resort to alternative sensing and control strategies for which the limitations do not apply. A key part of our objective has therefore been to investigate conditions under which dynamic feedback control can relax known limitations to distributed static feedback control.

(a) Platoon of trucks

(b) The Nordic power transmission grid

Figure 1.7: The problems of vehicle platooning and frequency synchronization of electric power grids are both treated in this thesis. While these systems appear very different, some of their key dynamical features can be described by our prototypical model (1.7), see Examples 1.1–1.2. (a) Photo courtesy of Scania (b) Picture adapted from Nordel [115].

1.3 Outline and contributions

The seven technical chapters of this thesis address different aspects of the problem that was just set up. The first four treat general first- and second-order consensus dynamics and also include a discussion on higher-order dynamics. The last three focus on power systems. The chapters also differ in their feedback scenarios and in the consideration of static versus dynamic controllers. An overview is provided in Table 1.1.

The overall outline of the remainder of the thesis and its technical contributions can be summarized as follows:

Chapter 2: Background

The background chapter provides an overview of the research that motivates or otherwise relates to the work in this thesis. We put particular emphasis on describing various notions of control performance in networks and their use in literature.

Chapter 3: Technical preliminaries

Chapter 3 reviews important technical and mathematical concepts that are used throughout the thesis. We review aspects of signal and system norms, graph theory and some operator theory that is useful for studying systems over regular networks.

Chapter 4: Coherence of large-scale networks with distributed PI and PD control

Here, we address known limitations of the standard second-order consensus protocol (1.8), which cause the variance of nodal fluctuations in (1.9) to scale unboundedly with network size for a large class of networks. We propose distributed proportional-integral (PI) and proportional-derivative (PD) controllers that relax these limitations to achieve a bounded variance, that is, fully coherent networks, in cases where absolute feedback of *one* of the two states is available. As we have seen in Section 1.1.1, this case applies, for example, in frequency control of power networks and in vehicular formation control with limited sensing. We also discuss optimal tuning of the controllers with respect to network coherence.

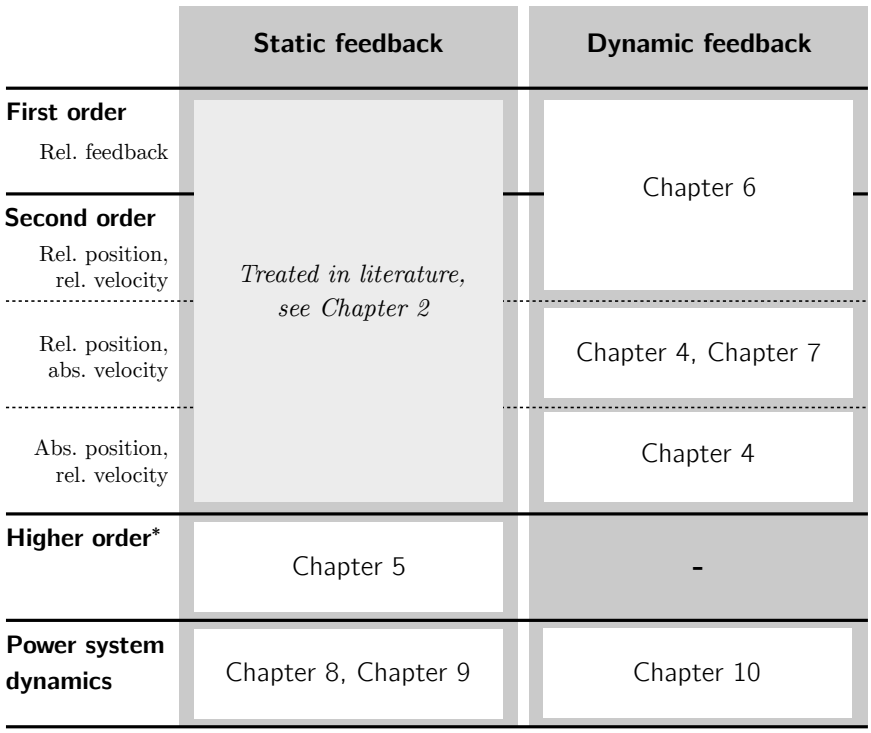
Chapter 4 is based on:

> E. Tegling and H. Sandberg, "On the coherence of large-scale networks with distributed PI and PD control," in *IEEE Control Systems Letters*, vol. 1, no. 1, pp. 170–175, July 2017.

Chapter 5: Inadmissibility of localized high-order consensus

Chapter 5 differs from remaining technical chapters of the thesis in that the question of interest is not primarily one of performance, but of stability. We consider

Table 1.1: Overview of the thesis chapters with respect to 1) the order of the local integrator dynamics, 2) the availability of relative vs. absolute feedback from position x_i and velocity $\dot{x}_i = v_i$ and 3) the consideration of static or dynamic feedback control. We remark that Chapter 5 studies admissibility (a problem of stability) rather than performance.



*The admissibility question

distributed consensus in systems where the local agents have high-order ($n \geq 3$) integrator dynamics (compare to (1.6) where $n = 1$ and (1.7) where $n = 2$), and where the feedback is *localized* in that each agent has a bounded number of neighbors. We prove that no consensus algorithm based on only relative feedback can then achieve consensus in networks of any size. That is, while a given algorithm may allow a small network to converge to consensus, the same algorithm will lead to instability if agents are added to the network so that it grows beyond a certain finite size. In our terminology, this renders the algorithm *inadmissible*.

This result is shown to hold in classes of network graphs whose algebraic connectivity, that is, the smallest non-zero Laplacian eigenvalue, is decreasing towards zero in network size. This applies, for example, to all planar graphs. The proof, which relies on Routh-Hurwitz criteria for complex-coefficient polynomials, holds true for weighted, directed graphs with normal graph Laplacians. We survey classes

of graphs where this issue arises.

We also consider a case of inadmissibility of second-order consensus in toric lattice graphs (the same type of networks as considered in Chapter 6). We show that with relative feedback, any amount of asymmetry in the feedback around each node will render the controller inadmissible. That is, if the lattice graph is directed. We discuss the implications of both of these results for the implementation of distributed integral action in first- and second-order integrator networks that is the topic of Chapter 6.

Chapter 5 is partly based on:

> E. Tegling, B. Bamieh and H. Sandberg, "Localized high-order consensus destabilizes large-scale networks," Submitted to *2019 American Control Conference (ACC)*.

Chapter 6: Fundamental limitations of dynamic feedback control in regular networks

Chapter 6 addresses the question of whether distributed dynamic feedback can perform better than static feedback in terms of network coherence. A positive answer to this question was given in Chapter 4 for second-order integrator networks with partial access to absolute feedback. In Chapter 6, we consider both first- and second-order systems (we here term them consensus and vehicular formation control problems) and focus on the case with only relative feedback.

We model the consensus and vehicular formation control problems over toric lattice networks. For the resulting *spatially invariant* systems, we study the large scale asymptotics (in network size) of the global performance metric (1.9). With static, relative feedback, such metrics are known to scale unfavorably in lattices of low spatial dimensions, preventing, for example, a 1-dimensional string of vehicles to move like a rigid object. Here, we show that the same limitations in general apply also to dynamic feedback control that is locally of first order. This means that the addition of one local state to the controller gives a similar asymptotic performance to the memoryless case.

To derive our results, we present a general technical framework for the evaluation of stability and performance of spatially invariant systems in the limit of large networks. This framework, which is one of the key contributions of this work, allows for an $\mathcal{H}_2$ performance analysis of large classes of dynamic feedback protocols that would otherwise be intractable.

Chapter 6 is based on:

> E. Tegling, P. Mitra, B. Bamieh and H. Sandberg, "On fundamental limitations of dynamic feedback control in regular large-scale networks," *IEEE Transactions on Automatic Control*, provisionally accepted, August 2018.

E. Tegling, P. Mitra, H. Sandberg and B. Bamieh, "Coherence and stability in large-scale networks with distributed dynamic feedback," *22nd International Symposium on Mathematical Theory of Networks and Systems*, Minneapolis, MN, July 2016.

Chapter 7: Noise-induced limitations to distributed integral control

In Chapter 7 we address the question of how the performance of the distributed integral controller introduced in Chapter 4 (called *DAPI* control) is affected by measurement noise on the absolute feedback measurements. As in Chapter 6, we consider the second-order vehicular formation control problem modeled over a toric lattice network and analyze asymptotic scalings of the metric (1.9). While the results in Chapter 4 and 6 suggested that DAPI control fundamentally improves these performance scalings compared to static feedback control, this chapter shows that an explicit inclusion of measurement noise leads to the opposite conclusion. That is, under measurement noise, the asymptotic scaling of both global and local performance is worse than with static feedback control.

We show that the noise's impact on performance decreases with an increased inter-nodal alignment of the local integral states. However, even though the controller can be tuned for acceptable performance for any given network size, performance will degrade as the network grows, limiting the scalability of any such controller tuning. In particular, the requirement for inter-nodal alignment increases with network size. We show that this in practice implies that large and sparse networks will require any integral control to be centralized or have all-to-all communication. In this case, the best-achievable performance scaling, which is shown to be that of static feedback control, is retrieved.

Chapter 7 is based on:

E. Tegling and H. Sandberg, "Noise-induced limitations to the scalability of distributed integral control," *Systems & Control Letters*, under review, August 2018.
Also presented at:
23rd International Symposium on Mathematical Theory of Networks and Systems, Hong Kong, China, July 2018.

Chapter 8: The Price of Synchrony: resistive losses in synchronizing power networks

Chapter 8 introduces the "Price of Synchrony" as a notion of transient performance in electric power networks. That is, the resistive power losses that are incurred in keeping a network of synchronous generators at a synchronous state. We show how the total network's transient resistive losses can be quantified through an $\mathcal{H}_2$ norm of a linear system of coupled swing equations with an appropriately defined

performance output. We evaluate this norm for general network topologies and provide a number of relevant bounds on performance, as well as numerical examples.

The chapter's main result shows that the Price of Synchrony scales unboundedly with the number of generator nodes in the network. However, and in sharp contrast to the results in Chapters 4–7, it is only weakly dependent on the network topology. These results point to a fundamental limitation to performance in networks where the frequency synchronization is reliant on power flow mechanisms. They also provide a basis for the analysis in Chapters 9–10.

Chapter 8 is based on:

> E. Tegling, B. Bamieh and D. F. Gayme, "The Price of Synchrony: Evaluating the resistive losses in synchronizing power networks," in *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 254–266, Sept. 2015.

Chapter 9: Performance impact of variable voltage dynamics

The results in Chapter 8 were derived under the standard linear power flow assumptions, under which the voltage profile is assumed uniform across the network. In Chapter 9, we study the impact of variable voltages on performance. The scenario differs from Chapter 8 in that we here consider a microgrid with droop-controlled inverters. We obtain a networked system with *cross-coupled* voltage and phase-frequency dynamics (the latter being qualitatively identical to the swing dynamics considered in Chapter 8), and we adapt the Price of Synchrony performance output to quantify transient losses for this system.

The results demonstrate that while frequency synchronization results in losses that scale with a network's size but only weakly depend on its connectivity, the losses associated with voltage droop control will be larger in a highly interconnected network than in a loosely interconnected one. We discuss interpretations of this somewhat counterintuitive result and provide bounds for different network topologies. Furthermore, we show that while the cross-couplings between the voltage and phase-frequency dynamics lead to conditions on the droop controllers that must be observed to guarantee stability, they only have a small impact on performance. A decoupled analysis is therefore well motivated in systems with reasonable stability margins.

Chapter 9 is partly based on:

> E. Tegling, D. F. Gayme and H. Sandberg, "Performance metrics for droop-controlled microgrids with variable voltage dynamics," *IEEE 54th Annual Conference on Decision and Control (CDC)*, Osaka, 2015, pp. 7502–7509.

Chapter 10: Advantages and limitations of distributed secondary frequency control

In this chapter, we address the problem of *secondary* frequency control in power networks, the objective of which is to eliminate the stationary control error that arises through standard droop control (which is essentially proportional control). The DAPI controller that we introduced in Chapter 4 is one of the controllers that has been proposed as a distributed solution to this problem. Here, we evaluate its performance in terms of the Price of Synchrony metric compared to both standard droop control as well as to a corresponding centralized approach called CAPI.

In line with Chapter 4, we find that the DAPI controller can greatly reduce transient losses compared both to droop and CAPI control, provided it has access to noiseless frequency measurements. However, additive measurement noise has a significant negative impact on the DAPI controller's performance. In fact, the total transient losses can be shown to scale faster than linearly in network size. The negative impact of measurement noise, however, decreases with an increased internodal alignment of integral states. We prove that for any level of noise and any network size, there is in fact a controller tuning that allows DAPI to perform better than droop and CAPI control. Therefore, if the scaling of performance in network size is not a priority, but rather the transient performance of a given network, our results show that DAPI control is a good choice, but requires careful tuning.

Chapter 10 is based on:

> E. Tegling, M. Andreasson, J. W. Simpson-Porco and H. Sandberg, "Improving performance of droop-controlled microgrids through distributed PI-control," *American Control Conference (ACC)*, Boston, MA, 2016, pp. 2321–2327.
>
> H. Flamme, E. Tegling and H. Sandberg, "Performance limitations of distributed integral control in power networks under noisy measurements," *American Control Conference (ACC)*, Milwaukee, WI, 2018, pp. 5380–5386.
>
> M. Andreasson, E. Tegling, H. Sandberg and K. H. Johansson, "Coherence in synchronizing power networks with distributed integral control," *IEEE 56th Annual Conference on Decision and Control (CDC)*, Melbourne, VIC, 2017, pp. 6327–6333.

Chapter 11: Conclusions

In this chapter, we conclude the findings of the thesis. In particular, we discuss possible implications of the fundamental performance limitations for applications like those presented in the introduction. We discuss certain limitations of the methods and the results, and some open problems.

Other publications

The contributions of the following article are included in this thesis in the form of an illustrative example:

> M. Andreasson, E. Tegling, H. Sandberg and K. H. Johansson, "Performance and scalability of voltage controllers in multi-terminal HVDC networks," *American Control Conference (ACC)*, Seattle, WA, 2017, pp. 3029–3034.

The contributions of the following articles have inspired parts of this thesis, but they are not explicitly covered.

> E. Sjödin[2], U. Topcu and D. F. Gayme, "Risk-mitigated optimal power flow for wind powered grids," *American Control Conference (ACC)*, Montreal, QC, 2012, pp. 4431–4437.
>
> E. Sjödin and D. F. Gayme, "Transient losses in synchronizing renewable energy integrated power networks," *American Control Conference (ACC)*, Portland, OR, 2014, pp. 5217–5223.
>
> N. Govindarajan, H. Arbabi, L. van Blargian, T. Matchen, E. Tegling and I. Mezić, "An operator-theoretic viewpoint to non-smooth dynamical systems: Koopman analysis of a hybrid pendulum," *IEEE 55th Annual Conference on Decision and Control (CDC)*, Las Vegas, NV, 2016, pp. 6477–6484.

A number of the results presented in this thesis have previously appeared in the following theses:

> E. Tegling, *On performance limitations of large-scale networks with distributed feedback control*, Licentiate thesis, KTH Royal Institute of Technology, 2016.
>
> E. Sjödin, *The Price of Synchrony: Evaluating transient power losses in renewable energy integrated power networks*, M.Sc. thesis, KTH Royal Institute of Technology, 2013.

Remark 1.1 (The author's contributions) Wherever listed as the first author of a publication, the author of this thesis had the most significant role in developing the results, and has completed all or the majority of the writing. Remaining authors have contributed to problem formulations and taken advisory or supervisory roles.

In the collaboration with H. Flamme, this thesis' author defined the problem and had an advisory role, with a leading role in the writing. In the collaborations with M. Andreasson, the author contributed to the problem formulation and solution methods, and completed a majority of the writing. Finally, in the collaboration with N. Govindarajan *et al.*, the publication's first author contributed most to the analysis and the writing. The order of the second to fifth authors is alphabetical and reflects an equal contribution.

[2]The name Sjödin was changed to Tegling in 2014.

Chapter 2

Background

To describe the mechanisms by which local control and interaction rules lead to globally desirable behaviors is at the core of the field of networked control. This is a rich and challenging problem, and it is therefore not surprising that various aspects of it have become some of the most active topics of control theory over the last decades. These topics range from basic theoretical research, where objectives include the description of fundamental limitations to performance and robustness, to highly applied problems, such as the coordination of robots for specific missions, or the control of power electronic inverters in microgrids. This thesis is motivated by work on both ends of this spectrum. In this chapter we provide a brief overview of the research that motivates and relates to the thesis.

2.1 Characterizing limitations of networked control systems

In the introduction of this thesis, we presented a number of control challenges in networked and multi-agent systems. Overall, a common theme in addressing them is the design of control and interaction rules for a given system, and then showing that these rules lead to the desired performance, stability or robustness properties. An issue with this approach is that when performance guarantees for the proposed interaction rules cannot be found, it is not evident whether this is due to a fundamental limitation, or simply lack of ingenuity on the part of the control designer. A systematic approach to understanding fundamental limitations would instead be to solve optimal control problems and to check whether the best achievable performance is acceptable or fundamentally limited. In fully centralized control, textbook theory applies and several performance criteria based on the plant's unstable poles and zeros are well known [166, 200]. In networked control problems, however, there are *architectural constraints* on the controller, typically described by a graph structure. Under such constraints, the optimal control problem is, with few exceptions, non-convex and difficult to solve.

It is known that for large-scale networks, the centralized solutions to linear-

quadratic optimal control problems have an inherent degree of locality [15, 111]. However, the optimal control problem with a *prescribed* degree of locality tends to be non-convex [184]. An important exception is the class of quadratically-invariant problems [88, 146], which includes the subclass of funnel-causal problems in which the propagation speed in the controller is as fast as in the plant [16]. A novel approach to imposing controller sparsity while maintaining convexity was introduced in [186] and applied to the type of locality-constrained spatially invariant networks considered in parts of this thesis in [80]. Otherwise, the problem also gains tractability for positive systems, which can be exploited for a distributed controller synthesis [139].

For large-scale networks, a fruitful approach to understanding fundamental limitations while overcoming the difficulty of non-convexity is to derive bounds on the best-achievable performance in terms of various relevant metrics. Examples of such metrics include notions of controllability and $\mathcal{H}_2$ or $\mathcal{H}_\infty$ performance. Though the corresponding optimization problems would be non-convex, informative performance bounds that depend on the network size, topological properties and nodal dynamics can often be derived. This will also be the approach taken in this thesis. We will next review some of the notions of network performance and their use in recent literature to describe fundamental limitations. These performance notions are then summarized in Table 2.1. First, we make a remark on terminology.

Remark 2.1 (Demarcation of the performance problem) In this thesis, we adhere to the standard terminology with respect to the terms "performance", "stability" and "robustness" defined in, for example, [200, p. 211]. Consider therefore the standard feedback control system depicted in Figure 2.1. This closed-loop system is said to have nominal stability (NS) if the controller K internally stabilizes the nominal plant P, and nominal performance (NP) if performance objectives (usually in terms of signal and closed-loop system norms) are satisfied for P. It is said to have robust stability (RS) if K also stabilizes every plant in a larger model set $\Pi \ni P$ and robust performance (RP) if the performance objectives are also satisfied for every such plant.

In this thesis, we treat nominal stability and performance problems, and will therefore use the term "robustness" and its antonym "fragility" sparingly. We remark, however, that related literature sometimes refers to a controller's failure to attenuate noise as lack of "robustness (towards noise)." We term this a problem of (nominal) performance throughout. Similarly, we do not consider a small stability margin a performance problem in and of itself. It is, however, a sign of poor robustness and may lead to unfavorable performance. □

2.1.1 Notions of network controllability

The standard notion of controllability due to Kalman [83] is a binary one; it is either possible to drive a system to any given target state by means of some control input, or it is not. In a network context, it may for example be relevant to consider

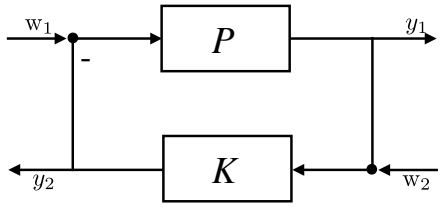


Figure 2.1: Standard feedback control system with (disturbance) input signals w_1, w_2, output signal y_1 and control signal y_2.

how many actuator nodes that are required to gain controllability over a network, see for example [94, 119], or [46] for an accessible overview. However, driving a system in general and a network in particular to a given state that is theoretically reachable may require a large control effort. A quantitative approach to studying network controllability is therefore called for. Such an approach is presented by Pasqualetti *et al.* in [126], where the worst-case energy required to steer a network to a given state is considered.

Recall that the minimum signal energy required to steer a system on the form $\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t)$ to a state x_f from $x(0) = 0$ in the time t_f is (provided the system is controllable)

$$E(u, t_f) = \int_0^{t_f} \|u(\tau)\|_2^2 \mathrm{d}\tau = x_f^T W_{C,t_f}^{-1} x_f,$$

where

$$W_{C,t_f} = \int_0^{t_f} e^{\mathcal{A}\tau} \mathcal{B}\mathcal{B}^T e^{\mathcal{A}^T \tau} \mathrm{d}\tau$$

is called the *controllability Gramian* at time t_f [166, Chapter 4]. If we let $\|x_f\|_2 = 1$ then it holds that

$$E(u, t_f) = x_f^T W_{C,t_f}^{-1} x_f \leq \lambda_{\min}^{-1}(W_{C,t_f}),$$

that is, the control energy is upper bounded by the inverse of the smallest eigenvalue of the controllability Gramian, with equality if the target state x_f is parallel to the corresponding eigenvector. Based on these facts, [126] studies worst-case controllability in terms of $\lambda_{\min}(W_{C,t_f})$ and investigates in particular the relation between the number of control nodes and the control energy.[1] They prove, for example, that to guarantee an upper bound on the control energy, the number of control nodes must be a linear function of the total number of network nodes. Otherwise, the energy will grow exponentially with network size.

[1] In fact, [126] considers a discrete-time system and the discrete time equivalent of W_{C,t_f}.

The same notion of controllability is also studied in a recent paper [125], where there is shown to be a trade-off between the controllability Gramian eigenvalues and the network robustness. That is, there is an inverse relation between $\lambda_{\min}(W_{C,\infty})$ and the stability radius

$$r(\mathcal{A}) = \min\{||\Delta|| \ : \ \mathcal{A} + \Delta \in \mathbb{C}^{N\times N} \text{ is not Hurwitz stable}\},$$

proving that more robust networks are also intrinsically more difficult to control. If the network size grows while the number of control inputs remains constant, controllability or robustness, or both, will decrease.

The controllability Gramian can also be used to quantify the degree of controllability in other ways. The trace of its inverse

$$\text{tr}(W_{C,t_f}^{-1})$$

is proportional to the average minimal control energy required to reach any target state x_f with unit norm. Tzoumas *et al.* [179] seek to minimize this measure through optimal placement of actuators in the network. However, as this problem is NP-hard, they consider a perturbed version of the trace that permits an efficient algorithm.

Another approach had earlier been taken by Summers and Lygeros in [171]. Motivated by the energy interpretation of $\text{tr}(W_{C,t_f}^{-1})$, they choose to maximize

$$P_C := \text{tr}(W_{C,\infty})$$

in a sensor and actuator placement problem, as this effectively minimizes the average energy at the infinite horizon $t_f = \infty$. Considering the trace of the Gramian itself rather than its inverse allows the optimization problem to be solved efficiently. It is noted in [126], however, that the selection of nodes that maximizes this metric does not necessarily minimize the worst-case control energy. We will see similar conflicts between seemingly related performance objectives later in this thesis. At this point, we also remark that the weighted trace $\text{tr}(\mathcal{C}P_C\mathcal{C}^T)$ corresponds to the $\mathcal{H}_2$ norm of the system with the output $y(t) = \mathcal{C}x(t)$. There is therefore a link between this notion of controllability and the $\mathcal{H}_2$ norm, which is the main performance measure used in this thesis and which will be discussed next.

2.1.2 Notions of sensitivity and coherence: $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance

A natural way to characterize a control system's performance is through the size of signals of interest, measured through norms. Indeed, the controllability measure discussed above bounded the L_2 norm of the control signal. By letting the system's output be the key signal of interest, the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ system norms will capture how this signal is impacted by several classes of inputs. A derivation of hard bounds on these norms therefore reveals some of the most important fundamental limitations to the control system in question.

A prototypical problem for characterizing bounds on the best-achievable performance in networks was formulated by Bamieh *et al.* in [14] (with a preliminary version appearing already in [13]). With a setup like the one in Section 1.2, they inquire as to the level of network *coherence* achievable through distributed static state feedback that is subject to locality constraints. The notion of coherence, which is also central to this thesis, is captured through the variance of deviations from a state of consensus:

$$V_i^{\mathrm{dav}}(t) = \mathbb{E}\left\{ \left(x_i(t) - \frac{1}{N}\sum_{j=1}^{N} x_j(t) \right)^2 \right\}.$$

In other words, if this variance is small, the feedback interactions are successful in upholding a state of consensus, and the network stays coherent. The same metric (then termed mean-square deviation) had earlier been used as a performance objective for the design of optimal interaction weights in the first-order consensus problem by Xiao *et al.* [190]. Having modeled the system as being subject to a white noise disturbance, the variance above is (at steady state) equivalent to an $\mathcal{H}_2$ norm.

Bamieh *et al.* [14] consider large-scale networks built over lattices. They prove a network dimensionality dependence in asymptotic (in network size) performance scalings, which show that lower-dimensional lattices have worse levels of coherence. This implies for example that vehicular platoon formations, being one-dimensional lattices, exhibit limitations in terms of the feasibility of constructing a formation that moves like a rigid object.

The particular aspect of the dependence on dimensionality of performance bounds was also observed in [19] in the context of distributed state estimation over lattice graphs. Their proofs rely on notions of effective resistance of graphs; a concept that we will later return to. Seeking a generalization of the dimensionality dependence observed in [14, 19], Patterson and Bamieh [128] studied the coherence problem for fractal graphs. Though fractal graphs also have a well-defined dimensionality, their coherence properties could be shown *not* to coincide with those of lattice graphs. For d-regular random graphs, however, Jadbabaie and Olshevsky [79] showed that good coherence properties can be expected provided d is sufficiently large, which suggests a generalization of the lattice dimension to other regular graphs. A similar result was shown in [28]. The proper translation of the notion of dimensionality to characterize coherence properties in general graphs, however, has remained an open question.

For consensus networks over general graphs, the question of coherence, and sensitivity to noise in general, has been addressed in a number of works. Young *et al.* [192] compared directed and undirected graphs and were among the first to discuss the relation between the coherence properties, captured through the $\mathcal{H}_2$ norm, and the *convergence speed* of the consensus algorithm. Interestingly, the convergence speed, which is proportional to the graph's algebraic connectivity, need

not scale well with network size in graphs with good coherence scaling, for example a 3-dimensional lattice. Conditions for the two performance criteria coinciding are given in [134] in the case of leader-follower consensus. The convergence speed, in turn, is in the case of standard first-order consensus also proportional to the system's $\mathcal{H}_\infty$ norm with respect to the error $y(t) = x(t) - \mathbf{1}\mathbf{1}^T x(t)$. This norm can be evaluated as [155]:

$$\|G\|_\infty = \frac{1}{\lambda_2},$$

where λ_2 is the algebraic connectivity of the graph (see Section 3.2). The $\mathcal{H}_\infty$ norm, which can be interpreted as the system's maximum $\mathcal{L}_2$ gain and thus quantifies the worst-case deviation from consensus, will therefore not necessarily be small for systems with small $\mathcal{H}_2$ norms.

In a series of works including [155,156], Siami and Motee characterize the coherence behavior of a number of classes of network graphs such as trees and bipartite graphs. They also consider noise sensitivity in terms of alternative performance outputs and norms ($\mathcal{H}_\infty$ in particular) and discuss overall trade-offs between network sparsity and performance. Also Grunberg and Gayme [62] address the $\mathcal{H}_2$ performance problem in general graphs, by exploiting a connection to their effective resistances. As such, there is a strong connection to the earlier work in [57], which studied the minimization of effective resistances of graphs from a general standpoint. This connection allows [62] to provide an alternative proof of the performance bounds in [14]. They also discuss the impact of localized disturbances and propose a notion of inter-nodal (as opposed to global) performance and its scalings. Finally, Pirani *et al.* [133, 134] discuss scalings of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance of consensus networks. In particular, they derive asymptotic performance bounds for random graphs, which are expressed in terms of their edge probability.

Though the best-achievable performance of networks in terms of notions of coherence is known to be fundamentally limited, optimization of the network and feedback design with respect to these performance criteria remains a relevant problem. The literature is extensive and often focused on algorithmic aspects, but we note a few examples. The problem of optimal leader selection with respect to network coherence has been studied in [53,93,128,130]. Optimal edge weight selection was addressed in [91] and optimal edge additions in [171] (note the correction [170]), both considering coherence as the performance objectve. Algorithms for optimal topology design with a joint sparsity and $\mathcal{H}_2$ performance objective was studied in, for example, [69,92].

The studies listed so far have focused on the dependence of the best-achievable performance on the local node dynamics and the network topology. However, this work has been limited to distributed *static* feedback control which does not alter the structure of local dynamics. Hence, the possible impact of controller dynamics on performance bounds, which is one of the key questions addressed in this thesis, has remained an open question.

An exception is the recent work by Pates *et al.* [127] that proves underlying

fundamental limitations to control based on noisy relative measurements between neighboring nodes of a string formation. Regardless of the transfer function of the local controllers, there will be a non-zero lower bound on the $\mathcal{H}_\infty$ norm of the closed-loop system. This means that certain disturbances will inevitably cause unfavorable macroscopic behaviors in, for example, vehicular platoons. The setup in [127] is not immediately comparable to the one used in this thesis, and the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance objectives do not always coincide. However, this work generalizes some of the disturbance amplification results from the platooning literature, which we will discuss next.

2.1.3 Performance of vehicles platoons: disturbance amplification and string stability

Apart from their applicational relevance, the fixed topological structure of vehicle platoons has led to a particular research interest in their performance limitations under different feedback protocols. The most well-known performance issue in platoons is *string instability*, see for example [17, 106, 149, 169, 172]. String instability implies that disturbances are amplified along the string of vehicles, though the overall system can be stable in the classical sense (it is therefore a notion of *performance* rather than of stability, see also [21]). A key objective in the string stability literature is to find a bound on relevant error signals, typically an input-to-output gain, that holds independently of the platoon size, and renders the control scalable.

For example, Seiler *et al.* [149] prove fundamental limitations in terms of error scalings in both predecessor-follower and bidirectional platoons (see Figure 2.2). They show that the $\mathcal{H}_\infty$ norm from a disturbance d at the first vehicle to the error e at a vehicle further down the string will grow unboundedly with the size of the platoon. That is, there is a platoon size $\bar{N}$ such that the transfer function

$$\|G_{d,e}\|_\infty \geq M$$

for all $N \geq \bar{N}$, given any number M. They show this to be due to a constraint on the complementary sensitivity function. The limitation can, however, be relaxed if the control uses the vehicle's absolute position with respect to the leader, which agrees with the results on $\mathcal{H}_2$ performance limitations found in [14]. We will elaborate on the importance of absolute measurements in this thesis.

An extension to [149] in the case of bidirectional control was presented by Barooah and Hespanha in [17]. For a string of identical vehicles they show that errors will grow without bound as vehicles are added, regardless of the (linear) controller. Their proof relies partly on the observation that if the local dynamics has three or more integrators, the platoon is eventually rendered unstable as it grows. Yadlapalli *et al.* [191] generalize these results to platoons where feedback interactions are extended to the q nearest neighbors in both directions. They demonstrate that scalability in terms of a bounded spacing error without absolute measurements can only be achieved if at least one vehicle allows q to increase linearly in the platoon

(a) Predecessor-follower protocol

(b) Symmetric bidirectional protocol

(c) Asymmetric bidirectional protocol

Figure 2.2: A common theme in the platooning literature is the comparison of performance bounds, typically in terms of string stability, with different types of feedback protocols. In the predecessor-follower protocol (a), only measurements with respect to the vehicle in front are used. In the bidirectional ones, measurements with respect to the vehicle behind are also used. If the front and back measurements are given equal weight in the feedback, the protocol is termed symmetric (b), else it is asymmetric (c).

size. In Chapter 5 of this thesis, we consider high-order integrator networks and further generalize the stability results observed in [17, 191].

Stability of vehicle platoons, more precisely the scaling of the stability margin defined as the real part of the least stable eigenvalue, was studied in [20, 67]. These studies conclude that allowing for an asymmetry in the feedback protocol with respect to the vehicle ahead and the vehicle behind (termed *mistuning*), improves this scaling as the platoon size grows. Overall, comparing symmetric bidirectional platooning protocols to asymmetric ones, or to the predecessor-follower protocol, is a common theme in the platooning literature. For example, Hao and Barooah [68] compared their respective $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance, proving that the symmetric protocol achieves the best scalings. They also show that a *nonlinear* feedback protocol can alleviate some of the limitations of the predecessor-follower algorithm. In a later study by Herman *et al.* [71], asymmetric protocols were shown to have unfavorable $\mathcal{H}_\infty$ performance scalings, regardless of the controller used. However, they can be alleviated if the local controllers have access to the lead vehicle's velocity and thus eliminate one of the two integrators in the local dynamics [72]. In this thesis, we will also discuss certain aspects of symmetric versus asymmetric feedback protocols, but mainly consider symmetric ones.

In summary, issues concerning the scalability of distributed controllers in terms of stability and performance have received a fair amount of attention in the platooning literature, and a number of fundamental limitations have been described. These issues are less well studied in the consensus and multi-agent systems literature, where the set of agents tends to be assumed fixed, but network topologies can be more complicated, see for example [117]. There have therefore been some

recent efforts to generalize particularly the notion of string stability to general networks [21, 169]. In [21], Besselink and Knorn propose the concept of scalable input-to-state stability, which can be applied to any networked system, also with nonlinear dynamics. It implies that state perturbations at all nodes remain upper bounded for any bounded disturbance input anywhere in the network, regardless of the network size. They present necessary conditions for this to hold true in spatially invariant networks, but more general results are yet to be developed.

2.1.4 A note on notions of centrality

The term *centrality* often arises in the context of performance of networked systems. Centrality indices are used to distinguish the relative importance of individual nodes for the network's global performance, and has been of particular importance in the study of social networks. In traditional graph theory, the simplest and most well-known characterization of centrality, apart from nodal degree, is that of *closeness*. In this case, a node is central if its greatest distance to any other node is small [35, Chapter 1]. However, depending on the performance objectives, other centrality notions may be more relevant.

For example, in the context of network controllability, [171] defines an *energy controllability centrality* as $\mathrm{tr}(W_{C,i})$, where $W_{C,i}$ is the controllability Gramian obtained with only the i^{th} input channel. *Information centrality*, which is a purely graph theoretical measure proposed in [168], is shown in [53] to be of importance for optimal leader selection with respect to coherence. In [153], a node's relative contribution to the network coherence is proposed as a measure of centrality. It is evaluated as $\eta_i = \frac{\partial \mathbf{V}}{\partial \sigma_i^2}$, where σ_i^2 is the noise intensity at the i^{th} node and $\mathbf{V}$ is the total steady-state variance:

$$\mathbf{V} = \lim_{t\to\infty} \sum_{i=1}^{N} \mathbb{E}\left\{ \left(x_i(t) - \frac{1}{N} \sum_{j=1}^{N} x_j(t) \right)^2 \right\}.$$

Finally, the resistance distance, a concept used to derive asymptotic error scalings in [19], characterizes node centrality in [32].

Centrality notions will, however, not be important for the developments in this thesis. With few exceptions, the networks we consider are homogeneous in terms of nodal dynamics, and we do not treat placement or leader selection problems (apart from the illustrative Example 8.3). Nevertheless, it may become relevant for future extensions of topics considered in this thesis.

Table 2.1: Summary of performance characterizations discussed in Section 2.1, and examples of their interpretation and evaluation in a network context.

Performance characterization	**Network interpretation**	**Evaluation**	**Example reference**
Convergence speed	Time to reach consensus	For 1^{st}-order consensus conv. speed $\sim \frac{1}{\lambda_2}$	[192]
Degree of controllability	Energy required to steer network to a given state	E.g., $\text{tr}(W_C^{-1})$ or $\lambda_{\min}^{-1}(W_C)$	[126]
$\mathcal{H}_2$ performance	Coherence of network driven by white noise	E.g., $\text{tr}(\mathcal{C}P_C\mathcal{C}^T)$. See Section 3.1.	[14]
$\mathcal{H}_\infty$ performance	Worst-case disturbance amplification.	For 1^{st}-order consensus $\|G\|_\infty = \frac{1}{\lambda_2}$	[127], [134], [155]
String stability	Amplification of a disturbance d entering at lead vehicle of a platoon	$\|G_{d,e}\|_\infty$.	[17], [149], [191]

2.2 Scalable frequency control for a new power system paradigm

Maintaining the frequency close to its nominal value is one of the key control objectives in alternating current (AC) power systems. This is traditionally achieved through a hierarchy of control actions, starting with decentralized primary control (droop control), via the secondary control layer with the automatic generation control (AGC) to the tertiary control layer, where an optimal economic dispatch of generators takes place [76, 98]. These control layers operate at different time scales and, traditionally, with different degrees of centralization. Power systems are, however, entering a new paradigm where local, small-scale generation resources are increasingly replacing high-capacity centralized power plants [49,108]. This has motivated considerable research efforts in developing, first, a formal understanding for frequency synchronization, and second, more flexible and scalable distributed schemes for generation planning and frequency control.

In Chapters 8–10 of this thesis, we join this effort by addressing questions of performance and scalability of distributed frequency controllers. This section highlights a small subset of the related research. The surrounding literature is, however, vast and the interested reader is referred to the recent review articles [44,110], and the references therein.

2.2.1 A systems and control perspective on power network synchronization

Power networks, from having been viewed as well-understood (albeit complex) systems with little room for innovations, have received renewed research attention from a systems and control perspective over the last decade. This is in particular true for the problem of rotor-angular, or synchronous stability. Synchronous stability, being one of several characterizations of power system stability, is the ability of the system to recover synchrony after a disturbance [120]. In this context, synchrony refers to the alignment of the phase and frequency of all generators within a particular power network. In other words, it is when all of the frequencies are equal [112] and the phase differences are at an equilibrium state corresponding to balanced power flows throughout the network. Maintaining synchrony thus depends on a network's ability to sustain or restore this condition when perturbed from its nominal operating point.

Synchronous stability properties of power systems are often studied using a so-called *Kron reduced* model where constant-impedance loads are absorbed into the "transmission lines" of the reduced network along with any phase-shifting transformers, see [30, 39, 112, 123, 183]. The resulting system is a set of coupled swing equations:

$$m_i\ddot{\theta}_i + d_i\dot{\theta}_i = P_{m,i} - P_{e,i},$$

where the power flow is a trigonometric function of phase angle differences:

$$P_{e,i} \;=\; \bar{g}_i V_i^2 + \sum_{j\in\mathcal{N}_i} g_{ij} V_i\, V_j \cos(\theta_i - \theta_j) + \sum_{j\in\mathcal{N}_i} b_{ij} V_i\, V_j \sin(\theta_i - \theta_j).$$

Here, V_i are voltage magnitudes[2] and $b_{ij},\ g_{ij},\ \bar{g}_i$ are constants (see Section 8.1 for details). This system can then be analyzed to determine conditions under which the synchronized state is stable [40, 112].

In particular, Dörfler and Bullo have analyzed this problem in a series of works [36–38, 40] by drawing connections between power networks and coupled *Kuramoto oscillators*:

$$d_i\dot{\theta} = \omega_i - \sum_{j=1}^{N} a_{ij} \sin(\theta_i - \theta_j),$$

in which ω_i is a natural frequency, d_i is a constant and the a_{ij} are coupling strengths. Stability conditions for this model had earlier been derived by, for example, Jadbabaie *et al.* in [78]. This modeling framework provides a first-order approximation of the power system, and properties of the well-understood Kuramoto oscillator problem can be exploited to provide analytical conditions for frequency synchronization. An analogous analysis was presented for droop-controlled inverters in microgrids in [158, 159]. Mallada and Tang [101] used similar first-order models to investigate the effects of power flow scheduling and increasing network connectivity (that is, addition of transmission lines) on the rate of synchronization.

The question of optimal control design for synchronizing oscillator networks was addressed in [48] and applied to so-called wide-area control to prevent inter-area oscillations in transmission networks in [42]. These authors use sparsity-promoting constraints to limit the number of exogenous control inputs and employ quadratic performance objectives, equivalent to system $\mathcal{H}_2$ norms, for their control design. We also analyze the performance of synchronizing power networks through an $\mathcal{H}_2$ norm metric in this thesis. However, the system dynamics and performance outputs are defined differently, and the performance metrics are therefore distinct.

2.2.2 Novel approaches to frequency control

The new requirements posed by large-scale integration of distributed generation sources has led to a questioning of the traditional hierarchical structure of frequency control. This is especially true for microgrids, which are envisioned as networks composed of distributed generation units, loads and energy storage elements, which can either connect to the main power grid or operate independetly from it [87, 103]. Traditional centralized planning and generation control conflicts with the microgrid

[2] In a power systems context, the notation V_i, $i = 1, \ldots, N$ will always be used to denote voltage magnitudes. The notation for steady-state variance in a network coherence context: $V_N = \mathbf{V}/N$, is unrelated to voltages. The meaning of the notation in each case will be clear from the context.

paradigm of autonomous operation and scalability, so more flexible, distributed approaches are called for.

On the one hand, this problem has been addressed through online optimization techniques, typically based on primal-dual gradient methods, that exploit the frequency dynamics of the power system [90, 100, 102, 195, 197]. One approach is through load-side frequency control, that is, the use of controllable loads rather than generators to compensate for power imbalances. This is attractive in a scenario with high penetration of renewable energy, which is often non-dispatchable and variable in nature. Zhao *et al.* [197] address the primary frequency control problem and prove that the swing dynamics of the system uniquely solves an optimal load control problem. This means that the problem can be solved in a completely decentralized fashion. In [100, 102], this problem is extended to include secondary frequency control, that is, a restoration of the synchronous frequency to the desired setpoint. In this case, signals must be exchanged between neighboring nodes, but the control remains distributed.

In an approach similar to [197], Li *et al.* [90] incorporate the optimal dispatch problem (that is, the tertiary frequency control layer) to the AGC (the secondary layer). The resulting optimization problem is again implemented online as frequency control dynamics and solved in a distributed fashion. A slightly more detailed model is used by Zhang and Papachristodoulou [195], where the physics of the power system is shown to solve an optimal power flow problem at steady state.

Another line of research has studied various integral control strategies to address the secondary frequency control problem [6, 7, 43, 151, 159, 178, 196]. Among the proposed controllers are fully decentralized proportional-integral (PI) controllers, which have been shown to suffer from poor robustness properties and to lead to suboptimal power injections [6, 43]. When complemented by a distributed averaging of the integral state, however, these controllers can eliminate frequency errors while maintaining optimality properties, as shown independently by Andreasson *et al.* [7] and Simpson-Porco *et al.* [159]. In particular, these *distributed averaging integral controllers* (we will later refer to them as DAPI) maintain *power sharing*, by which a fair allocation of power injections in accordance with the generators' ratings is meant. They can also be shown to solve the optimal dispatch problem [43, 196], and thus eliminate the need for the tertiary frequency control layer. In this thesis, we will discuss some of these integral controllers from a transient performance perspective.

2.2.3 Characterizing transient performance

The most novel aspect of this thesis in a power systems context lies in the analysis of transient performance, not in stability or convergence analysis. Transient performance refers to how fast and successfully the power network rejects disturbances in the form of faults, contingencies or rapid generation or load variations. It can be described through a number of characteristics. The most important ones are described in Figure 2.3. They include the maximum rate of change of frequency

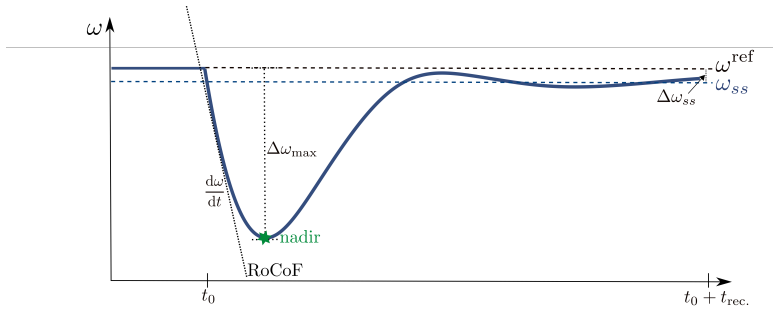


Figure 2.3: Schematic transient response of a power system after a disturbance or contingency. The rate of change of frequency (RoCoF) is defined as the maximum frequency gradient and the nadir is the maximal deviation from the frequency setpoint. Other performance indicators are the steady-state frequency error $\Delta\omega_{ss}$ and the time to recovery, here $t_{\text{rec.}}$. $\mathcal{H}_2$-norm metrics (depending on the output definition) is related to the area between the frequency curve and the ω_{ss}-line.

(RoCoF), the maximal frequency deviation (frequency nadir), the recovery time, and the frequency deviation at steady state (stationary error). An $\mathcal{H}_2$ norm metric, like the one we consider here, would in Figure 2.3 be related to the area of the response. It thus captures many of these characteristics at once.

We study performance in terms of the resistive power losses incurred in regulating frequency under persistent stochastic disturbances. These losses arise due to non-equilibrium power flows that occur whenever the system is out of synchrony, and can be quantified through the $\mathcal{H}_2$ norm of a system of coupled swing equations with an appropriately defined output. We term this particular notion of performance *the Price of Synchrony*. This type of analysis was first proposed by Bamieh and Gayme [12], and was shortly thereafter extended into the more detailed analysis in [175] that also appears in Chapter 8 of this thesis.

Since these studies, there has been an increased interest in using $\mathcal{H}_2$ norm-based performance criteria for power system analysis and control design. Motivated by an expected loss of inertia in renewable energy-integrated power networks, Poolla *et al.* use various $\mathcal{H}_2$ norm based performance criteria in a problem of optimal allocation of virtual inertia, and Pirani et al. [135] analyze the effect of loss of inertia through system norms. The use of select control nodes for optimization of $\mathcal{H}_2$ performance metrics similar to the Price of Synchrony was studied in [61].

The previously mentioned works focus on standard droop controlled networks, but alternative frequency controllers have also been analyzed. For example, the $\mathcal{H}_2$ performance of primal-dual methods for secondary frequency control (see Section 2.2.2) was analyzed in [162] and its extension [161]. Their results show that these methods may result in controllers with unfavorable performance scalings, but propose regularization methods that lead to improved scalability. Secondary fre-

quency control is also treated by Wu *et al.* [188], who analyze optimal topology design for the distributed averaging integral controller that we also consider in this thesis. Their results demonstrate that the $\mathcal{H}_2$-optimal centralized controller barely outperforms a sparse one. Finally, a dynamic droop controller inspired by classic lead-lag control proposed in [99] was analyzed for $\mathcal{H}_2$ performance in [82], and found to outperform the regular droop controller.

The aforementioned studies, as well as the work to be presented in this thesis, are subject to a number of limitations. First, power system dynamics are inherently non-linear, so a use of system norms only allows for a small-signal analysis. A meaningful generalization of the results to nonlinear dynamics remains an open question. Second, power systems can be subject to a variety of disturbance scenarios that do not lend themselves to the current modeling framework. To gain further insights Coletta and Jaquod have recently analyzed the performance under line contingencies [33], and colored noise in [32]. Third, much of the analysis requires an assumption of uniformity in the generators, in terms of damping and inertia. This assumption is partly relaxed in [122], which unlike most of the related work takes a transfer function approach to the analysis. They also propose and analyze norm-based measures that closely resemble the RoCoF and frequency nadir. We will return to these references and elaborate on the limitations of our study later in the thesis.

2.3 Concluding summary

This thesis is motivated by two areas of research whose overarching objectives are, respectively, 1) to characterize fundamental limitations of feedback control under architectural constraints, and 2) to develop new control strategies for a modern, highly distributed power system paradigm. Our approach to these problems is to analyze limitations to the performance and the scalability of distributed static and dynamic feedback controllers. This is done in terms of $\mathcal{H}_2$-norm based performance metrics. Our review of related literature has shown that this is not the only meaningful way to characterize the performance of a networked system. Others include convergence times, the degree of controllability, maximum error signals, and (in a power systems context) the rate of change of frequency. Stability and robustness properties are also highly important. We note that most of the reviewed literature on the performance aspects of networked control is less than a decade old. It is therefore reasonable that a consensus on the most relevant aspects of networked control performance and their relations to each other have yet to be developed.

This chapter's literature review describes an increasing interest for $\mathcal{H}_2$-norm based performance measures in both general networked systems and in power systems. We argue that this use is well motivated, not only from a standpoint of mathematical tractability. On the one hand, the $\mathcal{H}_2$ norm is closely related to notions of network controllability via the controllability Gramian. On the other hand, it captures *expected* deviations from a control objective, rather than worst-

case scenarios. This means that a limitation on $\mathcal{H}_2$ performance is a limitation on the system's behavior under ordinary circumstances. While a control design should also take worst-case and possibly very unlikely scenarios into consideration, such scenarios do not necessarily give the best understanding for system behaviors such as network coherence. Finally, in the context of transient frequency responses in power networks, the $\mathcal{H}_2$ norm gives combined information of both amplitudes and durations of transients, both of which are important. Provided stability conditions can be met, the use of $\mathcal{H}_2$-norm criteria for control and topology design, which has gained increased popularity, is therefore reasonable.

Chapter 3

Technical Preliminaries

Before proceeding to this thesis' technical contributions, we review parts of the mathematical framework that will be needed. We first review the signal and system norms that will be used for performance evaluations throughout. In Section 3.2 we introduce relevant concepts from graph theory and define entities like the graph Laplacian matrix. Section 3.3 goes through definitions and notation for the analysis of spatially invariant systems and Section 3.4 contains other important definitions including the notion of asymptotic scalings.

3.1 Signal and system norms

Performance of control systems is usually described in terms of the "size" of certain signals of interest. In this thesis, we will consider signals that, in a sense, quantify the level of agreement, or order, in a network. Their size is measured through relevant signal and system norms. In this section, we review $\mathcal{L}_2$ and $\mathcal{H}_2$ norms and discuss how they can be computed and interpreted. This theory along with more details can be found in [200, Chapter 4].

3.1.1 The $\mathcal{L}_2$ norm

Consider a vector-valued signal $v(t) = [v_1(t), \ldots, v_n(t)]^T$ defined for $t \geq 0$. Its $\mathcal{L}_2$ norm is defined as the square root of the time integral of $v^*(t)v(t)$:

$$||v||_2 = \left(\int_0^\infty v^*(t)v(t)\mathrm{d}t \right)^{1/2}. \tag{3.1}$$

Here, and throughout the thesis, we use * to denote the complex conjugate transpose of a vector or matrix[1]. The $\mathcal{L}_2$ norm can also be evaluated in the frequency

[1] We will often use the notation for the complex conjugate transpose v^* even when the vector or matrix is real valued and the ordinary transpose v^T could have been used instead. This choice

domain. We then use the Laplace-transformed signal $\hat{v}(s)$, evaluated on the imaginary axis $s = \mathbf{j}\omega$:

$$||\hat{v}||_2 = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{v}^*(\mathbf{j}\omega) v(\mathbf{j}\omega) \mathrm{d}\omega \right)^{1/2}. \tag{3.2}$$

By Parseval's theorem, the time-domain and frequency-domain norms are equal, that is,

$$||v||_2 = ||\hat{v}||_2.$$

3.1.2 The $\mathcal{H}_2$ norm

Now, consider a general linear MIMO system in the Laplace domain: $\hat{y}(s) = G(s)\hat{\mathrm{w}}(s)$, with input w, output y and transfer matrix $G(s)$. If the transfer matrix $G(s)$ is strictly proper ($G(\infty) = 0$) and input-output stable, the $\mathcal{H}_2$ norm is defined in analogy with the $\mathcal{L}_2$ norm in (3.2) as

$$||G||_2 = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{tr}(G^*(\mathbf{j}\omega) G(\mathbf{j}\omega)) \mathrm{d}\omega \right)^{1/2}. \tag{3.3}$$

The $\mathcal{H}_2$ norm characterizes the system's input-output behavior by quantifying the size of the output y, as a sum over all frequencies, given certain inputs w. In control design, a typical control objective is to keep the $\mathcal{H}_2$ norm below a given limit, and the feedback is chosen accordingly.

The $\mathcal{H}_2$ norm of a system G is rarely evaluated from its definition in (3.3). It is instead convenient to use its state space realization.

$$\begin{aligned} \dot{\psi}(t) &= \mathcal{A}\psi(t) + \mathcal{B}\mathrm{w}(t) \\ y(t) &= \mathcal{C}\psi(t), \end{aligned} \tag{3.4}$$

where $G(s) = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B}$. Through Parseval's theorem, it is straightforward to show that $||G||_2$ is equal to the time-domain norm (3.1) of the impulse response $g(t) = \mathcal{C}e^{\mathcal{A}t}\mathcal{B}$ for $t \geq 0$ (setting $g(t) = 0$ for $t < 0$). That is,

$$\|G\|_2^2 = \mathrm{tr} \left(\int_0^{\infty} \mathcal{B}^* e^{\mathcal{A}^* t} \mathcal{C}^* \mathcal{C} e^{\mathcal{A}t} \mathcal{B} \mathrm{d}t \right). \tag{3.5}$$

Now, we define the *observability Gramian* as the matrix

$$P = \int_0^{\infty} e^{\mathcal{A}^* t} \mathcal{C}^* \mathcal{C} e^{\mathcal{A}t} \mathrm{d}t, \tag{3.6}$$

which allows (3.5) to be written as

$$\|G\|_2^2 = \mathrm{tr}\left(\mathcal{B}^* P \mathcal{B}\right). \tag{3.7}$$

of notation is made for the sake of consistency, as some quantities will sometimes be real valued and sometimes complex valued. This holds, for example, for $\hat{\mathcal{A}}_n$ in (3.9).

The observability Gramian is given as the solution to the *Lyapunov equation*:

$$\mathcal{A}^* P + P\mathcal{A} = -\mathcal{C}^*\mathcal{C}, \tag{3.8}$$

which is unique if $\mathcal{A}$ is Hurwitz stable (that is, if all its eigenvalues have strictly negative real parts). We can also calculate $||G||_2^2$ using the *controllability Gramian* P_C; $||G||_2^2 = \mathrm{tr}(\mathcal{C}P_C\mathcal{C}^*)$, with $\mathcal{A}P_C + P_C\mathcal{A}^* = -\mathcal{B}\mathcal{B}^*$.

Exploting the unitary invariance

The $\mathcal{H}_2$ norm is unitarily invariant. This means that it is unaffected by a unitary change of variables in the system. Throughout this thesis, we will exploit this property when evaluating system $\mathcal{H}_2$ norms through spectral decompositions.

For this purpose, consider a unitary matrix U and assume that it (block) diagonalizes the system matrix $\mathcal{A}$ in (3.4), so that $U^*\mathcal{A}U = \hat{\mathcal{A}}$, with $\hat{\mathcal{A}} = \mathrm{diag}\{\hat{\mathcal{A}}_n\}$, where $\hat{\mathcal{A}}_n$, $n = 1, \ldots, N$ are the eigenvalues of $\mathcal{A}$ (or small block matrices, whose eigenvalues are the eigenvalues of $\mathcal{A}$). We use the matrix U to transform the state vector $\psi(t)$ according to $\psi(t) =: U\hat{\psi}(t)$. Recalling that a unitary matrix has the property $U^*U = I$, the system dynamics in terms of the transformed states $\hat{\psi}$ become:

$$\begin{aligned}\dot{\hat{\psi}}(t) &= U^*\mathcal{A}U\hat{\psi}(t) + U^*\mathcal{B}\mathrm{w}(t)\\ y(t) &= \mathcal{C}U\hat{\psi}(t).\end{aligned}$$

Since the $\mathcal{H}_2$ norm is unitarily invariant, we can multiply the input and output by the unitary matrix U (assuming $w, y \in \mathbb{R}^N$, else zeros can be added). We therefore define $\hat{\mathrm{w}}(t) := U^*\mathrm{w}(t)$ and $\hat{y}(t) := U^*y(t)$, and obtain the system $\hat{G}$ from $\hat{\mathrm{w}}$ to $\hat{y}$ as

$$\begin{aligned}\dot{\hat{\psi}}(t) &= U^*\mathcal{A}U\hat{\psi}(t) + U^*\mathcal{B}U\hat{\mathrm{w}}(t) =: \hat{\mathcal{A}}\hat{\psi}(t) + \hat{\mathcal{B}}\hat{\mathrm{w}}(t)\\ \hat{y}(t) &= U^*\mathcal{C}U\hat{\psi}(t) =: \hat{\mathcal{C}}\hat{\psi}(t).\end{aligned}$$

Now, assume that $\mathcal{B}$ and $\mathcal{C}$ are such that they are, like $\mathcal{A}$, (block) diagonalized by U. That is, $\hat{\mathcal{B}} = \mathrm{diag}\{\hat{\mathcal{B}}_n\}$ and $\hat{\mathcal{C}} = \mathrm{diag}\{\hat{\mathcal{C}}_n\}$. The norm (3.5), which in terms of the transformed matrices becomes

$$||\hat{G}||_2^2 = \mathrm{tr}\left(\int_0^\infty \hat{\mathcal{B}}^* e^{\hat{\mathcal{A}}^* t}\hat{\mathcal{C}}^*\hat{\mathcal{C}}e^{\hat{\mathcal{A}}t}\hat{\mathcal{B}}\mathrm{d}t\right),$$

is then equivalent to

$$||\hat{G}||_2^2 = \mathrm{tr}\left(\sum_{n=1}^{N}\int_0^\infty \hat{\mathcal{B}}_n^* e^{\hat{\mathcal{A}}_n^* t}\hat{\mathcal{C}}_n^*\hat{\mathcal{C}}_n e^{\hat{\mathcal{A}}_n t}\hat{\mathcal{B}}_n\mathrm{d}t\right), \tag{3.9}$$

since all the matrices are (block) diagonal. Since the trace and the sum can switch order, we recognize this as the sum of the respective squared $\mathcal{H}_2$ norms of N decoupled subsystems $\hat{G}_n$:

$$\begin{aligned}\dot{\hat{\psi}}_n(t) &= \hat{\mathcal{A}}_n\hat{\psi}_n(t) + \hat{\mathcal{B}}_n\hat{\mathrm{w}}_n(t)\\ \hat{y}_n(t) &= \hat{\mathcal{C}}_n\hat{\psi}_n(t),\end{aligned}$$

which allows us to simply evaluate the $\mathcal{H}_2$ norm of (3.4) as

$$||G||_2^2 = ||\hat{G}||_2^2 = \sum_{n=1}^{N} ||\hat{G}_n||_2^2. \tag{3.10}$$

The individual subsystem norms can be calculated as $\|\hat{G}_n\|_2^2 = \mathrm{tr}\left(\hat{\mathcal{B}}_n^* P_n \hat{\mathcal{B}}_n\right)$ where the observability Gramians P_n are obtained through the Lyapunov equations

$$\hat{\mathcal{A}}_n^* P_n + P_n\hat{\mathcal{A}}_n = -\hat{\mathcal{C}}_n^*\hat{\mathcal{C}}_n, \tag{3.11}$$

for $n = 1, \ldots, N$.

3.1.3 Interpretations of the $\mathcal{H}_2$ norm

Throughout this thesis, $\mathcal{H}_2$ norms will be used to characterize system performance in various settings. This usage is supported by some of the $\mathcal{H}_2$ norm's standard interpretations, which we recap here. Denote by G the LTI system (3.4) and consider the following three scenarios:

i. Response to a white stochastic input. When the input w is a white second order process with unit covariance, that is, $\mathbb{E}\{\mathrm{w}(\tau)\mathrm{w}^*(t)\} = \delta(t-\tau)I$, the (squared) $\mathcal{H}_2$ norm of the system is the steady-state total variance of all of the output components, that is,

$$\|G\|_2^2 \;=\; \lim_{t\to\infty} \mathbb{E}\{y^*(t)y(t)\}. \tag{3.12}$$

ii. Response to a random initial condition. With zero input and an initial condition that is a random variable ψ_o with zero mean and correlation $\mathbb{E}\{\psi_o\psi_o^*\} = \mathcal{B}\mathcal{B}^*$, the (squared) $\mathcal{H}_2$ norm is the time integral of the resulting response y:

$$\|G\|_2^2 \;=\; \int_0^\infty \mathbb{E}\{y^*(t)y(t)\}\ \mathrm{d}t. \tag{3.13}$$

This interpretation is closely related to interpretation (iii):

iii. Sum of impulse responses. Let e_i be the vector with a 1 in the i^{th} component and zero everywhere else. Consider N experiments where in each experiment,

the system is fed an impulse at the i^{th} input channel, that is, $\mathrm{w}_i(t) = e_i\delta(t)$. Denote the corresponding output by $y_{e_i} \in \mathbb{R}^N$. The (squared) $\mathcal{H}_2$ norm is then the total sum of the $\mathcal{L}_2$ norms of these outputs, that is,

$$\|G\|_2^2 = \sum_{i=1}^{N} \int_0^\infty y_{e_i}^*(t) y_{e_i}(t)\, \mathrm{d}t. \tag{3.14}$$

A stochastic version of this scenario corresponds to a system where the inputs w_i can occur with equal probability. Under this assumption $\|G\|_2^2$ becomes the expected value of this integral: $\|G\|_2^2 = N\mathbb{E}\left\{\int_0^\infty y^*(t)y(t)\, \mathrm{d}t\right\}$, where the expectation is taken over the input channels.

We will also consider the following non-standard scenario:

iv. Response to random steps. Let the individual inputs w_i, $i = 1, \ldots, N$ be step signals at time $t = 0$, whose sizes $\mathrm{w}_{0,i}$ are independent with zero mean and unit correlation. That is, $\mathrm{w}(t) = \mathrm{w}_0\Theta(t)$ where $\mathbb{E}\{\mathrm{w}_0\mathrm{w}_0^*\} = I_N$ and $\Theta(t)$ is the unit step function. The expected $\mathcal{L}_2$ norm of the output:

$$\mathbb{E}\left\{\int_0^\infty y^*(t)y(t)\, \mathrm{d}t\right\}, \tag{3.15}$$

where the expectation is taken over the input signals, is obtained as the $\mathcal{H}_2$ norm of the system from w to the signal $z(t) = \int_0^t y(\tau)\mathrm{d}\tau$, provided it is $\mathcal{L}_2$-bounded. That is, the $\mathcal{H}_2$ norm of the transfer matrix $\frac{1}{s}G(s)$, provided this is an input-output stable system.

The interpretation iv. follows from the stochastic interpretation of iii., noting that a step is an integrated impulse, and that we consider one in every channel. This scenario is particularly useful in cases where the system state is $\psi = [x,\ \dot{x}]^T$ and $y = \mathcal{C}\dot{x}$. In this case, the measure (3.15) under a step input (which, note, is not an $\mathcal{H}_2$ norm), can be obtained as the $\mathcal{H}_2$ norm with respect to the output $z = \mathcal{C}x$.

3.2 Graph theory

The interactions between nodes, or agents, in a networked dynamical system are modeled through graphs. Many of the results in this thesis are based on properties of the underlying network graphs, in particular, spectral properties of the graph Laplacian. The graph Laplacian is one of the standard matrices used to represent graph structures. In this section, we review relevant definitions and properties of the graph Laplacian. This theory can be found in textbooks like [35,59,105], though we note that the notation tends to vary.

3.2.1 Definitions

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a graph, where $\mathcal{V} = \{1, 2, \ldots, N\}$ is the set of nodes (vertices), and $\mathcal{E} = \{(i,j)\} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges. We will consider *weighted* graphs throughout. This means that each edge $(i,j) \in \mathcal{E}$ has an associated nonnegative weight $w(i,j) : \mathcal{E} \mapsto \mathbb{R}$, which we will simply denote w_{ij}. We define $w_{ij} = 0$ if $(i,j) \notin \mathcal{E}$. An *unweighted* graph is equivalent to one where $w_{ij} = 1$ for all $(i,j) \in \mathcal{E}$.

If the graph $\mathcal{G}$ is *directed*, the edge (i,j) points from node i (the tail) to node j (the head). The neighbor set $\mathcal{N}_i$ of node i is then the set of nodes j to which there is an edge $(i,j) \in \mathcal{E}$. The outdegree of node i is defined as $d_i^+ = \sum_{j=1}^N w_{ij}$ and its indegree is $d_i^- = \sum_{j=1}^N w_{ji}$. The graph $\mathcal{G}$ is *balanced* if $d_i^+ = d_i^-$ for all $i \in \mathcal{V}$. It is *strongly connected* if there is a directed path connecting any two nodes $i, j \in \mathcal{V}$ and has a *connected spanning tree* if there is a path from some node $i \in \mathcal{V}$ to any other node $j \in \mathcal{V} \backslash \{i\}$.

The graph $\mathcal{G}$ is *undirected* if $(i,j) \in \mathcal{E} \Rightarrow (j,i) \in \mathcal{E}$ for all $i, j \in \mathcal{V}$ and $w_{ij} = w_{ji}$. The *degree* of a node $i \in \mathcal{V}$ is then $d_i = \sum_{j \in \mathcal{N}_i} w_{ij}$, where the neighbor set $\mathcal{N}_i$ is defined as for the directed graph. If the graph is unweighted, $d_i = |\mathcal{N}_i|$. An undirected graph is called *connected* if there is a path connecting any two nodes $i, j \in \mathcal{V}$. It is always balanced.

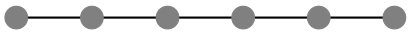
There exist several named, special graphs, some of which will be used in this thesis. Figure 3.1 shows a number of such graphs.

3.2.2 The (weighted) graph Laplacian

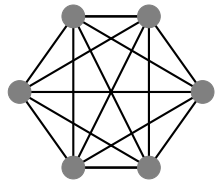
Assume that $\mathcal{G}$ is *undirected*. The weighted graph Laplacian $\mathcal{L} \in \mathbb{R}^{N \times N}$ of $\mathcal{G}$ is defined as follows

$$\mathcal{L}_{ij} = \begin{cases} -w_{ij} & \text{if } j \neq i \text{ and } j \in \mathcal{N}_i \\ \sum_{k \in \mathcal{N}_i} w_{ik} & \text{if } j = i \\ 0 & \text{otherwise.} \end{cases} \tag{3.16}$$

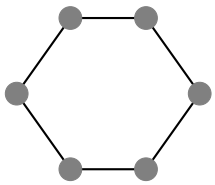
By this definition, $\mathcal{L} = D - A$, where $D = \text{diag}\{d_1, d_2, \ldots, d_N\}$ is the diagonal matrix of node degrees and A is what is called the *adjacency matrix* of the graph.



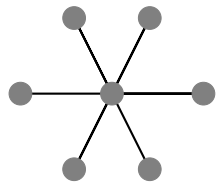
(a) Path graph



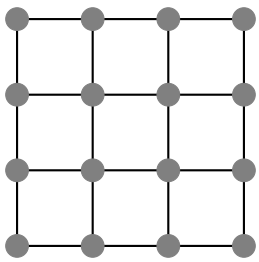
(b) Complete graph



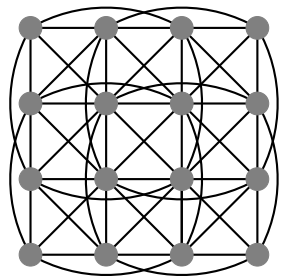
(c) Ring graph



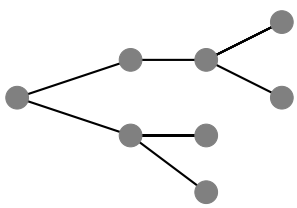
(d) Star graph



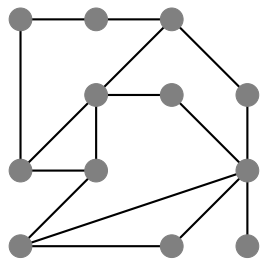
(e) Lattice graph (2D)



(f) 2-fuzz lattice graph (2D)



(g) Example of tree graph



(h) Example of planar graph

Figure 3.1: Some special graphs.

Properties of the graph Laplacian

The weighted graph Laplacian $\mathcal{L}$ defined as in (3.16) has the following properties:

i. Symmetry. Since $\mathcal{G}$ is undirected, the edge from node i to node j is identical to the edge from node j to node i. Therefore, $\mathcal{L}_{ij} = \mathcal{L}_{ji} \quad \forall i, j \in \mathcal{V}$, and $\mathcal{L}$ is symmetric.

ii. Zero row/column sums. Since $\mathcal{L}_{ii} = -\sum_{j \neq i} \mathcal{L}_{ij}$, all rows and columns of $\mathcal{L}$ sum to 0. That means that the vector $\mathbf{1}$ with all components equal to 1 is an eigenvector of $\mathcal{L}$ corresponding to the eigenvalue 0. That is,

$$\mathcal{L}\mathbf{1} = 0.$$

Graph Laplacians are thus singular.

iii. Positive semidefiniteness. If the graph underlying the Laplacian is connected, then, zero is a simple eigenvalue of $\mathcal{L}$. Remaining $N-1$ eigenvalues are positive. If the graph is not connected, the multiplicity of the zero eigenvalue will equal the number of isolated subgraphs.

iv. Diagonalizability by unitary matrix. Since $\mathcal{L}$ is symmetric, it can be diagonalized by a unitary matrix[2] U whose columns are orthonormal, such that $\mathcal{L} = U^* \Lambda U$, where $\Lambda = \text{diag}\{\lambda_1, \ \lambda_2, \ldots, \lambda_N\}$ is a diagonal matrix of the eigenvalues of $\mathcal{L}$. Throughout, we let them be ordered so that

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N.$$

Example 3.1 (Average consensus) Consider a network of N agents, depicted in Figure 3.2, each with a scalar information state $x_i \in \mathbb{R}$, that is controlled according to: $\dot{x}_i = u_i$ (omitting time dependence in the notation). The control objective is for all agents to reach consensus over this state, so that $x_1 = x_2 = \cdots = x_N = x^{\text{avg}}$, where x^{avg} is the average of the initial states. This can be achieved through the standard average consensus algorithm:

$$u_i = \sum_{j \in \mathcal{N}_i} w_{ij}(x_j - x_i).$$

Now, if we define the state vector $x = (x_1, \ldots, x_N)^T$, we can use the graph Laplacian (3.16) to write the control dynamics as:

$$\dot{x} = -\mathcal{L}x. \tag{3.17}$$

If the graph is connected, that is, if there is a path between any two agents in the network, then it is well known that (3.17) achieves the control objective, that is, consensus (see for example [141]).

[2] In fact, U will be a real-valued orthonormal matrix since $\mathcal{L}$ is real symmetric.

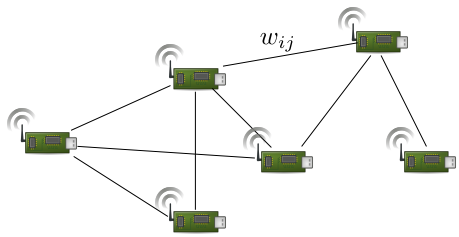


Figure 3.2: A network of $N = 6$ sensors, where the lines symbolize communication links with positive weights w_{ij}.

Graph Laplacians for directed graphs

If $\mathcal{G}$ is directed, the graph Laplacian can be defined just as in (3.16). Each diagonal entry $\mathcal{L}_{ii}$ then represents the outdegree of the i^{th} node. Clearly, $\mathcal{L}$ is not symmetric in this case. However, the rows still sum to zero, so $\mathcal{L}\mathbf{1} = 0$ and the zero eigenvalue is simple if $\mathcal{G}$ has a connected spanning tree. Remaining eigenvalues are in the right half of the complex plane and the eigenvalues can be numbered as $0 = \lambda_1 < \mathrm{Re}\{\lambda_2\} \leq \ldots \leq \mathrm{Re}\{\lambda_N\}$.

The graph Laplacian of a directed graph is not necessarily diagonalizable. However, if the graph is such that $\mathcal{L}$ is *normal*, that is, if $\mathcal{L}^T\mathcal{L} = \mathcal{L}\mathcal{L}^T$, then it is unitarily diagonalizable as described under point iv. above. Normality of $\mathcal{L}$ implies that $\mathcal{G}$ must be balanced.

3.2.3 The grounded graph Laplacian

It will sometimes be useful to consider the *grounded* graph Laplacian $\bar{\mathcal{L}}$. This is the $(N-1) \times (N-1)$ matrix that is obtained by deleting the k^{th} row and column of $\mathcal{L}$ for some k.

For undirected graphs, the grounded Laplacian $\bar{\mathcal{L}}$ is symmetric positive definite, provided the graph $\mathcal{G}$ underlying $\mathcal{L}$ is connected. See [107] for a proof. That is, unlike the ordinary Laplacian, it lacks a zero eigenvalue. We will generally denote the eigenvalues of $\bar{\mathcal{L}}$ by $\bar{\lambda}_n$ for $n = 1, \ldots, N-1$.

Consider the following example for the use of the grounded Laplacian.

Example 3.2 (Leader-follower consensus) Consider the system from Example 3.1, but now assume that Agent number 1 is a leader and that the objective is for all agents to converge to its state $x_1^\star$. Without loss of generality, we can transfer this state to the origin so that $x_1(t) \equiv x_1^\star(t) = 0$. This is equivalent to the slack bus control in the multi-terminal HVDC network from Example 1.4.

The dynamics of the remaining agents ("the followers") can now be written

$$\dot{\bar{x}} = -\bar{\mathcal{L}}\bar{x},$$

where $\bar{\mathcal{L}}$ is the grounded graph Laplacian obtained by removing the first row and column of $\mathcal{L}$, and $\bar{x} = [x_2, \ldots, x_N]^T$. Since $-\bar{\mathcal{L}}$ is negative definite, this system converges exponentially to $0 = x_1^\star$.

The electrical analogy of fixing the state at a given node to zero is to connect it to ground. This is what motivated the term grounded Laplacian.

3.2.4 The algebraic connectivity of a graph

The second-smallest eigenvalue of the graph Laplacian, λ_2, is of particular importance. It is called the *algebraic connectivity* of the graph $\mathcal{G}$, as it reflects how well interconnected the graph is. It can also be called the Fiedler eigenvalue after Miroslav Fiedler who introduced the concept in [52].

The algebraic connectivity is a spectral property of the graph. The term "algebraic" should be understood in contrast to other notions of connectivity in traditional graph theory, such as edge or vertex connectivity, see [35, Chapter 1].

We will discuss the algebraic connectivity and its scaling in network size for select graphs in Chapter 5.

3.2.5 Some important theorems

The following theorems from algebraic graph theory will be useful for the results in this thesis.

Interlacing theorems

The following theorems reveal that eigenvalues can only increase if edges are added to a graph or if their weights increase.

Theorem 3.1 ([109], Theorem 3.2) *Let $\mathcal{G}$ be a graph to which we add an edge e to create the new graph $\mathcal{G}' = \{\mathcal{V}, \mathcal{E} \cup \{e\}\}$. Then, the Laplacian eigenvalues of $\mathcal{G}$ and $\mathcal{G}'$ interlace, that is,*

$$0 = \lambda_1 = \lambda_1' < \lambda_2 \leq \lambda_2' \leq \lambda_3 \leq \ldots \leq \lambda_N \leq \lambda_N', \tag{3.18}$$

where λ_n, $n = 1, \ldots, N$ are the eigenvalues of the graph Laplacian $\mathcal{L}$ of $\mathcal{G}$ and λ_n' are those of the graph Laplacian $\mathcal{L}'$ of $\mathcal{G}'$. At least one inequality is strict.

The following corollary follows trivially.

Corollary 3.2 *Let $\mathcal{G}'$ be a connected, weighted graph with the same vertex set as $\mathcal{G}$, but only a subset of the edges. That is, $\mathcal{G}' = \{\mathcal{V}, \mathcal{E}'\}$, $\mathcal{E}' \subset \mathcal{E}$. Then,*

$$0 = \lambda_1' = \lambda_1 < \lambda_2' \leq \lambda_2 \leq \lambda_3' \leq \ldots \leq \lambda_N' \leq \lambda_N. \tag{3.19}$$

Corollary 3.2 is relevant for the subgraphs of a graph $\mathcal{G}$.

A similar result as in Theorem 3.1 holds if we instead of adding an edge consider increasing the weight of an edge:

Theorem 3.3 *Consider a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$. Construct the graph $\mathcal{G}'$ by increasing the weight of an edge $e' = (i', j') \in \mathcal{E}$. Then the Laplacian eigenvalues of $\mathcal{G}'$ satisfy*

$$\lambda'_n \geq \lambda_n$$

for each $n = 2, \ldots, N$, where λ_n are the eigenvalues of $\mathcal{G}$. At least one inequality is strict.

Proof: Increasing the weight of an edge (i', j') by Δw means that the new graph Laplacian can be written $\mathcal{L}' = \mathcal{L} + \Delta\mathcal{L}$, where $\Delta\mathcal{L}$ is also a positive semidefinite graph Laplacian (of a disconnected graph). By the Courant-Weyl inequalities, see [25, Theorem 2.8.1], this implies that $\lambda'_n \geq \lambda_n$ for each $n = 1, \ldots, N$. The diagonal elements $\mathcal{L}'_{i'i'} > \mathcal{L}_{i'i'}$ and $\mathcal{L}'_{j'j'} > \mathcal{L}_{j'j'}$. Therefore $\mathrm{tr}(\mathcal{L}') = \sum_{n=1}^{N} \lambda'_n > \sum_{n=1}^{N} \lambda_n = \mathrm{tr}(\mathcal{L})$, so at least one inequality must be strict. ■

Effective resistance and Rayleigh's monotonicity law

There is a connection between graph Laplacian eigenvalues and the notion of the *effective resistance* of a graph, see [45, 47, 57]. Consider a network of resistors as in Figure 3.3. The effective resistance between any two nodes i and j is defined as R_{ij}^{eff}. This would correspond to the voltage one would measure if one were to close a circuit with a one Ampere current source between nodes i and j. It can also be called the resistance distance [85].

The *total* effective resistance of the network graph is called the *Kirchhoff index* and is defined as

$$K_f := \sum_{i<j} R_{ij}^{\mathrm{eff}} = \frac{1}{2} \sum_{i,j=1}^{N} R_{ij}^{\mathrm{eff}}. \tag{3.20}$$

It is a scalar measure of how well "connected" the network is, or how "large" it is in terms of resistance distance [57].

The Kirchhoff index has been shown, for example in [64], to equal:

$$K_f = N \sum_{n=2}^{N} \frac{1}{\lambda_n^G}, \tag{3.21}$$

where λ_n^G are the eigenvalues of the *conductance matrix* $\mathcal{L}_G$ of the resistor network. This is the weighted graph Laplacian defined as in (3.16) with the weights $w_{ij} = 1/r_{ij} = g_{ij}$ (g_{ij} is the conductance of the line (i, j)). The following relation will also be of importance:

$$\mathrm{tr}(\mathcal{L}_G^{\dagger}) = \sum_{n=2}^{N} \frac{1}{\lambda_n^G}, \tag{3.22}$$

where $\dagger$ denotes the Moore-Penrose pseudo inverse. The relation (3.22) holds for any graph Laplacian.

We are now ready to state the following Lemma.

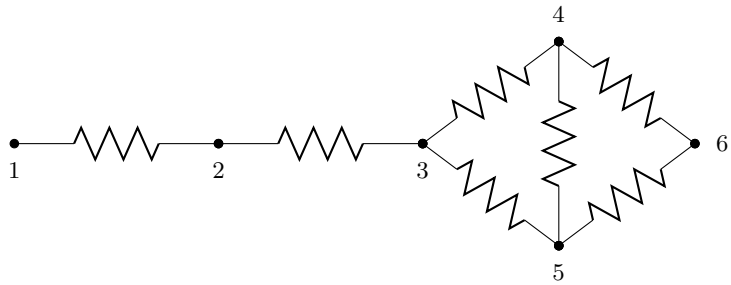


Figure 3.3: A resistor network. The pairwise effective resistance between any two nodes i and j is denoted R_{ij}^{eff}. For example, $R_{13}^{\text{eff}} = r_{12} + r_{23}$. The total effective resistance $\sum_{i<j} R_{ij}^{\text{eff}}$ is called the *Kirchhoff index* and can be evaluated through the Laplacian eigenvalues of the corresponding network graph.

Lemma 3.4 (Rayleigh's monotonicity law) *Removing an edge from a graph, or increasing its resistance, can only increase the effective resistance between any two nodes in the network. Conversely, adding edges or decreasing their resistance can only decrease the effective resistance between any two nodes.*

Proof: See, for example, [45]. ■

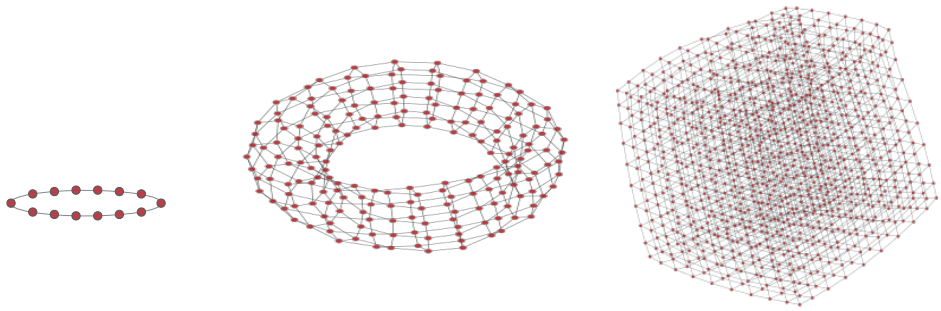
The following corollary follows from Rayleigh's monotonicity law as well as from Theorems 3.1 and 3.3:

Corollary 3.5 *Removing an edge from a graph, or increasing its resistance, can only increase the Kirchhoff index K_f of a graph, and vice versa.*

3.3 Spatially invariant systems

In several parts of this thesis, we will consider networked systems modeled over regular lattice structures. These systems will then be assumed to be *spatially invariant* with respect to the network, which essentially means that the control dynamics are invariant to spatial translations. Spatial invariance can be viewed as a counterpart to time invariance in spatio-temporal systems, and like the assumption of time invariance, it provides a powerful set of tools for a tractable analysis.

The most important of those tools will be spatial Fourier transforms, which allow for an analysis of system properties in what can be thought of as a spatial frequency domain. Later in this thesis, we will introduce a technical framework for evaluating the performance of spatially invariant systems based on such properties. Here, we review some necessary preliminaries. More details to spatially invariant systems can be found in [15].



(a) $d = 1$ (b) $d = 2$ (c) $d = 3$ (without periodicity)

Figure 3.4: Illustration of the d-dimensional toric lattice $\mathbb{Z}_L^d$ for $d = 1, 2, 3$. In the one-dimensional case $\mathbb{Z}_L$ is a ring graph and in the two-dimensional case $\mathbb{Z}_L^2$ is a torus. For $d = 3$ the lattice is drawn without the periodic boundary conditions.

3.3.1 Topology

We will consider problems modeled over the d-dimensional torus $\mathbb{Z}_L^d$ with a total of $N = L^d$ nodes and d assumed finite. In the one-dimensional case ($d = 1$), $\mathbb{Z}_L$ is simply the L node ring graph, which we can represent by the set of integers $\{-\frac{L}{2}, \ldots, 0, 1, \ldots, \frac{L}{2} - 1\}$ mod L for L even, and $\{-\frac{L-1}{2}, \ldots, 0, 1, \ldots, \frac{L-1}{2}\}$ mod L for L odd. $\mathbb{Z}_L^d$ is the direct product of d such rings. See also Figure 3.4.

Going forward, it will also be useful to define the infinite d-dimensional lattice $\mathbb{Z}^d$, which is the direct product of d copies of the integers.

3.3.2 Operations

We will define real-valued function arrays over the network, such as $a : \mathbb{Z}_L^d \mapsto \mathbb{R}$, where we will use multi-index notation to denote the k^{th} array entry $a_k = a_{(k_1,\ldots,k_d)}$. Similarly, the state at node $k = (k_1, \ldots, k_d)$ in the d-dimensional torus is denoted

$$x_{(k_1,\ldots,k_d)}(t),$$

which in our case will either be a scalar in $\mathbb{C}$ or a vector-valued signal in $\mathbb{R}^d$. We will in most cases omit the time dependence in the notation.

Linear operators, denoted by upper case letters, will be used to define multi-dimensional circular convolutions with function arrays over $\mathbb{Z}_L^d$. For example, the

convolution operator A associated with the array a is defined as follows:

$$\begin{aligned} h &= Ax = a \star x \\ &\Updownarrow \\ h_{(k_1,\ldots,k_d)} &= \sum_{(l_1,\ldots,l_d)\in\mathbb{Z}_L^d} a_{(k_1,\ldots,k_d)-(l_1,\ldots,l_d)} x_{(l_1,\ldots,l_d)}, \end{aligned} \tag{3.23a}$$

or, in short,

$$h_k = (Ax)_k = \sum_{l\in\mathbb{Z}_L^d} a_{k-l}x_l. \tag{3.23b}$$

In cases where the state $x \in \mathbb{R}^d$, the array element a_k is a $d \times d$ matrix, which in this thesis will be assumed to be diagonal due to coordinate decoupling. The addition of multi-indices in the $\mathbb{Z}_L^d$ arithmetic is done as follows:

$$k + l = (k_1, \ldots, k_d) + (l_1, \ldots, l_d) = (k_1 + l_1, \ldots, k_d + l_d)_{\bmod L}.$$

Here, $\bmod L$ implies that the operation is circulant. All operations on $\mathbb{Z}_L^d$ will be assumed spatially invariant. They can therefore be represented by convolution operators with single-index arrays as in (3.23).

3.3.3 Spatial Fourier transforms

The spatial discrete Fourier transform (DFT) of the array a will be denoted with $\hat{a}$, and we will use the letter n to denote the index, or *wavenumber*, of the spatial Fourier transform. For example, the function array $a_{(k_1,\ldots,k_d)}$ has $\hat{a}_{(n_1,\ldots,n_d)}$ as its Fourier transform, where the wavenumber $(n_1, \ldots, n_d)$ can be thought of as a spatial frequency variable. We will use the DFT that is defined as:

$$\hat{a}_n := \sum_{k\in\mathbb{Z}_L^d} a_k e^{-\mathbf{j}\frac{2\pi}{L}n\cdot k}, \tag{3.24}$$

where $n \cdot k = n_1k_1 + \cdots + n_dk_d$.

Function arrays can also be defined over the infinite d-dimensional lattice $\mathbb{Z}^d$. We then use the subscript ∞ for the array, as in a_∞, with entries $a_{(k_1,\ldots,k_d)}$ for $k \in \mathbb{Z}^d$. The corresponding convolution operator is denoted A_∞. The Z-transform of a_∞ evaluated on the unit circle $e^{\mathbf{j}\theta}$ is:

$$\hat{a}_\infty(\theta) := \sum_{k\in\mathbb{Z}^d} a_k e^{-\mathbf{j}\theta\cdot k}, \tag{3.25}$$

where $\theta = (\theta_1, \ldots, \theta_d)$ denotes a spatial frequency, which takes values in the multi-variable rectangle $\mathcal{R}^d := [-\pi, \pi]^d$.

We will use the term *(generalized) Fourier symbol* of convolution operators for the DFT or Z-transform of the corresponding function array. For example, $\hat{a}$ in (3.24) is the Fourier symbol of the operator A. The values that $\hat{a}$ takes are exactly the eigenvalues of A. In cases where a is matrix valued, the eigenvalues of A are the union of all eigenvalues of $\hat{a}_{(n_1,\ldots,n_d)}$ as $(n_1, \ldots, n_d)$ runs through $\mathbb{Z}_L^d$.

3.4 Miscellaneous definitions

3.4.1 Notation for scalings

Throughout this thesis, we derive what we term *(asymptotic) scalings* of performance measures with respect to the size of networks. The symbol $\sim$ will be used to denote scalings in the following manner:

$$u(N) \sim v(N) \quad \Leftrightarrow \quad \underline{c}v(N) \leq u(N) \leq \bar{c}v(N), \tag{3.26}$$

for any $N \in \mathbb{N}$, where the fixed constants $\underline{c}$, $\bar{c}$ are independent of the variable N. When a scaling is said to hold *asymptotically*, the relation (3.26) holds for all $N > \bar{N}$ for some fixed $\bar{N}$.

The same notation can be used in the same manner for functions defined on other domains than the natural numbers. We can for example write $u(\theta) \sim v(\theta)$ for $\theta \in [-\pi, \pi] \subset \mathbb{R}$, which again means that $\underline{c}v(\theta) \leq u(\theta) \leq \bar{c}v(\theta)$ for some fixed constants $\underline{c}$, $\bar{c}$.

3.4.2 Representation of AC power networks

An alternating-current (AC) power network can be modeled as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the nodes in $\mathcal{V}$ represent generator or load buses, and the edges in $\mathcal{E}$ represent power lines. Each power line has an associated impedance $z_{ij} = r_{ij} + \mathbf{j}x_{ij}$, where r_{ij} is the line's resistance and x_{ij} its reactance. An example of such a network for $N = 7$ is found in Figure 8.1. The inverse of the line impedance is called line *admittance*:

$$y_{ij} = \frac{1}{z_{ij}} = g_{ij} - \mathbf{j}b_{ij},$$

where $g_{ij} = \frac{r_{ij}}{r_{ij}^2 + x_{ij}^2}$ and $b_{ij} = \frac{x_{ij}}{r_{ij}^2 + x_{ij}^2}$ are , respectively, the *conductance* and *susceptance* of the line. Furthermore, each node $i \in \mathcal{V}$ may have a shunt conductance $\bar{g}_i$ and a shunt susceptance $\bar{b}_i$ which represent the node's connections to ground.

Power networks are generally modeled through the network admittance matrix Y defined as:

$$Y_{ij} := \begin{cases} \bar{g}_i + \bar{b}_i + \sum_{k \in \mathcal{N}_i} (g_{ik} - \mathbf{j}b_{ik}), & \text{if } i = j, \\ -(g_{ij} - \mathbf{j}b_{ij}), & \text{if } i \neq j \text{ and } j \in \mathcal{N}_i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.27}$$

where $\mathcal{N}_i$ denotes the neighbor set of node i. The diagonal elements Y_{ii} of the admittance matrix is the *self-admittance* of node i and is equal to the sum of the admittances of all lines incident to that node (including the shunt).

The admittance matrix can be partitioned into a real and an imaginary part and we define

$$Y = (\mathcal{L}_G + \operatorname{diag}\{\bar{g}_i\}) - \mathbf{j}(\mathcal{L}_B + \operatorname{diag}\{\bar{b}_i\}) \tag{3.28}$$

where $\mathcal{L}_G$ is called the conductance matrix and $\mathcal{L}_B$ the susceptance matrix. The matrices $\mathcal{L}_G$ and $\mathcal{L}_B$ are weighted graph Laplacians where the weights are, respectively, the conductance and susceptance of each edge in the graph.

Chapter 4

Coherence of Large-Scale Networks With Distributed PI and PD Control

In this first technical chapter of the thesis, we consider the problem of distributed consensus among agents with second-order dynamics. The consensus objective is to drive the network of agents to a state of agreement or onto a common trajectory. However, when the system is subject to external disturbances, there are limitations to the achievable performance. In particular, there are limitations to how close the agents of a large-scale network can come to a state of consensus, or what we term the *coherence* of the network. Such fundamental limitations were studied in [14], where scalings of measures of network coherence with network size were established for first- and second-order consensus networks. It was shown that reasonable performance in sparse networks, such as vehicle platoons, requires that each agent can access measurements of its own states with respect to a global reference frame; what we term absolute feedback.

The importance of absolute feedback for network coherence has also been recognized in [93, 128, 133], where the impact of leaders – select agents with access to absolute measurements – was studied in first-order consensus networks. Here, we focus on second-order consensus, and do not consider adding absolute state feedback to improve performance, which would require increasing the agents' sensory capabilities. Instead, we propose the use of alternative controller structures, namely controllers with integral and derivative action in addition to the standard proportional control.

Proportional-integral-derivative (PID) control has been studied for consensus networks in [95] and the role of integral action in disturbance rejection for these systems was described in [5, 54, 150]. While these works have focused on proving convergence of the respective control strategies, we address the question of performance and scalability. We will show that in cases where agents have access to absolute measurements of one of their two states, PI and PD control can relax the performance limitations that apply to systems with only proportional control.

The idea behind this result is simple; integral or derivative action, when applied to the available absolute state measurement, emulates absolute feedback from the other state. For example, the derivative of a position measurement corresponds to a velocity measurement. Therefore, it is intuitive that a performance improvement would be obtained with *ideal* integral or derivative action.

However, the controllers we consider here are modified to account for imperfections. In particular, we first consider a *distributed averaging PI (DAPI)* controller, in which the integral state is passed through a consensus filter to prevent destabilizing drift due to measurement noise. A version of this controller has been proposed in the context of electric power networks in [6,163,196]. Second, we model low-pass filtering of the derivative action in the proposed *filtered distributed PD (F-DPD)* controller. Interestingly, a fundamental performance improvement is achieved with any design of these filters, yet this chapter also provides insights into how their design impacts performance, as well as to their optimal tunings.

4.1 Problem setup

The problem treated in this chapter was already introduced in Section 1.2, but we repeat the setup here in more detail. We also define the performance output that allows us to measure network coherence through the system's input-output $\mathcal{H}_2$ norm.

4.1.1 Network model and definitions

Consider a network of N identical agents modeled by the undirected, connected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, 2, \ldots, N\}$ is the set of nodes and $\mathcal{E} = \{(i, j)\}$ is the set of edges. Assume that each edge has an associated weight $w_{ij} = w_{ji} > 0$ and denote by $\mathcal{L}$ the weighted graph Laplacian of $\mathcal{G}$, defined as in (3.16). Remaining notation in this chapter adheres to the definitions in Section 3.2.

4.1.2 Second-order consensus

We consider a set $\mathcal{V}$ of agents, each governed by double-integrator dynamics and subject to stochastic disturbances:

$$\begin{aligned} \dot{x}_i(t) &= v_i(t) \\ \dot{v}_i(t) &= u_i(t) + \mathrm{w}_i(t). \end{aligned} \tag{4.1}$$

Here, $u_i(t)$ is a control input and $\mathrm{w}_i(t)$ is assumed to be a zero-mean Gaussian white noise process that is uncorrelated across nodes (see Section 1.2). Henceforth, we will often drop the time-dependence in the notation. Without loss of generality, we assume that the states $x_i, v_i \in \mathbb{R}$ for each agent i represent deviations from a desired trajectory $\bar{x}_i$ with common constant velocity $\bar{v}$, so that

$$\bar{x}_i(t) := \bar{v}t + \delta_i,$$

where δ_i is a given setpoint. The states x_i, v_i may carry different meanings depending on the application. We recall two examples at the end of this section.

The control objective is for all agents to follow the desired trajectory, in our case, to drive the states to zero. We model full state feedback control and first consider a standard linear consensus algorithm (see, for example, [140])

$$u_i = -\sum_{j\in\mathcal{N}_i} f_{ij}(x_i - x_j) - \sum_{j\in\mathcal{N}_i} g_{ij}(v_i - v_j) - f_0 x_i - g_0 v_i, \tag{4.2}$$

where $f_{ij},\ g_{ij},\ f_0,\ g_0$ are nonnegative gains. In this chapter, we refer to (4.2) as *proportional (P) control*.

Recall that we distinguish between two types of state measurements and feedback: *relative* and *absolute*. Relative measurements are taken with respect to neighbors, as in $(x_i - x_j),\ (v_i - v_j)$ for $j \in \mathcal{N}_i$, while absolute measurements imply that agent i can access its own state x_i or v_i. In (4.2), we say that *absolute feedback* from the state x_i (v_i) exists if $f_0 > 0$ ($g_0 > 0$). In this chapter, we will consider cases in which absolute measurements are only available for *either* x_i or v_i, and let this apply to all agents. Regarding the assumption on uniform gains $f_0,\ g_0$, see Remark 4.1.

By defining the state vectors $x = [x_1, \dots, x_N]^T$ and $v = [v_1, \dots, v_N]^T$, we can write the system (4.1) with control (4.2) as

$$\begin{bmatrix}\dot{x}\\ \dot{v}\end{bmatrix} = \begin{bmatrix}0 & I\\ -\mathcal{L}_F - f_0 I & -\mathcal{L}_G - g_0 I\end{bmatrix}\begin{bmatrix}x\\ v\end{bmatrix} + \begin{bmatrix}0\\ I\end{bmatrix}\mathrm{w}, \tag{4.3}$$

where $\mathcal{L}_{F(G)}$ is the weighted graph Laplacian defined as in (3.16), with edge weights f_{ij} (g_{ij}) for $(i,j) \in \mathcal{E}$. The vector w contains the noise.

In order to provide tractable closed-form solutions in what follows, we will impose the following assumption:

Assumption 4.1 (Proportional gains) The ratio f_{ij}/g_{ij} is uniform across all $(i,j) \in \mathcal{E}$. We write $f_{ij} = f w_{ij}$ and $g_{ij} = g w_{ij}$, so that $\mathcal{L}_F = f\mathcal{L}$ and $\mathcal{L}_G = g\mathcal{L}$, with $f, g \geq 0$.

Example 4.1 (Vehicular formation) Consider a set of N vehicles in a formation, where the control objective for each vehicle is to follow the trajectory $\bar{x}_i(t) := \bar{v}t + i\Delta$, despite being subject to random forcings w_i. Here, $\bar{v}$ is a common cruising velocity and Δ is the desired inter-vehicle spacing. Each vehicle controls its velocity according to (4.2), with nearest-neighbor interactions, so that $\mathcal{N}_i = \{i+1, i-1\}$ for $i = 2, \dots, N-1$, $\mathcal{N}_1 = \{2\}$ and $\mathcal{N}_N = \{N-1\}$, resulting in a 1-dimensional string formation. The closed-loop system becomes (4.3).

Consider a scenario in which the vehicles are not equipped with speedometers, but have radars to measure relative positions and velocities with respect to neighbors. Furthermore, the position of the lead vehicle is broadcast across the network, so each vehicle can calculate its own position. In this example, there is therefore

absolute feedback from the position x, but not from the velocity v, so $f_0 > 0$ while $g_0 = 0$ in (4.2).

Example 4.2 (Frequency control in power network) Synchronization in power networks is typically studied through a system of coupled swing equations. Under some simplifying assumptions which we will discuss in Chapter 8, the linearized swing equation can be written as:

$$m\ddot{\theta}_i + d\dot{\theta}_i = -\sum_{j\in\mathcal{N}_i} b_{ij}(\theta_i - \theta_j) + P_{m,i}, \tag{4.4}$$

where θ_i is the phase angle deviation at node i, $\dot{\theta}_i = \omega_i$ is the frequency deviation, and m_i and d_i are, respectively, inertia and damping coefficients. The term $P_{m,i}$ can be seen as the net power injection at the node, and b_{ij} is the susceptance of the $(i,j)^{\text{th}}$ power line. In the absence of any additional control input, we also refer to (4.4) as *frequency droop control*.

The system (4.4) can be cast as the P-controlled system (4.3), with $x \mathrel{\widehat{=}} \theta$, $v \mathrel{\widehat{=}} \omega$, $g_0 \mathrel{\widehat{=}} \frac{d}{m}$, $\mathcal{L}_F \mathrel{\widehat{=}} \frac{b_{ij}}{w_{ij}m}\mathcal{L}$, $\mathcal{L}_G = 0$ and $f_0 = 0$. The power injection P_m can be absorbed into the disturbance input w, which we take to represent random fluctuations in generation and load. In this problem, there is absolute feedback from the frequency ω, but only relative feedback from the phase angles θ (absolute measurement of phase angles would require phasor measurement units (PMUs), which in general are not available).

4.1.3 Performance metric

We will evaluate control performance in terms of a network coherence metric. This can be understood as a measure of network disorder, or in other words, how well the control objective of consensus is achieved. Similar to [14, 128], we define the coherence metric as the steady-state variance of the agents' deviation from the network average:

$$\mathbf{V} = \lim_{t\to\infty} \sum_{i\in\mathcal{V}} \mathbb{E}\left\{ \left(x_i(t) - \frac{1}{N}\sum_{j\in\mathcal{V}} x_j(t) \right)^2 \right\}. \tag{4.5}$$

The quantity $\mathbf{V}$ can be evaluated analytically as the squared $\mathcal{H}_2$ norm of an input-output system H from the disturbance input $\mathrm{w}(t)$ in (4.1) to a performance output $y(t)$ defined as

$$y(t) = \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T \right) x(t), \tag{4.6}$$

where $\mathbf{1}$ is the $(N \times 1)$ vector of all ones.

To better analyze the scaling of (4.5) with the network size, we will normalize it by the total number of agents:

$$V_N = \frac{1}{N}\mathbf{V}. \tag{4.7}$$

It is the scaling of V_N as N increases that we will refer to as the level of coherence in the network. For a given control law to be *scalable*, V_N should be uniformly bounded, in which case the system can be regarded as fully coherent.

4.2 Limitations of proportional control

We begin the discussion on limitations of the proportional control law (4.2) by stating the closed-form expression for its performance.

Theorem 4.1 *The scaled performance output variance* (4.7) *for the P-controlled system* (4.3) *is given by*

$$V_N^{\mathrm{P}} = \frac{1}{2N}\sum_{n=2}^{N}\frac{1}{(f_0+f\lambda_n)(g_0+g\lambda_n)}. \tag{4.8}$$

Proof: See Appendix 4.A.

Clearly, in the absence of absolute feedback (that is, if $f_0 = 0$ and/or $g_0 = 0$), the sum in (4.8) tends towards infinity if one or more of the Laplacian eigenvalues λ_n, $n \geq 2$ approaches zero. Laplacian eigenvalues typically tend to zero as a network grows large, unless it is densely interconnected. If the sum increases faster than linearly in N, the performance metric V_N^P will scale badly and the network lacks coherence.

This was the problem considered in [14], where asymptotic scalings of expressions such as (4.8) were derived for networks built over d-dimensional toric lattices. Their results for the most problematic 1-dimensional network (a ring graph) show that V_N^P scales linearly in N if one of $f_0,\ g_0$ is zero, and as N^3 if both are zero. The same scalings can be derived using arguments based on effective resistances for networks that can be embedded in lattice networks, see [18], or for a recent application in power systems [9]. We will return to such regular networks and prove these scalings formally in Chapter 6.

This points at a fundamental limitation to the performance of the P controller (4.3), as the variance (4.8) is only bounded for any network if there is absolute feedback from both the states x_i and v_i. That is, both measurements are required for P control to be scalable. An important question has therefore been whether alternative controller structures can alleviate these limitations. In the next section, we present linear controllers for which it suffices to have an absolute measurement of either x_i or v_i to achieve bounded scalings for any network.

Remark 4.1 The assumption that the absolute feedback gains f_0, g_0 are uniform across the network is needed to derive the closed-form expression (4.8), but can be shown not to be important for the main conclusion that V_N^P scales badly in the absence of absolute feedback. With non-uniform gains f_0, g_0, a lower bound on V_N^P on the form (4.8) can instead be stated in terms of the maximum available gains, see Appendix 4.B. □

4.3 Distributed PI and PD control

We now introduce the distributed PI and PD controllers with which we propose to address the limitations posed by proportional control. We show that these controllers achieve bounded output variances (4.7), and demonstrate the performance improvement through numerical examples.

4.3.1 Absolute v-feedback: Distributed PI control

Suppose that an absolute measurement of v_i is available, while one of x_i is not. For this case, we propose to use the following *distributed averaging proportional-integral (DAPI)* controller:

$$\begin{aligned} u_i &= -\sum_{j\in\mathcal{N}_i} f_{ij}(x_i - x_j) - \sum_{j\in\mathcal{N}_i} g_{ij}(v_i - v_j) - g_0 v_i + K_I z_i \\ \dot{z}_i &= -v_i - \sum_{j\in\mathcal{N}_i} c_{ij}(z_i - z_j), \end{aligned} \tag{4.9}$$

where z_i is the integral state, K_I is a positive gain (integral gain), and we will discuss the averaging filter $-\sum_{j\in\mathcal{N}_i} c_{ij}(z_i - z_j)$ shortly. First, note that if $c_{ij} = 0$ for all (i,j), it holds that $z_i(t) = \int_0^t v_i(\tau)d\tau + z_i(0)$ and since, by definition, $\int_0^t v_i(\tau)d\tau = x_i(t) - x_i(0)$, the integral state z_i would correspond to an absolute measurement of the state x_i, modulo initial values. In this case, we would therefore expect that the controller (4.9) has a performance similar to (4.2) with absolute feedback from both v_i and x_i. However, as mentioned in this chapter's introduction, an averaging filter on the integral states with positive gains c_{ij} is required to ensure system stability [5]. We will show that also with this filter, the desired performance scaling can be achieved.

By inserting (4.9) into (4.1) we obtain, in vector form:

$$\begin{bmatrix} \dot{x} \\ \dot{v} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 \\ -\mathcal{L}_F & -\mathcal{L}_G - g_0 I & K_I \\ 0 & -I & -\mathcal{L}_C \end{bmatrix} \begin{bmatrix} x \\ v \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} \mathrm{w}. \tag{4.10}$$

In line with Assumption 4.1, we make the following assumption on the weighted graph Laplacian $\mathcal{L}_C$:

Assumption 4.2 The gains c_{ij} are proportional to f_{ij}, g_{ij} for all $(i,j) \in \mathcal{E}$. We write $c_{ij} = cw_{ij}$, so that $\mathcal{L}_C = c\mathcal{L}$, with $c \geq 0$.

Under Assumptions 4.1–4.2, the system (4.10) is input-output stable with respect to the output (4.6) [174] and its performance can be stated as follows:

Proposition 4.2 *The scaled performance output variance* (4.7) *for the DAPI-controlled system* (4.10) *is given by*

$$V_N^{\mathrm{DAPI}} = \frac{1}{2N} \sum_{n=2}^{N} \frac{1}{fg\lambda_n^2 + \frac{K_I f(g_0+\lambda_n(c+g))+g_0 f\lambda_n(c^2\lambda_n+f+cg_0)}{f+cg_0+c\lambda_n(c+g)}}. \tag{4.11}$$

Proof: The result (4.11) is derived in the same manner as in the proof of Theorem 4.1. ■

While the expression (4.11) is convoluted, it allows for the following important conclusion:

Corollary 4.3 *For any positive and finite gains K_I and c, V_N^{DAPI} in* (4.11) *is uniformly bounded in N. It holds that*

$$0 < V_N^{\mathrm{DAPI}} < \frac{f+cg_0}{2K_I f g_0}. \tag{4.12}$$

Proof: It is readily verified that the expression

$$s_n := \frac{1}{fg\lambda_n^2 + \frac{K_I f(g_0+\lambda_n(c+g))+g_0 f\lambda_n(c^2\lambda_n+f+cg_0)}{f+cg_0+c\lambda_n(c+g)}} \tag{4.13}$$

from the terms in (4.11) is monotonically decreasing in $\lambda_n \geq 0$. Its supremum $\bar{s}_n$ is therefore obtained as $\lambda_n \to 0$ and is $\bar{s}_n = \frac{f+cg_0}{K_I f g_0}$. Thus, the sum $\sum_{n=2}^{N} s_n < (N-1)\bar{s}_n < N\frac{f+cg_0}{K_I f g_0}$, which inserted in (4.11) gives the upper bound in (4.12). The lower bound is obtained when $\lambda_n \to \infty$ for all $n = 2, \ldots, N$. This is, for example, the case when $\mathcal{G}$ is the complete graph $\mathcal{K}_N$, as $N \to \infty$. ■

Corollary 4.3 implies that the DAPI controller, unlike the P controller, is scalable and allows any network to be fully coherent.

Remark 4.2 If $c \to 0$ in (4.11), the output variance for P control with absolute feedback from both x_i and v_i is retrieved, substituting the integral gain K_I for f_0 in (4.8). This is in line with the discussion at the beginning of this section. Interestingly though, letting $c \to 0$ does not necessarily minimize V_N^{DAPI}. We will discuss how to optimize c for performance in Section 4.4. □

Example 4.2 (Continued) Applying DAPI control to the power system dynamics (4.4) yields the closed-loop system:

$$\begin{aligned}\dot{\omega}_i &= -\frac{d}{m}\omega_i - \frac{1}{m}\sum_{j\in\mathcal{N}_i} b_{ij}(\theta_i - \theta_j) + K_I z_i + P_{m,i} \\ \dot{z}_i &= -\omega_i - \sum_{j\in\mathcal{N}_i} c_{ij}(z_i - z_j).\end{aligned} \tag{4.14}$$

In this case, the integral action is also referred to as *secondary frequency control*, the role of which is to eliminate any stationary frequency control errors that arise with only P control [6, 163]. As we have shown in this section, it also improves transient performance.

On the one hand, the fact that V_N^{DAPI} is bounded implies that the network can remain fully coherent in terms of phase angles under white power injection noise $P_m = \mathrm{w}$. On the other hand, Interpretation iv. in Section 3.1.3 allows an alternative scenario where

$$V_N^{\mathrm{DAPI}} = \mathbb{E}\left\{\int_0^\infty \frac{1}{N}\sum_{i=1}^{n}\left(\omega_i(t) - \frac{1}{N}\sum_{j=1}^{N}\omega_j(t)\right)^2 \mathrm{d}t\right\},$$

and the expectation is taken over the net power injection (equivalently, the net power load), which for $t \geq 0$ is modeled as a zero-mean random variable with correlation $\mathbb{E}\{P_m P_m^*\} = I_N$. This interpretation relates directly to the control objective of frequency synchronization; a large integral indicates a lack of frequency coherence over the transient. Step changes in loads are also a realistic input scenario in a power network.

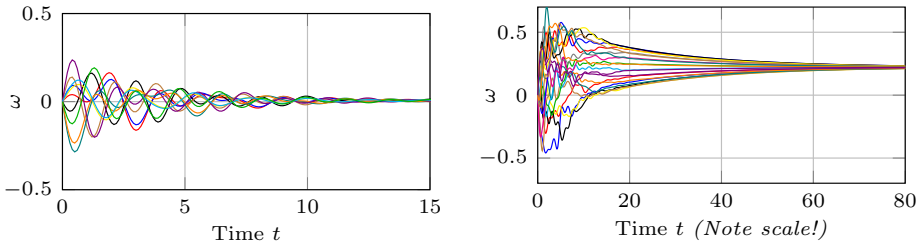
In Figure 4.1, we show a simulation of P control (4.4) and DAPI control (4.14) on a theoretical radial power network (here modeled by a path graph) with, respectively, 10 and 100 nodes under random step changes in the load. The figures show that DAPI control scales significantly better to the larger network than P control.

In the simulation, we have set $m = \frac{20}{\omega^{\mathrm{ref}}}$, $d = \frac{10}{\omega^{\mathrm{ref}}}$, with $\omega^{\mathrm{ref}} = 2\pi 60$ Hz, $K_I = 1$, $c = 0.1$ and $b_{ij} = 0.3$ for all $(i, j) \in \mathcal{E}$.

A simulation of the DAPI controller on a 469 generator model of the Nordic power system depicted in Figure 1.7b is found in [8]. That simulation supports the conclusions from this example.

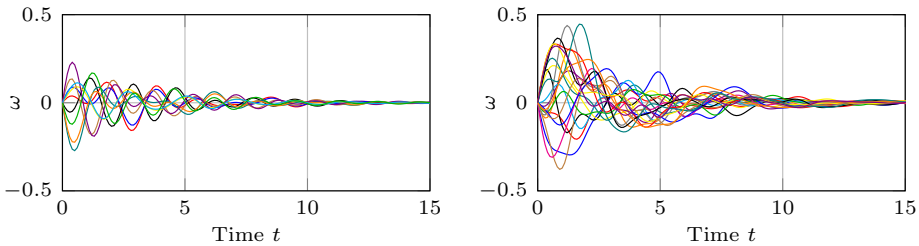
4.3.2 Absolute x-feedback: Distributed PD control

Suppose now that an absolute measurement of x_i is available, while one of v_i is not. In this case, distributed proportional-derivative (PD) control can be used to improve performance compared to the P controller. With ideal derivative action,



(a) P control (Droop), $N = 10$

(b) P control (Droop), $N = 100$



(c) DAPI control, $N = 10$

(d) DAPI control, $N = 100$

Figure 4.1: Frequency trajectories in power network with path graph topology with $N = 10$ and $N = 100$, under random step changes in the net power load at time $t = 0$ (only subset shown for $N = 100$). We note that with droop control, performance deteriorates significantly as N is increased, which is not the case with DAPI control.

the controller would take the form:

$$u_i = -\sum_{j\in\mathcal{N}_i} f_{ij}(x_i - x_j) - \sum_{j\in\mathcal{N}_i} g_{ij}(v_i - v_j) - f_0 x_i - K_D \frac{\mathrm{d}x_i}{\mathrm{d}t}, \tag{4.15}$$

which, since $\frac{\mathrm{d}x_i}{\mathrm{d}t} = v_i(t)$, is identical to (4.2) with absolute feedback from both x_i and v_i, substituting the derivative gain K_D for g_0. Clearly, this controller would therefore also have the same performance. Ideal derivative action is, however, neither possible nor desirable to implement, partly due to the sensitivity to high-frequency noise. We therefore consider a controller where the derivative action is low-pass filtered:

$$\begin{aligned} u_i &= -\sum_{j\in\mathcal{N}_i} f_{ij}(x_i - x_j) - \sum_{j\in\mathcal{N}_i} g_{ij}(v_i - v_j) - f_0 x_i + z_i \\ \dot{z}_i &= -\frac{1}{\tau} z_i - \frac{1}{\tau} K_D \frac{\mathrm{d}x_i}{\mathrm{d}t}. \end{aligned} \tag{4.16}$$

Here, the state z_i corresponds to the derivative action, and $\tau > 0$ is the time constant of the filter. We will refer to the controller (4.16) as *filtered distributed*

PD control (F-DPD). Note that the ideal PD controller is retrieved for $\tau = 0$.

Remark 4.3 An alternative approach would be to apply ideal derivative action and low-pass filter the entire control signal u_i in (4.15). The resulting performance scaling is similar. We choose the controller (4.16) to enable a better comparison to the P controller (4.2), as the proportional action is unfiltered in both. □

The system (4.1) with the controller (4.16) becomes:

$$\begin{bmatrix} \dot{x} \\ \dot{v} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 \\ -\mathcal{L}_F - f_0 I & -\mathcal{L}_G & I \\ 0 & -\frac{1}{\tau} K_D I & -\frac{1}{\tau} I \end{bmatrix} \begin{bmatrix} x \\ v \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} w. \tag{4.17}$$

The performance of this system, under Assumption 4.1, is given by the following proposition:

Proposition 4.4 *The scaled performance output variance* (4.7) *for the F-DPD-controlled system* (4.17) *is given by*

$$V_N^{\mathrm{F-DPD}} = \frac{1}{2N} \sum_{n=2}^{N} \frac{1}{(f_0 + f\lambda_n)\left(g\lambda_n + \frac{K_D(\tau g\lambda_n + 1)}{\tau^2(f_0 + f\lambda_n) + \tau g\lambda_n + 1}\right)} \tag{4.18}$$

Proof: This result is derived in the same manner as in the proof of Theorem 4.1. The system's stability can in this case be verified through the characteristic polynomial $\xi^3 + (g\lambda_n + \frac{1}{\tau})\xi^2 + (\frac{K_D + g\lambda_n}{\tau} + f_0 + f\lambda_n)\xi + \frac{f_0 + f\lambda_n}{\tau} = 0$, which satisfies the Routh-Hurwitz criterion. ■

Again, the expression is convoluted, but it reveals that the output variance is bounded for any network:

Corollary 4.5 *For any positive and finite gains K_D and τ, the variance $V_N^{\mathrm{F-DPD}}$ in* (4.18) *is uniformly bounded in N. It holds that*

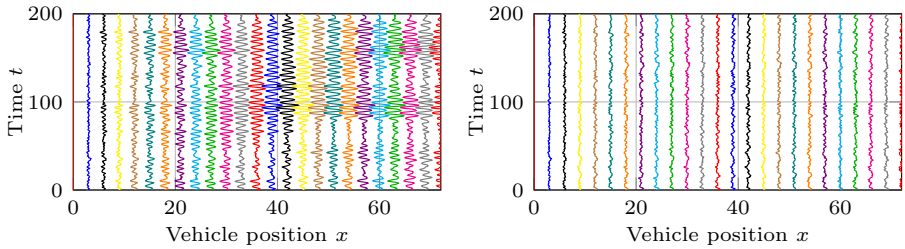
$$0 < V_N^{\mathrm{F-DPD}} < \frac{\tau^2 f_0 + 1}{2 f_0 K_D} \tag{4.19}$$

Proof: Similar to the proof of Corollary 4.3, it can be verified that

$$s_n := \frac{1}{(f_0 + f\lambda_n)\left(g\lambda_n + \frac{K_D(\tau g\lambda_n + 1)}{\tau^2(f_0 + f\lambda_n) + \tau g\lambda_n + 1}\right)}$$

from the terms in (4.18) is monotonically decreasing in $\lambda_n > 0$. An upper bound is thus obtained as $\lambda_n \to 0$, which gives $s_n < \frac{\tau^2 f_0 + 1}{f_0 K_D}$ for all n. Therefore, $V_N^{\mathrm{F-DPD}} = \frac{1}{2N} \sum_{n=2}^{N} s_n < \frac{N-1}{2N} \frac{\tau^2 f_0 + 1}{f_0 K_D} < \frac{\tau^2 f_0 + 1}{2 f_0 K_D}$. As with Corollary 4.3, the lower bound of 0 is obtained, for example, by letting $\lambda_n \to \infty$. ■

Corollary 4.5 implies that the F-DPD controller, unlike the P controller, is scalable and allows any network to be fully coherent.



(a) P control (b) F-DPD control

Figure 4.2: Subset of position trajectories (relative to leader) for a 100 vehicle formation, subject to noise. With F-DPD control, trajectories are more coherent.

Example 4.1 (Continued) Consider again the formation of vehicles. Since no absolute feedback from velocities v_i was available, we implement the F-DPD controller (4.16).

Figure 4.2 shows a simulation of a 100 vehicle formation under a white noise disturbance input. While the spacings between neighbors seem well-regulated in both systems, the P-controlled system appears less coherent.

In the simulation, we have set $f = g = f_0 = K_D = 1$, and $\tau = 0.1$.

4.4 Controller tuning for improved coherence

Neither the DAPI controller (4.9), nor the F-DPD controller (4.16) represents ideal integral or derivative action. In both cases, we have modeled filters to mitigate well-known issues related to noise and uncertainties. In this section, we discuss how the design of those filters impacts performance.

4.4.1 Optimal distributed averaging in DAPI

As already discussed, robust stability of the DAPI controlled system (4.10) requires an alignment between the agents' integral states z_i through a distributed averaging filter. While the control objectives will be reached for any non-zero gains c_{ij} [163], an important control design question is how to choose these gains to optimize network performance. In our case, this choice is reflected through the constant c (see Assumption 4.2). Consider the following proposition.

Proposition 4.6 *For a given DAPI controlled system* (4.10)*, there is* either *a finite* $c^\star > 0$ *that minimizes* V_N^{DAPI} *in* (4.11)*,* or V_N^{DAPI} *is minimized by* $c^\star = 0$ *and is monotonically increasing for* $c > 0$*.*

The case $c^\star > 0$ holds if

$$f > \frac{1}{\lambda_n}(g\lambda_n + g_0)^2 \qquad \forall n = 2, \dots, N$$

and the case $c^\star = 0$ holds if

$$f \leq \frac{1}{\lambda_n}(g\lambda_n + g_0)^2 \qquad \forall n = 2, \dots, N.$$

If none of these applies, $c^\star$ must be evaluated case by case.

Proof: See Appendix 4.C.

Finding a closed form expression for $c^\star$ is in general not tractable due to the many terms in the sum (4.11), and must be done on a case-by-case basis. However, it is possible in the special case of a complete network graph:

Proposition 4.7 *If the network graph $\mathcal{G}$ is complete and the edge weights $w_{ij} = w$ for all $(i,j) \in \mathcal{E}$, then $c^\star$ is given by*

$$c^\star = \sqrt{\frac{f}{Nw}} - g + \frac{g_0}{Nw} \tag{4.20}$$

if this is a positive number. Otherwise $c^\star = 0$.

Proof: In this case, the eigenvalues $\lambda_n = Nw$ for $n \geq 2$, and (4.20) follows from the proof of Proposition 4.6. ■

Remark 4.4 An explicit inclusion of measurement noise in the model increases the output variance V_N^{DAPI}, and shifts the optimal $c^\star$. Such a model reveals that $c = 0$ is always suboptimal, since V_N^{DAPI} is then infinite. We will discuss in detail how measurement noise impacts the performance of DAPI in Chapter 7. □

It is interesting to note that in cases where $c^\star > 0$, the performance of the DAPI controlled system at the optimum is better than with P control and absolute feedback from both x_i and v_i (as that performance is retrieved for $c = 0$, see Remark 4.2). Looking at the expression (4.20), we note that this occurs in particular if g is small relative to f, so that there is little alignment in the state v between agents. This is the case in the power network example (Example 4.2), where $g = 0$. We will return to analyzing optimal tuning of DAPI controllers for such systems in Chapter 10.

4.4.2 Impact of low-pass filter in F-DPD

The low pass filter in the F-DPD controller (4.16) is included as a more realistic implementation of derivative action, which is otherwise well-known to be highly sensitive to high frequency variations and noise. While, by Corollary 4.5, F-DPD achieves bounded output variance for any finite filter constant τ, choosing a small value (corresponding to a high bandwidth) gives better performance:

Proposition 4.8 *For a given F-DPD controlled system* (4.17)*, the output variance* $V_N^{\mathrm{F-DPD}}$ *in* (4.18) *is minimized by* $\tau = 0$ *and is monotonically increasing for* $\tau > 0$*.*

Proof: It holds that

$$\frac{\mathrm{d}}{\mathrm{d}\tau} V_N^{\mathrm{F-DPD}} = \frac{1}{2N} \sum_{n=2}^{N} \frac{K_D g \lambda_n^2 \tau^2 + 2 K_D \lambda_n \tau}{2(g^2 \lambda_n^2 \tau + f g \lambda_n^2 \tau^2 + f_0 g \lambda_n \tau^2 + K_D g \lambda_n \tau + g \lambda_n + K_D)^2}.$$

Since $K_D, \lambda_n, f_0 > 0$, $f, g \geq 0$, it holds $\frac{\mathrm{d}}{\mathrm{d}\tau} V_N^{\mathrm{F-DPD}} = 0$ for $\tau = 0$ and $\frac{\mathrm{d}}{\mathrm{d}\tau} V_N^{\mathrm{F-DPD}} > 0$ for all $\tau > 0$. ■

This is an intuitive result, as $\tau = 0$ would give ideal derivative action and therefore the most accurate substitute for absolute feedback from v_i. However, a greater value for τ would make the system less sensitive to noise. This trade-off must be done based on system-specific knowledge.

Remark 4.5 We have limited the analysis here to a first order filter, as higher order filters make the expressions for output variances even more convoluted. However, numerical results indicate that higher order filters increase $V_N^{\mathrm{F-DPD}}$, though it remains bounded. □

4.5 Discussion

In this chapter, we have addressed limitations to the performance of standard static, or proportional, feedback control in double-integrator networks. These limitations imply that the variance V_N defined in (4.5) that characterizes network coherence may scale unboundedly in network size unless absolute feedback from both the systems' states is available. Here, we addressed these shortcomings by proposing distributed PI and PD controllers, which radically improve performance by making V_N bounded if absolute measurements from one of the system's states are available.

As we could see in this thesis' introduction, the second-order consensus-type dynamics considered here is relevant for a variety of applications ranging from biological networks to coordination of robots, see, for example, [140]. Here, we treated examples from vehicular formation and power networks. Our results imply that if agents have limited access to absolute feedback, then the proposed PI and PD controllers are preferable to P control, in particular for large-scale networks. We remark that this conclusion holds even though the proposed controllers do not model *ideal* integral or derivative action, but have filters that mitigate effects that would arise in a practical setting.

However, an explicit inclusion of measurement noise on the available absolute state measurements changes the conclusion that distributed PI control can fundamentally improve performance compared to P control. This does, namely, cause additional variance that is small, but scales unboundedly with network size. We treat this particular problem in detail in Chapter 7. The impact of other practical

network aspects, such as communication delays, non-reliability of the channels and non-symmetries, is an important research question that is left for future work.

Appendix to Chapter 4

4.A Proof of Theorem 4.1

The output variance $\mathbf{V}$ in (4.5) is obtained as the squared $\mathcal{H}_2$ norm of the system H from input w in (4.3) to the output y in (4.6). We derive the norm through a unitary state transformation: $x =: U\hat{x}$, $v =: U\hat{v}$, where U is the unitary matrix that diagonalizes the Laplacian matrix $\mathcal{L}$ (and by Assumption 4.1 also $\mathcal{L}_F, \mathcal{L}_G$), so that $\mathcal{L} = U^*\Lambda U$ with $\Lambda = \mathrm{diag}\{\lambda_1, \cdots, \lambda_N\}$. Due to the unitary invariance of the $\mathcal{H}_2$ norm (see Section 3.1.2), we can also transform the input and output according to $\hat{w} := U^*w$, $\hat{y} := U^*y$.

This state transformation block-diagonalizes the system (4.3):

$$\begin{bmatrix}\dot{\hat{x}}\\ \dot{\hat{v}}\end{bmatrix} = \begin{bmatrix}0 & I\\ -f\Lambda - f_0 & -g\Lambda - g_0\end{bmatrix}\begin{bmatrix}\hat{x}\\ \hat{v}\end{bmatrix} + \begin{bmatrix}0\\ I\end{bmatrix}\hat{w}$$
$$\hat{y} = \begin{bmatrix}I^{\mathrm{red}} & 0\end{bmatrix}\begin{bmatrix}\hat{x}\\ \hat{v}\end{bmatrix},$$

where the matrix $I^{\mathrm{red}} = \begin{bmatrix}0 & 0_{N-1}^T\\ 0_{N-1} & I_{N-1}\end{bmatrix}$. This corresponds to N decoupled subsystems $\hat{H}_n$:

$$\begin{aligned}\begin{bmatrix}\dot{\hat{x}}_n\\ \dot{\hat{v}}_n\end{bmatrix} &= \underbrace{\begin{bmatrix}0 & 1\\ -f\lambda_n - f_0 & -g\lambda_n - g_0\end{bmatrix}}_{=:\hat{\mathcal{A}}_n}\begin{bmatrix}\hat{x}_n\\ \hat{v}_n\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ 1\end{bmatrix}}_{=:\hat{\mathcal{B}}_n}\hat{w}_n\\ \hat{y}_n &= \underbrace{\begin{bmatrix}1 & 0\end{bmatrix}}_{=:\hat{\mathcal{C}}_n}\begin{bmatrix}\hat{x}_n\\ \hat{v}_n\end{bmatrix}\end{aligned} \tag{4.21}$$

for $n = 2, \ldots, N$. The mode $n = 1$ corresponds to the average mode in x, which is unobservable from the output (4.6), so $\hat{y}_1 \equiv 0$. Therefore, even though this mode would be undamped if $f_0 = 0$, the overall system remains input-output stable since $\hat{\mathcal{A}}_n$ is Hurwitz for $n \geq 2$ (its characteristic polynomial is $\xi^2 + (g_0 + g\lambda_n)\xi + f_0 + f\lambda_n$, which has only positive coefficients).

It holds that

$$\mathbf{V} = ||H||_2^2 = \sum_{n=1}^{N} ||\hat{H}_n||_2^2 = \sum_{n=2}^{N} ||\hat{H}_n||_2^2,$$

where the last equality is due to the fact that $\hat{y}_1 \equiv 0$ and therefore $||\hat{H}_1||_2^2 = 0$. Each subsystem norm is obtained as $||\hat{H}_n||_2^2 = \mathrm{tr}\{\hat{\mathcal{B}}_n^T P_n \hat{\mathcal{B}}_n\}$, where P_n is the solution to the Lyapunov equation (3.11). Straightforward calculations give that

$$||\hat{H}_n||_2^2 = \frac{1}{2(f\lambda_n + f_0)(g\lambda_n + g_0)}$$

for each n, and the result (4.8) follows. ■

4.B Performance bound in the case of non-uniform absolute feedback gains

Consider Remark 4.1. With non-uniform absolute feedback gains, the control law (4.2) reads

$$u_i = -\sum_{j\in\mathcal{N}_i} f_{ij}(x_i - x_j) - \sum_{j\in\mathcal{N}_i} g_{ij}(v_i - v_j) - f_{0,i}x_i - g_{0,i}v_i. \tag{4.22}$$

Assume $f_{0,i} \leq f_0^{\max}$, $g_{0,i} \leq g_0^{\max}$ for $i \in \mathcal{V}$ and for technical reasons let $f_0^{\max} > 0$. We prove the lower bound

$$V_N^{\mathrm{P}} \geq \frac{1}{2N}\sum_{n=2}^{N} \frac{1}{(f_0^{\max} + f\lambda_n)(g_0^{\max} + g\lambda_n)}.$$

On vector form, we can write the system with the control law (4.22) as

$$\begin{aligned}\begin{bmatrix}\dot{x}\\ \dot{v}\end{bmatrix} &= \Big(\underbrace{\begin{bmatrix}0 & I\\ -\mathcal{L}_F - f_0^{\max}I & -\mathcal{L}_G - g_0^{\max}I\end{bmatrix}}_{\mathcal{A}_0} + \underbrace{\begin{bmatrix}0 & I\\ F' & G'\end{bmatrix}}_{\mathcal{A}'}\Big)\begin{bmatrix}x\\ v\end{bmatrix} + \begin{bmatrix}0\\ I\end{bmatrix}\mathrm{w},\\ y &= \begin{bmatrix}0 & I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\end{bmatrix}\begin{bmatrix}x\\ v\end{bmatrix}\end{aligned} \tag{4.23}$$

where $F' = \mathrm{diag}\{f_0^{\max} - f_{0,i}\} \geq 0$ and $G' = \mathrm{diag}\{g_0^{\max} - g_{0,i}\} \geq 0$. Now, assume P_0 solves the Lyapunov equation

$$\mathcal{A}_0^T P_0 + P_0\mathcal{A}_0 = -\mathcal{C}^T\mathcal{C}. \tag{4.24}$$

This is the equation we solved in Appendix 4.A. Let $P = P_0 + P'$ be the controllability Gramian for the system (4.23), that is, the solution to: $(\mathcal{A}_0 + \mathcal{A}')^T(P_0 + P') + (P_0 + P')(\mathcal{A}_0 + \mathcal{A}') = -\mathcal{C}^T\mathcal{C}$. This is equivalent to

$$\begin{aligned}&\mathcal{A}_0^T P_0 + P_0\mathcal{A}_0 + \mathcal{A}'^T P_0 + P_0\mathcal{A}' + (\mathcal{A}_0 + \mathcal{A}')^T P' + P'(\mathcal{A}_0 + \mathcal{A}')^T = -\mathcal{C}^T\mathcal{C}\\ &\quad\Leftrightarrow (\mathcal{A}_0 + \mathcal{A}')^T P' + P'(\mathcal{A}_0 + \mathcal{A}')^T = -\left(\mathcal{A}'^T P_0 + P_0\mathcal{A}'\right) =: -Q,\\ &\quad\Leftrightarrow \mathcal{A}^T P' + P'\mathcal{A} = -Q\end{aligned} \tag{4.25}$$

where in the first step we subtracted (4.24) and in the second we re-introduced $\mathcal{A} = \mathcal{A}_0 + \mathcal{A}'$. The matrix $Q \geq 0$ since $P_0 \geq 0$ and $\mathcal{A}' \geq 0$ ($\mathcal{A}'$ is lower triangular with nonnegative elements on the diagonal). Given that $\mathcal{A}$ is Hurwitz if $f_0^{\max} > 0$, this means that the solution P' to (4.25) must be positive semidefinite, that is, $P' \geq 0$. This means that the diagonal elements $\{P'\}_{ii} \geq 0$, and therefore that

$$\begin{aligned}\|H\|_2^2 &= \mathrm{tr}(\mathcal{B}^T P\mathcal{B}) = \mathrm{tr}(\mathcal{B}^T P_0\mathcal{B}) + \mathrm{tr}(\mathcal{B}^T P'\mathcal{B})\\ &\geq \mathrm{tr}(\mathcal{B}^T P_0\mathcal{B}) = \sum_{n=2}^{N}\frac{1}{2(f_0^{\max} + f\lambda_n)(g_0^{\max} + g\lambda_n)}.\end{aligned} \tag{4.26}$$

The bound on V_N^P follows.

4.C Proof of Proposition 4.6

Clearly, it holds that $\frac{\mathrm{d}}{\mathrm{d}c}V_N^{\mathrm{DAPI}} = \frac{1}{2N}\sum_{n=2}^{N}\frac{\mathrm{d}}{\mathrm{d}c}s_n$, where s_n was defined in (4.13). Each term s_n can be written as a fraction $s_n = \frac{p_n}{q_n}$. By the quotient rule, $\frac{\mathrm{d}s_n}{\mathrm{d}c} = \frac{p_n' q_n - p_n q_n'}{(q_n)^2}$, where $(q_n)^2 > 0$ for any c, and

$$p_n' q_n - p_n q_n' = K_I \lambda_n^2 \left[c^2 + 2\left(g + \frac{g_0}{\lambda_n}\right) c + \left(g + \frac{g_0}{\lambda_n}\right)^2 - \frac{f}{\lambda_n} \right].$$

Inspecting this expression reveals that, since $K_I, \lambda_n, g_0 > 0$, and $f, g \geq 0$, it holds that $\frac{\mathrm{d}s_n}{\mathrm{d}c} > 0$ for any $c > 0$ if $\frac{f}{\lambda_n} \leq (g + \frac{g_0}{\lambda_n})^2$ for all $n = 2, \ldots, N$. In this case $\frac{\mathrm{d}}{\mathrm{d}c}V_N^{\mathrm{DAPI}} > 0$ for all $c > 0$, and $c^\star = 0$.

Conversely, if $\frac{f}{\lambda_n} > (g + \frac{g_0}{\lambda_n})^2$ for all $n = 2, \ldots, N$ then $\frac{\mathrm{d}s_n}{\mathrm{d}c}\big|_{c=0} < 0$, and each s_n is minimized by the positive root of the quadratic function above. In this case, $c_n^\star > 0$. Therefore, $\frac{\mathrm{d}}{\mathrm{d}c}V_N^{\mathrm{DAPI}}\big|_{c=0} < 0$ and some $c^\star > 0$ minimizes V_N^{DAPI}.

Finally, if $\frac{f}{\lambda_n} > (g + \frac{g_0}{\lambda_n})^2$ for some, but not all n, the existence of a positive minimizer $c^\star > 0$ depends on remaining system parameters. ■

Chapter 5

Inadmissibility of Localized High-Order Consensus

The main focus of this thesis is to characterize performance limitations in networked systems where the local dynamics are of order one or two. These are also the systems that most of the consensus literature has focused on. However, higher-order algorithms, which we will look into in this chapter, have also received attention, for example in [81, 113, 138, 142–144, 201]. Here, the local dynamics of each agent is modeled as an n^{th} order integrator ($n \geq 3$), and the control signal – the consensus algorithm – is a weighted sum of relative differences between states of neighboring agents. This can be viewed as an important theoretical generalization of the first- and second-order algorithms [81], but also has practical relevance. For example, not only position and velocity, but also acceleration feedback play a role in flocking behaviors, leading to a model where $n = 3$ [142]. Systems that can be re-cast as the high-order consensus algorithm can also be constructed by applying dynamic feedback control to a first- or second-order system. This case is of particular importance for this thesis, and will be revisited in Chapter 6.

Throughout the thesis, we are interested in the scalability of control algorithms to large networks. In remaining chapters, this problem is posed as one of *performance*, subject to a scaling of the network. We show in this chapter that the problem in high-order consensus is more fundamental: can *stability* be maintained as the network grows?

The first result we present here is clear-cut: the high-order ($n \geq 3$) consensus algorithm treated in, for example, [142, 143] can *not* allow the network size to scale in graphs where the algebraic connectivity is decreasing towards zero in network size. We prove that at some finite network size, the closed-loop stability criteria will then inevitably be violated, rendering the consensus algorithm *inadmissible* in our terminology. While this can be avoided if the controller is re-tuned as the network size grows, such a re-tuning would require global knowledge of the network and thus prevent a truly distributed implementation.

The algebraic connectivity, that is, the smallest non-zero eigenvalue of the weighted network graph Laplacian, decreases towards zero in classes of graphs where the interactions are *localized*, in that the size of each agent's neighborhood is bounded. In leader-follower consensus over undirected graphs, the locality property alone is sufficient to cause inadmissibility. This latter result was shown by Yadlapalli *et al.* in [191] using a different method than the one we present here. We generalize their result to leaderless consensus and directed, weighted graphs and describe an issue that, to the best of our knowledge, has otherwise not been observed in literature.

We also discuss the standard second-order consensus algorithm in systems over toric lattice graphs. Here, we show that localized consensus will render the system unstable as the network grows if the graph is *directed*. That is, if the feedback interactions are not symmetric around each node. This points to an important fragility in second-order consensus in these networks; stability can only be guaranteed for any network size if feedback interactions are perfectly symmetric. We remark that while this result has not been observed in the wider consensus or formation control literature, it recently appeared in [26] in the context of flocking.

In this chapter, we give some numerical examples to illustrate the results, and also discuss their consequence for distributed dynamic feedback control that will be treated in the following chapters.

5.1 The n^{th}-order consensus problem

We begin by introducing the modeling framework for the n^{th}-order consensus algorithm. The algorithm we consider adheres to the ones considered in [113, 142–144] and is a straightforward extension to the standard first- and second-order consensus algorithms introduced in Chapter 1.

5.1.1 Network model and definitions

Consider a network of N agents modeled by the graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, 2, \ldots, N\}$ is the set of nodes and $\mathcal{E} = \{(i, j)\}$ is the set of edges, each with an associated weight $w_{ij} > 0$. We will in general let $\mathcal{G}$ be directed, but will assume throughout that the weighted graph Laplacian $\mathcal{L}$ of $\mathcal{G}$ is *normal* (see Section 3.2.2). We also assume that $\mathcal{G}$ has a connected spanning tree and number the eigenvalues of $\mathcal{L}$ so that

$$0 = \lambda_1 < \text{Re}\{\lambda_2\} \leq \ldots \leq \text{Re}\{\lambda_N\}.$$

All notation and terminology used in this chapter adhere to the definitions in Section 3.2.

5.1.2 n^{th} order consensus algorithm

The local dynamics at each node $i \in \mathcal{V}$ is modeled as a chain of n integrators:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}x_i^{(0)}(t) &= x_i^{(1)}(t)\\ &\vdots\\ \frac{\mathrm{d}}{\mathrm{d}t}x_i^{(n-2)}(t) &= x_i^{(n-1)}(t)\\ \frac{\mathrm{d}}{\mathrm{d}t}x_i^{(n-1)}(t) &= u_i(t),\end{aligned}$$

where we let the information state $x_i(t) \in \mathbb{R}$ (see Remark 5.1). The notation for time derivatives is such that $x_i^{(0)}(t) = x_i(t)$, $x_i^{(1)}(t) = \frac{\mathrm{d}}{\mathrm{d}t}x_i(t) = \dot{x}_i(t)$ etc. until $x_i^{(n)}(t) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}x_i(t)$. Going forward, we will drop the time dependence in the notation.

We consider the following n^{th} order consensus algorithm:

$$u_i = -\sum_{k=0}^{n-1} a_k \sum_{j\in\mathcal{N}_i} w_{ij}(x_i^{(k)} - x_j^{(k)}), \tag{5.1}$$

where the a_k are nonnegative fixed gains. All feedback is assumed relative throughout this chapter. We will discuss the impact of absolute feedback in Section 5.5.

Defining the full state vector $\xi = [x^{(0)}, x^{(1)}, \ldots, x^{(n-1)}]^T$, we can write the system's closed-loop dynamics as

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi = \underbrace{\begin{bmatrix} 0 & I_N & 0 & \cdots & 0\\ 0 & 0 & I_N & \cdots & \vdots\\ 0 & 0 & 0 & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & I_N\\ -a_0\mathcal{L} & -a_1\mathcal{L} & -a_2\mathcal{L} & \cdots & -a_{\mathrm{n-1}}\mathcal{L}\end{bmatrix}}_{\mathcal{A}}\xi, \tag{5.2}$$

where the graph Laplacian $\mathcal{L}$ was defined in (3.16) and I_N denotes the $N \times N$ identity matrix.

Remark 5.1 We limit the analysis to a scalar information state, though an extension to $x_i(t) \in \mathbb{R}^m$ is straightforward provided the system is controllable in the m coordinate directions. In this case, the system dynamics can be written $\dot{\xi} = (\mathcal{A} \otimes I_m)\xi$, where $\otimes$ denotes the Kronecker product. This would not affect the chapter's main result concerning the stability of $\mathcal{A}$.

Remark 5.2 This chapter focuses on the stability of $\mathcal{A}$, rather than the system's performance under a disturbance input. Therefore, we need not consider an exogenous input, nor define a performance output. □

Leader-follower consensus

We may also consider leader-follower consensus as in [191]. Here, the state of agent 1 is assumed fixed, meaning that it acts as a leader for remaining agents (under the assumption that there is a directed path to each of them from agent 1). Without loss of generality we can then set $x_1 = \dot{x}_1 = \ldots, x_1^n \equiv 0$. The closed-loop dynamics for remaining agents can be written

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\xi} = \underbrace{\begin{bmatrix} 0 & I_{N-1} & 0 & \cdots & 0 \\ 0 & 0 & I_{N-1} & \cdots & \vdots \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_{N-1} \\ -a_0\bar{\mathcal{L}} & -a_1\bar{\mathcal{L}} & -a_2\bar{\mathcal{L}} & \cdots & -a_{n-1}\bar{\mathcal{L}} \end{bmatrix}}_{\bar{\mathcal{A}}} \bar{\xi}, \tag{5.3}$$

where $\bar{\mathcal{L}}$ is the *grounded* graph Laplacian obtained by deleting the first row and column of $\mathcal{L}$ and $\bar{\xi}$ is obtained by removing the states of agent 1. Note that $\bar{\mathcal{L}}$ unlike $\mathcal{L}$ has all of its eigenvalues in the right half plane [189].

5.1.3 Conditions for consensus and admissibility

Consensus among the agents is said to be reached if $x_i^{(k)} \to x_j^{(k)}$ for all $i, j \in \mathcal{V}$ and for $k = 0, 1, \ldots, n-1$. It is known that the algorithm (5.1) achieves consensus if the eigenvalues of $\mathcal{A}$ are in the left half plane, apart from exactly n zero eigenvalues that are associated with the drift of the network average. This condition is in line with standard results for first- and second-order consensus, and is shown in [143] for $n = 3$:

Theorem 5.1 ([143], Theorem 3.1) *In the case of $n = 3$, the algorithm (5.1) achieves consensus exponentially if and only if $\mathcal{A}$ has exactly three zero eigenvalues and all of the other eigenvalues have negative real parts.*

We also require the following lemma:

Lemma 5.2 ([143], Lemma 3.1) *In the case of $n = 3$, the matrix $\mathcal{A}$ has exactly three zero eigenvalues if and only if $\mathcal{L}$ has a simple zero eigenvalue.*

The proofs in [143] can be straightforwardly extended to $n > 3$.

This means that it is sufficient to verify that the $(N-1)\cdot n$ non-zero eigenvalues of $\mathcal{A}$ have negative real parts. Here, we will be dealing with systems where this is true for small network sizes N, but where one or more eigenvalues leave the left half plane and cause instability as the network grows beyond some network size $\bar{N}$. In these cases, we call the control algorithm *inadmissible*.

Definition 5.1 (Admissibility) A control design u is *admissible* if the resulting closed-loop system reaches consensus for *any* finite network size N.

5.2 Inadmissibility of high-order consensus

This section is devoted to the chapter's main result. We first describe the key underlying assumptions, before proving that the high-order consensus algorithm will be inadmissible if the network graph has what we term a *decreasing algebraic connectivity*. This property applies to several classes of graphs, and we end this section by listing a few of them.

5.2.1 Underlying assumptions

The following assumptions are important for the upcoming analysis.

Assumption 5.1 (Locality) The feedback is *localized*, meaning that the controller uses measurements only from a neighborhood of size at most q, where q is fixed and independent of N. That is,

$$|\mathcal{N}_i| \leq q \quad \forall i \in \mathcal{V}. \tag{5.4}$$

Assumption 5.2 (Finite weights and gains) The system gains and edge weights are finite, that is, $w_{ij} \leq w_{\max} < \infty$ for all $(i,j) \in \mathcal{E}$ and $a_k \leq a_{\max} < \infty$ for all $k = 0, 1, \ldots, n$.

Assumption 5.3 (Fixed parameters) The gains a_k for all $k = 0, 1, \ldots, n$, the maximum edge weight $w_{\max}$, and the locality parameter q do not change if a node (with connecting edges) is added to the graph $\mathcal{G}$. That is, these parameters are all independent of the total network size N.

In the following, the notion of an increase in the network size N should be understood as the addition of nodes to the network (along with connecting edges) in such a manner that Assumptions 5.1–5.3 remain satisfied. These assumptions contribute to the key property; that the algebraic connectivity of $\mathcal{G}$ decreases towards zero. This is clarified through examples in Section 5.2.3.

5.2.2 Main result

The chapter's main result is negative, and states that if $n \geq 3$, a consensus algorithm on the form (5.1) can never be admissible in certain graphs.

Theorem 5.3 *If $n \geq 3$, no control on the form* (5.1) *is admissible under Assumptions 5.1–5.3 if the graph $\mathcal{G}$ is such that* $\mathrm{Re}\{\lambda_2\} \to 0$ *as* $N \to \infty$.

Proof: The first step of the proof is to block-diagonalize the matrix $\mathcal{A}$. Let U be the unitary matrix that diagonalizes the graph Laplacian $\mathcal{L}$, so that $U^*\mathcal{L}U = \Lambda = \mathrm{diag}\{0, \lambda_2, \ldots, \lambda_N\}$. By pre- and post-multiplying $\mathcal{A}$ by the $(Nn \times Nn)$ matrix

$\mathcal{U} = \text{diag}\{U, U, \ldots, U\}$, we get

$$\mathcal{U}^*\mathcal{A}\mathcal{U} = \underbrace{\begin{bmatrix} 0 & I_N & 0 & \cdots & 0 \\ 0 & 0 & I_N & \cdots & \vdots \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_N \\ -a_0\Lambda & -a_1\Lambda & -a_2\Lambda & \cdots & -a_{\mathrm{n}-1}\Lambda \end{bmatrix}}_{\hat{\mathcal{A}}}. \tag{5.5}$$

This can be re-arranged into N decoupled submatrices $\hat{\mathcal{A}}_l$:

$$\hat{\mathcal{A}}_l = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \vdots \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0\lambda_l & -a_1\lambda_1 & -a_2\lambda_l & \cdots & -a_{\mathrm{n}-1}\lambda_l \end{bmatrix},$$

for $l = 1, \ldots, N$. The eigenvalues of $\mathcal{A}$ are the union of the eigenvalues of all $\hat{\mathcal{A}}_l$. Clearly, the n zero eigenvalues are obtained from $\hat{\mathcal{A}}_1$ since $\lambda_1 = 0$. Therefore, we must require all eigenvalues of all $\hat{\mathcal{A}}_l$, $l = 2, \ldots, N$ to have negative real parts for any N to ensure admissibility.

The characteristic polynomial of each $\hat{\mathcal{A}}_l$ is

$$p_l(s) = s^n + a_{n-1}\lambda_l s^{n-1} + \ldots + a_1\lambda_l s + a_0\lambda_l. \tag{5.6}$$

In general, the eigenvalues λ_l are complex-valued. Consider therefore the Routh-Hurwitz criteria for polynomials with complex coefficients. As these criteria do not appear frequently in literature, we include a detailed derivation here.

Consider the polynomial

$$p(\mu) = \mu^n + (f_{n-1} + \mathbf{j}g_{n-1})\mu^{n-1} + \ldots (f_0 + \mathbf{j}g_0) = 0, \tag{5.7}$$

where $\mathbf{j} = \sqrt{-1}$ denotes the imaginary number. The roots μ will be such that $\text{Im}(\mu) > 0$ if and only if all inequalities

$$-\Delta_2 = -\begin{vmatrix} 1 & f_{n-1} \\ 0 & g_{n-1} \end{vmatrix} > 0, \quad \Delta_4 = \begin{vmatrix} 1 & f_{n-1} & f_{n-2} & f_{n-3} \\ 0 & g_{n-1} & g_{n-2} & g_{n-3} \\ 0 & 1 & f_{n-1} & f_{n-2} \\ 0 & 0 & g_{n-1} & g_{n-2} \end{vmatrix} > 0, \ \cdots,$$

$$(-1)^n\Delta_{2n} = (-1)^n \begin{vmatrix} 1 & f_{n-1} & \cdots & f_0 & 0 & \cdots & \cdots & 0 \\ 0 & g_{n-1} & \cdots & g_0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & f_1 & f_0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & g_1 & g_0 & 0 & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & \cdots & \cdots & 0 & 1 & \cdots & f_1 & f_0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & g_1 & g_0 \end{vmatrix} > 0 \tag{5.8}$$

are satisfied [177, pp 21f]. Evaluating the determinants, the first two conditions become

$$g_{n-1} < 0, \tag{5.9}$$

$$f_{n-1}g_{n-1}g_{n-2} - f_{n-2}g_{n-1}^2 + g_{n-3}g_{n-1} - g_{n-2}^2 > 0. \tag{5.10}$$

We are interested in the polynomial $p_l(s)$ in (5.6) and seek a condition for $\mathrm{Re}\{s\} < 0$ for all $l = 2, \ldots, N$. Therefore, we substitute $\mu = -\mathbf{j}s$ in (5.7) and identify the coefficients from (5.6). The coefficients that appear in (5.9)–(5.10) are then

$$\begin{aligned} f_{n-1} &= a_{n-1}\mathrm{Im}\{\lambda_l\}, & g_{n-1} &= -a_{n-1}\mathrm{Re}\{\lambda_l\}, \\ f_{n-2} &= -a_{n-2}\mathrm{Re}\{\lambda_l\}, & g_{n-2} &= -a_{n-2}\mathrm{Im}\{\lambda_l\}, \\ f_{n-3} &= -a_{n-3}\mathrm{Im}\{\lambda_l\}, & g_{n-3} &= a_{n-3}\mathrm{Re}\{\lambda_l\}. \end{aligned} \tag{5.11}$$

Note that these relations hold regardless of n. Now, the condition (5.9) reads

$$a_{n-1}\mathrm{Re}\{\lambda_l\} > 0,$$

which is always true for $l = 2, \ldots, N$ if $a_{n-1} > 0$ since $\mathrm{Re}\{\lambda_l\} > 0$. The second condition (5.10) can after some manipulation be written as

$$\begin{aligned} a_{n-1}(\mathrm{Re}\{\lambda_l\})^2(a_{n-1}a_{n-2}\mathrm{Re}\{\lambda_l\} - a_{n-3}) + \\ + a_{n-2}(\mathrm{Im}\{\lambda_l\})^2(a_{n-1}^2\mathrm{Re}\{\lambda_l\} - a_{n-2}) > 0, \end{aligned} \tag{5.12}$$

for all $l = 2, \ldots, N$. While the factors in front of the brackets remain positive for all λ_l (provided $a_k > 0$), the brackets will eventually become negative if $\mathrm{Re}\{\lambda_l\} \to 0$ for some l. Thus, if $\mathrm{Re}\{\lambda_2\}$, where λ_2 is the eigenvalue with smallest real part, is decreasing in N towards zero, the condition (5.12) will eventually be violated.

This implies that if $n > 3$, at least one root of the characteristic polynomial $p_2(s)$ will not have a negative real part for any N. Therefore, it is not true that all eigenvalues of $\mathcal{A}$ except the n zero eigenvalues have negative real parts (as required by Theorem 5.1). The control is therefore not admissible. ■

Remark 5.3 If the graph is undirected, then the polynomial (5.6) has real-valued coefficients. The result can then be derived using the standard Routh-Hurwitz criteria. This gives the simpler condition

$$a_{n-1}a_{n-2}\lambda_2 - a_{n-3} > 0, \tag{5.13}$$

which can never remain satisfied if $\lambda_2 \to 0$ as $N \to \infty$. □

Remark 5.4 The condition that $\mathcal{L}$ be normal can be relaxed if $\mathcal{L}$ is diagonalizable as in (5.5) by some (non-unitary) matrix. The remainder of the proof would hold true. □

Theorem 5.3 implies that the high-order consensus algorithm can never allow the network size of certain graphs to increase indefinitely without leading to instability. Instability will occur at the smallest N for which the Routh-Hurwitz criteria (5.8) in the proof are not satisfied, and at least one eigenvalue leaves the open left half plane. We will term this critical network size $\bar{N}$. In Figure 5.1 we display $\bar{N}$ for $n = 3, 4, 5$ in a path graph.

Inadmissibility of high-order leader-follower consensus

High-order leader-follower consensus on the form (5.3) in undirected networks will always be inadmissible under the given assumptions. This was also observed in [191], and before that in [17] for a path graph topology. We first require the following Lemma:

Lemma 5.4 *Consider the grounded Laplacian matrix $\bar{\mathcal{L}}$ of an undirected graph $\mathcal{G}$. Let the relation* (5.4) *hold. The smallest eigenvalue $\bar{\lambda}_1$ of $\bar{\mathcal{L}}$ then satisfies*

$$\bar{\lambda}_1 \leq \frac{q}{N-1} w_{\max}, \tag{5.14}$$

where $w_{\max}$ is the largest edge weight in $\mathcal{E}$.

Proof: By the Rayleigh-Ritz theorem [73, Theorem 4.2.2] it holds

$$\bar{\lambda}_1 \leq \frac{v^T \bar{\mathcal{L}} v}{v^T v}, \quad \forall v \in \mathbb{C}^{N-1} \backslash \{0\}.$$

This implies in particular that

$$\bar{\lambda}_1 \leq \frac{\mathbf{1}_{N-1}^T \bar{\mathcal{L}} \mathbf{1}_{N-1}}{\mathbf{1}_{N-1}^T \mathbf{1}_{N-1}} = \frac{\sum_{k \in \mathcal{N}_1} w_{1k}}{N-1} \leq \frac{q w_{\max}}{N-1},$$

where $\mathbf{1}_{N-1}^T \bar{\mathcal{L}} \mathbf{1}_{N-1} = \sum_{k \in \mathcal{N}_1} w_{1k}$ is the weighted sum of all edges leading to the leader node 1. The equality holds since each row k of the grounded Laplacian $\bar{\mathcal{L}}$ sums to zero if the corresponding node k has no connection to the leader, and otherwise to $w_{1k} \leq w_{\max}$. ■

Theorem 5.5 *Let the graph $\mathcal{G}$ be undirected. The leader-follower consensus algorithm* (5.3) *is then inadmissible for $n \geq 3$ under Assumptions 5.1–5.3.*

Proof: The arguments in the proof of Theorem 5.3 apply. In this case, $N-1$ real-valued characteristic polynomials like (5.6) are obtained. The condition (5.12) reduces to

$$a_{n-1}a_{n-2}\bar{\lambda}_l - a_{n-3} > 0, \tag{5.15}$$

for $l = 1, \ldots, N-1$. Using Lemma 5.4, we see that (5.15) requires $a_{n-1}a_{n-2} > \frac{1}{qw_{\max}}a_{n-3}(N-1)$, which cannot be satisfied under Assumption 5.3. The algorithm is thus inadmissible. ■

5.2.3 Affected classes of graphs

The inadmissibility of high-order consensus applies to any network whose underlying graph is such that $\mathrm{Re}\{\lambda_2\}$ is decreasing towards zero as N increases. The second-smallest Laplacian eigenvalue λ_2 of an undirected graph is real-valued and known as the *algebraic connectivity* of the graph. While the correct generalization of algebraic connectivity to directed graphs is not clear-cut, we know the following:

Lemma 5.6 *If $\mathcal{L}$ is normal then*

$$\mathrm{Re}\{\lambda_2\} = \lambda_2^s,$$

where λ_2^s is the smallest non-zero eigenvalue of $\mathcal{L}^s = (\mathcal{L} + \mathcal{L}^T)/2$, that is, the symmetric part of $\mathcal{L}$.

Proof: See Lemma 9.1.2 in [63]. ■

The matrix $\mathcal{L}^s$ is the graph Laplacian corresponding to the *mirror graph* $\hat{\mathcal{G}}$ of $\mathcal{G}$, which is the *undirected* graph obtained as $\hat{\mathcal{G}} = \{\mathcal{V}, \mathcal{E} \cup \hat{\mathcal{E}}\}$, where $\hat{\mathcal{E}}$ is the set of all edges in $\mathcal{E}$, but *reversed*, and whose edge weights are $\hat{w}_{ij} = \hat{w}_{ij} = (w_{ij}+w_{ji})/2$ [118]. Clearly, the mirror graph of an undirected graph is the graph itself. Lemma 5.6 implies that $\mathrm{Re}\{\lambda_2\}$ of $\mathcal{G}$ is the algebraic connectivity of its mirror graph $\hat{\mathcal{G}}$. We introduce the following terminology:

Definition 5.2 The graph $\mathcal{G}$ is said to have *decreasing algebraic connectivity* if, for its mirror graph $\hat{\mathcal{G}}$, the algebraic connectivity $\lambda_2 \to 0$ as $N \to \infty$.

This means that Theorem 5.3 will apply to graphs with decreasing algebraic connectivity, and it suffices to identify this property in undirected graphs. We next give a (non-exhaustive) account of classes of graphs with this property.

Lattices, fuzzes, and their embedded graphs

Consider the d-dimensional toric lattice $\mathbb{Z}_L^d$ with $N = L^d$ nodes, and let each node be connected to its r neighbors in each direction (letting $q = 2rd$). Such a lattice is called an r-fuzz.

Lemma 5.7 (Algebraic connectivity of r-fuzz) *For an r-fuzz lattice of d dimensions,*

$$\lambda_2 \leq \frac{c}{N^{2/d}}, \tag{5.16}$$

where c is a constant that depends on the fixed parameters r (that is, q), $w_{\max}$ and d, but not on N.

Proof: Follows from the upcoming Lemma 6.8. There, the edge weights are assumed uniform, but the result (5.16) follows from setting all edge weights to $w_{\max}$ and then considering Lemma 5.8. ■

The bound (5.16) also holds for any subgraph of the r-fuzz lattice, that is, any graph that is *embeddable* in the lattice. This follows from the following important lemma:

Lemma 5.8 *Adding an edge to a graph or increasing the weight of an edge increases (or leaves unchanged) λ_2 of the corresponding graph Laplacian, and vice versa.*

Proof: Follows from Theorems 3.1 and 3.3. ■

Planar graphs

Planar graphs are embeddable in two-dimensional lattices so Lemma 5.7 applies. For this important case, however, a more precise bound is available:

Lemma 5.9 (Algebraic connectivity of planar graphs) *For a planar graph,*

$$\lambda_2 \leq \frac{8qw_{\max}}{N}, \tag{5.17}$$

Proof: See [167, Theorem 6]. ■

Constant-genus graphs

Planar graphs can be generalized to graphs with constant *genus*. The genus of a planar graph is $g = 0$. Higher genus implies that the graph can be drawn on a surface with g handles (or "holes") without any one edge crossing another. For example, a torus would correspond to $g = 1$ and a pretzel shape to $g = 3$.

Lemma 5.10 (Algebraic connectivity of constant-genus graphs) *Let $\mathcal{G}$ have constant and bounded genus g. Then*

$$\lambda_2 \leq \frac{c_2}{N},$$

where c_2 is a constant that depends on q, g and $w_{\max}$, but not on N.

Proof: See [84, Theorem 2.3]. ■

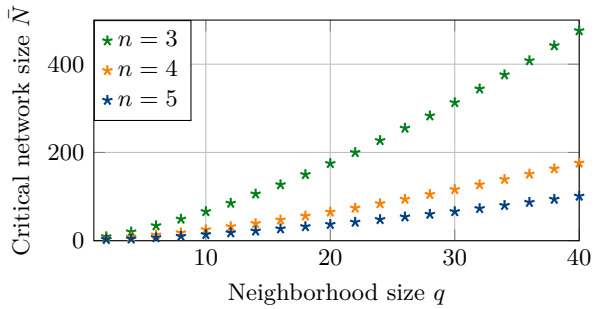


Figure 5.1: Critical network size $\bar{N}$ at which the stability conditions are violated for an n^{th} order consensus algorithm. The graph is an undirected path graph where each node is connected to its q nearest neighbors. Increasing the neighborhood size q here increases $\bar{N}$ faster than linearly. Also note that for higher n, the stability conditions are violated at smaller $\bar{N}$.

Tree graphs with growing diameter

The *diameter* $\operatorname{diam}\{\mathcal{G}\}$ of the graph $\mathcal{G}$ is defined as the longest distance between any two nodes in the graph. If we let $\mathcal{G}$ be a *tree graph*, then, by [60, Corollary 4.4] it holds $\lambda_2 \leq 2w_{\max}\left(1-\cos\left(\frac{\pi}{\operatorname{diam}(\mathcal{G})+1}\right)\right)$. This allows us to derive the following lemma.

Lemma 5.11 (Algebraic connectivity of tree graphs) *Let $\mathcal{G}$ be a tree graph. Then*

$$\lambda_2 \leq \frac{\pi^2 w_{\max}}{(\operatorname{diam}\{\mathcal{G}\}+1)^2}, \tag{5.18}$$

and if $\operatorname{diam}\{\mathcal{G}\} \to \infty$ *as* $N \to \infty$, *then* $\lambda_2 \to 0$.

Proof: Follows from the relation above, noting that $1-\cos x \leq \frac{x^2}{2}$ for any x. Clearly, the right hand side is decreasing in $\operatorname{diam}\{\mathcal{G}\}$. ■

5.3 Numerical examples

In this section, we present two simple numerical examples to illustrate this chapter's main result.

Example 5.1 (Locality and critical network size) Assumption 5.1 of locality, that is, a fixed upper bound on the size of each agent's neighborhood, is key for this chapter's main result. Indeed, if each agent's neighborhood were allowed to grow as more and more agents are added to the network, the high-order consensus algorithm could stay admissible.

As an example, consider an undirected path graph where each node is connected to its $q/2$ neighbors in each direction (q assumed even). If all edge weights are 1, the algebraic connectivity of this graph is

$$\lambda_2 = \sum_{k=1}^{q/2} 2\left(1 - \cos\left(\frac{\pi k}{N}\right)\right),$$

which for any given N is larger, the greater q is. Increasing q thus delays the violation of the stability criteria as N grows.

In Figure 5.1, we depict the critical network size $\bar{N}$ as a function of the neighborhood size q in the undirected path graph. Here, we have selected a consensus algorithm where $a_0 = 0.1$, $a_1 = 0.8$, $a_2, a_3, a_4 = 1$. The plot shows that increasing q pushes the critical network size upwards faster than linearly.

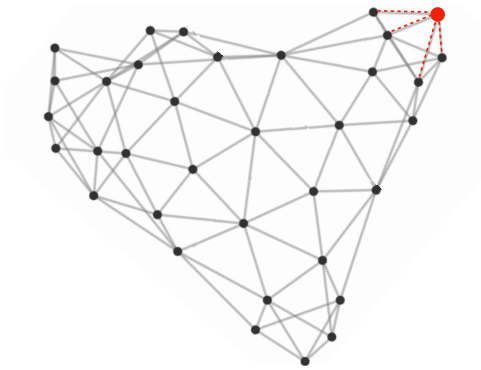
We also note that the system becomes unstable at smaller $\bar{N}$ for higher n. This is because the higher-order conditions in (5.8) are violated sooner than the lower-order ones (though only one condition, $\Delta_4 > 0$, was needed to prove inadmissibility in Theorem 5.3).

Example 5.2 (Instability through node addition) The second example illustrates the phase transition – from consensus to instability – that the system experiences as the critical network size is reached. Figure 5.2a illustrates a graph that has been randomly generated by means of triangulation. Here, the maximum neighborhood size is $q = 8$ while the median is 5. All edge weights are set to 1.

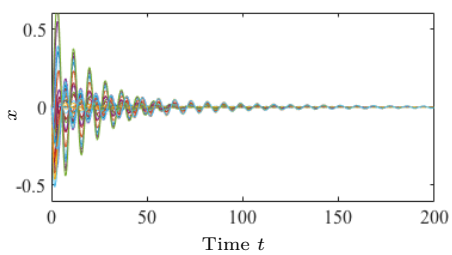
We consider a third order consensus algorithm:

$$x_i^{(3)} = -\sum_{j\in\mathcal{N}_i} \left[0.5(x_i - x_j) + (\dot{x}_i - \dot{x}_j) + (\ddot{x}_i - \ddot{x}_j)\right],$$

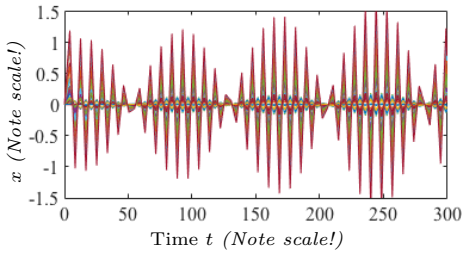
which by the condition (5.13) ensures stability if $\lambda_2 > 0.5$. With 34 nodes, the graph in Figure 5.2a has $\lambda_2^{34} = 0.536$ and the system achieves consensus, as seen from the simulation in Figure 5.2b. We then add a 35^{th} node along with 4 connecting edges, as indicated in red color in the graph in Figure 5.2a. Now, $\lambda_2^{35} = 0.493$ and the system becomes unstable. Figure 5.2c shows how the agents' positions x oscillate at an increasing amplitude.



(a) Network graph



(b) $N = 34$



(c) $N = 35$

Figure 5.2: Simulation of 3rd order consensus in graph depicted in (a), subject to random initial accelerations. In (b) the network's 34 agents converge to an equilibrium. In (c) a 35th node has been added, indicated by red color in the graph. This addition leads to instability. The plots (b) and (c) show position trajectories relative to Agent no. 1. Note the different scales.

5.4 Inadmissibility of second-order consensus in directed lattices

We now turn the attention to second-order integrator networks ($n = 2$). These are particularly important as they are the standard model for, for example, vehicular formation networks. For these systems, it is also easy to guarantee admissibility if the underlying graph is undirected. However, we show here that in the case of a directed, balanced, periodic lattice network graph, the standard consensus algorithm is in fact inadmissible under the locality constraint.

To formalize the result, let us assume in this section that the network graph is the d-dimensional toric lattice of period L, $\mathbb{Z}_L^d$. In Chapter 6 we will express feedback operations on $\mathbb{Z}_L^d$ using the operator formalism from Section 3.3. To simplify the reading of this Chapter, however, we adhere to the graph Laplacian-based notation used thus far.

Therefore, consider first the case $d = 1$ and number the nodes as $1, 2, \ldots, L$.

The control algorithm is assumed to be identical at each node i in the network:

$$u_i = -a_0 \sum_{j=-r}^{r} w_j(x_i - x_{i+j}) - a_1 \sum_{j=-r}^{r} w_j(\dot{x}_i - \dot{x}_{i+j}), \tag{5.19}$$

where r is a locality parameter such that $q = 2rd$. Enforcing the periodicity of the lattice implies, for example, that $x_{1-1} = x_N$ and ensures normality of the graph Laplacian. This way, the graph Laplacian $\mathcal{L}$ is a *circulant matrix* with the weight w_j on the j^{th} diagonal. In the d-dimensional case, the Laplacian is the Cartesian product of d such circulant matrices. The following assumption is important.

Assumption 5.4 (Asymmetric consensus) The feedback around each node $i \in \mathbb{Z}_L^d$ is *asymmetric*, meaning that $w_{-j} \neq w_j$ in (5.19) for at least one $j \in \{1, \ldots, r\}$.

Now, consider the following theorem.

Theorem 5.12 *Let $n \geq 2$ and let Assumptions 5.1–5.4 hold. Then, the consensus algorithm* (5.19) *is inadmissible.*

Proof: We proceed as in the proof of Theorem 5.3 to obtain the characteristic polynomials

$$p_l(s) = s^2 + a_1\lambda_l s + a_0\lambda_l, \tag{5.20}$$

for $l = 1, \ldots, N$, and the Routh-Hurwitz criterion derived from $\Delta_4 > 0$ in (5.8) becomes

$$a_0 a_1 \mathrm{Re}\{\lambda_l\}[a_1(\mathrm{Im}\{\lambda_l\})^2 + a_0\mathrm{Re}\{\lambda_l\}] - a_0(\mathrm{Im}\{\lambda_l\})^2 > 0, \tag{5.21}$$

which must be satisfied for $l = 2, \ldots, N$ to grant admissibility. The smallest eigenvalue is:

$$\lambda_2 = \sum_{\substack{i=-r \\ i\neq 0}}^{r} w_i(1 - \cos\left(\frac{2\pi i}{L}\right)) - \mathbf{j} \sum_{\substack{i=-r \\ j\neq 0}}^{r} w_i \sin\left(\frac{2\pi i}{L}\right),$$

where, recall, $L = N^{1/d}$ is the lattice size. A series expansion of, respectively, the real part and the squared imaginary part of λ_2 around zero reveals that both $\mathrm{Re}\{\lambda_2\} \sim 1/L^2$ and $(\mathrm{Im}\{\lambda_2\})^2 \sim 1/L^2$ (the notation $\sim$ is defined in (3.26)). See Chapter 6 and Appendix 6.F for further details to these bounds. In (5.21) this means that the first term $a_0a_1\mathrm{Re}\{\lambda_2\}(a_1(\mathrm{Im}\{\lambda_2\})^2 + a_0\mathrm{Re}\{\lambda_2\}) \sim 1/L^4$, while the second term $-a_0(\mathrm{Im}\{\lambda_l\})^2 \sim -1/L^2$. Since $1/L^4$ approaches zero at a faster rate than $1/L^2$ as $L \to \infty$, it means that (5.21) cannot be satisfied for sufficiently large L (that is, sufficiently large N). The algorithm is thus inadmissible. ■

Remark 5.5 By studying the proof of Theorem 5.12 carefully, we note that the second-order consensus algorithm will in fact be inadmissible in any directed graph (with diagonalizable Laplacian) where both $\mathrm{Re}\{\lambda_2\} \to 0$ and $\mathrm{Im}\{\lambda_2\} \to 0$ as $N \to \infty$, but where the *rate* of decrease is *higher* in $\mathrm{Re}\{\lambda_2\}$. A characterization of graphs where this property applies is an open problem. □

This result implies that there is a *fragility* in the second-order consensus algorithm (5.19), in that it can guarantee stability for any network size N if it is symmetric ($w_j = w_{-j}$ for $j = -r, \dots, r$). However, even the slightest asymmetry will eventually cause instability in a periodic lattice if N is allowed to increase. We remark again that the same result has been reported in [26] for nearest-neighbor feedback ($r = 1$).

5.5 Alternative system dynamics

In this section, we discuss systems whose dynamics can not be written on the form (5.2). We begin by discussing the impact of absolute feedback. To simplify the exposition, we limit the analysis here to the case $n = 3$. We then consider implications of this Chapter's results for the possibility to apply distributed dynamic feedback to systems with consensus dynamics of first and second order.

5.5.1 Impact of absolute feedback

The admissibility issues of the consensus algorithm (5.1) would not arise if the control included *absolute* state feedback. In this case, the control algorithm would be

$$u_i = -\sum_{k=0}^{2} \left(a_k \sum_{j \in \mathcal{N}_i} w_{ij}(x_i^{(k)} - x_j^{(k)}) - a_k^{\text{abs}} x^{(k)} \right). \tag{5.22}$$

We say that absolute feedback from the state $x^{(k)}$ is available if one can set $a_k^{\text{abs}} > 0$. Consider the following proposition:

Proposition 5.13 *There is a choice of a_k, a_k^{abs}, for $k = 0, 1, 2$ so that the algorithm (5.22) is admissible for any undirected graph $\mathcal{G}$ if at least one of $a_1^{\text{abs}}, a_2^{\text{abs}} > 0$. If $\mathcal{G}$ is directed, there is such a choice only if $a_2^{\text{abs}} > 0$.*

Proof: In this case, the characteristic polynomial corresponding to (5.6) becomes

$$p_l(s) = s^3 + (a_2\lambda_l + a_2^{\text{abs}})s^2 + (a_1\lambda_l + a_1^{\text{abs}})s + (a_0\lambda_l + a_0^{\text{abs}}) \tag{5.23}$$

Since a_k^{abs} are real-valued, the real parts of the coefficients in the polynomial (5.23) are now $a_k\text{Re}\{\lambda_l\} + a_k^{\text{abs}}$, while the imaginary parts remain $a_k\text{Im}\{\lambda_l\}$. We can therefore substitute $(\text{Re}\{\lambda_l\} + a_k^{\text{abs}}/a_k)$ for $\text{Re}\{\lambda_l\}$ in the relations (5.11) and obtain the relevant stability criterion from (5.10) as:

$$\begin{aligned}(a_2\text{Re}\{\lambda_l\} + a_2^{\text{abs}})\big[a_1a_2\text{Re}\{\lambda_l\}^2 + (a_1^{\text{abs}}a_2 + a_1a_2^{\text{abs}})\text{Re}\{\lambda_l\} + a_1^{\text{abs}}a_2^{\text{abs}} \\ - a_0\text{Re}\{\lambda_l\} - a_0^{\text{abs}}\big] + a_1\text{Im}\{\lambda_l\}^2[a_2^2\text{Re}\{\lambda_l\} + a_2a_2^{\text{abs}} - a_1] > 0,\end{aligned} \tag{5.24}$$

for $l = 2, \dots, N$.

Table 5.1: Summary of cases where absolute feedback can grant admissibility in third-order consensus ($n = 3$). There is absolute feedback from the state $x^{(k)}$ if $a_k^{\text{abs}} > 0$ in the algorithm (5.22). Else, we term the feedback from $x^{(k)}$ relative. Note that availability of absolute feedback as listed here is a necessary but not sufficient condition for admissibility.

Position x	**Velocity** $\dot{x}$	**Acceleration** $\ddot{x}$	**Admissibility**
relative or absolute	relative	relative	No
relative or absolute	absolute	absolute	Yes
relative	absolute	relative	Yes[1]
relative	relative	absolute	Yes
absolute	absolute	relative	No
absolute	relative	absolute	No

[1] in undirected graphs

If *both* a_1^{abs} and a_2^{abs} are non-zero, they can clearly be selected so that both brackets can remain positive even as $\text{Re}\{\lambda_l\} \to 0$ (setting $a_2 a_2^{\text{abs}} > a_1$, $a_1^{\text{abs}} a_2^{\text{abs}} > a_0 \text{Re}\{\lambda_N\} + a_0^{\text{abs}}$ is sufficient).

If $a_2^{\text{abs}} > 0$ while $a_1^{\text{abs}} = 0$, then the first bracket can only remain positive as $\text{Re}\{\lambda_l\} \to 0$ if $a_0^{\text{abs}} = 0$. If so, setting $a_1 a_2^{\text{abs}} > a_0$ and $a_2 a_2^{\text{abs}} > a_1$ is sufficient.

If only $a_1^{\text{abs}} > 0$ while $a_2^{\text{abs}} = 0$, then the first bracket can only remain positive as $\text{Re}\{\lambda_l\} \to 0$ if $a_0^{\text{abs}} = 0$. If so, setting $a_1^{\text{abs}} a_2 > a_0$ is sufficient. The second bracket will, however, become negative as $\text{Re}\{\lambda_l\} \to 0$. Then, if $\text{Im}\{\lambda_2\}^2 \to 0$ at a rate slower than $\text{Re}\{\lambda_2\}^2 \to 0$ as $N \to \infty$, the condition (5.24) is eventually violated. This is, for example, the case for periodic lattice graphs (see the proof of Theorem 5.12). This means that $a_1^{\text{abs}} > 0$ can grant admissibility if $a_0^{\text{abs}} = 0$ and $\text{Im}\{\lambda_2\} = 0$ for $l = 2, \ldots, N$, that is, if $\mathcal{G}$ undirected. ■

This implies that absolute feedback from the high-order terms, that is, velocity or acceleration, is necessary to grant admissibility to the high-order consensus algorithm. Reading the proof in detail further reveals that if absolute measurements exist of position, but only of one out of the remaining two states, then the algorithm will be inadmissible if it includes absolute position feedback. This limitation can likely be relaxed by applying derivative action in the manner described in Chapter 4. In this chapter, however, we limit the analysis to static feedback. The results of this section are summarized in a concise manner in Table 5.1.

5.5.2 Implications for distributed dynamic feedback control

One of this thesis' objectives is to address the question of whether distributed dynamic feedback control can alleviate performance limitations that apply to static feedback control in networks where the local dynamics are of first or second order. The dynamic feedback we consider increases the order of the system dynamics, meaning that not all designs are admissible.

We consider single-state dynamic feedback controllers that can be seen as a generalization of the distributed integral controller (DAPI) from Chapter 4. That is, for the first-order consensus problem ($n = 1$), we introduce the controller state z and obtain the closed-loop dynamics

$$\begin{aligned} \dot{x} &= -a_0\mathcal{L}x + z \\ \dot{z} &= -b_0\mathcal{L}x - (b_z\mathcal{L} + b_z^{\text{abs}})z, \end{aligned} \tag{5.25}$$

where b_0, b_z are nonnegative gains. Note that we also allow absolute feedback, $b_z^{\text{abs}} \geq 0$ in the controller state z. Note, however that if $b_z^{\text{abs}} > 0$, the controller state is stabilized and no integral action is eventually exerted on the system. For the second-order problem ($n = 2$), we obtain:

$$\begin{aligned} \ddot{x} &= -a_0\mathcal{L}x - a_1\mathcal{L}\dot{x} + z \\ \dot{z} &= -b_0\mathcal{L}x - b_1\mathcal{L}\dot{x} - (b_z\mathcal{L} + b_z^{\text{abs}})z. \end{aligned} \tag{5.26}$$

We assume that $a_0, a_1 > 0$, which is necessary for admissibility if $z = 0$.

Through a linear state transformation, the systems (5.25)–(5.26) can be cast onto the form of (5.2). In the dynamically controlled first-order case, the transformed system matrix becomes:

$$\mathcal{A} = \begin{bmatrix} 0 & I_N \\ -b_0\mathcal{L} - a_0\mathcal{L}(b_z\mathcal{L} + b_z^{\text{abs}}I) & -a_0\mathcal{L} - (b_z\mathcal{L} + b_z^{\text{abs}}I) \end{bmatrix},$$

and in the second-order case:

$$\mathcal{A} = \begin{bmatrix} 0 & I_N & 0 \\ 0 & 0 & I_N \\ -b_0\mathcal{L} - a_0\mathcal{L}(b_z\mathcal{L} + b_z^{\text{abs}}I) & -(a_0+b_1)\mathcal{L} - a_1\mathcal{L}(b_z\mathcal{L} + b_z^{\text{abs}}I) & -a_1\mathcal{L} - (b_z\mathcal{L} + b_z^{\text{abs}}I) \end{bmatrix}.$$

The analysis in Sections 5.2–5.4 can now be applied to these systems. To begin with, the proof of our main result in Theorem 5.3 reveals the following corollary:

Corollary 5.14 *Let the graph $\mathcal{G}$ be undirected and let Assumptions 5.1–5.3 hold. If $\mathcal{G}$ is such that $\lambda_2 \to 0$ as $N \to \infty$, a necessary condition for admissibility of the dynamic feedback protocol* (5.26) *is that at least one of the following conditions holds:*

a) $b_0 = 0$,

b) $b_z^{\text{abs}} > 0$.

Proof: Since $\mathcal{G}$ is assumed undirected, the eigenvalues λ_l are real-valued. The Routh-Hurwitz criteria for the system (5.26) require that: $b_0 + a_0(b_z\lambda_l + b_z^{\text{abs}}) > 0$, for $l = 2, \ldots, N$, which is satisfied if we do *not* have $b_0 = 0$ and $b_z = 0$, and $b_z^{\text{abs}} = 0$ simultaneously, and that

$$(a_1\lambda_l + b_z\lambda_l + b_z^{\text{abs}})(a_0 + b_1 + a_1(b_z\lambda_l + b_z^{\text{abs}})) - b_0 - a_0(b_z\lambda_l + b_z^{\text{abs}}) > 0. \quad (5.27)$$

If $\lambda_2 \to 0$, then (5.27) will eventually be violated unless $b_z^{\text{abs}} > 0$ or $b_0 = 0$, or if both of these conditions hold. ■

Corollary 5.14 implies that relative position measurements cannot be included in the dynamic feedback, unless the controller state is stabilized through absolute feedback.

When it comes to the dynamically controlled first-order consensus problem (5.25), its admissibility is limited by Theorem 5.12. We therefore again consider the system over the toric lattice $\mathbb{Z}_L^d$, so that $\mathcal{L}$ is circulant, and provide the following corollary:

Corollary 5.15 *Let Assumptions 5.1–5.4 hold, so that the graph Laplacian $\mathcal{L}$ is an asymmetric, circulant matrix. Then, a necessary condition for admissibility of the dynamic feedback protocol* (5.25) *is that $b_z^{\text{abs}} > 0$.*

Proof: Proceeding as in the proof of Theorem 5.12, we derive the Routh-Hurwitz criterion for the system (5.25) (corresponding to (5.21)) as

$$\begin{aligned}(a_1 + b_z)\text{Re}\{\lambda_l\}\big[&(a_0 + b_z)(a_1 + b_z)\text{Im}\{\lambda_l\}^2\text{Re}\{\lambda_l\} + b_z^{\text{abs}}(a_0 + b_z)\text{Im}\{\lambda_l\}^2 \\ &+ (a_0 + b_z)(a_1 + b_z)\text{Re}\{\lambda_l\}^3 + (b_z^{\text{abs}}(a_1 + b_z) + 2b_z^{\text{abs}}(a_0 + b_z))\text{Re}\{\lambda_l\}^2 \\ &+ ((b_z^{\text{abs}})^2\frac{a_0 + b_z}{a_1 + b_z} + 2(b_z^{\text{abs}})^2)\big] + (b_z^{\text{abs}})^3 - (a_0 + b_z)^2\text{Im}\{\lambda_l\}^2 > 0, \quad (5.28)\end{aligned}$$

for $l = 2, \ldots, N$. If $b_z^{\text{abs}} = 0$, the condition (5.28) will, like in the proof of Theorem 5.12, eventually be violated since $\text{Re}\{\lambda_2\}^2$ approaches zero at a rate faster than $\text{Im}\{\lambda_2\}^2$ as $L \to \infty$. It is, however, possible to choose a $b_z^{\text{abs}} > 0$, so that (5.28) remains satisfied for any $\lambda_l > 0$. ■

Corollary 5.15 says that it is not possible to implement localized integral action to a first-order consensus network that is built over a directed, periodic lattice.

These results will have implications in the next chapter, where we will consider the performance of this type of dynamic feedback controllers. The set-up and notation will differ, which also allows us to derive a slightly more general result compared to Corollary 5.15. The main conclusion, however, is that dynamic feedback control in networks must be carefully designed to guarantee admissibility if only relative measurements are available.

5.6 Discussion

This chapter's main result shows that there is an important difference between the well-studied standard first- and second-order consensus algorithms, and the corresponding higher-order algorithm, in that the latter is not always scalable to large networks. When subject to locality constraints, it will cause instability at some finite network size in an undirected leader-follower network. The same holds for leaderless consensus in large classes of both undirected and directed graphs, including planar graphs and graphs that are embeddable in d-dimensional lattices. An interesting consequence of this result is that at some given network size, the addition of only one agent to a multi-agent network renders the previously converging system unstable. This can be thought of as a type of phase transition, though the physical intuition behind this result is yet to be described.

A further interesting consequence is that an analysis of asymptotic (in network size) performance of localized, relative state feedback is only possible in first- and second-order integrator networks. This means that the analysis in [14] cannot, as was conjectured there, be extended to chains of $n > 2$ integrators.

We have also shown that the second-order consensus algorithm will be inadmissible in any *directed* graph where the real part of at least one Laplacian eigenvalues decreases faster towards zero than the imaginary part as the network size grows. One class of graphs where this holds is periodic lattice graphs of the type we treat in Chapters 6–7, but it is not unlikely that other classes of graphs share the property. If so, it means that second-order consensus with only relative feedback can be highly non-robust in the sense that a small asymmetry in the network can render the system inadmissible. This robustness issue is worth investigating further, along with other classes of graphs where it may appear.

Absolute state feedback partly alleviates the limitations of high-order consensus in terms of admissibility. For example, absolute feedback from acceleration can render third-order consensus admissible, but not absolute position feedback. It is an interesting and open question whether the ideas from Chapter 4, that is, applying integral or derivative action to the available absolute state measurement, can be used to gain admissibility in cases where partial absolute state measurements are available.

In order to discuss scalability of controllers, and admissibility in particular, an assumption that the control algorithm be fixed is necessary. That is, the control cannot be allowed to be re-tuned as the network size grows. This assumption also underlies the performance analyses in the upcoming chapters. By re-tuning the consensus algorithms in this chapter, for example by relaxing the locality assumption, as shown in Example 5.1, or by adjusting the gains, they can achieve consensus also as the network grows. However, such a re-tuning requires global knowledge of the network and prevents "plug-and-play" distributed control.

Chapter 6

Fundamental Limitations of Dynamic Feedback Control in Regular Networks

The central problem in this thesis is to understand and quantify how architectural controller constraints affect global performance. In our case, these constraints can be described by a network graph structure and imply that each controller's connectivity is limited and localized, compared to the overall large-scale network. The question that motivates the work in this chapter is whether dynamic feedback controllers perform better than static ones in large-scale networks under such constraints. In Chapter 4, we answered this question for the particular case where the local controllers had partial access to absolute states. Here, the scope is larger, and we address the performance limitations arising from constraining the feedback to only relative state measurements.

We study comparatively the large-scale asymptotics (in network size) of global performance of localized static versus dynamic feedback control in regular lattice networks. It has previously been shown in [14] that localized static state feedback is subject to limitations in terms of its asymptotic performance scalings in lattices of low spatial dimensions. These limitations do, for example, prevent vehicle platoon formations, which resemble 1-dimensional lattices, from exhibiting a coherent behavior. An important question arises as to whether the use of localized dynamic feedback may alleviate these limitations, that is, whether the controller's additional memory may compensate for the lack of global sensing.

This question of static versus dynamic feedback is a version of an old question in the area of decentralized control [157]. It can be motivated by recalling the following important fact about state feedback control. In fully centralized optimal linear quadratic control (such as LQR or state-feedback H_∞ control), *static state feedback* is optimal. In other words, there is no additional advantage in using dynamic or time-varying controllers over static gains when the full state of the system is available for feedback. This is, however, no longer true when architectural constraints are imposed on the controller, such as diagonal or banded structures [3,

185, 187]. In our case, architectural constraints corresponding to bandedness are imposed through a lattice-network structure, which motivates the use of dynamic feedback in search for the best-achievable performance.

The dynamic feedback controllers we consider are modeled with first-order dynamics and can share their state locally. The resulting control laws can be seen as generalized distributed proportional-integral (PI) controllers. The DAPI controller, which we considered in Chapter 4 and which was previously proposed for distributed frequency regulation in power networks, describes a subclass of such control laws. Here, we consider two classes of systems with respectively first- and second-order local integrator dynamics: consensus and vehicular formation control problems. The systems are modeled over a class of regular networks, specifically toric lattices, with a fixed number of neighbor interactions. In line with Chapter 4, we characterize performance through the achievable network coherence that is described through a nodal variance matric.

The chapter's main result shows that the fundamental limitations in terms of asymptotic performance scalings that apply to localized static feedback in these networks in general carry over also to dynamic feedback control that is locally of first order. This means that while additional memory in the controller may offer other advantages in architecturally constrained control problems, it will not alleviate the unfavorable scaling of nodal variance in low-dimensional lattice networks. An important exception to this result is the case treated in Chapter 4, where a bounded variance could be achieved through distributed PI control if the controller can access noiseless absolute measurements of the local velocity.

As the problems we consider are modeled over toric lattice networks, the resulting systems are *spatially invariant*, see Section 3.3. The topological restriction is a consequence of the aim of the study; to characterize performance scalings in network size. This requires a possibility to grow the network while preserving certain topological properties, such as locality. Hypercubic lattices (including those with periodicity) is one of a few graph families with such topological invariance properties. Others include triangular lattices and fractals, see for example [129]. The periodicity of the lattices allows feedback protocols to be defined using multi-dimensional circulant operators. These enable a tractable characterization through Fourier analysis. At large system sizes, however, the periodic boundary condition will have little or no effect on behaviors in the interior of the network. This intuitive reasoning can be attributed to exponential spatial decay rates of local perturbations [15]. This in particular implies that the lack of coherence that our results predict for 1-dimensional ring-shaped lattices will also be observed in, for example, vehicular platoons without the periodic boundary condition. This is also demonstrated through a case study in Section 6.6.

Chapter overview

Next, we set up the consensus and vehicular formation problems with static versus dynamic feedback control. In Section 6.2, we re-visit the performance measure,

which is defined through the variance of nodal state fluctuations and corresponds to a (scaled) system $\mathcal{H}_2$ norm. Section 6.3 introduces a novel framework for evaluating such $\mathcal{H}_2$ norms and their asymptotic scalings in spatially invariant systems. This framework allows for an analysis of large classes of dynamic feedback protocols whose $\mathcal{H}_2$ norm expressions are otherwise intractable. It is also useful for analyzing stability of these protocols, to which we devote Section 6.4. Building on the results from Chapter 5, we show here that several control designs are inadmissible, that is, they do not allow for a scaling of the network size without rendering the system unstable. The performance of admissible feedback protocols is then analyzed in Section 6.5, where our main result is derived. In Section 6.6 we discuss practical implications of our results and present a numerical simulation. We end this chapter in Section 6.7 with a discussion of our findings as well as some open problems.

6.1 Consensus and vehicular formations with static vs. dynamic feedback

We now formulate models for two types of problems: (first-order) consensus and vehicular formations. Both problems are modeled on networks that are built over the discrete torus $\mathbb{Z}_L^d$ of spatial period L and in d dimensions. The total network size is $N = L^d$. The resulting systems will be assumed *spatially invariant*, and we will use the formalism and notation introduced in Section 3.3 throughout.

In the consensus problem, there is one local information state at each network site, while there are two such states (position and velocity) in the vehicular formation case. For both models, we introduce a *static controller*, as was considered in [14], which we will compare to a *dynamic controller* with an auxiliary memory state at each network site, see Figure 6.1.

6.1.1 Consensus

We first consider the first-order consensus algorithm in continuous time. The single-integrator dynamics at each site k in the network is given by

$$\dot{x}_k = u_k + \mathrm{w}_k, \quad k \in \mathbb{Z}_L^d, \tag{6.1}$$

where u_k denotes the control signal. The process disturbance w_k, modeling random insertions and deletions, is as usual assumed mutually uncorrelated across nodes and modeled as white noise (see Section 1.2).

We next introduce the two types of linear time-invariant feedback control for the system (6.1).

Static feedback

In the case of static feedback, the control input is a linear function of the current state:

$$u_k = (Fx)_k. \tag{6.2}$$

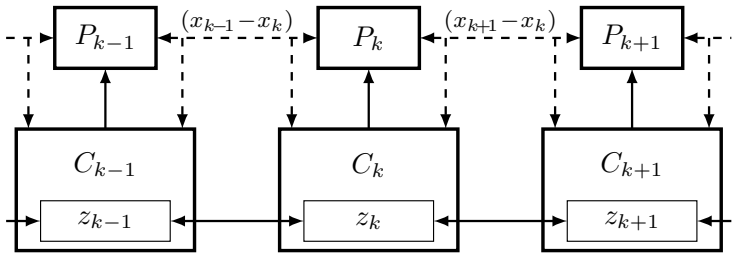


Figure 6.1: Structure of the controller $\{C_k\}$ and plant $\{P_k\}$ interactions. For diagrammatic simplicity only nearest neighbor interaction is depicted, though our analysis is applicable to any fixed number of neighbor interactions. Dashed arrows indicate relative state measurements and interactions. The controller states z_k, rather than just their relative values, can be shared between sub-controllers.

The feedback operator F, can be suitably designed to fulfill the control objectives. A common example of such a control scheme is the one where the control signal at each node is the weighted average of the differences between that node and its $2d$ neighbors, that is,

$$\begin{aligned} u_k =& \tilde{f}[(x_{(k_1-1,\ldots,k_d)} - x_k) + (x_{(k_1+1,\ldots,k_d)} - x_k) + \cdots \\ &+ (x_{(k_1,\ldots,k_d-1)} - x_k) + (x_{(k_1,\ldots,k_d+1)} - x_k)], \end{aligned} \tag{6.3}$$

where $\tilde{f}$ is a positive scalar. The algorithm (6.3) will be referred to as the *standard consensus algorithm*. The associated function array is:

$$f_{(k_1,\ldots,k_d)} = \begin{cases} -2d\tilde{f} & k_1 = \cdots = k_d = 0 \\ \tilde{f} & k_i = \pm 1, \text{ and } k_j = 0, \text{ for } i \neq j \\ 0 & \text{otherwise.} \end{cases} \tag{6.4}$$

In general, we can write the consensus algorithm (6.1) with static feedback as

$$\dot{x} = Fx + \mathrm{w}. \tag{6.5}$$

Dynamic feedback

To model dynamic feedback, we let the controller have access to an auxiliary controller state $z_{(k_1,\ldots,k_d)}$, which is a scalar at each network site k:

$$\begin{aligned} u_k &= z_k + (Fx)_k \\ \dot{z}_k &= (Az)_k + (Bx)_k, \end{aligned}$$

where A, B, F are linear feedback operators, the properties of which will be discussed shortly. We can now write the consensus algorithm (6.1) with dynamic

feedback as:

$$\begin{bmatrix} \dot{z} \\ \dot{x} \end{bmatrix} = \begin{bmatrix} A & B \\ I & F \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{w}. \tag{6.6}$$

Remark 6.1 The static control laws considered here correspond to what we termed proportional (P) control in Chapter 4 and compared to PI and PD controllers. The dynamic control laws considered here are more general, and we therefore use the terms "dynamic" versus "static" in this chapter. □

6.1.2 Structural assumptions for the consensus problem

We now list the assumptions imposed on the system and on the feedback operators A, B, F in the consensus algorithm. Assumptions 6.1–6.2 will also carry over to the vehicular formation problems.

Assumption 6.1 (Spatial invariance) All feedback operators are spatially invariant and fixed with respect to $\mathbb{Z}_L^d$, and are therefore circular convolution operators, as defined in (3.23).

For example, the standard consensus algorithm (6.3) on the 1-dimensional ring graph $\mathbb{Z}_L$ can be written as the convolution of the state x with the array $f = \{0, \ldots, 0, \tilde{f}, -2\tilde{f}, \tilde{f}, 0, \ldots, 0\}$.

Assumption 6.2 (Locality) All feedback operators use only local information from a neighborhood of width $2q$, where q is independent of L. For the function array f associated with the operator F, this means that

$$f_{(k_1,\ldots,k_d)} = 0 \quad \text{if} \quad |k_i| > q, \tag{6.7}$$

for any $i \in \{1, 2, \ldots, d\}$. The same condition holds for all other operators. The situation is illustrated in Figure 6.2.

Assumption 6.3 (Relative state measurements) All controllers can only access *relative* measurements of the physical state x. Hence, the feedback can only involve differences between states of neighboring nodes. This means that each term of the form $\tilde{f}x_k$ in the convolution Fx is accompanied by another term $-\tilde{f}x_l$, for some other index l, so that we obtain $\tilde{f}(x_k - x_l)$. In particular, this implies that the operators F and B in (6.5) and (6.6) have the property

$$\sum_{k \in \mathbb{Z}_L^d} f_k = 0, \quad \sum_{k \in \mathbb{Z}_L^d} b_k = 0. \tag{6.8}$$

Since the state z is internal to each controller, we need *not* impose this requirement on A in (6.6).

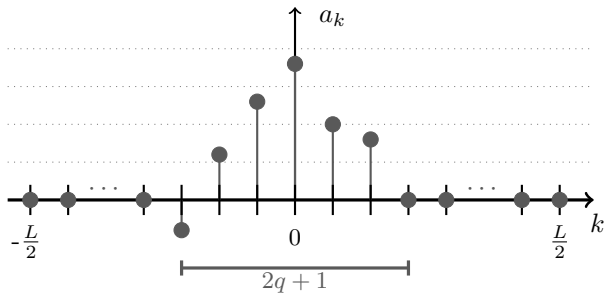


Figure 6.2: Spatial interactions are defined by convolution with an array $\{a_k\}$. As the lattice size L increases, the locality property ($a_k = 0$ for $|k| > q$) ensures that the interactions are unambiguously defined.

6.1.3 Vehicular formations

For the vehicular formation problem, consider $N = L^d$ identical vehicles arranged in $\mathbb{Z}_L^d$. The double-integrator dynamics at each site $k = (k_1, \ldots, k_d) \in \mathbb{Z}_L^d$ is then

$$\ddot{x}_k = u_k + \mathrm{w}_k, \tag{6.9}$$

where, as before, u_k is the control signal and w_k is white process noise, which models random forcings at each site.

The position vector $x_k = [x_k^1 \ \cdots \ x_k^d]^T$ at each network site, and its time derivative, the velocity vector $v_k = [v_k^1 \ \cdots \ v_k^d]^T$, are both d-dimensional vectors. Without loss of generality, we will assume that they each represent absolute deviations from a desired trajectory $\bar{x}_k$ and constant heading velocity $\bar{v}$, with

$$\bar{x}_k := \bar{v}t + k\Delta_x.$$

Here, Δ_x is the constant spacing between the vehicles in $\mathbb{Z}_L^d$.

In analogy to the consensus case, we now introduce the two types of linear feedback control for the system (6.9).

Static feedback

The control input is here assumed to be full state feedback that is linear in the variables x and v:

$$u_k = (Fx)_k + (Gv)_k.$$

An example of such feedback is the combined look-ahead and look-behind controller in a 1-dimensional string:

$$u_k = f_+(x_{k+1} - x_k) + f_-(x_{k-1} - x_k) + g_+(v_{k+1} - v_k) + g_-(v_{k-1} - v_k) - g_o v_k, \tag{6.10}$$

where the g's and f's are positive design parameters. If g_o is zero, this control law satisfies Assumption 6.3 of relative state measurements. If $g_o \neq 0$, we will refer to that term in the feedback law as absolute feedback from velocity.

In practice, absolute velocity measurements can be made available through a speedometer. The presence of viscous damping can also be treated as a special case of absolute velocity feedback. The model (6.9) can then be modified so that

$$\dot{v}_k = -\mu v_k + u_k + \mathrm{w}_k,$$

where $\mu \geq 0$ is the drag coefficient. Comparing this to (6.10) we can identify μ with g_o.

In this chapter, we do not consider the case where absolute feedback is available from the position x_k but not from the velocity v_k. Such a scenario was described in Example 4.1, and can be addressed through the F-DPD controller from Chapter 4. See also Remark 6.4.

In summary, the vehicular formation problem (6.9) with the static feedback law becomes

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & I \\ F & G \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{w}. \tag{6.11}$$

Dynamic feedback

To model the dynamic feedback laws, we introduce the auxiliary controller state z_k at each network site k, which is a d-dimensional vector containing a memory of past position and velocity errors in each coordinate direction. We get:

$$\begin{aligned} u_k &= z_k + (Fx)_k + (Gv)_k \\ \dot{z}_k &= (Az)_k + (Bx)_k + (Cv)_k. \end{aligned}$$

An example of dynamic feedback control for double-integrator systems is the DAPI controller from Chapter 4, which requires absolute velocity feedback. In the notation of this chapter, the DAPI controller on a 1-dimensional string can be written:

$$\begin{aligned} u_k &= z_k + f_+(x_{k+1} - x_k) + f_-(x_{k-1} - x_k) - g_o v_k \\ \dot{z}_k &= a_+(z_{k+1} - z_k) + a_-(z_{k-1} - z_k) - c_o v_k, \end{aligned} \tag{6.12}$$

where the operator A achieves the weighted averaging of the internal state z across nodes, which prevents de-stabilizing drift in the memory states at different nodes.

In general, we can write the equations of motion for the closed loop system with dynamic feedback as:

$$\begin{bmatrix} \dot{z} \\ \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} A & B & C \\ 0 & 0 & I \\ I & F & G \end{bmatrix} \begin{bmatrix} z \\ x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} \mathrm{w}. \tag{6.13}$$

6.1.4 Structural assumptions for the vehicular formation problem

For the vehicular formation systems, we impose the following assumptions *in addition to* Assumptions 6.1–6.2 from before.

Assumption 6.4 (Relative position measurements) The controllers can only access relative measurements of the position states x. This means that the operators F and B in (6.11) and (6.13) have the property (6.8).

Remark 6.2 We will consider the case where the velocity feedback operators G and C have the relative measurement property (6.8) as well as when they do not. We refer to these cases as, respectively, *relative* and *absolute velocity feedback*. □

Assumption 6.5 (Reflection symmetry) The interactions between the vehicles on $\mathbb{Z}_L^d$ are symmetric around each site k. This implies that the arrays associated with the operators A, B, C, F, G have even symmetry, so that for each array element $f_{(k_1,\ldots,k_d)} = f_{(-k_1,\ldots,-k_d)}$. For example, in (6.12) this condition requires $a_+ = a_-$, $b_+ = b_-$, $f_+ = f_-$ and $g_+ = g_-$. A particular implication of this assumption is that the Fourier symbols of the operators will be real valued.

The property of reflection symmetry, will be relevant (but not enforced) also in the consensus case. By slight abuse of terminology, we will in the following refer to a feedback operator as *symmetric* if the associated array has this property, and *asymmetric* if it does not.

Assumption 6.6 (Coordinate decoupling) The feedback in each of the d coordinate directions is entirely decoupled from the vector components in the other coordinates. Furthermore, the array elements associated with the operators A, B, C, F, G are isotropic. By this assumption, the array elements a, b, c, f, g are diagonal and the convolution in (3.23) will turn into d decoupled, identical, scalar convolutions.

Assumptions 6.1–6.4 will be important for the upcoming analysis. Assumption 6.5 is imposed due to the inadmissibility of asymmetric static feedback on toric lattice networks shown in Section 5.4, while Assumption 6.6 is mainly made to simplify the calculations.

6.2 Performance measure and main result

We are concerned with the performance of the consensus and vehicular formation problems in terms of the system's coherence at steady state. This can be quantified as the steady-state variance of nodal state fluctuations, which are caused by persistent stochastic disturbances. In particular, we are interested in the *scaling* of this performance measure with the system size, as it grows asymptotically. We call a system that exhibits a better scaling *more coherent* than a system with bad scaling, as the former will form a more rigid formation when the system grows. If the scaling is such that the variance per node is bounded, the system is said to be *fully coherent*.

As before, we define the relevant performance output as:

Definition 6.1 (Deviation from average performance measure)

$$y_k := x_k - \frac{1}{N}\sum_{l\in\mathbb{Z}_L^d} x_l. \tag{6.14}$$

In operator form, this becomes

$$y = (I - \frac{1}{N}J_{\mathbf{1}})x =: Hx, \tag{6.15}$$

where $J_{\mathbf{1}}$ is the convolution operator corresponding to the array with all elements equal to 1.

Recall that each of the systems we consider is assumed to be driven by zero mean white noise. Provided that it is input-output stable, the system's squared $\mathcal{H}_2$ norm therefore gives the total steady-state variance of the output (see Section 3.1.3):

$$\mathbf{V} := \sum_{k\in\mathbb{Z}_L^d} \lim_{t\to\infty} \mathbb{E}\{y_k^*(t)y_k(t)\}. \tag{6.16}$$

We are now considering spatially invariant systems. This implies that the output variance $\mathbb{E}\{y_k^*(t)y_k(t)\}$ will be independent of the site k in the network. We obtain this *per-site variance* by simply dividing the total $\mathcal{H}_2$ norm by the system size $N = L^d$:

$$V_N = \lim_{t\to\infty} \mathbb{E}\{y_k^*(t)y_k(t)\} = \frac{\mathbf{V}}{N}. \tag{6.17}$$

Remark 6.3 The consensus type dynamics considered here will typically have a single marginally stable mode at the origin corresponding to the motion of the average (in each of the d coordinate directions). This is a consequence of Assumption 6.3 of relative measurements. The $\mathcal{H}_2$ norm (6.16) is only finite if this mode is unobservable from the system output. Here, the output operator H has the relative measurement property (6.8), that is, $\sum_{k\in\mathbb{Z}_L^d} h_k = 0$, implying that the average mode is indeed unobservable. Provided remaining system modes are stable, $\mathbf{V}$ in (6.16) will thus be finite for any finite system size N; a condition equivalent to bounded-input, bounded-output (BIBO) stability. □

6.2.1 Performance scalings with static and dynamic feeedback

Our main objective is to determine whether dynamic feedback may improve performance compared to the static feedback laws that were also evaluated in [14]. The following sections will introduce the methodology that is used to establish asymptotic scalings of performance. At this point, we summarize the chapter's main results as follows.

Theorem 6.1 (Asymptotic performance scalings) *Consider the first-order consensus problem under Assumptions 6.1–6.3 and 6.7 and the vehicular formation problem under Assumptions 6.1, 6.2, and 6.4–6.6. The steady-state per-site variance V_N defined in* (6.17) *then scales asymptotically as follows:*

1. *Consensus*

   Static *feedback* or dynamic *feedback*

$$V_N \sim \frac{1}{\beta} \begin{cases} N & d = 1 \\ \log N & d = 2 \\ 1 & d \geq 3, \end{cases} \tag{6.18}$$

2. *Vehicular formations*

   a) *Relative feedback:*
   Static *feedback* or dynamic *feedback*

$$V_N \sim \frac{1}{\beta^2} \begin{cases} N^3 & d = 1 \\ N & d = 2 \\ N^{1/3} & d = 3 \\ \log N & d = 4 \\ 1 & d \geq 5, \end{cases} \tag{6.19}$$

   b) *Absolute velocity (but relative position) feedback:*
   Static *feedback*

$$V_N \sim \frac{1}{\beta} \begin{cases} N & d = 1 \\ \log N & d = 2 \\ 1 & d \geq 3, \end{cases} \tag{6.20}$$

   Dynamic *feedback*

$$V_N \sim 1, \tag{6.21}$$

where $N = L^d$ is the network size, $\beta = \max\{||f||_\infty, ||g||_\infty\}$ is an algorithm parameter reflecting the magnitude of feedback gains, and the symbol $\sim$ denotes scaling up to a factor that is independent of N and β in the manner defined in (3.26).

Therefore, if only relative state measurements are available (Assumption 6.3), no dynamic feedback laws on the forms (6.6) and (6.13) exhibit better coherence properties than static, memoryless feedback under the given assumptions.

However, a dynamic feedback law can theoretically achieve full coherence in any spatial dimension using absolute feedback from velocities, even though position measurements are relative. As previously shown in [14], and as will be evident from the developments in Section 6.5, a static feedback protocol would require absolute measurements of *both* states to achieve the same performance.

Remark 6.4 The case with absolute position but relative velocity feedback is not considered in this chapter. The interested reader is referred to [174] where it is shown that V_N then scales as in (6.20) for both the static and dynamic feedback laws modeled in this chapter. A distributed PD controller was, however, shown in Chapter 4 to give full coherence in this case. □

6.3 The $\mathcal{H}_2$-norm density and asymptotic scalings

We now introduce the technical framework that will be used to determine the $\mathcal{H}_2$ performance scalings in Theorem 6.1. This novel framework is based on the idea of mapping the operators that define the system dynamics onto an infinite lattice. Usually, $\mathcal{H}_2$ norms are calculated using traces of system Gramians that lead to sums involving system eigenvalues. In the limit of large systems, they can instead be estimated through integrals over a continuous function which we call the $\mathcal{H}_2$-norm density. We show that simple properties of this $\mathcal{H}_2$-norm density determine the asymptotic performance scalings.

6.3.1 The limit from finite to infinite lattices

All feedback operators considered in this chapter define convolutions with *local* arrays on $\mathbb{Z}_L^d$, by Assumption 6.2. Hence, for a given operator A we have that $a_k = 0$ if $|k| > q$ for some fixed q. This means that any such array a can be unambiguously re-defined on $\mathbb{Z}_{L'}^d$ for *any* given $L' > 2q$ by filling it with zero components wherever $|k| > q$. This also means that the same array can be used to define a convolution over the *infinite* lattice $\mathbb{Z}^d$. As we shall see, such a re-definition proves useful when analyzing the systems asymptotically.

Let a be a local array defined over $\mathbb{Z}_L^d$ and a_∞ its counterpart defined on $\mathbb{Z}^d$, in which the elements $\{a_k\}$ have been filled out with zeros for $|k| > q$ up until infinity. The discrete Fourier transform (DFT) of a, denoted $\hat{a}_n$ is given by (3.24), while the Z-transform of the array a_∞, denoted $\hat{a}_\infty(\theta)$, is given by (3.25). Comparing (3.24) with (3.25) it is clear that the DFT of a are simply subsamples of the Z-transform of a_∞:

$$\hat{a}_n = \hat{a}_\infty\left(\frac{2\pi}{L}n\right), \quad n \in \mathbb{Z}_L^d. \tag{6.22}$$

Given that we are interested in system behaviors as $N \to \infty$, it will be convenient to consider these Z-transforms of operators over the infinite lattice $\mathbb{Z}^d$, and their behavior in the continuous spatial frequency variable $\theta \in \mathcal{R}^d := [-\pi, \pi]^d$, rather than the DFTs at discrete spatial wavenumbers.

For this purpose, let us take a general state space system on the form (3.4) and map the system operators $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ onto $\mathbb{Z}^d$ to obtain $\mathcal{A}_\infty$, $\mathcal{B}_\infty$, $\mathcal{C}_\infty$. By virtue of the spatial invariance property, $\mathcal{A}_\infty$, $\mathcal{B}_\infty$ and $\mathcal{C}_\infty$ are circulant convolution operators and the Z-transform can be used to *(block) diagonalize* them, see e.g. [15]. Then, at each $\theta \in \mathcal{R}^d$, we obtain the transforms $\hat{\mathcal{A}}_\infty(\theta)$, $\hat{\mathcal{B}}_\infty(\theta)$ and $\hat{\mathcal{C}}_\infty(\theta)$, which are

matrix-valued in general. The DFTs $\hat{\mathcal{A}}_n$, $\hat{\mathcal{B}}_n$, $\hat{\mathcal{C}}_n$ of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are now precisely the values of $\hat{\mathcal{A}}_\infty(\theta)$, $\hat{\mathcal{B}}_\infty(\theta)$ and $\hat{\mathcal{C}}_\infty(\theta)$ at $\theta = \frac{2\pi}{L}n$, for all wavenumbers $n \in \mathbb{Z}_L^d$. Consider the following example:

Example 6.1 In the static consensus system (6.5), we have $\mathcal{A} = F$. If we let F represent the standard consensus algorithm (6.3), then $\mathcal{A}_\infty = F_\infty$ has the associated function array f_∞, defined just as in (6.4) (note that it has infinitely many zero components for $|k_i| > 1$, $i = 1, \dots, d$).

The Z-transform of f_∞ is:

$$\hat{f}_\infty(\theta) = \tilde{f}(-2 + e^{\mathrm{j}\theta} + e^{-\mathrm{j}\theta}) = -2\tilde{f}(1 - \cos\theta), \tag{6.23}$$

for $\theta \in \mathcal{R}^d$. The DFT of f, corresponding to the N eigenvalues of F, is therefore given by $\hat{f}_n = -2\tilde{f}(1 - \cos\frac{2\pi}{L}n)$ for $n \in \mathbb{Z}_L^d$.

6.3.2 $\mathcal{H}_2$ norm evaluation in the spatial frequency domain

From now on, let us assume that the system under consideration is input-output stable, so that its $\mathcal{H}_2$ norm (6.16) exists. Let the system be on the general form (3.4). Its $\mathcal{H}_2$ norm can then be calculated as

$$\mathbf{V} = \mathrm{tr}\left(\int_0^\infty \mathcal{B}^* e^{\mathcal{A}^* t}\mathcal{C}^*\mathcal{C}e^{\mathcal{A}t}\mathcal{B}\mathrm{d}t\right). \tag{6.24}$$

Now, recall that the system could be (block) diagonalized by the DFT, where the Fourier symbols $\hat{\mathcal{A}}_n, \hat{\mathcal{B}}_n, \hat{\mathcal{C}}_n$ correspond to decoupled diagonal elements. This is a unitary transformation, so in line with the discussion in Section 3.1.2, the trace in (6.24) can be re-written as:

$$\mathbf{V} = \mathrm{tr}\left(\sum_{n\in\mathbb{Z}_L^d}\int_0^\infty \hat{\mathcal{B}}_n^* e^{\hat{\mathcal{A}}_n^* t}\hat{\mathcal{C}}_n^*\hat{\mathcal{C}}_n e^{\hat{\mathcal{A}}_n t}\hat{\mathcal{B}}_n\mathrm{d}t\right). \tag{6.25}$$

Now, consider the output operator H defined in (6.15). It is easy to verify that its Fourier symbol is $\hat{h}_0 = 0$, and $\hat{h}_n = 1$ for $n \neq 0$. This implies that the output matrix $\hat{\mathcal{C}}_0 = 0$ for all systems (that is, the zero mode is unobservable). Consequently, we can obtain the $\mathcal{H}_2$ norm in (6.25) by summing only over $n \in \mathbb{Z}_L^d \backslash \{0\}$.

Furthermore, following the discussion in the previous section, we can regard the Fourier symbols in (6.25) as subsamples of $\hat{\mathcal{A}}_\infty(\theta), \hat{\mathcal{B}}_\infty(\theta)$, and $\hat{\mathcal{C}}_\infty(\theta)$. Given this relationship, we can now state the per-site variance $V_N = \mathbf{V}/N$ from (6.17) as

$$V_N = \frac{1}{N}\sum_{\substack{\theta=\frac{2\pi}{L}n \\ n\in\mathbb{Z}_L^d\backslash\{0\}}} \mathrm{tr}\left(\hat{\mathcal{B}}_\infty^*(\theta)\hat{P}(\theta)\hat{\mathcal{B}}_\infty(\theta)\right), \tag{6.26}$$

where the individual time integrals are defined as follows:

Definition 6.2

$$\hat{P}(\theta) := \int_0^\infty e^{\hat{\mathcal{A}}_\infty^*(\theta)t}\hat{\mathcal{C}}_\infty^*(\theta)\hat{\mathcal{C}}_\infty(\theta)e^{\hat{\mathcal{A}}_\infty(\theta)t}\mathrm{d}t. \tag{6.27}$$

We call $\hat{P}(\theta)$ the *observability Gramian* at θ.

The observability Gramian at each $\theta \neq 0$ is obtained by solving the Lyapunov equation

$$\hat{\mathcal{A}}_\infty^*(\theta)\hat{P}(\theta) + \hat{P}(\theta)\hat{\mathcal{A}}_\infty(\theta) = -\hat{\mathcal{C}}_\infty^*(\theta)\hat{\mathcal{C}}_\infty(\theta), \tag{6.28}$$

and is unique and finite provided $\hat{\mathcal{A}}_\infty(\theta)$ is Hurwitz.

For all problem formulations considered here, $\hat{\mathcal{B}}_\infty(\theta)$ is a vector where one element[1] is 1 and remaining elements are zero. Hence, $\mathrm{tr}(\hat{\mathcal{B}}_\infty^*(\theta)\hat{P}(\theta)\hat{\mathcal{B}}_\infty(\theta))$ in (6.26) is just one element of the matrix $\hat{P}(\theta)$[2]. This is a quantity that will be used often in the following, and we make the definition:

Definition 6.3 (Per-site $\mathcal{H}_2$-norm density)

$$\hat{p}(\theta) := \mathrm{tr}\left(\hat{\mathcal{B}}_\infty^*(\theta)\hat{P}(\theta)\hat{\mathcal{B}}_\infty(\theta)\right). \tag{6.29}$$

This quantity captures the distribution of the per-site variance V_N over the spatial frequency variable θ and we therefore refer to it as the *per-site $\mathcal{H}_2$-norm density*.

Now, notice that if the value of $\hat{p}(\theta)$ is bounded for all $\theta \in \mathcal{R}^d$, then V_N in (6.26) will remain bounded as $N \to \infty$ and the system in question is *fully coherent*. For the consensus and vehicular formation problems, however, there is typically a single zero eigenvalue at wavenumber $n = 0$ that corresponds to the spatial average mode (see Section 6.2). This makes $\hat{\mathcal{A}}_\infty(0)$ non-Hurwitz, and in turn causes a singularity in $\hat{p}(\theta)$ at $\theta = 0$. Even though the mode at $\theta = 0$ itself is unobservable from the system output, the singularity makes the $\mathcal{H}_2$-norm density grow unboundedly for small θ, that is, for small wavenumbers.

For this reason, we use the following appropriate integral to estimate the value of the sum in (6.26):

$$S(\Delta) := \int_{\Delta \leq |\theta_1| \leq \pi} \cdots \int_{\Delta \leq |\theta_d| \leq \pi} \hat{p}(\theta)\ \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_d, \tag{6.30}$$

where the argument Δ indicates the size of a deleted neighborhood around $\theta = 0$. We recognize the sum in (6.26) as a Riemann sum approximation of the integral (6.30) with volume element $1/N = 1/L^d$. The integral can therefore be used to bound the sum asymptotically. Consider the following lemma:

[1] In the vehicular formation case, each "element" is a $d \times d$ diagonal matrix, where each of the d diagonal elements is equal by Assumption 6.6.

[2] Or the sum of d identical such elements.

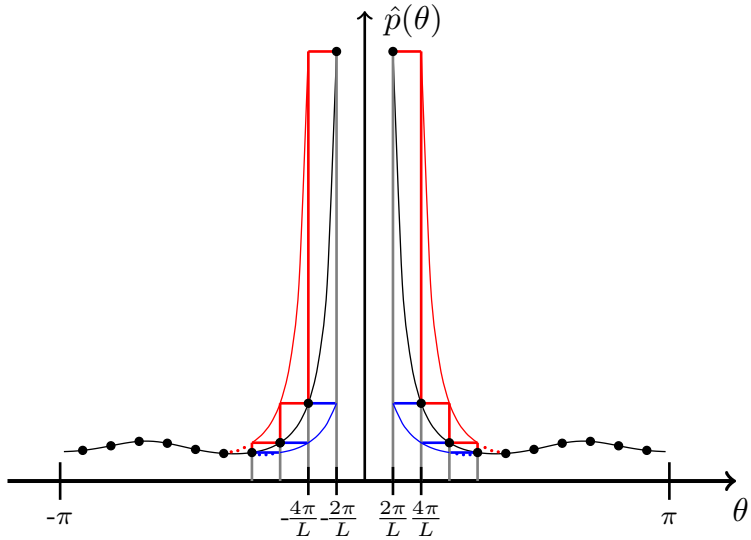


Figure 6.3: Illustration of the upper and lower bounds in Lemma 6.2. The per-site variance V_N is a sum that can be bounded by the lower (upper) Riemann integrals of the $\mathcal{H}_2$-norm density $\hat{p}(\theta)$ shifted by $\frac{2\pi}{L}$ ($\frac{4\pi}{L}$) represented in red (blue). The systems we consider have a zero eigenvalue at $\theta = 0$, causing a singularity in $\hat{p}(\theta)$. The order of this singularity, that is, the rate at which $\hat{p}(\theta) \xrightarrow{\theta \to 0} \infty$ determines how fast the integrals, and thereby V_N, grows as $L \to \infty$. This growth corresponds to the asymptotic performance scaling.

Lemma 6.2 *The per-site variance V_N in* (6.26) *is upper and lower bounded by the integral* (6.30) *as*

$$S\left(\frac{4\pi}{L}\right) \leq V_N \leq S\left(\frac{2\pi}{L}\right), \tag{6.31}$$

for all $L > \bar{L}$, for some fixed $\bar{L}$.

Proof: See Appendix 6.C.

The integral and the Riemann sum approximations are illustrated in Figure 6.3.

The performance of the consensus and vehicular formation systems can now be evaluated as follows. First, the system operators are re-defined on $\mathbb{Z}^d$ and (block) diagonalized using the Z-transform (3.25). Second, the Lyapunov equation (6.28) is solved to determine $\hat{p}(\theta)$. Bounds on the variances V_N are then found through Lemma 6.2. Next, we derive general expressions for the scaling (in L) of the integral (6.30).

Remark 6.5 It is important to note that the systems we consider remain of *finite* size N throughout. This is a preqrequisite for the finiteness of the $\mathcal{H}_2$ norm. Only

the system operators are re-defined onto the infinite lattice, to facilitate an estimation of the $\mathcal{H}_2$ norm by the integrated $\mathcal{H}_2$-norm density according to Lemma 6.2. □

6.3.3 Bounds on asymptotic scalings

We are interested in the scaling of the per-site variance V_N in (6.26) with the total number of nodes N as this number grows large. Using the integral in (6.30) and the bounds in Lemma 6.2, we can now derive asymptotic scalings of V_N by exploiting bounds on the per-site $\mathcal{H}_2$-norm density $\hat{p}(\theta)$. We begin by a simple example.

Example 6.2 Consider the standard consensus algorithm (6.3) and for simplicity let the dimension $d = 1$. The Lyapunov equation (6.28) is scalar and solved by

$$\hat{P}(\theta) = \hat{p}(\theta) = \frac{1}{2}\frac{-1}{\hat{f}_\infty(\theta)} \tag{6.32}$$

for all $\theta \in [-\pi, \pi]\backslash\{0\}$, where we have used that $\hat{\mathcal{C}}_\infty(\theta) = \hat{h}_\infty(\theta) = 1$ for $\theta \neq 0$. The Fourier symbol $\hat{f}_\infty(\theta)$ was given by (6.23) in Example 6.1. Substituting it into the $\mathcal{H}_2$-norm density (6.32), the integral from (6.30) becomes:

$$S(\Delta) = \frac{1}{4\tilde{f}}\int_{\Delta\leq|\theta|\leq\pi}\frac{1}{1-\cos\theta}\mathrm{d}\theta.$$

The lower bound in from Lemma 6.2 is

$$S\left(\frac{4\pi}{L}\right) = \frac{-1}{2\tilde{f}}\left[\cot\frac{\theta}{2}\right]_{\frac{4\pi}{L}}^{\pi} = \frac{1}{2\tilde{f}}\cot\frac{2\pi}{L},$$

and the upper bound has the same form. A series expansion of the cotangent function reveals that this expression scales as $\frac{1}{\tilde{f}}L = \frac{1}{\tilde{f}}N$ asymptotically.

This result is represented in case 1a) of Theorem 6.1 ($\tilde{f}$ here corresponds to the algorithm parameter β).

In general, let us assume that the $\mathcal{H}_2$-norm density is such that

$$\hat{p}(\theta) \sim \frac{1}{\beta^{r/2}}\frac{1}{(\theta_1^2+\theta_2^2+\cdots+\theta_d^2)^{r/2}}, \tag{6.33}$$

for some non-negative r. The number r characterizes the order of the $\mathcal{H}_2$-norm density's singularity at $\theta = 0$. In the upcoming analysis, we will show that any admissible controller for the systems considered here results in $\mathcal{H}_2$-norm densities that satisfy (6.33) with $r \in \{0, 2, 4\}$.

We have also introduced the algorithm parameter β, which reflects the size of the system's feedback gains (c.f. $\tilde{f}$ in Example 6.2). In particular, let $\beta :=$

$\max\{||f||_\infty, ||g||_\infty\}$. All feedback array elements, which are bounded by assumption, are then proportional to β. We show in Section 6.5.5 that the parameter β is bounded by the system's total control effort. It can therefore be considered a proxy for control effort.

The number r, determines the coherence properties for a given system. If $r = 0$, the system is fully coherent. Otherwise, the level of coherence depends on the spatial dimension d of the network. We now state the main result of this section:

Lemma 6.3 *Assume that the per-site $\mathcal{H}_2$-norm density $\hat{p}(\theta)$ defined in* (6.29) *satisfies* (6.33)*. The steady-state per-site variance then scales asymptotically as*

$$V_N \sim \frac{1}{\beta^{r/2}} \begin{cases} L^{r-d} & \text{if } d < r \\ \log L & \text{if } d = r \\ 1 & \text{if } d > r \end{cases} \tag{6.34}$$

up to some constant, which is independent of the lattice size L and the algorithm parameter β.

Proof: First, substitute the approximation (6.33) into the integral $S(\Delta)$ in (6.30) and denote the resulting integral $\tilde{S}(\Delta)$. We transform this to hyperspherical coordinates by defining $\rho = (\theta_1^2 + \cdots + \theta_d^2)^{1/2}$ and the $d-1$ coordinates $\phi_1, \ldots, \phi_{d-2} \in [0, \pi]$ and $\phi_{d-1} \in [0, 2\pi]$ which are such that $\theta_i = \rho \cos\phi_i \prod_{j=1}^{i-1} \sin\phi_j$ for $i = 1, \ldots, d-1$ and $\theta_d = \rho \prod_{j=1}^{d-1} \sin\phi_j$. We obtain:

$$\begin{aligned} \tilde{S}(\Delta) &= \int_{\Delta \le |\theta_1| \le \pi} \cdots \int_{\Delta \le |\theta_d| \le \pi} \frac{1}{\beta^{r/2}} \frac{1}{(\theta_1^2 + \cdots + \theta_d^2)^{r/2}} \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_d \\ &= \frac{c_d}{\beta^{r/2}} \int_\Delta^\pi \int_0^{2\pi} \int_0^\pi \cdots \int_0^\pi \frac{1}{\rho^r} \rho^{d-1} \sin^{d-2}\phi_1 \cdots \sin\phi_{d-2} \mathrm{d}\rho \mathrm{d}\phi_{d-1} \cdots \mathrm{d}\phi_1 \\ &= \frac{c_d}{\beta^{r/2}} \mathcal{S}_d \int_\Delta^\pi \rho^{d-r-1} \mathrm{d}\rho, \end{aligned} \tag{6.35}$$

where $\mathcal{S}_d$ is the (generalized) surface area of the d-dimensional unit sphere and c_d is a non-zero and finite scaling factor arising from integrating over a hypersphere rather than a hypercube.

Now, by Lemma 6.2 we know that V_N is bounded as

$$\underline{c}\tilde{S}\left(\frac{4\pi}{L}\right) \le S\left(\frac{4\pi}{L}\right) \le V_N \le S\left(\frac{2\pi}{L}\right) \le \bar{c}\tilde{S}\left(\frac{2\pi}{L}\right),$$

for all $L \ge \bar{L}$ for some $\bar{L}$, and with the constants $\underline{c}, \bar{c}$ from the scaling bounds in (6.33). Substituting for Δ in (6.35) the values $\frac{2\pi}{L}$ and $\frac{4\pi}{L}$ from these upper and lower bounds and defining new constants $\underline{c}', \bar{c}'$, the solution to the integral gives

that

$$V_N \leq \bar{c}' \mathcal{S}_d \frac{1}{\beta^{r/2}} \begin{cases} \frac{1}{r-d}\pi^{d-r}\left(\left(\frac{L}{2}\right)^{r-d} - 1\right) & \text{if } d \neq r \\ \log L - \log 2 & \text{if } d = r \end{cases}$$

$$V_N \geq \underline{c}' \mathcal{S}_d \frac{1}{\beta^{r/2}} \begin{cases} \frac{1}{r-d}\pi^{d-r}\left(\left(\frac{L}{4}\right)^{r-d} - 1\right) & \text{if } d \neq r \\ \log L - \log 4 & \text{if } d = r \end{cases}$$

Noticing that these bounds are identical up to a constant for any given d, the result (6.34) follows. ■

6.4 Admissibility of dynamic feedback laws

We now turn to the question of stability of the consensus and vehicular formation systems with dynamic feedback, which is a prerequisite for the $\mathcal{H}_2$ performance evaluation laid out in the previous section. In particular, we must require the underlying systems to be BIBO stable for any network size N to allow for the asymptotic performance analysis. In Chapter 5, we learnt that this is often not possible for localized high-order algorithms. Some feedback configurations will namely cause instability beyond a certain network size and therefore be *inadmissible* with respect to our analysis. In order to rule these out, this section presents necessary conditions for the admissibility of our controllers.

6.4.1 Conditions for input-output stability

The stability of a given LTI system on the form (3.4) can be verified by ensuring that its individual Fourier symbols are stable in their own right. We begin by re-stating the following Theorem from related work:

Theorem 6.4 *[15, Corollary 1] The system* (3.4) *on* $\mathbb{Z}_L^d$ *is exponentially stable if and only if the matrix* $\hat{\mathcal{A}}_n$ *is Hurwitz stable for every* $n \in \mathbb{Z}_L^d$.

Proof: See [15, Theorem 1] and note that the group $\mathbb{Z}_L^d$ is compact. ■

Now, we are evaluating these systems asymptotically, and must therefore require that they remain stable for *any* lattice size L, as this number grows. Since the Fourier symbols $\hat{\mathcal{A}}_n$ can be seen as subsamples of $\hat{\mathcal{A}}_\infty(\theta)$ (see Section 6.3.1), the only way to ensure stability for any lattice size L is to make sure that $\hat{\mathcal{A}}_\infty(\theta)$ is stable for every θ. In our case, though, the mode at $n = 0$ is unobservable from the considered output (see Remark 6.3). BIBO stability is therefore guaranteed if $\hat{\mathcal{A}}_\infty(\theta)$ is stable for every θ away from zero:

Corollary 6.5 *The system* (3.4) *on* $\mathbb{Z}_L^d$ *with output defined as in* (6.15) *is BIBO stable for any network size* $N = L^d$ *if and only if the matrix* $\hat{\mathcal{A}}_\infty(\theta)$ *is Hurwitz stable for all* $\theta \in \mathcal{R}^d \backslash \{0\}$.

The upcoming performance analysis should be constrained to feedback laws that guarantee stability for any network size N according to Corollary 6.5. We therefore make the following specific admissibility definition:

Definition 6.4 (Admissible feedback law) A feedback control law defined on $\mathbb{Z}_L^d$ is *admissible* if the corresponding closed-loop system is BIBO stable with respect to the output (6.15) for any network size $N = L^d$.

Remark 6.6 Note that the considered systems are finite-dimensional for any given lattice size L. Their BIBO stability is therefore equivalent to the total variance $\mathbf{V}$ being bounded. □

Under relative feedback, admissibility of the dynamic feedback laws is not straightforward. We next present some necessary conditions.

6.4.2 Admissibility conditions under relative feedback

First, consider the consensus problem with dynamic feedback (6.6) with feedback operators A, B, F. Using Corollary 6.5, we derive the following theorem:

Theorem 6.6 *Consider the consensus system* (6.6). *A necessary condition for admissibility is that at least one of the following conditions holds:*

a) The operator B *is* symmetric,

b) The operator A *involves* absolute feedback, *that is,* A *does not satisfy* (6.8).

Proof: See Appendix 6.D.

In the vehicular formation case with relative velocity feedback, a similar admissibility condition holds:

Theorem 6.7 *Consider the vehicular formation system* (6.13), *where the feedback operators* F, G, B, C *have the relative measurement property* (6.8). *A necessary condition for admissibility is that at least one of the following conditions holds:*

a) The operator $B = 0$, *while* $A \neq 0$,

b) The operator A *involves* absolute feedback, *that is,* A *does not satisfy* (6.8).

Proof: See Appendix 6.E.

Theorem 6.7 implies that integral control based on position measurements cannot be implemented for large networks, unless there is built-in "damping" of the auxiliary state through an absolute feedback term in A. Note, however, that if the purpose of the dynamic feedback law is to eliminate stationary errors through

integral action, including such a term in A would defeat the purpose. In this case, the auxiliary state z is namely stabilized and the integral action reduced to zero.

Remark 6.7 Theorems 6.6–6.7 imply that a system with a given feedback protocol may be stable for small lattice sizes L, but becomes unstable at some lattice size L^{crit} unless the criteria are satisfied. As long as the control effort (feedback gains) is bounded, L^{crit} will always exist and be finite. □

Remark 6.8 A result that is analogous to Theorem 6.7 was presented in Corollary 5.14 in a general network setting. This means that Theorem 6.7 is not an artefact of this chapter's topological restriction to a toric lattice network. □

6.5 Performance scalings with dynamic feedback

We established in Section 6.3 that the asymptotic performance scaling depends on properties of the per-site $\mathcal{H}_2$-norm density. We now evaluate the $\mathcal{H}_2$ norm densities for admissible feedback laws, and derive this chapter's main result that was previewed in Theorem 6.1. In order to establish results for dynamic feedback, we first need to consider the respective problem under static feedback.

6.5.1 Consensus: performance with static feedback

We begin by deriving the performance scaling for the static consensus problem (6.2). As seen in Example 6.2, the Lyapunov equation (6.28) is a scalar equation, which is solved by

$$\hat{P}(\theta) = \hat{p}(\theta) = \frac{-1}{2\text{Re}\{\hat{f}_\infty(\theta)\}}. \tag{6.36}$$

Now, consider the following lemma:

Lemma 6.8 *Consider the static consensus system* (6.2). *Provided the feedback operator F is admissible, it holds*

$$\text{Re}\{\hat{f}_\infty(\theta)\} \sim -\beta(\theta_1^2 + \ldots + \theta_d^2). \tag{6.37}$$

Proof: By the definition of the Z-transform (3.25) it holds

$$\text{Re}\{\hat{f}_\infty(\theta)\} = \sum_{k\in\mathbb{Z}^d} f_k \cos(\theta\cdot k) = \sum_{k\in\mathbb{Z}^d} f_k[1-(1-\cos(\theta\cdot k))] = -\sum_{k\in\mathbb{Z}^d} f_k(1-\cos(\theta\cdot k))\,, \tag{6.38}$$

where we have used the relative measurement property (6.8), which implies $\sum_{k\in\mathbb{Z}^d} f_k = \sum_{k\in\mathbb{Z}^d_N} f_k = 0$. A Taylor series expansion of (6.38) around $\theta = 0$ is

$$\sum_{k\in\mathbb{Z}^d} f_k(1-\cos(\theta\cdot k)) = \sum_{k\in\mathbb{Z}^d} f_k \left(\frac{(\theta\cdot k)^2}{2} - \frac{(\theta\cdot k)^4}{4!} + \cdots\right), \tag{6.39}$$

which is upper bounded by its first term, that is, $\sum_{k\in\mathbb{Z}^d} f_k\,(1-\cos(\theta\cdot k)) \leq \sum_{k\in\mathbb{Z}^d} f_k \frac{(\theta\cdot k)^2}{2}$ for all θ. Thus,

$$\begin{aligned}\sum_{k\in\mathbb{Z}^d} f_k\,(1-\cos(\theta\cdot k)) &\leq \frac{1}{2}\sum_{k\in\mathbb{Z}^d} |f_k|\,(k_1\theta_1+\cdots+k_d\theta_d)^2\\ &\leq \frac{1}{2}\sum_{0\neq k\in\mathbb{Z}^d} ||f||_\infty q^2\,(|\theta_1|+\cdots+|\theta_d|)^2\\ &\leq 2^{d-1}q^{d+2}||f||_\infty(2d+1)(\theta_1^2+\cdots+\theta_d^2), \quad (6.40)\end{aligned}$$

where the second inequality follows from the locality assumption (6.7) and the third from straightforward algebra.

Next, the Taylor expansion (6.39) reveals that $\mathrm{Re}\{\hat{f}_\infty(\theta)\}$ goes to zero at a quadratic rate. We can therefore always find a fixed, nonnegative $\underline{c}$ so that

$$-\mathrm{Re}\{\hat{f}_\infty(\theta)\} = \sum_{k\in\mathbb{Z}^d} f_k\,(1-\cos(\theta\cdot k)) \geq \underline{c}(\theta_1^2+\cdots+\theta_d^2) \qquad (6.41)$$

in some interval near zero; $\theta\in[-\Delta,\Delta]^d$ for a small Δ. Note that no lower-degree polynomial in θ (apart from the zero polynomial) could serve as a lower bound in (6.41). Furthermore, given that the feedback law is admissible, it must hold $-\mathrm{Re}\{\hat{f}_\infty(\theta)\} > \varepsilon > 0$ for all $\theta\in\mathcal{R}^d\backslash[-\Delta,\Delta]^d =: \mathcal{R}^d_\Delta$ with any fixed Δ. We can therefore always adjust $\underline{c}$ so that (6.41) holds for the entire region $\mathcal{R}^d$.

Defining the algorithm parameter $\beta = ||f||_\infty$, and noticing that remaining parameters of (6.40) and (6.41) are independent of θ and L, the result (6.37) follows. ■

Inserting the scaling from Lemma 6.8 into the $\mathcal{H}_2$-norm density (6.36) shows that

$$\hat{p}(\theta) = \frac{-1}{2\mathrm{Re}\{\hat{f}_\infty(\theta)\}} \sim \frac{1}{\beta}\frac{1}{(\theta_1^2+\cdots+\theta_d^2)}, \qquad (6.42)$$

that is, the $\mathcal{H}_2$-norm density for the static consensus system (6.2) satisfies (6.33) with $r=2$. The per-site variance thus scales according to Lemma 6.3 with $r=2$.

6.5.2 Consensus: performance with dynamic feedback

Before turning to the case of dynamic feedback, note that the performance of the consensus system with static feedback is independent of any imaginary part of the Fourier symbol $\hat{f}_\infty(\theta)$. It is therefore independent of whether the feedback operator F is symmetric or not. In the upcoming evaluation of dynamic feedback, we will therefore limit the analysis to F being symmetric:

Assumption 6.7 The operator F in the dynamic consensus protocol (6.6) is symmetric, that is, it satisfies the properties listed in Assumption 6.5. It follows that $\hat{f}_\infty(\theta) = \mathrm{Re}\{\hat{f}_\infty(\theta)\}$.

Remark 6.9 Assumption 6.7 is made to simplify the exposition by limiting the number of possible feedback configurations that must be considered. It is our belief, based on computer-aided evaluation, that the main result would hold also without this assumption. □

Let us now assume that the choice of operators A, B, F is admissible. The solution to the Lyapunov equation (6.28) then gives that

$$\hat{p}(\theta) = \frac{-1}{2\hat{f}_\infty(\theta) + 2\varphi^c(\theta)}, \tag{6.43}$$

where $\varphi^c(\theta)$ is a function of the Fourier symbols of A, B and F. This $\mathcal{H}_2$-norm density would scale differently from (6.42) if the function $\varphi^c(\theta)$ were non-zero and *scaled differently* in θ than $\hat{f}_\infty(\theta)$, for which we established a scaling in Lemma 6.8. This is, however, not the case for any admissible configuration of the feedback operators A and B. Consider the following lemma:

Lemma 6.9 *For any admissible choice of the operators A, B, F in* (6.6) *satisfying Assumptions 6.1–6.3, 6.7, the function $\varphi^c(\theta)$ in* (6.43) *is such that*

$$\hat{f}_\infty(\theta) + \varphi^c(\theta) \sim -\beta(\theta_1^2 + \ldots + \theta_d^2). \tag{6.44}$$

Therefore, the $\mathcal{H}_2$-norm density $\hat{p}(\theta)$ in (6.43) *will satisfy* (6.33) *with $r = 2$ for any design of the dynamic feedback law.*

Proof: See Appendix 6.F.

The asymptotic performance scaling will thus be unchanged compared to static feedback. Rewriting the asymptotic scalings from Lemma 6.3 in terms of total network size $N = L^d$ gives the result in Theorem 6.1.

6.5.3 Vehicular formations: performance with static feedback

Consider the vehicular formation problem under static feedback (6.11). The solution to the Lyapunov equation (6.28) gives the $\mathcal{H}_2$-norm density

$$\hat{p}(\theta) = \frac{d}{2\hat{f}_\infty(\theta)\hat{g}_\infty(\theta)}. \tag{6.45}$$

The following lemma is used to bound this $\mathcal{H}_2$-norm density:

Lemma 6.10 *Consider the feedback operators F and G in the static vehicular formation problem* (6.11)*, and assume they are admissible. It holds $\hat{f}_\infty(\theta) \sim -\beta(\theta_1^2 + \ldots + \theta_d^2)$. If G has the relative measurement property* (6.8)*, then also $\hat{g}_\infty(\theta) \sim -\beta(\theta_1^2 + \ldots + \theta_d^2)$. Otherwise, $\hat{g}_\infty(\theta) \sim \hat{g}_0$, for some constant $\hat{g}_0$.*

Proof: By Assumption 6.5, $\hat{f}_\infty(\theta), \hat{g}_\infty(\theta)$ are real valued. If they satisfy the relative measurement property (6.8), they therefore have the same properties as $\mathrm{Re}\{\hat{f}_\infty(\theta)\}$ from the consensus case, and scale as in (6.37). If G has absolute feedback, it follows from (6.38) that $\hat{g}_\infty(\theta) = \hat{g}_0 - \sum_{k\in\mathbb{Z}^d} g_k\,(1-\cos(\theta\cdot k))$, where $\hat{g}_0 = \sum_{k\in\mathbb{Z}^d} g_k < 0$. Due to Assumption 6.2 of locality, this number is uniformly bounded for all $\theta \in \mathcal{R}^d$. We can thus write $\hat{g}_\infty(\theta) \sim \hat{g}_0$. ■

In the case of only *relative* feedback, Lemma 6.10 bounds the $\mathcal{H}_2$-norm density from (6.45) as

$$\hat{p}(\theta) = \frac{d}{2\hat{f}_\infty(\theta)\hat{g}_\infty(\theta)} \sim \frac{1}{\beta^2(\theta_1^2 + \ldots + \theta_d^2)^2}. \tag{6.46}$$

The per-site variance thus scales as in Lemma 6.3 with $r = 4$.

With *absolute velocity* feedback we instead get that

$$\hat{p}(\theta) \sim \frac{1}{\beta\,(\theta_1^2 + \ldots + \theta_d^2)}.$$

In this case, the per-site variance thus scales as in Lemma 6.3 with $r = 2$.

We can also note that relaxing Assumption 6.4 and allowing absolute feedback from both position and velocity would let $\hat{f}_\infty(\theta) \sim \hat{f}_0$ and $\hat{g}_\infty(\theta) \sim \hat{g}_0$, making the $\mathcal{H}_2$-norm density (6.45) uniformly bounded in θ. That is, $r = 0$ in Lemma 6.3 and the system would be fully coherent.

The results for the static case outlined above, which are in line with those in [14, Table 1], are summarized in Theorem 6.1.

6.5.4 Vehicular formations: performance with dynamic feedback

Now, consider the vehicular formation system with dynamic feedback. Provided the feedback configuration is admissible, the Lyapunov equation (6.28) gives

$$\hat{p}(\theta) = \frac{d}{2\hat{f}_\infty(\theta)\hat{g}_\infty(\theta) + 2\varphi^v(\theta)}, \tag{6.47}$$

where $\varphi^v(\theta)$ is a function of the Fourier symbols of the operators A, B, C, F and G. We now analyze (6.47) for the case with both relative and absolute velocity feedback.

Relative feedback

In order for the $\mathcal{H}_2$-norm density in (6.47) to scale differently from the static case (6.45), the function $\varphi^v(\theta)$ would need to scale differently in θ from the product $\hat{f}_\infty(\theta)\hat{g}_\infty(\theta)$, whose scaling was established in (6.46). This is, however, not possible with only relative feedback:

Lemma 6.11 *For any admissible choice of the operators A, B, C, F, G in (6.13) with only relative feedback in B, C, F, G, the function $\varphi^v(\theta)$ in (6.47) is such that*

$$\hat{f}_\infty(\theta)\hat{g}_\infty(\theta) + \varphi^v(\theta) \sim \beta^2(\theta_1^2 + \ldots + \theta_d^2)^2, \tag{6.48}$$

Therefore, the $\mathcal{H}_2$-norm density $\hat{p}(\theta)$ in (6.47) will satisfy (6.33) with $r = 4$ for any design of the dynamic feedback.

Proof: See Appendix 6.G.

We conclude that in the case of only relative feedback, dynamic feedback on the form (6.13) *cannot* improve the asymptotic performance scaling compared to static feedback.

Remark 6.10 Certain choices of A, B, C, F, G in (6.13) may appear as though one can achieve $\varphi^v(\theta) \sim -\beta(\theta_1^2 + \ldots + \theta_d^2)$, and thereby improve performance. For example, if $A = 0$, it holds $\varphi^v(\theta) = \hat{b}_\infty(\theta) + \hat{c}_\infty(\theta)\hat{g}_\infty(\theta)$ and one may wish to set B as the standard consensus operator (6.3), thus obtaining $r = 2$. However, by Theorem 6.7, such a choice is inadmissible. □

Absolute velocity feedback

In this case, we first consider the distributed-averaging proportional-integral (DAPI) controller (6.12) for the 1-dimensional vehicular platoon. The solution to the Lyapunov equation yields

$$\hat{p}_{\text{DAPI}}(\theta) = \frac{1}{2\hat{f}\hat{g} - 2\frac{\hat{c}\hat{f}(\hat{a}+\hat{g})}{\hat{a}^2+\hat{g}\hat{a}-\hat{f}}}, \tag{6.49}$$

where we have left out the $\infty-$subscript and the arguments of the individual Fourier symbols for notational compactness.

Now, A and F in DAPI are standard consensus operators whose Fourier symbols look like (6.23), while $G = -g_o I$ and $C = -c_o I$, which gives $\hat{g}_\infty(\theta) = -g_o$, and $\hat{c}_\infty(\theta) = -c_o$. Inserting into (6.49) yields (after some simplifications):

$$\hat{p}_{\text{DAPI}}(\theta) = \frac{1}{4g_o f_+(1-\cos\theta) + 2\frac{c_o g_o f_+ + 2c_o f_+ a_+(1-\cos\theta)}{f_+ + a_+ g_o + 2a_+^2(1-\cos\theta)}},$$

which recognize as being uniformly bounded in $\theta \in \mathcal{R}$. The same conclusion holds for $d > 1$. This implies that the DAPI controller, as we already found in Chapter 4, yields *full coherence*.

If absolute velocity measurements are available, several designs of the dynamic feedback in (6.13) can be shown to give the same result as the DAPI controller. In particular, G and C can also include relative feedback and B can be non-zero.

The asymptotic performance scalings for the vehicular formation problem with dynamic feedback are summarized in Theorem 6.1, where they have been re-written in terms of total network size $N = L^d$.

6.5.5 Control effort bounds

We have introduced the algorithm parameter $\beta = \max\{||f||_\infty, ||g||_\infty\}$ in the derivations above. This parameter affects the performance scaling, as evident from our main result in Theorem 6.1. In particular, if β were allowed to increase unboundedly, full coherence could be achieved in any spatial dimension. This is not feasible in any realistic control problem, where the amount of control effort is bounded. We now show that the size of the feedback array elements and therefore β are bounded by the total control effort at each network site. We quantify the control effort through:

$$\mathbb{E}\{u_k^* u_k\}, \tag{6.50}$$

that is, the variance of the control signal at each network site. In [14, Lemma 5.1], such bounds are presented for the case of static feedback. Here, we present bounds for the dynamic feedback case, but limit the analysis to the consensus algorithm with symmetric feedback for the sake of brevity:

Lemma 6.12 *Consider the consensus problem with dynamic feedback* (6.6)*, where the feedback operators A, B, F satisfy Assumptions 6.2 and 6.5. The following bounds hold:*

$$\mathbb{E}\{u_k^* u_k\} \geq \frac{1}{2}||f||_\infty \tag{6.51a}$$

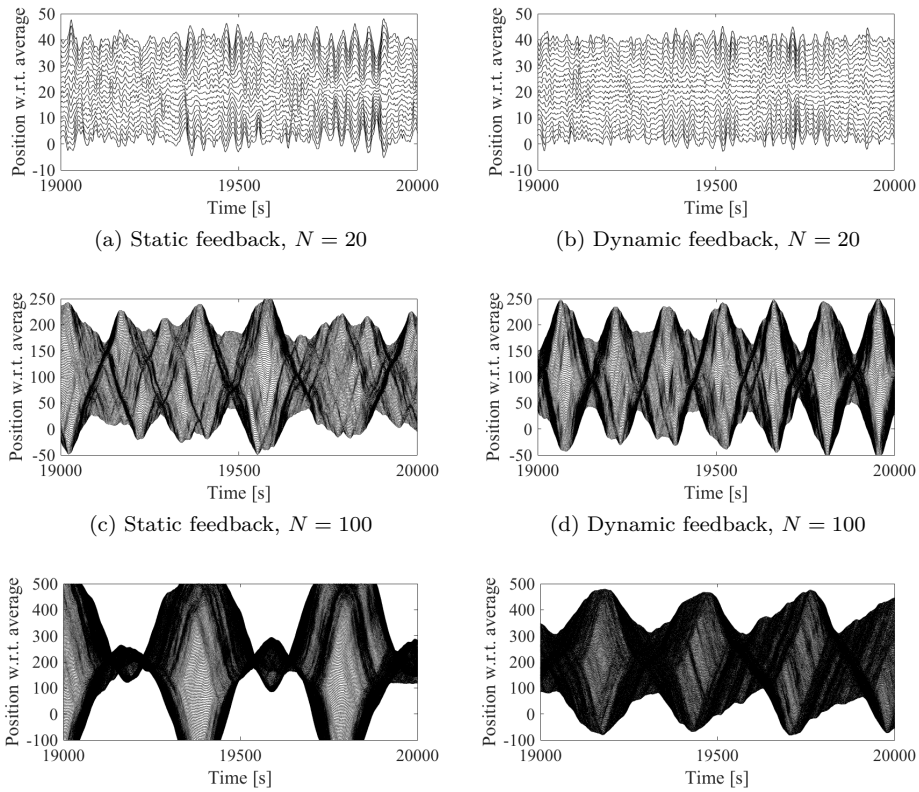
$$\mathbb{E}\{u_k^* u_k\} \geq \sqrt{\left(\frac{||a||_\infty}{4}\right)^2 + \frac{||b||_\infty}{4(2q)^d}} - \frac{||a||_\infty}{4} \tag{6.51b}$$

Proof: See Appendix 6.H.

Note that the constants in the bounds are independent of network size. Since we have set $\beta = ||f||_\infty$ and $||a||_\infty, ||b||_\infty \sim \beta$, we can conclude that the asymptotic scalings for the consensus problem in Theorem 6.1 will apply to any algorithm with control effort bounds.

6.6 Implications and numerical example

The performance limitations discussed in this chapter are in terms of scalings of global $\mathcal{H}_2$ performance, with respect to an output defined through nodal state fluctuations. We argued that a better scaling implies that the network remains more coherent, or rigid, when subjected to a process noise disturbance. Figure 6.4 shows simulations of strings of vehicles (that is, platoons) with both static and dynamic feedback from relative measurements. As the platoons grow, they exhibit an increasing lack of coherence. This is manifested through slow and large-scale fluctuations of the platoon length, clearly indicating that the platoon does not move like a rigid body. While the shape and size of these fluctuations are different with dynamic feedback compared to static, the relative performance deterioration



(a) Static feedback, $N = 20$

(b) Dynamic feedback, $N = 20$

(c) Static feedback, $N = 100$

(d) Dynamic feedback, $N = 100$

(e) Static feedback, $N = 200$

(f) Dynamic feedback, $N = 200$

Figure 6.4: Simulation of an N-vehicle platoon with static feedback (6.11) and dynamic feedback (6.13) from relative measurements. At each time step of 0.1 s the independent inputs w_k are sampled from a Gaussian distribution. We display the time trajectories of all vehicles' positions, with the average motion of the platoon subtracted and a reference spacing $\Delta_x = 2$ units inserted between vehicles. Under perfect control, the trajectories would be N straight horizontal lines separated by Δ_x. Note that the times displayed are 19000 s $\leq t \leq$ 20000 s (approx. steady state), and that the scales on the vertical axes are proportional to N. The platoon exhibits an accordion-like motion for large N with both static and dynamic feedback, showcasing the lack of coherence predicted by Theorem 6.1.

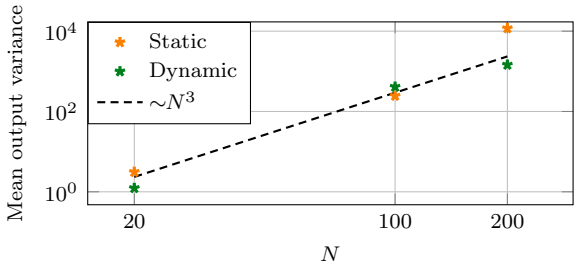


Figure 6.5: Mean variance (over the platoon) of the performance output (6.14) for the system trajectories displayed in Figure 6.4. The data points agree with the $V_N \sim N^3$ scaling predicted by Theorem 6.1.

is similar, as predicted by Case 2a in Theorem 6.1. This can also be seen from the corresponding variances displayed in Figure 6.5.

The origin of these slowly varying mode shapes in vehicle platoons was discussed in [14] and more recently in [127]. Our introduced concept of per-site $\mathcal{H}_2$-norm density provides additional insights. The $\mathcal{H}_2$-norm density is largest near $\theta = 0$, revealing that the low spatial frequency modes are most energetic (see Figure 6.3). As these correspond to the smallest system eigenvalues, they are also temporally slow. This results in slowly varying modes that have long spatial wavelengths and therefore span across the entire platoon.

Our derivations are made for spatially invariant systems, that is, lattices with periodic boundary conditions. The simulation here, however, is done for a string of vehicles where the first is not connected to the last. For large platoons, the boundary condition has a limited effect on the interior of the network. The corresponding simulation for a ring of vehicles can indeed be verified to have a very similar appearance to Figure 6.4.

While the relation (6.17) does not hold if the assumption of spatial invariance is relaxed, the quantity V_N can be evaluated as the mean variance over the network. This is also what is displayed in Figure 6.5. Through graph embedding (that is, noting that the string can be embedded in a ring graph) [18] it is straightforward to show that the mean variance for the string will be *at least as large* as for the ring graph case. It is therefore subject to the same limitations. Embedding arguments can also be made in higher spatial dimensions, in particular to relate the performance of 2D lattices to networks described by planar graphs.

The simulation in Figures 6.4–6.5 also demonstrates why it is the *scaling* of the per-site variance, rather than its actual value for a given N, that is meaningful for describing fundamental limitations. Even though a particular controller can achieve lower per-site variance in a given finite-size network (here, for example, the dynamic one at $N = 200$), the fact that it scales with network size implies that performance inevitably degrades as the network grows. This holds regardless of scaling coefficients. The result of Theorem 6.1 thus implies that neither static nor

dynamic feedback from only relative measurements is *scalable* to large networks. They are therefore both fundamentally limited.

6.7 Discussion

In this chapter, we have presented limitations to the achievable performance of distributed static and dynamic feedback in networked systems built over toric lattices. These limitations are described in terms of asymptotic scalings of $\mathcal{H}_2$ performance (that is, of upper and lower bounds) that are derived under a number of assumptions. In this section, we discuss some of these assumptions and possible generalizations along with some general implications of our results.

6.7.1 Non-regular networks

The results in this chapter are derived for systems defined on toric lattice networks, under certain restrictive assumptions. Accepting a generalization of the coherence metric V_N in (6.17) to a notion of mean variance, the assumptions of symmetry, uniformity in gains, and isotropy can be relaxed at the cost of analytic tractability, or by giving looser bounds on performance. Graph embeddings can, as already discussed, also be used to bound performance of more general networks through the lattices in which they can be embedded. The principle for this argument is that the removal of any network connection can only decrease the graph Laplacian eigenvalues (corresponding to the Fourier symbols in this study) and therefore increases V_N (see Section 3.2.5 for details). Any subgraph of a lattice (that is, any embedded graph) would thus have greater V_N than the lattice. We note that our theory allows for q neighbor connections in each lattice direction, making such embedding arguments less restrictive than they may seem.

Other concepts that are important for this chapter's results, such as locality, spatial dimension and a consistent notion of growing the network, are not straightforwardly generalized. For families of graphs where the behavior of the graph Laplacian eigenvalues (that is, the Laplacian eigenvalue density) is known, the ideas in Section 6.3 (for example, the $\mathcal{H}_2$-norm density) can be applied. A treatment of Laplacian eigenvalue densities in the limit of large networks is found, for example, in [2,50]. In this case, the notion of spatial dimension can possibly be generalized to one of *spectral* dimension. However, such considerations would only apply to graph families that can be scaled consistently, preserving properties like locality and dimensionality. Relevant contributions on performance limitations under alternative topological assumptions have been made in [62, 127, 129, 156]. The proper generalization of the topological properties that cause the limitations described here, however, remains an open research question.

6.7.2 Performance improvement with distributed integral control

As in Chapter 4, we established here that dynamic feedback such as the DAPI algorithm (6.12) can yield a fully coherent vehicular formation in any spatial dimension, provided that it has access to absolute measurements of velocities with respect to a global reference frame. The intuition behind this result is that absolute measurements of velocities are integrated in time to yield a substitute for absolute positions. With absolute feedback from both position and velocity, formations are known to be fully coherent [14]. However, as such a strategy is essentially so-called "dead reckoning," it can be sensitive to noisy measurements. In the next chapter, we will investigate how measurement noise affects the performance of this controller.

6.7.3 The temporal aspect

This chapter demonstrated unfavorable asymptotic scalings in network size of the *steady-state* variance of state fluctuations under a stochastic disturbance input. This is one of the interpretations of the $\mathcal{H}_2$ norm introduced in Section 3.1.3. Another aspect of these scalings that we did not consider here is the temporal one; the time to reach the steady state also scales in network size. For the consensus problem, the variance is an increasing function of time during the transient [190]. For very large, sparse networks, this means that the steady-state value may take a long time to reach, but it can be seen as a worst-case variance. A detailed analysis of the temporal aspects of performance scalings is a potential direction for future work.

6.7.4 Higher-order dynamic feedback controllers

The dynamic feedback controllers considered in this chapter all contain a single local memory state z, as illustrated in Figure 6.1. They thus describe a class of distributed proportional-integral (PI) controllers with respect to the system's states. While we show that this type of controller cannot improve performance scalings compared to static, memoryless controllers as long as they are limited to relative state feedback, it is an open question whether a higher-order controller, with an arbitrary number of local states, can.

The results from Chapter 5 reveal that increasing the order of local dynamics leads to admissibility limitations. Though it may be possible to impose conditions on the feedback similar to those in Theorem 6.7, it is likely increasingly difficult with a higher number of states (compare to Figure 5.1).

Even if an admissible controller with a higher number of auxiliary states can be designed, the limitation to relative state feedback implies that the marginally stable mode at the origin remains. Therefore, the Fourier symbol $\hat{\mathcal{A}}(\theta)$ is singular at $\theta = 0$. As a consequence, the $\mathcal{H}_2$-norm density will scale badly near $\theta = 0$ (consider the Lyapunov equation (6.28), and note that the right hand side is identity).

We therefore conjecture that the unfavorable scaling of performance in low spatial dimensions remains.

Performance scalings do, however, depend on the network's spatial dimension, with higher dimensions offering improved performance. It is an open and ongoing research question whether a higher-dimensional structure can be emulated through a high-order dynamic controller structure. Such an approach would, however, likely require that the controller order grows asymptotically along with the physical network, thereby creating a very high controller complexity.

Appendix to Chapter 6

6.A Scalings of sums and products

Many of the proofs in this appendix are based on the behaviors, or scalings, of functions of Fourier symbols in θ. Here we make some preliminary remarks on such scalings.

Recall that the notation $u(\theta) \sim v(\theta)$ implies $\underline{c}v(\theta) \leq u(\theta) \leq \bar{c}v(\theta)$ for all $\theta \in \mathcal{R}^d = [-\pi, \pi]^d$, where $\underline{c}, \bar{c}$ are fixed, positive constants. For example, we write $\hat{f}_\infty(\theta) \sim \beta(\theta_1^2 + \ldots + \theta_d^2)$, or $\hat{f}_\infty(\theta) \sim \beta|\theta|^2$ for short (the modulus $|\theta| = (\theta_1^2 + \ldots + \theta_d^2)^{1/2}$).

For products and sums of such functions, it holds $u'(\theta) = u_1(\theta)u_2(\theta) + u_3(\theta) \sim v_1(\theta)v_2(\theta) + v_3(\theta)$, implying that the bounds are $\underline{c}_1\underline{c}_2 v_1(\theta)v_2(\theta) + \underline{c}v_3 \leq u'(\theta) \leq \bar{c}_2 v_2(\theta)\bar{c}_2 v_2(\theta) + \bar{c}v_3(\theta)$. For a quotient: $u'(\theta) = u_1(\theta)/u_2(\theta) \sim v_1(\theta)/v_2(\theta)$ implies $(\underline{c}_1/\bar{c}_2)v_1(\theta)/v_2(\theta) \leq u'(\theta) \leq (\bar{c}_1/\underline{c}_2)v_1(\theta)/v_2(\theta)$.

Therefore, the scalings of functions of Fourier symbols can be determined simply by inserting the individual Fourier symbols' scalings. For example, if $\hat{f}_\infty(\theta) \sim \beta|\theta|^2, \hat{g}_\infty(\theta) \sim \beta|\theta|^2$, then $\hat{f}_\infty(\theta)\hat{g}_\infty(\theta) \sim \beta^2|\theta|^4$ and $\hat{f}_\infty(\theta)/\hat{g}_\infty(\theta) \sim 1$. This is used throughout to determine scalings of $\mathcal{H}_2$-norm densities.

6.B Maclaurin expansions of Z-transforms

The Maclaurin series expansions of Z-transforms will be used to derive admissibility conditions in Theorems 6.6–6.7. Consider an operator A, and its Z-transform $\hat{a}_\infty(\theta)$ given in (3.25). The Maclaurin expansion of $\hat{a}_\infty(\theta)$ in the coordinate direction $\theta = (\theta_1, 0, \ldots, 0)$ is

$$\hat{a}_\infty(\theta_1, 0, \ldots, 0) = \bar{a}_0 + \mathbf{j}\bar{a}_1\theta_1 + \bar{a}_2\theta_1^2 + \cdots . \tag{6.52}$$

Note that if A fulfills Assumption 6.3 of relative measurements, then $\bar{a}_0 = 0$. If A fulfills Assumption 6.5 of symmetry, then $\hat{a}_\infty(\theta)$ is real-valued and $\bar{a}_{1,3,\ldots} = 0$.

6.C Proof of Lemma 6.2

Given that $\hat{\mathcal{A}}_\infty(\theta)$ is Hurwitz for $\theta \neq 0$, the $\mathcal{H}_2$-norm density $\hat{p}(\theta)$ is continuous and bounded over the compact domain given by $\delta \leq |\theta_i| \leq \pi$ for $i = 1, \ldots, d$, and any fixed $\delta > 0$. It is therefore Riemann integrable on that domain.

On the interval $\Delta < |\theta_i| < \delta$, allowing for $\Delta \to 0$, $\hat{p}(\theta)$ will instead be monotonic. For simplicity, we show this through the scalar case in which $\hat{\mathcal{A}}_\infty(\theta) = \hat{a}_\infty(\theta) = \sum_{k\in\mathbb{Z}^d} a_k e^{-\mathbf{j}\theta\cdot k}$, which is negative for $\theta \neq 0$ as $\hat{\mathcal{A}}_\infty(\theta)$ is Hurwitz. Solving the Lyapunov equation (6.28) then gives $\hat{P}(\theta) = \hat{p}(\theta) = -1/2 \sum_{k\in\mathbb{Z}^d} a_k \cos(\theta \cdot k)$ for $\theta \in \mathcal{R}^d \backslash \{0\}$. Its derivative in each coordinate direction $i = 1, \ldots, d$ is

$$\frac{\mathrm{d}\hat{p}(\theta)}{\mathrm{d}\theta_i} = \frac{-2\sum_{k\in\mathbb{Z}^d} a_k k_i \sin(k_1\theta_1 + \cdots + k_d\theta_d)}{(2\sum_{k\in\mathbb{Z}^d} a_k \cos(k_1\theta_1 + \cdots + k_d\theta_d))^2}. \tag{6.53}$$

Now, note that $\mathrm{sgn}(\sin(kx)) = \mathrm{sgn}(x)$ for for$|x| \leq \frac{\pi}{k}$. Therefore, by the locality asssumption (6.7), the derivative (6.53) satisfies $\frac{\mathrm{d}\hat{p}(\theta)}{\mathrm{d}\theta_i} < 0$ for $\theta_i \in (0, \delta)$ and $\frac{\mathrm{d}\hat{p}(\theta)}{\mathrm{d}\theta_i} > 0$ for $\theta_i \in (-\delta, 0)$ with $\delta \geq \pi/q$. The $\mathcal{H}_2$-norm density $\hat{p}(\theta)$ is thus monotonically decreasing away from zero for $|\theta_i| \leq \delta$, where δ can always be fixed. A similar argument can be construed for when $\hat{\mathcal{A}}_\infty(\theta)$ is matrix-valued, in which case one considers matrix-valued coefficients of the Z-transform.

It is well-known that integrals of monotonic functions $f(x)$ can be estimated by upper and lower Riemann sums according to:

$$\int_m^{n+1} f(x)\mathrm{d}x \leq \sum_{k=m}^{n} f(k) \leq \int_{m-1}^{n} f(x)\mathrm{d}x$$

if $f(x)$ decreasing (and vice versa if $f(x)$ increasing). We use this to bound the *monotonic* part of the sum in (6.26):

$$V_N^\delta = \frac{1}{L_\delta^d} \sum_{\substack{\theta = \frac{2\pi}{L} n \\ |n_i| < \delta \frac{L}{2\pi}}} \mathrm{tr}\left(\hat{\mathcal{B}}_\infty^*(\theta)\hat{P}(\theta)\hat{\mathcal{B}}_\infty(\theta)\right) \tag{6.54}$$

by the integral from Δ to δ:

$$S^\delta(\Delta) := \int_{\Delta \leq |\theta_1| \leq \delta} \cdots \int_{\Delta \leq |\theta_d| \leq \delta} \hat{p}(\theta)\mathrm{d}\theta_1 \cdots \mathrm{d}\theta_d$$

as

$$S^\delta\left(\frac{4\pi}{L}\right) \leq V_N^\delta \leq S^\delta\left(\frac{2\pi}{L}\right), \tag{6.55}$$

since $2\pi/L$ and $4\pi/L$ are the first two wavenumbers, or sampling points in the sum. Here, L_δ is the number of summands for which $|n_i| < \delta\frac{L}{2\pi}$, corresponding to the domain where $\hat{p}(\theta)$ is known to be monotonic.

For the remainder of the sum, we use the Riemann integrability away from zero. That is, let

$$V_N^\pi = \frac{1}{(L - L_\delta)^d} \sum_{\substack{\theta = \frac{2\pi}{L} n \\ |n_i| \geq \delta \frac{L}{2\pi}}} \mathrm{tr}\left(\hat{\mathcal{B}}_\infty^*(\theta)\hat{P}(\theta)\hat{\mathcal{B}}_\infty(\theta)\right)$$

and note that $\lim_{L\to\infty} V_N^\pi = S^\pi$, where $S^\pi := \int_{\delta \leq |\theta_1| \leq \pi} \cdots \int_{\delta \leq |\theta_d| \leq \pi} \hat{p}(\theta)\mathrm{d}\theta_1 \cdots \mathrm{d}\theta_d$. In other words, the sum converges to the integral. Therefore, at some $\bar{L}$, we will have that $|V_{\bar{N}}^\pi - S^\pi| < S^\delta(\frac{2\pi}{\bar{L}}) - S^\delta(\frac{4\pi}{\bar{L}})$, so that

$$S^\delta\left(\frac{4\pi}{L}\right) + S^\pi \leq V_N^\delta + V_N^\pi \leq S^\delta\left(\frac{2\pi}{L}\right) + S^\pi,$$

for all $L \geq \bar{L}$, or $N > \bar{N}$, which is precisely equivalent to the statement of Lemma 6.2. ■

6.D Proof of Theorem 6.6

Each matrix $\hat{\mathcal{A}}_\infty(\theta) = \begin{bmatrix} \hat{a}_\infty(\theta) & \hat{b}_\infty(\theta) \\ 1 & \hat{f}_\infty(\theta) \end{bmatrix}$ has the characteristic equation

$$p(s,\theta) = s^2 + (-\hat{a}_\infty(\theta) - \hat{b}_\infty(\theta))s - \hat{b}_\infty(\theta) + \hat{a}_\infty(\theta)\hat{f}_\infty(\theta) = 0. \tag{6.56}$$

Going forward, we omit the $\infty-$subscript and the argument θ of the individual Fourier symbols for notational compactness. The system is input-output stable if and only if $\mathrm{Re}\{s_{1,2}(\theta)\} < 0$ for every $\theta \neq 0$ by Corollary 6.5. First note that if all operators are symmetric, then $\hat{b}, \hat{a}, \hat{f}$ are real-valued, and $\mathrm{Re}\{s_{1,2}(\theta)\} < 0$ if all coefficients of (6.56) are positive. This is achieved, for example, by setting $\hat{b}, \hat{a}, \hat{f} < 0$.

Otherwise, to find necessary conditions for stability under asymmeric feedback, it suffices to study this condition along one of the coordinate directions, so we let $\theta = (\theta_1, 0, \dots, 0)$. We use the Routh-Hurwitz criteria for complex-valued polynomials from the proof of Theorem 5.3. For the polynomial (5.7) with $n = 2$, the conditions $\Delta_2 < 0$, $\Delta_4 > 0$ for $\mathrm{Im}\{\mu\} > 0$ read: $g_1 < 0$, $f_1 g_1 g_0 - f_o g_1^2 - g_0^2 > 0$. Identifying the coefficients from (6.56) and transforming the condition to one for $\mathrm{Re}\{s\} < 0$ gives first that $\mathrm{Re}\{-\hat{a} - \hat{b}\} > 0$ (which is satisfied, for example, if $\mathrm{Re}\{\hat{a}\}, \mathrm{Re}\{\hat{b}\} < 0$), and second,

$$\begin{aligned} \mathrm{Re}\{\hat{a}+\hat{b}\}\mathrm{Im}\{\hat{a}+\hat{b}\}\mathrm{Im}\{-\hat{b}+\hat{a}\hat{f}\} + \Big(\mathrm{Re}\{\hat{a}+\hat{b}\}\Big)^2 \mathrm{Re}\{-\hat{b}+\hat{a}\hat{f}\} \\ - \Big(\mathrm{Im}\{-\hat{b}+\hat{a}\hat{f}\}\Big)^2 > 0 \end{aligned} \tag{6.57}$$

To determine whether this can be satisfied for all θ, and in particular the small θ that correspond to large N, we substitute the first terms of the Maclaurin expansions of the Fourier symbols $\hat{a}, \hat{b}, \hat{f}$ from (6.52) into (6.57):

$$\begin{aligned} &\underbrace{(\bar{a}_0 + (\bar{a}_2 + \bar{b}_2)\theta_1^2)(-\bar{b}_1\theta_1 + \bar{a}_0\bar{f}_1\theta_1 + (\bar{a}_1\bar{f}_2 + \bar{f}_1\bar{a}_2)\theta_1^3)(\bar{a}_1 + \bar{b}_1)\theta_1}_{m_1} \\ &\qquad + \underbrace{(\bar{a}_0\bar{f}_2 - \bar{a}_1\bar{f}_1 + \bar{a}_2\bar{f}_2\theta_1^2)(\bar{a}_2 + \bar{b}_2)^2\theta_1^6}_{m_2} \\ &\qquad\qquad - \underbrace{(-\bar{b}_1\theta + \bar{a}_0\bar{f}_1\theta_1 + (\bar{a}_1\bar{f}_2 + \bar{f}_1\bar{a}_2)\theta_1^3)^2}_{m_3^2} > 0, \end{aligned} \tag{6.58}$$

or in short $m_1 + m_2 - m_3^2 > 0$. Note that $m_3^2 > 0$, so $m_1 + m_2$ must be sufficiently large for (6.58) to be satisfied.

Assume first $\bar{b}_1 \neq 0$. In this case $m_3^2 \sim \theta_1^2$. We must then require $m_1 + m_2$ to scale as θ_1^p with $p \leq 2$, else $m_1 + m_2$ becomes too small for sufficiently small θ_1. While $m_2 \sim \theta_1^6$ or $\sim \theta^8$, we can have $m_1 \sim \theta_1^2$, but only if $\bar{a}_0 \neq 0$. We conclude that we must set

$$\bar{a}_0 \neq 0 \ \text{ if } \ \bar{b}_1 \neq 0.$$

This means that absolute feedback in A is a necessary condition for admissibility if B is asymmetric.

If $\bar{b}_1 = 0$, then if also $\bar{a}_0 = 0$, $m_3^2 \sim \theta_1^6$, while $m_1 \sim \theta_1^5$ or θ_1^6. If $\bar{a}_0 \neq 0$, we have $m_3^2 \sim \theta_1^2$ while $m_1 \sim \theta_1$ or $\sim \theta_1^2$, so it is possible to satisfy (6.58). Setting one out of $\bar{f}_1$ or $\bar{a}_1$ to zero will not affect these scalings (if they are both zero, $m_3 = 0$ and (6.58) is also satisfied). We conclude that

$$\bar{b}_1 = 0$$

is a necessary condition if $\bar{a}_0 = 0$, meaning that B must be symmetric if A has no absolute feedback. This concludes the proof. ■

Remark 6.11 An alternative proof for Theorem 6.6 is given in [176]. □

6.E Proof of Theorem 6.7

The characteristic polynomial of the matrix $\hat{\mathcal{A}}_\infty(\theta)$ is

$$p(s,\theta) = s^3 - (\hat{a}+\hat{g})s^2 + (\hat{a}\hat{g} - \hat{f} - \hat{c})s + \hat{a}\hat{f} - \hat{b}, \tag{6.59}$$

where we have again omitted the $\infty-$subscript and the argument θ of the Fourier symbols. Recall that all Fourier symbols are now real by Assumption 6.5. We can therefore use the standard Routh-Hurwitz stability criteria which state that; given a characteristic polynomial $p(s) = m_3 s^3 + m_2 s^2 + m_1 s + m_0$, then necessary and sufficient conditions for stability are that

(i) $m_i > 0$, and

(ii) $m_2 m_1 > m_3 m_0$.

In the case of (6.59), a necessary condition for satisfying (i) is that we do *not* have $\hat{a} = 0$, $\hat{b} = 0$ simultaneously. That is, if $B = 0$, then we must have $A \neq 0$. Otherwise, the condition (i) can easily be satisfied, for example by ensuring $\hat{a}, \hat{b}, \hat{c}, \hat{f}, \hat{g} < 0$. Assuming (i) is satisfied, consider (ii), which says that:

$$-(\hat{a}+\hat{g})(\hat{a}\hat{g} - \hat{f} - \hat{c}) > \hat{a}\hat{f} - \hat{b}.$$

First, we note that if $\hat{b} = 0$, then this reduces to $-\hat{a}^2\hat{g} - \hat{a}\hat{g}^2 + \hat{a}\hat{c} + \hat{g}\hat{f} + \hat{g}\hat{c} > 0$, which is also satisfied if $\hat{a}, \hat{c}, \hat{f}, \hat{g} < 0$. For the case where $\hat{b} \neq 0$, we follow the approach in the previous proof and expand the inequality with the first terms of the Maclaurin expansions along θ_1:

$$-(\bar{a}_0 + (\bar{a}_2 + \bar{g}_2)\theta_1^2)(\bar{a}_2\bar{g}_2\theta_1^4 + (\bar{a}_0\bar{g}_2 - \bar{f}_2 - \bar{c}_2)\theta_1^2) > -\bar{b}_2\theta_1^2 + \bar{a}_0\bar{f}_2\theta_1^2 + \bar{a}_2\bar{f}_2\theta_1^4 \tag{6.60}$$

Both sides of this inequality are positive if condition (i) above is satisfied. Now, if the RHS of (6.60) scales in lower powers of θ_1 than the LHS, then near $\theta_1 = 0$ it becomes arbitrarily many times larger than the LHS, and (6.60) cannot be satisfied. In particular, if $\bar{b}_2 \neq 0$, then the RHS scales as θ_1^2, which is only true for the LHS if $\bar{a}_0 \neq 0$. Therefore, absolute feedback in A is necessary if $B \neq 0$. This concludes the proof. ■

6.F Proof of Lemma 6.9

To prove Lemma 6.9 we treat the two admissible feedback configurations given by Theorem 6.6 separately.

Case a) B symmetric

If $\hat{b}_\infty(\theta)$ is real, then

$$\varphi^c = \frac{\hat{b}\mathrm{Re}\{\hat{a}\}(\mathrm{Re}\{\hat{a}\}+\hat{f})}{\hat{b}\hat{f}+\mathrm{Re}\{\hat{a}\}(\hat{b}-\mathrm{Im}\{\hat{a}\}^2-(\mathrm{Re}\{\hat{a}\}+\hat{f})^2)}. \tag{6.61}$$

We notice immediately, that if $\hat{a}=0$, that is, if $A=0$, then $\varphi^c \equiv 0$, and $\hat{f}+\varphi^c$ scales just as $\hat{f}$.

Otherwise, recall that $\hat{f} \sim -\beta(\theta_1^2+\ldots+\theta_d^2)$ (for short: $\hat{f} \sim -\beta|\theta|^2$) by Lemma 6.8. B now has the same properties as F, so $\hat{b} \sim -\beta|\theta|^2$. A on the other hand, may be asymmetric and have absolute feedback. Therefore, we in general have $\sum_{k\in\mathbb{Z}_L^d} a_k = \hat{a}_0$, where $\hat{a}_0 \leq 0$, and in line with (6.38) we obtain $\mathrm{Re}\{\hat{a}\} = \hat{a}_0 - \sum_{k\in\mathbb{Z}^d} a_k(1-\cos(\theta\cdot k))$, so $\mathrm{Re}(\hat{a}) \sim \hat{a}_0 - \beta\theta^2$. If A is asymmetric, the imaginary part of its Fourier symbol is $\mathrm{Im}\{\hat{a}\} = -\sum_{k\in\mathbb{Z}^d} a_k \sin(\theta\cdot k)$. Through similar calculations as in the proof of Lemma 6.8, we can derive the bound

$$\begin{aligned}\mathrm{Im}\{\hat{a}\}^2 \leq (2(2q)^d+1)\sum_{k\in\mathbb{Z}^d} a_k^2\sin^2(\theta\cdot k) &= ((2q)^d+\frac{1}{2})\sum_{k\in\mathbb{Z}^d} a_k^2(1-\cos(2\theta\cdot k)) \\ &\leq ((2q)^d+\frac{1}{2})(2q)^{(d+2)}||a||_\infty^2(\theta_1^2+\cdots+\theta_d^2).\end{aligned}$$

We can thus write $\mathrm{Im}\{\hat{a}\}^2 \leq \bar{c}_a\beta|\theta|^2$ with $\bar{c}_a \geq 0$. Clearly, it also holds $\mathrm{Im}\{\hat{a}\}^2 \geq 0$.

Now, consider the terms $\hat{b}-(\mathrm{Im}\{\hat{a}\})^2$ in the denominator of (6.61). By the arguments in Appendix 6.A, it holds $\hat{b}-\mathrm{Im}\{\hat{a}\}^2 \sim -\beta|\theta|^2$. Inserting this, together with $\hat{f},\hat{b} \sim -\beta|\theta|^2$, $\mathrm{Re}\{\hat{a}\} \sim \hat{a}_0 - \beta|\theta|^2$ into (6.61) gives

$$\varphi^c \sim \beta|\theta|^2\frac{-2\beta|\theta|^2+2\hat{a}_0}{2\beta^2|\theta|^4+\beta(1-3\hat{a}_0)|\theta|^2+\hat{a}_0^2}.$$

This can be written as $\varphi^c \sim -\bar{\varphi}\beta|\theta|^2$, and we note that $\bar{\varphi}$ will be a bounded, positive constant for any β and all $\theta \in \mathcal{R}^d$. In fact, $\bar{\varphi} \leq 2$ if $\hat{a}_0 = 0$, or $\bar{\varphi} \leq \frac{2}{|\hat{a}_0|}$ if $\hat{a}_0 < 0$. Therefore, $\hat{f}+\varphi^c \sim -\beta|\theta|^2 - \bar{\varphi}\beta|\theta|^2 \sim -\beta|\theta|^2$, which is precisely (6.44).

Case b) B asymmetric

If B is not symmetric, we must by Theorem 6.6 require A to have absolute feedback, so that $\hat{a}_\infty(\theta) \sim \hat{a}_0 < 0$. Inserting this into φ^c gives

$$\varphi^c \sim \frac{\hat{a}_0^2\mathrm{Re}\{\hat{b}\}+\mathrm{Im}\{\hat{b}\}^2+\hat{a}_0\mathrm{Re}\{\hat{b}\}\hat{f}}{\hat{a}_0(\mathrm{Re}\{\hat{b}\}-(\hat{a}_0+\hat{f})^2)+\hat{f}\mathrm{Re}\{\hat{b}\}}. \tag{6.62}$$

Now, $\mathrm{Im}\{\hat{b}\}^2$ satisfies the same inequality as $\mathrm{Im}\{\hat{a}\}^2$ above. Since $\hat{f} \sim -\beta|\theta|^2$, and $\mathrm{Re}\{\hat{b}\} \sim -\beta|\theta|^2$, the numerator terms from (6.62) satisfy $\hat{a}_0^2\mathrm{Re}\{\hat{b}\} + \mathrm{Im}\{\hat{b}\}^2 \sim -\hat{a}_0^2\beta|\theta|^2$, provided that $\hat{a}_0$ is sufficiently large to ensure admissibility.

Inserting all scalings into (6.62) gives

$$\varphi^c \sim \beta|\theta|^2 \frac{\hat{a}_0^2 + \hat{a}_0}{\beta^2(\hat{a}_0-1)|\theta|^4 + \beta\hat{a}_0(-2\hat{a}_0+1)|\theta|^2+\hat{a}_0^3} =: -\bar{\varphi}\beta|\theta|^2$$

Here, $\bar{\varphi}$ is a positive constant, which for any β and all $\theta \in \mathcal{R}^d$ satisfies $\bar{\varphi} \leq \frac{1}{|\hat{a}_0|}$. We can again conclude that $\hat{f} + \varphi^c \sim -\beta|\theta|^2$, which proves the lemma. ■

6.G Proof of Lemma 6.11

The function φ^v in (6.47) is given as

$$\varphi^v = \frac{\hat{b}^2 + \hat{b}(\hat{a}\hat{c} + \hat{c}\hat{g} - \hat{a}\hat{f} - \hat{a}\hat{g}^2 - \hat{a}^2\hat{g}) - \hat{c}\hat{f}\hat{a}(\hat{a} + \hat{g})}{\hat{b} - \hat{a}\hat{f} + \hat{a}^2(\hat{a} + \hat{g})} \tag{6.63}$$

Now, the feedback operators B, C, F, G have the relative measurement property (6.8), while A need not, so in line with Lemma 6.10, we have $\hat{b}, \hat{c}, \hat{f}, \hat{g} \sim -\beta|\theta|^2$ and $\hat{a} \sim \hat{a}_0 - \beta|\theta|^2$ with $\hat{a}_0 \leq 0$. We consider the two cases given by the admissibility Theorem 6.7 separately.

Case a) $B = 0$

Substituting the scalings of the individual Fourier symbols into (6.63) gives:

$$\varphi^v \sim \beta^2|\theta|^4 \frac{2\beta|\theta|^2 - \hat{a}_0}{2\beta^2|\theta|^4 + \beta(1 - 3\hat{a}_0)|\theta|^2 + \hat{a}_0^2} = \tilde{\varphi}\beta^2|\theta|^4.$$

For any β and for all $\theta \in \mathcal{R}^d$, we identify $\tilde{\varphi}$ as a positive constant, with $\tilde{\varphi} \leq \frac{1}{|\hat{a}_0|}$ if $\hat{a}_0 \neq 0$, $\tilde{\varphi} \leq 2$ if $\hat{a}_0 = 0$. Therefore, $\hat{f}\hat{g} + \varphi^v \sim \beta^2|\theta|^4 + \tilde{\varphi}\beta^2|\theta|^4 \sim \beta^2|\theta|^4$, which is precisely (6.48).

Case b) $B \neq 0$

If the operator B is nonzero, A is required by Theorem 6.7 to have absolute feedback, so $\hat{a}_0 < 0$. We can then set $\hat{a} \sim \hat{a}_0 < 0$ and:

$$\varphi^v \sim \beta^2|\theta|^4 \frac{\beta(1 - 2\hat{a}_0)|\theta|^2 + 2\hat{a}_0^2 - 1}{\beta\hat{a}_0(\hat{a}_0 + 1)|\theta|^2 - \hat{a}_0^3} = \tilde{\varphi}\beta^2|\theta|^4.$$

Again, $\tilde{\varphi}$ can be identified as a bounded positive constant, so $\hat{f}\hat{g} + \varphi^v \sim \beta^2|\theta|^4$ also in this case (provided $\hat{a}_0 \geq 1$, which signifies that the amount of absolute feedback in A is sufficient to guarantee admissibility).

It remains to consider the case in which the feedback operator $C = 0$. This does not give a meaningful control design if also $B = 0$, so it was not considered under case a) above. Substituting $\hat{c} = 0$ and the scalings of remaining Fourier symbols into (6.63) gives

$$\varphi^v \sim \beta^2|\theta|^4 \frac{2\beta^2|\theta|^4 - \beta(3\hat{a}_0+1)\theta^2 + \hat{a}_0^2 + \hat{a}_0 - 1}{2\beta^3|\theta|^6 + \beta^2(1-5\hat{a}_0)|\theta|^4 + \beta(4\hat{a}_0^2 - \hat{a}_0 + 1)|\theta|^2 - \hat{a}_0^3},$$

and the same conclusion as with $C \neq 0$ holds. ■

6.H Proof of Lemma 6.12

Consider the dynamics (6.6), but define the control signal $u = z + Fx$ as the output. We can then obtain the control signal variance in (6.50) through the $\mathcal{H}_2$ norm from w to u, divided by the total network size N. We use the DFT (3.24) to block-diagonalize the system, and solve a Lyapunov equation for each wavenumber n. We obtain that

$$\sum_{k\in\mathbb{Z}_L^d} \mathbb{E}\{u_k^* u_k\} = \sum_{n\in\mathbb{Z}_L^d\setminus\{0\}} \frac{\hat{b}_n - \hat{f}_n(\hat{a}_n + \hat{f}_n)}{2(\hat{a}_n + \hat{f}_n)},$$

which is equivalent to

$$N\mathbb{E}\{u_k^* u_k\} = \frac{1}{2}\left(||\hat{f}||_1 + ||\frac{\hat{b}}{\hat{a}+\hat{f}}||_1\right). \tag{6.64}$$

The equivalence of the sum and the l_1-norm follows from the fact that we must have $\hat{f}_n, \hat{b}_n < 0$ and $\hat{f}_n + \hat{a}_n < 0$ for all n in order to guarantee stability (see Theorem 6.4). Now, if $\hat{f}$ is the Fourier transform of a function array f, then $||\hat{f}||_\infty \leq ||f||_1$ and $||f||_\infty \leq \frac{1}{N}||\hat{f}||_1$ (see [14]). Inserting in (6.64) gives the first bound of the Lemma:

$$||f||_\infty \leq \frac{1}{N}||\hat{f}||_1 \leq 2\mathbb{E}\{u_k^* u_k\}.$$

It also holds that

$$2N\mathbb{E}\{u_k^* u_k\} \geq ||\frac{\hat{b}}{\hat{a}+\hat{f}}||_1 \geq \frac{||\hat{b}||_1}{||\hat{a}+\hat{f}||_\infty} \geq \frac{||\hat{b}||_1}{||\hat{a}||_\infty + ||\hat{f}||_\infty},$$

where the last equality follows from the triangle inequality. Now, we can use the fact that $||\hat{a}||_\infty \leq ||\hat{a}||_1 \leq (2q)^d||a||_\infty$ and substitute the bound above on $||\hat{f}||_\infty$ to get that

$$2N\mathbb{E}\{u_k^* u_k\} \geq \frac{||\hat{b}||_1}{(2q)^d(||a||_\infty + 2\mathbb{E}\{u_k^* u_k\})}.$$

Now, we use that $||b||_\infty \leq \frac{1}{N}||\hat{b}||_1$ to rewrite this as

$$4(2q)^d(\mathbb{E}\{u_k^* u_k\})^2 + 2(2q)^d||a||_\infty\mathbb{E}\{u_k^* u_k\} - ||b||_\infty \geq 0,$$

which leads to the second bound of the Lemma. ■

Chapter 7

Noise-Induced Limitations to Distributed Integral Control

In Chapters 4 and 6 of this thesis, we have shown that an addition of appropriately filtered distributed derivative or integral action can alleviate performance limitations that apply to distributed static feedback controllers. This holds in double-integrator networks in cases where the controllers have access to an absolute measurement of one of the two local states (position or velocity). In this chapter, we will examine this result closer, and focus on the implementation of integral control in the form of DAPI in the case where absolute velocity measurements are available.

Here, we investigate to which extent the superior performance of distributed integral control compared to standard distributed static feedback is robust to measurement noise in the controller. The apparent reason for the improved performance is namely that integration of the absolute velocity measurements emulates absolute position feedback. Any noise and bias in the velocity measurements is prevented from causing destabilizing drifts in this position feedback by a *distributed averaging filter* in the DAPI controller. Yet, we show here that noisy measurements may still have a significant impact on performance.

We consider the same setup as in Chapter 6, that is, we model the vehicular formation control problem over a toric lattice network. However, we consider asymptotic performance scalings not only in terms of the global error variance considered so far, but also of the local error variance that was also introduced as part of our initial problem formulation in Chapter 1. We show that while the performance of noiseless DAPI control scales well, the addition of measurement noise gives rise to its own contribution to the error variance with an unfavorable scaling in low-dimensional lattices. Even though this contribution, which is also proportional to the noise intensity, may be small in absolute terms for small networks, it limits the overall scalability of the controller. In fact, it becomes *worse* than with distributed static feedback.

These unfavorable scalings hold under the assumptions imposed in Chapter 6, in

particular, the assumptions of locality and of fixed feedback gains. Having observed that the distributed averaging filter in the DAPI controller reduces the impact of measurement noise, we study how a relaxation of those assumptions with regards to this filter may improve performance. That is, as the lattice size grows, we show how 1) increasing the size of each agent's neighborhood in the communication topology, and 2) increasing the distributed averaging gains affect the performance scalings. Our main conclusion is that these manipulations can restore the best-achievable performance scaling, which is shown to be that of distributed static feedback. That, however, requires an increasing amount of inter-nodal alignment through the distributed averaging filter. In a low-dimensional lattice, this eventually implies either very large distributed averaging gains that in principle require a centralized implementation, or all-to-all communication.

7.1 Noiseless vs. noisy distributed integral control

We will now consider the vehicular formation control problem from Chapter 6 and focus on the distributed integral controller DAPI. The network setting in the d-dimensional toric lattice along with all assumptions carry over to this chapter.

7.1.1 Static feedback

Recall from Section 6.1.3, that the vehicular formation control problem could be written

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & I \\ F & G \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{w} + \begin{bmatrix} 0 \\ I \end{bmatrix} u, \tag{7.1}$$

where F and G as before define convolution operators over $\mathbb{Z}_L^d$, and w is a process noise disturbance. Here, we have introduced the (secondary) control input u, which will exert the integral control action. We refer to the system (7.1) as subject to static feedback if $u = 0$.[1]

7.1.2 DAPI control under noisy measurements

Throughout this chapter, we assume absolute velocity measurements to be available, and therefore consider the use of DAPI control, which we write here as:

$$\begin{aligned} u &= z \\ \dot{z} &= -c_o v^m + Az, \end{aligned} \tag{7.2}$$

where z is the auxiliary controller state, v^m is the velocity *measured by the controller*, $c_o > 0$ is a fixed (integral) gain and A is a feedback operator subject to the

[1] Alternatively, any control on the form $u = F^u x + G^u v$, where F^u and G^u satisfy the same assumptions as F and G, is possible. Without loss of generality we can then assume $u = 0$ and absorb F^u and G^u in (7.1).

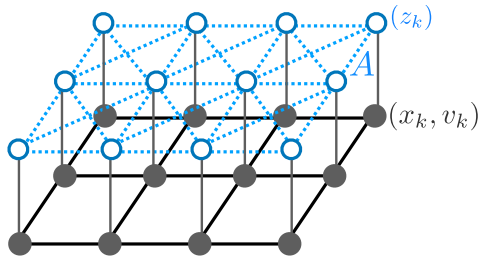


Figure 7.1: Example structure of the distributed integral control. The inter-nodal alignment of integral states z_k takes place over a communication network (dashed lines), while the state feedback interactions take place over the physical network (solid lines). Note that this schematic network is depicted without periodicity.

same assumptions as F, including the relative measurement and symmetry properties (see also Remark 7.1). Recall that the operator A represents the *distributed averaging filter*, over which an alignment of the integral states takes place. The purpose of this alignment is to prevent drifts in the integral states z_k (due to noise or bias), which would otherwise destabilize the system [6]. It is useful to think of the information exchange through A as taking place over a communication network layer, separate from the physical network. The setup is illustrated in Figure 7.1.

We have shown in Chapters 4 and 6 that the controller (7.2) improves performance compared to static feedback. This could be attributed to the fact that integration of absolute velocity measurements can provide a substitute for the otherwise lacking absolute position feedback. It turns out, however, that this result is sensitive to the accuracy of the absolute velocity measurements.

Here, let us therefore model additive measurement noise and let the velocity measurement in (7.2) be

$$v^m = v + \eta,$$

where the vector η contains uncorrelated white noise with the relative intensity ε defined through $\mathbb{E}\{\eta\eta^T\} = \varepsilon\mathbb{E}\{\mathrm{w}\mathrm{w}^T\}$. We will refer to the DAPI control as *noiseless* if $\varepsilon = 0$, and as *noisy* if $\varepsilon > 0$. Inserting this into (7.1) gives the closed-loop system:

$$\begin{bmatrix} \dot{z} \\ \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} A & 0 & -c_o I \\ 0 & 0 & I \\ I & F & G \end{bmatrix} \begin{bmatrix} z \\ x \\ v \end{bmatrix} + \begin{bmatrix} 0 & -c_o\varepsilon \\ 0 & 0 \\ I & 0 \end{bmatrix} \bar{\mathrm{w}}, \tag{7.3}$$

where $\bar{\mathrm{w}} \in \mathbb{R}^{2N}$ is a vector of uncorrelated white noise with unit covariance.

Remark 7.1 In this chapter, we will not consider DAPI control with absolute feedback in the operator A. This would namely stabilize the state z and no integral action would then be exerted on the system (7.1). □

Remark 7.2 We have assumed that the velocity enters without noise in the system dynamics (7.1). It may also be reasonable to model the same noise there, so that

$$\dot{v} = Fx + G(v + \eta) + u + \mathrm{w}.$$

This can, however, be shown not to affect the qualitative system behavior discussed here. See Appendix 7.A for a mathematical argument. □

7.2 Scalings of local and global performance

We will now summarize the performance of the vehicular formation control problem under static feedback and noiseless DAPI control, and compare it to the new results under noisy DAPI control. As usual, we consider performance in terms of the scaling of the per-site performance output variance V_N defined in (6.17). Here, we will, however, consider *two* performance outputs; one global and one local. The global error, or deviation from the network average, is defined as in previous chapters:

Definition 7.1 (Global error)

$$y_k = x_k - \frac{1}{N}\sum_{l\in\mathbb{Z}_L^d} x_l. \tag{7.4}$$

The local error is here defined as follows:

Definition 7.2 (Local error)

$$y_k = x_k - x_{k-1}. \tag{7.5}$$

As before, it is only if the per-site variance $V_N = \mathbb{E}\{y_k^*(t)y_k(t)\} = \frac{\mathbf{V}_N}{N}$ is bounded in the network size N that we can say that a control law is scalable to large networks. Consider the following result:

Result 7.1 (Performance scalings) *Let Assumptions 6.1, 6.2, and 6.4–6.6 hold. Table 7.1 lists the asymptotic scaling of the per-site variance V_N of the local and global error under*

(i) Static feedback, that is, the system (7.1) *with the input $u = 0$,*

(ii) Noiseless *DAPI control, that is, the system* (7.3) *with noise intensity $\varepsilon = 0$, and*

(iii) Noisy *DAPI control, that is, the system* (7.3) *with noise intensity $\varepsilon > 0$.*

Proof: The result (i) appeared in [14, Corollary 3.2], and (ii) follows from Corollary 4.3 (for the local error, see also Proposition 7.2 here). The result (iii) follows from the upcoming Proposition 7.2 and Corollaries 7.3–7.4. ■

Table 7.1: Asymptotic performance scalings for the system (7.1) with (i) $u = 0$ (static feedback), (ii). u as in (7.2) with $v^m = v$ (noiseless DAPI) and (iii) u as in (7.2) with $v^m = v + \eta$ (noisy DAPI). Scalings are up to a constant independent of network size N, algorithm parameter $\beta = \max\{||f||_\infty, ||g||_\infty\}$ and relative noise intensity ε.

	Local error	**Global error**
(i) Static feedback	$V_N \sim \frac{1}{\beta}$ for any d	$V_N \sim \frac{1}{\beta}\begin{cases} N & d=1 \\ \log N & d=2 \\ 1 & d\geq 3 \end{cases}$
(ii) *Noiseless* DAPI	$V_N \sim \frac{1}{\beta}$ for any d	$V_N \sim \frac{1}{\beta}$ for any d
(iii) *Noisy* DAPI	$V_N \sim \frac{1}{\beta} + \frac{\varepsilon^2}{\beta}\begin{cases} N & d=1 \\ \log N & d=2 \\ 1 & d\geq 3 \end{cases}$	$V_N \sim \frac{1}{\beta} + \frac{\varepsilon^2}{\beta}\begin{cases} N^3 & d=1 \\ N & d=2 \\ N^{1/3} & d=3 \\ \log N & d=4 \\ 1 & d\geq 5 \end{cases}$

Remark 7.3 The DAPI controller offers no improvement in terms of the scaling of the per-site variance of the *local* error compared to static feedback. In absolute terms, however, the variance is reduced (see Proposition 7.2 and note that $|\varphi(\theta)| > 0$). □

Result 7.1 reveals that the measurement noise η leads to an unfavorable scaling of both local and global error variance – even worse than with static feedback. This may not be an issue for small networks, as the variance is scaled by the factor ε^2, which can be very small (recall, ε represents the intensity of the measurement noise η relative to the process disturbance w). However, performance will deteriorate as the network size grows, thus limiting the scalability of the DAPI controller.

7.3 Improving the scalability of DAPI

Let us now consider a situation where the static feedback in (7.1) is fixed and the design of the DAPI controller (7.2) for performance is of interest. In Proposition 7.2 below, we show that the error variance consists of two terms due to, respectively, disturbances and measurement noise. For any given system of a fixed network size, it is possible to trade off these terms and to optimize the control design. We will return to this problem in Chapter 10. However, the unfavorable scaling of the error variance due to measurement noise sets fundamental limitations to the scalability of any such control design to large networks. A numerical example showcasing this issue is shown in Figure 7.2. The objective in this chapter, rather than to solve a

performance optimization problem for a given system, is to point to the underlying limitations.

However, we also analyze how a relaxation of some of the assumptions on the DAPI control design may improve performance scalings. We therefore allow for a re-tuning of the distributed averaging filter, which is parameterized through the operator A, subject to the network size. This allows us to derive network size-dependent conditions on A for improved scalability, and to discuss their implications in the limit of a large network.

To enable this analysis, we use the technical framework from Section 6.3. Consider the $\mathcal{H}_2$ norm density expressions in the proposition below.

Proposition 7.2 *The $\mathcal{H}_2$ norm density of the system* (7.3) *with respect to the* global *error measurement* (7.4) *is:*

$$\hat{p}(\theta) = \underbrace{\frac{1}{2\hat{f}_\infty(\theta)} \cdot \frac{d}{\varphi(\theta) + \hat{g}_\infty(\theta)}}_{\hat{p}^{\mathrm{w}}(\theta)} + \underbrace{\frac{\varepsilon^2}{\hat{a}_\infty(\theta)} \cdot \frac{c_o}{2\hat{f}_\infty(\theta)} \cdot \frac{d}{1 + \hat{g}_\infty(\theta)\varphi^{-1}(\theta)}}_{\hat{p}^{\eta}(\theta)}. \tag{7.6}$$

The $\mathcal{H}_2$ norm density with respect to the local *error measurement* (7.5) *is:*

$$\hat{p}(\theta) = \underbrace{\frac{\hat{l}(\theta)}{2\hat{f}_\infty(\theta)} \cdot \frac{d}{\varphi(\theta) + \hat{g}_\infty(\theta)}}_{\hat{p}^{\mathrm{w}}(\theta)} + \underbrace{\frac{\varepsilon^2}{\hat{a}_\infty(\theta)} \cdot \frac{c_o \hat{l}(\theta)}{2\hat{f}_\infty(\theta)} \cdot \frac{d}{1 + \hat{g}_\infty(\theta)\varphi^{-1}(\theta)}}_{\hat{p}^{\eta}(\theta)} \tag{7.7}$$

where

$$\varphi(\theta) = \frac{c_o(\hat{a}_\infty(\theta) + \hat{g}_\infty(\theta))}{\hat{a}_\infty^2(\theta) + \hat{g}_\infty(\theta)\hat{a}_\infty(\theta) - \hat{f}_\infty(\theta)},$$

and $\hat{l}(\theta) = 2(1-\cos\theta)$.

Here, $\hat{p}^{\mathrm{w}}(\theta)$ corresponds to the $\mathcal{H}_2$ norm density of the system with noiseless DAPI control and $\hat{p}^{\eta}(\theta)$ represents the contribution from the measurement noise.

Proof: The result follows from block-diagonalizing the system (7.3) through Fourier transforms in line with Example 6.1. For the global performance output (7.4) we have that $\hat{\mathcal{C}}_\infty(\theta) = \begin{bmatrix} 1 & 0 \end{bmatrix}$ for $\theta \neq 0$. For the local one (7.5) it holds $\hat{\mathcal{C}}_\infty^*(\theta)\hat{\mathcal{C}}_\infty(\theta) = \begin{bmatrix} \hat{l}(\theta) & 0 \\ 0 & 0 \end{bmatrix}$ and it is easily verified using (3.25) that $\hat{l}(\theta) = 2(1-\cos\theta)$. In both cases, $\hat{\mathcal{C}}_\infty(\theta = 0) = 0$ and the subsystem $(\hat{\mathcal{A}}_\infty(0), \hat{\mathcal{C}}_\infty(0))$ is therefore unobservable.

Solving the Lyapunov equation (6.28) for all $\theta \in \mathcal{R}^d \backslash \{0\}$ gives the $\mathcal{H}_2$-norm densities above. The contributions from the disturbance inputs w and η can be separated since they are uncorrelated. ■

The following corollaries lead to the results in Table 7.1.

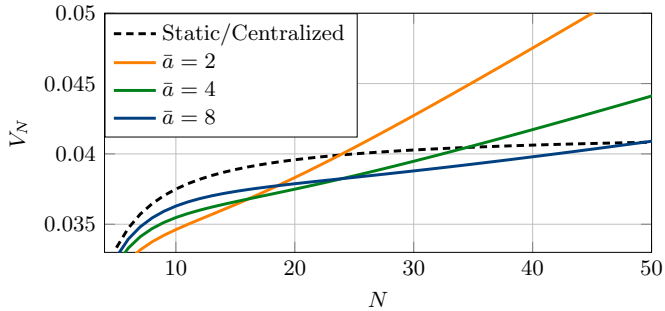


Figure 7.2: Scaling of local error variance with static feedback vs. noisy DAPI control in 1D lattice. For a given network size N, it is possible to set the gain $\bar{a}$ in the distributed averaging filter A so that the integral controller performs better than static feedback. Yet, no such controller scales well in N. A centralized integral controller on the form (7.8) that corresponds to $\bar{a} \to \infty$ will, however, have the same performance as static feedback for any N.

Corollary 7.3 *It holds that $\hat{p}^{\mathrm{w}}(\theta)$ is uniformly bounded with respect to both global and local error, that is, $r = 0$ in Lemma 6.3.*

Proof: Substituting the scalings from Lemma 6.10 into the expressions in Proposition 7.2 reveals that $\varphi(\theta) \sim 1/\beta|\theta|^2$, where the modulus $|\theta| := (\theta_1^2 + \ldots + \theta_d^2)^{1/2}$. The product $\hat{f}_\infty(\theta)\varphi(\theta)$ is thus bounded away from zero and the result follows. ■

Corollary 7.4 *It holds $\hat{p}^{\eta}(\theta) \sim \varepsilon^2/\beta|\theta|^4$, or $r = 4$ (global error), and $\hat{p}^{\eta}(\theta) \sim \varepsilon^2/\beta|\theta|^2$, or $r = 2$ (local error).*

Proof: Lemma 6.10 gives that $\varphi^{-1}(\theta) \sim \beta|\theta|^2$. Since $\hat{f}_\infty(\theta) \sim -\beta|\theta|^2$, $\hat{a}_\infty(\theta) \sim \bar{a}|\theta|^2$, the product $\hat{f}_\infty(\theta)\hat{a}_\infty(\theta) \sim -\beta|\theta|^4$ and the result follows. ■

7.3.1 From distributed to centralized integral control

Under the given assumptions, the performance scalings in Table 7.1 hold. With regards to the DAPI control design (7.2), this means for any fixed, finite gain c_o and any operator A with fixed, finite array elements subject to a locality constraint. We now inquire as to whether better scalings than in Table 7.1 can be achieved if these assumptions were relaxed. And if so, what requirements would this impose on the DAPI controller? To begin with, the following conclusions can be drawn from Proposition 7.2:

a. It is not possible to set $\hat{a}_\infty(\theta) = 0$, since in that case $\hat{p}^{\eta}(\theta) = \infty$.

b. If $\hat{a}_\infty(\theta) \to \infty$, or $c_o \to 0$, then $\hat{p}^\eta(\theta) \to 0$, that is, the noise contribution to the variance vanishes. At the same time, $\varphi(\theta) \to 0$ and $\hat{p}^{\mathrm{w}}(\theta)$ becomes as with static feedback (compare (7.6)–(7.7) to (6.45)).

c. If $\hat{a}_\infty(\theta)$ is bounded away from zero, then $\hat{p}^\eta(\theta) \sim \varepsilon^2/\beta|\theta|^2$ (global error) and $\hat{p}^\eta(\theta) \sim \varepsilon^2/\beta$ (local error).
At the same time, $\varphi(\theta)$ becomes uniformly bounded. Then $\hat{p}^{\mathrm{w}}(\theta) \sim 1/\beta|\theta|^2$ (global error), and $\hat{p}^{\mathrm{w}}(\theta) \sim 1/\beta$ (local error), that is, the same as with static feedback.

Using these observations, the following result is derived:

Proposition 7.5 *The best-achievable performance scaling for the noisy integral controlled system* (7.3) *is that of distributed static feedback in Table 7.1.*

Proof: First, note that for any fixed $c_o > 0$ and $\hat{a}_\infty(\theta)$, the scalings in Table 7.1 hold. For a better performance scaling, the behavior of $\hat{p}^\eta(\theta)$ in θ must change for the better (r in Lemma 6.3 must decrease). This can only happen if $c_o \to 0$, $\hat{a}_\infty(\theta) \to \infty$ or if $\hat{a}_\infty(\theta)$ becomes bounded away from zero. Since $\hat{p}^{\mathrm{w}}(\theta)$ and $\hat{p}^\eta(\theta)$ have inverse dependencies on the function $\varphi(\theta)$ in which both c_o and $\hat{a}_\infty(\theta)$ appear, this will lead to cases b and c above. ■

This means that the system can never have bounded variance in terms of the global error measurement (7.4). However, a bounded variance and thus scalability in terms of the local error (7.5) can be achieved by a re-tuning of the DAPI controller.

Based on cases b and c above, the best-achievable performance scaling can be retrieved by tuning the DAPI controller in three ways:

Decreasing the integral gain c_o

Decreasing the value of the gain c_o reduces the impact of the measurement noise η. To counteract the unfavorable scaling of $\hat{p}^\eta(\theta)$, it must be ensured that $c_o/\hat{a}_\infty(\theta)$ is uniformly bounded in θ. Since by Lemma 4, $\hat{a}_\infty(\theta) \sim -\bar{a}|\theta|^2$, this requires $c_o \sim \min|\theta|^2$.

The smallest wavenumber that contributes to the error variance in (6.26) corresponds to $\theta_{\min} = 2\pi/L$. This implies that c_o must be *decreased as* $1/L^2$.

As the network grows, this implies $c_o \to 0$ and the integral action is eliminated. In this case, the control input u is simply not used.

Increasing the distributed averaging gain

For a fixed c_o, the distributed averaging gain can be increased so that $\hat{a}_\infty(\theta)$ becomes bounded away from zero even as L increases. Recall that $\hat{a}_\infty(\theta) \sim -\bar{a}|\theta|^2$ where $\bar{a} = ||a||_\infty$. This need not approach zero if $\bar{a} \sim 1/|\theta|^2$.

Again, $\theta_{\min} = 2\pi/L$, meaning that $\bar{a}$ must be *increased as* L^2. This implies that we must require $\bar{a} \to \infty$ when the lattice size L grows.

While an infinite gain in distributed averaging is not feasible in practice, the same result can be realized as *centralized* averaging integral control (CAPI). Here, a central controller has instantaneous access to the integral states at all nodes. The control signal u_k is then the same for all $k \in \mathbb{Z}_L^d$:

$$\begin{aligned} u_k &= z; \\ \dot{z} &= \frac{1}{N} \sum_{k \in \mathbb{Z}_L^d} v_k^m. \end{aligned} \tag{7.8}$$

It is not difficult to show that this controller has the same performance with respect to the errors (7.4) and (7.5) as static feedback. We will elaborate on the CAPI controller and its performance in Chapter 10.

Increasing communication network connectivity

By relaxing Assumption 6.2 of locality for A, we can also bound $\hat{a}_\infty(\theta)$ away from zero. To see this, recall that $\hat{a}_\infty(\theta) = -\sum_{k \in \mathbb{Z}^d} a_k(1 - \cos(\theta \cdot k))$. Again, we know that $\theta_{\min} = 2\pi/L$, and therefore, the argument in the cosine $|\theta \cdot k| \geq 2\pi k/L$. This need not approach zero if we allow the feedback window $q_A := \max_{a_k \neq 0} |k|$ to grow as $q_A \sim L$. Then, $\hat{a}_\infty(\theta)$ stays bounded away from zero as $\theta \to 0$.

Allowing q_A to scale with the lattice size L in principle implies allowing an all-to-all connection in the communication network. For large networks this is practically challenging (if not impossible), and a centralized approach may be preferable.

7.3.2 Implications for distributed integral control

The previous section showed that the distributed averaging filter A in the DAPI controller (7.2) is important for performance. Recall that the role of the filter A is to align the controllers' integral states z_k across the network, in order to gain robustness to measurement noise and bias. Our results in Chapter 4, as well as those in related work [188], have indicated that "little" inter-nodal alignment (that is, small gains $\bar{a}$ and few interconnections in the communication network) is optimal for performance in the absence of measurement noise. It is intuitively clear that the inter-nodal alignment through A becomes increasingly important if measurement noise is considered explicitly.

These results, however, reveal that it is not enough to scale the distributed averaging gain $\bar{a}$ with the noise intensity, here parameterized through ε. Perhaps surprisingly, the need for inter-nodal alignment instead grows with the network size. It is required that the distributed averaging gain $\bar{a} \sim L^2 = N^{2/d}$, or that the feedback window $q_A \sim L = N^{1/d}$.

This means that large and sparse lattice networks require so much alignment that centralized integral control will in principle be necessary. So, while the distributed integral controller (7.2) is not scalable under noisy measurements, the centralized integral controller (7.8) is as scalable as distributed static feedback.

7.4 Discussion

7.4.1 Control design for finite size networks

Our focus has been to characterize limitations of distributed integral control in terms of the *scaling* of $\mathcal{H}_2$ performance to large networks. We showed that such limitations arise due to noisy measurements, and can only be alleviated by asymptotically increasing the amount of inter-nodal alignment between controllers.

Naturally, any real-world application will have a finite number of nodes, and the controller can thus always be tuned for acceptable performance. Our results imply, however, that such a tuning cannot be done independently of the network size. Therefore, even though the controller is implemented in a distributed fashion, its tuning requires global knowledge. In other words, our results prevent "plug-and-play" integral control.

7.4.2 General distributed integral control

A natural question to ask is whether alternative distributed integral control designs are also subject to the limitations of DAPI control. In general, we can consider any dynamic feedback from Chapter 6 and let: $u = z$, $\dot{z} = Az + Bx + Cv^m$, where A, B, C are feedback operators. Even with a controller on this form it is possible to show a result in line with Proposition 7.5.

With such a controller, one must be aware that the important property of *load sharing* is lost if $B \neq 0$. Load sharing implies that the control effort in regulating a constant disturbance is shared equally (or proportionally) between controllers. This is particularly important in electric power networks (where it is also referred to as power sharing). Therefore, even though a controller with $B \neq 0$ can achieve the same performance as static feedback, it may not be desirable in practice.

A similar issue applies if A contains absolute feedback. Again, the same performance as with static feedback can be achieved, but the integral action is also lost (see Remark 7.1).

The problem of scalability of distributed dynamic feedback controllers under noise is also addressed in [161], which investigated saddle-point methods implemented as distributed control algorithms. Similar to our results, they found that careful algorithm design is necessary to prevent unfavorable scalings of $\mathcal{H}_2$ performance. In particular, they find that an appropriately designed distributed algorithm can achieve the same performance scaling as the corresponding centralized one. To the best of our knowledge, however, the design of a fully scalable (in $\mathcal{H}_2$ sense) distributed feedback controller that satisfies the same objectives as DAPI control (elimination of stationary errors, load sharing) remains an open problem.

Appendix to Chapter 7

7.A Alternative noise model

Consider the case described in Remark 7.2, where the static feedback controller is subject to the same measurement noise as the DAPI controller. In this case, the $\mathcal{H}_2$-norm density with respect to the *global* error measurement can be verified to be:

$$\hat{p}(\theta) = \frac{1}{2\hat{f}_\infty(\theta)} \cdot \frac{d}{\varphi(\theta) + \hat{g}_\infty(\theta)} (1 + \varepsilon \hat{g}^2(\theta)) + \\ + \frac{\varepsilon^2}{\hat{a}_\infty(\theta)} \cdot \frac{c_o}{2\hat{f}_\infty(\theta)} \cdot \frac{d}{1 + \hat{g}_\infty(\theta)\varphi^{-1}(\theta)} \left(1 + \frac{2\hat{g}_\infty(\theta)\hat{a}_\infty(\theta)}{c_o}\right). \quad (7.9)$$

Comparing this to (7.6), we note that the term $\hat{p}^{\mathrm{w}}(\theta)$ has been multiplied by the factor $(1 + \varepsilon \hat{g}^2(\theta))$ and the term $\hat{p}^{\eta}(\theta)$ by the factor $\left(1 + \frac{2\hat{g}_\infty(\theta)\hat{a}_\infty(\theta)}{c_o}\right)$. Both of these factors are uniformly bounded away from zero in θ and do therefore not affect the scaling behaviors of $\hat{p}^{\mathrm{w}}(\theta)$ and $\hat{p}^{\eta}(\theta)$. The same situation applies to the case with the local error.

The $\mathcal{H}_2$ norm density and thus V_N does, however, increase slightly in absolute terms. This is to be expected since the system is subjected to additional noise.

Chapter 8

The Price of Synchrony: Resistive Losses in Synchronizing Power Networks

We now turn to a different type of networked dynamical system and study the problem of frequency control in power networks. In this chapter, we model a synchronous generator network as a system of coupled oscillators like in Example 1.2, and evaluate performance in terms of the resistive losses that are incurred in keeping this network at a synchronous state. These losses, which we refer to as transient losses, are in addition to the static losses associated with steady-state flows in the power network. They arise due to the transient power flow fluctuations that occur when the system is perturbed from a synchronous state by a small transient event or in the face of small stochastic disturbances. We term this performance measure the "Price of Synchrony", as it reflects the cost, in terms of real power losses, associated with lack of synchrony.

We show how the total network's transient power losses can be quantified using the $\mathcal{H}_2$ norm of a linear system of coupled swing equations with an appropriately defined performance output. This output accounts for local phase deviations analogous to the local errors defined in (1.10). We show that the total transient losses over the network will scale unboundedly with the network's size. Furthermore, they are shown to be only weakly dependent on network topology. It is therefore no longer true, as in Chapters 4–7, that increasing network connectivity improves performance.

These limitations imply that even though the power losses that arise during synchronization are typically a small percentage of the total real power flow, our results (based on a simplified network model) indicate that these losses may become significant as power networks evolve toward increasingly distributed systems. Furthermore, merely adding links to the network to increase connectivity is unlikely to alleviate the increases in transient power losses as the network grows. We illustrate these findings through a number of numerical examples.

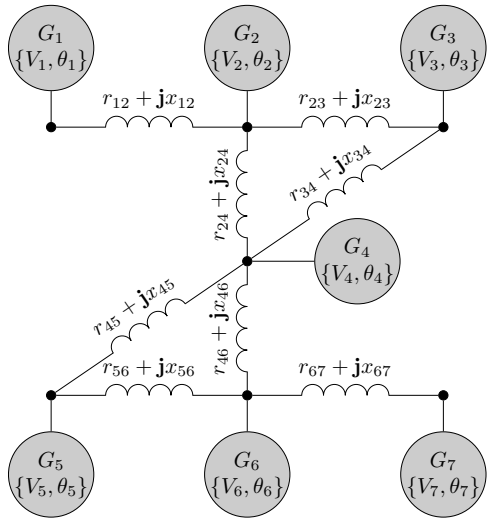


Figure 8.1: An example of a network of $N = 7$ generator nodes.

8.1 Formulation of coupled oscillator dynamics

8.1.1 Network model and swing equation

Consider a network described by the undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $|\mathcal{V}| = N$ is the total number of nodes, or *buses*. $\mathcal{E} = \{(i, j)\}$ represents the set of edges, or network lines. We will assume a Kron-reduced model of the power network (see, for example, [29, 39, 114, 124]) where the reduction procedure eliminates the constant-impedance loads and absorbs their effects into the network line models $\mathcal{E}$, along with any phase-shifting transformers. Thus, at every node $i \in \mathcal{V}$, there is a generator with inertia constant m_i and damping coefficient d_i. The voltage at node i is $v_i = V_i e^{\mathbf{j}\theta}$, where V_i is the voltage magnitude and θ_i the phase angle. The angular frequency is $\omega_i = \dot{\theta}_i$. Such a network is depicted in Figure 8.1 for a system where $N = 7$. In the absence of any external control input, the dynamics of the i^{th} generator can be described using the following classical machine model [123]:

$$m_i \ddot{\theta}_i + d_i \dot{\theta}_i = P_{m,i} - P_{e,i}, \tag{8.1}$$

where $P_{m,i}$ is the mechanical power input from the turbine and $P_{e,i}$ is the real power injected into the grid at node i, for which we will shortly provide an expression.

The swing equation (8.1) describes the physical acceleration or deceleration that arises in a synchronous generator as soon as there is a power imbalance and which, provided the network is stable, will allow the network to resynchronize.[1]

[1]This thesis will not include a detailed review of power system dynamics and stability. The interested reader will find an accessible introduction to the relevant concepts in [164, Chapter 2].

This control is *spontaneous*, in the sense that it is a consequence of the physics of the system [112]. Still, it will be useful for us to regard the swing equation as a type of control actuation ("droop control") in line with our prototypical problem formulation in Section 1.2.

8.1.2 The power flow equation

The real electric power flow injected to the grid at each node $i \in \mathcal{V}$ is given by

$$P_{e,i} \;=\; \bar{g}_i V_i^2 + \sum_{j\in\mathcal{N}_i} g_{ij} V_i\, V_j \cos(\theta_i - \theta_j) + \sum_{j\in\mathcal{N}_i} b_{ij} V_i\, V_j \sin(\theta_i - \theta_j), \tag{8.2}$$

where $\mathcal{N}_i$ denotes the neighbor set of node i in the Kron-reduced network $\mathcal{G}$. The parameters g_{ij} and b_{ij} are, respectively, the conductance and susceptance of the line (i,j) and $\bar{g}_i$ is the shunt conductance of node i (see Section 3.4.2 for their definitions). As per convention in power flow analysis, we assume that all quantities in (8.2) have been normalized by system constants and are measured in per unit (p.u.).

In what follows, we will use a simplified, linear model in which we consider small deviations from a stable operating point $\left[\theta^{\mathrm{ss}},\quad \omega^{\mathrm{ss}}\right]^T$. Without loss of generality we can transfer this operating point to the origin through a change of variables. This linearization allows us to investigate the effects of small disturbances or persistent small amplitude noise within a neighborhood of the operating point. The standard linear power flow assumptions include assuming constant voltage amplitudes, $V_i = 1$ p.u. for all $i \in \mathcal{V}$ and retaining only the linear terms in (8.2), which leads to

$$P_{e,i} \approx \sum_{j\in\mathcal{N}_i} b_{ij}\,(\theta_i - \theta_j)\,. \tag{8.3}$$

See for example [137] for a detailed analysis of the applicability of such assumptions. Having transferred the operating point to the origin, we may let deviations in the constant $P_{m,i}$ in (8.1) be absorbed into the disturbance input w_i, which characterizes, for example, fluctuations in generation and loads. Substituting (8.3) into (8.1) then gives

$$m_i\ddot{\theta}_i + d_i\dot{\theta}_i \approx -\sum_{j\in\mathcal{N}_i} b_{ij}\,(\theta_i - \theta_j) + \mathrm{w}_i. \tag{8.4}$$

Now, we can use the system's susceptance matrix $\mathcal{L}_B$ from (3.28) to rewrite this on vector form as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\theta\\ \omega\end{bmatrix} = \begin{bmatrix}0 & I\\ -M^{-1}\mathcal{L}_B & -M^{-1}D\end{bmatrix}\begin{bmatrix}\theta\\ \omega\end{bmatrix} + \begin{bmatrix}0\\ M^{-1}\end{bmatrix}\mathrm{w}, \tag{8.5}$$

where $\theta = \left[\theta_1, \theta_2, \ldots, \theta_N\right]^T$, $\omega = \left[\omega_1, \omega_2, \ldots, \omega_N\right]^T$, $M = \mathrm{diag}\{m_i\}$, and $D = \mathrm{diag}\{d_i\}$.

8.2 The Price of Synchrony performance measure

The question of interest here is not to characterize conditions for the stability of the power network, but to evaluate its performance in terms of losses associated with transient power flows. We therefore assume that the system matrices in (8.5) are such that the dynamics are stable around the equilibrium manifold for which all phases are equal. We now define the performance output that will be used to evaluate the real power losses arising from the fluctuating phase angle differences.

To define the relevant performance measure, we adopt the approach first presented in [12]. Consider the real power loss (the "copper losses") over the line (i, j), given by Ohm's law as

$$P_{ij}^{\text{loss}} = g_{ij}|v_i - v_j|^2,$$

where $v_i = V_i e^{\mathbf{j}\theta_i}$ denotes complex voltage. If we enforce the linear power flow assumptions and retain only the terms that are quadratic in the state variables, standard trigonometric methods give that $P_{ij}^{\text{loss}} \approx g_{ij}(\theta_i - \theta_j)^2$. Since the quantity $\theta_i - \theta_j$ represents a deviation from an operating point, P_{ij}^{loss} is equivalent to the power loss over the edge during the transient (the excursion from the operating point). The corresponding sum of instantaneous, transient resistive power losses over all links in the network can therefore be approximated as

$$\mathbf{P}^{\text{loss}} = \sum_{(i,j)\in\mathcal{E}} g_{ij}(\theta_i - \theta_j)^2. \tag{8.6}$$

We can now make use of the conductance matrix $\mathcal{L}_G$ from (3.28) to rewrite (8.6) as the quadratic form

$$\mathbf{P}^{\text{loss}} = \theta^T \mathcal{L}_G \theta. \tag{8.7}$$

Since $\mathcal{L}_G$ is a positive semidefinite graph Laplacian, it has a unique positive semidefinite square-root $\mathcal{L}_G^{1/2}$. We can therefore define an output of the system (8.5) as follows:

Definition 8.1 (Price of Synchrony performance output) The instantaneous resistive power losses in a network are measured as $\mathbf{P}^{\text{loss}} = y^T y$ with the performance output

$$y := \mathcal{L}_G^{1/2}\theta. \tag{8.8}$$

Note that this $\mathbf{P}^{\text{loss}}$ represents a squared sum of weighted differences in states between neighboring nodes. That is, a sum of squared *local* errors over the network. This is most easily seen from (8.6). The expected power losses can therefore be written as a sum over local error variances on the form (1.10).

For ease of reference we rewrite the state dynamics (8.5) and the output equation (8.8) together as the MIMO LTI system H:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\theta\\ \omega\end{bmatrix} = \underbrace{\begin{bmatrix}0 & I\\ -M^{-1}\mathcal{L}_B & -M^{-1}D\end{bmatrix}}_{\mathcal{A}}\begin{bmatrix}\theta\\ \omega\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ M^{-1}\end{bmatrix}}_{\mathcal{B}}\mathrm{w} \tag{8.9a}$$

$$y = \underbrace{\begin{bmatrix}\mathcal{L}_G^{1/2} & 0\end{bmatrix}}_{\mathcal{C}}\begin{bmatrix}\theta\\ \omega\end{bmatrix}. \tag{8.9b}$$

This system is a *linear-quadratic* approximation of the full nonlinear problem in the sense that the dynamics have been linearized around an equilibrium corresponding to the condition where the power flows in the system are balanced and all generators are operating at a nominal frequency. The instantaneous resistive power losses are quadratically approximated by the squared Euclidean norm of the output signal y. We next describe scenarios that allow these losses to be evaluated through the system's input-output $\mathcal{H}_2$ norm.

Remark 8.1 The system (8.9a) represents linearized dynamics in which line resistances are not present in the first approximation, having been assumed small compared to the line reactances. The output (8.9b) represents a quadratic approximation of the power losses and measures the effect of non-zero line resistances, given the state trajectories arising from the system dynamics. A justification for these assumptions is given in Chapter 9, where line resistances are also accounted for in the dynamics. □

8.2.1 $\mathcal{H}_2$ norm interpretations for swing dynamics

The LTI system (8.9) is formulated so that the square of the Euclidean norm of the output $y^T(t)y(t)$ approximates the instantaneous resistive power loss at each time t. The $\mathcal{H}_2$ norm of this system can be interpreted as the average (per time t) power loss in a setting with persistent disturbances, or alternatively, as the total (over time) energy loss due to a transient event. These interpretations of the $\mathcal{H}_2$ norm relate to its standard interpretations given in Section 3.1. Here, we summarize physical scenarios which permit the input-output $\mathcal{H}_2$ norm of (8.9) to quantify resistive losses.

i. Response to a white stochastic input. When the input w is a white noise disturbance, the system's $\mathcal{H}_2$ norm is

$$||H||_2^2 \;=\; \lim_{t\to\infty}\mathbb{E}\{y^T(t)y(t)\}. \tag{8.10}$$

For the swing dynamics (8.9) the disturbance vector w can be thought of as persistent stochastic forcing at each generator. These disturbances, which are uncorrelated across generators, can be due to uncertainties in local generator

conditions, such as changes in local load or supplied mechanical power. The expression (8.10) is then exactly the expectation of the total (over the entire network) instantaneous power losses due to the system's maintaining synchrony.

ii. Response to a random initial condition. For the system (8.9) it holds that

$$\mathcal{B}\mathcal{B}^T = \begin{bmatrix} 0 & 0 \\ 0 & M^{-2} \end{bmatrix},$$

which is diagonal. Therefore, an initial condition ψ_o that is a random variable with zero mean and correlation $\mathbb{E}\{\psi_o \psi_o^T\} = \mathcal{B}\mathcal{B}^T$ corresponds to each generator having a random initial frequency perturbation that is uncorrelated across generators and zero initial phase perturbation. In this case, the $\mathcal{H}_2$ norm

$$\|H\|_2^2 \;=\; \int_0^\infty \mathbb{E}\{y^T(t)y(t)\}\ \mathrm{d}t \tag{8.11}$$

quantifies the total (over all time and the entire network) expected resistive energy loss due to the system returning to a synchronized state. The expectation is taken over the initial conditions.

iii. Impulse responses. Consider a scenario where impulses can be fed at each generator node with equal probability. For the system (8.9), this input scenario corresponds to each generator being subject to impulse force disturbances (note, the disturbance w enters into the momentum equation of each generator). Such disturbances could occur, for example, due to changed operation of the generator, a sudden lost load at the bus or a fault event. In these cases, the $\mathcal{H}_2$ norm

$$\|H\|_2^2 \;=\; N\mathbb{E}\left\{\int_0^\infty y^T(t)y(t)\ \mathrm{d}t\right\} \tag{8.12}$$

quantifies the total (over all time and the entire network) expected energy loss due to the system returning to a synchronized state. The expectation is taken over the set of generators.

8.2.2 Relations to network coherence

In Chapters 4, 6, and 7, we studied measures of *network coherence*, which in the present context translates to phase coherence, that is, the tightness of the phases of all generators. More precisely, the quantity

$$\mathbb{E}\left\{\left(\theta_i - \frac{1}{N}\sum_{j=1}^{N}\theta_j\right)^2\right\}, \tag{8.13}$$

expresses the variance of the deviation of the i^{th} generator's phase from the average over all generators in the network. This quantity is never zero when there are stochastic disturbance inputs.

Asymptotic scalings (in network size N) of the performance metric (8.13) were derived in Chapter 6 for systems with regular network structures. It was, among other things, found that performance depends strongly on the lattice dimension (see Theorem 6.1), which implies that more interconnected networks tend to be more coherent and vice versa.

While phase coherence is not directly related to the control objective of synchronization, the $\mathcal{H}_2$ norm used to quantify (8.13) is equivalent to a measure of *frequency* coherence under step disturbances, see Example 4.2. The same scaling behaviors therefore apply in both scenarios.

We note that our previous analyses considered performance as *per node*, while in the present context, it is the *total* or *aggregate* transient resistive power loss over the entire network that is of concern. Thus, although the two settings have analogous dynamics, the performance objectives differ. We point out that the coherence measure (8.13) is *not* the Euclidean norm of the output y defined in (8.9b). In other words, the amount of phase disorder in a network as measured by (8.13) is not necessarily related to resistive power losses, and in particular may not scale similarly with network size N or have the same dependence on network connectivity. While networks with high phase coherence may be desirable for other reasons (such as stability of the nonlinear model), the results to be presented shortly indicate that the Price of Synchrony (total transient resistive power losses) can be large even in highly coherent networks.

8.3 Evaluating resistive losses

In this section, we derive a formula for the $\mathcal{H}_2$ norm of the system (8.9) in terms of the system's parameters. We then consider the implications for some important special cases. Throughout this section we make the following simplifying assumption, which allows us to derive analytic expressions for the $\mathcal{H}_2$ norm:

Assumption 8.1 (Identical generators) All synchronous generators have identical inertia and damping coefficients, that is, $m_i = m$ and $d_i = d$ for all $i \in \mathcal{V}$. This gives $M = mI$ and $D = dI$.

8.3.1 System reduction

As previously discussed, $\mathcal{L}_G$ and $\mathcal{L}_B$ are weighted graph Laplacians and as such, they each have a zero eigenvalue, see Section 3.2. The zero eigenvalue implies that these matrices are singular and that the system (8.9) is not asymptotically stable. However, as before, the corresponding mode is not observable from the performance output y. This is shown formally for the system at hand in Appendix 8.A. Assuming the network to be connected, remaining eigenvalues are strictly in the left half of the complex plane, and the system therefore has a finite $\mathcal{H}_2$ norm.

In order to rigorously evaluate the $\mathcal{H}_2$ norm of (8.9) we perform a system reduction, or *grounding*, procedure that effectively removes the unobservable mode at

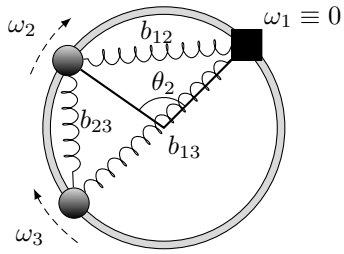


Figure 8.2: Mechanical analogy of the grounded swing dynamics in (8.14).

zero and enables us to investigate a reduced system that is asymptotically stable. We derive this reduced system by first defining a reference state $k \in \mathcal{V}$. By then deleting the k^{th} rows and columns of $\mathcal{L}_G$ and $\mathcal{L}_B$, we obtain the *grounded* Laplacians $\bar{\mathcal{L}}_G$ and $\bar{\mathcal{L}}_B$ (see Section 3.2.3). The states of the reduced system $\bar{\theta}$ and $\bar{\omega}$ are then obtained by discarding the k^{th} elements of each state vector. This leads to a system that is equivalent to one where $\theta_k = \omega_k \equiv 0$ for some node $k \in \mathcal{V}$. The physical interpretation of the reduced system is that the k^{th} node is connected to ground. A corresponding mechanical analogy is shown in Figure 8.2. We call our reduced, or grounded, system $\bar{H}$ and rewrite it as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\bar{\theta}\\ \bar{\omega}\end{bmatrix} = \underbrace{\begin{bmatrix}0 & I\\ -\frac{1}{m}\bar{\mathcal{L}}_B & -\frac{d}{m}I\end{bmatrix}}_{\bar{\mathcal{A}}}\begin{bmatrix}\bar{\theta}\\ \bar{\omega}\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ \frac{1}{m}I\end{bmatrix}}_{\bar{\mathcal{B}}}\bar{\mathrm{w}} \tag{8.14a}$$

$$\bar{y} = \underbrace{\begin{bmatrix}\bar{\mathcal{L}}_G^{\frac{1}{2}} & 0\end{bmatrix}}_{\bar{\mathcal{C}}}\begin{bmatrix}\bar{\theta}\\ \bar{\omega}\end{bmatrix}, \tag{8.14b}$$

where the reduced state vector is $\bar{\psi} = \begin{bmatrix}\bar{\theta} & \bar{\omega}\end{bmatrix}^T$. Assuming a network where the underlying graph is connected, the grounded Laplacians $\bar{\mathcal{L}}_G$ and $\bar{\mathcal{L}}_B$ are positive definite Hermitian matrices. All eigenvalues of the system $\bar{H}$ are thus strictly in the left half plane and the input-output system from $\bar{\mathrm{w}}$ to $\bar{y}$ has a finite $\mathcal{H}_2$ norm.

Remark 8.2 The dynamics of the grounded system (8.14) are not equivalent to those of the original system. In particular, the eigenvalues, and thereby the swing modes, differ. This can be understood intuitively by comparing Figure 8.2 to Figure 1.3. However, as we will show, the two systems are equivalent in terms of the performance metric considered here. □

8.3.2 $\mathcal{H}_2$ norm calculation

The squared $\mathcal{H}_2$ norm of the system $\bar{H}$ can be calculated through the equations (3.7)–(3.8). Here, we call denote the observability Gramian X and partition it into four

submatrices. The Lyapunov equation (3.8) expanded for the system (8.14) then becomes

$$\begin{bmatrix} 0 & -\frac{1}{m}\bar{\mathcal{L}}_B \\ I & -\frac{d}{m}I \end{bmatrix} \begin{bmatrix} X_1 & X_0 \\ X_0^T & X_2 \end{bmatrix} + \begin{bmatrix} X_1 & X_0 \\ X_0^T & X_2 \end{bmatrix} \begin{bmatrix} 0 & I \\ -\frac{1}{m}\bar{\mathcal{L}}_B & -\frac{d}{m}I \end{bmatrix} = - \begin{bmatrix} \bar{\mathcal{L}}_G & 0 \\ 0 & 0 \end{bmatrix},$$

from which we extract the following two equations:

$$X_0 - \frac{d}{m}X_2 + X_0^T - X_2\frac{d}{m} = 0 \tag{8.15a}$$

$$-\frac{1}{m}\bar{\mathcal{L}}_B X_0^T - X_0\frac{1}{m}\bar{\mathcal{L}}_B = -\bar{\mathcal{L}}_G. \tag{8.15b}$$

Then, using (8.15a) it is straightforward to compute $\frac{d}{m}\mathrm{tr}(X_2) = \mathrm{tr}(\mathrm{Re}\{X_0\})$. Equation (8.15b) can be rearranged to yield

$$\bar{\mathcal{L}}_B X_0^T \bar{\mathcal{L}}_B^{-1} + X_0 = m\bar{\mathcal{L}}_G\tilde{\mathcal{L}}_B^{-1},$$

where we make use of the fact that $\bar{\mathcal{L}}_B$ is nonsingular. Combining these expressions and using standard matrix trace relationships leads to the following expression

$$\mathrm{tr}\,(X_2) = \frac{m^2}{2d}\mathrm{tr}\left(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G\right). \tag{8.16}$$

Finally, noting that $\mathrm{tr}(\mathcal{B}^T X \mathcal{B}) = \frac{1}{m^2}\mathrm{tr}(X_2)$, we derive the following Lemma.

Lemma 8.1 *The squared $\mathcal{H}_2$ norm of the input-output mapping of the system* (8.14) *is given by*

$$||\bar{H}||_2^2 = \frac{1}{2d}\mathrm{tr}\left(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G\right), \tag{8.17}$$

where $\bar{\mathcal{L}}_B$ and $\bar{\mathcal{L}}_G$ are the grounded Laplacians obtained using the procedure described in Section 8.3.1, and d is each generator's self damping.

The choice of grounded node k has no influence on the $\mathcal{H}_2$ norm in (8.17). We illustrate this point through the following lemmas, which are then used to derive the main result of Theorem 8.4.

Lemma 8.2 *Let H denote the input-output mapping (8.9) under Assumption 8.1, and let $\bar{H}$ denote the corresponding reduced system (8.14). Then, the norm $||H||_2^2$ exists and*

$$||H||_2^2 = ||\bar{H}||_2^2.$$

Proof: See Appendix 8.A.

Lemma 8.3 *Let $\bar{\mathcal{L}}_G$ and $\bar{\mathcal{L}}_B$ be the reduced, or grounded, Laplacians obtained by deleting the k^{th} row and column of $\mathcal{L}_G$ and $\mathcal{L}_B$ respectively. Then*

$$\mathrm{tr}\left(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G\right) = \mathrm{tr}\left(\mathcal{L}_B^{\dagger}\mathcal{L}_G\right), \tag{8.18}$$

where † denotes the Moore-Penrose pseudo inverse.

Proof: See Appendix 8.B.

Remark 8.3 Lemma 8.3 extends to *any* pair of weighted graph Laplacians $\mathcal{L}_G, \mathcal{L}_B \in \mathbb{C}^{N\times N}$, even if they do not have the same underlying graph. It is, however, a requirement that the graph underlying $\mathcal{L}_B$ is connected (so that $\bar{\mathcal{L}}_B$ is nonsingular). This implies that the network grounding approach introduced here is amenable to any performance output that satisfies $y^T y = \theta^T \mathcal{L}_{G'}\theta$ where $\mathcal{L}_{G'}$ is some graph Laplacian. In particular, one could use this result to measure performance over a subnetwork of $\mathcal{G}$. □

Our main result can now be stated in the following theorem, which was also independently derived in [154].

Theorem 8.4 *Consider a network of N synchronous generators whose input-output response is given by (8.9). Under Assumption 8.1, the squared $\mathcal{H}_2$ norm of the system is*

$$\|H\|_2^2 = \frac{1}{2d}\mathrm{tr}\left(\mathcal{L}_B^{\dagger}\mathcal{L}_G\right). \tag{8.19}$$

Thus, the total transient losses of the system are a function of what we term the generalized Laplacian ratio *of $\mathcal{L}_G$ to $\mathcal{L}_B$.*

Proof: The result follows directly from Lemmas 8.1 - 8.3. ■

In (8.9), we assumed that the mechanical input $P_{m,i}$ to each generator i is lumped into the input w. If, instead, one chooses to scale the input by the generator's inertia, that is, define $\mathrm{w}' := \frac{1}{m}\mathrm{w}$ and $\mathcal{B}' := [0 \quad I]^T$, then the squared $\mathcal{H}_2$ norm of the resulting system can be derived in an analogous manner, as shown in the following Corollary.

Corollary 8.5 *Consider the modified input-output mapping*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\theta\\ \omega\end{bmatrix} &= \begin{bmatrix}0 & I\\ -\frac{1}{m}\mathcal{L}_B & -\frac{d}{m}I\end{bmatrix}\begin{bmatrix}\theta\\ \omega\end{bmatrix} + \begin{bmatrix}0\\ I\end{bmatrix}\mathrm{w}'\\ y &= \begin{bmatrix}\mathcal{L}_G^{1/2} & 0\end{bmatrix}\begin{bmatrix}\theta\\ \omega\end{bmatrix}.\end{aligned} \tag{8.20}$$

The squared $\mathcal{H}_2$ norm of this system is

$$\|H'\|_2^2 = \frac{m^2}{2d}\mathrm{tr}\left(\mathcal{L}_B^{\dagger}\mathcal{L}_G\right).$$

Proof: Following the proof of Lemma 8.1, we first note that for this modified system $\mathrm{tr}(\mathcal{B}'^T X \mathcal{B}') = \mathrm{tr}(X_2)$. The result then follows directly from Lemmas 8.2 and 8.3. ■

Theorem 8.4 states that the Price of Synchrony (the transient resistive power loss) is proportional to what can be thought of as a *generalized ratio* between the conductance and susceptance matrices. The ratio of line conductance to susceptance, or equivalently, resistance to reactance, is generally small for transmission systems and is therefore often neglected in power flow calculations [137] and stability analyses. However, the generalized ratio in equation (8.19) implies that the transient resistive losses will increase with network size (that is, the number of generators). Therefore, they may become significant in large networks with highly distributed generation even when line resistances are small. In low to medium voltage networks where distributed generation units are typically connected, the resistance-to-reactance ratios are higher than in transmission systems.[2] In such networks, this trend would be doubly problematic, as both the network size and the resistance-to-reactance ratio are larger. The next section explores these effects directly for the important special case of uniform resistance-to-reactance ratios across the network.

8.3.3 Special Case: Uniform line ratios

We now consider the special case when the generalized Laplacian ratio in (8.19) is a scalar matrix αI, where

$$\alpha := \frac{g_{ij}}{b_{ij}} = \frac{r_{ij}}{x_{ij}}.$$

In other words, all network lines $(i,j) \in \mathcal{E}$ have equal resistance-to-reactance ratios. This assumption implies that the conductance matrix is a scalar multiple of the susceptance matrix,

$$\mathcal{L}_G = \alpha \mathcal{L}_B, \tag{8.21}$$

and, of course, the same holds for the reduced Laplacians: $\bar{\mathcal{L}}_G = \alpha \bar{\mathcal{L}}_B$. In this case, the Lemmas 8.1 and 8.2 give that

$$||H||_2^2 = \frac{1}{2d}\mathrm{tr}(\bar{\mathcal{L}}_B^{-1}\alpha\bar{\mathcal{L}}_B) = \frac{\alpha}{2d}(N-1), \tag{8.22}$$

which is the result presented in [12]. This result is remarkable in that it says that, for this special case, the size of the transient losses depends only on the network's size and is entirely independent of its topology.

A choice of $\alpha_{\max} \geq \frac{g_{ij}}{b_{ij}}$ for all edges $(i,j) \in \mathcal{E}$, can be used to define a conservative bound based on (8.22). One can also define a lower bound $\alpha_{\min} \leq \frac{g_{ij}}{b_{ij}}$, and bound the $\mathcal{H}_2$ norm of the system as:

$$\frac{\alpha_{\min}}{2d}(N-1) \leq ||H||_2^2 \leq \frac{\alpha_{\max}}{2d}(N-1), \tag{8.23}$$

[2] Typically, this ratio is 1/16 in 400 kV lines but 2/3 in 11 kV systems [55].

where $\alpha_{\min(\max)}$ are the respective upper and lower bounds. These bounds, which were also proposed in [154], increase unboundedly with the number of generators and are independent of the network topology.

Motivation for uniform line ratio assumption

The equal line ratio assumption is not unreasonable for power systems, as the ratio of resistances to reactances of typical transmission links tend to lie within a small interval. This assumption is used in many related studies, such as [42, 96], and it can be motivated by a uniformity in the physical line properties in the network. The study [112] also found that the node degrees of Kron-reduced networks tend to be much more uniform than than those of the full power networks that they are derived from. Those results suggest that the "lines" of such reduced systems are also more uniform than those found in actual power networks and therefore the equal line ratio assumption is suitable for the reduced network considered here.

In Example 8.1, we demonstrate numerically that the result for uniform line ratios (8.22) gives a good approximation of (8.19) when is α is chosen as the average line ratio. In other words, the uniform line ratio assumption only leads to small errors with respect to the performance measure of interest. In Chapters 9–10 we will return to this assumption in order to make the developments and results more tractable.

8.3.4 Performance scaling and topology independence

In the notation from Chapter 6, the bounds expressed in (8.23) implies that the Price of Synchrony scales as

$$\|H\|_2^2 \sim \frac{1}{\beta} N, \tag{8.24}$$

where we can identify the algorithm parameter β with the generator damping d. We can therefore regard this result as a fundamental limitation to the performance of (Kron reduced, linearized) power networks, in which the swing equation (8.4) is the means by which synchrony is achieved. It is very important to note that the scaling (8.24) is independent of the network topology, unlike the coherence scalings in Theorem 6.1.

This topology independence is also in contrast to other measures of power system stability and performance metrics. For example, the topology of the system plays an important role in determining whether a system of this kind can synchronize [36,37,132,147]. The network connectivity of a power system is also directly related to its rate of convergence and damping properties [101]. Some intuition for this topology independence can be given as follows. We expect a highly interconnected network to have much more phase coherence than a loosely interconnected network with the same number of nodes. Consequently the power flows per link in a highly connected network are relatively small, but there are many more links than in the loosely connected network. Thus in the aggregate, the total transient power losses

are the same for both networks. A more coherent network is, however, in many ways more desirable for stability reasons.

Remark 8.4 The conclusion that total transient losses scale as in (8.24) is contingent on the modeling assumption that each generator is subject to a disturbance input w_i. If this is not the case, and only a subset $\mathcal{V}^{\mathrm{dist}}$ are subject to disturbances, then the total losses are instead proportional to the size of that subset $|\mathcal{V}^{\mathrm{dist}}| = N^{\mathrm{dist}}$. That is, if (8.21) holds, then $\|H\|_2^2 = \frac{\alpha}{2d}(N^{\mathrm{dist}} - 1)$. As long as the ratio N^{dist}/N remains approximately constant as the network size grows, the scaling in (8.24) will, however, hold. □

8.4 Generalizations and bounds

In this section, we provide some further bounds on the $\mathcal{H}_2$ norm in Theorem 8.4 and discuss their implications. We also address the more general case of systems with non-identical generators.

8.4.1 Loss bounds

As previously mentioned, the term $\mathrm{tr}(\mathcal{L}_B^\dagger \mathcal{L}_G)$ in Theorem 8.4 can be interpreted as a generalized ratio between the power network's conductance matrix $\mathcal{L}_G$ and its susceptance matrix $\mathcal{L}_B$. That is, a generalized ratio between the real and imaginary part of the network admittance matrix (without the shunt admittances). Let us henceforth denote the respective eigenvalues of $\mathcal{L}_G$ as $\lambda_N^G \geq ... \geq \lambda_2^G > 0$ and of $\mathcal{L}_B$ as $\lambda_N^B \geq ... \geq \lambda_2^B > 0$. The generalized ratio of these two Laplacians can then be lower bounded in terms of their eigenvalues as

$$\mathrm{tr}\left(\mathcal{L}_B^\dagger \mathcal{L}_G\right) \geq \sum_{i=2}^{N} \frac{\lambda_i^G}{\lambda_i^B}, \tag{8.25}$$

(see, for example, [194] for a proof). In the case of uniform line ratios, each eigenvalue ratio is equal to α, and the inequality in (8.25) turns into an equality. The loss scaling with network size in (8.24) can then also be seen from (8.25). This scaling is evident because the number of eigenvalues, and thus the number of summands, grows with each added node. We illustrate this growth and the tightness of the bound in (8.25) through Example 8.2.

The inequality (8.25) also provides some insight into why the $\mathcal{H}_2$ norm does not have a strong dependence on network connectivity even for networks of non-identical line ratios. Although the eigenvalues of the Laplacian are difficult to characterize precisely for general graphs, it is well known that they relate strongly to the node degrees (see, for example, [31,193]). In (8.25) however, we consider the ratio between the eigenvalues of $\mathcal{L}_G$ and $\mathcal{L}_B$. Since these are two graph Laplacians describing the same topology, their node degrees g_{ii} and b_{ii} can be related through $\bar{\alpha}_i$, the average ratio of line conductances to susceptances of the lines incident to

node i. This $\bar{\alpha}_i$ is independent of *how many* such incident lines there are. It is therefore reasonable to infer that each eigenvalue ratio λ_i^G/λ_i^B is also more strongly related to $\bar{\alpha}_i$ than to the number of lines connected to each node, which would be a measure the network connectivity. We will further explore this notion through the examples in Section 8.5.

The resistive losses can also be bounded as in (8.23), which is a simple way to express the performance scaling with network size. However, the bound may become loose if the system is heterogeneous in terms of the line resistance-to-reactance ratios. This may be the case if a combined transmission and distribution network is considered, or in cases of highly varying impedance loads. In some cases, it is then better to bound the losses in terms of graph-theoretical quantities. This can, for example, also be done in the following manner.

$$\lambda_2^G \mathrm{tr}(\mathcal{L}_B^\dagger) \leq \mathrm{tr}(\mathcal{L}_B^\dagger \mathcal{L}_G) \leq \frac{\mathrm{tr}(\mathcal{L}_G)}{\lambda_2^B}, \tag{8.26}$$

where λ_2^G and λ_2^B are the *algebraic connectivities* of the graphs weighted by, respectively, line conductances and susceptances. See [194] for a proof of (8.26).

In general, it holds that $\lambda_2^G \leq \frac{N}{N-1} g_{ii,\min}$ and $\lambda_2^B \leq \frac{N}{N-1} b_{ii,\min}$, where $g_{ii},\ b_{ii}$ are the respective node degrees in the weighted conductance and susceptance graphs. Furthermore, the quantity $\mathrm{tr}(\mathcal{L}_B^\dagger)$ is proportional to what we can interpret as the *total effective reactance* of the network, in analogy with the concept of graph total effective resistance, see Section 3.2.5. By Rayleigh's monotonicity law (Lemma 3.4), the total effective reactance can be decreased by adding lines and increasing line susceptances. However, the algebraic connectivity λ_2 is very small for weakly connected networks and often decreases with network size (see Section 5.2.3). Therefore, while the bounds (8.26) where λ_2^B appears in the denominator, can be accurate for small and well-connected networks, they become loose for large, sparsely interconnected networks. The usefulness of the bound (8.26), compared for example to (8.23), therefore depends on the network type. In Example 8.2, we evaluate these bounds for different network topologies.

8.4.2 Systems of non-identical generators

The results derived under Assumption 8.1, that is, by considering a network of identical generators, show that transient losses scale with the network size as in (8.24). In order to put these results into context it is desirable to understand the extent to which these scaling properties apply to systems of non-uniform generators. In this section we explore these ideas and use the results from previous sections to gain some insight. We begin by examining the special case where one non-uniform generator is added to the network.

From Theorem 8.4 we can deduce that

$$\frac{1}{2d_{\max}} \mathrm{tr}\left(\mathcal{L}_B^\dagger \mathcal{L}_G\right) \leq \|H\|_2^2 \leq \frac{1}{2d_{\min}} \mathrm{tr}\left(\mathcal{L}_B^\dagger \mathcal{L}_G\right), \tag{8.27}$$

where $d_{\min} = \min_{i\in\mathcal{V}} d_i$ and $d_{\max} = \max_{i\in\mathcal{V}} d_i$. The losses are thus lower and upper bounded by the properties of the most strongly and lightly damped generators respectively, and the overall scaling in (8.24) remains valid. Therefore, some interesting questions that arise from this observation are: (1) How does adding a generator to an existing network affect the total resistive losses? and (2) What are the important parameters in determining this incremental cost? The next result addresses one such scenario.

Lemma 8.6 *Consider a network of N generators with transient resistive losses given by $\|H_0\|_2^2 = \|\bar{H}^k\|_2^2$, where $\bar{H}^k$ is the corresponding reduced system with node k grounded. If one connects an additional generator with damping d_{N+1} and inertia m_{N+1} to node k by a single link with line ratio of $\alpha_{k,N+1} = \frac{r_{k,N+1}}{x_{k,N+1}}$, then the losses of the new, reduced, system are given by*

$$\|\bar{H}_1\|_2^2 = \|H_0\|_2^2 + \frac{1}{2d_{N+1}}\alpha_{k,N+1}.$$

If the dynamics are as per (8.20) *the additive term is instead $\frac{m_{N+1}^2}{2d_{N+1}}\alpha_{k,N+1}$.*

Proof: See Appendix 8.C.

This fairly intuitive result can be interpreted as follows. The additional losses incurred through the connection of a "light" (low inertia) or well damped generator are smaller than those incurred due to adding a "heavy" (high inertia) or poorly damped generator. In the face of increasingly distributed generation, this result implies that while the synchronization losses do scale with the network size, the impact of low inertia or small-scale distributed generators is relatively low, compared to that of large conventional generators.

In this thesis, we will not present further analytical results for systems with non-identical generators, but carry out most of the analysis under Assumption 8.1. The numerical Example 8.3, however, provides some further insight into performance of networks with non-uniform generation. That example shows that, although the scaling relationships and topology independence results hold for limited parameter variations, judicious sizing and placement of new generators can in fact improve performance.

8.5 Numerical examples

The results derived in this chapter show that the Price of Synchrony is highly dependent on the number of generators in the system. It was also found to depend on the system's resistance-to-reactance ratios and the generator properties, but only weakly on the network topology. In this section, we provide some numerical examples to illustrate these results and to explore more general networks.

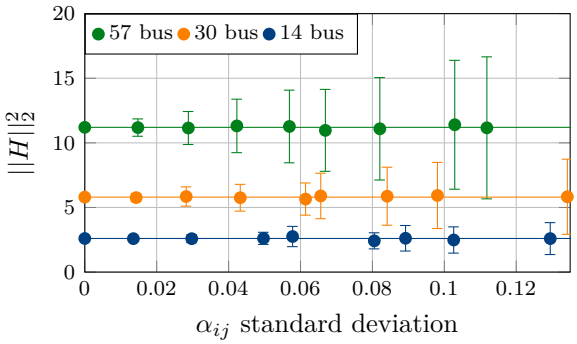


Figure 8.3: $\mathcal{H}_2$ norms from Example 8.1 for the IEEE 14 bus, 30 bus, and 57 bus benchmark networks. The edge weights are randomly generated using increasingly varying resistance-to-reactance ratios $\alpha_{ij} = \frac{r_{ij}}{x_{ij}}$. The bars illustrate the bounds in (8.23).

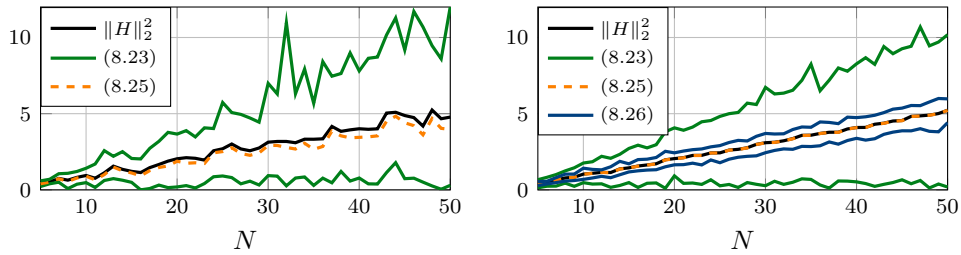
Example 8.1 (Non-uniform line ratios: increasing line ratio variance) We first investigate the Price of Synchrony in systems with increasingly non-uniform network line ratios α_{ij} (defined in Section 8.3.3). We consider a hypothetical set of identical generators with $d = 1$ placed at each node of, respectively, the IEEE 14 bus, 30 bus and 57 bus benchmark topologies [116]. We take the values for the reactances x_{ij} for each $(i, j) \in \mathcal{E}$ from the benchmark system data. We then define a random series of heterogeneous line ratios by setting $r_{ij} = \alpha_{ij} x_{ij}$, where the α_{ij} are each drawn once from uniform distributions on the following range of intervals: $0.4,\ 0.4 \pm 0.025,\ 0.4 \pm 0.05, ..., 0.4 \pm 0.2$.

Figure 8.3 shows the resistive losses computed from the result in Theorem 8.4 for a number of these systems. The horizontal axis indicates the standard deviation of the line ratios and the bars represent the upper and lower bounds of the inequality (8.23).

This example shows that increasing the standard deviation of the line ratios leads to a looser bound in (8.23). However, the actual resistive losses show only small variations as long as the average line ratio remains constant. We can also note that the transient losses depend strongly on the network size (here 14, 30 or 57 nodes), which is consistent with the relationship in (8.22).

Example 8.2 (Scalings and bounds for topology extremes) We now demonstrate the fact that the Price of Synchrony is only weakly dependent on network topology, but instead scales with the total number of network nodes. We compare the $\mathcal{H}_2$ norm in (8.19) and the bounds discussed in Section 8.4 for two systems where the underlying topology is (a) a path graph, and (b) a complete graph (see Figure 3.1) and let their respective system size increase.

These particular graph topologies are chosen because they represent the two



(a) Path graph (b) Complete graph

Figure 8.4: $\mathcal{H}_2$ norms from Example 8.2 for a (a) path graph and a (b) complete graph network of N nodes, with some of the bounds from in Section 8.4. Despite some variation due to the randomness in the line parameters, the $\mathcal{H}_2$ norm scales directly with the network size and is roughly the same for the path as for the complete graph. The bound related to the Laplacian eigenvalue ratios (8.25) is the tightest bound, and for the complete graph, the inequality (8.26) linked to the algebraic connectivity λ_2, also gives tight bounds (for the path graph, the latter have been left out since they are off by orders of magnitude due to small connectivity).

extremes with respect to connectivity, as well as with respect to the coherence metric (8.13) [156]. Therefore, they demonstrate our somewhat counterintuitive result that the Price of Synchrony does not depend on network connectivity (that is, neither algebraic connectivity, nor edge and vertex connectivity).

We simplify the example by assuming uniform dampings $d = 1$ for all generators in the network and then assign random line parameters to each line in the following manner. We draw both the line reactance x_{ij} and line ratio α_{ij} from a normal distribution with mean 0.2 and standard deviation 0.1 (negative values are not allowed). Based on Example 8.1, one can then expect the norm for each network to lie close to the result of equation (8.22) with the average ratio $\bar{\alpha} = 0.2$, and we should expect a linear growth of the $\mathcal{H}_2$ norm in N. This is also the main trend in Figure 8.4.

Figure 8.4 shows the $\mathcal{H}_2$ norms for the path and complete graph topologies as the network size increases from a 5 node to a 50 node system. We also indicate the bounds (8.23) and (8.25) on both panels a and b. We show the bounds from (8.26) only for the complete graph as this bound is very loose in the case of path graphs. For both types of networks, the eigenvalue ratio bound in (8.25) provides the tightest bound. We also note that the graph-theoretical dependent bounds (8.26) are more accurate than the line ratio bounds (8.23) for the complete graph, in line with the discussion in Section 8.4.1.

Example 8.3 (Placement of non-uniform generators) Our final example further relaxes Assumption 8.1 of equal generator parameters. We consider again our 7 bus example network from Figure 8.1, and let the lines have equal line impedances

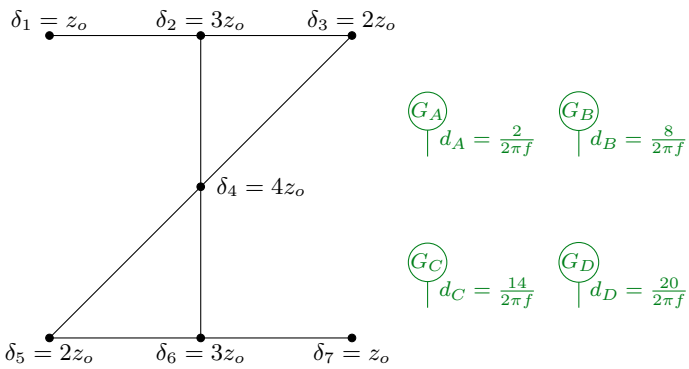


Figure 8.5: Illustration of the situation in Example 8.3. We place 4 generators with different dampings in the 7-node example network with respect to their node degrees δ_i.

$z_{ij} = 0.1 + \mathbf{j}0.6 = z_o$. The nodes in this network have the following nodal degrees δ_i in terms of the impedance z_o:

$$\delta_1 = \delta_7 = z_o, \quad \delta_2 = \delta_6 = 3z_o, \quad \delta_3 = \delta_5 = 2z_o, \quad \delta_4 = 4z_o.$$

The nodes thus have *four* different degrees, and we will now place *four different generators* $A - D$ at them. These generators all have inertia $m = \frac{20}{2\pi f}$, but we let their dampings be

$$d_A = \frac{2}{2\pi f}, \quad d_B = \frac{8}{2\pi f}, \quad d_C = \frac{14}{2\pi f}, \quad d_D = \frac{20}{2\pi f}.$$

The situation is depicted in Figure 8.5. We study this system under two conditions:

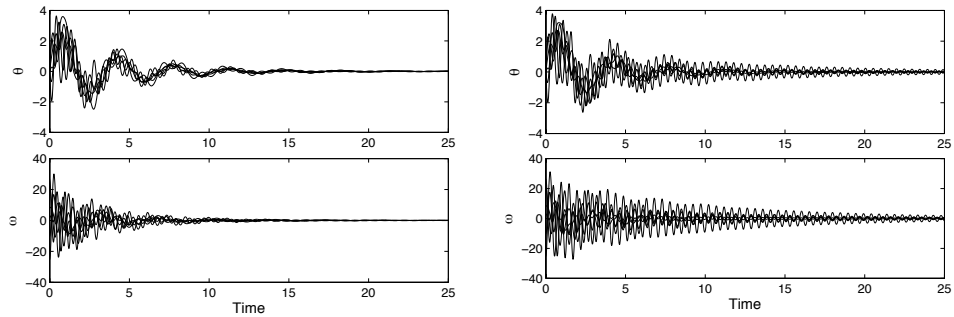
(a) The strongly damped generators placed at the *most* interconnected nodes:

$$\begin{aligned} d_A &\rightarrow \text{nodes 1 and 7} \\ d_B &\rightarrow \text{nodes 3 and 5} \\ d_C &\rightarrow \text{nodes 2 and 6} \\ d_D &\rightarrow \text{node 4} \end{aligned}$$

Here, dampings and node degrees are *matched*, and we call the resulting system (with node 1 grounded) $\bar{H}_{\text{match}}$.

(b) The strongly damped generators placed at the *least* interconnected nodes:

$$\begin{aligned} d_A &\rightarrow \text{node 4} \\ d_B &\rightarrow \text{nodes 2 and 6} \\ d_C &\rightarrow \text{nodes 3 and 5} \\ d_D &\rightarrow \text{nodes 1 and 7} \end{aligned}$$



(a) Damping matched to node degrees

(b) Damping mismatched to node degrees

Figure 8.6: Simulation from Example 8.3 with network in Figure 8.5. In (a) the strongly damped generators are placed at the most well-interconnected nodes, and in (b) the strongly damped generators placed are at the least interconnected nodes. The system in (b) is less coherent and experiences larger resistive losses during the transient response: 27.9 compared to 13.2 in system (a) for these particular trajectories.

Here, dampings and node degrees are *mismatched*, and we call the resulting system (with node 1 grounded) $\bar{H}_{\text{mismatch}}$.

We simulate the systems when they are subjected to a random initial angular velocity disturbance (according to the $\mathcal{H}_2$ norm interpretation (ii) in Section 8.2.1). The $\mathcal{H}_2$ norms can be evaluated numerically to:

$$||\bar{H}_{\text{match}}||_2^2 = 18.9, \text{ while } ||\bar{H}_{\text{mismatch}}||_2^2 = 20.7.$$

This shows that transient energy losses are *lower* for the system corresponding to case (a) where the dampings are *matched* to the node degrees.

Figure 8.6 shows the state trajectories of the two systems for a particular input sequence. The figure shows that the transient behavior of the system $\bar{H}_{\text{mismatch}}$ (b) is less "coherent" than that of $\bar{H}_{\text{match}}$ (a). Since the underlying graph and therefore the matrix $\mathcal{L}_G$ in the performance output is the same, it is clear in this case that the additional oscillations in the mismatched case (b) lead to increased transient losses. For the particular trajectories shown, the random input is such that the "mismatched" network is particularly excited compared to the "matched" network. Therefore, when we compute the losses for the particular trajectories shown, they are for the matched case (a) 13.2, and for the mismatched case (b) 27.9.

These results and similar case studies have led us to conclude that for systems with non-uniform generator parameters, judicious network design that places well-damped generators at highly interconnected nodes can reduce transient power losses. A heuristic explanation to this fact is that a well-damped generator is able to exert a larger effect on the entire network if it is well-connected than if it is

remotely located. However, although we are considering an extreme case where the best damped generator has as much as a 10 times larger damping coefficient than the most poorly damped one, the best placement only achieves 10% lower losses compared to the worst one (the optimality was confirmed numerically for this example).

We remark that in general, node degree may not be the most relevant centrality index[3] with respect to the Price of Synchrony. In fact, in a similar problem of inertia placement studied by Poolla *et al.* [136], some standard heuristics failed to optimize $\mathcal{H}_2$ performance. Our example is therefore somewhat anecdotal, and more work is needed to gain full insight to the optimal placement problem.

8.6 Discussion

In this chapter, we quantified the resistive line losses that occur due to the power flows that are required to maintain synchrony in a power network in the presence of persistent disturbances or transient events. These losses are the cost of using power flow through transmission lines as the signaling mechanism for synchronization, which motivates the term "Price of Synchrony". In the special case of a network of identical generators, we derived a formula for these losses, showing that they can be expressed as a generalized ratio of graph Laplacians weighted by, respectively, line conductances and susceptances. We showed that this quantity scales unboundedly with the number of nodes (generators) in the system, but that it is only weakly dependent on network topology. In this section, we discuss some of the implications of these results as well as certain technical aspects of our analysis.

8.6.1 Topology independence

In the special case where all of the transmission lines have equal conductance to susceptance ratios, we showed that the total transient resistive losses are independent of network topology, and directly proportional to the number of nodes in the network. This topological independence implies that while highly interconnected networks may have better phase coherence and stability properties than a sparsely interconnected one, the two networks are equivalent in terms of the power losses associated with maintaining synchrony. This conclusion may at first seem surprising, but some intuition can be gained by considering the following contrast between highly versus sparsely interconnected networks. A highly connected, and therefore highly phase-coherent network has smaller phase fluctuations than a sparsely interconnected one, and therefore less non-equilibrium power flows. So, while the "per-link" resistive losses are smaller in the former, it has many more links than the latter, and thus the total losses summed over all links can be the same for both networks.

[3] See Section 2.1.4 for a brief discussion on centrality indices.

8.6.2 Loss scaling

We evaluated the transient losses over the *entire* network, and not *per node* as with previous performance metrics. This is motivated by the fact that the losses arise over the power lines, so an evaluation per node is not physically motivated. However, if normalized by the number of nodes, the transient losses are clearly upper bounded for any network size. Additionally, we remarked that a generator contributes to the total transient losses only if it is also subjected to a disturbance. Hence, if we regard the issue of transient losses as an issue *per generator*, the scaling is reasonable.

The unbounded growth of the total transient losses with the number of generators, however, has interesting implications for the design of future power grids. This growth is, namely, unlikely to be mitigated by increasing network coherence through additional transmission links, unless these links focus on optimal matching of generator properties to node centrality. Such strategic addition of generation may, however, be impractical for an existing system or when connecting new distributed generation. Furthermore, our numerical examples indicate that even a strategic network design has relatively small impact on performance.

These results point to a fundamental limitation to a system where power flow is the mechanism by which the system resynchronizes or maintains a synchronous state. This performance limitation may be particularly relevant to future power networks that are likely to have orders of magnitude more generators than today's networks due to a high integration of renewable distributed generation units. The network as a whole will then experience a large increase in transient losses. Our results can be construed as an argument in support of investigating the use of communication links (for phase and frequency information) as an additional means of stabilizing control in power networks. Strategies that compensate for disturbances locally to prevent unnecessary inter-nodal power flows, or co-location of generators at a common bus would also be beneficial for performance.

8.6.3 Model limitations

The linearized coupled oscillator model for the power network that we use here is subject to a number of limitations. Some of them can be relaxed; variable voltage dynamics are accounted for in Chapter 9 and the Kron-reduced network model can be generalized to a structure-preserving Bergen-Hill model [165]. Neither of these model generalizations affect the results herein qualitatively.

In the derivation of our closed-form results, we assumed uniformity in generator damping and inertia. For non-uniform sets of generators, bounds like (8.27) that are qualitatively equivalent to our main result can be shown to hold. This means that our results do give insights to, for example, the scaling behavior of transient losses also in more heterogeneous networks. Bounds of this type may, however, be loose. As we could see in Example 8.3, heterogeneity in generator parameters can also interact with network topology and affect performance within the bounds. An

extended study that characterizes optimality of generator parameters with respect to transient losses is part of ongoing work. The studies [122] and [136] also address various aspects of performance in heterogeneous generator networks.

The Price of Synchrony metric allows an evaluation of total transient losses under a variety of input scenarios, including persistent stochastic noise, impulses, and initial perturbations. An alternative input scenario; step changes in loads, was considered in Example 4.2 for the evaluation of frequency coherence. In all cases, disturbances have been assumed uncorrelated across nodes. Possibly more realistic disturbance scenarios have recently been addressed in [32], where performance under colored (rather than white) noise is considered. The impact of line contingencies on synchronization performance was also studied in [33].

Lastly, the applicability of the linearized model for power network synchronization can be questioned. In our case, it limits the study to a small-signal analysis. In this type of analysis, an $\mathcal{H}_2$-norm metric (rather than ∞-norms) is well suited to characterize performance, since worst-case disturbances may drive the system outside the basin of attraction of the synchronous state. Overall, a meaningful generalization of our study to nonlinear power system dynamics is an open problem.

Appendix to Chapter 8

8.A Proof of Lemma 8.2

Consider the following state transformation of the system H in (8.9):

$$\begin{bmatrix}\theta\\ \omega\end{bmatrix} =: \begin{bmatrix}U & 0\\ 0 & U\end{bmatrix}\begin{bmatrix}\theta'\\ \omega'\end{bmatrix},$$

where U is the unitary matrix which diagonalizes $\mathcal{L}_B$, that is, $U^*\mathcal{L}_B U = \Lambda_B = \mathrm{diag}\{0,\ \lambda_2^B, ..., \lambda_N^B\}$, with $0 = \lambda_1^B \le \lambda_2^B \le ... \le \lambda_N^B$ being the eigenvalues of $\mathcal{L}_B$. We have assumed, without loss of generality, that $U = [\frac{1}{\sqrt{N}}\mathbf{1}\ \ u_2\ ...\ u_N]$, where u_i, $i = 2, ..., N$ are the eigenvectors corresponding to the aforementioned eigenvalues.

Since the $\mathcal{H}_2$ norm is unitarily invariant, see Section 3.1.2, we can also define $\mathrm{w}' = U^*\mathrm{w}$ and $y' = U^*y$ to obtain the system

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\theta'\\ \omega'\end{bmatrix} &= \begin{bmatrix}0 & I\\ -\frac{1}{m}\Lambda_B & -\frac{d}{m}I\end{bmatrix}\begin{bmatrix}\theta'\\ \omega'\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{m}I\end{bmatrix}\mathrm{w}'\\ y' &= \begin{bmatrix}U^*\mathcal{L}_G^{1/2}U & 0\end{bmatrix}\begin{bmatrix}\theta'\\ \omega'\end{bmatrix}.\end{aligned} \tag{8.28}$$

Now, observe that

$$U^*\mathcal{L}_G^{1/2}U = \begin{bmatrix}0 & \cdots & 0\\ \vdots & \hat{\mathcal{L}}_G^{1/2} & \\ 0 & & \end{bmatrix}, \tag{8.29}$$

which implies that the first rows and columns of both $U^*\mathcal{L}_G^{1/2}U$ and Λ_B are zero. We thus have that the states $\theta_1' = \frac{1}{\sqrt{N}}\sum_{i=1}^N \theta_i$ and $\omega_1' = \frac{1}{\sqrt{N}}\sum_{i=1}^N \omega_i$ satisfy the dynamics

$$\dot{\theta_1'} = \omega_1' \tag{8.30a}$$

$$\dot{\omega_1'} = -\frac{d}{m}\omega_1' + \frac{1}{m}\mathrm{w}_1' \tag{8.30b}$$

$$y_1' = 0. \tag{8.30c}$$

We note the subsystem (8.30) by H_1'. It corresponds to the single zero eigenvalue of $\mathcal{L}_B$ and is clearly unobservable as $y_1' \equiv 0$. The remaining eigenvalues of the system (8.28) lie strictly in the left half of the complex plane since $\lambda_2^B, \ldots, \lambda_N^B > 0$. It follows that the input-output transfer function from w' to y' is stable and has finite $\mathcal{H}_2$ norm. By the equivalence of the system (8.28) and H, we have thus established the existence of the $\mathcal{H}_2$ norm for the system H.

We can now partition the system (8.28) into the respective subsystems H_1' in (8.30) and $\hat{H}$. We take $\hat{\mathcal{L}}_G$ as the symmetric positive definite submatrix in (8.29) and define $\hat{\Lambda}_B = \mathrm{diag}\{\lambda_2^B,\ \lambda_3^B, ..., \lambda_N^B\}$. We then write the input-output

mapping $\hat{H}$ as:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\hat{\theta}\\ \hat{\omega}\end{bmatrix} &= \underbrace{\begin{bmatrix}0 & I\\ -\frac{1}{m}\hat{\Lambda}_B & -\frac{d}{m}I\end{bmatrix}}_{\hat{\mathcal{A}}}\begin{bmatrix}\hat{\theta}\\ \hat{\omega}\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ \frac{1}{m}I\end{bmatrix}}_{\hat{\mathcal{B}}}\hat{\mathrm{w}} \\ \hat{y} &= \underbrace{\begin{bmatrix}\hat{\mathcal{L}}_G^{1/2} & 0\end{bmatrix}}_{\hat{\mathcal{C}}}\begin{bmatrix}\hat{\theta}\\ \hat{\omega}\end{bmatrix} \end{aligned} \quad . \tag{8.31}$$

Note that the systems $\hat{H}_1$ and $\bar{H}$ are completely decoupled and we therefore have that $||H||_2^2 = ||H_1'||_2^2 + ||\hat{H}||_2^2 = ||\hat{H}||_2^2$.

The $\mathcal{H}_2$ norm of $\hat{H}$ can now be calculated in perfect analogy to the derivations in Section 8.3.2 and we obtain that

$$\|H\|_2^2 = \frac{1}{2d}\mathrm{tr}(\hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G). \tag{8.32}$$

Next, we show that the result of Lemma 8.1, which is in terms of the reduced graph Laplacians, can be written in terms of the state transformed matrices $\hat{\Lambda}_B$ and $\hat{\mathcal{L}}_G$. Define the $N \times (N-1)$ and the $(N-1) \times N$ matrices R and P by:

$$R = \begin{bmatrix}0 & \cdots & 0\\ & I_{N-1} & \end{bmatrix}, \quad P = \begin{bmatrix}I_{k-1} & & 0\\ & -\mathbf{1}_{N-1} & \\ 0 & & I_{N-k}\end{bmatrix},$$

where k is the index of the grounded node and $\mathbf{1}_{N-1}$ is the $(N-1) \times 1$ vector with all entries equal to 1. By this design, $\hat{\Lambda}_B = R^*\Lambda_B R$, $\hat{\mathcal{L}}_G = R^*U^*\mathcal{L}_G U R$ and $\mathcal{L}_{B(G)} = P^*\bar{\mathcal{L}}_{B(G)}P$. Further, to simplify notation, we define the $(N-1)\times(N-1)$ non-singular matrix $V = PUR$. Then we can write

$$\mathrm{tr}(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G) = \mathrm{tr}(VV^{-1}\bar{\mathcal{L}}_B^{-1}(V^*)^{-1}V^*\bar{\mathcal{L}}_G),$$

since $VV^{-1} = (V^*)^{-1}V^* = I$. By the cyclic properties of the trace:

$$\begin{aligned} \mathrm{tr}(VV^{-1}\bar{\mathcal{L}}_B^{-1}(V^*)^{-1}V^*\bar{\mathcal{L}}_G) &= \mathrm{tr}(V^{-1}\bar{\mathcal{L}}_B^{-1}(V^*)^{-1}V^*\bar{\mathcal{L}}_G V) \\ &= \mathrm{tr}((V^*\bar{\mathcal{L}}_B V)^{-1}V^*\bar{\mathcal{L}}_G V). \end{aligned} \tag{8.33}$$

But $V^*\bar{\mathcal{L}}_B V = R^*U^*P^*\bar{\mathcal{L}}_B PUR = \hat{\Lambda}_B$ and $V^*\bar{\mathcal{L}}_G V = R^*U^*P^*\bar{\mathcal{L}}_G PUR = \hat{\mathcal{L}}_G$. Hence,

$$\mathrm{tr}(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G) = \mathrm{tr}(\hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G).$$

In conclusion,

$$\|H\|_2^2 = \frac{1}{2d}\mathrm{tr}(\hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G) = \frac{1}{2d}\mathrm{tr}(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G) = \|\bar{H}\|_2^2,$$

which proves the Lemma. ■

8.B Proof of Lemma 8.3

By the proof of Lemma 8.2, we have that $\mathrm{tr}(\bar{\mathcal{L}}_B^{-1}\bar{\mathcal{L}}_G) = \mathrm{tr}(\hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G)$. Now,

$$\mathrm{tr}(\hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G) = \mathrm{tr}\left(\begin{bmatrix}0 & 0\\ 0 & \hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G\end{bmatrix}\right) = \mathrm{tr}\left(\begin{bmatrix}0 & 0\\ 0 & \hat{\Lambda}_B^{-1}\end{bmatrix} U^*\mathcal{L}_G U\right). \tag{8.34}$$

By definition, see e.g. [65],

$$U^*\mathcal{L}_B^\dagger U = \mathrm{diag}\{0,\ \frac{1}{\lambda_2^B},...,\frac{1}{\lambda_N^B}\} = \begin{bmatrix}0 & 0\\ 0 & \hat{\Lambda}_B^{-1}\end{bmatrix},$$

which makes (8.34) equivalent to $\mathrm{tr}(U^*\mathcal{L}_B^\dagger UU^*\mathcal{L}_G U) = \mathrm{tr}(U^*\mathcal{L}_B^\dagger\mathcal{L}_G U)$. But since the trace is unitarily invariant, it follows that $\mathrm{tr}(\hat{\Lambda}_B^{-1}\hat{\mathcal{L}}_G) = \mathrm{tr}(\mathcal{L}_B^\dagger\mathcal{L}_G)$, which is precisely the statement of the lemma. ■

8.C Proof of Lemma 8.6

Without loss of generality, number the nodes so that $k = N$. Let $\bar{M} := \mathrm{diag}\{m_1,...,m_N\}$, $\bar{D} = \mathrm{diag}\{d_1,...,d_N\}$ and denote the new $(N+1)^{\text{th}}$ node as NI for notational compactness. The reduced system $\bar{H}_1$ can then be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\bar{\theta}\\ \theta_{NI}\\ \bar{\omega}\\ \omega_{NI}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & I_{N-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1\\ -\bar{M}^{-1}\bar{\mathcal{L}}_B & \mathbf{0} & -\bar{M}^{-1}\bar{D} & \mathbf{0}\\ \mathbf{0} & -\frac{b_{N,NI}}{m_{NI}} & \mathbf{0} & -\frac{d_{NI}}{m_{NI}}\end{bmatrix}\begin{bmatrix}\bar{\theta}\\ \theta_{NI}\\ \bar{\omega}\\ \omega_{NI}\end{bmatrix} + \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & 0\\ \bar{M}^{-1} & \mathbf{0}\\ \mathbf{0} & \frac{1}{m_{NI}}\end{bmatrix}\begin{bmatrix}\bar{\mathrm{w}}\\ \mathrm{w}_{NI}\end{bmatrix} \tag{8.35}$$

$$\begin{bmatrix}\bar{y}\\ y_{NI}\end{bmatrix} = \begin{bmatrix}\bar{\mathcal{L}}_G^{1/2} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \sqrt{g_{N,NI}} & \mathbf{0} & 0\end{bmatrix}\begin{bmatrix}\bar{\theta}\\ \theta_{NI}\\ \bar{\omega}\\ \omega_{NI}\end{bmatrix},$$

where $b_{N,NI}$, $g_{N,NI}$ are, respectively, the susceptance and conductance of the line (N, NI). Let the input-output mapping H_{NI} be the SISO subsystem of (8.35) given by:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\theta_{NI}\\ \omega_{NI}\end{bmatrix} &= \begin{bmatrix}0 & 1\\ -\frac{b_{N,NI}}{m_{NI}} & -\frac{d_{NI}}{m_{NI}}\end{bmatrix}\begin{bmatrix}\theta_{NI}\\ \omega_{NI}\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{m_{NI}}\end{bmatrix}\mathrm{w}_{NI}\\ y_{NI} &= \begin{bmatrix}\sqrt{g_{N,NI}} & 0\end{bmatrix}\begin{bmatrix}\theta_{NI}\\ \omega_{NI}\end{bmatrix}.\end{aligned}$$

From (8.35), it is clear that the systems $\bar{H}_0$ and H_{NI} are entirely decoupled. We can therefore write $\bar{H}_1 = \mathrm{diag}\{\bar{H}_0,\ \bar{H}_1\}$, and have that

$$||\bar{H}_1||_2^2 = ||H_0||_2^2 + ||H_{NI}||_2^2.$$

Now, the $\mathcal{H}_2$ norm of H_{NI} can be calculated as in Section 8.3.2 and becomes:

$$||H_{NI}||_2^2 = \frac{1}{2d_{NI}} \frac{g_{N,NI}}{b_{N,NI}} = \frac{\alpha_{N,NI}}{2d_{NI}},$$

and the result follows. ■

Chapter 9

Performance Impact of Variable Voltage Dynamics

The key motivation for studying power systems in this thesis is the transition to a distributed generation paradigm. This implies that numerous smaller, local generation sources replace large, centralized power plants. Much of this generation will be connected at low- to medium-voltage grids. In such grids, the assumptions of a flat voltage profile and negligible resistance-to-reactance ratios are no longer valid, and a tight voltage control becomes important. In this chapter, we will analyze how the results from Chapter 8 are impacted by a coupling of the phase-frequency dynamics to variable voltage dynamics.

We address this question by studying a microgrid network. Microgrids, which are envisioned as small, autonomously operated networks composed of distributed generation units, loads and storage elements, have been proposed as a strategy to facilitate the transition to a distributed generation paradigm [87, 103]. The distributed generation units within a microgrid are typically interfaced with the network via DC/AC or AC/AC power converters, or *inverters*. The network's stability, synchronization and power balance depend on control actions taken in these inverters [131, 199]. *Droop control*; essentially a decentralized proportional controller, is a widely proposed control scheme in this context. In Chapter 2, we reviewed some research focused on deriving analytical conditions for synchronization and power sharing in droop-controlled power networks. Most of these works have assumed constant voltage profiles, though droop control for voltage and reactive power stabilization in microgrids has recently attracted attention [56, 147, 148, 160]. In particular, Schiffer *et al.* derive conditions on controller gains for frequency and voltage stability in [147] and for reactive power sharing in [148].

In this chapter, we study the same type of inverter-based microgrid as in [148] and characterize transient performance in terms of a Price of Synchrony metric that is adapted to account for losses due to voltage variations. We will show, on one hand, that the cross-coupling of phase-frequency and voltage dynamics leads to

conditions on the respective droop controllers that must be observed to guarantee stability of even the linearized model. Provided these conditions are satisfied with a reasonable stability margin, we prove, on the other hand, that the cross-coupling only has a small effect on the total transient losses. This means that frequency and voltage droop control can be decoupled with respect to our performance analysis. While frequency synchronization was shown to produce losses that scale with network size but only have a weak topology dependence, the additional losses due to voltage control will be larger in a well-interconnected network than in a sparsely interconnected one. This means that the typically higher rate of convergence and degree of coherence in a well-interconnected network comes at the cost of higher transient losses.

9.1 Microgrid model with frequency and voltage droop control

In this chapter, we consider a model of an inverter-based microgrid. The inverters are the interface between the typically asynchronous or DC renewable generation sources and the synchronous AC power grid. They will be modeled as voltage sources whose voltage amplitude and frequency outputs can be regulated, in our case through droop controllers. The resulting closed-loop system is one where the phase-frequency dynamics, which are analogous to the previous Chapter, are coupled with voltage dynamics. Under certain conditions, which we discuss at the end of this section, a decoupling of the two can, however, be motivated.

9.1.1 Network model and definitions

The microgrid is modeled through a Kron-reduced network model $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ as in Chapter 8. Consequently, every node $i \in \mathcal{V}$ represents a generation unit with a power inverter as its grid interface. Each of these has an associated complex voltage $v_i = V_i e^{\mathbf{j}\theta_i}$, where $V_i = |v_i|$ is the voltage magnitude and θ_i is the phase angle. The network itself is represented through its admittance matrix, introduced in Section 3.4.2.

Special case: complete network graph

Many results in this chapter are derived for the special case when the graph underlying the network $\mathcal{G}$ is *complete*, that is, there is a line (i, j) connecting each node pair $i, j \in \mathcal{V}$. Our focus on this case is partly motivated by the study [112], which showed that Kron-reduction of power networks typically results in complete network models.

It will often be useful to study these networks in the limit where $N \to \infty$, that is, when they are large. While the notion of large microgrids may appear an oxymoron, it should be thought of as a highly distributed system scenario with a large number of end points with controllable generation units.

9.1.2 Inverter and droop control model

We now introduce the models of the power inverters, adopting the framework from [147]. We assume that the inverters are voltage sources, whose amplitude and frequency output are regulated according to:

$$\dot{\theta}_i = u_i^{\theta} \tag{9.1a}$$

$$\tau_{V_i}\dot{V}_i = -V_i + u_i^{V}, \tag{9.1b}$$

where u_i^V and u_i^θ are the respective control signals. The voltage regulation is modeled with a small lag, represented by a filter with time constant $\tau_{V_i} \geq 0$.

The controls u_i^θ and u_i^V are then implemented as droop controllers; proportional controllers based, respectively, on active and reactive power deviations:

$$u_i^{\theta} = \omega^{\text{ref}} - k_{P,i}(\hat{P}_i - P_i^{\text{ref}}) \tag{9.2a}$$

$$u_i^{V} = V_i^{\text{ref}} - k_{Q,i}(\hat{Q}_i - Q_i^{\text{ref}}), \tag{9.2b}$$

where ω^{ref}, V_i^{ref}, P_i^{ref} and Q_i^{ref} are the respective setpoints for the frequency, voltage magnitude, active and reactive power. The parameters $k_{P,i},\ k_{Q,i} > 0$ are the *droop gains* and $\hat{P}_i$ and $\hat{Q}_i$ are the active and reactive powers measured by the power electronics at the inverter. These measurements are assumed to be processed through low-pass filters given by:

$$\tau_{P,i}\dot{\hat{P}}_i = -\hat{P}_i + P_i \tag{9.3a}$$

$$\tau_{Q,i}\dot{\hat{Q}}_i = -\hat{Q}_i + Q_i, \tag{9.3b}$$

where $\tau_{P,i},\ \tau_{Q,i} > 0$ are the filter time constants and P_i and Q_i are the actual power injections to the network at node i. The relations (9.1)–(9.3) can now be used to formulate the closed-loop system dynamics. For this purpose, it first assumed that the time constant for the voltage control, τ_{V_i} is small compared to $\tau_{Q,i}$, and can be neglected [147]. We therefore set $\tau_{V_i} = 0$ in (9.1b), and then by substituting (9.2) into (9.1), we obtain:

$$\dot{\theta}_i = \omega_i \tag{9.4a}$$

$$\omega_i = \omega^{\text{ref}} - k_{P,i}(\hat{P}_i - P_i^{\text{ref}}) \tag{9.4b}$$

$$V_i = V_i^{\text{ref}} - k_{Q,i}(\hat{Q}_i - Q_i^{\text{ref}}), \tag{9.4c}$$

where we have introduced the inverter frequency ω_i. Taking the derivatives of (9.4b) and (9.4c) with respect to time gives $\dot{\omega}_i = -k_{P,i}\dot{\hat{P}}_i$ and $\dot{V}_i = -k_{Q,i}\dot{\hat{Q}}_i$, in which we can insert the equations (9.3). We then substitute $\hat{P}_i$ and $\hat{Q}_i$ using (9.4b) and (9.4c) and obtain the control dynamics for the phase angles and voltages as:

$$\begin{aligned} \dot{\theta}_i &= \omega_i \\ \tau_{P,i}\dot{\omega}_i &= -\omega_i + \omega^{\text{ref}} - k_{P,i}(P_i - P_i^{\text{ref}}) \\ \tau_{Q,i}\dot{V}_i &= -V_i + V_i^{\text{ref}} - k_{Q,i}(Q_i - Q_i^{\text{ref}}). \end{aligned} \tag{9.5}$$

We next derive the equations for P_i and Q_i.

Remark 9.1 The procedure outlined above results in a second-order model for the phase-frequency dynamics, which is analogous to the swing dynamics in Chapter 8. Frequency droop control is sometimes, for example in [159], studied through a first-order model. That model is obtained in the limit where $\tau_P \to 0$. Its performance expressions also correspond to those of the second-order model in the same limit We refer to [174, Chapter 5] for an elaboration on the first-order model. □

9.1.3 Linearized complex power flows

Introducing $\theta_{ij} = (\theta_i - \theta_j)$ as the phase angle difference between neighboring nodes, the active and reactive powers injected to the grid at each node $i \in \mathcal{V}$ are

$$P_i = -g_{ii}V_i^2 + \sum_{j\in\mathcal{N}_i} V_iV_j(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij}) \tag{9.6a}$$

$$Q_i = b_{ii}V_i^2 + \sum_{j\in\mathcal{N}_i} V_iV_j(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij}). \tag{9.6b}$$

At the i^{th} node, $g_{ii} = \bar{g}_i + \sum_{j\in\mathcal{N}_i} g_{ij}$ and $b_{ii} = \bar{b}_i + \sum_{j\in\mathcal{N}_i} b_{ij}$ represent respectively the shunt conductance and shunt susceptance. We make the common assumption [58, 148] that the shunt elements are purely inductive so that, in our notation, $\bar{g}_i = 0$ and $\bar{b}_i \geq 0$ for all $i \in \mathcal{V}$.

As usual, we will be considering the system under the assumption of small deviations from an operating point. For these reasons, we can approximate the power flows by a linearization around the point $P_i^0(\theta_{ij}^0, V_i^0, V_j^0)$ and $Q_i^0(\theta_{ij}^0, V_i^0, V_j^0)$, where $V_i^0 = V_j^0 = V^0 = 1$ p.u. and $\theta_{ij} = 0$ for all $i, k \in \mathcal{N}$. This procedure gives the linearized power injections at node i as:

$$\Delta P_i = \sum_{k\sim i} (-g_{ij}(\Delta V_i - \Delta V_j) + b_{ij}\Delta\theta_{ij}) \tag{9.7a}$$

$$\Delta Q_i = 2\bar{b}_i\Delta V_i + \sum_{k\sim i} (b_{ij}(\Delta V_i - \Delta V_j) + g_{ij}\Delta\theta_{ij})\,. \tag{9.7b}$$

9.1.4 Closed-loop dynamics

We now summarize the relations derived so far into one closed-loop system subject to distributed disturbances that represent generation and load fluctuations. For this purpose, we let the linearization points used in (9.7a)–(9.7b) be the setpoints of the droop controllers in (9.2), so that $\Delta P_i = P_i - P_i^{\text{ref}}$ and $\Delta Q_i = Q_i - Q_i^{\text{ref}}$ for all $i \in \mathcal{V}$. Without loss of generality we transfer this equilibrium to the origin. In an effort to avoid cumbersome notation, we then omit the difference operator Δ and let the state variables $(\theta_{ij},\ \omega_i,\ V_i)$ represent deviations from the operating point. We assume that the system is driven by a disturbance input w.

Inserting the power flow equations (9.7a)–(9.7b) to the control dynamics (9.5) of the i^{th} inverter gives:

$$\begin{aligned}\dot{\theta}_i &= \omega_i \\ \tau_{P,i}\dot{\omega}_i &= -\omega_i - k_{P,i}(-\sum_{k\sim i} g_{ij}(V_i - V_j) + \sum_{k\sim i} b_{ij}\theta_{ij}) + \mathrm{w}_i^{\omega} \\ \tau_{Q,i}\dot{V}_i &= -V_i - k_{Q,i}(2\bar{b}_i V_i + \sum_{k\sim i} b_{ij}(V_i - V_j) + \sum_{k\sim i} g_{ij}\theta_{ij}) + \mathrm{w}_i^{V}.\end{aligned} \tag{9.8}$$

Now, by defining $\theta,\ \omega,\ V$ as column vectors containing the states $\theta_i,\ \omega_i,\ V_i,\ i \in \mathcal{V}$ and using the susceptance and conductance matrices $\mathcal{L}_B$ and $\mathcal{L}_G$ defined in (3.28), we can summarize the above as follows:

$$\begin{bmatrix}\dot{\theta}\\ \dot{\omega}\\ \dot{V}\end{bmatrix} = \underbrace{\begin{bmatrix}0 & I & 0\\ -K_P T_P^{-1}\mathcal{L}_B & -T_P^{-1} & K_P T_P^{-1}\mathcal{L}_G\\ -K_Q T_Q^{-1}\mathcal{L}_G & 0 & -C_Q T_Q^{-1} - K_Q T_Q^{-1}\mathcal{L}_B\end{bmatrix}}_{\mathcal{A}}\begin{bmatrix}\theta\\ \omega\\ V\end{bmatrix} + \underbrace{\begin{bmatrix}0 & 0\\ T_P^{-1} & 0\\ 0 & T_Q^{-1}\end{bmatrix}}_{\mathcal{B}}\mathrm{w}, \tag{9.9}$$

where $\mathrm{w} = [\mathrm{w}^{\omega},\ \mathrm{w}^{V}]^T$ is the disturbance input vector, and we define $K_{P(Q)} = \mathrm{diag}\{k_{P,i(Q,i)}\}$, $T_{P(Q)} = \mathrm{diag}\{\tau_{P,i(Q,i)}\}$ and $C_Q = \mathrm{diag}\{c_{Q,i}\}$ with $c_{Q,i} = 1 + 2k_{Q,i}\bar{b}_i$.

We emphasize that the phase-frequency dynamics and the voltage dynamics in (9.9) are in general *cross-coupled*. Even though the droop controllers (9.2) are decoupled, that is, the voltage V is controlled based on reactive power flow while phase θ and frequency ω are controlled based on active power flow, these respective power flows depend on *both* V and θ according to (9.7). The strength of this cross-coupling depends on the network lines' resistance-to-reactance ratio. Throughout this chapter, we make the following assumption:

Assumption 9.1 The conductance-to-susceptance (equivalently, resistance-to-reactance) ratio α is uniform for all network lines, that is,

$$\alpha := \frac{g_{ij}}{b_{ij}}, \quad \forall (i,j) \in \mathcal{E}.$$

It thus holds $\mathcal{L}_G = \alpha \mathcal{L}_B$.

This assumption, which is used in several related studies [42, 97, 159], is not unreasonable for power networks. See Section 8.3.3 for an elaboration.

Under Assumption 9.1, the cross-coupling between the phase-frequency and the voltage dynamics in (9.9) will be proportional to α (note $\mathcal{L}_G = \alpha\mathcal{L}_B$ in the elements $\mathcal{A}_{(2,3)}$ and $\mathcal{A}_{(3,1)}$ in (9.9)). The situation is illustrated in Figure 9.1. In the next section, we will show that this cross-coupling will be important for the stability properties of the system.

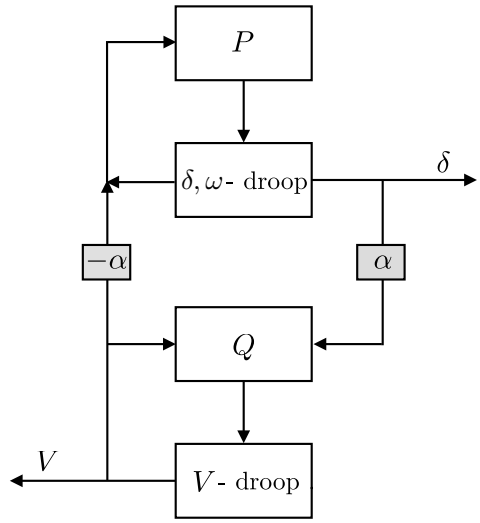


Figure 9.1: Illustration of cross-coupling between voltage and phase-frequency dynamics in (9.8). The strength of the cross-coupling is parameterized through the conductance-to-susceptance (resistance-to-reactance) ratio α. If α is small, the cross-coupling is weak and system properties can be studied through decoupled dynamics where one assumes $P(\theta, V) \approx P(\theta)$ and $Q(\theta, V) \approx Q(V)$.

9.1.5 System decoupling

The dominating behavior of the system in (9.9) can be analyzed in the limit where the network's resistances are small compared to its reactances, that is, when $\alpha \to 0$ (we show this formally with respect to our analysis in Section 9.4). Under this common assumption, the active power flow is a function only of the phase angles and the reactive power flow is a function only of the voltage magnitudes. That is, $P(\theta, V) \approx P(\theta)$, $Q(\theta, V) \approx Q(V)$, see for example [23, 56, 147, 163]. This *decouples* the phase-frequency and voltage dynamics. The decoupled system $H^{\text{dec}} = (\mathcal{A}^{\text{dec}}, \mathcal{B}, \mathcal{C})$ is formulated by setting $\mathcal{L}_G = 0$ in the matrix $\mathcal{A}$ in (9.9):

$$\mathcal{A}^{\text{dec}} := \begin{bmatrix} 0 & I & 0 \\ -K_P T_P^{-1} \mathcal{L}_B & -T_P^{-1} & 0 \\ 0 & 0 & -C_Q T_Q^{-1} - K_Q T_Q^{-1} \mathcal{L}_B \end{bmatrix}.$$

In power networks, and in particular in low-voltage distribution networks, the assumption of small resistance-to-reactance ratios is not generally applicable. The cross-couplings in the system therefore become increasingly relevant. In highly resistive networks, it may even reasonable to consider the opposite dependencies; $P(\theta, V) \approx P(V)$ and $Q(\theta, V) \approx Q(\theta)$ [23], though such considerations are not made in this thesis.

We are, however, modeling an inverter-based network. Given that inverter output impedances are highly inductive they tend to dominate line resistances in Kron-reduced network models [147,148]. Therefore, α will be relatively small in our model, and we will indeed be able to show relevant characteristics of our problem by considering the simpler, decoupled system H^{dec}.

9.2 The Price of Synchrony with variable voltages

We now extend the Price of Synchrony performance measure introduced in Section 8.2 to also account for losses associated with fluctuating voltages. For this purpose, consider again the real power loss over an edge (i, j), which is given by Ohm's law as

$$P_{ij}^{\mathrm{loss}} = g_{ij}|v_i - v_j|^2, \tag{9.10}$$

where $v_i,\ v_j$ are the complex voltages at nodes i and j. We can now enforce the common linearized system assumption of small phase angle differences. Standard trigonometric methods then give that $|v_i - v_j|^2 \approx (V_i - V_j)^2 + (V_i(\theta_i - \theta_j))^2$. Since we also assume $V_i \approx 1$ p.u. around the linearization point for all $i \in \mathcal{V}$, an approximation of the power loss over the edge (i, j) is $P_{ij}^{\mathrm{loss}} = g_{ij}\left[(V_i - V_j)^2 + (\theta_i - \theta_j)^2\right]$. The total instantaneous power losses over the network are then approximately

$$\mathbf{P}_{\mathrm{loss}} \approx \sum_{j \in \mathcal{N}_i} g_{ij}\left[(V_i - V_j)^2 + (\theta_i - \theta_j)^2\right]. \tag{9.11}$$

Making use of the conductance matrix $\mathcal{L}_G$, we can write (9.11) as the quadratic form

$$\mathbf{P}_{\mathrm{loss}} = V^T \mathcal{L}_G V + \theta^T \mathcal{L}_G \theta, \tag{9.12}$$

which allows us to make the following definition:

Definition 9.1 (Price of Synchrony performance output with variable voltages) The instantaneous resistive power losses in a power network are measured as $\mathbf{P}_{loss} = y^T y$, where the performance output y is

$$y := \mathcal{L}_G^{1/2}\theta + \mathcal{L}_G^{1/2} V. \tag{9.13}$$

Here V and θ are the state vectors that were defined in the Section 9.1.4.

9.3 Admissible frequency and voltage droop controllers

We have now derived a model for an inverter-based microgrid with both frequency and voltage droop control. Although there is a large body of literature that treats the stability properties of these droop controllers, they are typically treated separately from each other, implicitly or explicitly assuming the problems to be decoupled. In this section, we will discuss stability conditions for the system (9.9), and show that the cross-coupling of voltage and frequency dynamics indeed limits the set of admissible droop control designs.

We are here considering the linearized microgrid dynamics, and can therefore clearly not perform a complete stability analysis of the cross-coupled system. However, our analysis provides necessary conditions for the full nonlinear system, and can therefore give insights to the problem. In particular we show that certain relations between the voltage and frequency droop controllers must be observed, and that these depend on the network's resistance-to-reactance ratio α.

Remark 9.2 The term "admissibility" is used here in a slightly different meaning than in Chapters 5–6. Here, admissible frequency and voltage droop controllers refer to decoupled control designs that also guarantee stability of the cross-coupled system. □

9.3.1 Conditions on droop gains versus α

To provide tractable closed-form stability conditions, we first consider the system (9.9) in the hypothetical case of a system with identical inverters. Consider the following assumptions:

Assumption 9.2 The power measurement time constants $\tau_{P,i}$, and $\tau_{Q,i}$ are uniform across all inverters, that is, $\tau_{P,i} = \tau_P$ and $\tau_{Q,i} = \tau_Q$, $\forall i \in \mathcal{V}$.

Assumption 9.3 All nodes have equal shunt susceptances $\bar{b}_i = \bar{b} \geq 0$, and zero shunt conductances $\bar{g}_i = 0$, $\forall i \in \mathcal{V}$.

Assumption 9.4 The droop gains $k_{P,i}$, and $k_{Q,i}$ are uniform across all inverters, that is, $k_{P,i} = k_P$ and $k_{Q,i} = k_Q$, $\forall i \in \mathcal{V}$.

The system (9.9) has a single zero eigenvalue and the matrix $\mathcal{A}$ will therefore not be Hurwitz. As usual, however, the corresponding mode that describes the drift of the mean phase angle is unobservable from the output. Here, we will therefore consider input-output (IO) stability of the system (9.9) with respect to the output from Definition 9.1. This is equivalent to ensuring that remaining eigenvalues of $\mathcal{A}$ have negative real parts. A condition for this is given in the following theorem:

Theorem 9.1 *Let Assumptions 9.1–9.4 hold. Then, the system* (9.9) *is input-output (IO) stable if and only if*

$$\frac{k_Q^2 \tau_P \lambda_N^2 + (k_P \tau_Q^2 + k_Q \tau_Q + 2 c_Q k_Q \tau_P)\lambda_N + \tau_P c_Q^2 + \tau_Q c_Q}{k_P k_Q \tau_P \tau_Q \lambda_N^2} > \alpha^2,$$

where λ_N is the greatest eigenvalue of $\mathcal{L}_B$.

Proof: See Appendix 9.B.

Theorem 9.1 shows that when $\alpha > 0$, not any choice of k_P and k_Q is admissible. Instead, a relation between these gains must be observed. This relation depends on for example λ_N, that is, the underlying network topology.

From Theorem 9.1 we derive a more tractable conservative stability condition, which is tight in the limit where the eigenvalue λ_N is large. A case where λ_N is large is when the graph $\mathcal{G}$ is complete and the number of nodes N becomes large (see Section 9.1.1).

Corollary 9.2 *Under Assumptions 9.1–9.4, the system* (9.9) *is IO stable if*

$$\frac{k_Q}{k_P \tau_Q} > \alpha^2.$$

Proof: Follows from Theorem 9.1, noting that the left hand side is decreasing in $\lambda_N > 0$, and letting $\lambda_N \to \infty$. ■

In cases where λ_N is not large, the following alternative sufficient condition is relevant:

Corollary 9.3 *Under Assumptions 9.1–9.4* (9.9) *is IO stable if*

$$\alpha^2 \leq \frac{\tau_Q}{\tau_P} \frac{1}{k_Q \lambda_N} =: \alpha_{\text{crit}}^2.$$

This implies that in case $\alpha \leq \alpha_{\text{crit}}$ due to k_Q being chosen sufficiently small in relation to λ_N, any k_P is admissible. This is easier to satisfy if λ_N is small, that is, in sparsely interconnected networks. Conversely, if $\alpha > \alpha_{\text{crit}}$, sufficiently large gains k_P will violate the stability condition in Theorem 9.1.

9.3.2 Stability of decoupled system

Let us now consider the special case of the decoupled system H^{dec}. Clearly, since the parameter $\alpha = 0$ in $\mathcal{A}^{\text{dec}}$, the condition in Theorem 9.1 is always fulfilled. Therefore, H^{dec} is IO stable for any choice of the parameters. It also follows from the following lemma:

Lemma 9.4 *Under Assumptions 9.1–9.4, the set of eigenvalues of the decoupled system matrix* $\mathcal{A}^{\text{dec}}$ *is:*

$$\sigma(\mathcal{A}^{\text{dec}}) = \bigcup_{n=2}^{N} \left\{ -\frac{1}{2\tau_P} \left(1 \pm \sqrt{1 - k_P \tau_P \lambda_n} \right), -\frac{c_Q}{\tau_Q} - \frac{k_Q}{\tau_Q} \lambda_n \right\} \cup \left\{ 0, -\frac{1}{\tau_P}, -\frac{c_Q}{\tau_Q} \right\}.$$

If the parameters $k_P, \tau_P, k_Q, \tau_Q > 0$ *and the shunt susceptance satisfies* $c_Q > 0 \Leftrightarrow b > \frac{-1}{2k_Q}$, *all eigenvalues apart from the zero eigenvalue lie strictly in the left half of the complex plane and the system* $H^{\text{dec}} = (\mathcal{A}^{\text{dec}}, \mathcal{B}, \mathcal{C})$ *is IO stable.*

Proof: Follows from block-diagonalizing the matrix $\mathcal{A}^{\text{dec}}$ as in Appendix 9.A and then evaluating the characteristic polynomials of $\hat{\mathcal{A}}_n^{\text{dec}}$. Since the Laplacian eigenvalues $\lambda_n > 0$ for $n = 2, \ldots, N$, it is easy to see that if $k_P, \tau_P, k_Q, \tau_Q, c_Q > 0$, all eigenvalues of $\mathcal{A}^{\text{dec}}$ (apart from the zero eigenvalue) have negative real parts. ■

Remark 9.3 The system H^{dec} can be shown to be IO stable also with non-uniform system parameters. See, for example, [4]. □

9.4 Performance of frequency and voltage droop control

In this section, we present the Chapter's main result. We begin by analyzing the performance in terms of transient resistive losses for the cross-coupled system. After showing that the error made by considering the decoupled system H^{dec} is small, we proceed to analyzing the topology-dependence of that system's performance.

9.4.1 System performance and error due to decoupling

Consider now the system (9.9). We noted in Section 9.1.4 that it has cross-couplings between the voltage and phase-frequency dynamics that are proportional to the resistance-to-reactance ratio α. These cross-couplings make the closed-form performance expressions more or less intractable. However, we show here that it can be expressed as a series expansion in α:

Proposition 9.5 *Let Assumptions 9.1–9.4 hold. The squared $\mathcal{H}_2$ norm of* (9.9) *with respect to the output from Definition 9.1 is, for sufficiently small α, given by:*

$$\|H\|_2^2 = c_0^\Lambda \alpha + c_1^\Lambda \alpha^3 + c_2^\Lambda \alpha^5 + \dots, \tag{9.14}$$

where c_k^Λ, $k = 0, 1, 2, \dots$, are positive, scalar functions of the eigenvalues λ_n of $\mathcal{L}_B$.

The first term in the expansion corresponds to the squared $\mathcal{H}_2$ norm of the decoupled system H^{dec}:

$$c_0^\Lambda \alpha = \|H^{\text{dec}}\|_2^2,$$

which is given by:

$$\|H^{\text{dec}}\|_2^2 = \frac{\alpha}{2k_P}(N-1) + \frac{\alpha}{2\tau_Q}\sum_{n=2}^{N}\frac{1}{\frac{c_Q}{\lambda_n}+k_Q}. \tag{9.15}$$

Proof: See Appendix 9.C.

Remark 9.4 The requirement that α be sufficiently small is equivalent to requiring IO stability. It is thus satisfied if the condition in Theorem 9.1 holds. □

The following corollary follows immediately from (9.14)–(9.15):

Corollary 9.6

$$\|H\|_2^2 > \|H^{\text{dec}}\|_2^2.$$

This means that the cross-coupling between voltage and phase-frequency dynamics is strictly detrimental for performance in terms of transient power losses.

Notice that the part of the $\mathcal{H}_2$ norm that can be attributed to the cross-coupling is in high-order terms of the resistance-to-reactance ratio α. Since α is (much) smaller than 1 in the model considered here, the error made by decoupling the system is small. Define the relative error η made by decoupling the system as:

$$\eta = \frac{\|H\|_2^2 - \|H^{\mathrm{dec}}\|_2^2}{\|H^{\mathrm{dec}}\|_2^2} \tag{9.16}$$

and consider the following corollary:

Corollary 9.7 *The relative error η is given by*

$$\eta = \frac{c_1^\Lambda}{c_0^\Lambda}\alpha^2 + \frac{c_2^\Lambda}{c_0^\Lambda}\alpha^4 + \ldots \;=: \bar{c}_1^\Lambda \alpha^2 + \bar{c}_2^\Lambda \alpha^4 + \ldots$$

This means that the relative error $\eta = \mathcal{O}(\alpha^2)$. This quadratic trend is also seen from the numerical example in Figure 9.2.

In the important special case where the graph $\mathcal{G}$ underlying the network is complete and where the number of nodes N is large, the eigenvalues λ_n will become large. For this case, we can derive the following result:

Proposition 9.8 *If $\lambda_n \to \infty$ for all $n = 2, \ldots, N$, then*

$$\|H^{\mathrm{dec}}\|_2^2 \approx \frac{\alpha}{2}(N-1)\left(\frac{1}{k_P} + \frac{1}{\tau_Q k_Q}\right), \tag{9.17}$$

and

$$\eta \approx \frac{k_P \tau_Q}{k_Q}\alpha^2 + \left(\frac{k_P \tau_Q}{k_Q}\right)^2 \alpha^4 + \ldots. \tag{9.18}$$

Proof: See Appendix 9.D.

We conjecture, based on numerical evaluation, that (9.18) provides an upper bound on the relative error for any underlying network graph.

It is interesting to note that the coefficients in (9.18) are the inverted ratio from the stability condition in Corollary 9.2. Both hold in the limit of a large, complete graph. This is in line with the intuitition that the larger the stability margin, the smaller is the error made by considering the decoupled system (which is always stable) in the performance analysis.

In conclusion, the relative error η will be small provided a reasonable stability margin is maintained. We therefore proceed to analyzing the transient power losses of the decoupled system further.

9.4.2 Topology dependence of decoupled system performance

Consider now the norm expression (9.15). This expression is monotonically increasing in the Laplacian eigenvalues λ_n, implying that transient losses increase

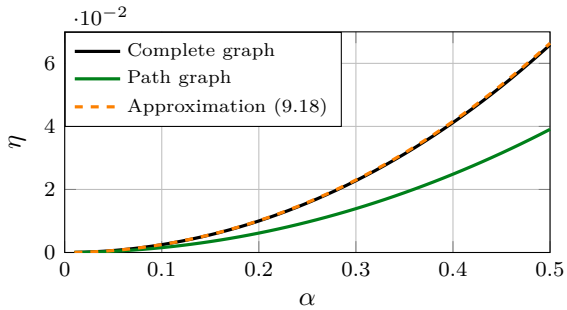


Figure 9.2: Norm error η in (9.16) as a function of α for networks of size $N = 50$ with complete graph and path graph structures. We also display the approximation (9.18). The errors are small and decrease quadratically as $\alpha \to 0$, as predicted by Proposition 9.5. Here, $x_{ij} = 0.2$, $k_P = 1$, $k_Q = 2$, $\tau_P = \tau_Q = 0.5$ and $c_Q = 1$.

with increasing network connectivity and line susceptances. While microgrid network structures may vary, in terms of connectivity they will fall somewhere between the two extremes given by the complete graph and tree graphs. We next present results for these two cases.

Consider first the special case of a complete graph.

Proposition 9.9 *Let Assumptions 9.1–9.4 hold. If the graph underlying the network $\mathcal{G}$ is complete, then the transient power losses are bounded from above by:*

$$\|H^{\mathrm{dec}}\|_2^2 \leq \frac{\alpha}{2}(N-1)\left(\frac{1}{k_P} + \frac{1}{\tau_Q\left(\frac{c_Q}{N\underline{b}} + k_Q\right)}\right), \tag{9.19}$$

where $\underline{b}$ is the arithmetic mean of the susceptances b_{ij} of all network lines $(i,j) \in \mathcal{E}$. They are bounded from below by:

$$\|H^{\mathrm{dec}}\|_2^2 \geq \frac{\alpha}{2}(N-1)\left(\frac{1}{k_P} + \frac{1}{\tau_Q\left(\frac{c_Q}{Nb_{\min}} + k_Q\right)}\right), \tag{9.20}$$

where $b_{\min} = \min_{\mathcal{E}} b_{ij}$. If $b_{ij} = \underline{b} = b_{\min}$ for all $(i,j) \in \mathcal{E}$, (9.19)–(9.20) turn into equalities.

Proof: See Appendix 9.E.

While a complete graph is a reasonable model for existing power networks after Kron-reduction (see Section 9.1.1), future microgrids may arise through the addition of generation units at many or all nodes in distribution grids. Distribution grids typically have a radial network structure, meaning that a tree graph (a graph without cycles) is the underlying topology. With generation units at most nodes,

the tree graph structure can be maintained also after Kron-reduction. The following proposition bounds the transient losses for this case:

Proposition 9.10 *Let Assumptions 9.1–9.4 hold. If the graph underlying the network $\mathcal{G}$ is a tree graph, then the transient power losses are bounded by:*

$$\|H^{\mathrm{dec}}\|_2^2 \leq \frac{\alpha}{2}(N-1)\left(\frac{1}{k_P} + \frac{1}{\tau_Q(\frac{c_Q}{2\underline{b}} + k_Q)}\right). \tag{9.21}$$

Here, $\underline{b}$ is again the mean line susceptance. A lower bound is given by:

$$\|H^{\mathrm{dec}}\|_2^2 \geq \frac{\alpha}{2}(N-1)\left(\frac{1}{k_P} + \frac{1}{\tau_Q(\frac{c_Q N^2}{16 b_{\min}} + k_Q)}\right). \tag{9.22}$$

Proof: See Appendix 9.F.

In all cases, losses scale with the network size N and will grow unboundedly for large networks, in line with the conclusions in Chapter 8. The bound (9.22) is loose, but supports this conclusion. This can be seen from Figure 9.3, which displays the values of the $\mathcal{H}_2$ norm as a function of network size N for the topologies discussed in this section (a path graph is used as an example of a tree graph). We also include randomly generated planar graphs, whose performance lies in between the two, in the figure.

Finally, we show that increasing the network connectivity always leads to increased losses:

Proposition 9.11 *Adding an edge to the set $\mathcal{E}$ or increasing the susceptance b_{ij} of any edge $(i,j) \in \mathcal{E}$ (at constant α) can only increase the power losses given by $\|H^{\mathrm{dec}}\|_2^2$.*

Proof: Note that the function $\phi_n = (\frac{c_Q}{\lambda_n} + k_Q)^{-1}$ is monotonically increasing in λ_n (since $\frac{d}{d\lambda_n}\phi_n = \frac{1}{c_q\lambda_n + k_Q\lambda^2} > 0$). Remaining terms in $||H^{\mathrm{dec}}||_2^2$ are independent of λ_n. The proposition thus holds if it is true that at least one λ_n increases while the others do not decrease under the operations. This is precisely what is shown in Theorem 3.1 (adding an edge) and in Theorem 3.3 (increasing the susceptance of an edge). ■

Remark 9.5 Proposition 9.11 implies that (9.22) provides a lower bound for *any* underlying graph topology. □

The fact that a highly interconnected network incurs *larger* power losses in recovering or maintaining synchrony than a sparsely interconnected network stands in sharp contrast to our earlier result as well as to typical notions of power system stability, which we already discussed in Section 8.3.3. Figure 9.4 emphasizes the counterintuitiveness of this result. It shows the transient behaviors obtained from simulations of a 5 node network with, respectively, complete and path graph

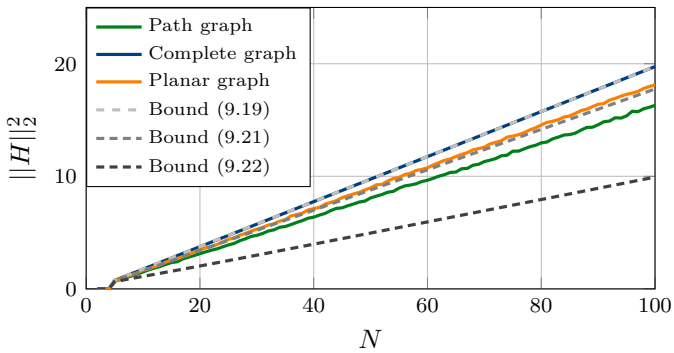


Figure 9.3: $\mathcal{H}_2$ norm in (9.15) for sample networks with path graph, complete graph, and random planar graph topologies, along with the upper bound (9.19), for complete graphs and (9.21) for tree graphs, and the lower bound (9.22). We note that performance in terms of transient losses scales linearly in network size in all cases. In this example, $k_P = k_Q = c_Q = 1$, $\alpha = 0.2$ and the line susceptances are uniformly distributed on the interval $[0.5, 3.25]$. The random planar graphs were generated using Delaunay triangulation.

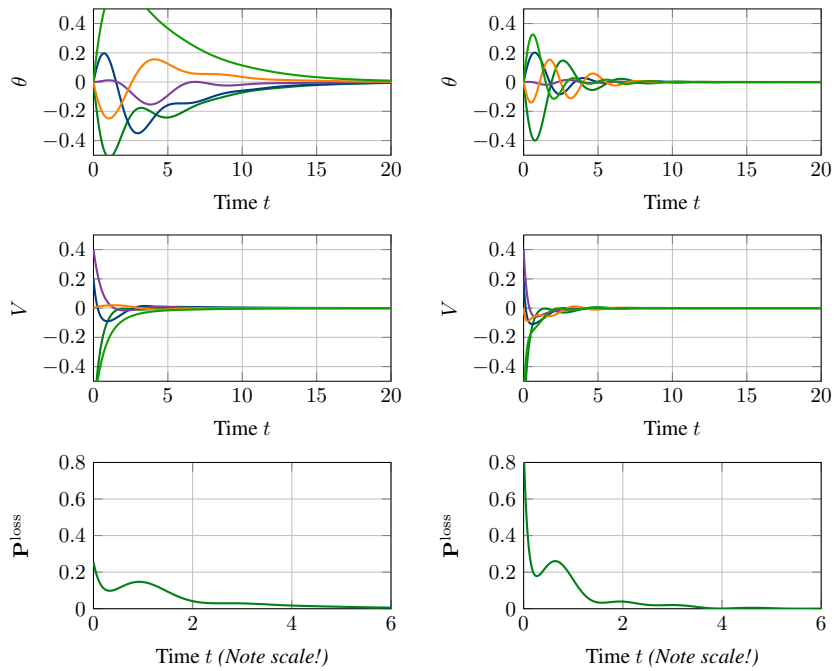
topologies. The plot clearly shows a faster convergence in the complete graph case. This faster convergence, however, comes at a greater cost in terms of power losses.

Some intuition behind this result may be obtained in terms of self-damping, or absolute feedback, which was a central topic in Chapter 4. Note that the voltage dynamics in (9.5), which are of first order, include an absolute feedback term $-V_i$. This means that a disturbance could theoretically be attenuated even if node i were isolated, since the system $\dot{V}_i = V_i^{\text{ref}} - V_i$ is in itself asymptotically stable. However, the network must be connected, and the voltage dynamics at node i therefore inevitably depend on neighboring nodes. The neighbors can support each other in attenuating disturbances and achieve faster voltage control, but this control is exerted through power flows, which give rise to losses. This means that any additional line connected to node i increases the number of flows and losses, and therefore, a more highly interconnected network leads to larger transient losses than a more loosely interconnected one.

9.5 Discussion

In this Chapter, we relaxed some of the simplifying assumptions from Chapter 8 and studied the impact of non-zero resistances and variable voltages on performance in terms of the Price of Synchrony. We used an inverter-based microgrid model, in which distributed generation sources are assumed grid-connected via inverters. These are, in turn, modeled as voltage sources subject to both voltage and frequency droop control.

Unsurprisingly, an explicit accounting for variable voltages strictly increases the



(a) Path graph (b) Complete graph

Figure 9.4: Simulations of frequency and voltage droop control (cross-coupled) with $N = 5$ inverters in, respectively, a path and a complete graph topology. Despite its higher rate of convergence, the complete graph incurs greater energy losses. These losses correspond to the area under the curves on the bottom panels.

transient losses compared to our previous results. This is true both for the case when the voltage dynamics are considered decoupled from the phase-frequency dynamics and for the case when the two systems are cross-coupled. The cross-couplings can be shown to only have a small contribution to the overall transient losses, in that it is proportional to high-order terms of the ratio α, which is (much) smaller than 1. On the other hand, the cross-coupling limits the set of admissible voltage and frequency droop controllers. For example, in a network with a large, complete graph as the underlying topology, the active power droop gain cannot be set too large in relation to the reactive power gain.

The additional transient losses that arise through voltage control were shown to increase with an increasing connectivity of the underlying network. That is, while additional links in the network improve network coherence and give faster convergence, they lead to greater transient losses. This is in sharp contrast to the previous results in this thesis. It can be thought of as a consequence of the performance measure on the one hand, which penalizes inter-nodal flows, and the

presence of absolute feedback on the other, which stabilizes the local dynamics independently of such flows. In the next chapter, we will see that a similar topology dependence holds with (noiseless) integral control that emulates absolute feedback.

Appendix to Chapter 9

9.A Diagonalization of system dynamics

Under Assumptions 9.1–9.4 the system (9.9) with the output y from Definition 9.1 can be diagonalized by a unitary state transformation as outlined in Section 3.1.2. As before, we use the matrix U that diagonalizes $\mathcal{L}_B$ (and $\mathcal{L}_G$, by Assumption 9.1) for this transformation.

We obtain the block-diagonalized system $\hat{H}$, which corresponds to N decoupled subsystems, or modes, $\hat{H}_n$:

$$\begin{aligned}
\begin{bmatrix} \dot{\hat{\theta}}_n \\ \dot{\hat{\omega}}_n \\ \dot{\hat{V}}_n \end{bmatrix} &= \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -\frac{k_P}{\tau_P}\lambda_n & -\frac{1}{\tau_P} & \frac{k_P}{\tau_P}\alpha\lambda_n \\ -\frac{k_Q}{\tau_Q}\alpha\lambda_n & 0 & -\frac{c_Q}{\tau_Q} - \frac{k_Q}{\tau_Q}\lambda_n \end{bmatrix}}_{\hat{\mathcal{A}}_n} \begin{bmatrix} \hat{\theta}_n \\ \hat{\omega}_n \\ \hat{V}_n \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 \\ \frac{1}{\tau_P} & 0 \\ 0 & \frac{1}{\tau_Q} \end{bmatrix}}_{\hat{\mathcal{B}}_n} \hat{\mathrm{w}}_n \\
\hat{y}_n &= \underbrace{\sqrt{\alpha\lambda_n}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\hat{\mathcal{C}}_n} \begin{bmatrix} \hat{\theta}_n \\ \hat{\omega}_n \\ \hat{V}_n \end{bmatrix}
\end{aligned} \tag{9.23}$$

for $n = 1, 2, \ldots, N$. Note that since $\lambda_1 = 0$, $\hat{y}_1 \equiv 0$. Therefore, even though $\hat{\mathcal{A}}_1$ is non-Hurwitz, the mode is not observable.

9.B Proof of Theorem 9.1

The system (9.9) is IO stable if and only if all observable (and controllable) system modes are asymptotically stable. After performing the diagonalization outlined in Appendix 9.A, this becomes equivalent to ensuring that $\hat{\mathcal{A}}_n$, $n = 2, \ldots, N$, from (9.23) are all Hurwitz. The characteristic equation $\det(zI - \hat{\mathcal{A}}_n) = 0$ for $n = 2, \ldots, N$ reads

$$a_3 z^3 + a_2 z^2 + a_1 z + a_0 = 0,$$

where

$$\begin{aligned}
a_3 &= \tau_P \tau_Q, \\
a_2 &= \tau_Q + \tau_P(c_Q + k_Q\lambda_n), \\
a_1 &= k_P\lambda_n(c_Q + k_Q\lambda_n(\alpha^2 + 1)), \\
a_0 &= c_Q + k_Q\lambda_n(\tau_Q + 1).
\end{aligned}$$

The coefficients a_i are all are positive since $\tau_P, \tau_Q, k_P, k_Q, c_Q, \lambda_n > 0$. Therefore, by the Routh-Hurwitz stability criterion all eigenvalues z lie strictly in the LHP if and only if $a_i > 0$ for $i = 0, \ldots, 3$ (clearly satisfied), and $a_1 a_2 > a_0 a_3$. Re-arranging these terms gives the condition

$$\frac{k_Q^2\tau_P\lambda_n^2 + (k_P\tau_Q^2 + k_Q\tau_Q + 2c_Q k_Q\tau_P)\lambda_n + \tau_P c_Q^2 + \tau_Q c_Q}{k_P k_Q \tau_P \tau_Q \lambda_n^2} > \alpha^2$$

for each $n = 2, \ldots, N$. The left hand side is decreasing in λ_n and the inequality is thus satisfied for all n if it is with λ_N. ■

9.C Proof of Proposition 9.5

After the block-diagonalization laid out in Appendix 9.A, it holds $\|H\|_2^2 = \sum_{n=2}^{N} \|\hat{H}_n\|_2^2$. Each subsystem norm is obtained as $\|\hat{H}_n\|_2^2 = \mathrm{tr}(\hat{\mathcal{B}}_n^T X_n \hat{\mathcal{B}}_n)$, where the Gramian X_n satisfies $\hat{\mathcal{A}}_n^T X_n + X_n \hat{\mathcal{A}}_n + \hat{C}_n^T \hat{C}_n = 0$. Define

$$X_n = X_n^{\mathrm{dec}} + X_n^1, \tag{9.24}$$

where X_n^{dec} is the observability Gramian of the decoupled subsystem, obtained as the solution to $\hat{A}_n^{\mathrm{dec}T} X_n^{\mathrm{dec}} + X_n^{\mathrm{dec}} \hat{A}_n^{\mathrm{dec}} + \hat{C}_n^T \hat{C}_n = 0$. Straightforward calculations give that

$$X_n^{\mathrm{dec}} = \frac{\alpha}{2} \begin{bmatrix} \lambda_n \tau_P + \frac{1}{k_P} & \frac{\tau_P}{k_P} & 0 \\ \frac{\tau_P}{k_P} & \frac{\tau_P}{k_P} & 0 \\ 0 & 0 & \frac{\tau_Q}{c_Q/\lambda_n + k_Q} \end{bmatrix}$$

and

$$\|\hat{H}_n^{\mathrm{dec}}\|_2^2 = \mathrm{tr}(\hat{\mathcal{B}}_n^T X_n^{\mathrm{dec}} \hat{\mathcal{B}}_n) = \frac{\alpha}{2k_P} + \frac{\alpha}{2\tau_Q} / (\frac{c_Q}{\lambda_n} + k_Q).$$

Using the partitioning (9.24) of X_n, the Lyapunov equation becomes $\hat{\mathcal{A}}_n^T X_n^1 + X_n^1 \hat{\mathcal{A}}_n + \hat{\mathcal{A}}_n^T X_n^{\mathrm{dec}} + X_n^{\mathrm{dec}} \hat{\mathcal{A}}_n + \hat{C}_n^T \hat{C}_n = 0$. Substituting X_n^{dec} and $\hat{C}_n^T \hat{C}_n$ into this gives that

$$\hat{\mathcal{A}}_n^T X_n^1 + X_n^1 \hat{\mathcal{A}}_n + \alpha^2 Z_n = 0, \tag{9.25}$$

where

$$Z_n = \frac{\lambda_n}{2} \begin{bmatrix} 0 & 0 & \frac{c_Q}{c_Q + k_Q \lambda_n} \\ 0 & 0 & \tau_P \\ \frac{c_Q}{c_Q + k_Q \lambda_n} & \tau_P & 0 \end{bmatrix}.$$

Now, $\|\hat{H}_n\|_2^2 = \|\hat{H}_n^{\mathrm{dec}}\|_2^2 + \mathrm{tr}(\hat{\mathcal{B}}_n^T X_n^1 \hat{\mathcal{B}}_n)$. Note that (9.25) is a Lyapunov equation and X_n^1 will be positive definite and $\mathrm{tr}(\hat{\mathcal{B}}_n^T X_n^1 \hat{\mathcal{B}}_n)$ bounded provided $\hat{\mathcal{A}}_n$ is Hurwitz, that is, if the condition from Theorem 9.1 is satisfied. We derive

$$\mathrm{tr}(\hat{\mathcal{B}}_n^T X_n^1 \hat{\mathcal{B}}_n) = \frac{1}{\tau_P^2} X_{22}^1 + \frac{1}{\tau_Q^2} X_{33}^1 = \alpha^3 \frac{\tilde{c}_1 \alpha^2 + \tilde{c}_2}{\tilde{c}_3 \alpha^4 + \tilde{c}_4 \alpha^2 + \tilde{c}_5},$$

where the coefficients $\tilde{c}_i$ are somewhat involved functions of λ_n and the system parameters, but not of α:

$$\begin{aligned}
\tilde{c}_1 &= k_P k_Q^2 \lambda_n^3 \tau_P (k_P \lambda_n + c_Q \tau_Q + k_Q \lambda_n \tau_Q) \\
\tilde{c}_2 &= (\tau_P k_P^2 k_Q^2 + \tau_P \tau_Q k_P k_Q^3)\lambda_n^4 + (2c_Q \tau_P k_P^2 k_Q + 2c_Q \tau_P \tau_Q k_P k_Q^2 - c_Q \tau_P k_Q^3)\lambda_n^3 \\
&\quad + (\tau_P c_Q^2 k_P^2 + \tau_P \tau_Q c_Q^2 k_P k_Q - 2\tau_P c_Q^2 k_Q^2 + \tau_Q c_Q k_P^2 - \tau_Q c_Q k_Q^2)\lambda_n^2 \\
&\quad - (k_Q \tau_P c_Q^3 + k_Q \tau_Q c_Q^2)\lambda_n \\
\tilde{c}_3 &= -2k_P^2 k_Q^2 \lambda_n^3 \tau_P \tau_Q (c_Q + k_Q \lambda_n) \\
\tilde{c}_4 &= 2k_P k_Q \lambda_n (c_Q + k_Q \lambda_n)(\tau_P c_Q^2 + 2\tau_P c_Q k_Q \lambda_n - k_P \tau_P c_Q \lambda_n \tau_Q + c_Q \tau_Q + \\
&\quad + \tau_P k_Q^2 \lambda_n^2 - k_P \tau_P k_Q \lambda_n^2 \tau_Q + k_Q \lambda_n \tau_Q + k_P \lambda_n \tau_Q^2) \\
\tilde{c}_5 &= 2k_P (c_Q + k_Q \lambda_n)^2 (\tau_P c_Q^2 + 2\tau_P c_Q k_Q \lambda_n + c_Q \tau_Q + \tau_P k_Q^2 \lambda_n^2 + k_Q \lambda_n \tau_Q + k_P \lambda_n \tau_Q^2)
\end{aligned}$$

For sufficiently small α we can perform a series expansion around $\alpha = 0$:

$$\operatorname{tr}(\hat{\mathcal{B}}_n^T X_n^1 \hat{\mathcal{B}}_n) = \frac{\tilde{c}_2}{\tilde{c}_5}\alpha^3 + \frac{\tilde{c}_1\tilde{c}_5 - \tilde{c}_2\tilde{c}_4}{\tilde{c}_5^2}\alpha^5 + \cdots =: c_{1,n}\alpha^3 + c_{2,n}\alpha^5 + \cdots ,$$

where the $c_{i,n}$, $i = 1, 2, \ldots$ are functions of λ_n. The full norm is now written

$$\|H\|_2^2 = \sum_{n=2}^{N} \|\hat{H}_n\|_2^2 = \sum_{n=2}^{N} \|\hat{H}^{\mathrm{dec}}\|_2^2 + \alpha^3 \sum_{n=2}^{N} c_{1,n} + \alpha^5 \sum_{n=2}^{N} c_{2,n} + \cdots .$$

Defining

$$c_0^\Lambda = \frac{1}{\alpha}\sum_{n=2}^{N} \|\hat{H}_n^{\mathrm{dec}}\|_2^2 = \frac{1}{k_P}(N-1) + \frac{1}{2\tau_Q}\sum_{n=2}^{N} \frac{1}{\frac{c_Q}{\lambda_n} + k_Q} = \frac{1}{\alpha}\|H^{\mathrm{dec}}\|_2^2$$

and $c_i^\Lambda = \sum_{n=2}^{N} c_{i,n}$ for $i = 1, 2, \ldots$, the result follows. ■

9.D Proof of Proposition 9.8

If $\lambda_n \to \infty$, then $\sum_{n=2}^{N} 1/\left(\frac{c_Q}{\lambda_n} + k_Q\right) \to \frac{1}{k_Q}(N-1)$ and (9.17) follows.

The higher order terms in $||H||_2^2$ are derived in line with the proof of Proposition 9.5, but noting that if $\lambda_n \to \infty$, then $\frac{c_Q}{c_Q + k_Q \lambda_n} \to 0$. Using this to solve (9.25) gives that

$$\operatorname{tr}\{\hat{B}_n^T X_n^1 \hat{B}_n\} = \alpha^3 \frac{k_P + k_Q \tau_Q}{2k_Q^2} \cdot \frac{1}{1 - \frac{k_P \tau_Q}{k_Q}\alpha^2} \tag{9.26}$$

in the limit where $\lambda_n \to \infty$. Now, provided $\alpha^2 \frac{k_P \tau_Q}{k_Q} < 1$ (equivalent to Corollary 9.2), the series expansion of (9.26) becomes

$$\mathrm{tr}\{\hat{B}_n^T X_n^1 \hat{B}_n\} = \alpha^3 \frac{k_P + k_Q \tau_Q}{2k_Q^2} \left(1 + \frac{k_P \tau_Q}{k_Q}\alpha^2 + \left(\frac{k_P \tau_Q}{k_Q}\right)^2 \alpha^4 + \cdots\right) = \frac{1}{N-1}(||\hat{H}||_2^2 - ||\hat{H}^{\mathrm{dec}}||_2^2).$$

Therefore,

$$\eta = \frac{\frac{\alpha^3}{2}\frac{k_P + k_Q\tau_Q}{k_Q^2}\sum_{k=0}^{\infty}\alpha^{2k}\left(\frac{k_P\tau_Q}{k_Q}\right)^k}{\frac{\alpha}{2}\left(\frac{1}{k_P}+\frac{1}{\tau_Q k_Q}\right)} = \alpha^2\frac{k_P\tau_Q}{k_Q}\sum_{k=0}^{\infty}\alpha^{2k}\left(\frac{k_P\tau_Q}{k_Q}\right)^k,$$

which is equivalent to (9.18). ■

9.E Proof of Proposition 9.9

Consider the function $\phi(x) = \frac{1}{\frac{1}{x}+k}$, which is concave for $x > 0$, $k \geq 0$ (since $\phi''(x) = \frac{-2k}{(1+kx)^3} < 0$). We have that $\lambda_n/c_Q > 0$ for $n = 2, \ldots, N$ and can therefore apply Jensen's inequality of the form $\sum_{i=1}^m \phi(x_i) \leq m\phi\left(\frac{1}{m}\sum_{i=1}^m x_i\right)$ to (9.15) to obtain:

$$||H^{\mathrm{dec}}||_2^2 \leq \frac{\alpha}{2k_P}(N-1) + \frac{\alpha}{2\tau_Q}(N-1)\frac{1}{\frac{c_Q}{\frac{1}{N-1}\sum_{n=2}^N \lambda_n} + k_Q}. \tag{9.27}$$

By the definition of $\mathcal{L}_B$, the average of its $N-1$ non-zero eigenvalues is

$$\frac{1}{N-1}\sum_{n=2}^N \lambda_n = \frac{\mathrm{tr}\{\mathcal{L}_B\}}{N-1} = \frac{2\sum_{\mathcal{E}} b_{ij}}{N-1} = \frac{2\frac{N(N-1)}{2}\underline{b}}{N-1} = N\underline{b},$$

where $\underline{b}$ is the arithmetic mean of the susceptances of the $N(N-1)/2$ edges in the complete graph. Substituting the above into (9.27) yields the result (9.19).

Given that $\phi(x)$ is monotonically increasing in x, the inequality (9.20) is derived by setting $\mathcal{L}_B = b_{\min}\mathcal{L} + \Delta\mathcal{L}_B$. Here, $\mathcal{L}$ is an unweighted complete graph Laplacian, and $\Delta\mathcal{L}_B$ is a complete graph Laplacian with edge weights $b_{ij} - b_{\min} \geq 0$. Since $\mathcal{L}$ and $\Delta\mathcal{L}_B$ are simultaneously diagonalizable [164, Lemma A.1], $\lambda_n = b_{\min}N + \lambda_n^{\Delta B} \geq b_{\min}N$. If $\mathcal{L}_B = b_{\min}\mathcal{L}$, $\Delta\mathcal{L}_B = 0$ and (9.20) holds with equality. ■

9.F Proof of Proposition 9.10

Again, the Jensen inequality from the proof of Theorem 9.9 is applied. Here, the average of the $N-1$ non-zero eigenvalues in (9.27) is

$$\frac{1}{N-1}\sum_{n=2}^N \lambda_n = \frac{\mathrm{tr}\{\mathcal{L}_B\}}{N-1} = \frac{2\sum_{\mathcal{E}} b_{ij}}{N-1} = \frac{2(N-1)\underline{b}}{N-1} = 2\underline{b},$$

where $\underline{b}$ is the mean susceptance of the $(N-1)$ edges of any tree graph.

Noting that the function $\phi(x)$ is monotonically increasing in x, in our case, in λ_n, we obtain the lower bound by setting $\lambda_n = \lambda_{\min} = b_{\min}\lambda_2^{\mathcal{P}_N}$ for all $n = 2, \ldots, N$. Here, $\lambda_2^{\mathcal{P}_N} = 2(1 - \cos\frac{2\pi}{N})$ is the algebraic connectivity of an unweighted path graph $\mathcal{P}_N$, which is the tree graph with the smallest algebraic connectivity (see, for example, [152]). Using the fact that for any $x \in [-\pi, \pi]$, $1 - \cos x \geq \frac{2}{\pi^2}x^2$, we obtain that $\lambda_2^{\mathcal{P}_N} \geq \frac{16}{N^2}$ and the bound follows. ■

Chapter 10

Advantages and Limitations of Distributed Secondary Frequency Control

So far, we have considered synchronization performance in power networks with static feedback control. That is, with the frequency droop controller. While droop control, under reasonable conditions (see for example [159]), is successful at stabilizing the network while achieving power sharing, it generally causes the steady-state network frequency to deviate from its nominal value (50 Hz or 60 Hz) [131]. This deficiency motivates so-called secondary frequency control, the objective of which is to eliminate the stationary error. To achieve this objective, architectures with various degrees of centralization have been proposed. Completely decentralized secondary controllers tend to destroy the power sharing properties established by droop control and may lead to a violation of generation constraints [6, 42]. Conversely, centralized control requires a network architecture that conflicts with a scalable distributed generation paradigm. This has motivated the development of distributed control algorithms that simultaneously eliminate frequency errors and maintain the optimality properties of droop control [6, 159, 163, 196].

The DAPI controller, which we have treated earlier in this thesis, is one such controller. We will re-visit this controller in this chapter, and evaluate its performance in terms of transient power losses, that is, the Price of Synchrony. In particular, we will compare it to a corresponding centralized averaging proportional-integral controller (CAPI), which we briefly introduced in Chapter 7. This chapter's main results show that DAPI control has the potential to significantly reduce transient power losses compared to the static droop controller, as well as to the CAPI controller, which has the same performance. The room for performance improvement is shown to be greater in a sparsely connected network.

However, the DAPI controller's performance and scalability are significantly limited by measurement noise. This means that unless it is carefully tuned – and re-tuned as the network grows – the DAPI controller can cause much larger transient losses than droop or CAPI control. In fact, frequency measurement noise leads to

transient losses that, though initially small, can grow faster than linearly in network size. In very large and sparse networks, we conclude in line with Chapter 7 that the centralized CAPI controller may be preferable for secondary frequency control.

10.1 Distributed vs. centralized secondary frequency control

We will now introduce models for secondary frequency control in power networks. The purpose of secondary frequency control is to eliminate any stationary frequency errors. It is achieved through integral action. We therefore re-introduce the DAPI controller from Chapter 4 and the CAPI controller from Chapter 7 and discuss their implementation in a frequency control context.

The setup in this chapter is the same as in the previous, that is, we consider the frequency droop controller introduced there for inverter-based networks. Recall, however, that while the notation differs slightly, this system is qualitatively equivalent to the swing dynamics from Chapter 8.

10.1.1 Droop-controlled power network

Consider again a Kron-reduced model of a power network and denote the underlying graph by $\mathcal{G}^P = \{\mathcal{V}, \mathcal{E}^P\}$. Under standard droop control, and assuming linearized power flows, the phase θ_i and frequency deviation ω_i at each generator (or inverter) $i \in \mathcal{V}$ evolves as:

$$\begin{aligned} \dot{\theta}_i &= \omega_i \\ \tau_i \dot{\omega}_i &= -\omega_i - k_i \sum_{j \in \mathcal{N}_i^P} b_{ij}(\theta_i - \theta_j) + P_i + u_i, \end{aligned} \tag{10.1}$$

where $\mathcal{N}_i^P$ is the neighbor set of node i in $\mathcal{G}^P$ and u_i is the input from the secondary controller, to be introduced shortly. The net power injection P_i will, as before, be modeled as a white stochastic disturbance input that is uncorrelated across the nodes (though other input scenarios are also meaningful, see Section 8.2.1).

10.1.2 CAPI control

The droop-controlled power system (10.1) can be shown to return to a stable operating point. However, since droop control is effectively a proportional control law, the equilibrium will in general *not* be the desired operating point at $\omega = 0_N$ (recall, $\omega \in \mathbb{R}^N$ is the vector of frequency deviations). This issue can be overcome by designing the secondary control input u to make $\omega = 0_N$ the only stable equilibrium.

One such design is to add an averaging integral control term, resulting in a *centralized averaging proportional integral (CAPI)* controller, which was proposed in [5], and generalized to arbitrary positive weights in [41,43]. The CAPI controller requires the control signal u to be computed centrally through an integration of the

average frequency deviation and, subsequently, to be broadcast to all generators. The strategy thus resembles traditional automated generation control (AGC). The controller takes the form:

$$\begin{aligned} u_i^{\mathrm{CAPI}} &= z \\ q_I \dot{z} &= -\frac{1}{N}\sum_{i=1}^{N} \hat{\omega}_i \end{aligned} \tag{10.2}$$

where $\hat{\omega}_i$ is the frequency measured at node i and $q_I > 0$ is a controller gain (an inverted integral gain). Notice that the control input u_i is identical at all nodes $i \in \mathcal{V}$.

10.1.3 DAPI control

DAPI controllers have been proposed for frequency control of synchronous generator networks as well as for microgrids [7,159,178]. By including a distributed averaging of the integral states, this controller eliminates the need for a central control entity and can be implemented in a distributed fashion. The distributed averaging takes place over an additional *communication layer* that needs to be introduced on top of the physical power network layer, see Figure 7.1. We model this layer by the graph $\mathcal{G}^C = \{\mathcal{V}, \mathcal{E}^C\}$, which is assumed to be undirected and connected.

Here, we write the DAPI controller as

$$\begin{aligned} u_i^{\mathrm{DAPI}} &= -z_i \\ q_{I,i}\dot{z}_i &= \omega_i - \sum_{j \in \mathcal{N}_i^C} c_{ij}(z_i - z_j), \end{aligned} \tag{10.3}$$

where $c_{ij} = c_{ji} > 0$ and $q_{I,i} > 0$ are controller gains and $\mathcal{N}_i^C$ is the set of generators that generator i can communicate with in $\mathcal{G}^C$.

The DAPI controller has been shown to achieve the important property of stationary power sharing, implying that the generated power of all generators is equal at steady state [7, 159]. It can also be modified to achieve *weighted* power sharing [159].

We remark that the DAPI controller corresponds to the CAPI controller in the limit of infinite interaction gains c_{ij}. In this case, the distributed averaging is infinitely fast, and the integral states z_i in (10.3) are identical to z in (10.2). In the other limit, when $c_{ij} \to 0$, the control is entirely decentralized. In this case, the slightest measurement errors cause a drift in the integral states that destabilizes the system (see also Section 4.3.1). In power networks, decentralized secondary frequency control on this form is only implementable with phasor measurement units (PMUs) at every node [6].

Remark 10.1 Any distributed control law can also be implemented through a centralized controller architecture. Therefore, if the communication structure for the CAPI controller is in place, it can also be used to implement the DAPI control

law in (10.3). In this chapter, a comparison of distributed and centralized integral control should, however, be understood as a comparison of the particular DAPI and CAPI controllers. □

10.1.4 Noise model

Similar to Chapter 7, we will here consider the effect of imperfect frequency measurements and therefore introduce the additional noise term η_i, which is modeled as additive white measurement noise on the frequency, so that

$$\hat{\omega}_i = \omega_i + \eta_i. \tag{10.4}$$

As we will see, this noise has a large impact on the relative performance of the secondary controllers.

The power injection fluctuations P_i were modeled as uncorrelated white disturbance inputs. We choose to relate their intensity to the white measurement noise η_i through the constant $\varepsilon > 0$, so that

$$\mathbb{E}\{\eta_i(t')\eta_i^T(t)\} = \varepsilon\mathbb{E}\{P_i(t')P_i^T(t)\},$$

where $P = \begin{bmatrix}P_1, \dots, P_N\end{bmatrix}^T$ and $\eta = \begin{bmatrix}\eta_1, \dots, \eta_N\end{bmatrix}^T$. This allows us to define the process $\mathrm{w} \in \mathbb{R}^{2N}$ with $\mathbb{E}\{\mathrm{w}(t')\mathrm{w}^T(t)\} = \delta(t-t')I$, and construct the input vector

$$\begin{bmatrix}P\\ \eta\end{bmatrix} = \begin{bmatrix}I & 0\\ 0 & \varepsilon I\end{bmatrix}\mathrm{w}. \tag{10.5}$$

10.1.5 Closed-loop system dynamics

By substituting the secondary controller inputs u_i^{DAPI} from (10.3) and u_i^{CAPI} from (10.2) along with the measurement noise model (10.4) into the system dynamics (10.1), and by making use of the input vector w in (10.5), we obtain the closed-loop systems on vector form as:

$$\begin{bmatrix}\dot{\theta}\\ T\dot{\omega}\\ q_I\dot{z}\end{bmatrix} = \begin{bmatrix}0 & I & 0\\ -K\mathcal{L}_B & -I & I\\ 0 & -\frac{1}{N}\mathbf{1}^T & 0\end{bmatrix}\begin{bmatrix}\theta\\ \omega\\ z\end{bmatrix} + \begin{bmatrix}0 & 0\\ I & 0\\ 0 & \varepsilon\mathbf{1}^T\end{bmatrix}\mathrm{w} \tag{10.6}$$

for the CAPI controlled system, and

$$\begin{bmatrix}\dot{\theta}\\ T\dot{\omega}\\ Q_I\dot{z}\end{bmatrix} = \begin{bmatrix}0 & I & 0\\ -K\mathcal{L}_B & -I & I\\ 0 & -I & -\mathcal{L}_C\end{bmatrix}\begin{bmatrix}\theta\\ \omega\\ z\end{bmatrix} + \begin{bmatrix}0 & 0\\ I & 0\\ 0 & \varepsilon I\end{bmatrix}\mathrm{w} \tag{10.7}$$

for the DAPI controlled system. At this point, we have (re-)introduced the state vectors $\theta = \begin{bmatrix}\theta_1, \dots, \theta_N\end{bmatrix}^T$, $\omega = \begin{bmatrix}\omega_1, \dots, \omega_N\end{bmatrix}^T$, and in the DAPI case $z = \begin{bmatrix}z_1, \dots, z_N\end{bmatrix}^T$ (note that z is scalar in the CAPI case), as well as the matrices $K = \mathrm{diag}\{k_i\}$, $T = \mathrm{diag}\{\tau_i\}$, and $Q_I = \mathrm{diag}\{q_{I,i}\}$.

Remark 10.2 In this model, we have assumed that the frequency ω_i enters without noise in the system dynamics (10.1). This is meaningful in a setting with synchronous machines, but a power electronic inverter that interfaces a renewable generation source may exert droop control using a noisy measurement of the frequency. In this case, the DAPI controlled system would become

$$\begin{bmatrix} \dot{\theta} \\ T\dot{\omega} \\ Q_I \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 \\ -K\mathcal{L}_B & -I & I \\ 0 & -I & -\mathcal{L}_C \end{bmatrix} \begin{bmatrix} \theta \\ \omega \\ z \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ I & \varepsilon I \\ 0 & \varepsilon I \end{bmatrix} \text{w}. \tag{10.8}$$

This leads to correlations in the noise input between the input channels. These correlations do, however, not affect the qualitative behavior of the system (see Appendix 10.B). We therefore limit the upcoming analysis to the uncorrelated inputs modeled in (10.6) and (10.7). □

10.2 Performance analysis

We now evaluate the performance of the systems under, respectively, CAPI and DAPI control, with and without frequency measurement noise, and compare it to the droop-controlled system. We will again characterize performance in terms of the transient resistive power losses. That is, we use the Price of Synchrony performance metric from Definition 8.1.

The closed-form expressions in this section will be derived under the following assumptions:

Assumption 10.1 (Identical generators and controller parameters) All synchronous generators have identical inertia and damping coefficients, and the secondary frequency controllers have identical settings. That is, $K = \text{diag}\{k\}$, $T = \text{diag}\{\tau\}$, and $Q_I = \text{diag}\{q_I\}$.

Assumption 10.2 (Uniform resistance-to-reactance ratios) The resistance-to-reactance, equivalently conductance-to-susceptance, ratio of all lines are uniform and constant. That is,

$$\alpha := \frac{g_{ij}}{b_{ij}}, \tag{10.9}$$

for all $(i, j) \in \mathcal{E}$. This implies $\mathcal{L}_G = \alpha \mathcal{L}_B$.

Assumption 10.3 (Topology for $\mathcal{G}^C$) The topology of the communication network $\mathcal{G}^C$ used in the DAPI control law (10.3) is identical to that of the physical network $\mathcal{G}$. We also assume

$$\mathcal{L}_C = \gamma \mathcal{L}_B, \tag{10.10}$$

that is, $\gamma = \frac{c_{ij}}{b_{ij}}$, with $\gamma \geq 0$, for all $(i, j) \in \mathcal{E} = \mathcal{E}^C$.

Assumption 10.3 implies that the communication network layer for the DAPI controller is set up along the physical network lines, and is shown in [6] to constitute a

sufficient criterion for load sharing with minimized generation costs. Note that this assumption was not imposed in Chapter 7, which considered a spatially invariant network. The property (10.10) says that the distributed averaging gains are set in proportion to the line susceptances b_{ij}, and will help us obtain tractable analytic expressions for the $\mathcal{H}_2$ norm of (10.7).

10.2.1 Expressions for the Price of Synchrony

Consider now the input-output mappings H^{CAPI} given by (10.6) with the performance output y from Definition 8.1, and H^{DAPI} given by (10.7) with the same output. Let H^{droop} be the corresponding system with a zero secondary control input. Their respective $\mathcal{H}_2$ performance is given in the following proposition:

Theorem 10.1 (Performance of DAPI and CAPI) *Let Assumptions 10.1–10.3 hold. The $\mathcal{H}_2$ norm for the DAPI-controlled system H^{DAPI} is given by*

$$||H^{\mathrm{DAPI}}||_2^2 = ||H_P^{\mathrm{DAPI}}||_2^2 + ||H_\eta^{\mathrm{DAPI}}||_2^2, \tag{10.11}$$

where

$$||H_P^{\mathrm{DAPI}}||_2^2 = \frac{\alpha}{2k} \sum_{i=2}^{N} \frac{1}{1+\varphi(\lambda_i,\gamma,k,q_I,\tau)^{-1}} \tag{10.12}$$

is the expected power loss associated with the power injection noise P and

$$||H_\eta^{\mathrm{DAPI}}\|_2^2 = \varepsilon^2 \frac{\alpha}{2k} \sum_{i=2}^{N} \frac{1}{\gamma\lambda_i} \cdot \frac{1}{1+\varphi(\lambda_i,\gamma,k,q_I,\tau)} \tag{10.13}$$

is the expected loss associated with the frequency measurement noise η. The function φ is defined as

$$\varphi(\lambda_i,\gamma,k,q_I,\tau) := \frac{kq^2\lambda_i + q_I\gamma\lambda_i + \tau(\gamma\lambda_i)^2}{q_I+\tau\gamma\lambda_i}, \tag{10.14}$$

where λ_i with $0=\lambda_1<\lambda_2\leq\ldots\leq\lambda_N$ are the eigenvalues of $\mathcal{L}_B$.

The $\mathcal{H}_2$ norm for the CAPI-controlled system H^{CAPI} is given by

$$\|H^{\mathrm{CAPI}}\|_2^2 = \|H_P^{\mathrm{CAPI}}\|_2^2 = \frac{\alpha}{2k}\,(N-1)\,.$$

$\|H^{\mathrm{CAPI}}\|_2^2$ is independent of any measurement noise η and equal to $\|H^{\mathrm{droop}}\|_2^2$.

Proof: See Appendix 10.A.

Remark 10.3 The norm expression (10.13) is positive and represents the additional power losses arising due to the frequency measurement noise η in the DAPI controlled system. In the CAPI case, η does not give rise to any additional power losses. This can be explained by the fact that the CAPI controller affects all generators equally. Any error caused by the measurement noise may affect the synchronous frequency, but will not induce additional power flows between generators. □

We note that the function $\varphi_i := \varphi(\lambda_i, \gamma, k, q_I, \tau)$ that appears in $\|H^{\mathrm{DAPI}}\|_2^2$ is positive. In the absence of any measurement noise, the following corollary is therefore readily verified:

Corollary 10.2 *For any $k, q_I, \tau, \gamma > 0$, and independently of the underlying network graph, it holds*

$$\|H_P^{\mathrm{DAPI}}\|_2^2 < \|H_P^{\mathrm{CAPI}}\|_2^2,$$

that is, the expected power losses due to the power injection noise P are smaller with the DAPI controller than with the CAPI controller.

This result is in line with the findings in Chapter 4, though the performance metric here differs. It shows that, in the absence of noise, the Price of Synchrony performance of the DAPI controller is superior to that of the CAPI controller. Perhaps surprisingly, the CAPI controller does not offer a benefit compared to the droop controller. This can be understood by noting that CAPI exerts integral action only on the average state, and affects all nodes equally. It therefore does not affect the inter-nodal flows and the synchronization itself, but it shifts the equilibrium point.

The measurement noise gives rise to the term $\|H_\eta^{\mathrm{DAPI}}\|_2^2$, which is strictly detrimental to the DAPI controller's performance. The careful reader will have noticed the factor $\frac{1}{\lambda_i}$ appearing the expression, which causes an unfavorable scaling of performance in network size. We will discuss this scaling in Section 10.4. Until then, we focus on characterizing performance in networks with a fixed number of generators.

10.2.2 The role of network connectivity

With standard droop control and under the given assumptions, the expected transient losses are entirely independent of network topology. This observation was discussed in detail in Chapter 8. Theorem 10.1 shows that the same is true under CAPI control. However, the losses incurred under DAPI control (10.11) do depend on network topology through the eigenvalues λ_n of $\mathcal{L}_B$. We now analyze this topology dependence for noiseless and noisy case separately.

Noiseless DAPI

In the absence of measurement noise, the transient losses will grow with increased network connectivity:

Proposition 10.3 *Adding an edge to the network $\mathcal{G}^P$ or increasing its susceptance (at constant α) can only* increase *the expected power losses associated with the power injection noise $\|H_P^{\mathrm{DAPI}}\|_2^2$ in* (10.12)*, and vice versa.*

Proof: It is readily verified that

$$\frac{\mathrm{d}\varphi_i}{\mathrm{d}\lambda_i} = \frac{2\lambda_i\gamma^2\tau q_I + \gamma q_I^2 + k q_I^3}{(q_I + \tau q_I \lambda_i)^2} > 0,$$

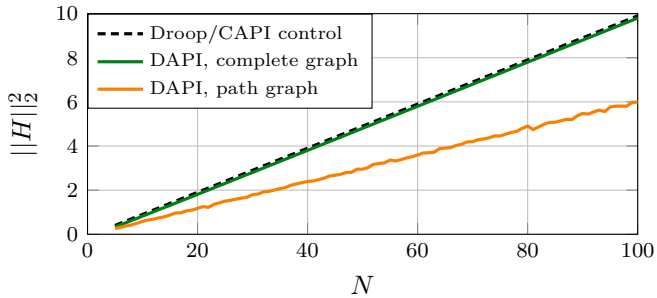


Figure 10.1: $\mathcal{H}_2$ norms with droop, CAPI and noiseless DAPI control for sample networks of size N with path graph and complete graph topologies. Note that $||H_P^{\mathrm{CAPI}}||_2^2$ in (10.1) is topology-independent. In all cases, losses increase with network size, but they are smaller for DAPI control in a path graph. Here, we have set $k = \gamma = m = 1$ and drawn the line susceptances b_{ij} from a uniform distribution on $[0.5, 1.5]$.

for all $\lambda_i > 0$ and therefore that $\frac{\mathrm{d}}{\mathrm{d}\lambda_i}\|H_P^{\mathrm{DAPI}}\|_2^2 > 0$ for all $i = 2, \ldots, n$. The proposition therefore holds true if at least one eigenvalue λ_i of $\mathcal{L}_B$ increases while the others do not decrease. This is precisely what is shown in Theorem 3.1 (adding an edge) and Theorem 3.3 (increasing the susceptance of an edge). ■

Proposition 10.3 implies, in particular, that the relative performance improvement of DAPI control over droop and CAPI control will be largest for sparse network topologies, such as those found in radial distribution networks. The best performance is achieved in a tree graph, as they have the fewest possible edges: $N - 1$. The least performance improvement can be expected for a complete graph topology, as it has every possible edge. In Figure 10.1 we compare such a topology to a path graph with respect to the results in Theorem 10.1. Although the losses for both topologies can be seen to grow with the network size, the comparison confirms the lower losses obtained in the path graph case.

The fact that a sparsely interconnected network outperforms a highly interconnected network by incurring smaller power losses is somewhat counterintuitive. We discussed in Section 9.4 that the result can be attributed to absolute state feedback. In the case of DAPI control, absolute phase feedback is emulated through an integration of frequency measurements. This topology dependence holds if frequency measurement noise is absent or small. If it is not, the conclusions change.

DAPI under measurement noise

Next, consider the losses associated with frequency measurement noise. This expression has the *inverse* dependence on network connectivity compared to the noiseless loss term:

Proposition 10.4 *Adding an edge to the network $\mathcal{G}^P$ or increasing its susceptance can only* decrease *the expected power losses associated with the frequency measurement noise $\|H_\eta^{\mathrm{DAPI}}\|_2^2$ in* (10.13)*, and vice versa.*

Proof: Since $\frac{\mathrm{d}\varphi_i}{\mathrm{d}\lambda_i} > 0$ and $\lambda_i > 0$ for $i = 2, \dots, N$, it holds $\frac{\mathrm{d}}{\mathrm{d}\lambda_i}[\gamma\lambda_i(1+\varphi_i)]^{-1} < 0$. Since α is invariant to edge addition, but may decrease if susceptances increase, we have that $\frac{\mathrm{d}}{\mathrm{d}\lambda_i}\|H_\eta^{\mathrm{DAPI}}\|_2^2 < 0$ for all $i = 2, \dots, n$, and the proposition follows. ■

Remark 10.4 Proposition 10.4 does not require conductances g_{ij} to be constant, but holds if the ratio $\alpha = \frac{g_{ij}}{b_{ij}} = \text{const.}$ or decreasing. □

This result is in line with those in Chapter 7, where we showed that increasing the connectivity of the communication network in DAPI improved its performance. It shows that inter-nodal alignment reduces the impact of measurement noise. At the same time, it leads to more power flows that induce losses.

Propositions 10.3–10.4 imply that the *total* transient losses' dependence on the network connectivity is not straightforward in the presence of measurement noise. The best network topology for loss reduction thus depends on remaining system parameters, and in particular, on the relative noise intensity parameterized by ε.

10.3 Control design for loss reduction

In the previous section, we established that the noiseless DAPI controller outperforms the droop and the CAPI controller in terms of transient resistive losses. We will now focus on optimizing the performance of the DAPI controller by tuning its parameters. In general, this amounts to minimizing the closed-form expressions in Theorem 10.1. Such an exercise is straightforward if the system parameters, in particular the eigenvalues of the susceptance matrix, are known. This is, however, typically not the case. The objective of this section is therefore to provide general insights to the optimal controller design. We approach the problem in two steps: first, we optimize the noiseless DAPI controller and then discuss how the optimum is affected by measurement noise.

10.3.1 Optimizing noiseless DAPI performance

Let us assume that the droop controller parameters are fixed and the design of the DAPI controller for performance is of interest. By Assumptions 10.1 and 10.3 the problem then reduces to minimizing $\|H_P^{\mathrm{DAPI}}\|_2^2$ over (positive) γ and q_I, assuming a fixed inertia constant τ and droop gain k. Consider therefore the function

$$f_i(q_I, \gamma) := \varphi_i^{-1} = \frac{q_I + \tau\gamma\lambda_i}{kq_I^2\lambda_i + q_I\gamma\lambda_i + \tau(\gamma\lambda_i)^2}, \tag{10.15}$$

that appears in the denominator of $\|H_P^{\mathrm{DAPI}}\|_2^2$ in (10.12). We note that $\lim_{q_I \to 0} f_i(q_I, 0) = \infty$, implying that for a communication gain $\gamma = 0$ and an

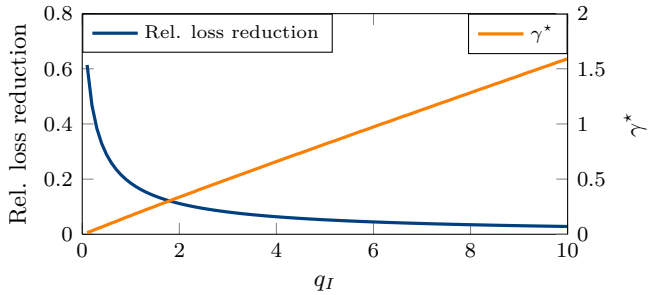


Figure 10.2: Relative loss reduction with DAPI control (as given by $(\|H_P^{\mathrm{CAPI}}\|_2^2 - \|H^{\mathrm{DAPI}}\|_2^2)/\|H^{\mathrm{CAPI}}\|_2^2$) for a test network based on the IEEE 57 bus benchmark system topology, at $\gamma = \gamma^\star$, as function of q_I. Here, $k = \tau = 1$.

inverse integral gain $q_I \to 0$,

$$\|H_P^{\mathrm{DAPI}}\|_2^2 = 0.$$

That is, the losses are reduced to zero. Similarly, $\lim_{q_I \to \infty} f_i(q_I, \gamma) = 0$, so for any fixed γ,

$$\|H_P^{\mathrm{DAPI}}\|_2^2 = \|H^{\mathrm{CAPI}}\|_2^2 = \frac{\alpha}{2k}(N-1),$$

that is the performance of the noiseless DAPI controller approaches that of the CAPI (or, equivalently, droop) controller. The same is also true in the limit where $\gamma \to \infty$ for any q_I. This is expected, as increasing the communication gain γ in the DAPI controller to infinity implies arbitrarily fast distributed averaging, ultimately resembling the centralized averaging in CAPI control.

Setting $q_I = 0$, which can achieve zero synchronization cost, is not practically feasible as it would require an infinite control effort. A question of practical relevance is therefore that of choosing the optimal γ given a fixed q_I, corresponding to the desired or practically feasible integral control effort.

A numerical example is shown in Figure 10.2. It displays the relative performance improvement achieved through DAPI control as a function of q_I, along with the corresponding optimal value for γ, for a hypothetical network based on the IEEE 57-bus benchmark system topology [116]. To address the question analytically, we begin by considering a special case.

Special case: complete graphs

We first consider the special case of a complete graph with uniform susceptances $b_{ij} = b$ as the underlying topology. In this case, the eigenvalues of $\mathcal{L}_B$ satisfy $\lambda_2 = \lambda_3 = \ldots = \lambda_N = bN$. Therefore, minimizing $\|H_P^{\mathrm{DAPI}}\|_2^2$ with respect to γ is equivalent to maximizing $f_i(q_I, \gamma)$ from (10.15), with $\lambda_i = bN$. We can thus derive a closed-form expression for the optimal value of γ, denoted $\gamma^\star$.

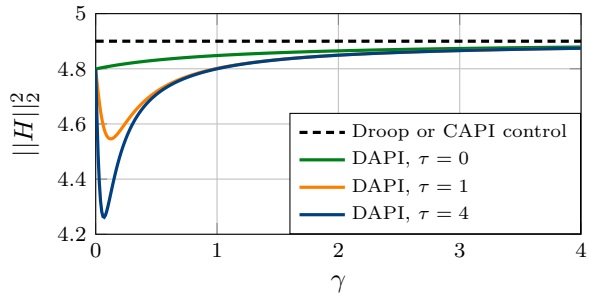


Figure 10.3: $\mathcal{H}_2$ norms in (10.12) as a function of γ for a complete graph with $N = 50$ nodes. Here, $k = q_I = 1$, and the inertia constant $\tau \in \{0, 1, 4\}$.

Proposition 10.5 *Let Assumptions 10.1–10.3 hold. If the graph $\mathcal{G}$ underlying the power network is complete and the line susceptances $b_{ij} = b$ for all $(i, j) \in \mathcal{E}$, then the $\gamma^\star$ that minimizes $\|H_P^{\mathrm{DAPI}}\|_2^2$ is given by*

$$\gamma^\star = \frac{q_I}{Nb\tau}\left(\sqrt{Nbk\tau} - 1\right) \tag{10.16}$$

if $Nbk\tau > 1$. Otherwise, $\gamma^\star = 0$.

Proof: First, let us re-write the function $f_i(q_I, \gamma)$ from (10.15) as

$$f_i(q_I, \gamma) = \frac{1}{\gamma\lambda_i + \frac{kq^2\lambda_i}{q_I + \tau\lambda_i}} =: \frac{1}{g_i(q_I, \gamma)}.$$

Maximizing $f_i(q_I, \gamma)$ is now equivalent to minimizing $g_i(q_I, \gamma)$. Differentiating with respect to γ, and setting the derivative equal to zero gives

$$\frac{\mathrm{d}g_i(q_I, \gamma)}{\mathrm{d}\gamma} = \lambda_i - \frac{\tau k q_I^2 \lambda_i}{(q_I + \tau\gamma\lambda)^2} = 0 \quad \Leftrightarrow \quad (q_I + \tau\gamma\lambda)^2 - \tau k q_I^2 \lambda_i = 0,$$

which gives that

$$\gamma_i^\star = \frac{-q_I + \sqrt{\tau k \lambda_i} q_I}{\tau\lambda_i}.$$

From this, we conclude that there is a positive extreme point $\gamma_i^\star$ if and only if $\sqrt{\tau k \lambda_i} > 1$. If it exists, this positive extreme point will minimize $g_i(q_I, \gamma)$ since $\frac{\mathrm{d}^2 g_i(q_I,\gamma)}{d\gamma^2} = 2\frac{\tau^2 k q_I^2 \lambda_i^2}{(q_I + \tau\gamma\lambda)^3} > 0$ for $\gamma > 0$. Otherwise, $g_i(q_I, \gamma)$ is minimized by the lower end point of the interval in γ, i.e., $\gamma = 0$. Recalling now that $\lambda_i = bN$ for $i = 2, \ldots, N$, it holds that $\gamma_i^\star$ is equal for all i, and the result follows. ■

Proposition 10.5 answers the question of when the DAPI controller has a theoretical performance that is superior to completely decentralized PI control where $\gamma = 0$.

Recall that $\gamma = 0$ is infeasible since it would lead to instability (as clearly seen from the upcoming Corollary 10.7.) The case $\gamma^\star > 0$ holds, for example, if the inertia constant τ or the droop gain k is large. The optimum's dependence on τ is also illustrated in Figure 10.3.

We also note that the optimal $\gamma^\star$ is generally small. Yet, it increases, not only in q_I as we already discussed, but also in the droop gain k. An intuitive explanation for this can be given as follows. Larger gains k and q_I cause disturbances to spread across the network rather than to be suppressed locally. A DAPI controller with larger communication gain γ anticipates this spread by faster distributing the integral action across the generators.

Finally, we note that for large N, the optimal $\gamma^\star$ again approaches zero. This means that, the larger the graph, the less important is the distributed averaging in *noiseless* DAPI. This conclusion is, however, no longer true when the controller is subject to measurement noise, as we will discuss shortly.

Remark 10.5 Proposition 10.5 is equivalent to Proposition 4.7, though the performance objectives differ. This means that $\gamma^\star$ is optimal with respect to both the network coherence and the Price of Synchrony. □

General graphs with arbitrary susceptances

We now generalize the results of the previous section to power networks with general underlying graph structures and present a result similar to Proposition 4.6. In this case, the eigenvalues λ_i for $i = 2, \dots, N$ will in general not be equal, so minimizing $\|H_P^{\mathrm{DAPI}}\|_2^2$ is no longer equivalent to maximizing each $f_i(q_I, \gamma)$. The following proposition summarizes the cases where there is a positive $\gamma^\star$, when it is zero, and when the optimum must be evaluated on a case-by-case basis.

Proposition 10.6 *Let Assumptions 10.1–10.3 hold. The choices of γ that minimize $\|H_P^{\mathrm{DAPI}}\|_2^2$ are as follows:*
If $\sqrt{\tau k \lambda_i} \leq 1$ for all $i = 2, \dots, N$, then

$$\gamma^\star = 0,$$

if $\sqrt{\tau k \lambda_i} \leq 1$ for $2 \leq i \leq l < N$, then

$$0 \leq \gamma^\star \leq \max_i \frac{q_I}{\tau \lambda_i}(\sqrt{\tau k \lambda_i} - 1),$$

and if $\sqrt{\tau k \lambda_i} > 1$ for all $i = 2, \dots, N$, then

$$\min_i \frac{q_I}{\tau \lambda_i}(\sqrt{\tau k \lambda_i} - 1) \leq \gamma^\star \leq \max_i \frac{q_I}{\tau \lambda_i}(\sqrt{\tau k \lambda_i} - 1).$$

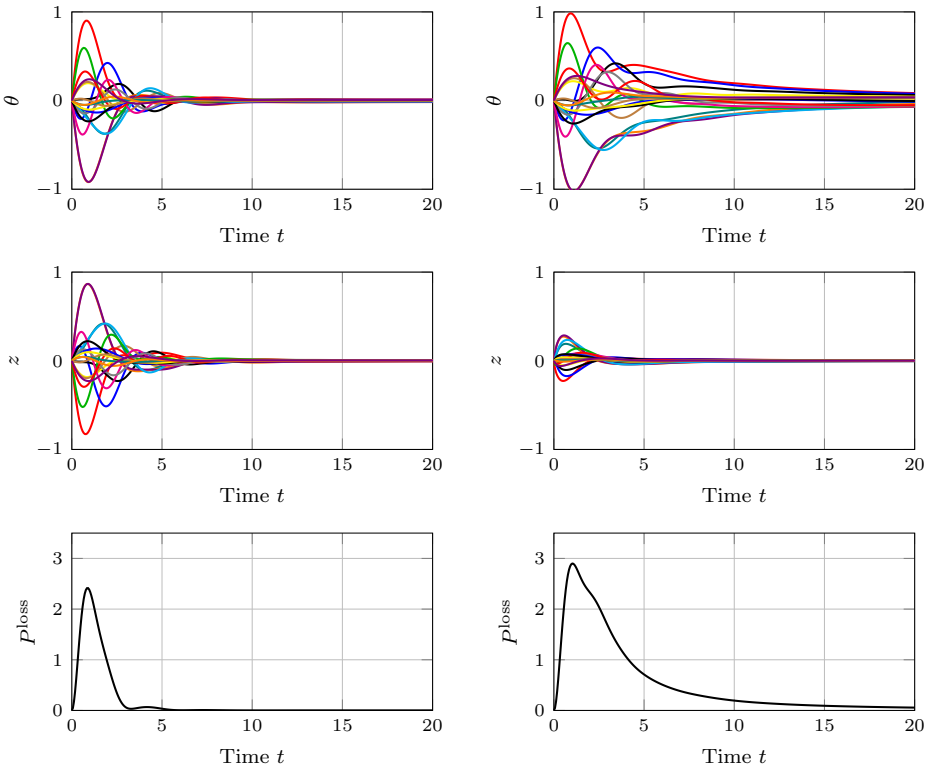


Figure 10.4: Simulations of the *noiseless* DAPI-controlled system (10.7) under initial frequency perturbations on a 20-node network with path graph topology. With a carefully selected communication gain γ, the convergence rate of the integral state matches that of the phase angles, and transient losses are minimized (they correspond to the area under the curves on the bottom panels).

Proof: The case where $\sqrt{\tau k \lambda_i} \leq 1$ for all $i = 2, \ldots, N$ is simple, since $f_i(q_I, \gamma)$ from (10.15) is then minimized by $\gamma^\star = 0$ for all $i = 2, \ldots, N$ by the proof of Proposition 10.5. Now, consider the case when $\sqrt{\tau k \lambda_i} > 1$ for all $i = 2, \ldots, N$ and assume for the sake of argument that $\gamma^\star < \min_i \frac{q_I}{\tau \lambda_i}(\sqrt{\tau k \lambda_i} - 1)$. However, in this case we would, by the proof of Proposition 10.5, have $\frac{\partial g_i(\gamma, q_I)}{\partial \gamma}\big|_{\gamma=\gamma^\star} < 0$ for all $i = 2, \ldots, N$ and $\gamma^\star$ cannot be the minimizer of $\|H_P^{\text{DAPI}}\|_2^2$, which gives a contradiction. The remaining cases are proven similarly. ■

The simulation in Figure 10.4 shows the performance improvement attainable (in noiseless DAPI) by selecting an appropriate communication gain γ. If γ is

selected too large, the distributed averaging converges too fast compared to the phase angles and reduces the absolute feedback effect. This results in increased inter-nodal power flows and greater transient losses. Recall, however that losses remain smaller than with droop or CAPI control as long as the DAPI control is noiseless.

10.3.2 The impact of measurement noise

From the previous section, we concluded that a small value for γ, or even $\gamma = 0$, minimizes the DAPI controller's losses in the absence of noise. However, when measurement noise η is added to the model, it is immediately obvious that $\gamma = 0$ is never an optimal, or even feasible choice (as predicted also by the stability analyses in [6,159]). The following corollary to Proposition 10.1 follows directly from (10.13).

Corollary 10.7 *It holds that*

$$\lim_{\gamma \to 0} ||H_\eta^{\mathrm{DAPI}}||_2^2 = \infty$$

and therefore $\lim_{\gamma \to 0} ||H^{\mathrm{DAPI}}||_2^2 = \infty$.

This implies that if γ is chosen too small, the losses in the presence of noise may become very large.

It turns out that an analytic expression for the optimal choice of γ in the presence of measurement noise is difficult to obtain even for special graph structures, and a numerical evaluation for each case is necessary. We can, however, give the following proposition.

Proposition 10.8 *Let $\gamma^{\star,\eta}$ be optimal with respect to the power losses $\|H^{\mathrm{DAPI}}\|_2^2$ in* (10.11) *in the presence of measurement noise η, and let $\gamma^\star$ be optimal in its absence (that is, let $\gamma^\star$ minimize $\|H_P^{\mathrm{DAPI}}\|_2^2$). Then,*

$$\gamma^{\star,\eta} > \gamma^\star.$$

Proof: See Appendix 10.C.

Figure 10.5 shows a numerical example that illustrates the results of this section. Overall, our results imply that the distributed averaging of integral states that takes place in the DAPI controller is increasingly important in the presence of measurement noise. Selecting a too small γ may cause the transient losses under DAPI to become very large, and in particular, they may vastly exceed those under CAPI control. An important question therefore becomes whether there is always a control design that allows DAPI to outperform CAPI, even under measurement noise. The answer is positive and given in the next section.

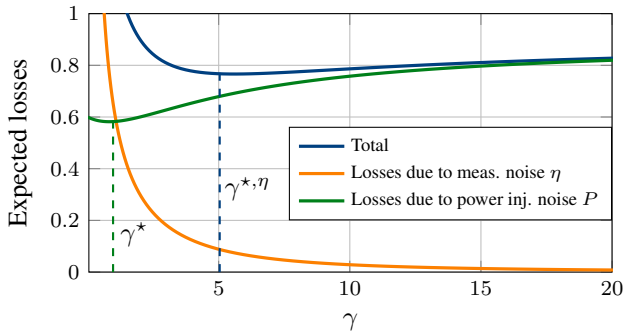


Figure 10.5: Shift of optimal $\gamma^\star$ in DAPI due to measurement noise as predicted by Proposition 10.8. The figure displays the norm $\|H^{\mathrm{DAPI}}\|_2^2$ (blue), corresponding to the total expected transient power losses, and its two components $\|H_\eta^{\mathrm{DAPI}}\|_2^2$ (orange) and $\|H_P^{\mathrm{DAPI}}\|_2^2$ (green). As $\gamma \to 0$, transient losses approach infinity in the presence of noise. A sufficiently large γ must therefore be chosen. Here, we have modeled a complete graph with $n = 10$ nodes and set $k = 5$, $q_I = \tau = 0.8$, $\alpha = \varepsilon = 1$ and $b_{ij} = 0.05$ for all $(i,j) \in \mathcal{E}^P$.

Existence of optimal control design

With the optimal choice of $\gamma^{\star,\eta}$, DAPI control outperforms CAPI control also under measurement noise. Consider the following proposition:

Proposition 10.9 *For any network $\mathcal{G}^P$ it holds that*

$$||H^{\mathrm{DAPI}}||_2^2 < ||H^{\mathrm{CAPI}}||_2^2 \,, \quad \text{if} \quad \hat{\gamma} < \gamma < \infty, \tag{10.17}$$

where $\hat{\gamma} = \frac{\varepsilon^2}{\lambda_2}$, and λ_2 is the smallest non-zero eigenvalue of $\mathcal{L}_B$.

Proof: The DAPI $\mathcal{H}_2$ norm (10.11) can be written as

$$||H^{\mathrm{DAPI}}||_2^2 = \frac{\alpha}{2k} \sum_{i=2}^{N} \left(\frac{1}{1+\varphi_i^{-1}} + \frac{\varepsilon^2}{\gamma \lambda_i} \frac{1}{1+\varphi_i} \right).$$

This expression can be upper bounded by $(N-1)$ times the largest summand:

$$||H^{\mathrm{DAPI}}||_2^2 \leq \frac{\alpha}{2k}(N-1) \cdot \sup_{\lambda_i > 0} \left\{ \frac{1}{1+\varphi_i^{-1}} + \frac{\varepsilon^2}{\gamma \lambda_i} \frac{1}{1+\varphi_i} \right\}.$$

Let $\lambda^\star$ be the maximizing eigenvalue. We are looking to choose γ so that $||H^{\mathrm{DAPI}}||_2^2 < ||H^{\mathrm{CAPI}}||_2^2 = \frac{\alpha}{2k}(N-1)$ holds. Some simplifications reveal that this holds if $\gamma > \frac{\varepsilon^2}{\lambda^\star} \geq \frac{\varepsilon^2}{\lambda_2} =: \hat{\gamma}$, and (10.17) follows. ■

Proposition 10.9 implies that for $\gamma > \hat{\gamma}$, and in particular for the optimum $\gamma^{\star,\eta}$, DAPI will perform better than CAPI. Thus, the CAPI controller's transient losses

can still be seen as an upper bound for DAPI. Nonetheless, one should be aware that selecting a too small γ in DAPI can lead to much worse performance.

It is very important to note that $\hat{\gamma}$ in Proposition 10.9 depends on λ_2, the algebraic connectivity of the network. This means that the choice of γ that allows DAPI to outperform CAPI cannot in general be made independently of the network size. In other words, the optimal control design may not be scalable. This is due to the unfavorable scaling of $\|H_\eta^{\mathrm{DAPI}}\|_2^2$ in (10.13). We address this issue in the next section.

10.4 Limitations to the scalability of DAPI control in power networks

Already in Chapter 7, we observed that the DAPI controller does not scale well to large networks when subjected to noise. Here, we will look closer into what this result implies in terms of the Price of Synchrony in power networks.

To discuss the scalability of the DAPI controller, we need to assume that all parameters and gains of the controller (10.3) are fixed. That is, they do not change as the network size grows. Further, we will consider the transient power losses from Definition 8.1 *normalized* by the total number of nodes N. These *per-node* losses should at least remain bounded in order for the controller to be regarded as scalable.

The expected power losses associated with measurement noise are given by $\|H_\eta^{\mathrm{DAPI}}\|_2^2$ in (10.13). Notice the factor $\frac{1}{\lambda_i}$, which tends to infinity for small eigenvalues. As we have seen throughout this thesis, this causes an unfavorable scaling of the losses in sparse networks. Scalings of this type of expression for lattice graphs and their fuzzes were derived in Chapter 6, and we also noted that those scalings will apply to any graph that can be embedded in such lattices. To facilitate the upcoming discussion, we repeat the relevant scalings here.

Proposition 10.10 (Transient loss scaling in lattices) *Let the graph $\mathcal{G}^P$ be a lattice or its r-fuzz in one or two dimensions ($d = 1$ or $d = 2$), and let the line susceptances b_{ij} be bounded. Then, the* per-node *losses associated with measurement noise scale asymptotically according to:*

$$\frac{1}{N}||H_\eta^{\mathrm{DAPI}}||_2^2 \sim \begin{cases} \varepsilon^2 N & \text{if } d = 1 \\ \varepsilon^2 \log N & \text{if } d = 2, \end{cases}$$

The per-node losses associated with power injection noise are, on the other hand, upper bounded as:

$$\frac{1}{N}||H_P^{\mathrm{DAPI}}||_2^2 \leq \frac{\alpha}{2k}.$$

Proof: See Appendix 10.D.

This result means that while the losses, when evaluated per node, were upper bounded for DAPI in the absence of noise, they may in theory grow unboundedly in large 1- or 2-dimensional lattice networks in the presence of noise. The situation is the same as the one depicted in Figure 7.2. That is, with a suitable tuning, the DAPI controller may perform well in a network of any given size, but losses will increase unboundedly (per generator) as the network grows.

The unbounded growth of the power losses per generator can be understood as caused by the secondary control input becoming increasingly distorted as the network grows. In practice, however, there is clearly a limit on how large the transient losses can become, which depends on the generators' power ratings. The scaling result in Proposition 10.10 should therefore be interpreted as setting a limit on the feasible network size for each controller tuning.

With CAPI control, on the other hand, the per-generator losses are bounded in network size also in the presence of noise. This means that the CAPI controller is preferable to the DAPI controller from a scalability perspective.

10.5 Discussion

This chapter's results show that the distributed DAPI controller can improve the frequency control performance in power networks, in that transient losses are reduced compared to both the static droop controller and the centralized CAPI controller. However, the DAPI controller's sensitivity to measurement noise implies that it requires very careful tuning to maintain its advantage. We now discuss some interpretations and implications of our results.

10.5.1 Distributed versus centralized control

In line with previous results, the (noiseless) DAPI controller was found to improve performance compared to the static droop controller. It is perhaps surprising that the centralized CAPI approach does not achieve the same performance improvement in terms of power losses, given that centralized strategies are often expected to be superior from a performance perspective (their downside being difficulty to implement and that they may have a single point of failure). For example, a centralized averaging PI controller for frequency control was shown in [7] to have a higher rate of convergence than a comparable distributed controller. A similar result was derived in [27] for wireless sensor networks in HVAC systems. We show here that, unless subjected to noise, the DAPI controller is superior both in terms of transient power losses and network coherence. We remark again, however, that the CAPI controller is lonly a special centralized control strategy. It is, for example, limited in that it affects all nodes equally.

Also under measurement noise, the DAPI controller can outperform CAPI provided its distributed averaging filter is carefully tuned. We found that the commu-

nication gain for the distributed averaging should then satisfy $\gamma > \hat{\gamma}$. However, the parameter $\hat{\gamma}$ is inversely proportional to the network's algebraic connectivity, which we saw in Section 5.2.3 to be a decreasing function in network size for many graphs. This means that if the network size grows, the local DAPI-controllers may need to be re-tuned to maintain good performance. For very large and sparse networks it holds $\hat{\gamma} \to \infty$, meaning that one would have to revert to CAPI control.

10.5.2 Practical implications of scaling results

The results on performance scalings considered throughout this thesis reveal that certain controllers are fundamentally limited in their scalability to large networks. In this chapter, the most severe limitations of this type apply to the DAPI controller under measurement noise. They imply that the total transient losses may even scale faster than the number of generators (that is, faster than the number of disturbance inputs) in network size.

From a practical perspective, it is worth pointing out two aspects of this result. First, there is an upper bound on the feasible power flows in a network, so a faster-than-linear growth of losses is eventually impossible. Second, and more importantly, the losses due to measurement noise are scaled by the factor ε^2. Recall that ε is the relative intensity of the measurement noise compared to noise in the net power injection. It is reasonable to assume that these differ by orders of magnitude. The practical implications of the scaling results are therefore likely limited, at least for moderately large networks.

10.5.3 Relaxing limitations through phasor measurements

A promising strategy to alleviate the limitations of DAPI control is through PMUs. These measure the phase directly and thereby eliminate the need for integral states that are distorted by noisy frequency measurements. Preliminary numerical results indicate that equipping a fixed ratio of the generators with PMUs can indeed counteract the unfavorable performance scaling due to measurement noise. A derivation of analytical results, as well as a design of suitable control strategies, are part of future work.

Appendix to Chapter 10

10.A Proof of Theorem 10.1

For the DAPI case, the calculation of the $\mathcal{H}_2$ norm is done in the same manner as in the proof of Theorem 4.1 in Appendix 4.A, and the details are therefore omitted. Here, the fact that the $\mathcal{H}_2$ norm can be partitioned into one part associated with the input P and one part associated with the input η is due to these inputs being uncorrelated. The total $\mathcal{H}_2$ norm is therefore the sum of the contributions of each of these inputs: $\|\hat{H}_n^{\mathrm{DAPI}}\|_2^2 = \mathrm{tr}(\hat{\mathcal{B}}_n^T X_n \hat{\mathcal{B}}_n) = \frac{1}{\tau^2} X_{22} + \frac{\varepsilon^2}{q_I^2} X_{33} = \|\hat{H}_{P,n}^{\mathrm{DAPI}}\| + \|\hat{H}_{\eta,n}^{\mathrm{DAPI}}\|$.

In the CAPI case, we also perform a unitary transformation using the matrix $\mathcal{U} = \mathrm{blkdiag}\{U, U, 1\}$ (U being the unitary matrix that diagonalizes $\mathcal{L}_B$). This gives the decoupled subsystems:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\hat{\theta}_1\\ \hat{\omega}_1\\ \hat{z}_1\end{bmatrix} = \begin{bmatrix}0 & 1 & 0\\ 0 & -\frac{1}{\tau} & \frac{\sqrt{N}}{\tau}\\ 0 & -\frac{1}{q_I\sqrt{N}} & 0\end{bmatrix}\begin{bmatrix}\hat{\theta}_1\\ \hat{\omega}_1\\ \hat{z}_1\end{bmatrix} + \begin{bmatrix}0 & 0\\ \frac{1}{\tau} & 0\\ 0 & \sqrt{N}\varepsilon/q_I\end{bmatrix}\hat{\mathrm{w}}_1, \quad \hat{y}_1 = 0$$

for $n = 1$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\hat{\theta}_n\\ \hat{\omega}_n\\ \hat{z}_n\end{bmatrix} = \begin{bmatrix}0 & 1 & 0\\ -\frac{k}{\tau}\lambda_n & -\frac{1}{\tau} & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\hat{\theta}_n\\ \hat{\omega}_n\\ \hat{z}_n\end{bmatrix} + \begin{bmatrix}0 & 0\\ 1/\tau & 0\\ 0 & 0\end{bmatrix}\hat{\mathrm{w}}_n, \quad \hat{y}_n = \begin{bmatrix}\sqrt{\alpha\lambda_n} & 0 & 0\end{bmatrix}\begin{bmatrix}\hat{\theta}_n\\ \hat{\omega}_n\\ \hat{z}_n\end{bmatrix},$$

for $n = 2, \ldots, N$, since the first eigenvector u_1 of U that corresponds to $\lambda_1 = 0$ is parallel to $\mathbf{1}$, and remaining eigenvectors $u_n \perp \mathbf{1}$. The subsystem norms for $\hat{H}_n^{\mathrm{CAPI}}$ can now be evaluated in the same manner as in Appendix 4.A. Clearly, it is only the power injection noise (corresponding to the first element of $\hat{\mathrm{w}}_n$) that contributes to the norm. □

10.B Alternative noise model

If the additional noise term enters the dynamic equation as in (10.8), the $\mathcal{H}_2$ norm becomes $||\tilde{\mathcal{S}}^{\mathrm{DAPI}}||_2^2 = \frac{a}{2k}\sum_{i=2}^{n}\left((1+\varepsilon^2)\frac{1}{1+\varphi_i^{-1}} + \varepsilon^2\left(2+\frac{1}{\gamma\lambda_i}\right)\frac{1}{1+\varphi_i}\right)$. Clearly, the factor $(1+\varepsilon^2)$ does not change the qualitative behavior of the first term. For the second term it is the scaling in N that is of most interest (see Section 10.4). This depends on the factor $\frac{1}{\gamma\lambda_i}$ and the qualitative result of Proposition 10.10 is not affected by adding the term 2 to this factor.

10.C Proof of Proposition 10.8

The case of $\gamma^\star = 0$ is trivial due to Corollary 10.7. Assume therefore $\gamma^\star > 0$.

Define $\bar{\chi}_i(\gamma) = \frac{1}{1+\varphi_i(\gamma)}$ and $\underline{\chi}_i(\gamma) = \frac{1}{1+\varphi_i^{-1}(\gamma)}$, and note that $\underline{\chi}_i + \bar{\chi}_i = 1$. Therefore $\bar{\chi}_i$ (that occurs in the expression for $||H_\eta^{\mathrm{DAPI}}||_2^2$) is increasing whenever $\underline{\chi}_i$ (that occurs in $||H_P^{\mathrm{DAPI}}||_2^2$) is decreasing and vice versa.

Now, since $\gamma^\star$ is a minimizer for $||H_P^{\mathrm{DAPI}}||_2^2$, the expression $\sum_{i=2}^N \bar{\chi}_i$ in $||H_\eta^{\mathrm{DAPI}}||_2^2$ will have a maximum at $\gamma^\star$. We thus know that

$$0 < \sum_{i=2}^{N} \bar{\chi}_i(\gamma) < \sum_{i=2}^{N} \bar{\chi}_i(\gamma^\star)$$

for any γ in the open interval $(0, \gamma^\star)$ and that

$$\sum_{i=2}^{N} \bar{\chi}_i(\gamma^\star) \geq \lim_{\gamma \to \infty} \sum_{i=2}^{N} \bar{\chi}_i(\gamma) = 0.$$

That means that the set of values that $\sum_{i=2}^N \bar{\chi}_i(\gamma)$ can assume in the interval $(0, \gamma^\star)$ are also contained in the set of values that it can assume in $(\gamma^\star, \infty)$. The same argument can be applied to $\sum_{i=2}^N \underline{\chi}_i(\gamma)$. Therefore, for any $\gamma^- \in (0, \gamma^\star)$, there exists a $\gamma^+ > \gamma^\star$ such that $\sum_{i=2}^N \bar{\chi}_i(\gamma^-) = \sum_{i=2}^N \bar{\chi}_i(\gamma^+)$ while also $\sum_{i=2}^N \underline{\chi}_i(\gamma^-) = \sum_{i=2}^N \underline{\chi}_i(\gamma^+)$, since $\underline{\chi}_i + \bar{\chi}_i = 1$.

From $\frac{\varepsilon^2}{\gamma^- \lambda_i} > \frac{\varepsilon^2}{\gamma^+ \lambda_i}$ follows that

$$\sum_{i=2}^{N} \frac{\varepsilon^2}{\gamma^- \lambda_i} \bar{\chi}_i(\gamma^-) > \sum_{i=2}^{N} \frac{\varepsilon^2}{\gamma^+ \lambda_i} \bar{\chi}_i(\gamma^+)$$

while $\sum_{i=2}^N \underline{\chi}_i(\gamma^-) = \sum_{i=2}^N \underline{\chi}_i(\gamma^+)$ still holds. This means that

$$||H^{\mathrm{DAPI}}||_2^2\big|_{\gamma=\gamma^-} > ||H^{\mathrm{DAPI}}||_2^2\big|_{\gamma=\gamma^+}.$$

Therefore, if we assume γ^- locally minimizes $||H^{\mathrm{DAPI}}||_2^2$ in the interval $(0, \gamma^\star)$, then γ^+ would give an even smaller value for $||H^{\mathrm{DAPI}}||_2^2$. We conclude that $\gamma^{\star,\eta} \geq \gamma^\star$.

It remains to show that $\gamma^\star$ is not optimal in the presence of η. Since

$$\frac{d}{d\gamma} ||H^{\mathrm{DAPI}}||_2^2\big|_{\gamma=\gamma^\star} = -\frac{a\varepsilon^2}{2k(\gamma^\star)^2} \sum_{i=2}^{N} \frac{1}{\lambda_i} \frac{1}{1+\varphi_i(\gamma^\star)} < 0,$$

it follows that $\gamma^\star$ is not optimal. Hence,

$$\gamma^{\star,\eta} > \gamma^\star.$$

■

10.D Proof of Proposition 10.10

First we need to show that the factors $\frac{1}{1+\varphi_i}$ from (10.13) are uniformly bounded. Recall that $\frac{d\varphi_i}{d\lambda_i} > 0$ for $\lambda_i \geq 0$. Thus, $\lambda_{\min} = 0$ and $\lambda_{\max} = \max_i \lambda_i$ give the

lower and upper bounds for φ_i, respectively. In a 1-dimensional lattice the largest possible eigenvalue is $4b_{\max}$ and in a 2-dimensional lattice it is $16b_{\max}^2$, where $b_{\max} = \max_{(i,j)\in\mathcal{E}^P} b_{ij}$, since the eigenvalues of an unweighted path graph are $\lambda_i = 2(1 - \cos\frac{\pi i}{N})$. Therefore, $0 \leq \varphi_i \leq \varphi(\lambda_{\max}, \gamma, k, q_I, \tau)$ and all $\frac{1}{1+\varphi_i}$ are uniformly bounded with respect to N.

Now, we can bound $\frac{1}{N}||H_\eta^{\mathrm{DAPI}}||_2^2$ as:

$$\underline{c}\frac{\varepsilon^2}{N}\sum_{i=2}^{N}\frac{1}{\lambda_i^1} := \frac{\alpha\varepsilon^2}{2kn\gamma b_{\max}(1+\varphi(\lambda_{\max}))}\sum_{i=2}^{N}\frac{1}{\lambda_i^1} \leq \frac{1}{N}||H_\eta^{\mathrm{DAPI}}||_2^2$$
$$\leq \frac{\alpha\varepsilon^2}{2kn\gamma b_{\min}(1+\varphi(\lambda_{\min}))}\sum_{i=2}^{N}\frac{1}{\lambda_i^1} =: \bar{c}\frac{\varepsilon^2}{N}\sum_{i=2}^{N}\frac{1}{\lambda_i^1},$$

where λ_i^1 are the Laplacian eigenvalues of the corresponding unweighted graph (see Section 3.2).

The scaling of the sum $\frac{1}{N}\sum_{i=2}^{N}\frac{1}{\lambda_i^1}$ in a (periodic) d-dimensional r-fuzz lattice (with $r = q/2$) was derived in Chapter 6 in terms of Fourier symbols. We can also make use of the Kirchoff index K_f of a lattice graph. We then use the relation (3.21) to write

$$\frac{\underline{c}\varepsilon^2}{N^2}K_f \leq \frac{1}{N}||H_\eta^{\mathrm{DAPI}}||_2^2 \leq \frac{\bar{c}\varepsilon^2}{N^2}K_f.$$

It is proven in [18] that the Kirchhoff index for infinite lattices, including r-fuzzes, scales like $K_f \sim N^3$ if $d = 1$ and $K_f \sim N^2 \log N$ if $d = 2$. The result follows. ■

Chapter 11

Conclusions

The constraints imposed by a distributed controller structure lead to limitations on the achievable global performance in large-scale networked systems. In this thesis, we have described and quantified such limitations in terms of scalings of nodal variance, that is, $\mathcal{H}_2$-norm metrics, on one hand, and in terms of network stability on the other. In this final chapter of the thesis, we summarize the conclusions from our findings. We will discuss how the various problems we have considered answer the questions we set up in the introduction and which problems remain open.

11.1 Performance limitations and their dependence on system properties

The problems we have studied in this thesis have been set up in different ways (for example, over lattice networks or general networks), considered different applications (such as frequency synchronization or vehicular formation control problems) and different performance objectives (network coherence or transient resistive losses). In many cases, we have revealed unfavorable scalings of performance in network size. In principle, we have found that they can be attributed to constraining the feedback to localized, relative state measurements. Yet the performance limitations can be more or less severe depending on other properties of the system and of the controller. We next discuss some of these properties and the roles they play in different settings.

11.1.1 Network topology

Topological properties of the underlying network graphs are important for determining behaviors of networked systems. In our case, this is also where the locality constraints are relevant; they limit the size of each node's neighborhood and thereby the connectivity of the network. In the problems we have considered, the network

connectivity has played different roles. Our main conclusions can be summarized as follows:

i. *A high connectivity improves network coherence.* The unfavorable scaling of the global error variance is, simply put, due to the behavior of sums over inverted graph Laplacian eigenvalues like $\sum_{i=2}^{N} \frac{1}{\lambda_i}$. Under locality constraints, some of these eigenvalues typically become very small as the network size grows, causing the sum to grow fast. Increasing the network connectivity increases the size of the eigenvalues, thus improving coherence. Full coherence (a bounded global error variance per site) can, for example, be achieved if they all remain bounded away from zero.

   In lattice networks, full coherence is achieved in high spatial dimensions ($d = 3$ in first-order and $d = 5$ in second-order consensus). Note that this holds even though the algebraic connectivity λ_2 approaches zero in those networks. This means that the algebraic connectivity is not necessarily a good indicator of the scaling of network coherence, while the spatial dimension of a lattice is. Therefore, arranging nodes in a grid is fundamentally better with respect to coherence than arranging them in a string. The correct generalization of the spatial dimension to general graphs is, however, an open research question. Still, a bound on the coherence scaling for a general graph can be obtained if it can be embedded in a d-dimensional lattice.

   Local error variance may also have a scaling that depends on terms like $\sum_{i=2}^{N} \frac{1}{\lambda_i}$, and where the above conclusions therefore carry over. This is the case with noisy distributed integral control, for example in power networks.

ii. *A low connectivity renders high-order consensus inadmissible.* The connectivity of a graph is important for the admissibility of high-order ($n \geq 3$) consensus algorithms that are limited to relative feedback. If the algebraic connectivity approaches zero as the network size grows, then the system becomes unstable beyond some critical network size. This property applies to many network graphs where locality constraints apply. This means, for example, that even though a 5-dimensional lattice allows full coherence in second-order consensus, third-order consensus will be inadmissible in such a lattice. The same conclusion holds for certain dynamic feedback protocols that are applied in a second-order integrator network.

iii. *The Price of Synchrony does not depend on network connectivity.* Under standard frequency droop control, the Price of Synchrony metric, which can be cast as a local error variance, only depends weakly on the network topology. Under the assumptions of uniform generator parameters and resistance-to-reactance ratios in the network, it is entirely independent of topology.

iv. *A low connectivity improves the Price of Synchrony with variable voltages.* If voltage droop control is considered, the Price of Synchrony metric is shown to be an *increasing* function of network connectivity. This means that the lowest

transient power losses are achieved for a sparse network topology. The result can be attributed to the fact that voltage droop control involves absolute feedback from the full local state.

The above results mean that while the two performance objectives we have considered; coherence and the Price of Synchrony, often depend differently on the network topology, they do not in general conflict. The only case where they have opposite dependencies on connectivity is where absolute feedback from the full local state is available, as in voltage droop control. In this case, the control already has good performance and scalability properties, so the conflicting dependencies need not be a great concern in the control design.

11.1.2 Order of local dynamics

We have considered networks of agents with integrator dynamics of different orders n. If all state feedback is relative and static, the order of the dynamics has a significant impact on the system's stability and performance in large-scale networks:

$n = 1$: The case $n = 1$ corresponds to the standard consensus problem. Here, the per-site variance of the global error (coherence) has the worst-case asymptotic scaling

$$V_N \sim \frac{1}{\beta} N$$

for a 1-dimensional lattice (see Theorem 6.1). Recall that β is an algorithm parameter that reflects the control effort.

$n = 2$: The case $n = 2$ corresponds to the vehicular formation control problem. Under relative feedback,

$$V_N \sim \frac{1}{\beta^2} N^3$$

for a 1-dimensional lattice (except if the feedback is *asymmetric* and the lattice has periodic boundary conditions, in which case the feedback is inadmissible).

$n \geq 3$: If $n \geq 3$, then localized relative feedback is inadmissible and

$$V_N = \infty$$

if $N > \bar{N}$ for some finite $\bar{N}$. This holds for graphs whose algebraic connectivity is decreasing in network size, including lattice graphs of any finite dimension.

We have also considered dynamic feedback controllers, which introduce additional states, thus increasing the model order n. We saw that the dynamic feedback must be designed carefully for the control to remain admissible, but that it need not deteriorate the system's performance scaling. On the other hand, it cannot improve the scaling as long as the feedback remains relative.

If (partial) absolute feedback is available, the order of the system dynamics plays a different role. This is evident in the consideration of frequency synchronization dynamics. Here $n = 2$, but only the phase feedback is relative while frequency feedback is absolute. With the static droop controller, this means that the scaling of phase coherence is as for the standard consensus problem where $n = 1$. Frequency synchronization is also often described in literature by a first-order model that approximates the swing dynamics. Our analysis also reveals that the Price of Synchrony metric, that is, the resistive power loss incurred in the synchronization transient, is unchanged by such an approximation.

11.1.3 Absolute state measurements and distributed dynamic feedback

The availability of absolute state measurements is important to ensure scalability of a distributed control design, in terms of both performance and admissibility. For second-order integrator networks, we showed that if absolute measurements of the partial local state are available, then a dynamic feedback controller can fundamentally improve performance in terms of the scaling of global error variance compared to static feedback. It can also reduce the local error variance and in particular the Price of Synchrony metric in absolute terms, thereby improving transient performance in power networks. This is done by applying integral or derivative action to the available state measurement to emulate absolute feedback from the other state.

In the case of derivative control, this is straightforward, and the performance improvement is attained even with filtered derivative action. Distributed integral control, however, is sensitive to measurement noise and bias. One of the issues is if different controllers' memory states diverge due to slight measurement errors. This issue appears in completely decentralized integral control and leads to instability, but is solved through an inter-nodal alignment of the memory states through a distributed averaging filter. In some cases, an optimal design of this filter can even improve performance *beyond* that of a system with absolute measurements of both states. A second issue is, however, that noisy absolute measurements deteriorate performance.

Measurement noise deteriorates the performance scaling of distributed integral control in terms of both global and local error, and increases the need for inter-nodal alignment of the integral states. As long as the relative intensity of the measurement noise compared to the process disturbance remains small, the implications for moderately sized networks are limited. However, the results imply that distributed integral control is fundamentally limited in its scalability. The need for inter-nodal alignment grows with network size, and our results indicate that very large and sparse networks require all-to-all communication or centralized integral control to be scalable.

11.2 Implications for a distributed control design in large-scale networks

Throughout this thesis, we have revealed fundamental limitations to the performance of distributed control in large-scale networks. These have been in terms of asymptotic scalings of global error variance in first- and second-order consensus networks, which determine the achievable level of coherence of the network. In the context of electric power networks, we have shown limitations in terms of resistive losses that arise in the synchronization transient, which scale with the total network size. We have also discussed secondary frequency control and revealed benefits of using distributed integral control for this purpose, as well as its limitations in terms of scalability. Finally, we have also revealed fundamental limitations to localized consensus algorithms in high-order integrator networks that prevent them from stabilizing large-scale networks.

The previous section summarized how various properties of the networked system affect these limitations. Many of these properties, as well as the limitations, are highly relevant in real-world networked systems like those mentioned in the introduction of the thesis. Yet, we have treated them here using idealized models. In this final section, we discuss how our results can be understood in the context of practical distributed control design.

11.2.1 Expectations on a distributed control design

The fundamental limitations we describe show what performance can be expected from a distributed feedback control design. This allows a control designer to distinguish whether a failure to meet a given specification is due to bad design or to infeasibility. Fundamental limitations can also indicate a need to re-design the network or the controller structure. Though the models we have used in this thesis are simplified, and performance objectives in a practical system may be more complex, our results reveal what properties are important for the performance and scalability properties of a control design.

For example, our results give insights to questions like "What performance is expected if the network size is doubled?," "What is the effect of adding a certain number of links?," or "How much noise can be tolerated in the local measurements?" Which particular properties or performance objectives that are most relevant, or whether the performance limitations are tolerable, depends on the system in question. For example, our results on distributed integral control show, on one hand, that it can provide a large performance improvement compared to standard static feedback, but on the other, that its scalability is limited when subject to noise. In a practical situation, for example when designing controllers for power-electronic inverters in a microgrid, one would need to determine the suitability of distributed integral control based on estimations of the noise intensity and the size of the network. Similarly, knowing the benefit of absolute measurements, a consideration of whether to prioritize a sensing and communication infrastructure that give vehicles

access to their absolute positions can weigh in the importance of scalability for the particular system.

We have shown that the limitations to distributed control designs and their dependencies on system properties are not always intuitive. Examples of partly counterintuitive results include the inadmissibility of high-order consensus, given that first- and second-order consensus are generally admissible (see Example 5.2), and the topology (in)dependence of the Price of Synchrony metric in droop-controlled power networks (see Example 8.2). Knowledge of such counterintuitive behaviors is particularly important for imposing reasonable expectations on a networked system.

Model simplifications, including linearizations, topological restrictions, and, in the case of power networks, Kron reductions, have been necessary to derive the analytical results that, for example, allow us to draw the conclusions in Section 11.1. This is not an unusual approach in control theory. Indeed, the most well-known and applied results (including those on fundamental limitations) are derived for linear time-invariant systems, though no such exact systems appear in the physical world. Our results predict behaviors and limitations of networked control systems. The predictions are evidently more accurate, the more similar the physical system is to the idealized one. For highly nonlinear or very heterogeneous systems, our results, like many others, have limited applicability. For such cases, further analysis is required, or insights can be obtained through simulations or experiments.

11.2.2 Tuning and re-tuning of local controllers

Throughout this thesis, an underlying assumption has been that the local control design is fixed and independent of a scaling of the network size. In practice, the motivation for this assumption is to enable plug-and-play control, that is, the possibility to add subsystems to the network without re-tuning remaining subsystems' controllers. In some of the applications we have considered, such as vehicle platooning or power networks with distributed generation, subsystems tend to have different owners and therefore independent control designers. In such cases, plug-and-play control is particularly relevant.

We have also shown that allowing for an appropriate re-tuning of local controllers as the network grows can relax performance limitations. The downside of such a re-tuning is the requirement for global model knowledge, which rules out plug-and-play control. An interesting research question is therefore whether it is possible to design a distributed re-tuning protocol, possibly using ideas from system identification, that relies on limited, localized model knowledge, yet still relaxes limitations.

11.2.3 The bigger picture

Though our results provide important insights to limitations of distributed feedback control in networks, they do not give a complete picture. Our results have, for example, shown that certain controllers perform well in terms of the network

coherence. However, they may still be limited in terms of, say, $\mathcal{H}_\infty$ performance, or lack robustness to changes in the model. The opposite case is also possible. To a certain extent, this is acceptable and a control design should consider the most relevant performance objectives for the system at hand. As a background to this thesis, we therefore reviewed a number of performance characterizations for networked systems that have been proposed in literature. If the objective, however, is to gain a full understanding for the type of local control and interaction rules that lead to globally desirable behaviors in networked systems, a more holistic approach to the performance and robustness questions may be required.

One direction for such an analysis would be to take inspiration from traditional control theory and study all relevant transfer functions of a system (the Gang of Four, or Six [11]) jointly, under the architectural constraints imposed by a network structure. For example, through gap metrics. This may reveal further fundamental limitations, but also give insights to inter-dependencies and trade-offs between performance objectives. In this case, our analyses, which have considered transfer functions one by one, provide a good starting point.

Bibliography

[1] A. Alam, B. Besselink, V. Turri, J. Mårtensson, and K. H. Johansson. Heavy-duty vehicle platooning for sustainable freight transportation: A cooperative method to enhance safety and efficiency. *IEEE Control Syst. Mag.*, 35(6): 34–56, Dec 2015.

[2] S. Alexander and R. Orbach. Density of states on fractals: "fractons". *Journal de Physique Lettres*, 43(17):625–631, 1982.

[3] B. Anderson and J. B. Moore. Time-varying feedback laws for decentralized control. *IEEE Trans. Autom. Control*, 26(5):1133–1139, 1981.

[4] M. Andreasson. *Control of Multi-Agent Systems with Applications to Distributed Frequency Control Power Systems.* Licentiate thesis, KTH Royal Institute of Technology, 2013. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-118638.

[5] M. Andreasson, D. V. Dimarogonas, H. Sandberg, and K. H. Johansson. Distributed control of networked dynamical systems: Static feedback, integral action and consensus. *IEEE Trans. Autom. Control*, 59(7):1750–1764, July 2014.

[6] M. Andreasson, D. V. Dimarogonas, H. Sandberg, and K. H. Johansson. Distributed PI-control with applications to power systems frequency control. In *American Control Conf. (ACC)*, pages 3183–3188, June 2014.

[7] M. Andreasson, D.V. Dimarogonas, K. H. Johansson, and H. Sandberg. Distributed vs. centralized power systems frequency control. In *European Control Conf. (ECC)*, pages 3524–3529, July 2013.

[8] M. Andreasson, E. Tegling, H. Sandberg, and K. H. Johansson. Coherence in synchronizing power networks with distributed integral control. In *IEEE 56th Annual Conf. on Decision and Control (CDC)*, pages 6327–6333, Dec 2017.

[9] M. Andreasson, E. Tegling, H. Sandberg, and K. H. Johansson. Performance and scalability of voltage controllers in multi-terminal HVDC networks. In *American Control Conf. (ACC)*, May 2017.

[10] K. J. Åström. *Stochastic control theory*. Academic Press, Inc., New York, 1970.

[11] K. J. Åström and R. M. Murray. *Feedback Systems*. Princeton University Press, New Jersey, 2008.

[12] B. Bamieh and D. F. Gayme. The price of synchrony: Resistive losses due to phase synchronization in power networks. In *American Control Conf. (ACC)*, pages 5815–5820, June 2013.

[13] B. Bamieh, M. Jovanović, P. Mitra, and S. Patterson. Effect of topological dimension on rigidity of vehicle formations: Fundamental limitations of local feedback. In *IEEE 47th Annual Conf. on Decision and Control (CDC)*, pages 369–374, Dec 2008.

[14] B. Bamieh, M. R. Jovanović, P. Mitra, and S. Patterson. Coherence in large-scale networks: Dimension-dependent limitations of local feedback. *IEEE Trans. Autom. Control*, 57(9):2235–2249, Sept 2012.

[15] B. Bamieh, F. Paganini, and M. A. Dahleh. Distributed control of spatially invariant systems. *IEEE Trans. Autom. Control*, 47(7):1091–1107, Jul 2002.

[16] B. Bamieh and P. G. Voulgaris. A convex characterization of distributed control problems in spatially invariant systems with communication constraints. *Syst. Control Lett.*, 54(6):575–583, 2005.

[17] P. Barooah and J. P. Hespanha. Error amplification and disturbance propagation in vehicle strings with decentralized linear control. In *IEEE 44th Annual Conf. on Decision and Control (CDC)*, pages 4964–4969, Dec 2005.

[18] P. Barooah and J. P. Hespanha. Estimation on graphs from relative measurements. *IEEE Control Syst. Mag.*, 27(4):57–74, Aug 2007.

[19] P. Barooah and J. P. Hespanha. Estimation from relative measurements: Electrical analogy and large graphs. *IEEE Trans. Signal Process.*, 56(6): 2181–2193, June 2008.

[20] P. Barooah, P. G. Mehta, and J. P. Hespanha. Mistuning-based control design to improve closed-loop stability margin of vehicular platoons. *IEEE Trans. Autom. Control*, 54(9):2100–2113, Sept 2009.

[21] B. Besselink and S. Knorn. Scalable input-to-state stability for performance analysis of large-scale networks. *IEEE Control Systems Letters*, 2(3):507–512, July 2018.

[22] B. Besselink, V. Turri, S. H. van de Hoef, K. Liang, A. Alam, J. Mårtensson, and K. H. Johansson. Cyber-physical control of road freight transport. *Proc. IEEE*, 104(5):1128–1141, May 2016.

[23] K. De Brabandere, B. Bolsens, J. Van den Keybus, A. Woyte, J. Driesen, and R. Belmans. A voltage and frequency droop control method for parallel inverters. In *IEEE 35th Annual Power Electronics Specialists Conf.*, volume 4, pages 2501–2507 Vol.4, 2004.

[24] P. Bresesti, W. L. Kling, R. L. Hendriks, and R. Vailati. HVDC connection of offshore wind farms to the transmission system. *IEEE Transactions on Energy Conversion*, 22(1):37–43, 2007.

[25] A. E. Brouwer and W. H. Haemers. *Spectra of Graphs*. Springer, New York, NY, 2012.

[26] C. E. Cantos, J. J. P. Veerman, and D. K. Hammond. Signal velocity in oscillator arrays. *Eur. Phys. J., Special Topics*, 225(6):1115–1126, Sept 2016.

[27] X. Cao, J. Chen, Y. Xiao, and Y. Sun. Building-environment control with wireless sensor and actuator networks: Centralized versus distributed. *IEEE Transactions on Industrial Electronics*, 57(11):3596–3605, Nov 2010.

[28] R. Carli, F. Garin, and S. Zampieri. Quadratic indices for the analysis of consensus algorithms. In *2009 Information Theory and Applications Workshop*, pages 96–104, Feb 2009.

[29] H.-D. Chiang. *Direct Methods for Stability Analysis of Electric Power Systems: Theoretical Foundation, BCU Methodologies, and Applications*. Wiley, 2011.

[30] H.-D. Chiang, F. F. Wu, and P. P. Varaiya. Foundations of the potential energy boundary surface method for power system transient stability analysis. *IEEE Trans. on Circuits and Systems*, 35(6):712 –728, June 1988.

[31] F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1994.

[32] T. Coletta, B. Bamieh, and P. Jacquod. Transient performance of electric power networks under colored noise. *ArXiv preprint, arXiv:1807.09048*, 2018. Available: http://arxiv.org/abs/1807.09048.

[33] T. Coletta and P. Jacquod. Performance measures in electric power networks under line contingencies. *ArXiv preprint, arXiv:1711.10348*, 2017. Available: http://arxiv.org/abs/1711.10348.

[34] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *Journal of parallel and distributed computing*, 7(2):279 – 301, 1989.

[35] R. Diestel. *Graph theory*. Springer, Heidelberg, 2010.

[36] F. Dörfler and F. Bullo. Synchronization and transient stability in power networks and non-uniform Kuramoto oscillators. In *American Control Conf. (ACC)*, pages 930–937, Baltimore, MD, 2010.

[37] F. Dörfler and F. Bullo. Topological equivalence of a structure-preserving power network model and a non-uniform Kuramoto model of coupled oscillators. In *IEEE 50th Annual Conf. on Decision and Control (CDC)*, pages 7099–7104, Orlando, FL, 2011.

[38] F. Dörfler and F. Bullo. Synchronization and transient stability in power networks and non-uniform Kuramoto oscillators. *SIAM Journal on Control and Optimization*, 50(3):1616–1642, 2012.

[39] F. Dörfler and F. Bullo. Kron reduction of graphs with applications to electrical networks. *IEEE Trans. on Circuits and Systems I*, 60(1):150–163, Jan 2013.

[40] F. Dörfler, M. Chertkov, and F. Bullo. Synchronization in complex oscillator networks and smart grids. *Proc. of the National Academy of Sciences*, 110 (6):2005–2010, 2013.

[41] F. Dörfler and S. Grammatico. Amidst centralized and distributed frequency control in power systems. In *American Control Conf. (ACC)*, pages 5909–5914, July 2016.

[42] F. Dörfler, M. R. Jovanović, M. Chertkov, and F. Bullo. Sparsity-promoting optimal wide-area control of power networks. *IEEE Trans. Power Syst.*, 29 (5):2281–2291, Sept 2014.

[43] F. Dörfler, J. Simpson-Porco, and F. Bullo. Breaking the hierarchy: Distributed control & economic optimality in microgrids. *IEEE Trans. Control Netw. Syst.*, 3(3):241–253, 2016.

[44] F. Dörfler, J. W. Simpson-Porco, and F. Bullo. Electrical networks and algebraic graph theory: Models, properties, and applications. *Proc. IEEE*, 106 (5):977–1005, May 2018.

[45] P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks*. The Mathematical Association of America, 1984.

[46] M. Egerstedt, S. Martini, M. Cao, K. Camlibel, and A. Bicchi. Interacting with networks: How does structure relate to controllability in single-leader, consensus networks? *IEEE Control Syst. Mag.*, 32(4):66–73, Aug 2012.

[47] W. Ellens, F. M. Spieksma, P. Van Mieghem, A. Jamakovic, and R. E. Kooij. Effective graph resistance. *Linear Algebra Appl.*, 435(10):2491–2506, 2011.

[48] M. Fardad, F. Lin, and M. R. Jovanović. Design of optimal sparse interconnection graphs for synchronization of oscillator networks. *IEEE Trans. Autom. Control*, 59(9):2457–2462, Sept 2014.

[49] H. Farhangi. The path of the smart grid. *IEEE Power Energy Mag.*, 8(1): 18–28, Jan 2010.

[50] I. J. Farkas, I. Derényi, A. L. Barabási, and T. Vicsek. Spectra of "real-world" graphs: beyond the semicircle law. *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.*, 64 2 Pt 2:026704, 2001.

[51] J. A. Fax and R. M. Murray. Information flow and cooperative control of vehicle formations. *IEEE Trans. Autom. Control*, 49(9):1465–1476, Sept 2004.

[52] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[53] K. Fitch and N. E. Leonard. Information centrality and optimal leader selection in noisy networks. In *IEEE 52nd Annual Conf. on Decision and Control (CDC)*, pages 7510–7515, Dec 2013.

[54] R. A. Freeman, P. Yang, and K. M. Lynch. Stability and convergence properties of dynamic average consensus estimators. In *IEEE 45th Annual Conf. on Decision and Control (CDC)*, pages 338–343, Dec 2006.

[55] L. Freris and D. Infield. *Renewable energy in power systems.* John Wiley & Sons, Chichester, U.K., 2008.

[56] B. Gentile, J. W. Simpson-Porco, F. Dörfler, S. Zampieri, and F. Bullo. On reactive power flow and voltage stability in microgrids. In *American Control Conf. (ACC)*, pages 759–764, June 2014.

[57] A. Ghosh, S. Boyd, and A. Saberi. Minimizing Effective Resistance of a Graph. *SIAM Rev.*, 50(1):37–66, Feb 2008.

[58] J. D. Glover, M. S. Sarma, and T. Overbye. *Power System Analysis and Design.* Cengage, Stamford, 5th edition, 2011.

[59] C. Godsil and G. Royle. *Algebraic Graph Theory*, volume 207 of *Graduate Texts in Mathematics.* Springer, 2001.

[60] R. Grone, R. Merris, and V. Sunder. The Laplacian spectrum of a graph. *SIAM J. Matrix Anal. Appl.*, 11(2):218–238, 1990.

[61] T. W. Grunberg and D. F. Gayme. Minimizing interactions in mixed oscillator networks. In *IEEE 53rd Annual Conf. on Decision and Control (CDC)*, pages 3209–3215, Los Angeles, CA, Dec 2014.

[62] T. W. Grunberg and D. F. Gayme. Performance measures for linear oscillator networks over arbitrary graphs. *IEEE Trans. Control Netw. Syst.*, 5(1):456–468, Mar 2018.

[63] K. Guo. *Simple eigenvalues of graphs and digraphs*. PhD thesis, Simon Fraser University, 2015. Available: http://summit.sfu.ca/item/14931.

[64] I. Gutman and B. Mohar. The quasi-Wiener and the Kirchhoff indices coincide. *J. of Chem. Inf. and Comp. Sciences*, 36(5):982–985, 1996.

[65] I. Gutman and W. Xiao. Generalized inverse of the Laplacian matrix and some applications. *Bull. Cl. Sci. Math. Nat. Sci. Math.*, 129(29):15–23, 2004.

[66] T. M. Haileselassie and K. Uhlen. Impact of DC line voltage drops on power flow of MTDC using droop control. *IEEE Trans. Power Syst.*, 27(3):1441–1449, Aug 2012.

[67] H. Hao and P. Barooah. On achieving size-independent stability margin of vehicular lattice formations with distributed control. *IEEE Trans. Autom. Control*, 57(10):2688–2694, Oct 2012.

[68] H. Hao and P. Barooah. Stability and robustness of large platoons of vehicles with double-integrator models and nearest neighbor interaction. *Int. J. Robust Nonlinear Control*, 23(18):2097–2122, 2013.

[69] S. Hassan-Moghaddam and M. R. Jovanović. Topology design for stochastically forced consensus networks. *IEEE Trans. Control Network Syst.*, 5(3): 1075–1086, Sept 2018.

[70] J. K. Hedrick, M. Tomizuka, and P. Varaiya. Control issues in automated highway systems. *IEEE Control Syst. Mag.*, 14(6):21–32, Dec 1994.

[71] I. Herman, D. Martinec, Z. Hurák, and M. Šebek. Nonzero bound on Fiedler eigenvalue causes exponential growth of H-infinity norm of vehicular platoon. *IEEE Trans. Autom. Control*, 60(8):2248–2253, 2015.

[72] I. Herman, D. Martinec, Z. Hurák, and M. Sebek. Scaling in bidirectional platoons with dynamic controllers and proportional asymmetry. *IEEE Trans. Autom. Control*, 62(4):2034–2040, 2017.

[73] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Presss, New York, 1985.

[74] R. Hororitz and P. Varaiya. Control design of an automated highway system. *Proc. of the IEEE*, 88(7):913–925, 2000.

[75] IEEE Smart Grid / Zpryme Research & Consulting. Power systems of the future: The case for energy storage, distributed generation, and microgrids. Technical report, IEEE, Nov 2012.

[76] M. D. Ilić. From hierarchical to open access electric power systems. *Proc. IEEE*, 95(5):1060–1084, May 2007.

[77] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Autom. Control*, 48(6):988–1001, June 2003.

[78] A. Jadbabaie, N. Motee, and M. Barahona. On the stability of the Kuramoto model of coupled nonlinear oscillators. In *American Control Conf. (ACC)*, pages 4296–4301, June 2004.

[79] A. Jadbabaie and A. Olshevsky. Combinatorial bounds and scaling laws for noise amplification in networks. In *European Control Conf. (ECC)*, pages 596–601, July 2013.

[80] E. Jensen and B. Bamieh. Optimal spatially-invariant controllers with locality constraints: A system level approach. In *American Control Conf. (ACC)*, pages 2053–2058, June 2018.

[81] F. Jiang, L. Wang, and Y. Jia. Consensus in leaderless networks of high-order-integrator agents. In *American Control Conf. (ACC)*, pages 4458–4463, June 2009.

[82] Y. Jiang, R. Pates, and E. Mallada. Performance tradeoffs of dynamically controlled grid-connected inverters in low inertia power systems. In *IEEE 56th Annual Conf. on Decision and Control (CDC)*, pages 5098–5105, Dec 2017.

[83] R. E. Kalman. Contributions to the theory of optimal control. *Bol. Soc. Mat. Mexicana*, 5:102–119, 1960.

[84] J. Kelner. Spectral partitioning, eigenvalue bounds, and circle packings for graphs of bounded genus. *SIAM J. Comput.*, 35(4):882–902, 2006.

[85] D. J. Klein and M. Randić. Resistance distance. *J. Math. Chem.*, 12(1): 81–95, Dec 1993.

[86] P. Kundur. *Power System Stability and Control*. The EPRI Power System Engineering. McGraw-Hill Companies, Inc., New York, 1994. ISBN 9780070359581.

[87] R. H. Lasseter. Microgrids. In *IEEE Power Eng. Society Winter Meeting*, volume 1, pages 305–308 vol.1, 2002.

[88] L. Lessard and S. Lall. Convexity of decentralized controller synthesis. *IEEE Trans. Autom. Control*, 61(10):3122–3127, Oct 2016.

[89] W. S. Levine and Athans M. On the optimal error regulation of a string of moving vehicles. *IEEE Trans. Autom. Control*, AC-11(3):355–361, Jul. 1966.

[90] N. Li, C. Zhao, and L. Chen. Connecting automatic generation control and economic dispatch from an optimization view. *IEEE Trans. Control Network Syst.*, 3(3):254–264, Sept 2016.

[91] F. Lin, M. Fardad, and M. R. Jovanović. Optimal control of vehicular formations with nearest neighbor interactions. *IEEE Trans. Autom. Control*, 57 (9):2203–2218, Sept 2012.

[92] F. Lin, M. Fardad, and M. R. Jovanović. Design of optimal sparse feedback gains via the alternating direction method of multipliers. *IEEE Trans. Autom. Control*, 58(9):2426–2431, Sept 2013.

[93] F. Lin, M. Fardad, and M. R. Jovanović. Algorithms for leader selection in stochastically forced consensus networks. *IEEE Trans. Autom. Control*, 59 (7):1789–1802, July 2014.

[94] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási. Controllability of complex networks. *Nature*, 473:167–173, May 2011.

[95] D. A. Burbano Lombana and M. di Bernardo. Distributed PID control for consensus of homogeneous and heterogeneous networks. *IEEE Trans. Control Netw. Syst.*, 2(2):154–163, June 2015.

[96] L.-Y. Lu and C.-C. Chu. Consensus-based droop control synthesis for multiple power converters in lossy micro-grids. In *IEEE PES Asia-Pacific Power and Energy Engineering Conf.*, Dec 2013.

[97] L. Y. Lu and C. C. Chu. Consensus-based droop control synthesis for multiple DICs in isolated micro-grids. *IEEE Trans. Power Syst.*, 30(5):2243–2256, Sept 2015.

[98] J. Machowski, J. W. Bialek, and J. R. Bumby. *Power System Dynamics: Stability and Control*. Wiley, 2008.

[99] E. Mallada. iDroop: A dynamic droop controller to decouple power grid's steady-state and dynamic performance. In *IEEE 55th Annual Conf. on Decision and Control (CDC)*, pages 4957–4964, Dec 2016.

[100] E. Mallada and S. H. Low. Distributed frequency-preserving optimal load control. In *IFAC World Congress*, volume 19, pages 5411–5418, 2014.

[101] E. Mallada and A. Tang. Improving damping of power networks: Power scheduling and impedance adaptation. In *IEEE 50th Annual Conf. on Decision and Control (CDC)*, pages 7729 – 7734, Orlando, FL, 2011.

[102] E. Mallada, C. Zhao, and S. Low. Optimal load-side control for frequency regulation in smart grids. In *52nd Annual Allerton Conf. on Communication, Control, and Computing*, pages 731–738, Sept 2014.

[103] T. Markvart. Microgrids: Power systems for the 21st century? *Refocus*, 7(4): 44–48, 2006.

[104] S. M. Melzer and B. C. Kuo. Optimal regulation of systems described by a countably infinite number of objects. *Automatica*, 7(3):359–366, 1971.

[105] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods for Multiagent Networks*. Princeton University Press, Princeton, NJ, 2010.

[106] R. H. Middleton and J. H. Braslavsky. String instability in classes of linear time invariant formation control with limited communication range. *IEEE Trans. Autom. Control*, 55(7):1519–1530, July 2010.

[107] U. Miekkala. Graph properties for splitting with grounded Laplacian matrices. *BIT*, 33(3):485–495, 1993.

[108] M. Milligan, B. Frew, B. Kirby, M. Schuerger, K. Clark, D. Lew, P. Denholm, B. Zavadil, M. O'Malley, and B. Tsuchida. Alternatives no more: Wind and solar power are mainstays of a clean, reliable, affordable grid. *IEEE Power Energy Mag.*, 13(6):78–87, Nov 2015.

[109] B. Mohar. The Laplacian spectrum of graphs. In *Graph Theory, Combinatorics, and Applications*, pages 871–898. Wiley, 1991.

[110] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei. A survey of distributed optimization and control algorithms for electric power systems. *IEEE Trans. Smart Grid*, 8(6): 2941–2962, Nov 2017.

[111] N. Motee and A. Jadbabaie. Optimal control of spatially distributed systems. *IEEE Trans. Autom. Control*, 53(7):1616–1629, Aug 2008.

[112] A. E. Motter, S. A. Myers, M. Anghel, and T. Nishikawa. Spontaneous synchrony in power-grid networks. *Nat. Phys.*, 9(3):191–197, Feb 2013.

[113] W. Ni and D. Cheng. Leader-following consensus of multi-agent systems under fixed and switching topologies. *Syst. Control Lett.*, 59(3):209–217, 2010.

[114] T. Nishikawa and A. E. Motter. Comparative analysis of existing models for power-grid synchronization. *New J. Phys.*, 17(1):015012, 2015.

[115] O. Gjerde and G. Berglund and A. Hjelle. Annual report. Technical report, Nordel, 1999.

[116] University of Washington. Power systems test case archive, 2018. URL `https://egriddata.org/group/uw-power-systems-test-case-archive`.

[117] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. of the IEEE*, 95(1):215–233, Jan 2007.

[118] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, 49 (9):1520–1533, Sept 2004.

[119] A. Olshevsky. Minimal controllability problems. *IEEE Trans. Control Network Syst.*, 1(3):249–258, Sept 2014.

[120] P. Kundur *et al.* Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions. *IEEE Trans. Power Syst.*, 19(3):1387–1401, Aug 2004.

[121] K. R. Padiyar. *HVDC power transmission systems: Technology and system interactions*. New Age International, New Delhi, 1990.

[122] F. Paganini and E. Mallada. Global performance metrics for synchronization of heterogeneously rated power systems: The role of machine models and inertia. In *55th Annual Allerton Conf. on Communication, Control, and Computing*, pages 324–331, Oct 2017.

[123] M. A. Pai. *Power System Stability by Lyapunov's Method*. N.Holland Publishing Co., New York, NY, 1981.

[124] M. A. Pai. *Energy Function Analysis for Power System Stability*. Kluwer Academic Publishers, Norwell, MA, 1989.

[125] F. Pasqualetti, C. Favaretto, S. Zhao, and S. Zampieri. Fragility and controllability tradeoff in complex networks. In *American Control Conf. (ACC)*, pages 216–221, June 2018.

[126] F. Pasqualetti, S. Zampieri, and F. Bullo. Controllability metrics, limitations and algorithms for complex networks. *IEEE Trans. Control Network Syst.*, 1 (1):40–52, March 2014.

[127] R. Pates, C. Lidström, and A. Rantzer. Control using local distance measurements cannot prevent incoherence in platoons. In *IEEE 56th Annual Conf. on Decision and Control (CDC)*, pages 3461–3466, Dec 2017.

[128] S. Patterson and B. Bamieh. Leader selection for optimal network coherence. In *IEEE 49th Annual Conf. on Decision and Control (CDC)*, pages 2692–2697, Dec 2010.

[129] S. Patterson and B. Bamieh. Consensus and coherence in fractal networks. *IEEE Trans. Control Netw. Syst.*, 1(4):338–348, Dec 2014.

[130] S. Patterson, N. McGlohon, and K. Dyagilev. Optimal k-leader selection for coherence and convergence rate in one-dimensional networks. *IEEE Trans. Control Network Syst.*, 4(3):523–532, Sept 2017.

[131] J. A. Peas Lopes, C. L. Moreira, and A. G. Madureira. Defining control strategies for microgrids islanded operation. *IEEE Trans. Power Syst.*, 21(2): 916–924, May 2006.

[132] L. M. Pecora and T. Carroll. Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.*, 80(10):2109–2112, 1998.

[133] M. Pirani, E. M. Shahrivar, and S. Sundaram. Coherence and convergence rate in networked dynamical systems. In *IEEE 54th Annual Conf. on Decision and Control (CDC)*, pages 968–973, Dec 2015.

[134] M. Pirani, E. Moradi Shahrivar, B. Fidan, and S. Sundaram. Robustness of leader-follower networked dynamical systems. *IEEE Trans. Control Network Syst.*, 2018. To appear.

[135] M. Pirani, J. W. Simpson-Porco, and B. Fidan. System-theoretic performance metrics for low-inertia stability of power networks. In *IEEE 56th Annual Conf. on Decision and Control (CDC)*, pages 5106–5111, Dec 2017.

[136] B. K. Poolla, S. Bolognani, and F. Dörfler. Optimal placement of virtual inertia in power grids. *IEEE Trans. Autom. Control*, 62(12):6209–6220, Dec 2017.

[137] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans. Usefulness of DC power flow for active power flow analysis. In *IEEE Power Eng. Society General Meeting*, pages 2457–2462. IEEE, 2005.

[138] A. Radmanesh, A. Naghash, and A. Mohamadifard. Optimal distributed control of multi agents: Generalization of consensus algorithms for high-order state derivatives of SISO and MIMO systems. In *Int. Conf. on Control, Automation and Robotics*, pages 606–611, Apr 2017.

[139] A. Rantzer. Scalable control of positive systems. *Eur. J. Control*, 24:72–80, 2015.

[140] W. Ren and E. Atkins. Second-order consensus protocols in multiple vehicle systems with local interactions. In *AIAA Guidance, Navigation, and Control Conf.*, San Francisco, CA, Aug 2005.

[141] W. Ren, R. W. Beard, and E. M. Atkins. Information consensus in multivehicle cooperative control. *IEEE Control Syst. Mag.*, 27(2):71–82, Apr 2007.

[142] W. Ren, K. Moore, and Y. Chen. High-order consensus algorithms in cooperative vehicle systems. In *IEEE Int. Conf. on Networking, Sensing and Control*, pages 457–462, 2006.

[143] W. Ren, K. L. Moore, and Y. Chen. High-order and model reference consensus algorithms in cooperative control of multivehicle systems. *J. Dyn. Syst. Meas. Contr.*, 129(5):678–688, Sep 2007.

[144] H. Rezaee and F. Abdollahi. Average consensus over high-order multiagent systems. *IEEE Trans. Autom. Control*, 60(11):3047–3052, Nov 2015.

[145] B. P. Roberts and C. Sandberg. The role of energy storage in development of smart grids. *Proc. of the IEEE*, 99(6):1139–1144, 2011.

[146] M Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Trans. Autom. Control*, 51(2):274–286, 2006.

[147] J. Schiffer, R. Ortega, A. Astolfi, J. Raisch, and T. Sezi. Conditions for stability of droop-controlled inverter-based microgrids. *Automatica*, 50(10): 2457–2469, 2014.

[148] J. Schiffer, T. Seel, J. Raisch, and T. Sezi. Voltage stability and reactive power sharing in inverter-based microgrids with consensus-based distributed voltage control. *IEEE Trans. Control Syst. Technol.*, 24(1):96–109, Jan 2016.

[149] P. Seiler, A. Pant, and K. Hedrick. Disturbance propagation in vehicle strings. *IEEE Trans. Autom. Control*, 49(10):1835–1842, Oct 2004.

[150] G. S. Seyboth and F. Allgöwer. Output synchronization of linear multiagent systems under constant disturbances via distributed integral action. In *American Control Conf. (ACC)*, pages 62–67, July 2015.

[151] Q. Shafiee, J. C. Vasquez, and J. M. Guerrero. Distributed secondary control for islanded microgrids - a networked control systems approach. In *38th Annual Conf. on IEEE Industrial Electronics Society*, pages 5637–5642, Oct 2012.

[152] J.-Y. Shao, J.-M. Guo, and H.-Y. Shan. The ordering of trees and connected graphs by algebraic connectivity. *Linear Algebra Appl.*, 428(7):1421–1438, 2008.

[153] M. Siami, S. Bolouki, B. Bamieh, and N. Motee. Centrality measures in linear consensus networks with structured network uncertainties. *IEEE Trans. Control Network Syst.*, 5(3):924–934, Sept 2018.

[154] M. Siami and N. Motee. Fundamental limits on robustness measures in networks of interconnected systems. In *IEEE 52nd Annual Conf. on Decision and Control (CDC)*, Florence, Italy, 2013.

[155] M. Siami and N. Motee. Systemic measures for performance and robustness of large-scale interconnected dynamical networks. In *IEEE 53rd Annual Conf. on Decision and Control (CDC)*, pages 5119–5124, Dec 2014.

[156] M. Siami and N. Motee. Fundamental limits and tradeoffs on disturbance propagation in large-scale dynamical networks. *IEEE Trans. Autom. Control*, 61(12):4055–4062, 2016.

[157] D. D. Šiljak. *Large-scale dynamic systems: Stability and structure*, volume 2. North Holland, 1978.

[158] J. W. Simpson-Porco, F. Dörfler, and F. Bullo. Droop-controlled inverters are Kuramoto oscillators. In *IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pages 264–269, Santa Barbara, CA, USA, September 2012.

[159] J. W. Simpson-Porco, F. Dörfler, and F. Bullo. Synchronization and power sharing for droop-controlled inverters in islanded microgrids. *Automatica*, 49 (9):2603 – 2611, 2013.

[160] J. W. Simpson-Porco, F. Dörfler, and F. Bullo. Voltage stabilization in microgrids via quadratic droop control. In *IEEE 52nd Annual Conf. on Decision and Control (CDC)*, pages 7582–7589, Dec 2013.

[161] J. W. Simpson-Porco, B. Kameshwar Poolla, N. Monshizadeh, and F. Dörfler. Input-output performance of linear-quadratic saddle-point algorithms with application to distributed resource allocation problems. *ArXiv preprint, arXiv:1803.02182*, 2018. Available: https://arxiv.org/abs/1803.02182.

[162] J. W. Simpson-Porco, B. K. Poolla, N. Monshizadeh, and F. Dörfler. Quadratic performance of primal-dual methods with application to secondary frequency control of power systems. In *IEEE 55th Annual Conf. on Decision and Control (CDC)*, pages 1840–1845, Dec 2016.

[163] J. W. Simpson-Porco, Q. Shafiee, F. Dörfler, J. C. Vasquez, J. M. Guerrero, and F. Bullo. Secondary frequency and voltage control of islanded microgrids via distributed averaging. *IEEE Trans. Ind. Electron.*, 62(11):7025–7038, Nov 2015.

[164] E. Sjödin. *The Price of Synchrony: Evaluating Transient Power Losses in Renewable Energy Integrated Power Networks.* MSc. thesis, KTH Royal Institute of Technology, 2013. Available: https://www.diva-portal.org/smash/get/diva2:662281/FULLTEXT01.pdf.

[165] E. Sjödin and D.F. Gayme. Transient losses in synchronizing renewable energy integrated power networks. In *American Control Conf. (ACC)*, pages 5217–5223, June 2014.

[166] S. Skogestad and I. Postlethwaite. *Multivariable feedback control: analysis and design.* Wiley, Chichester, 1996.

[167] D. A. Spielman and S.-H. Teng. Spectral partitioning works: Planar graphs and finite element meshes. *Linear Algebra Appl.*, 421(2):284 – 305, 2007. Special Issue in honor of Miroslav Fiedler.

[168] K. Stephenson and M. Zelen. Rethinking centrality: Methods and examples. *Social Networks*, 11(1):1 – 37, 1989.

[169] S. Stüdli, M. M. Seron, and R. H. Middleton. From vehicular platoons to general networked systems: String stability and related concepts. *Annu. Rev. Control*, 44:157 – 172, 2017.

[170] T. H. Summers, F. L. Cortesi, and J. Lygeros. Corrections to "on submodularity and controllability in complex dynamical networks". *IEEE Trans. Control Network Syst.*, 5(3):1503–1503, Sept 2018.

[171] T. H. Summers and J. Lygeros. Optimal sensor and actuator placement in complex dynamical networks. In *IFAC World Congress*, volume 19, pages 3784–3789, 2014.

[172] D. Swaroop and J. K. Hedrick. String stability of interconnected systems. *IEEE Trans. Autom. Control*, 41(3):349–357, Mar 1996.

[173] D. Swaroop and J. K. Hedrick. Constant spacing strategies for platooning in automated highway systems. *Trans. ASME J. Dyn. Sys., Meas., Control*, 121(3):462–270, 1999.

[174] E. Tegling. *On performance limitations of large-scale networks with distributed feedback control*. Licentiate thesis, KTH Royal Institute of Technology, 2016. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-186180.

[175] E. Tegling, B. Bamieh, and D.F. Gayme. The price of synchrony: Evaluating the resistive losses in synchronizing power networks. *IEEE Trans. Control Netw. Syst.*, 2(3):254–266, Sept 2015.

[176] E. Tegling, P. Mitra, H. Sandberg, and B. Bamieh. On fundamental limitations of dynamic feedback control in regular large-scale networks. *ArXiv preprint, arXiv:1710.02880*, 2018. Available: https://arxiv.org/abs/1710.02880.

[177] A. Tondl. *Some problems of rotor dynamics*. Czechoslovak academy of sciences, Prague, 1965.

[178] S. Trip, M. Bürger, and C. De Persis. An internal model approach to (optimal) frequency regulation in power grids with time-varying voltages. *Automatica*, 64:240 – 253, 2016.

[179] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie. Minimal actuator placement with bounds on control effort. *IEEE Trans. Control Network Syst.*, 3(1):67–78, March 2016.

[180] U.S. Energy Information Administration. Annual energy outlooks 2010 with projections to 2035. Technical Report DOE/EIA-0383, U.S. Dept. of Energy, 2010. URL `http://www.eia.doe.gov/oiaf/aeo`.

[181] D. Van Hertem and M. Ghandhari. Multi-terminal VSC HVDC for the European supergrid: Obstacles. *Renewable and Sustainable Energy Rev.*, 14(9): 3156–3163, 2010.

[182] P. Varaiya. Smart cars on smart roads: Problems of control. *IEEE Trans. Autom. Control*, 38(2):195–207, Feb 1993.

[183] P. Varaiya, F. F. Wu, and R.-L. Chen. Direct methods for transient stability analysis of power systems: Recent results. *Proc. IEEE*, 73(12):1703 – 1715, Dec 1985.

[184] P. G. Voulgaris. A convex characterization of classes of problems in control with specific interaction and communication structures. In *American Control Conf. (ACC)*, pages 3128–3133, June 2001.

[185] S.-H. Wang and E. J. Davison. On the stabilization of decentralized control systems. *IEEE Trans. Autom. Control*, 18(5):473–478, 1973.

[186] Y. Wang, N. Matni, and J. C. Doyle. Separable and localized system level synthesis for large-scale systems. *IEEE Trans. Autom. Control*, 2018. ISSN 0018-9286. To appear.

[187] J. L. Willems. Time-varying feedback for the stabilization of fixed modes in decentralized control systems. *Automatica*, 25(1):127–131, 1989.

[188] X. Wu, F. Dörfler, and M. R. Jovanović. Topology identification and design of distributed integral action in power networks. In *American Control Conf. (ACC)*, pages 5921–5926, July 2016.

[189] W. Xia and M. Cao. Analysis and applications of spectral properties of grounded Laplacian matrices for directed networks. *Automatica*, 80:10–16, 2017.

[190] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *J. Parallel Distrib. Comput.*, 67(1):33–46, 2007.

[191] S. K. Yadlapalli, S. Darbha, and K. R. Rajagopal. Information flow and its relation to stability of the motion of vehicles in a rigid formation. *IEEE Trans. Autom. Control*, 51(8):1315–1319, Aug 2006.

[192] G. F. Young, L. Scardovi, and N. E. Leonard. Robustness of noisy consensus dynamics with directed communication. In *American Control Conf. (ACC)*, pages 6312–6317, June 2010.

[193] C. Zhan, G. Chen, and L. F. Yeung. On the distributions of Laplacian eigenvalues versus node degrees in complex networks. *Physica A*, 389(8):1779–1788, 2010.

[194] F. Zhang and Q. Zhang. Eigenvalue inequalities for matrix product. *IEEE Trans. Autom. Control*, 51(9):1506–1509, 2006.

[195] X. Zhang and A. Papachristodoulou. A real-time control framework for smart power networks. *Automatica*, 58(C):43–50, Aug 2015.

[196] C. Zhao, E. Mallada, and F. Dörfler. Distributed frequency control for stability and economic dispatch in power networks. In *American Control Conf. (ACC)*, pages 2359–2364, July 2015.

[197] C. Zhao, U. Topcu, N. Li, and S. H. Low. Design and stability of load-side primary frequency control in power systems. *IEEE Trans. Autom. Control*, 59(5):1177–1189, May 2014.

[198] J. Zhao and F. Dörfler. Distributed control and optimization in DC microgrids. *Automatica*, 61:18–26, 2015.

[199] Q. Zhong and T. Hornik. *Control of Power Inverters in Renewable Energy and Smart Grid Integration*. Wiley-IEEE Press, Chichester, UK, 2013.

[200] K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, New Jersey, 1996.

[201] Z. Zuo, B. Tian, M. Defoort, and Z. Ding. Fixed-time consensus tracking for multi-agent systems with high-order integrator dynamics. *IEEE Trans. Autom. Control*, 63(2):563–570, Feb 2018.