Structure identification and optimal design of large-scale networks of dynamical systems

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Fu Lin

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor of Philosophy

Professor Mihailo R. Jovanović

December, 2012

Acknowledgements

I would like to thank my advisor Professor Mihailo Jovanović. I am extremely fortunate to have the opportunity to work with Mihailo during my graduate study. His commitment to excellence in everything he does has always inspired me to work hard. His constant encouragement has helped me stay positive with research challenges, and his constructive criticism has made me be critical to my own work. I can not thank him enough for his guidance and support throughout my years in graduate school.

I am truly indebted to Professor Makan Fardad for his generosity and patience in our collaboration. His insightful comments and suggestions have always led to improvement of our joint work.

My gratitude extends to Professors Murti Salapaka, Zhi-Quan Luo, Georgios Giannakis, and Peter Seiler for serving on my dissertation committee. I am also thankful to Professors Tryphon Georgiou, Gary Balas, Yousef Saad, Huiqiang Jiang, and Daniel Spirn. I have benefited from interactions with them and from classes taught by them.

I am fortunate to have friends Xiao Shen, Qiang Fu, Xue Bai, Shahrouz Takyar, Rashad Moarref, Xianhua Jiang, Ke Zhou, Donatello Materassi, Zhiqiang Xing, Binh Lieu, Lipeng Ning, Armin Zare, Neil Dingra, Xiaofan Wu, Eric Dahlberg, Yongxin Chen, Wei Zhang, Yinglong Feng, and many others. They have made my stay in Minnesota a unique experience that is full of fun. I am particularly thankful to Binh Lieu for his help and support and for his organizing many memorable events.

I would also like to express my gratitude to Ann Erickson for helping me adjust to American culture when I arrived in Minneapolis, and to staff members of ECE department Linda Jagerson, Jeanine Maiden, Chimai Nguyen, and Jamie Halleckson for their help.

Finally, and most importantly, my deepest gratitude goes to my family. This dissertation would not be possible without their love. It was my grandfather Shunqin Lin who encouraged me to pursue my PhD in the US, it was my wife Sizhou Xue who accompanied me in this journey, and it was my son Tao Xue Lin who gave me strength to overcome difficult times. To them I dedicate this dissertation.

Dedication

To my grandfather Shunqin,

my wife Sizhou,

and my son Tao

Abstract

This dissertation is about structure identification and optimal control of large-scale networks of dynamical systems. It contains four parts. In Part I, we focus on identification of controller architectures that strike a balance between the performance of the system and the sparsity of the controller. This is achieved by solving a parameterized family of sparsity-promoting optimal control problems whose solution traces the trade-off curve that starts at the centralized controller and ends at the sparse controller of interest.

Part II is devoted to the design of sparse communication graphs for consensus networks. This class of problems is commonly seen in distributed estimation and control. We show that the sparsity-promoting control problem can be formulated as a semidefinite program whose globally optimal solution can be computed efficiently.

In Part III, we consider optimal localized control of vehicular formations with nearest neighbor interactions. We identify a class of convex problems by restricting the controller to symmetric feedback gains. For the design of non-symmetric gains, we solve a parameterized family of problems whose solution gradually changes from the spatially uniform gain to the optimal non-symmetric gain. We investigate the performance of localized controllers in large formations and show that the best performance is achieved with non-symmetric and spatially-varying controllers.

Finally, in Part IV, we consider the leader selection problem in consensus networks, in which leaders have access to their own state in addition to relative information exchange with neighbors. We are interested in selecting an a priori specified number of nodes as leaders such that the steady-state variance of the deviation from consensus is minimized. For this combinatorial optimization problem, we develop efficient algorithms to compute lower and upper bounds on the global optimal value.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Large-scale networks of dynamical systems are becoming increasingly important in science and technology [1–4]. These systems arise in a variety of applications ranging from economics, social networks, power systems, and robotics. Network science has emerged as an interdisciplinary field that draws from diverse disciplines including graph theory, matrix theory, dynamical systems, optimization, and statistical mechanics. Without a doubt, technological advancements form the main impetus for the rapid development of network science, with the advent of Internet, World Wide Web, and networking services such as Facebook and Twitter. On the engineering side, networks of dynamical systems have found wide applications in sensor networks, formations of automated vehicles, and space-borne optical interferometry, just to name a few.

In modeling, analysis, and control of networks of dynamical systems, it is of fundamental importance to understand the interplay between network structure and underlying dynamical properties. Network topology dictates information patterns of interconnected subsystems. These constraints on the flow of information between subsystems impose performance limitations in the control of networked systems. An important issue in the control design for networked systems is the selection of information exchange

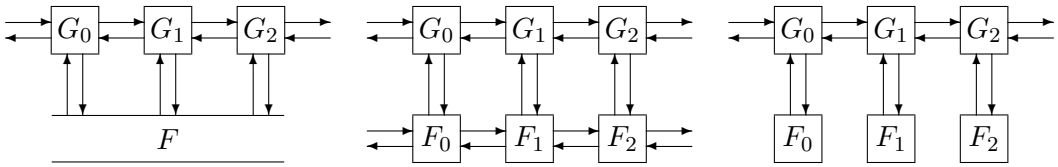


Figure 1.1: Distributed controllers with centralized, localized, and decentralized architectures.

network in the distributed controller. There is a trade-off between the achievable performance and the communication and computation cost. The best possible performance is attained if all controllers can communicate with each other, and as a whole decide on the control actions to be applied to each subsystem. However, this comes at the expense of excessive communication and computation. The other extreme is when every controller acts in isolation, and applies control actions to its corresponding subsystem based on the subsystem's output measurements. This fully decentralized scenario places minimal communication and computation requirement on the controller, but it may come at the expense of poor performance. A desired scenario, and a reasonable middle ground, is the localized information exchange in the distributed controller; see Figure 1.1 for an illustration of different controller architectures.

1.1 Main topics of the dissertation

1.1.1 Structured state feedback synthesis and sparsity-promoting optimal control

Design of distributed controllers with structural constraints is known to be a difficult problem. Research efforts have focused on identifying classes of distributed control problems that are tractable [5–18]. For spatially invariant cone- and funnel-causal systems, the design of quadratically optimal controllers can be cast into a convex problem

if the information in the controller propagates at least as fast as in the plant [7, 11]. A similar but more general algebraic characterization of the constraint set was introduced and convexity was established under the condition of quadratic invariance in [12]. Since these convex formulations are expressed in terms of the impulse response parameters, they do not lend themselves easily to state-space characterization. In [16], a state-space realization of optimal distributed controllers that satisfy cone-causality property was provided and methods for the design of sub-optimal controllers were developed. In [10] it was shown that the design of distributed controllers for systems with lower triangular structure is more amenable to convex analysis. Furthermore, explicit state-space formulae have been obtained for two-player linear-quadratic regulator problem [15, 18].

Design of structured state feedback gains

In the context of state feedback synthesis, localized architectures can be expressed in the form of *sparsity* constraints. For example, consider the mass-spring system with N masses shown in Fig. 1.2. Suppose that the control input $u_i(t) \in \mathbb{R}$ at the ith mass is determined by a linear combination of the states of the ith mass and the two neighboring masses

$$u_i(t) \;=\; -\,F_{i-1}\,x_{i-1}(t) \;-\; F_i\,x_i(t) \;-\; F_{i+1}\,x_{i+1}(t),$$

where the state $x_i(t) \in \mathbb{R}^2$ consists of the position and velocity of the ith mass and $F_i \in \mathbb{R}^{1\times 2}$ is the state feedback gain. The vector of control inputs can be expressed as

$$u(t) \;=\; -\,F\,x(t),$$

where the *structured* state feedback gain F belongs to the set of *block tridiagonal* matrices of size $\mathbb{R}^{N\times 2N}$ with block size $\mathbb{R}^{1\times 2}$.

For systems on general graphs, sparsity patterns of the feedback gain matrix F can

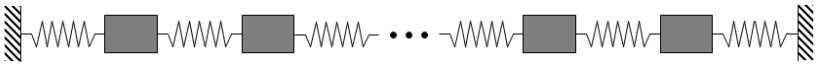


Figure 1.2: Mass-spring system on a line.

be more complex. Let the subspace $\mathcal{S}$ encapsulate these structural constraints on F and let J be a quadratic performance index with the feedback gain F. The structured optimal control problem can then be formulated as

$$\begin{aligned} \underset{F}{\text{minimize}} \quad & J(F) \\ \text{subject to} \quad & F \in \mathcal{S}. \end{aligned} \tag{1.1}$$

In what follows, we focus on the $\mathcal{H}_2$ norm of the closed-loop system (see Chapter 2). Both the structural constraint $F \in \mathcal{S}$ and the $\mathcal{H}_2$ performance J are easy to deal with when they are considered separately. In particular, the structural constraint $F \in \mathcal{S}$ forms a linear subspace which is a convex set of the state feedback gain F, and the closed-loop $\mathcal{H}_2$ norm J permits a convex formulation with an appropriate change of variables (e.g., see [19]). However, challenges arise when these convex characterizations are expressed in terms of different sets of optimization variables.

To illustrate this, consider the structured stabilizing feedback problem for the matrix pair (A, B), i.e., we seek an $F \in \mathcal{S}$ such that $A - BF$ is Hurwitz (under the assumption that such a stabilizing gain exists). The closed-loop stability is equivalent to the solvability of the inequality

$$(A - BF)X + X(A - BF)^T \prec 0 \tag{1.2}$$

with a positive definite matrix $X = X^T \succ 0$. A change of variables

$$Y = FX$$

converts (1.2) into a linear matrix inequality (LMI)

$$\left[\begin{array}{cc} A & B \end{array}\right] \left[\begin{array}{c} X \\ Y \end{array}\right] + \left[\begin{array}{cc} X & Y^T \end{array}\right] \left[\begin{array}{c} A^T \\ B^T \end{array}\right] \prec 0,$$

which is convex in X and Y. However, imposing structural constraint $F \in \mathcal{S}$ leads to

$$F = YX^{-1} \in \mathcal{S},$$

which is a nonconvex constraint on X and Y.

One simple but restrictive condition that retrieves convexity is to require that X be a diagonal matrix [20]. In that case, the structural constraint $F \in \mathcal{S}$ is equivalent to $Y \in \mathcal{S}$. Alternative characterizations of closed-loop stability have been obtained and several LMI-based algorithms have been developed; e.g., see [21–26].

Alternatively, we employ an augmented Lagrangian approach for the structured $\mathcal{H}_2$ problem (1.1) (see Chapter 4). This approach alleviates the difficulty of finding a structured stabilizing gain by solving a sequence of unstructured problems whose solution gradually converges to the structured optimal $\mathcal{H}_2$ controller. In addition, we utilize the sensitivity interpretation of the Lagrange multiplier to identify relaxations of structural constraints that are effective in improving the $\mathcal{H}_2$ performance.

Sparsity-promoting optimal control

In situations in which it is difficult to *a priori* assign the structural constraint set $\mathcal{S}$, it is useful to distinguish between different sparsity structures. Since the achievable performance depends on the constraint set $\mathcal{S}$, it is important to identify sparsity structures that strike a balance between the performance of the system and the sparsity of the controller. One approach is to minimize the $\mathcal{H}_2$ norm J subject to an upper bound on

the number of nonzero elements of F. This problem can be expressed as

$$\begin{array}{ll} \underset{F}{\text{minimize}} & J(F) \\ \text{subject to} & \mathbf{card}\,(F) \;\leq\; k, \end{array} \tag{1.3}$$

where k is a pre-specified positive integer and $\mathbf{card}(\cdot)$ is the *cardinality* function that counts the number of nonzero elements of a matrix. One approach is to check all possible sparsity patterns with cardinality less than k and to solve problem (1.1) for each sparsity pattern with a fixed structural constraint set $\mathcal{S}$. However, the computational cost of such an exhaustive search grows exponentially with the problem size (e.g., for $F \in \mathbb{R}^{5\times 10}$ and $k = 5$ the number of possible sparsity structures is approximately 2×10^6).

Instead, we minimize the weighted sum of the cost function J and the cardinality function of F

$$\underset{F}{\text{minimize}} \quad J(F) \,+\, \gamma\,\mathbf{card}\,(F) \tag{1.4}$$

where $\gamma \geq 0$ is the problem parameter. As γ varies over $[0, \infty)$, the solution to (1.4) traces a trade-off curve between the performance J and the cardinality of F; see Fig. 1.3 for an illustration. For $\gamma = 0$, the unique global solution to the $\mathcal{H}_2$ problem can be obtained from the positive definite solution to the algebraic Riccati equation [27]. This optimal solution usually leads to a centralized controller with a dense feedback gain matrix F. By gradually increasing γ, the optimal feedback gain follows a solution path from the centralized gain to the sparse gain of interest, until the desired balance between the performance of the system and the sparsity of the controller is achieved.

A challenging aspect of the optimal control problem (1.4) arises from the fact that both the $\mathcal{H}_2$ norm J and the cardinality function are nonconvex functions of F. We

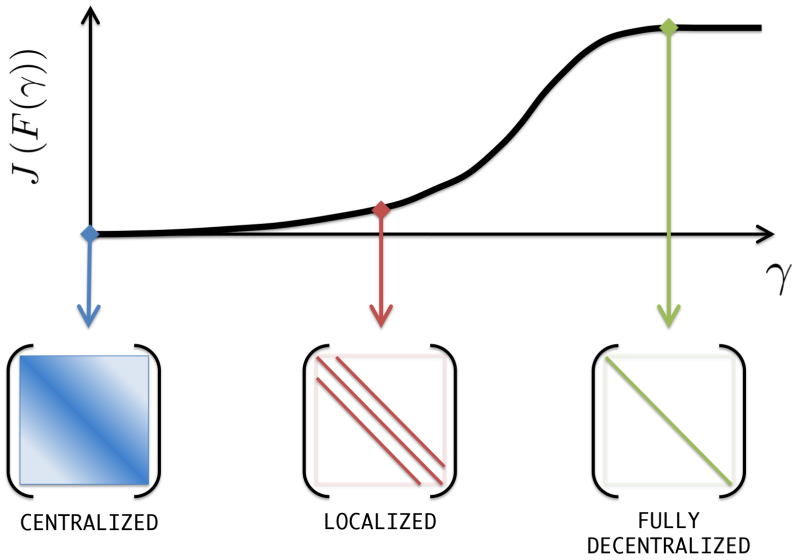


Figure 1.3: As γ increases, the cost function J increases and the feedback gain matrix F becomes sparser.

employ convex relaxations of the cardinality function such as the ℓ_1 norm

$$g(F) \;=\; \sum_{i,\,j} |F_{ij}|$$

and the weighted ℓ_1 norm

$$g(F) \;=\; \sum_{i,\,j} W_{ij}|F_{ij}|, \qquad W_{ij} \;\geq\; 0,$$

which can be more aggressive than the ℓ_1 norm in promoting sparsity. Furthermore, we also identify a class of distributed systems in consensus networks whose $\mathcal{H}_2$ performance permits convex formulations (see Chapter 6).

To solve (1.4) and its convex relaxations for each fixed γ, we use the alternating direction method of multipliers. This method alternates between promoting the sparsity of the controller and optimizing the performance of the system, which allows us to exploit

the structure of the corresponding objective functions. We demonstrate the effectiveness of this approach using several examples in Chapter 2. In addition, the MATLAB source codes can be found at

www.ece.umn.edu/∼mihailo/software/lqrsp/

1.1.2 Sparsity-promoting optimal control for consensus networks

Reaching consensus in a decentralized fashion is an important problem in network science [4]. This problem is often encountered in social networks where a group of individuals is trying to agree on a certain issue [28,29]. The related load balancing problem (i.e., to distribute computational load evenly over a network of processors) has been studied extensively in computer science [30–32]. Recently, consensus problems have received considerable attention in the context of distributed control [33–36].

Consider a network with N nodes in which each node updates its scalar state using a weighted sum of the differences between its own state and the states of other nodes

$$\dot{x}_i(t) \,=\, -\sum_{j\neq i} F_{ij}\,(x_i(t) \,-\, x_j(t)), \qquad i \,=\, 1,\ldots,N.$$

For undirected networks (see Chapter 6), we have

$$F_{ij} \,=\, F_{ji}, \qquad i \,\neq\, j.$$

Under the assumption that the network is connected, all node values will converge to the average of the initial conditions

$$\lim_{t\,\to\,\infty} x_i(t) \,=\, \frac{1}{N}\sum_{i\,=\,1}^{N} x_i(0), \qquad i \,=\, 1,\ldots,N.$$

In a number of applications, it is desired to maintain consensus in the presence of

uncertainty, e.g., introduced by modeling errors, measurement noise, or communication failures [4, 32, 37]. This motivates the study of robustness of stochastically forced consensus networks

$$\dot{x}_i(t) \;=\; -\sum_{j\neq i} F_{ij}\,(x_i(t) \;-\; x_j(t)) \;+\; d_i(t), \qquad i \;=\; 1,\ldots,N,$$

where $d_i(t) \in \mathbb{R}$ is the disturbance at node i. Under the influence of disturbances, the network will not reach a consensus value; each node will fluctuate around the average of node values.

We use the steady-state variance of the *deviation from average* to quantify the performance of stochastically forced consensus networks. We show that the steady-state variance can be expressed as

$$J(F) \;=\; \frac{1}{2}\,\mathrm{trace}\left(Q(F \;+\; \mathbb{1}\mathbb{1}^T/N)^{-1}\right),$$

where $Q = Q^T \succeq 0$ is an appropriately-chosen weight matrix and $\mathbb{1} \in \mathbb{R}^N$ is the vector of all ones. This result facilitates semidefinite programming (SDP) formulations for both the structured optimal control problem (1.1) and the sparsity-promoting problem with the (weighted) ℓ_1 norm as the penalty function. In particular, we show that

$$\underset{F}{\text{minimize}} \quad J(F) \;+\; \gamma \sum_{i,\,j} W_{ij}|F_{ij}|$$

can be cast as the following SDP (see Chapter 6 for details)

$$\begin{aligned} \underset{X,\,Y,\,F}{\text{minimize}} \quad & \frac{1}{2}\,\text{trace}\,(X) \;+\; \gamma \sum_{i,\,j} Y_{ij} \\ \text{subject to} \quad & \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & F + \mathbb{1}\mathbb{1}^T/N \end{bmatrix} \succeq 0 \\ & F\,\mathbb{1} \;=\; 0 \\ & -Y \;\leq\; W \circ F \;\leq\; Y. \end{aligned}$$

Thus, it can be solved efficiently using standard SDP solvers.

1.1.3 Optimal localized control of vehicular formations

A system of N identical vehicles moving along a straight line is shown in Fig. 1.4a. All vehicles are equipped with ranging devices that allow them to measure relative distances with respect to their immediate neighbors. The objective is to design an optimal controller that uses only local information (i.e., relative distances between the neighboring vehicles) to keep each vehicle at its global position on a grid of regularly spaced points moving with a constant velocity.

Recent work in this area has focused on fundamental performance limitations of both centralized and decentralized controllers for large-scale formations [38–40]. For centralized linear quadratic optimal control formulations based on penalizing relative position errors, it was shown in [39] that stabilizability and detectability deteriorate as formation size increases. In [40], it was shown that merge and split maneuvers can exhibit poor convergence rates even upon inclusion of absolute position errors in cost functionals. In [38], it was shown that sensitivity of spacing errors to disturbances increases with the number of vehicles for formations with localized symmetric controllers that utilize relative position errors between neighboring vehicles.

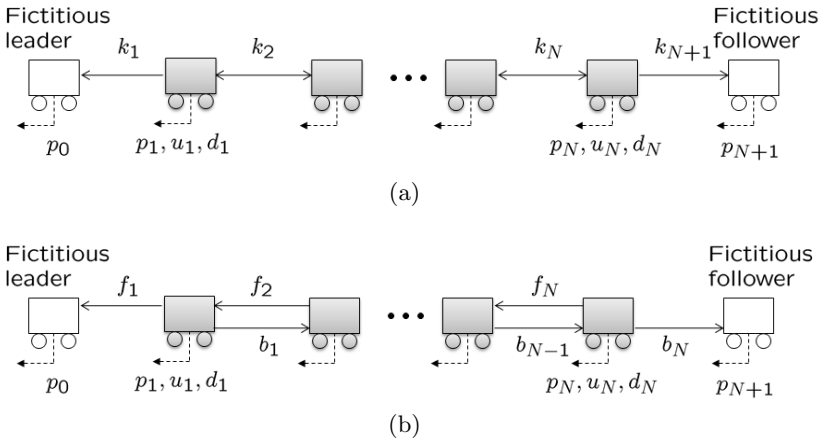


Figure 1.4: Formation of vehicles with nearest neighbors interactions: (a) symmetric gains and (b) non-symmetric gains.

The design of localized controllers can be formulated as the structured $\mathcal{H}_2$ problem (1.1). For symmetric feedback gains shown in Fig. 1.4a, we show that the optimal symmetric localized controller K can be obtained by solving the following SDP

$$\begin{array}{ll} \underset{X,\,K}{\text{minimize}} & \dfrac{1}{2}\,\text{trace}\,(X\,+\,K) \\ \text{subject to} & K\,\succ\,0, \quad K\in\mathcal{S}_K, \quad \left[\begin{array}{cc} K & Q^{1/2} \\ Q^{1/2} & X \end{array}\right]\,\succeq\,0, \end{array}$$

where $\mathcal{S}_K$ imposes linear equality constraints on K; see Chapter 9 for details.

For non-symmetric feedback gains shown in Fig. 1.4b, we solve a parameterized family of problems that ranges between an easily solvable problem and the problem of interest. In particular, we consider the parameterized weight matrix in the structured $\mathcal{H}_2$ problem

$$Q(\varepsilon)\;=\;Q_0\;+\;\varepsilon\,(Q_d\;-\;Q_0)$$

where Q_0 is an appropriately-selected initial weight, Q_d is the *desired* weight, and

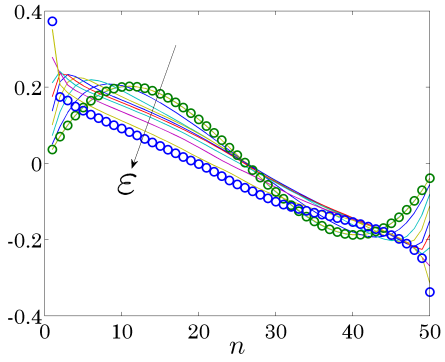


Figure 1.5: Normalized optimal forward gain profile changes from an almost sinusoidal shape at $\varepsilon = 10^{-4}$ to an almost piecewise linear shape at $\varepsilon = 1$ for a formation with $N = 50$ vehicles.

$\varepsilon \in [0, 1]$ is the homotopy parameter. Note that $Q = Q_0$ for $\varepsilon = 0$, and $Q = Q_d$ for $\varepsilon = 1$.

The homotopy-based Newton's method consists of three steps: (i) For $\varepsilon = 0$, we find the initial weight Q_0 with respect to which a spatially uniform gain F_0 is *inversely optimal*. This is equivalent to solving the structured optimal problem analytically with the performance weight Q_0. (ii) For $0 < \varepsilon \ll 1$, we employ perturbation analysis to determine the first few terms in the expansion

$$F(\varepsilon) \;=\; F_0 \;+\; \sum_{n=1}^{\infty} \varepsilon^n F_n.$$

(iii) We gradually increase ε and use the structured optimal gain obtained for the previous value of ε to initialize the next round of Newton iterations. This process is repeated until the desired value $\varepsilon = 1$ is reached; see Fig. 1.5.

We study the performance of the optimal symmetric and non-symmetric gains by examining the dependence on the formation size of global and local performance measures and the control effort. The global performance measure quantifies the resemblance of the formation to a rigid lattice, while the local performance measure quantifies relative

position errors between neighboring vehicles. Our results demonstrate that the best performance is achieved with the optimal localized controller that is both non-symmetric and spatially-varying.

1.1.4 Algorithms for leader selection in consensus networks

We consider consensus networks with two groups of nodes. Ordinary nodes, the so-called *followers*, form their control actions using relative information exchange with their neighbors; while special nodes, the so-called *leaders*, in addition to relative information from their neighbors also have access to their own states. This setting arises, for example, in the control of vehicular formations where all vehicles are equipped with ranging devices that provide relative distances from their neighbors, and the leaders additionally have GPS devices that provide their global position information.

Suppose that we want to equip a number of vehicles with GPS devices to keep all vehicles in the formation in their desired positions under the influence of noise. More precisely, we are interested in assigning an *a priori* specified number of nodes as leaders to minimize the steady-state variance of the deviation from consensus of the network.

For undirected networks in which all nodes including leaders are subject to stochastic disturbances, we show that the noise-corrupted leader selection problem has a convex objective function. In spite of this, the combinatorial nature of Boolean constraints (a node is either a leader or it is not) makes determination of the global minimum challenging for large networks. Instead, we focus on developing efficient algorithms to compute lower and upper bounds on the globally optimal value. We demonstrate that the developed algorithms significantly outperform simple degree-based-heuristics approach and they sometimes achieve tight bounds on global optimally value for small networks.

We also consider the noise-free leader selection problem in which leaders are assumed to be immune to noise and they follow their desired trajectories at all times. This idealized setting has recently be studied by several authors; e.g., see [41, 42]. For connected networks with at least one leader, it was shown in [41] that adding leaders always improves performance. In view of this, the one-at-a-time greedy algorithm was proposed in [41] for the noise-free leader selection problem. It was shown in [42] that the variance of deviation from consensus is a supermodular function of the set of noise-free leaders. This implies that the performance improvement by adding additional leaders is *diminishing* as the number of leaders increases. Furthermore, the supermodular optimization framework is then employed to show that the greedy algorithm provides selection of leaders that is within a provable bound from globally optimal solution [42].

In contrast to the above references, we use convex optimization to select noise-free leaders. We first provide an explicit expression for the objective function to identify the source of nonconvexity and to suggest an LMI-based convex relaxation. We then relax the *hard* Boolean constraint on the number of leaders with a *soft constraint* by augmenting the objective function with the ℓ_1 norm of the optimization variables. The ℓ_1 norm provides a means for obtaining a sparse solution whose nonzero elements identify the leaders. The developed algorithm produces a trade-off curve between the number of noise-free leaders and the variance of the deviation from consensus by solving a parameterized family of convex optimization problems.

1.2 Organization of the dissertation

This dissertation consists of four parts and an appendix. Each part centers on one of the four topics described above. At the end of each part, we summarize the main contributions and discuss future research directions.

Part I is devoted to sparsity-promoting optimal control problem, and it contains

four chapters. In Chapter 2, we consider the design of sparse and block sparse feedback gains that minimize variance amplification of distributed systems. In Chapter 3, we demonstrate that the alternating direction method of multipliers (ADMM) is well-suited to sparsity-promoting optimal control. In Chapter 4, we employ augmented Lagrangian method to design structured feedback gains with *a priori* assigned sparsity patterns. In Chapter 5, we summarize the main contributions of Part I, and discuss extensions and future research directions.

Part II examines sparsity-promoting optimal control for consensus networks. Building upon the framework developed in Part I, we employ tools from convex optimization and algebraic graph theory to further exploit the structure of optimal control problems. In Chapter 6, we show that both sparsity-promoting and structured optimal control problems can be cast as semidefinite programs. Therefore, the globally optimal controllers can be computed efficiently; for simple networks, even analytical expressions can be obtained. In Chapter 7, we use a graph-theoretic representation of the feedback gain matrix to develop efficient customized algorithms and to establish asymptotic scaling trends of performance measures in large networks. In Chapter 8, we conclude Part II and outline possible extensions.

Part III considers optimal control of vehicular formations and performance limitations in large formations. In Chapter 9, we design optimal localized feedback gains for one-dimensional formations with nearest neighbor interactions. In Chapter 10, we examine how performance of the optimally-controlled vehicular formation scales with the number of vehicles. In Chapter 11, we summarize Part III and discuss open research problems.

Part IV addresses the problem of selecting an *a priori* assigned number of leaders to most effectively improve the robustness of consensus networks, where leaders have access to global reference information. In Chapter 12, we consider the noise-corrupted leader

selection problem. For this combinatorial optimization problem, we obtain efficiently computable lower and upper bounds on the globally optimal value. In Chapter 13, we consider the noise-free leader selection problem, provide its connections to the noise-corrupted leader selection problem, and develop a soft-constraint method based on the sparsity-promoting optimal control framework developed in Part I. In Chapter 14, we conclude Part IV with remarks on applications in real-world networks and on future research directions.

Appendix A deals with least-squares approximation of structured covariances. We formulate the structured covariance least-squares problem and develop efficient unconstrained maximization methods for the corresponding dual problem.

1.3 Contributions of the dissertation

The contributions of the dissertation are summarized as follows.

Part I

Design of optimal sparse feedback gains. We design sparse and block sparse feedback gains that minimize the variance amplification of distributed systems. Our approach consists of two steps. First, we identify sparsity patterns of feedback gains by incorporating sparsity-promoting penalty functions into the optimal control problem, where the added terms penalize the number of communication links in the distributed controller. Second, we optimize state feedback gains subject to structural constraints determined by the identified sparsity patterns. This polishing step improves the quadratic performance of the distributed controller. In the first step, we identify sparsity structure of feedback gains using the alternating direction method of multipliers. This method alternates between promoting sparsity of the controller and optimizing performance of the system, which allows us to exploit the structure of the corresponding objective

functions. In particular, we take advantage of the separability of the sparsity-promoting penalty functions to decompose the minimization problem into sub-problems that can be solved analytically.

Design of structured optimal state feedback gains. We consider the design of optimal state feedback gains subject to structural constraints on the distributed controllers. These constraints are in the form of sparsity requirements for the feedback matrix, implying that each controller has access to information from only a limited number of subsystems. The minimizer of this constrained optimal control problem is sought using the augmented Lagrangian method. Notably, this approach does not require a stabilizing structured gain to initialize the optimization algorithm. Motivated by the structure of the necessary conditions for optimality of the augmented Lagrangian, we develop an alternating descent method to determine the structured optimal gain. We also utilize the sensitivity interpretation of the Lagrange multiplier to identify relaxations of structural constraints that are effective in improving the $\mathcal{H}_2$ performance.

Part II

Identification of sparse communication graphs in consensus networks. We consider the design of distributed controller architectures for undirected networks of single-integrators. In the presence of stochastic disturbances, we identify communication topologies that balance the variance amplification of the network with the number of communication links. This is achieved by solving a parameterized family of sparsity-promoting optimal control problems whose solution traces the optimal trade-off curve that starts at the centralized controller and ends at the controller with sparse communication links. We show that the optimal control problem can be formulated as a semidefinite program whose global solution can be computed efficiently.

Design of communication graphs for consensus networks. We obtain several expressions for the $\mathcal{H}_2$ norm in terms of edge weights of the identified communication

graph. We show that the $\mathcal{H}_2$ performance is a convex function over the convex set of stabilizing feedback gains. For several simple graphs, we derive explicit formulae for the solution to the structured $\mathcal{H}_2$ problem and provide asymptotic scalings of local and global performance measures with respect to the network size.

Part III

Optimal control of vehicular formations with nearest neighbor interactions. We consider the design of optimal localized feedback gains for one-dimensional formations in which vehicles only use information from their immediate neighbors. The control objective is to enhance coherence of the formation by making it behave like a rigid lattice. For the single-integrator model with symmetric gains, we establish convexity, implying that the globally optimal controller can be computed efficiently. We also identify a class of convex problems for double-integrators by restricting the controller to symmetric position and uniform diagonal velocity gains.

To obtain the optimal non-symmetric gains for both the single- and the double-integrator models, we solve a parameterized family of optimal control problems ranging from an easily solvable problem to the problem of interest as the underlying parameter increases. When this parameter is kept small, we employ perturbation analysis to decouple the matrix equations that result from the optimality conditions, thereby rendering the unique optimal feedback gain. This solution is used to initialize a homotopy-based Newton's method to find the optimal localized gain.

Performance of vehicular formations with nearest neighbor interactions. We investigate the performance of the optimal localized controllers by examining how the coherence of large-scale stochastically forced formations scales with the number of vehicles. We provide a spatially uniform non-symmetric controller that outperforms the optimal spatially varying symmetric controller in the scaling trend of macroscopic

performance measure. This result indicates that departure from symmetric gains can improve coherence of large-scale formations and that the controller structure may play a more important role than the optimal selection of feedback gains. Our results show that the localized controller that achieves the best performance is both non-symmetric and spatially-varying.

Part IV

Algorithms for noise-corrupted leader selection. We establish convexity of the objective function in the noise-corrupted leader selection problem. Based on this result, we introduce a convex relaxation of Boolean constraints to obtain a lower bound on the globally optimal value. We provide semidefinite formulation of the convex relaxation and develop an efficient customized interior point method. We also use a simple but efficient greedy algorithm and the alternating direction method of multipliers to compute upper bounds on the globally optimal value. We exploit the structure of low-rank modifications to significantly reduce the computational complexity.

Algorithms for noise-free leader selection. We provide an explicit expression for the objective function to identify the source of nonconvexity and to suggest an LMI-based convex relaxation. We then relax the Boolean constraint on the number of leaders with a soft constraint by augmenting the objective function with the ℓ_1 norm of the optimization variables. The ℓ_1 norm provides a means for obtaining a sparse solution whose nonzero elements identify the leaders. The developed algorithm produces a trade-off curve between the number of noise-free leaders and the variance of the deviation from consensus by solving a parameterized family of convex optimization problems.

Appendix A

Least-squares approximation of structured covariances. We consider the optimal least-squares approximation of structured covariances. State covariances of linear

systems satisfy certain constraints imposed by the underlying dynamics. These constraints dictate a particular structure of state covariances. However, sample covariances almost always fail to have the required structure. The renewed interest in using state covariances for estimating the power spectra of inputs gives rise to the approximation problem. We formulate the structured covariance least-squares problem and convert the Lyapunov-type matrix linear constraint into an equivalent set of trace constraints. Efficient unconstrained maximization methods capable of solving the corresponding dual problem are developed.

Part I

Sparsity-promoting optimal control

Chapter 2

Optimal sparse feedback synthesis

We consider the design of sparse and block sparse feedback gains that minimize variance amplification of distributed systems. The design procedure consists of two steps: the structure identification step and the polishing step. In the identification step, we search for sparsity patterns $\mathcal{S}$ that strike a balance between the variance amplification (i.e., the $\mathcal{H}_2$ norm) of the system and the sparsity of the controller. In the polishing step, we improve the $\mathcal{H}_2$ performance by solving the structured $\mathcal{H}_2$ problem with controllers restricted to the identified sparsity patterns $\mathcal{S}$.

In the first step, we formulate the sparsity-promoting optimal control problem by augmenting the $\mathcal{H}_2$ performance with penalty functions that promote sparsity of the controller. We consider several penalty functions that induce elementwise or blockwise sparsity structures of the state feedback gain. In the absence of sparsity-promoting terms, we recover the standard $\mathcal{H}_2$ problem whose solution leads to centralized controllers. As we increase the emphasis on the sparsity of controllers, the solution to the optimal control problem traces a trade-off curve between the $\mathcal{H}_2$ performance of the system and the sparsity of the controller. In the second step, we fix the sparsity structure identified in the first step and search for the structured feedback gain that optimizes the

$\mathcal{H}_2$ performance. This polishing step improves the $\mathcal{H}_2$ performance of the distributed controller.

This chapter is organized as follows. In Section 2.1, we formulate the sparsity-promoting optimal control problem. In Section 2.2, we consider several penalty functions that promote sparsity or block sparsity. In Section 2.3, we describe the design procedure of the optimal sparse synthesis. In Section 2.4, we use several examples to illustrate the utility of the developed sparsity-promoting framework.

2.1 Sparsity-promoting optimal control problem

Consider the following control problem

$$\begin{aligned} \dot{x} &= A\,x + B_1\,d + B_2\,u \\ z &= C\,x + D\,u \\ u &= -F\,x, \end{aligned} \tag{2.1}$$

where d and u are the disturbance and control inputs, z is the performance output, $C = \begin{bmatrix} Q^{1/2} & 0 \end{bmatrix}^T$ and $D = \begin{bmatrix} 0 & R^{1/2} \end{bmatrix}^T$ with standard assumptions that $Q = Q^T \geq 0$, $R = R^T > 0$, (A, B_2) is stabilizable, and $(A, Q^{1/2})$ is detectable. The matrix F is the state feedback gain and the closed-loop system is given by

$$\begin{aligned} \dot{x} &= (A - B_2F)\,x + B_1\,d \\ z &= \begin{bmatrix} Q^{1/2} \\ -R^{1/2}F \end{bmatrix} x. \end{aligned} \tag{2.2}$$

We are interested in identifying sparsity patterns of F that strike a balance between the number of nonzero elements of F and the $\mathcal{H}_2$ norm of the system from d to z. Let

us consider the following optimization problem

$$\text{minimize} \;\; J(F) \,+\, \gamma\, \mathbf{card}\,(F) \tag{2.3}$$

where

$$\mathbf{card}\,(F) \,:=\, \text{number of nonzero elements of } F. \tag{2.4}$$

The $\mathcal{H}_2$ norm is defined as

$$J(F) \,=\, \begin{cases} \text{trace}\left(B_1^T P(F) B_1\right), & F \text{ stabilizing} \\ \infty, & \text{otherwise} \end{cases} \tag{2.5}$$

where the matrix $P(F)$ denotes the closed-loop observability Gramian

$$P(F) \;=\; \int_0^\infty \mathrm{e}^{(A-B_2F)^T t}\,(Q + F^T R F)\,\mathrm{e}^{(A-B_2F)t}\,\mathrm{d}t$$

and it can be computed by solving the Lyapunov equation

$$(A - B_2F)^T\, P \,+\, P\,(A - B_2F) \,=\, -\left(Q + F^T R F\right). \tag{2.6}$$

In the absence of the sparsity-promoting term $\mathbf{card}(F)$, the optimal control problem (2.3) becomes the standard $\mathcal{H}_2$ problem whose solution results in centralized controllers with dense feedback matrices. By incorporating the sparsity-promoting term $\mathbf{card}(F)$ in (2.3), we promote sparsity of the feedback gain. The scalar $\gamma \geq 0$ characterizes our emphasis on the sparsity of F; a larger γ encourages a sparser F. By gradually increasing γ, the optimal feedback gain traces a solution path from the centralized gain to the sparse gain of interest, until the desired balance between the $\mathcal{H}_2$ performance of the system and the sparsity of the controller is achieved.

2.2 Sparsity-promoting penalty functions

Problem (2.3) is a combinatorial optimization problem whose solution usually requires an intractable combinatorial search. In optimization problems where sparsity is desired, the cardinality function is typically replaced by the ℓ_1 norm of the optimization variable [43, Chapter 6],

$$g(F) = \sum_{i,j} |F_{ij}|. \tag{2.7}$$

The widespread use of the ℓ_1 norm as a proxy for cardinality minimization can be attributed to its effectiveness in recovering sparse signals subject to measurement noise [44]. The ℓ_1 norm has been widely used in statistics for model selection [45], in signal processing for sparse signal representation [46], and in image processing for noise removal [47]; see [48] for a brief survey.

Recently, a *weighted* ℓ_1 norm was used to enhance sparsity in signal recovery [48],

$$g(F) = \sum_{i,j} W_{ij}|F_{ij}| \tag{2.8}$$

where weights $W_{ij} \in \mathbb{R}$ are positive numbers. The weighted ℓ_1 norm tries to bridge the difference between the ℓ_1 norm and the cardinality function. In contrast to the cardinality function that assigns the *same* cost to any nonzero element, the ℓ_1 norm penalizes more heavily the elements of larger magnitudes. The positive weights can be chosen to counteract this magnitude dependence of the ℓ_1 norm. For example, if the weights W_{ij} are chosen to be *inversely proportional* to the magnitude of F_{ij},

$$W_{ij} = \begin{cases} 1/|F_{ij}|, & F_{ij} \neq 0, \\ 1/\varepsilon, & F_{ij} = 0, \quad 0 < \varepsilon \ll 1, \end{cases} \tag{2.9}$$

then the weighted ℓ_1 norm of F and the cardinality function of F coincide

$$\sum_{i,\,j} W_{ij}\,|F_{ij}| \;=\; \mathbf{card}\,(F)\,.$$

The above scheme (2.9), however, cannot be implemented because the weights depend on the unknown feedback gains. A reweighted algorithm that solves a sequence of weighted ℓ_1 optimization problems in which the weights are determined by the solution of the weighted ℓ_1 problem in the previous iteration was proposed in [48].

Both the ℓ_1 norm and its weighted version are *convex* relaxations of the cardinality function. On the other hand, we also consider *nonconvex* alternatives that could be more aggressive in promoting sparsity. Suppose that we wish to find the sparsest feedback gain that provides a given level of $\mathcal{H}_2$ performance $\sigma > 0$,

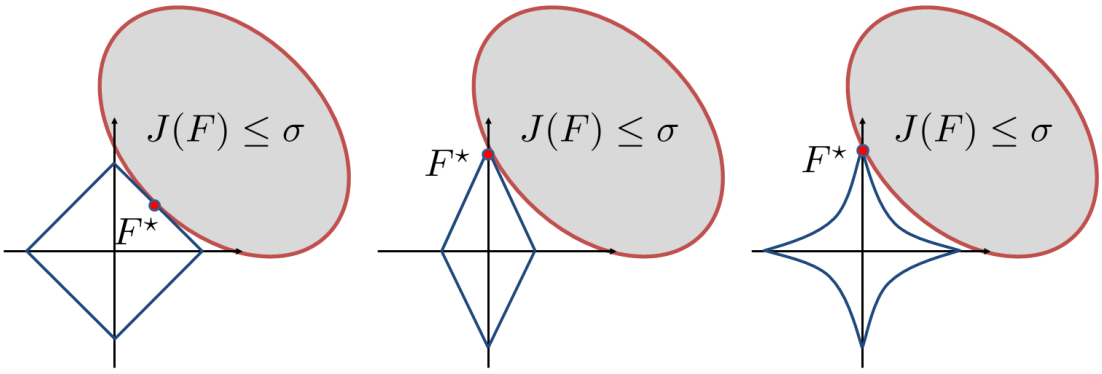
$$\begin{aligned} &\text{minimize} && \mathbf{card}\,(F) \\ &\text{subject to} && J(F) \;\leq\; \sigma. \end{aligned}$$

Approximating $\mathbf{card}\,(F)$ with a penalty function $g(F)$ yields

$$\begin{aligned} &\text{minimize} && g(F) \\ &\text{subject to} && J(F) \;\leq\; \sigma. \end{aligned} \tag{2.10}$$

Solution to (2.10) is the intersection of the constraint set $\mathcal{C} = \{F \,|\, J(F) \leq \sigma\}$ and the smallest sub-level set of g that touches $\mathcal{C}$; see Fig. 2.1. In contrast to the ℓ_1 norm whose sub-level sets are determined by the *convex* ℓ_1 ball, the sub-level sets of the *nonconvex* function (e.g., the ℓ_p norm with $0 < p < 1$) have a *star-like* shape. The *sum-of-logs* function

$$g(F) \;=\; \sum_{i,\,j} \log\left(1 \,+\, \frac{|F_{ij}|}{\varepsilon}\right), \quad 0 \,<\, \varepsilon \,\ll\, 1 \tag{2.11}$$



(a) (b) (c)

Figure 2.1: The solution $F^\star$ of the constrained problem (2.10) is the intersection of the constraint set $\mathcal{C} = \{F \,|\, J(F) \leq \sigma\}$ and the smallest sub-level set of g that touches $\mathcal{C}$. The penalty function g is (a) the ℓ_1 norm; (b) the weighted ℓ_1 norm with appropriate weights; and (c) the nonconvex function such as the sum-of-logs (2.11).

is another example of a nonconvex function with similar geometry of sub-level sets.

Remark 1. *Design of feedback gains that have block sparse structure can be achieved by promoting sparsity at the level of* submatrices *instead of at the level of individual elements. Let the feedback gain F be partitioned into submatrices $F_{ij} \in \mathbb{R}^{m_i \times n_j}$ that need not have the same size. The cardinality function, the weighted ℓ_1 norm, and the sum-of-logs can be generalized to matrix blocks by replacing the absolute value of F_{ij} in (2.4), (2.8) and (2.11) by the Frobenius norm $\|\cdot\|_F$ of F_{ij}, i.e.,*

$$\sum_{i,j} \mathbf{card}\left(\|F_{ij}\|_F\right), \qquad \sum_{i,j} W_{ij}\|F_{ij}\|_F, \qquad \sum_{i,j} \log\left(1 + \frac{\|F_{ij}\|_F}{\varepsilon}\right).$$

The use of the Frobenius norm $\|F_{ij}\|_F$ does not promote sparsity within the F_{ij} block; it instead promotes sparsity at the level of submatrices.

2.3 Design of sparse feedback gains

The sparsity-promoting feedback design makes use of the above discussed penalty functions. In order to obtain state feedback gains that strike a balance between the quadratic performance and the sparsity of the controller, we consider the following optimal control problem

$$\text{minimize} \;\; J(F) \,+\, \gamma\, g(F) \tag{SP}$$

where J is the closed-loop $\mathcal{H}_2$ norm and g is a sparsity-promoting penalty function, e.g., given by the cardinality function (2.4), the ℓ_1 norm (2.7), the weighted ℓ_1 norm (2.8), or the sum-of-logs (2.11). When the cardinality function in (2.4) is replaced by (2.7), (2.8), or (2.11), (SP) can be viewed as a relaxation of the combinatorial problem (2.3), obtained by approximating the cardinality function with corresponding penalty functions g.

As the parameter γ varies over $[0, +\infty)$, the solution of (SP) traces the trade-off curve between the $\mathcal{H}_2$ performance and the sparsity of the feedback gain. For $\gamma = 0$, (SP) becomes the standard LQR problem whose solution is given by the centralized gain

$$F_c \,=\, R^{-1} B_2^T P,$$

where P is the unique positive definite solution of the algebraic Riccati equation

$$A^T P \,+\, PA \,+\, Q \,-\, PB_2 R^{-1} B_2^T P \,=\, 0. \tag{2.12}$$

We then slightly increase γ and employ an iterative algorithm – the alternating direction method of multipliers (ADMM) – initialized by the optimal feedback matrix at the previous γ. The solution of (SP) becomes sparser as γ increases. After a desired level of sparsity is achieved, we fix the sparsity structure $\mathcal{S}$ and find the optimal structured

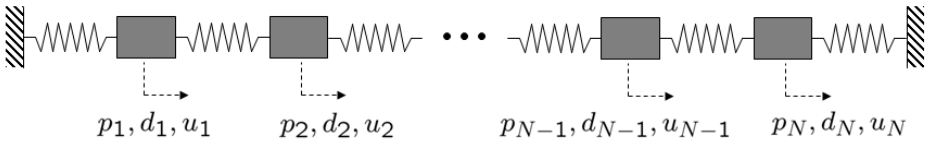


Figure 2.2: Mass-spring system on a line.

feedback gain by solving the structured $\mathcal{H}_2$ problem

$$\begin{aligned} &\text{minimize} && J(F) \\ &\text{subject to} && F \in \mathcal{S}. \end{aligned} \tag{2.13}$$

In Chapter 3, we develop effective algorithms that solve the sparsity-promoting optimal control problem (SP) and the structured $\mathcal{H}_2$ problem (2.13).

2.4 Examples

We next use three examples to illustrate the utility of the approach described in Section 2.3. The identified sparsity structures result in *localized* controllers in all three cases. Additional information about these examples, along with MATLAB source codes, can be found at

www.ece.umn.edu/∼mihailo/software/lqrsp/

2.4.1 Mass-spring system

For a mass-spring system with N masses shown in Fig. 2.2, let p_i be the displacement of the ith mass from its reference position and let the state variables be $x_1 := [\, p_1 \cdots p_N \,]^T$ and $x_2 := [\, \dot{p}_1 \cdots \dot{p}_N \,]^T$. For simplicity we consider unit masses and

spring constants.[1] The state-space representation is then given by (2.1) with

$$A = \begin{bmatrix} O & I \\ T & O \end{bmatrix}, \quad B_1 = B_2 = \begin{bmatrix} O \\ I \end{bmatrix},$$

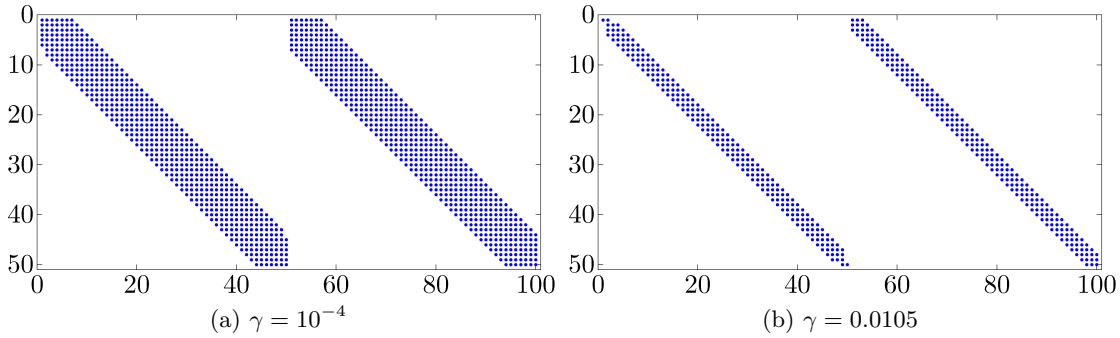
where T is an $N \times N$ symmetric tridiagonal matrix with -2 on its main diagonal and 1 on its first sub- and super-diagonal, and I and O are $N \times N$ identity and zero matrices. The state performance weight Q is the identity matrix and the control performance weight is $R = 10I$.

We use the weighted ℓ_1 norm as the sparsity-promoting penalty function. We set the weights W_{ij} using scheme (2.9) with the feedback gain being the solution $F^\star$ of (SP) at the previous value of γ. This places larger weight on smaller feedback gains; thus, they are more likely to be dropped in the next round of sparsity-promoting problem.

The optimal feedback gain at $\gamma = 0$ is computed from the solution of the algebraic Riccati equation (2.12). As γ increases, the number of nonzero sub- and super-diagonals of both optimal position gain $F_p^\star$ and optimal velocity gain $F_v^\star$ decreases; see Fig. 2.3. Eventually, both $F_p^\star$ and $F_v^\star$ become diagonal matrices. It is noteworthy that diagonals of both position and velocity feedback gains are nearly constant except for masses that are close to the boundary; see Fig. 2.4.

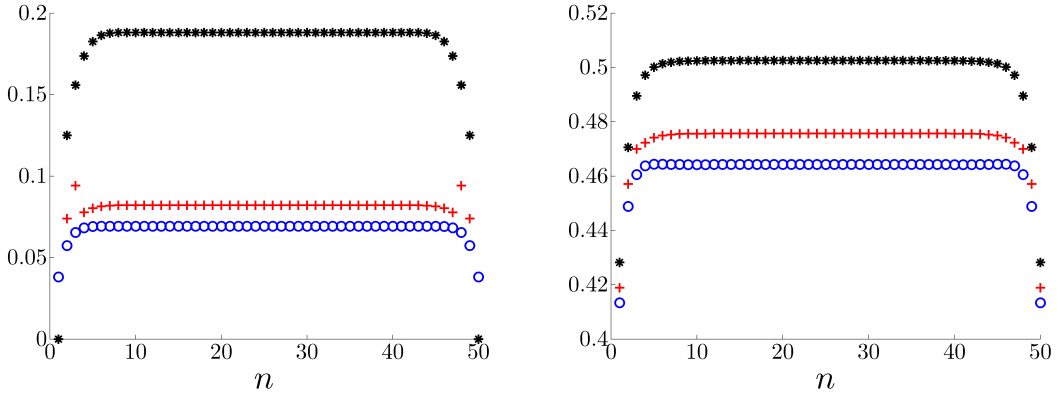
After sparsity structures of controllers are identified by solving (SP), we fix sparsity patterns and solve the structured $\mathcal{H}_2$ problem (2.13) to obtain the optimal *structured* controllers. Comparing the sparsity level and the performance of these controllers to those of the centralized controller F_c, we see that using only a *fraction* of nonzero elements, the sparse feedback gain $F^\star$ achieves an $\mathcal{H}_2$ performance that is comparable to the performance of F_c; see Fig. 2.5. In particular, using about 2% of nonzero elements, the $\mathcal{H}_2$ performance of $F^\star$ is only about 8% worse than that of F_c; see Table 2.1.

[1] Note that our method can be used to design controllers for arbitrary values of these parameters.



(a) $\gamma = 10^{-4}$ (b) $\gamma = 0.0105$

Figure 2.3: Sparsity patterns of $F^\star = [F_p^\star \quad F_v^\star] \in \mathbb{R}^{50\times 100}$ for the mass-spring system obtained using weighted ℓ_1 norm to promote sparsity. As γ increases, the number of nonzero sub- and super-diagonals of $F_p^\star$ and $F_v^\star$ decreases.

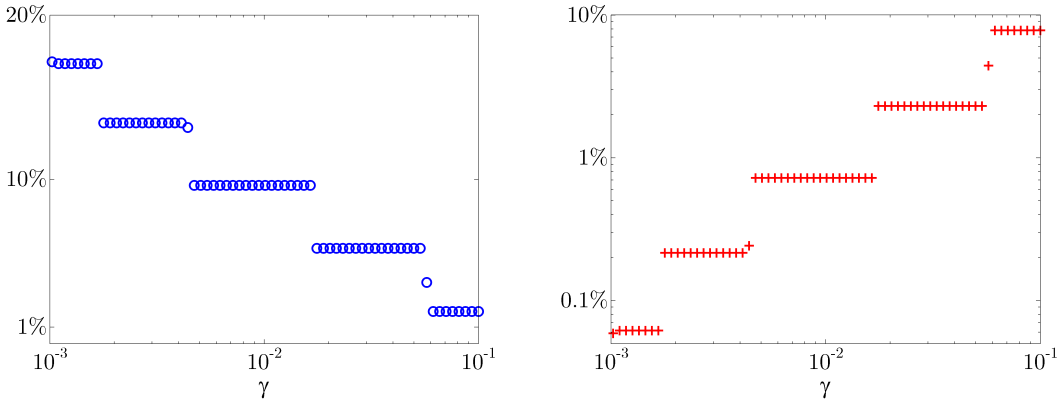


(a) (b)

Figure 2.4: (a) The diagonal of $F_p^\star$ and (b) the diagonal of $F_v^\star$ for different values of γ: 10^{-4} ($\circ$), 0.0281 ($+$), and 0.1 ($*$). The diagonals of the centralized position and velocity gains are almost identical to ($\circ$) for $\gamma = 10^{-4}$.

γ	0.01	0.04	0.10
$\mathbf{card}\,(F^\star)/\mathbf{card}\,(F_c)$	9.4%	5.8%	2.0%
$(J(F^\star) - J(F_c))/J(F_c)$	0.8%	2.3%	7.8%

Table 2.1: Sparsity vs. performance for mass-spring system. Using 2% of nonzero elements, the $\mathcal{H}_2$ performance of $F^\star$ is only 7.8% worse than the performance of the centralized gain F_c.



(a) **card** $(F^\star)$/**card** (F_c) (b) $(J(F^\star) - J(F_c))/J(F_c)$

Figure 2.5: (a) The sparsity level and (b) the performance loss of $F^\star$ compared to the centralized gain F_c.

2.4.2 A random network with unstable dynamics

Let $N = 100$ nodes be randomly distributed with a uniform distribution in a square region of 10×10 units. Each node is a linear system coupled with other nodes through the dynamics [13]

$$\dot{x}_i = A_{ii}\, x_i + \sum_{j \neq i}^{N} A_{ij}\, x_j + B_{1ii}\, d_i + B_{2ii}\, u_i, \qquad i = 1, \ldots, N.$$

Here, $(\cdot)_{ij}$ denotes the ijth block of a matrix and

$$A_{ii} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad B_{1ii} = B_{2ii} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad A_{ij} = \frac{1}{\mathrm{e}^{\alpha(i,j)}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{for} \quad i \neq j.$$

The coupling between two systems i and j is determined by the Euclidean distance $\alpha(i, j)$ between them. The performance weights Q and R are set to identity matrices.

We use the weighted ℓ_1 norm as the penalty function with the weights determined by (2.9). As γ increases, the communication architecture of distributed controllers

γ	12.6	26.8	68.7
$\mathbf{card}\,(F^\star)/\mathbf{card}\,(F_c)$	8.3%	4.9%	2.4%
$(J(F^\star)\,-\,J(F_c))/J(F_c)$	27.8%	43.3%	55.6%

Table 2.2: Sparsity vs. performance for the spatially distributed example. Using 8.3% of nonzero elements, $\mathcal{H}_2$ performance of $F^\star$ is only 27.8% worse than performance of the centralized gain F_c.

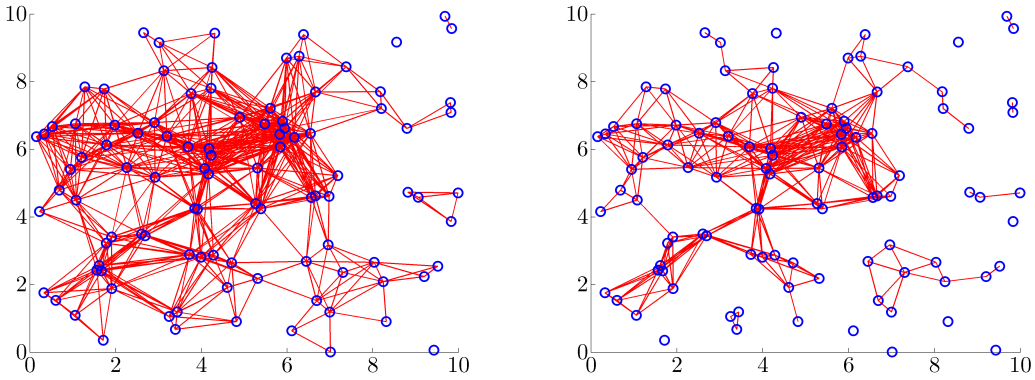
becomes sparser. Furthermore, the underlying communication graphs gradually attain a *localized* communication architecture; see Fig. 2.6. Note that, using about 8% of nonzero elements of F_c, $\mathcal{H}_2$ performance of $F^\star$ is only about 28% worse than performance of the centralized gain F_c; see Table 2.2. Figure 2.7 shows the optimal trade-off curve between the $\mathcal{H}_2$ performance and the feedback gain sparsity.

We note that the truncation of the centralized controller could result in a non-stabilizing feedback matrix. In contrast, our approach *gradually* modifies the feedback gain and increases the number of zero elements. Such an approach plays an important role in preserving the closed-loop stability. For example, keeping the largest 36.9% entries of F_c and setting the remaining entries of F_c to zero yields a non-stabilizing feedback gain. In comparison, using 35.6% nonzero elements of F_c the optimal feedback gain achieves $\mathcal{H}_2$ performance within 1.5% worse than that of F_c.

2.4.3 Block sparsity: An example from bio-chemical reaction

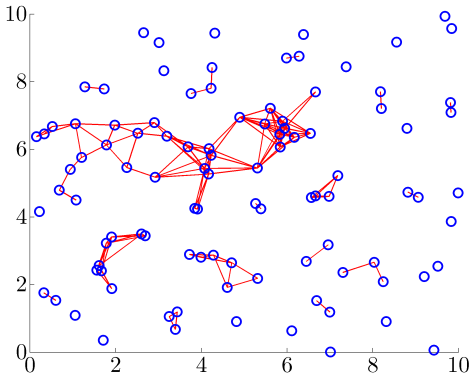
Consider a network of $N = 5$ systems coupled through the following dynamics

$$\dot{x}_i \,=\, A_{ii}\,x_i \,-\, \frac{1}{2}\sum_{j=1}^{N}(i-j)\,(x_i \,-\, x_j) \,+\, B_{1ii}\,d_i \,+\, B_{2ii}\,u_i,$$



(a) $\gamma = 12.6$

(b) $\gamma = 26.8$



(c) $\gamma = 68.7$

Figure 2.6: The localized communication graphs of distributed controllers obtained by solving (SP) for different values of γ. The communication structure becomes sparser as γ increases. Note that the communication graph does not have to be connected since the subsystems are (i) dynamically coupled to each other and (ii) allowed to measure their own states.

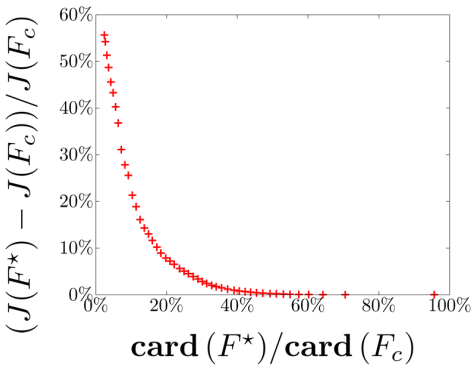


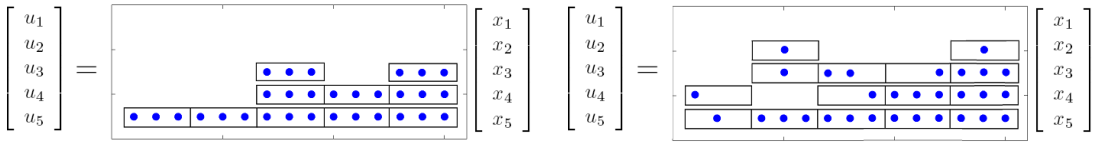
Figure 2.7: The optimal trade-off curve between the $\mathcal{H}_2$ performance loss and the sparsity level of $F^\star$ compared to the centralized gain F_c for the spatially distributed example.

where

$$A_{ii} = \begin{bmatrix} -1 & 0 & -3 \\ 3 & -1 & 0 \\ 0 & 3 & -1 \end{bmatrix}, \quad B_{1ii} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad B_{2ii} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}.$$

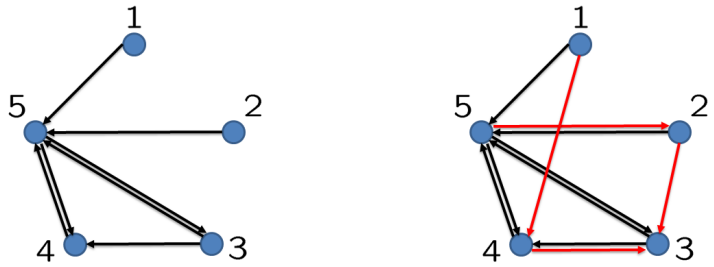
The performance weights Q and R are set to identity matrices. Systems of this form arise in bio-chemical reactions with a cyclic negative feedback [49].

We use the weighted sum of Frobenius norms as the sparsity-promoting penalty function and we set the weights W_{ij} to be inversely proportional to the Frobenius norm of the solution $F^\star_{ij}$ to (SP) at the previous value of γ, i.e., $W_{ij} = 1/(\|F^\star_{ij}\|_F + \varepsilon)$ with $\varepsilon = 10^{-3}$. As γ increases, the number of *nonzero blocks* in the feedback gain F decreases. Figure 2.8 shows sparsity patterns of feedback gains resulting from solving (SP) with sparse and block sparse penalty functions. Setting γ to values that yield the same number of nonzero elements in these feedback gains results in the block sparse feedback gain with a smaller number of nonzero blocks. In particular, the first two rows of the block sparse feedback gain in Fig. 2.8a are identically equal to zero (indicated by blank space). This means that the subsystems 1 and 2 do not need to be actuated. Thus, the



(a) (b)

Figure 2.8: The sparse feedback gains obtained by solving (SP) (a) using the weighted sum of Frobenius norms with $\gamma = 3.6$ and (b) using the weighted ℓ_1 norm (2.8) with $\gamma = 1.3$. Here, $F \in \mathbb{R}^{5\times 15}$ is partitioned into 25 blocks $F_{ij} \in \mathbb{R}^{1\times 3}$. Both feedback gains have the same number of nonzero elements (indicated by dots) and close $\mathcal{H}_2$ performance (less than 1% difference), but different number of nonzero blocks (indicated by boxes).



(a) (b)

Figure 2.9: Communication graphs of (a) the block sparse feedback gain in Fig. 2.8a and (b) the sparse feedback gain in Fig. 2.8b (red color highlights the additional links). An arrow pointing from node i to node j indicates that node i uses state measurement from node j.

communication graph determined by the block sparse feedback gain has fewer links; cf. Figs. 2.9a and 2.9b.

Chapter 3

Identification of sparsity patterns via alternating direction method of multipliers

In this chapter, we demonstrate that the alternating direction method of multipliers (ADMM) provides an effective tool for sparsity-promoting optimal control. This method alternates between promoting the sparsity of the feedback gain matrix and optimizing the performance of the system. The advantage of this alternating mechanism is threefold. First, it provides a flexible framework for incorporating different penalty functions introduced in Section 2.2. Second, it allows us to exploit the separable structure of sparsity-promoting penalty functions and to decompose optimization problems into sub-problems that can be solved analytically. Finally, it facilitates the use of a simple descent algorithm for the $\mathcal{H}_2$ optimization, in which the descent direction can be formed by solving two Lyapunov equations and one Sylvester equation.

This chapter is organized as follows. In Section 3.1, we introduce the ADMM algorithm and discuss the advantage of its alternating mechanism in sparsity-promoting

optimal control. In Section 3.2, by exploiting the separable structure of penalty functions, we decompose the corresponding optimization problem into sub-problems and provide the analytical solutions. In Section 3.3, we develop an efficient descent method for the optimization of the $\mathcal{H}_2$ performance. Finally, in Section 3.4, we solve the structured $\mathcal{H}_2$ problem using Newton's method in conjunction with the conjugate gradient scheme.

3.1 Alternating direction method of multipliers

Recall the sparsity-promoting optimal control problem

$$\underset{F}{\text{minimize}} \;\; J(F) \,+\, \gamma\, g(F) \tag{3.1}$$

where J is the closed-loop $\mathcal{H}_2$ norm (2.5) and g is a sparsity-promoting penalty function in Section 2.2. This problem is equivalent to the following constrained optimization problem

$$\begin{array}{ll} \underset{F,\,G}{\text{minimize}} & J(F) \,+\, \gamma\, g(G) \\ \text{subject to} & F \,-\, G \,=\, 0. \end{array} \tag{3.2}$$

It might appear that we have complicated the problem by introducing an additional variable G and an additional constraint $F - G = 0$. By doing this, however, we have in fact simplified problem (3.1) by decoupling the objective function into two parts that depend on two different variables. This decoupling allows us to exploit the separable structure of g and the differentiability of J in the ADMM algorithm introduced next.

We begin by forming the augmented Lagrangian associated with the constrained problem (3.2)

$$\mathcal{L}_\rho(F, G, \Lambda) \;\; = \;\; J(F) \,+\, \gamma\, g(G) \,+\, \text{trace}\left(\Lambda^T (F \,-\, G)\right) \,+\, \frac{\rho}{2}\, \|F \,-\, G\|_F^2,$$

where Λ is the dual variable (i.e., the Lagrange multiplier), ρ is a positive scalar, and $\|\cdot\|_F$ is the Frobenius norm. In order to find a minimizer of the constrained problem (3.2), the ADMM algorithm uses a sequence of iterations

$$F^{k+1} := \arg\min_{F} \mathcal{L}_\rho(F, G^k, \Lambda^k) \tag{3.3a}$$

$$G^{k+1} := \arg\min_{G} \mathcal{L}_\rho(F^{k+1}, G, \Lambda^k) \tag{3.3b}$$

$$\Lambda^{k+1} := \Lambda^k + \rho(F^{k+1} - G^{k+1}) \tag{3.3c}$$

until the primal residue $\|F^{k+1} - G^{k+1}\|_F$ and the dual residue $\|G^{k+1} - G^k\|_F$ are sufficiently small. In contrast to the *method of multipliers* explained in [50], in which F and G are *minimized jointly*,

$$(F^{k+1}, G^{k+1}) := \arg\min_{F,\,G} \mathcal{L}_\rho(F, G, \Lambda^k),$$

ADMM consists of an F-minimization step (3.3a), a G-minimization step (3.3b), and a dual variable Λ update step (3.3c). Thus, the optimal F and G are determined in an alternating fashion, which motivates the name *alternating direction*. Note that the dual variable update (3.3c) uses a step-size equal to ρ, which guarantees that one of the dual feasibility conditions is satisfied in each ADMM iteration; see [50, Section 3.3].

ADMM brings two major benefits to the sparsity-promoting optimal control problem (3.1):

- *Separability of g.* The penalty function g is *separable* with respect to the *individual* elements of the matrix. In contrast, the closed-loop $\mathcal{H}_2$ norm cannot be decomposed into componentwise functions of the feedback gain. By separating g and J in the minimization of the augmented Lagrangian $\mathcal{L}_\rho$, we can decompose G-minimization problem (3.3b) into sub-problems that only involve *scalar* variables.

This allows us to determine *analytically* the solution of (3.3b).

- *Differentiability of J.* The closed-loop $\mathcal{H}_2$ norm J is a *differentiable* function of the feedback gain matrix. This is in sharp contrast to g which is a *non-differentiable* function. By separating g and J in the minimization of the augmented Lagrangian $\mathcal{L}_\rho$, we can utilize descent algorithms that rely on the differentiability of J to solve the F-minimization problem (3.3a).

3.2 Separable solution to G-minimization problem

The completion of squares with respect to G in the augmented Lagrangian $\mathcal{L}_\rho$ can be used to show that (3.3b) is equivalent to

$$\underset{G}{\text{minimize}} \;\; \phi(G) \;=\; \gamma\, g(G) \;+\; \frac{\rho}{2}\,\|G \;-\; V^k\|_F^2 \tag{3.4}$$

where $V^k = (1/\rho)\Lambda^k + F^{k+1}$. To simplify notation, we drop the superscript k in V^k throughout this section. Since both g and the square of the Frobenius norm can be written as a summation of componentwise functions of a matrix, we can decompose (3.4) into sub-problems expressed in terms of the *individual* elements of G. For example, if g is the weighted ℓ_1 norm, then

$$\phi(G) \;=\; \sum_{i,\,j} \Big(\gamma\, W_{ij}\,|G_{ij}| \;+\; \frac{\rho}{2}\,(G_{ij} \;-\; V_{ij})^2\Big).$$

This facilitates the conversion of (3.4) to minimization problems that only involve *scalar* variables G_{ij}. By doing so, the solution of (3.4) for different penalty functions g, including the weighted ℓ_1 norm, the sum-of-logs function, and the cardinality function, can be determined *analytically*.

3.2.1 Weighted ℓ_1 norm

In this case, problem (3.4) is decomposed into sub-problems

$$\underset{G_{ij}}{\text{minimize}} \;\; \phi_{ij}(G_{ij}) \;=\; \gamma\, W_{ij}|G_{ij}| \;+\; \frac{\rho}{2}\,(G_{ij} \;-\; V_{ij})^2 \tag{3.5}$$

whose unique solution is given by the *soft thresholding* operator [50, Section 4.4.3]

$$G_{ij}^{\star} \;=\; \begin{cases} \left(1 \;-\; \dfrac{a}{|V_{ij}|}\right) V_{ij}, & |V_{ij}| \;>\; a \\ 0, & |V_{ij}| \;\leq\; a \end{cases} \tag{3.6}$$

where $a = (\gamma/\rho)W_{ij}$; see Fig. 3.1a for an illustration. For a given V_{ij}, $G_{ij}^{\star}$ is obtained by moving V_{ij} towards zero with the amount $(\gamma/\rho)W_{ij}$. In particular, $G_{ij}^{\star}$ is set to zero if $|V_{ij}| \leq (\gamma/\rho)W_{ij}$, implying that a more aggressive scheme for driving $G_{ij}^{\star}$ to zero can be obtained by increasing γ and W_{ij} and by decreasing ρ.

3.2.2 Cardinality function

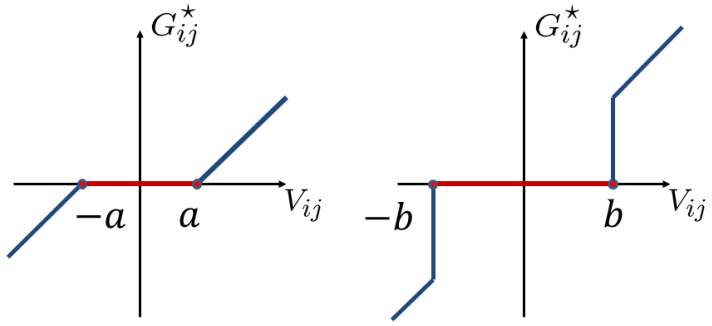
In this case, problem (3.4) is decomposed into sub-problems

$$\underset{G_{ij}}{\text{minimize}} \;\; \phi_{ij}(G_{ij}) \;=\; \gamma\, \mathbf{card}\,(G_{ij}) \;+\; \frac{\rho}{2}\,(G_{ij} \;-\; V_{ij})^2$$

whose unique solution is given by the *truncation* operator

$$G_{ij}^{\star} \;=\; \begin{cases} V_{ij}, & |V_{ij}| \;>\; b \\ 0, & |V_{ij}| \;\leq\; b \end{cases} \tag{3.7}$$

where $b = \sqrt{2\gamma/\rho}$; see Fig. 3.1b for an illustration. For a given V_{ij}, $G_{ij}^{\star}$ is set to V_{ij} if $|V_{ij}| > \sqrt{2\gamma/\rho}$ and to zero if $|V_{ij}| \leq \sqrt{2\gamma/\rho}$.



(a) $a = (\gamma/\rho)W_{ij}$ (b) $b = \sqrt{2\gamma/\rho}$

Figure 3.1: (a) The soft thresholding operator (3.6); (b) the truncation operator (3.7). The slope of the lines in both (a) and (b) is equal to one.

3.2.3 Sum-of-logs function

In this case, problem (3.4) is decomposed into sub-problems,

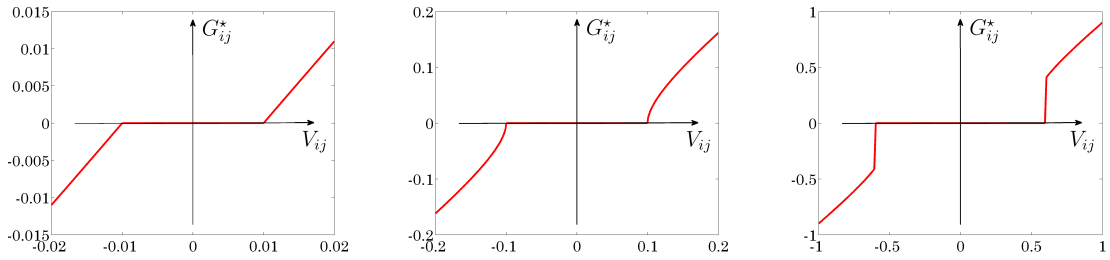
$$\underset{G_{ij}}{\text{minimize}} \;\; \phi_{ij}(G_{ij}) \;=\; \gamma \log\left(1 + \frac{1}{\varepsilon}\,|G_{ij}|\right) \;+\; \frac{\rho}{2}\,(G_{ij} - V_{ij})^2 \tag{3.8}$$

whose solution is given by

$$G_{ij}^{\star} \;=\; \begin{cases} 0, & \Delta \;\leq\; 0 \\ 0, & \Delta \;>\; 0 \text{ and } r^{+} \;\leq\; 0 \\ r^{+}V_{ij}, & \Delta \;>\; 0 \text{ and } r^{-} \;\leq\; 0 \text{ and } 0 \;<\; r^{+} \;\leq\; 1 \\ G^{0}, & \Delta \;>\; 0 \text{ and } 0 \;\leq\; r^{\pm} \;\leq\; 1 \end{cases} \tag{3.9}$$

where Δ is the discriminant and $r^{\pm}$ is the solution of a quadratic equation of r (see Section 3.2.4)

$$\begin{aligned} \Delta \;&=\; (|V_{ij}| + \varepsilon)^2 - 4(\gamma/\rho) \\ r^{\pm} \;&=\; \frac{1}{2|V_{ij}|}\left(|V_{ij}| \;-\; \varepsilon \;\pm\; \sqrt{\Delta}\right) \end{aligned} \tag{3.10}$$



(a) $\gamma = 0.1$ (b) $\gamma = 1$ (c) $\gamma = 10$

Figure 3.2: The operator (3.9) with $\{\rho = 100,\ \varepsilon = 0.1\}$ for different values of γ. For $\gamma = 0.1$, (3.9) resembles the soft thresholding operator (3.6) in Fig. 3.1a; for $\gamma = 10$, it resembles the truncation operator (3.7) in Fig. 3.1b; for $\gamma = 1$, operator (3.9) bridges the difference between the soft thresholding and truncation operators.

and

$$G^0 \ := \ \arg\min \{\phi_{ij}(r^+ V_{ij}),\ \phi_{ij}(0)\}.$$

For fixed ρ and ε, (3.9) is determined by the value of γ; see Figs. 3.2a-3.2c. For small γ, (3.9) resembles the soft thresholding operator (cf. Figs. 3.2a and 3.1a) and for large γ, it resembles the truncation operator (cf. Figs. 3.2c and 3.1b). In other words, (3.9) can be viewed as an intermediate step between the soft thresholding and the truncation operators. The detailed derivation of (3.9)-(3.10) is given in Section 3.2.4.

Remark 2. *In block sparse design, g is determined by*

$$\sum_{i,j} W_{ij}\|G_{ij}\|_F, \quad \sum_{i,j} \mathbf{card}\left(\|G_{ij}\|_F\right), \quad \sum_{i,j} \log\left(1 + \frac{\|G_{ij}\|_F}{\varepsilon}\right),$$

and the minimizers of (3.4) are obtained by replacing the absolute value of V_{ij} in (3.6), (3.7), and (3.10), respectively, with the Frobenius norm $\|\cdot\|_F$ of the corresponding block submatrix V_{ij}.

3.2.4 Derivation of (3.9)-(3.10)

The first step is to show that the minimizer of ϕ_{ij} in (3.8) can be written as

$$G_{ij}^{\star} \;=\; rV_{ij}, \qquad r \in [0,1].$$

To see this, note that when $V_{ij} > 0$, $G_{ij}^{\star}$ belongs to the interval $[0, V_{ij}]$ since both the logarithmic function $\log(1 + |G_{ij}|/\varepsilon)$ and the quadratic function $(G_{ij} - V_{ij})^2$ are monotonically increasing for $G_{ij} \geq V_{ij}$ and monotonically decreasing for $G_{ij} \leq 0$. A similar arguments shows that $G_{ij}^{\star}$ belongs to $[V_{ij}, 0]$ when $V_{ij} < 0$. Thus, minimizing $\phi_{ij}(G_{ij})$ is equivalent to minimizing

$$\phi_{ij}(r) \;=\; \gamma \log\left(1 + \frac{|V_{ij}|r}{\varepsilon}\right) \;+\; \frac{\rho}{2}\, V_{ij}^2 (r-1)^2$$

subject to the constraint $0 \leq r \leq 1$. Thus, we have converted a nondifferentiable unconstrained problem to a differentiable but constrained one.

We will now examine the sign of $\partial\phi_{ij}/\partial r$ for $r \in [0,1]$. Setting $\partial\phi_{ij}/\partial r = 0$ yields a quadratic equation (for $V_{ij} \neq 0$)

$$|V_{ij}|r^2 \;+\; (\varepsilon - |V_{ij}|)r \;+\; \frac{\gamma}{\rho|V_{ij}|} \;-\; \varepsilon \;=\; 0. \tag{3.11}$$

If the discriminant $\Delta \leq 0$, then $\partial\phi_{ij}/\partial r \geq 0$ and ϕ_{ij} is monotonically nondecreasing for $r \in [0,1]$; thus, the minimizer is $r^{\star} = 0$. Let $\Delta > 0$ and let $r^{\pm}$ be the solutions to the quadratic equation (3.11). Then

$$r^{\pm} \;=\; \frac{1}{2|V_{ij}|}\left(|V_{ij}| - \varepsilon \pm \sqrt{(|V_{ij}| + \varepsilon)^2 - 4(\gamma/\rho)}\right).$$

Note that $r^{\pm} \leq 1$. This can be verified from

$$\begin{aligned} (|V_{ij}| + \varepsilon)^2 - 4(\gamma/\rho) &\leq (|V_{ij}| + \varepsilon)^2 \\ |V_{ij}| - \varepsilon \pm \sqrt{(|V_{ij}| + \varepsilon)^2 - 4(\gamma/\rho)} &\leq 2|V_{ij}| \\ r^{\pm} &\leq 1. \end{aligned}$$

Then the minimum of ϕ_{ij} can be determined by checking the sign of $\partial\phi_{ij}/\partial r$ for $r \in [0, 1]$.

1. If $r^{\pm} \leq 0$, then $\partial\phi_{ij}/\partial r > 0$ for $r \in [0, 1]$. Thus, ϕ_{ij} is monotonically increase over $[0, 1]$ and the minimizer is $r^\star = 0$.

2. If $r^- \leq 0$ and $0 < r^+ \leq 1$, then $\partial\phi_{ij}/\partial r \leq 0$ for $r \in [0, r^+]$ and $\partial\phi_{ij}/\partial r > 0$ for $r \in (r^+, 1]$. In other words, ϕ_{ij} is monotonically nonincreasing over $[0, r^+]$ and monotonically increasing over $(r^+, 1]$. Thus, the minimizer of ϕ_{ij} for $r \in [0, 1]$ is $r^\star = r^+$.

3. Finally, if $0 \leq r^{\pm} \leq 1$, then $\partial\phi_{ij}/\partial r \geq 0$ for $r \in [0, r^-)$, $\partial\phi_{ij}/\partial r \leq 0$ for $r \in [r^-, r^+)$, and $\partial\phi_{ij}/\partial r \geq 0$ for $r \in [r^-, 1]$. Therefore, r^- is a local maximizer and r^+ is a local minimizer. Thus, the candidates for $r^\star$ are either 0 or r^+.

3.3 Anderson-Moore method for F-minimization problem

Anderson-Moore method is an iteratively descent algorithm originally proposed in [51] for unstructured output feedback design. Its advantage lies in its fast convergence (compared to the gradient descent method) and in its simple implementation (compared to second-order method such as Newton's method) [52, 53]. When applied to the F-minimization problem (3.3a), this method alternates between solving two Lyapunov equations and one Sylvester equation in each iteration.

By completing the squares with respect to F in the augmented Lagrangian $\mathcal{L}_\rho$, we

obtain the following equivalent problem to (3.3a)

$$\underset{F}{\text{minimize}} \;\; \varphi(F) \;=\; J(F) \;+\; \frac{\rho}{2}\,\|F \;-\; U^k\|_F^2 \tag{3.12}$$

where $U^k = G^k - (1/\rho)\Lambda^k$. Setting

$$\nabla\varphi \;:=\; \nabla J \;+\; \rho(F - U^k) \;=\; 0$$

yields the necessary conditions for optimality

$$2\left(RF \;-\; B_2^T P\right) L \;+\; \rho(F \;-\; U^k) \;=\; 0 \tag{NC-F}$$

$$(A \;-\; B_2F)\,L \;+\; L\,(A \;-\; B_2F)^T \;=\; -\,B_1B_1^T \tag{NC-L}$$

$$(A \;-\; B_2F)^T\,P \;+\; P\,(A \;-\; B_2F) \;=\; -\,(Q \;+\; F^T RF). \tag{NC-P}$$

Note that (NC-F), (NC-L), and (NC-P) are coupled matrix equations for F, L, and P. Here, the gradient of J is given by

$$\nabla J(F) \;=\; 2\left(RF \;-\; B_2^T P\right) L,$$

which can be obtained by expanding $J(F + \tilde{F})$ around F for small $\tilde{F}$ and collecting linear terms in $\tilde{F}$.

Starting with a stabilizing feedback F, Anderson-Moore method solves the Lyapunov equations (NC-L) and (NC-P), and then solves the Sylvester equation (NC-F) to obtain a new feedback gain $\bar{F}$. In other words, it alternates between solving (NC-L) and (NC-P) for L and P with F being fixed and solving (NC-F) for F with L and P being fixed. In Proposition 1, we show that the difference between two consecutive steps $\tilde{F} = \bar{F} - F$ forms a *descent direction* of $\varphi(F)$. As a consequence, standard step-size rules can be

employed to determine s in $F + s\tilde{F}$ to guarantee the convergence to a stationary point of φ. We summarize Anderson-Moore method in Algorithm 1.

Algorithm 1 Anderson-Moore method for F-minimization problem (3.3a) for the kth ADMM iteration

Require: The initial condition F^0. For $k \geq 1$, use the solution of (3.3a) from the previous ADMM iteration; for $k = 0$, use the minimizer of (3.2) from the previous γ value.

1: **for** $i = 0, 1, \ldots$ **do**
2: solve the Lyapunov equations (NC-L) and (NC-P) with $F = F^i$ to obtain L^i and P^i;
3: solve the Sylvester equation (NC-F) with $L = L^i$ and $P = P^i$ to obtain $\bar{F}^i$;
4: form $\tilde{F}^i = \bar{F}^i - F^i$ and update $F^{i+1} = F^i + s^i\tilde{F}^i$, with s^i determined by the Armijo rule;
5: **until** $\|\nabla\varphi(F^i)\|_F < \epsilon$.
6: **end for**
7: **Armijo rule** [54, Section 1.2] for the step-size s^i:
8: let $s^i = 1$, $\alpha, \beta \in (0, 1)$
9: **repeat**
10: $s^i = \beta s^i$
11: **until**

$$\varphi(F^i + s^i\tilde{F}^i) < \varphi(F^i) + \alpha\, s^i \langle \nabla\varphi(F^i), \tilde{F}^i\rangle.$$

Proposition 1. *The difference $\tilde{F} = \bar{F} - F$ in Anderson-Moore method (see step 4 in Algorithm 1) forms a descent direction of the objective function φ in* (3.12)*, i.e., $\langle \nabla\varphi(F), \tilde{F}\rangle < 0$. Moreover, $\langle \nabla\varphi(F), \tilde{F}\rangle = 0$ if and only if F is a stationary point of φ, i.e., $\nabla\varphi(F) = 0$.*

Proof. Since $\bar{F}$ satisfies (NC-F), i.e., $\nabla\varphi(F + \tilde{F}) = 0$, substituting $\bar{F} = F + \tilde{F}$ into (NC-F) yields the condition for descent direction

$$2R\tilde{F}L + \rho\tilde{F} + \nabla\varphi(F) = 0, \tag{3.13}$$

where

$$\nabla\varphi(F) = 2(RF - B_2^T P)L + \rho(F - U^k).$$

Computing the inner product between $\nabla\varphi(F)$ and $\tilde{F}$ yields

$$\langle \nabla\varphi(F), \tilde{F} \rangle \;=\; -\,2\,\langle R\tilde{F}L, \tilde{F}\rangle \;-\; \rho\,\langle \tilde{F}, \tilde{F}\rangle \;\leq\; 0, \tag{3.14}$$

where $-\langle R\tilde{F}L, \tilde{F}\rangle \leq 0$ follows from the positive definiteness of R and L. It remains to show that $\langle \nabla\varphi(F), \tilde{F}\rangle = 0$ is a necessary and sufficient condition for $\nabla\varphi(F) = 0$. The necessary condition is immediate and the sufficient condition follows from the fact that the equality in (3.14) implies $\tilde{F} = 0$, which in conjunction with (3.13) yields $\nabla\varphi(F) = 0$. This completes the proof. □

3.4 Polishing step: Solving structured $\mathcal{H}_2$ problem

We next turn to the $\mathcal{H}_2$ problem subject to structural constraints on the feedback matrix

$$\begin{aligned} \underset{F}{\text{minimize}} \quad & J(F) \\ \text{subject to} \quad & F \in \mathcal{S}. \end{aligned} \tag{3.15}$$

Here, we fix the sparsity patterns $\mathcal{S}$ identified using ADMM and then solve (3.15) to obtain the optimal feedback gain that belongs to the subspace $\mathcal{S}$. This *polishing* step, which is commonly used in cardinality optimization [43, Section 6.3.2], can improve the performance of sparse feedback gains resulting from the ADMM algorithm.

Since J is a smooth function that increases to infinity as one approaches the boundary of the set of stabilizing gains, the decreasing sequence of $\{\varphi(F^i)\}$ in Anderson-Moore method ensures that $\{F^i\}$ are stabilizing gains. Thus, the sparse feedback gains obtained in the ADMM algorithm are stabilizing. This feature of ADMM facilitates the use of Newton's method to solve the structured $\mathcal{H}_2$ problem (3.15).

3.4.1 Newton's method

Newton's method is an iterative *descent* algorithm for finding local minima of optimization problems [43]. Given an initial gain F^0, a decreasing sequence of the objective function $\{J(F^i)\}$ is generated by updating F according to $F^{i+1} = F^i + s^i \tilde{F}^i$. Here, $\tilde{F}^i$ is the Newton direction and s^i is the step-size. The Newton direction is the minimizer of the second-order approximation of the objective function [43]

$$J(F + \tilde{F}) \approx J(F) + \langle \nabla J(F), \tilde{F} \rangle + \frac{1}{2} \langle H(F, \tilde{F}), \tilde{F} \rangle, \tag{3.16}$$

where the matrix $H(F, \tilde{F})$ depends linearly on $\tilde{F}$.

To find ∇J for problem (3.15), we expand $J(F + \tilde{F})$ around F for the variation $\tilde{F} \in \mathcal{S}$, and collect linear terms in $\tilde{F}$ to get

$$\nabla J(F) = (2(RF - B_2^T P)L) \circ I_{\mathcal{S}}.$$

Here, L and P are the solutions of Lyapunov equations (NC-L) and (NC-P). Symbol $\circ$ denotes the elementwise multiplication of two matrices and the matrix $I_{\mathcal{S}}$ denotes the *structural identity* of subspace $\mathcal{S}$ under elementwise multiplication. Specifically, the ijth entry of $I_{\mathcal{S}}$ is given by

$$I_{\mathcal{S}ij} = \begin{cases} 1, & \text{if } F_{ij} \text{ is a free variable} \\ 0, & \text{if } F_{ij} = 0 \text{ is required} \end{cases} \tag{3.17}$$

implying that

$$F \circ I_{\mathcal{S}} = F \quad \text{for} \quad F \in \mathcal{S}.$$

To find $H(F, \tilde{F})$ in (3.16), we expand $\nabla J(F + \tilde{F})$ around F for the variation $\tilde{F} \in \mathcal{S}$,

and collect linear terms in $\tilde{F}$,

$$H(F,\tilde{F}) \;=\; \Big(2(R\tilde{F} - B_2^T\tilde{P})L \;+\; 2(RF - B_2^TP)\tilde{L}\Big) \circ I_{\mathcal{S}} \tag{3.18a}$$

where $\tilde{L}$ and $\tilde{P}$ are solutions of the Lyapunov equations

$$(A - B_2F)\tilde{L} \;+\; \tilde{L}(A - B_2F)^T \;\;=\;\; B_2\tilde{F}L \;+\; (B_2\tilde{F}L)^T, \tag{3.18b}$$

$$(A - B_2F)^T\tilde{P} \;+\; \tilde{P}(A - B_2F) \;\;=\;\; (PB_2 - F^TR)\tilde{F} \;+\; \tilde{F}^T(B_2^TP - RF). \tag{3.18c}$$

Note that H is a linear function of $\tilde{F}$ because both $\tilde{L}$ and $\tilde{P}$ linearly depend on $\tilde{F}$, and that both ∇J and H belong to $\mathcal{S}$ by construction.

3.4.2 Conjugate gradient method to compute Newton direction

Before we proceed, we set the following convention: henceforth, whenever we write $H(\tilde{F})$ it means that $\tilde{L}$ and $\tilde{P}$ have been expressed in terms of $\tilde{F}$ using (3.18b), (3.18c) and substituted into (3.18a), so that $H(\tilde{F})$ only depends on $\tilde{F}$.

The conjugate gradient (CG) method is an iterative scheme for minimization of (3.16), or equivalently, for minimization of the following quadratic form

$$\Phi(\tilde{F}) \;:=\; \frac{1}{2}\,\langle H(\tilde{F}),\tilde{F}\rangle \;+\; \langle \nabla J,\tilde{F}\rangle.$$

Starting with $\tilde{F}^0 = 0$, the CG method minimizes $\Phi(\tilde{F})$ along the *conjugate* directions $\Delta^k \in \mathcal{S}$ that are mutually orthogonal with respect to the inner product $\langle H(\cdot),\cdot\rangle$, that is, $\langle H(\Delta^i),\Delta^j\rangle \;=\; 0$ for $i \neq j$. As a consequence, the number of CG iterations to reach the minimizer of $\Phi(\tilde{F})$ is not larger than the number of unknown (nonzero) elements in $\tilde{F}$ [55, Chapter 5]. Furthermore, the direction Δ^k is generated from a linear combination

Algorithm 2 Computing Newton direction by conjugate gradient method

Require: the gradient direction ∇J
1: set $\tilde{F}^0 = 0$, $\Pi^0 = \nabla J$, $\Delta^0 = -\nabla J$,
2: **for** $k = 0$ to q **do**
3: (1) negative curvature test:
4: **if** $\langle H(\Delta^0), \Delta^0 \rangle \leq 0$ **then**
5: **return** $\tilde{F}^1 = -\nabla J$ as the approximate Newton direction;
6: **else if** $\langle H(\Delta^k), \Delta^k \rangle \leq 0$ and $k > 0$ **then**
7: **return** $\tilde{F}^k$ as the approximate Newton direction;
8: **end if**
9: (2) $\alpha^k = -\langle \Pi^k, \Delta^k \rangle / \langle H(\Delta^k), \Delta^k \rangle$;
10: (3) $\tilde{F}^{k+1} = \tilde{F}^k + \alpha^k \Delta^k$, $\Pi^{k+1} = \Pi^k + \alpha^k H(\Delta^k)$;
11: (4) $\beta^{k+1} = \langle H(\Delta^k), \Pi^{k+1} \rangle / \langle H(\Delta^k), \Delta^k \rangle$;
12: (5) $\Delta^{k+1} = -\Pi^{k+1} + \beta^{k+1} \Delta^k$;
13: **until:** The stopping criterion $\|\Pi^{k+1}\|_F < \epsilon$ is reached.
14: **end for**

of the previous direction Δ^{k-1} and the gradient of $\Phi(\tilde{F}^k)$,

$$\Pi^k := \nabla \Phi(\tilde{F}^k) = H(\tilde{F}^k) + \nabla J.$$

Thus, there is no need to save the entire history of conjugate directions $\{\Delta^k\}$ to obtain mutual orthogonality. This feature is appealing for large problems, as it saves on storage and computation.

We now summarize the CG method for computing the Newton direction in Algorithm 2 and then briefly elaborate on the purpose of each step of the algorithm. Note that the number of CG steps is no greater than the number of nonzero entries q enforced by the structure $\mathcal{S}$. Consequently, the CG method is computationally efficient when structure $\mathcal{S}$ is sparse.

Step (1), the negative curvature test checks whether given a direction Δ^k, we have $\langle H(\Delta^k), \Delta^k \rangle \leq 0$. If this inequality holds then $\Phi(\tilde{F})$ is not a convex function and the minimization of $\Phi(\tilde{F})$ is not well-defined, as the objective function can be driven

to minus infinity in the direction Δ^k. In that case the algorithm is terminated while ensuring that the latest update $\tilde{F}^k$ is still a descent direction of J.

Step (2) chooses α^k that minimizes $\Phi(\tilde{F}^k + \alpha\,\Delta^k)$ with respect to the scalar α (where $\tilde{F}^k$ and Δ^k are fixed). In other words, α^k is the exact, i.e., the optimal step-size in the direction Δ^k of minimizing $\Phi(\tilde{F}^k + \alpha\Delta^k)$.

Step (3) updates $\tilde{F}^{k+1}$ and the corresponding gradient Π^{k+1} for $\Phi(\tilde{F}^{k+1})$

$$\Pi^{k+1} \;=\; H(\tilde{F}^k + \alpha^k\,\Delta^k) \;+\; \nabla J \;=\; \Pi^k \;+\; \alpha^k\,H(\Delta^k).$$

Note that $H(\Delta^k)$ is already computed in Step (2).

Steps (4) and (5) generate a new direction Δ^{k+1} that is orthogonal to Δ^k, i.e., $\langle H(\Delta^k), \Delta^{k+1}\rangle = 0$. This is done by writing the new direction Δ^{k+1} as a linear combination of the negative gradient direction $-\Pi^{k+1}$ and the previous direction Δ^k, that is, $\Delta^{k+1} = -\Pi^{k+1} + \beta^{k+1}\,\Delta^k$, and then finding β^{k+1} such that $\langle H(\Delta^{k+1}), \Delta^k\rangle = 0$. In fact, it can be shown [55] that, as a result, we have $\langle H(\Delta^i), \Delta^{k+1}\rangle = 0$ and $\langle \Pi^i, \Pi^{k+1}\rangle = 0$ for $i = 0, 1, \ldots, k$. Finally, the Algorithm 2 terminates either when the gradient Π^k is sufficiently small or when the negative curvature test is satisfied.

Chapter 4

Augmented Lagrangian approach to structured feedback design

So far, we have identified sparsity patterns of the feedback gain matrix by solving sparsity-promoting optimal control problem. In this chapter, we consider the structured feedback design with *a priori* assigned sparsity patterns. We assume that the set of stabilizing structured gains is not empty and we are interested in finding the structured gain that minimize the $\mathcal{H}_2$ norm of the system.

We employ the augmented Lagrangian method that does not require knowledge of a stabilizing structured gain to initialize the algorithm. Instead, a sequence of unstructured minimization problems is solved and the resulting minimizers converge to the optimal structured gain. Furthermore, we use sensitivity interpretation of the Lagrange multiplier to identify effective relaxations of sparsity constraints for improving $\mathcal{H}_2$ performance.

This chapter is organized as follows. In Section 4.1, we recall the structured state feedback problem and discuss the smoothness property of the $\mathcal{H}_2$ norm. In Section 4.2,

we introduce the augmented Lagrangian approach and utilize the sensitivity interpretation of Lagrange multiplier to identify effective relaxations of sparsity patterns for performance improvement. In Section 4.3, we adapt the Anderson-Moore method for the minimization of the augmented Lagrangian. In Section 4.4, we illustrate the effectiveness of the proposed approach via two examples.

4.1 Structured feedback design

Recall the structured optimal control problem

$$\begin{aligned} &\text{minimize} && J(F) \\ &\text{subject to} && F \in \mathcal{S}, \end{aligned} \tag{4.1}$$

where J is the $\mathcal{H}_2$ norm defined in (2.5) and $\mathcal{S}$ is the prescribed sparsity structure that the feedback gain F belongs to. Our objective is to find $F \in \mathcal{S}$ that minimizes the $\mathcal{H}_2$ norm J. Note that we do not require the knowledge of a stabilizing structured feedback gain. This is in contrast to the structured $\mathcal{H}_2$ problem considered in Section 3.4 where a stabilizing gain is provided by ADMM in the structure identification step.

The closed-loop $\mathcal{H}_2$ norm of a stabilizable and detectable system increases to infinity as the least stable eigenvalue of $A_{\text{cl}} := A - B_2F$ goes towards the imaginary axis. Furthermore, J is a smooth function of F, since J is a linear function of the observability Gramian P and since P is a product of the exponential and polynomial functions of the feedback gain F. Therefore, the closed-loop $\mathcal{H}_2$ norm is a smooth function that increases to infinity as one approaches the boundary of the set of stabilizing feedback gains.

If a stabilizing $F \in \mathcal{S}$ is known, descent algorithms such as Newton's method in Section 3.4 can be employed to determine a local minimum of J. However, finding

a stabilizing structured gain is in general a challenging problem. To alleviate this difficulty, we employ the augmented Lagrangian method (also referred to as the *method of multipliers*), an algorithm closely related to ADMM. The augmented Lagrangian method minimizes a sequence of *unstructured* problems such that the minimizers of the unstructured problems converge to the minimizer of (4.1). Therefore, this method does not require a stabilizing *structured* feedback gain to initialize the optimization algorithm.

4.2 Augmented Lagrangian method

We begin by providing an algebraic characterization of the structural constraint $F \in \mathcal{S}$. Recall the structural identity $I_{\mathcal{S}}$ defined in (3.17). Let $I_{\mathcal{S}}^c := \mathbf{1} - I_{\mathcal{S}}$ denote the structural identity of the *complementary* subspace $\mathcal{S}^c$, where $\mathbf{1}$ is the matrix of all ones. Then $F \in \mathcal{S}$ is equivalent to the condition that the elementwise multiplication of F and $I_{\mathcal{S}}^c$ is zero

$$F \circ I_{\mathcal{S}}^c \;=\; 0.$$

Therefore, the structured $\mathcal{H}_2$ optimal control problem (4.1) can be rewritten as

$$\begin{aligned} \text{minimize} \quad & J(F) \;=\; \text{trace}\left(B_1^T P(F) B_1\right) \\ \text{subject to} \quad & F \circ I_{\mathcal{S}}^c \;=\; 0 \end{aligned} \tag{SH2}$$

where $P(F)$ is the solution of (2.6).

The Lagrangian function for (SH2) is given by

$$\mathcal{L}(F, V) \;=\; J(F) \;+\; \text{trace}\left(V^T (F \circ I_{\mathcal{S}}^c)\right).$$

From Lagrange duality theory [43, 54, 56], it follows that there exists a unique Lagrange

multiplier $V^\star \in \mathcal{S}^c$ such that the minimizer of $\mathcal{L}(F, V^\star)$ with respect to F is a local minimum of (SH2). The Lagrange dual approach minimizes $\mathcal{L}(F, V)$ with respect to *unstructured* F for a fixed V (the estimate of $V^\star$), and then updates V such that it converges to the Lagrange multiplier $V^\star$. Consequently, as V converges to $V^\star$, the minimizer of $\mathcal{L}(F, V)$ with respect to F converges to the minimizer of (SH2). This Lagrange dual approach is useful for convex problems [43]; for nonconvex problems, it relies on local convexity assumptions [56] that may not be satisfied in problem (SH2).

In what follows, a quadratic term is introduced to locally convexify the Lagrangian [54, 56] yielding the augmented Lagrangian for (SH2)

$$\mathcal{L}_\rho(F, V) \;=\; J(F) \;+\; \text{trace}\left(V^T(F \circ I_{\mathcal{S}}^c)\right) \;+\; \frac{\rho}{2}\,\|F \circ I_{\mathcal{S}}^c\|_F^2,$$

where the penalty weight ρ is a positive scalar and $\|\cdot\|_F$ is the Frobenius norm. Starting with an initial estimate of the Lagrange multiplier, e.g., $V^0 = 0$, the augmented Lagrangian method iterates between minimizing $\mathcal{L}_\rho(F, V^i)$ with respect to unstructured F (for a fixed V^i) and updating V as follows

$$V^{i+1} \;=\; V^i \;+\; \rho\,(F^i \circ I_{\mathcal{S}}^c),$$

where F^i is the minimizer of $\mathcal{L}_\rho(F, V^i)$. Note that, by construction, V^i belongs to the complementary subspace $\mathcal{S}^c$, that is,

$$V^i \circ I_{\mathcal{S}}^c \;=\; V^i.$$

Then the sequence $\{V^i\}$ converges to the Lagrange multiplier $V^\star$, and consequently, the sequence of the minimizers $\{F^i\}$ converges to the structured optimal feedback gain $F^\star$. We summarize this approach in Algorithm 3.

Algorithm 3 Augmented Lagrangian method for (SH2)

Require: $V^0 = 0$ and $\rho^0 > 0$.

1: **for** $i = 0, 1, \ldots,$ **do**
2: for a fixed V^i, minimize $\mathcal{L}_\rho(F, V^i)$ with respect to the unstructured F to have the minimizer F^i (see Section 4.3);
3: update $V^{i+1} = V^i + \rho^i (F^i \circ I_{\mathcal{S}}^c)$;
4: update $\rho^{i+1} = \sigma \rho^i$ with $\sigma > 1$;
5: **until**: the stopping criterion $\|F^i \circ I_{\mathcal{S}}^c\|_F < \epsilon$ is reached.
6: **end for**

The convergence rate of the augmented Lagrangian method depends on the penalty weight ρ; see [54]. In general, a large ρ results in a fast convergence rate. However, large values of ρ may introduce computational difficulty in minimizing the augmented Lagrangian. This is because the condition number of the Hessian matrix $\nabla^2 \mathcal{L}_\rho(F, V)$ becomes larger as ρ increases. It is thus recommended [54] to increase the penalty weight gradually until it reaches a certain threshold value τ. Our numerical experiments suggest that $\rho^0 \in [\,1,\, 5\,]$, $\sigma \in [\,3,\, 10\,]$, and $\tau \in [\,10^4,\, 10^6\,]$ work well in practice. Additional guidelines for choosing these parameters can be found in [54, Section 4.2].

4.2.1 Sensitivity interpretation of the Lagrange multiplier

The Lagrange multiplier provides useful information about the sensitivity of the optimal value with respect to the perturbation of constraints [43, 54, 56]. In particular, for the structured design problem, the Lagrange multiplier indicates how sensitive the optimal $\mathcal{H}_2$ norm is with respect to the change of structural constraints. We use this sensitivity interpretation to identify sparsity patterns for improving $\mathcal{H}_2$ performance.

Let $\langle \cdot, \cdot \rangle$ denote the standard inner product of matrices

$$\langle M_1, M_2 \rangle = \operatorname{trace}(M_1^T M_2).$$

It is readily verified that

$$\|F \circ I_{\mathcal{S}}^c\|_F^2 \;=\; \langle F \circ I_{\mathcal{S}}^c, F \circ I_{\mathcal{S}}^c \rangle \;=\; \langle F \circ I_{\mathcal{S}}^c, F \rangle$$

and

$$\langle V, F \circ I_{\mathcal{S}}^c \rangle \;=\; \langle V \circ I_{\mathcal{S}}^c, F \rangle \;=\; \langle V, F \rangle,$$

where we used $V \circ I_{\mathcal{S}}^c = V$. Thus, the augmented Lagrangian can be rewritten as

$$\mathcal{L}_\rho(F, V) \;=\; J(F) \;+\; \langle V, F \rangle \;+\; \frac{\rho}{2} \langle F \circ I_{\mathcal{S}}^c, F \rangle$$

and its gradient with respect to F is given by

$$\nabla \mathcal{L}_\rho(F, V) \;=\; \nabla J(F) \;+\; V \;+\; \rho\, (F \circ I_{\mathcal{S}}^c).$$

Since the minimizer $F^\star$ of $\mathcal{L}_\rho(F, V^\star)$ satisfies $\nabla \mathcal{L}_\rho(F^\star, V^\star) = 0$ and $F^\star \circ I_{\mathcal{S}}^c = 0$, we have

$$\nabla J(F^\star) \;+\; V^\star \;=\; 0.$$

Let the structural constraints $\{F_{ij} \;=\; 0,\; (i,j) \;\in\; \mathcal{S}^c\}$ be relaxed to $\{|F_{ij}| \;\leq\; w,\; (i,j) \in \mathcal{S}^c\}$ with $w > 0$, and let $\hat{F}$ be the minimizer of

$$\begin{aligned} &\text{minimize} \quad && J(F) \\ &\text{subject to} \quad && \{|F_{ij}| \leq w,\; (i,j) \in \mathcal{S}^c\}. \end{aligned} \tag{4.2}$$

Since the constraint set in (4.2) contains the constraint set in (SH2), $J(\hat{F})$ is smaller than or equal to $J(F^\star)$

$$J(\hat{F}) \;:=\; J(F^\star + \tilde{F}^\star) \;\leq\; J(F^\star), \tag{4.3}$$

where $\tilde{F}^\star$ denotes the difference between $\hat{F}$ and $F^\star$. Now, the Taylor series expansion of $J(F^\star + \tilde{F}^\star)$ around $F^\star$ yields

$$\begin{aligned} J(F^\star) \,-\, J(F^\star + \tilde{F}^\star) \,&=\, -\,\langle \nabla J(F^\star), \tilde{F}^\star \rangle \,+\, O(\|\tilde{F}^\star\|_F^2) \\ &=\, \langle V^\star, \tilde{F}^\star \rangle \,+\, O(\|\tilde{F}^\star\|_F^2). \end{aligned}$$

Furthermore,

$$\begin{aligned} \langle V^\star, \tilde{F}^\star \rangle \,&\leq\, \sum_{i,\,j} |V_{ij}^\star||\tilde{F}_{ij}^\star| \\ &=\, \sum_{(i,\,j)\,\in\,\mathcal{S}} |V_{ij}^\star||\tilde{F}_{ij}^\star| \,+\, \sum_{(i,\,j)\,\in\,\mathcal{S}^c} |V_{ij}^\star||\tilde{F}_{ij}^\star| \\ &\leq\, w \sum_{(i,\,j)\,\in\,\mathcal{S}^c} |V_{ij}^\star| \end{aligned}$$

where we have used the fact that $V_{ij}^\star = 0$ for $(i,j) \in \mathcal{S}$ and $|\tilde{F}_{ij}^\star| \leq w$ for $(i,j) \in \mathcal{S}^c$. Thus, up to the first order in $\tilde{F}^\star$, we have

$$J(F^\star) \,-\, J(F^\star + \tilde{F}^\star) \,\leq\, w \sum_{(i,j)\,\in\,\mathcal{S}^c} |V_{ij}^\star|.$$

Note that larger $|V_{ij}^\star|$ implies larger decrease in the $\mathcal{H}_2$ norm if the corresponding constraint $F_{ij} = 0$ is relaxed. This sensitivity interpretation can be utilized to identify controller architectures for performance improvement; see Section 4.4.2 for an illustrative example.

4.3 Anderson-Moore method for augmented Lagrangian minimization

Similar to the F-minimization step in Section 3.3, Anderson-Moore method can be employed for the minimization of the augmented Lagrangian. The expression for the gradient of $\mathcal{L}_\rho(F)$ is given by

$$\begin{aligned}\nabla\mathcal{L}_\rho(F) &= \nabla J(F) + V + \rho\,(F \circ I_{\mathcal{S}}^c) \\ &= 2(RF - B_2^T P)L + V + \rho\,(F \circ I_{\mathcal{S}}^c),\end{aligned}$$

where recall that L and P are the solution of the Lyapunov equations

$$(A - B_2F)L + L(A - B_2F)^T = -B_1B_1^T \tag{NC-L}$$

$$(A - B_2F)^T P + P(A - B_2F) = -(Q + F^T RF). \tag{NC-P}$$

Thus, the necessary condition for optimality of $\mathcal{L}_\rho(F)$ is given by

$$2(RF - B_2^T P)L + V + \rho\,(F \circ I_{\mathcal{S}}^c) = 0. \tag{NC-FS}$$

Solving the system of equations (NC-L), (NC-P), and (NC-FS) simultaneously for matrix variables F, L, and P is a non-trivial task. In the absence of structural constraints, setting

$$\nabla J(F) = 2(RF - B_2^T P)L = 0$$

yields the optimal unstructured feedback gain

$$F_c = R^{-1}B_2^T P$$

where the pair $(A - B_2F, B_1)$ is assumed to be controllable and therefore L is invertible. Here, P is the positive definite solution of the algebraic Riccati equation obtained by substituting F_c in (NC-P)

$$A^TP + PA + Q - PB_2R^{-1}B_2^TP = 0.$$

Starting with $F = F_c$, we can solve Lyapunov equations (NC-L) and (NC-P), and then solve (NC-FS) to obtain a new feedback gain $\bar{F}$. We can thus alternate between solving (NC-L), (NC-P) and solving (NC-FS) as described in Section 3.3.

Similar to the descent property established in Proposition 1, we show that the difference between two consecutive steps $\bar{F} - F$ is a descent direction of $\mathcal{L}_\rho$. Therefore, we can employ the Armijo rule to choose the step-size s in $F + s(\bar{F} - F)$ such that the alternating method converges to a stationary point of $\mathcal{L}_\rho$. We then update V and ρ in the augmented Lagrangian (see Section 4.2 for details), and use the minimizer of $\mathcal{L}_\rho$ to initialize another round of the alternating descent iterations. As V converges to $V^\star$, the minimizer of $\mathcal{L}_\rho$ converges to $F^\star$. Therefore, the augmented Lagrangian method traces a solution path (parameterized by V and ρ) between the unstructured optimal gain F_c and the structured optimal gain $F^\star$.

We summarize Anderson-Moore method for (SH2) in Algorithm 4. Although this algorithm is similar to Algorithm 1 with a few slight modifications, it is presented here for completeness.

Remark 3. *For a diagonal matrix R, the descent direction $\tilde{F} = \bar{F} - F$ can be computed efficiently in a row-by-row fashion. Substituting $\bar{F} = F + \tilde{F}$ in (NC-FS) yields*

$$2R\tilde{F}L + \rho\,(\tilde{F} \circ I_{\mathcal{S}}^c) + \nabla\mathcal{L}_\rho(F) = 0. \tag{4.4}$$

Algorithm 4 Anderson-Moore method to minimize augmented Lagrangian $\mathcal{L}_\rho(F, V^i)$

For $V^0 = 0$, start with the optimal unstructured feedback gain F_c.
For V^i with $i \geq 1$, start with the minimizer of $\mathcal{L}_\rho(F, V^{i-1})$.
for $k = 0, 1, \ldots,$ **do**
  solve Lyapunov equations (NC-L) and (NC-P) with fixed $F = F_k$ to obtain L_k and P_k;
  solve linear equation (NC-FS) with fixed $L = L_k$ and $P = P_k$ to obtain $\bar{F}_k$;
  update $F_{k+1} = F_k + s_k(\bar{F}_k - F_k)$ where s_k is determined by Armijo rule;
  until: the stopping criterion $\|\nabla\mathcal{L}_\rho(F_k)\|_F < \epsilon$ is reached.
end for
Armijo rule [54, Section 1.2] for step-size s_k:
Let $s_k = 1$, repeat $s_k = \beta s_k$
until

$$\mathcal{L}_\rho(F_k + s_k(\bar{F}_k - F_k)) < \mathcal{L}_\rho(F_k) + \alpha\, s_k \langle \nabla\mathcal{L}_\rho(F_k), \bar{F}_k - F_k\rangle,$$

where $\alpha, \beta \in (0, 1)$, e.g., $\alpha = 0.3$ and $\beta = 0.5$.

If R is a diagonal matrix, we can write the jth row of (4.4) as

$$\tilde{F}_j\big(2R_{jj}L + \rho\,\mathrm{diag}\,(I^c_{\mathcal{S}j})\big) + (\nabla\mathcal{L}_\rho)_j = 0 \tag{4.5}$$

where $(\cdot)_j$ denotes the jth row of a matrix and $\mathrm{diag}\,(I^c_{\mathcal{S}j})$ *is a diagonal matrix with $I^c_{\mathcal{S}j}$ on its main diagonal. Therefore, each row of $\tilde{F}$ can be computed independently by solving (4.5).*

4.4 Examples

We next demonstrate the utility of the augmented Lagrangian approach in the design of optimal structured controllers. The mass-spring system in Section 4.4.1 illustrates the efficiency of the augmented Lagrangian method, and the vehicular formation example in Section 4.4.2 illustrates the effectiveness of the Lagrange multiplier in identifying controller architectures for improving $\mathcal{H}_2$ performance.

Table 4.1: Mass-spring system with $N = 100$ masses and distributed controller using displacements of p masses from the left and p masses from the right. ALT# is the number of alternating steps and $F_c \circ I_{\mathcal{S}}$ is the projection of the centralized gain F_c onto the subspace $\mathcal{S}$.

p	ALT#	$J(F^\star)$	$J(F_c \circ I_{\mathcal{S}})$
0	92	499.9	546.5
1	83	491.2	497.2
2	71	488.0	489.6
3	70	486.8	487.6

4.4.1 Mass-spring system

We consider a mass-spring system example in Section 2.4.1. The control applied to the ith mass has access to displacement and velocity of the ith mass, and displacements of p neighboring masses on the left and p neighboring masses on the right. Thus, $I_{\mathcal{S}} = [\, S_p \;\; I \,]$ where S_p is a banded matrix with ones on p upper and p lower sub-diagonals. For $N = 100$ masses with $p = 0, 1, 2, 3$, the computational results are summarized in Table 4.1. Here, the stopping criterion for the augmented Lagrangian method is $\|F \circ I_{\mathcal{S}}^c\|_F < 10^{-6}$, and the stopping criterion for the alternating method is $\|\nabla \mathcal{L}_\rho(F)\|_F < 10^{-3}$.

We note that as the spatial spread p of the distributed controller increases (i) the improvement of $J(F^\star)$ becomes less significant; and (ii) $J(F_c \circ I_{\mathcal{S}}) \approx J(F^\star)$, i.e., near optimal performance can be achieved by the truncated optimal unstructured controller $F_c \circ I_{\mathcal{S}}$. These observations are consistent with the spatially decaying property of the optimal unstructured controller on the information from neighboring subsystems [5,13].

4.4.2 Formation of vehicles

We consider a formation of nine vehicles in a plane. The control objective is to keep constant distances between neighboring vehicles. Modeling these independently actuated

vehicles as double-integrators, in both horizontal and vertical directions, yields

$$A = \operatorname{diag}(A_i), \quad B_1 = \operatorname{diag}(B_{1i}), \quad B_2 = \operatorname{diag}(B_{2i}),$$
$$A_i = \begin{bmatrix} O_2 & I_2 \\ O_2 & O_2 \end{bmatrix}, \; B_{1i} = B_{2i} = \begin{bmatrix} O_2 \\ I_2 \end{bmatrix}, \; i = 1, \ldots, 9,$$

where I_2 and O_2 are 2×2 identity and zero matrices. The control weight R is set to identity, and the state weight Q is obtained by penalizing both the absolute and the relative position errors

$$x^T Q\, x = \sum_{i=1}^{9} \Big(p_{1i}^2 + p_{2i}^2 + 10 \sum_{j \in \mathcal{N}_i} \big((p_{1i} - p_{1j})^2 + (p_{2i} - p_{2j})^2 \big) \Big),$$

where p_{1i} and p_{2i} are the absolute position errors of the ith vehicle in the horizontal and vertical directions, respectively, and set $\mathcal{N}_i$ determines neighbors of the ith vehicle.

The decentralized control architecture with no communication between vehicles specifies the block diagonal structure $\mathcal{S}_d$; see Fig. 4.1a. We solve (SH2) for $F \in \mathcal{S}_d$ and obtain the Lagrange multiplier $V^\star \in \mathcal{S}_d^c$; see Fig. 4.1b. We next separate the entries of $V^\star$ into two groups depending on their magnitudes. For example, we put $V_{ij}^\star$ into group large if its magnitude is greater than half of the maximum absolute value of the entries of $V^\star$; Otherwise, we put $V_{ij}^\star$ into group small. In other words,

$$V_{ij}^\star \text{ is in group } \begin{cases} \text{small}, & \text{if } \; 0 < |V_{ij}^\star| \le 0.5\, V_M, \\ \text{large}, & \text{if } \; |V_{ij}^\star| > 0.5\, V_M, \end{cases} \tag{4.6}$$

where V_M is the maximum absolute value of the entries of $V^\star$. We solve (SH2) for $F \in \mathcal{S}_s$ or $F \in \mathcal{S}_l$, where $\mathcal{S}_s$ and $\mathcal{S}_l$ are the subspaces obtained from removing the constraints $\{F_{ij} = 0\}$ corresponding to $\{V_{ij}^\star\}$ in groups small and large, respectively. We also consider the performance of the optimal controller in the unstructured subspace $\mathcal{S}_u$

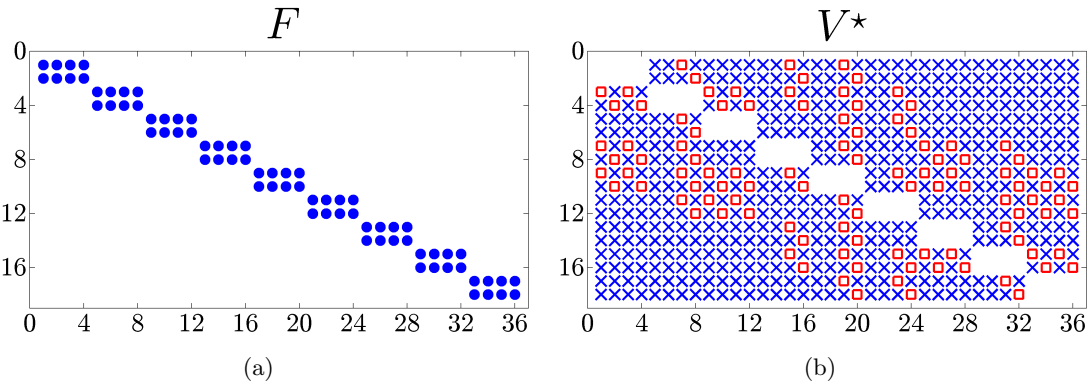


Figure 4.1: (a) Block diagonal feedback gain F where each block signifies that the two control inputs acting on each vehicle only have access to the four states of that vehicle; (b) Lagrange multiplier $V^\star$ with entries separated into groups small ($\times$) and large ($\square$) according to (4.6).

Table 4.2: Performance improvement, $\kappa = (J_d^\star - J^\star)/J_d^\star$, relative to the optimal $\mathcal{H}_2$ norm $J_d^\star = 65.4154$ with decentralized structure $\mathcal{S}_d$. Here, r is the number of extra variables in $\mathcal{S}_s$, $\mathcal{S}_l$, and $\mathcal{S}_u$ compared with $\mathcal{S}_d$, and κ/r is the performance improvement per variable.

	$J^\star$	κ	r	κ/r
$\mathcal{S}_s$	64.1408	1.95%	472	0.0041%
$\mathcal{S}_l$	64.2112	1.84%	104	0.0177%
$\mathcal{S}_u$	62.1183	5.04%	576	0.0088%

with no constraints on F.

Table 4.2 shows the influence of the number of optimization variables on the performance improvement. Note that $\mathcal{S}_l$ has the largest improvement *per variable* among all three structures $\mathcal{S}_s$, $\mathcal{S}_l$, and $\mathcal{S}_u$. As illustrated in Fig. 4.2, $\mathcal{S}_l$ determines a *localized* communication architecture in which each vehicle communicates only with its neighbors. Therefore, the Lagrange multiplier $V^\star$ identifies nearest neighbor interactions as the architecture of the distributed controller. We note that the nearest neighbor communication graph is also identified by solving the sparsity-promoting optimal control (SP) with the weighted ℓ_1 penalty.

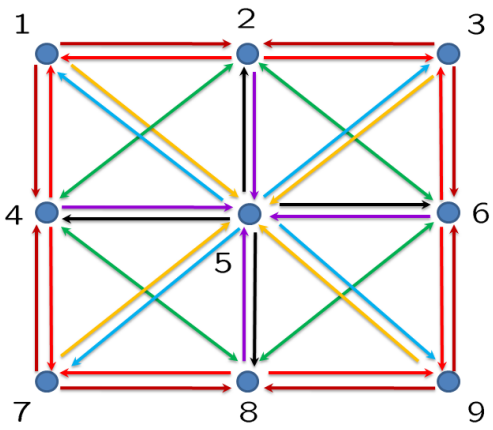


Figure 4.2: Localized controller architecture in which each vehicle communicates only with its neighbors. The arrow directed from node i to node j indicates that node i is sending information to node j. Priority order of communication channels is determined by the absolute values of $V_{ij}^{\star}$, ranging from the highest to the lowest: brown, red, orange, green, blue, purple, and black.

Chapter 5

Conclusions and future directions

Conclusions

We design sparse and block sparse state feedback gains that optimize the $\mathcal{H}_2$ performance of distributed systems. The design procedure consists of a structure identification step and a polishing step. In the identification step, we solve a parameterized family of sparsity-promoting optimal control problems whose solution gradually changes from the centralized gain to the sparse gain of interest. In the polishing step, we fix the identified sparsity patterns $\mathcal{S}$ and then solve the structured $\mathcal{H}_2$ problem to find the optimal feedback gain belonging to $\mathcal{S}$.

We demonstrate that the alternating direction method of multipliers (ADMM) is a simple but powerful algorithm for the sparsity-promoting optimal control problem. It alternates iteratively between promoting the sparsity of the controller and optimizing the performance of the system. This alternating mechanism facilitates the incorporation of different sparsity-promoting penalty functions that induce elementwise or blockwise sparsity patterns. By exploiting the separable structure of these penalty functions, we decompose the corresponding optimization problems into sub-problems that can be solved analytically.

We also employ the augmented Lagrangian method to design structured optimal controllers with *a priori* assigned structures. This approach does not require the knowledge of a structured stabilizing gain to initialize the algorithm. Furthermore, we use sensitivity interpretation of the Lagrange multiplier to identify effective relaxations of sparsity constraints for improving $\mathcal{H}_2$ performance.

Extensions and future directions

Alternative performance indices. Although we focus on the $\mathcal{H}_2$ performance, the developed framework based on ADMM can be extended to design problems with other performance indices. This is because the G-minimization step (3.3b) in ADMM is independent of the $\mathcal{H}_2$ objective function. Thus, the analytical solutions to G-minimization problem can be done exactly as in Section 3.2 for any performance index. On the other hand, the F-minimization step (3.3a) needs modifications for different objective functions.

Distributed implementation of ADMM. It is of interest to develop distributed schemes for the F-minimization step (3.3a) in the ADMM algorithm. Recently developed tools from distributed optimization [57, 58] may play an important role in this effort.

Uniqueness of the solution path. For a class of distributed systems studied in Chapter 6, the sparsity-promoting optimal control problem (SP) can be formulated as a convex optimization problem. In this case, the uniqueness of the γ-parameterized solution path to (SP) can be readily established. For broader classes of systems, the uniqueness property of the solution path is an important future research question.

Part II

Sparsity-promoting optimal control for consensus networks

Chapter 6

Identification of sparse communication graphs for consensus networks

Reaching consensus in a decentralized fashion is an important problem in network science [4]. This problem is often encountered in social networks where a group of individuals is trying to agree on certain issues [28, 29]. A related problem of distributing computational load evenly over a network of processors has been studied extensively in computer science [30, 31]. Recently, consensus problems have received considerable attention in the context of distributed control [33–36].

In cooperative control of vehicular formations, it is desired to use local interactions between vehicles to reach agreement on heading direction, velocity, and inter-vehicular spacing. Furthermore, vehicles have to maintain agreement in the presence of uncertainty introduced by modeling errors, measurement noise, and communication failures. It is therefore of importance to consider robustness of consensus networks.

In this chapter, we consider consensus networks subject to white stochastic disturbances. Under the influence of disturbances, the network will not reach a consensus value, but each node in the network will fluctuate around the average of node values. Therefore, the steady-state variance of the deviation from average can be used to quantify the robustness of stochastically forced consensus networks.

It is not surprising that the performance of networks depends heavily on the underlying interconnection topology [32, 37, 59]. A more connected network usually achieves better performance at the expense of more communication links. It is thus desired to identify networks that strike a balance between the number of communication links and the performance of the network. With this objective in mind, we build upon the sparsity-promoting optimal control framework developed in Part I. We solve a parameterized family of sparsity-promoting optimal control problems whose solution traces a trade-off curve between the performance of network and the sparsity of communication graph. The main contribution of this chapter is a semidefinite program formulation for the optimal network design problem. As a consequence, the globally optimal solution can be computed efficiently.

This chapter is organized as follows. In Section 6.1, we formulate the sparsity-promoting optimal control problem for the design of consensus networks. In Section 6.2, we provide a semidefinite program formulation for the optimal control problem. In Section 6.3, we use an illustrative example to demonstrate the utility of the developed approach.

6.1 Undirected consensus networks

Consider a network with N integrator nodes

$$\dot{x}_i = u_i + d_i, \qquad i = 1, \ldots, N, \tag{6.1}$$

where u_i is the control input acting on node i and d_i is the uncorrelated white stochastic disturbance with zero mean and unit variance. Each node forms its control action using a weighted sum of the differences between itself and other nodes

$$u_i \;=\; -\sum_{j \neq i} F_{ij}\,(x_i - x_j). \tag{6.2}$$

We focus on *undirected networks* in which the communication link between any pair of nodes (i, j) is bidirectional with the same feedback gain

$$F_{ij} \;=\; F_{ji}, \qquad i \neq j.$$

A zero element F_{ij} of the feedback gain matrix $F \in \mathbb{R}^{N \times N}$ implies that there is *no* communication link between nodes i and j. The communication architecture of the network is therefore determined by the sparsity pattern of F; in particular, the number of communication links is determined by the number of nonzero elements of F.

We are interested in identifying sparsity patterns of F that strike a balance between the number of communication links and a performance measure that quantifies the robustness of consensus in the stochastically forced network. Recall the sparsity-promoting optimal control problem

$$\underset{F}{\text{minimize}} \quad J(F) \;+\; \gamma\, g(F) \tag{6.3}$$

where $\gamma \geq 0$ is the problem parameter that controls the trade-off between the performance measure J (defined in Section 6.1.2) and the weighted ℓ_1 norm (introduced in Section 2.2)

$$g(F) \;=\; \sum_{i,\,j} W_{ij}\,|F_{ij}|, \qquad W_{ij} \;>\; 0.$$

Here, the weights W_{ij} are determined by (2.9) with the feedback gain being the solution $F^\star$ of (6.3) at the previous value of γ. (For $\gamma = 0$, problem (6.3) does not depend on the weighted ℓ_1 norm.)

6.1.1 Stochastically forced consensus networks

By substituting (6.2) into (6.1) and putting the variables in vectors, we can express the dynamics of stochastically forced consensus networks as follows

$$\dot{x} = -Fx + d. \tag{6.4}$$

Since F is a symmetric matrix $F = F^T$ for undirected networks, the stability of system (6.4) amounts to F being a positive definite matrix $F \succ 0$. Using only relative information exchange (6.2), however, the average mode of the network is not stabilizable. This can be verified by noting that F has a zero eigenvalue associated with the vector of all ones $\mathbb{1}$,

$$F\mathbb{1} = 0. \tag{6.5}$$

Hence, the control input u is invariant by adding a constant α to the state x, i.e.,

$$u = -Fx = -F(x + \alpha\mathbb{1}).$$

Note that property (6.5) is independent of the choice of F_{ij}. For connected networks, we can choose F_{ij} (see Section 7.2) such that the remaining $N-1$ eigenvalues of F are positive, i.e.,

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N.$$

In the absence of stochastic disturbances, all nodes converge asymptotically to the

average of the initial condition

$$\lim_{t\to\infty} x(t) \;=\; \mathbb{1}\,\bar{x}(0) \tag{6.6}$$

where the average mode is given by

$$\bar{x}(t) \;:=\; \frac{1}{N}\sum_{i=1}^{N} x_i(t) \;=\; \frac{1}{N}\,\mathbb{1}^T x(t).$$

To see (6.6), consider the eigenvalue decomposition $F = S\Lambda S^T$ with the diagonal matrix $\Lambda \;=\; \mathrm{diag}\{\lambda_i\}_{i=1}^N$ and the orthonormal matrix $S \;=\; [\,s_1 \;\cdots\; s_N\,]$ with $s_1 \;=\; \frac{1}{\sqrt{N}}\mathbb{1}$. Then (6.6) follows from

$$\lim_{t\to\infty} x(t) \;=\; \lim_{t\to\infty} \mathrm{e}^{-Ft}x(0) \;=\; \lim_{t\to\infty}\left(s_1 s_1^T \;+\; \sum_{i=2}^{N} \mathrm{e}^{-\lambda_i t} s_i s_i^T\right) x(0) \;=\; s_1 s_1^T x(0).$$

In the presence of disturbances, the average mode undergoes a random walk starting from the initial average value and its variance grows linearly with time. To see this, consider the dynamics of the average mode

$$\dot{\bar{x}} \;=\; \bar{d},$$

where $\bar{d} \;=\; \frac{1}{N}\mathbb{1}^T d$ is the average of disturbances. Since the trajectory of the average mode is determined by

$$\bar{x}(t) \;=\; \bar{x}(0) \;+\; \int_0^t \bar{d}(\tau)\,\mathrm{d}\tau,$$

and since $\bar{d}$ is white stochastic disturbance with zero and unit variance, we have

$$\mathbf{E}\{\bar{x}(t)\} \;=\; \bar{x}(0), \qquad \mathbf{E}\{(\bar{x}(t)-\bar{x}(0))^2\} \;=\; \int_0^t\int_0^t \mathbf{E}\{\bar{d}(\tau)\bar{d}(\sigma)\}\,\mathrm{d}\tau\,\mathrm{d}\sigma \;=\; t.$$

6.1.2 Performance of consensus networks

Several performance outputs can be considered to quantify the robustness of consensus. For example, consider the following global and local performance errors [37]:

- The global error quantifies the deviation of each node from the average mode

$$\tilde{x}_i \;=\; x_i \;-\; \bar{x}, \qquad i \;=\; 1, \ldots, N. \tag{6.7}$$

- The local error quantifies the difference between the neighboring nodes

$$\tilde{x}_{ij} \;=\; x_i \;-\; x_j \;\text{ for }\; (i,j) \;\in\; \mathcal{E} \tag{6.8}$$

  where $\mathcal{E}$ is the edge set of local interaction graph (see Section 6.3).

Note that the average mode $\bar{x}$ is unobservable from both global and local errors. Alternative performance outputs that render $\bar{x}$ unobservable can be considered as well [32, 37]. Let such a performance output be given by $z_1 = Cx$ with the output matrix C satisfying $C\mathbb{1} = 0$. The control objective is to keep the performance output z_1 and the control input $z_2 = u$ small under the influence of disturbances. This objective can be quantified using the $\mathcal{H}_2$ norm of the system from d to z

$$\begin{aligned} \dot{x} \;&=\; -Fx \;+\; d \\ z \;&=\; \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \;=\; \begin{bmatrix} C \\ -F \end{bmatrix} x. \end{aligned}$$

The $\mathcal{H}_2$ norm J is given by

$$J(F) \;=\; \operatorname{trace}\left(\int_0^\infty \mathrm{e}^{-Ft} \left(Q \;+\; FF\right) \mathrm{e}^{-Ft} \mathrm{d}t \right) \tag{6.9}$$

where $Q = C^T C$ and Q satisfies

$$Q\mathbb{1} = 0. \tag{6.10}$$

Note that the integral in (6.9) is bounded even though the average mode is not asymptotically stable. This is because the average mode is not observable from either Q or F.

6.2 Semidefinite programming formulation for the sparsity-promoting optimal control problem

In this section, we show that the sparsity-promoting optimal control problem (6.3) can be formulated as a semidefinite program (SDP). To this end, we first derive an alternative expression for the $\mathcal{H}_2$ norm.

Proposition 2. *Let the symmetric matrices F and Q satisfy*

$$\begin{aligned} F\mathbb{1} &= 0, \quad F + \mathbb{1}\mathbb{1}^T/N \succ 0, \\ Q\mathbb{1} &= 0, \quad Q + \mathbb{1}\mathbb{1}^T/N \succ 0. \end{aligned}$$

Then the $\mathcal{H}_2$ norm J in (6.9) can be expressed as

$$J(F) = \frac{1}{2}\operatorname{trace}\left(QF^\dagger + F\right) = \frac{1}{2}\operatorname{trace}\left(Q(F + \mathbb{1}\mathbb{1}^T/N)^{-1} + F\right), \tag{6.11}$$

where $F^\dagger$ denotes the Moore-Penrose pseudoinverse of F.

Proof. From the eigenvalue decomposition of the symmetric matrix F

$$F = S\Lambda S^T = \begin{bmatrix} \frac{1}{\sqrt{N}}\mathbb{1} & \bar{S} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \bar{\Lambda} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{N}}\mathbb{1}^T \\ \bar{S}^T \end{bmatrix}, \tag{6.12}$$

several formulae follow

$$F \;=\; \bar{S}\bar{\Lambda}\bar{S}^T \tag{6.13}$$

$$F^\dagger \;=\; \bar{S}\bar{\Lambda}^{-1}\bar{S}^T \tag{6.14}$$

$$\;=\; (F \,+\, \mathbb{1}\mathbb{1}^T/N)^{-1} \,-\, \mathbb{1}\mathbb{1}^T/N. \tag{6.15}$$

We now compute

$$J(F) \;=\; \mathrm{trace}\left(\int_0^\infty \mathrm{e}^{-\Lambda t}\,(S^TQS \,+\, \Lambda^2)\,\mathrm{e}^{-\Lambda t}\,\mathrm{d}t\right) \tag{6.16}$$

$$\;=\; \int_0^\infty \mathrm{trace}\left((\bar{S}^TQ\bar{S} \,+\, \bar{\Lambda}^2)\,\mathrm{e}^{-2\bar{\Lambda}t}\right)\mathrm{d}t \tag{6.17}$$

$$\;=\; \frac{1}{2}\,\mathrm{trace}\left(\bar{S}^TQ\bar{S}\,\bar{\Lambda}^{-1} \,+\, \bar{\Lambda}\right) \tag{6.18}$$

$$\;=\; \frac{1}{2}\,\mathrm{trace}\left(QF^\dagger \,+\, F\right) \tag{6.19}$$

$$\;=\; \frac{1}{2}\,\mathrm{trace}\left(Q(F \,+\, \mathbb{1}\mathbb{1}^T/N)^{-1} \,+\, F\right). \tag{6.20}$$

Here, step (6.16) follows from the eigenvalue decomposition (6.12); step (6.17) follows from the cyclic property of trace, i.e., $\mathrm{trace}(M_1M_2) = \mathrm{trace}(M_2M_1)$; step (6.18) follows from

$$\int_0^\infty \mathrm{e}^{-2\bar{\Lambda}t}\,\mathrm{d}t \;=\; \frac{1}{2}\,\bar{\Lambda}^{-1};$$

step (6.19) follows from (6.14); and finally, step (6.20) follows from (6.15) in conjunction with $Q\mathbb{1} = 0$. □

We now state the main result of this section.

Proposition 3. *For the objective function J in (6.11), the sparsity-promoting optimal control problem*

$$\underset{F}{\text{minimize}} \;\; J(F) \;+\; \gamma \sum_{i,\,j} W_{ij}\,|F_{ij}| \tag{6.21}$$

can be formulated as an SDP

$$\begin{aligned} \underset{X,Y,F}{\text{minimize}} \quad & \frac{1}{2}\operatorname{trace}(X + F) + \gamma \mathbb{1}^T Y \mathbb{1} \\ \text{subject to} \quad & \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & F + \mathbb{1}\mathbb{1}^T/N \end{bmatrix} \succeq 0 \\ & F\mathbb{1} = 0 \\ & -Y \leq W \circ F \leq Y. \end{aligned} \tag{6.22}$$

Here, a positive semidefinite matrix M is denoted as $M \succeq 0$ and an elementwise nonnegative matrix M is denoted as $M \geq 0$. The notation $M_1 \leq M_2$ is understood as $M_2 - M_1 \geq 0$ and $\circ$ is the elementwise multiplication of two matrices.

Proof. Transforming the weighted ℓ_1 norm in (6.21) to a linear function with linear inequality constraints yields

$$\begin{aligned} \underset{Y,F}{\text{minimize}} \quad & J(F) + \gamma \mathbb{1}^T Y \mathbb{1} \\ \text{subject to} \quad & -Y \leq W \circ F \leq Y \end{aligned}$$

where Y is a matrix with nonnegative elements $Y \geq 0$. The result then follows from Proposition 4 which provides the SDP formulation for the minimization of J. □

Proposition 4. *Suppose that*

$$Q\mathbb{1} = 0, \quad Q + \mathbb{1}\mathbb{1}^T/N \succ 0. \tag{6.23}$$

Then the optimization problem

$$\begin{aligned} \underset{F}{\text{minimize}} \quad & \frac{1}{2}\operatorname{trace}\left(Q(F + \mathbb{1}\mathbb{1}^T/N)^{-1} + F\right) \\ \text{subject to} \quad & F\mathbb{1} = 0, \quad F + \mathbb{1}\mathbb{1}^T/N \succ 0 \end{aligned} \tag{6.24}$$

can be formulated as an SDP

$$\begin{aligned} \underset{X,\,F}{\text{minimize}} \quad & \frac{1}{2}\,\text{trace}\,(X\,+\,F) \\ \text{subject to} \quad & \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & F + \mathbb{1}\mathbb{1}^T/N \end{bmatrix} \succeq 0 \\ & F\mathbb{1} \,=\, 0 \end{aligned} \tag{6.25}$$

where $Q^{1/2}$ denotes the matrix square-root of Q, i.e., $Q = Q^{1/2}Q^{1/2}$.

Proof. The key step of the proof is to show that the solution F from (6.25) satisfies

$$\bar{F} \,:=\, F \,+\, \mathbb{1}\mathbb{1}^T/N \,\succ\, 0, \tag{6.26}$$

i.e., $\bar{F}$ is a positive definite matrix. Clearly, $\bar{F} \succeq 0$ follows from the LMI in (6.25). Recall the generalized Schur complement [43, Appendix A.5.5]

$$\begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & \bar{F} \end{bmatrix} \succeq 0 \iff \bar{F} \succeq 0,\;\; (I - \bar{F}\bar{F}^\dagger)Q^{1/2} \,=\, 0,\;\; X \,-\, Q^{1/2}\bar{F}^\dagger Q^{1/2} \succeq 0$$

and consider the eigenvalue decomposition $Q = U\Lambda_Q U^T$ where $U = [\,\frac{1}{\sqrt{N}}\mathbb{1}\;\; \bar{U}\,]$ and

$$\Lambda_Q = \text{diag}\,(\lambda_Q), \qquad \lambda_Q \,=\, [\,0\;\lambda_{Q2}\;\cdots\;\lambda_{QN}\,], \tag{6.27}$$

with $\lambda_{Qi} > 0$ for $i = 2, \ldots, N$. Then multiplying U^T from the left and U from the right to $(I - \bar{F}\bar{F}^\dagger)Q^{1/2} = 0$ yields

$$U^T(I - \bar{F}\bar{F}^\dagger)U\Lambda_Q^{1/2} \,=\, 0.$$

From (6.27), it follows that the matrix $M := U^T(I - \bar{F}\bar{F}^\dagger)U$ has all zeros from the 2nd

to the Nth columns. Since M is a symmetric matrix, it follows that it has all zeros from the 2nd to the Nth rows as well. In other words, the only nonzero element of M is at its 1st row and 1st column, i.e.,

$$M = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} = U^T(I - \bar{F}\bar{F}^\dagger)U.$$

Therefore,

$$\bar{F}\bar{F}^\dagger = I - \begin{bmatrix} \frac{1}{\sqrt{N}}\mathbb{1} & \bar{U} \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{N}}\mathbb{1}^T \\ \bar{U}^T \end{bmatrix} = I - \frac{a}{N}\mathbb{1}\mathbb{1}^T.$$

Simple calculation

$$\bar{F} = \bar{F}\bar{F}^\dagger\bar{F} = \bar{F} - \frac{a}{N}\mathbb{1}\mathbb{1}^T\bar{F} = \bar{F} - \frac{a}{N}\mathbb{1}\mathbb{1}^T$$

shows that $a = 0$. Therefore, $\bar{F}\bar{F}^\dagger = I$ and (6.26) follows. Then the equivalence between (6.24) and (6.25) can be established by noting that

$$\begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & \bar{F} \end{bmatrix} \succeq 0 \iff X \succeq Q^{1/2}\bar{F}^{-1}Q^{1/2}$$

whenever $\bar{F} \succ 0$. To minimize the objective function in (6.24) for $\bar{F} \succ 0$, we simply take $X = Q^{1/2}\bar{F}^{-1}Q^{1/2}$, which yields the objective function in (6.25). □

6.2.1 SDP formulation for the structured $\mathcal{H}_2$ problem

After identifying the sparsity pattern $\mathcal{S}$ from the solution to (6.22), we next turn to the polishing step, i.e., solving the $\mathcal{H}_2$ problem subject to structural constraints on the

feedback matrix,

$$\begin{array}{ll} \underset{F}{\text{minimize}} & J(F) \\ \text{subject to} & F\mathbb{1} = 0, \quad F + \mathbb{1}\mathbb{1}^T/N \succ 0, \quad F \in \mathcal{S}. \end{array} \tag{6.28}$$

Here, we fix the sparsity pattern $F \in \mathcal{S}$ and then solve (6.28) to obtain the optimal feedback gain that belongs to $\mathcal{S}$. Problem (6.28) can be formulated as the following SDP

$$\begin{array}{ll} \underset{X,\,F}{\text{minimize}} & \dfrac{1}{2}\,\text{trace}\,(X + F) \\ \text{subject to} & \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & F + \mathbb{1}\mathbb{1}^T/N \end{bmatrix} \succeq 0 \\ & F\mathbb{1} = 0 \\ & F \circ I_{\mathcal{S}} = F \end{array}$$

where the structural identity $I_{\mathcal{S}}$ of $\mathcal{S}$ is given by (3.17).

6.3 An example

Consider $N = 50$ randomly distributed nodes in a region of 10×10 units. Let two nodes be neighbors if their Euclidean distance is not greater than 2 units; see Fig. 6.1a. We are interested in keeping both global and local errors in (6.7) and (6.8) small. Thus, the performance output z is given by

$$z = \begin{bmatrix} z_g \\ z_l \\ z_2 \end{bmatrix} = \begin{bmatrix} (I - \mathbb{1}\mathbb{1}^T/N)\,x \\ E^T x \\ -Fx \end{bmatrix}$$

where z_g and z_l denote the global and local errors, respectively, and E denotes the incidence matrix of the edge set $\mathcal{E}$ of the local interaction graph shown in Fig. 6.1a.

Each column of E is a vector of N elements representing an edge in $\mathcal{E}$; for an edge (i, j), the corresponding column of E has 1 and -1 at the ith and jth elements, and zero everywhere else. With the above choice of z, the matrix Q is determined by

$$Q = EE^T + (I - \mathbb{1}\mathbb{1}^T/N).$$

We solve the sparsity-promoting optimal control problem (6.3), followed by the polishing step described in Section 6.2.1, with 100 logarithmically-spaced points for $\gamma \in [10^{-3}, 1]$. As shown in Fig. 6.2, the number of nonzero elements of F decreases and the $\mathcal{H}_2$ norm J increases as γ increases. For $\gamma = 1$, the identified communication graph establishes long-range links between selected pairs of remote nodes, in addition to the interactions between neighbors; see Fig. 6.1b. Relative to the centralized gain F_c, the identified sparse gain F uses 7% nonzero elements, i.e.,

$$\mathbf{card}\,(F)/\mathbf{card}\,(F_c) = 7\%$$

and achieves a performance loss of only 14%, i.e.,

$$(J - J_c)/J_c = 14\%.$$

Here, F_c is the solution to (6.22) with $\gamma = 0$ and F is the solution to (6.22) with $\gamma = 1$, followed by the polishing step in Section 6.2.1.

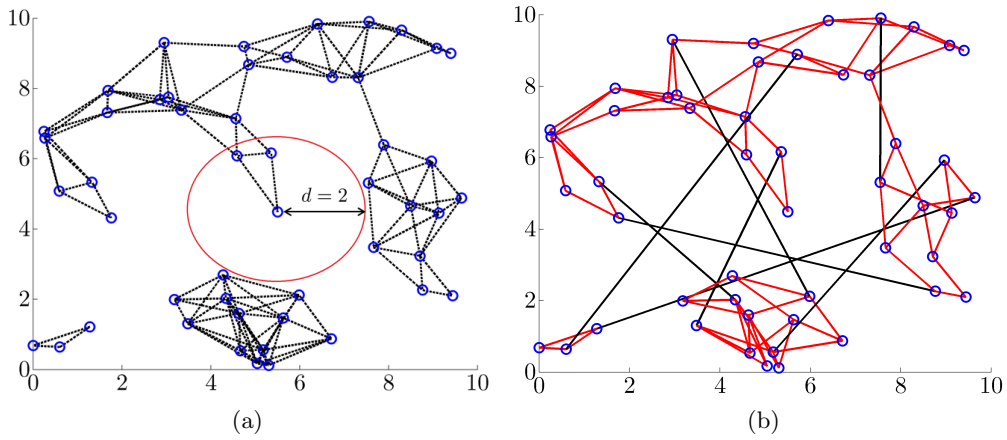


Figure 6.1: (a) Local performance graph where edges connect every pair of nodes with a distance not greater than 2 units. (b) Identified communication graph for $\gamma = 1$ where the long-range communication links are highlighted in black color.

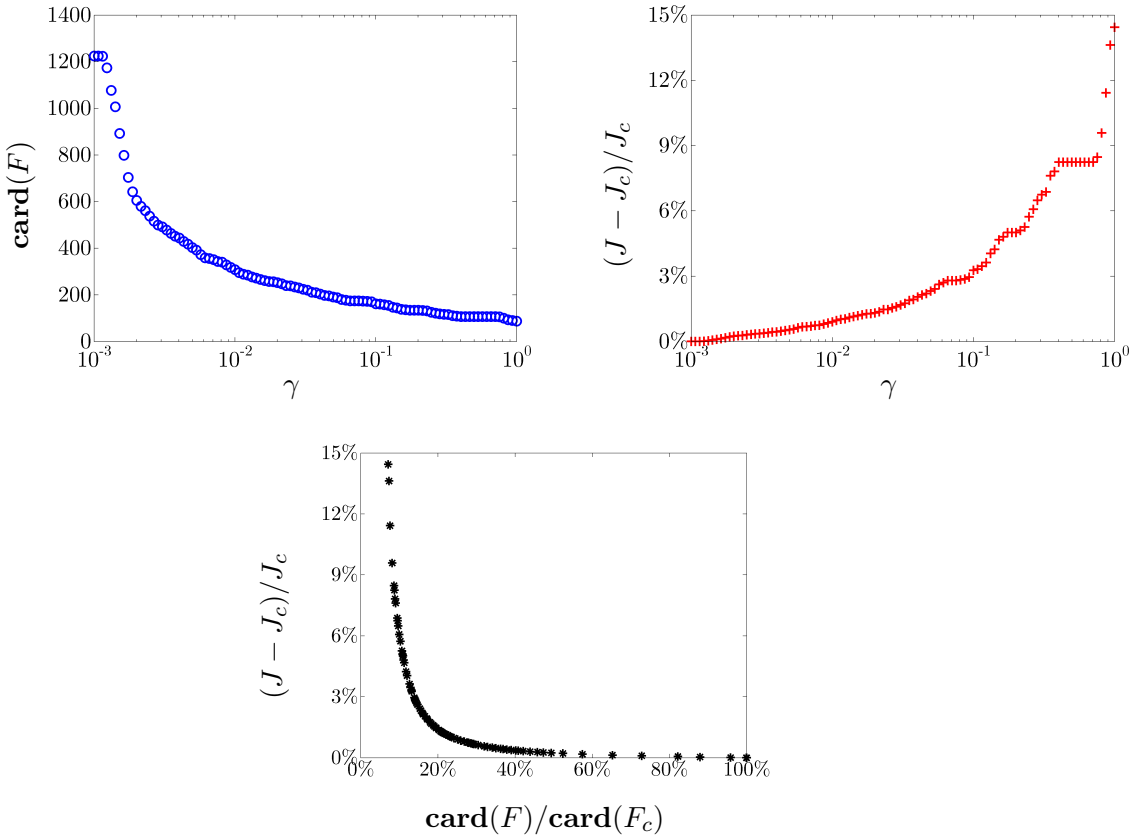


Figure 6.2: The solution to (6.3) as a function of γ, followed by the polishing step in Section 6.2.1, for the network shown in Fig. 6.1a.

Chapter 7

Design of communication graphs for consensus networks

In this chapter, we focus on the structured $\mathcal{H}_2$ problem for the design of consensus networks. We obtain several expressions for the $\mathcal{H}_2$ norm in terms of edge weights of the identified communication graph. We show that the $\mathcal{H}_2$ performance is a convex function over the convex set of stabilizing feedback gains over the edges. For several simple graphs, we derive explicit formulae for the solution to the structured $\mathcal{H}_2$ problem and provide asymptotic scalings of local and global performance measures with respect to the network size. The key step that facilitates the results in this chapter is the elimination of the average mode using a change of coordinates based on an algebraic representation of graphs.

This chapter is organized as follows. In Section 7.1, we determine the $\mathcal{H}_2$ norm of the system with a minimal realization obtained by removing the average mode. This is done by a similarity transformation based on the incidence matrix of graphs. Using the algebraic representation of the Laplacian matrix, we express the $\mathcal{H}_2$ objective function in terms of edge weights. In Section 7.2, we character the set of stabilizing gains and show

the convexity of the $\mathcal{H}_2$ performance with respect to edge weights. In Section 7.3, we derive analytical expressions for the solution to the structured $\mathcal{H}_2$ problem for several simple graphs. Furthermore, we provide the asymptotic scaling trends of local and global performance measures in large networks.

7.1 Alternative expressions for the $\mathcal{H}_2$ norm

Recall the dynamics of consensus networks from Chapter 6

$$\begin{aligned} \dot{x} &= -Fx + d \\ z &= \begin{bmatrix} C \\ -F \end{bmatrix} x. \end{aligned}$$

where both C and F have a zero eigenvalue associated with the eigenvector $\mathbb{1}$. Since the average mode $\bar{x} = \mathbb{1}^T x/N$ is not observable from z, we next separate $\bar{x}$ from the dynamics evolving in $\mathbb{1}^\perp$ (i.e., the subspace orthogonal to $\mathbb{1}$). Let $U \in \mathbb{R}^{N\times(N-1)}$ be a full rank matrix such that all columns of U are orthogonal to $\mathbb{1}$. Using a change of coordinates

$$\psi = \mathcal{T}x$$

where

$$\mathcal{T} = \begin{bmatrix} U^T \\ \mathbb{1}^T/N \end{bmatrix}, \qquad \mathcal{T}^{-1} = \begin{bmatrix} U(U^TU)^{-1} & \mathbb{1} \end{bmatrix}, \tag{7.1}$$

we have

$$\dot{\psi} = -\mathcal{T}F\mathcal{T}^{-1}\psi + \mathcal{T}d.$$

Straightforward algebra shows that

$$\dot{\psi} := \begin{bmatrix} \dot{\phi} \\ \dot{\bar{x}} \end{bmatrix} = \begin{bmatrix} -U^T F V & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \bar{x} \end{bmatrix} + \begin{bmatrix} U^T \\ \mathbb{1}^T/N \end{bmatrix} d$$

where

$$V = U(U^T U)^{-1}.$$

Under the new coordinates, the performance output is given by

$$z = \begin{bmatrix} CV & 0 \\ -FV & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \bar{x} \end{bmatrix}.$$

Removing the average mode $\bar{x}$, we obtain a minimal representation of the system containing only the state ϕ

$$\begin{aligned} \dot{\phi} &= -U^T F V \phi + U^T d \\ z &= \begin{bmatrix} CV \\ -FV \end{bmatrix} \phi. \end{aligned} \tag{7.2}$$

Thus, the $\mathcal{H}_2$ norm of the system from d to z can be expressed as

$$J(F) = \text{trace}\left(\Phi(V^T(Q + FF)V)\right), \tag{7.3}$$

where $Q = C^T C$ and Φ is the solution to the Lyapunov equation

$$(-U^T F V)\Phi + \Phi(-U^T F V)^T = -U^T U. \tag{7.4}$$

7.1.1 Incidence matrix and graph Laplacian

In the coordinate transformation (7.1), the only condition imposed on U is that its columns are orthogonal to $\mathbb{1}$. In this subsection, we choose a particular U based on the

incidence matrix of a graph. As we will see, this choice of U simplifies the Lyapunov equation (7.4) and consequently, the expression for the $\mathcal{H}_2$ norm (7.3). Furthermore, it facilitates the derivation of analytical expressions for the optimal feedback gains in Section 7.3.

We begin by introducing the incidence matrix and the Laplacian matrix of a graph. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected connected graph with $N = |\mathcal{V}|$ nodes and $m = |\mathcal{E}|$ edges. We associate with edge l between nodes i and j a vector $e_l \in \mathbb{R}^N$ that has 1 and -1 at its ith and jth entries, respectively, and 0 everywhere else. Let $k_l \in \mathbb{R}$ be the edge weight of l. The Laplacian matrix of the graph $\mathcal{G}$ is given by

$$F = \sum_{l=1}^{m} k_l \, e_l \, e_l^T .$$

Alternatively, it can be expressed as

$$F = EKE^T$$

where $E = [\, e_1 \; e_2 \; \cdots \; e_m \,] \in \mathbb{R}^{N \times m}$ is the incidence matrix of the graph $\mathcal{G}$ and $K = \operatorname{diag}\{k_l\}_{l=1}^{m} \in \mathbb{R}^{m \times m}$ is a diagonal matrix of edge weights.

Note that by construction all columns of the incidence matrix E are orthogonal to $\mathbb{1}$. Therefore, a particular choice of U is to pick $N-1$ independence columns of E. This can be done whenever the graph $\mathcal{G}$ is connected and in this case U will be the incidence matrix of a tree subgraph of $\mathcal{G}$; see [60]. Since the columns of U form the basis of $\mathbb{1}^\perp$ and since all columns of E belong to $\mathbb{1}^\perp$, the incidence matrix can be expressed as a linear combination of the columns of U,

$$E = UP.$$

As a consequence, the graph Laplacian can be expressed as

$$F \;=\; EKE^T \;=\; UPKP^TU^T.$$

Using $U^TV = U^TU(U^TU)^{-1} = I$, the Lyapunov equation (7.4) becomes

$$U^TUPKP^T\Phi \;+\; \Phi PKP^TU^TU \;=\; U^TU,$$

whose solution is given by

$$\Phi \;=\; \frac{1}{2}\,(PKP^T)^{-1}. \tag{7.5}$$

Therefore, the $\mathcal{H}_2$ norm (7.3) can be expressed as a function of the diagonal matrix K

$$J(K) \;=\; \frac{1}{2}\,\mathrm{trace}\left((PKP^T)^{-1}(V^TQV) \;+\; U^TU(PKP^T)\right). \tag{7.6}$$

7.1.2 Local and global performance measures

For the local performance output (6.8), we have

$$C \;=\; E^T.$$

Then

$$V^TQV \;=\; V^TC^TCV \;=\; PP^T$$

and the $\mathcal{H}_2$ norm (7.6) is given by

$$J_l(K) \;=\; \frac{1}{2}\,\mathrm{trace}\left(K_P^{-1}PP^T \;+\; U^TUK_P\right) \tag{7.7}$$

For the global performance output (6.7), we have

$$C = I - \frac{1}{N}\mathbb{1}\mathbb{1}^T.$$

Then

$$V^T Q V = V^T C^T C V = (U^T U)^{-1}$$

and the $\mathcal{H}_2$ norm (7.6) is given by

$$J_g(K) = \frac{1}{2}\operatorname{trace}\left((U^T U K_P)^{-1} + U^T U K_P\right). \tag{7.8}$$

7.2 Stability and convexity

This section contains the main results of this chapter, given by Propositions 5 and 6. In Proposition 5, we characterize the set of stabilizing gains K for the system (7.2). In Proposition 6, we show that J is a convex function on the set of stabilizing feedback gains.

Proposition 5. *The system (7.2) is stable if and only if*

$$K_P := PKP^T$$

is a positive definite matrix.

Proof. Since

$$U^T F V = U^T U K_P$$

is a product of a positive definite matrix $U^T U$ and a Hermitian matrix K_P, the proposition follows from a result in matrix analysis [61, Theorem 7.6.3] saying that the product $W_1 W_2$ of a positive definite matrix W_1 and a Hermitian matrix W_2 has the same number

of positive, negative, and zero eigenvalues as W_2. □

A sufficient condition for $K_P \succ 0$ is $K \succ 0$. Since K is the diagonal matrix, this sufficient condition is equivalent to $k_l > 0$ for all $l \in \{1, \ldots, m\}$, i.e., positive feedback gains on all edges of $\mathcal{G}$. This condition is also necessary for tree graphs, since in that case $\{E = U, P = I\}$ and $K_P = K$. Therefore, we conclude that for tree graphs, the system (7.2) is stable if and only if all feedback gains on edges are positive.

Proposition 6. *The $\mathcal{H}_2$ norm of system (7.2) is a convex function of the stabilizing feedback gain K.*

Proof. For two arbitrary stabilizing gains K_1 and K_2, we have

$$P(\theta K_1 + (1-\theta)K_2)P^T \succ 0, \qquad \theta \in (0,1),$$

that is, the convex combination $\theta K_1 + (1-\theta)K_2$ is also stabilizing. Thus, the set of stabilizing feedback gains forms a convex set. Since the $\mathcal{H}_2$ norm is given by

$$J(K) = \frac{1}{2}\operatorname{trace}\left(K_P^{-1}V^TQV + U^TUK_P\right)$$

and since $\operatorname{trace}(U^TUK_P)$ is linear (and thus convex) in K, it is sufficient to show convexity of the function $\operatorname{trace}(K_P^{-1}V^TQV)$. To this end, note that $\operatorname{trace}(K_P^{-1})$ is a convex function of the positive definite matrix K_P; see [43, Problem 3.18(a)]. Since $V^TQV \succ 0$ and the positive weighted sum preserves convexity [43, Section 3.2.1], it follows that $\operatorname{trace}(K_P^{-1}V^TQV)$ is a convex function of the stabilizing feedback gain K.

□

We conclude this section with the formulae for the gradient and Hessian of $J(K)$

$$\begin{aligned}\nabla J(K) &= \frac{1}{2}\operatorname{diag}(P^T(U^TU - K_P^{-1}V^TQVK_P^{-1})P),\\ \nabla^2 J(K) &= (P^TK_P^{-1}V^TQVK_P^{-1}P)\circ(P^TK_P^{-1}P).\end{aligned}$$

Since elementwise multiplication of two positive semidefinite matrices results in a positive semidefinite matrix [62, Theorem 5.2.1], the convexity of $J(K)$ also follows from the positive semidefiniteness of the Hessian matrix $\nabla^2 J(K)$.

7.3 Analytical solutions to the structured $\mathcal{H}_2$ problem

In this section, we consider the minimization of the global and local performance measures (7.8) and (7.7) with respect to the feedback gain matrix K for several simple graphs. By exploiting structure of path, star, circle, and complete graphs, we obtain analytical expressions for the optimal solutions.

7.3.1 Tree

A tree is a connected graph with no cycles. In this case, $P = I$ and (7.7) simplifies to

$$J_l(K) = \frac{1}{2}\operatorname{trace}\left(K^{-1} + U^TUK\right).$$

By the definition of the incidence matrix, the diagonal entries of the matrix U^TU are all equal to 2 yielding

$$J_l(K) = \frac{1}{2}\sum_{i=1}^{N-1}(k_i^{-1} + 2k_i).$$

The unique optimal feedback gains are thus *constant* for all the $N-1$ edges

$$k_l^* = 1/\sqrt{2}.$$

The optimal local performance measure is

$$J_l^* \;=\; (N-1)\sqrt{2}.$$

For the global performance measure (7.8), we have

$$\begin{aligned} J_g(K) &= \frac{1}{2}\,\mathrm{trace}\left((KU^TU)^{-1} \;+\; KU^TU\right) \\ &= \frac{1}{2}\sum_{i=1}^{N-1}(k_i^{-1}(U^TU)_{ii}^{-1} \;+\; 2k_i). \end{aligned}$$

The optimal feedback gain is thus given by

$$(k_g^*)_i \;=\; \sqrt{(U^TU)_{ii}^{-1}/2}, \qquad i \in \{1,\ldots,N-1\}.$$

The optimal global performance is

$$J_g^* \;=\; \sum_{i=1}^{N-1}\sqrt{2(U^TU)_{ii}^{-1}}.$$

Thus, in contrast to the local performance measure (7.7), the optimal feedback gain and the optimal global performance measure depend on the structure of the tree. We next consider two special cases where we can determine $(U^TU)^{-1}$ explicitly.

Path: In this case, $U^TU = T$ where T is a symmetric Toeplitz matrix with the first row given by $[\;2\;\;-1\;0\;\cdots\;0\;] \in \mathbb{R}^{N-1}$. The inverse of T can be determined analytically

$$(T^{-1})_{ij} \;=\; i(N-j)/N \quad \text{for} \quad j \;\geq\; i. \tag{7.9}$$

It follows that

$$(k_g^*)_i \;=\; \sqrt{i(N-i)/(2N)}, \qquad J_g^* \;=\; \sum_{i=1}^{N-1} \sqrt{2i(N-i)/N}.$$

For large N, $J_g^* \approx \dfrac{\pi N\sqrt{N}}{4\sqrt{2}}$.

Star: In this case, $U^TU \;=\; I + \mathbb{1}\mathbb{1}^T$ and $(U^TU)^{-1} \;=\; I - \mathbb{1}\mathbb{1}^T/N$. The diagonal entries of $(U^TU)^{-1}$ are all equal to $(N-1)/N$. The optimal feedback gain and the global performance measure are given by

$$k_g^* \;=\; \sqrt{(N-1)/(2N)}, \qquad J_g^* \;=\; (N-1)\sqrt{2(N-1)/N}.$$

7.3.2 Circle

A circle graph is an edge-transitive graph[1] and we use the result that the optimal solution for convex problems on edge-transitive graphs are constant [32], that is, $k_i^* = k$ for all $i \in \{1, \ldots, N\}$. The local performance measure (7.7) simplifies to

$$\begin{aligned} J_l(k) &= \frac{1}{2}\,\mathrm{trace}\left(PP^T(kPP^T)^{-1} + kPP^TU^TU\right) \\ &= \frac{1}{2}\,k^{-1}(N-1) \;+\; kN, \end{aligned}$$

where we used

$$\mathrm{trace}\left(PP^TU^TU\right) \;=\; \mathrm{trace}\left(E^TE\right) \;=\; 2N.$$

Thus,

$$k_l^* \;=\; \sqrt{(N-1)/(2N)}, \qquad J_l^* \;=\; \sqrt{2(N-1)^3/N}.$$

[1] An automorphism of a graph $\mathcal{G}$ is a permutation π of the vertex set $\mathcal{V}$ such that $(i,j) \in \mathcal{E}$ if and only if $(\pi(i), \pi(j)) \in \mathcal{E}$. A graph is edge-transitive if given any pair of edges there is an automorphism that transforms one edge into the other.

On the other hand, for global performance measure, we have

$$\begin{aligned} J_g(k) &= \frac{1}{2}\left(\frac{1}{k}\sum_{i=1}^{N-1}\frac{i(N-i)}{2N} + 2kN\right) \\ &= \frac{N^2-1}{24k} + kN. \end{aligned}$$

It follows that

$$k_g^* = \sqrt{(N^2-1)/(24N)}, \qquad J_g^* = \sqrt{N(N^2-1)/6}.$$

7.3.3 Complete graph

A complete graph has an edge between every pair of vertices; thus, it has $N(N-1)/2$ edges. A complete graph is also edge-transitive and with an identical procedure as done for the case of circle, for local performance measure (7.7) we have

$$k_l^* = 1/\sqrt{N}, \qquad J_l^* = (N-1)\sqrt{N}.$$

On the other hand, for the global performance measure (7.8) we have

$$\begin{aligned} J_g(k) &= \frac{1}{2}\,\mathrm{trace}\left((kPP^TU^TU)^{-1} + kPP^TU^TU\right) \\ &= \frac{1}{2}\,(N-1)((kN)^{-1} + kN). \end{aligned}$$

where we used the fact that the eigenvalues of PP^TU^TU are the nonzero eigenvalues of E^TE, which all equal to N. Thus, we have

$$k_g^* = 1/N, \qquad J_g^* = N-1.$$

Table 7.1: Optimal gains and performances for the path, star, circle, and complete graph.

	J_l^*	k_l^*	J_g^*	$(k_g^*)_i$
Path	$\sqrt{2}(N-1)$	$1/\sqrt{2}$	$(\pi/8)N\sqrt{2N}$	$\sqrt{(i(N-i))/(2N)}$
Star	$\sqrt{2}(N-1)$	$1/\sqrt{2}$	$\sqrt{2(N-1)^3/N}$	$\sqrt{(N-1)/(2N)}$
Circle	$\sqrt{2(N-1)^3/N}$	$\sqrt{(N-1)/(2N)}$	$\sqrt{N(N^2-1)/6}$	$\sqrt{(N^2-1)/(24N)}$
Complete	$(N-1)\sqrt{N}$	$1/\sqrt{N}$	$N-1$	$1/N$

Table 7.2: Comparison of the asymptotic scaling of optimal performance measures normalized by N, and the maximum value of feedback gains, for path, circle, star and complete graph.

	J_l^*/N	J_g^*/N	k_l^*	$\max(k_g^*)_i$
Path	$O(1)$	$O(\sqrt{N})$	$O(1)$	$O(\sqrt{N})$
Star	$O(1)$	$O(1)$	$O(1)$	$O(1)$
Circle	$O(1)$	$O(\sqrt{N})$	$O(1)$	$O(\sqrt{N})$
Complete	$O(\sqrt{N})$	$O(1)$	$O(1/\sqrt{N})$	$O(1/N)$

We summarize the formulas for the optimal feedback gains and the optimal performance measures (7.7) and (7.8) in Table 7.1.

7.3.4 Asymptotic scaling of performance measures with the number of nodes

We next consider how the optimal performance measures scale with the network size N. We also consider the asymptotic scaling of the optimal feedback gain $\{k_i^*\}_{i=1}^m$, which serves as an indication of the control effort of the optimal design. The results for the graphs considered are summarized in Table 7.2, where the optimal local and global performance measures are normalized by N. The normalized global performance measures are illustrated in Fig. 7.1.

In [37], it was established that using constant feedback gain for all edges of the circle, the global performance measure normalized by the formation size scales *linearly*

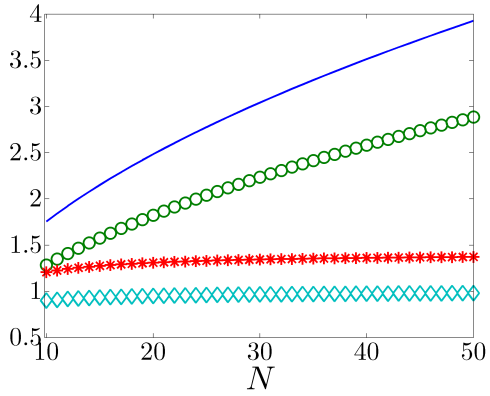


Figure 7.1: N-normalized global performance measure J_g^* for path ($-$), circle ($\circ$), star ($*$) and complete graph ($\Diamond$).

with N. This result was derived under the assumption that the amount of control effort is formation-size-independent. Note that the scaling of J_g^*/N can be reduced to a square-root dependence of N at the expense of k_g^* also increasing as a square-root function of N.

Chapter 8

Conclusions and future directions

Conclusions

We consider the design of distributed controller architectures for undirected networks of integrators. In the presence of stochastic disturbances, we identify communication topologies that strike a balance between the variance amplification of the network and the number of communication links. We solve a parameterized family of sparsity-promoting optimal control problems whose solution traces the optimal trade-off curve that starts at the centralized controller and ends at the controller with sparse graphs. We show that the optimal control problem can be formulated as a semidefinite program whose global solution can be computed efficiently.

We consider the design of edge weights on communication graphs for consensus networks. We show that the $\mathcal{H}_2$ performance is a convex function over the convex set of stabilizing feedback gains over the edges. For several simple graphs, we derive explicit formulae for the solution to the structured $\mathcal{H}_2$ problem and provide asymptotic scalings of local and global performance measures with respect to the network size.

Extensions and future directions

Consensus networks with directed graphs. For simple directed graphs such as path, circle, and star, we also obtain analytical solution to the structured $\mathcal{H}_2$ problem. The extension to special classes of directed graphs, such as directed trees, is still an open research problem.

Consensus networks with normal Laplacian matrices. In this case, we show that the $\mathcal{H}_2$ norm of consensus networks is also a convex function of edge weights. This result can be established by expressing the $\mathcal{H}_2$ norm in terms of the eigenvalues of the Laplacian matrix (e.g., see [59]) and then by employing the spectral function theory (e.g., see [32]). However, normal Laplacian matrices for directed graphs in general do not form a convex set of edge weights. This is because summation of two normal (but not symmetric) matrices in general is no longer a normal matrix. As such, it limits the utility of the convex characterization of the $\mathcal{H}_2$ norm. It is thus of interest to find a subclass of normal Laplacian matrices (different from symmetric matrices) that form a convex set of edge weights.

Part III

Optimal localized control of vehicular formations

Chapter 9

Optimal control of vehicular formations with nearest neighbor interactions

The control of vehicular platoons has attracted considerable attention since the mid sixties [63–65]. Recent technological advances in developing vehicles with communication and computation capabilities have spurred renewed interest in this area [33, 37–40, 66–69]. The simplest control objective for the one-dimensional (1D) formation shown in Fig. 9.1 is to maintain a desired cruising velocity and to keep a pre-specified constant distance between neighboring vehicles. This problem is emblematic of a wide range of technologically relevant applications including the control of automated highways, unmanned aerial vehicles, swarms of robotic agents, and satellite constellations.

We consider the design of optimal localized feedback gains for one-dimensional formations in which each vehicle only uses relative distances from its immediate neighbors and its own velocity. This nearest neighbor interaction imposes structural constraints on the feedback gains. We formulate the structured optimal control problem for both

the single- and the double-integrator models. For single-integrators, we show that the structured optimal control problem is convex when we restrict the feedback gain to be a symmetric positive definite matrix. In this case, the global minimizer can be computed efficiently, and even analytical expressions can be derived. For double-integrators, we also identify a class of convex problems by restricting the controller to symmetric position and uniform diagonal velocity gains.

We then remove this symmetric restriction for both the single- and the double-integrator models and begin the design process with a spatially uniform controller. We develop a homotopy-based Newton's method that traces a continuous solution path from this controller to the optimal localized gain. Along this homotopy path, we solve a parameterized family of the structured optimal control problems and obtain *analytical* solutions when the homotopy parameter is small. We employ perturbation analysis to decouple the matrix equations that result from optimality conditions, thereby rendering the unique optimal structured gain. This solution is used to warm-start Newton's method in order to efficiently compute the desired optimal gains as the homotopy parameter is gradually increased.

This chapter is organized as follows. In Section 9.1, we formulate the structured optimal control problem for both single- and double-integrator models. In Section 9.2, we show convexity of the symmetric gain design for the single-integrator model. In Section 9.3, we develop the homotopy-based Newtons method for non-symmetric gains. In Section 9.4, we examine the design of localized controllers for the double-integrator model.

9.1 Problem formulation

A system of N identical vehicles moving along a straight line is shown in Fig. 9.1. All vehicles are equipped with ranging devices that allow them to measure relative

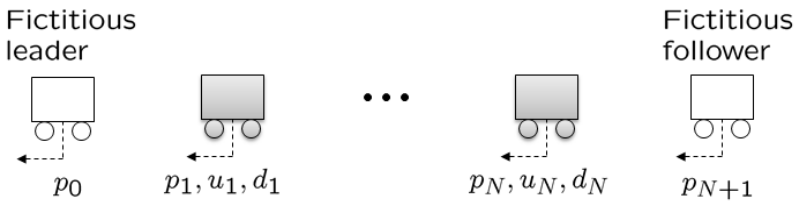


Figure 9.1: One-dimensional formation of vehicles.

distances with respect to their immediate neighbors. The objective is to design an *optimal* controller that uses *only local* information (i.e., relative distances between the neighboring vehicles) to keep each vehicle at its *global* position on a grid of regularly spaced points moving with a constant velocity.

9.1.1 Single-integrator model

We first consider the kinematic model in which control input $\bar{u}_n$ directly affects the velocity,

$$\dot{\bar{p}}_n \;=\; \bar{d}_n \,+\, \bar{u}_n, \quad n \,\in\, \{1, \ldots, N\}$$

where $\bar{p}_n$ is the position of the nth vehicle and $\bar{d}_n$ is the disturbance. The desired position of the nth vehicle is given by $p_{d,n} = v_d\, t + n\delta$ where v_d is the desired cruising velocity and δ is the desired distance between the neighboring vehicles. Every vehicle is assumed to have access to both v_d and δ.

The localized controller utilizes *relative position errors* between nearest neighbors,

$$\bar{u}_n \;=\; -\, f_n(\bar{p}_n \,-\, \bar{p}_{n-1} \,-\, \delta) \,-\, b_n(\bar{p}_n \,-\, \bar{p}_{n+1} \,+\, \delta) \,+\, v_d$$

where the design parameters f_n and b_n denote the forward and backward feedback gains of the nth vehicle. In deviation variables, $\{p_n := \bar{p}_n - p_{d,n},\, u_n := \bar{u}_n - v_d,\, d_n := \bar{d}_n\}$

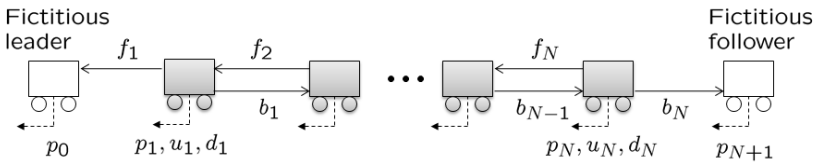


Figure 9.2: Formation of vehicles with localized non-symmetric gains.

the single-integrator model with nearest neighbor interactions is given by

$$\dot{p}_n \;=\; d_n \,+\, u_n \tag{9.1a}$$

$$u_n \;=\; -\, f_n\,(p_n \,-\, p_{n-1}) \,-\, b_n\,(p_n \,-\, p_{n+1}) \tag{9.1b}$$

where the relative position errors $p_n - p_{n-1}$ and $p_n - p_{n+1}$ can be obtained by ranging devices.

As illustrated in Fig. 9.2, fictitious lead and follow vehicles, respectively indexed by 0 and $N+1$, are added to the formation. These two vehicles are assumed to move along their desired trajectories, implying that $p_0 = p_{N+1} = 0$ and they are not considered to belong to the formation. Hence, the control actions for the 1st and the Nth vehicles are given by

$$u_1 \;=\; -\, f_1\, p_1 \,-\, b_1\,(p_1 \,-\, p_2)$$

$$u_N \;=\; -\, f_N\,(p_N \,-\, p_{N-1}) \,-\, b_N\, p_N.$$

In other words, the first and the last vehicles have access to their own global position errors p_1 and p_N, which can be obtained by equipping them with GPS devices.

9.1.2 Double-integrator model

For the double-integrator model,

$$\ddot{\bar{p}}_n \;=\; \bar{d}_n \,+\, \bar{u}_n, \qquad n \,\in\, \{1, \ldots, N\}$$

we consider the controller that has an access to the relative position errors between the neighboring vehicles and the absolute velocity errors,

$$\bar{u}_n \;=\; -\,f_n\,(\bar{p}_n \,-\, \bar{p}_{n-1} \,-\, \delta) \,-\, b_n\,(\bar{p}_n \,-\, \bar{p}_{n+1} \,+\, \delta) \,-\, g_n\,(\dot{\bar{p}}_n \,-\, v_d)$$

where g_n denotes the velocity feedback gain. In deviation variables, $\{p_n := \bar{p}_n - p_{d,n}$, $v_n := \dot{\bar{p}}_n - v_d$, $u_n := \bar{u}_n$, $d_n := \bar{d}_n\}$, the double-integrator model is given by

$$\ddot{p}_n \;=\; d_n \,+\, u_n \tag{9.2a}$$

$$u_n \;=\; -\,f_n\,(p_n \,-\, p_{n-1}) \,-\, b_n\,(p_n \,-\, p_{n+1}) \,-\, g_n\,v_n. \tag{9.2b}$$

In matrix form, control laws (9.1b) and (9.2b) can be written as

$$u \;=\; -\,FC\,p \;=\; -\,\begin{bmatrix} F_f & F_b \end{bmatrix} \begin{bmatrix} C_f \\ C_f^T \end{bmatrix} p$$

$$u \;=\; -\,FC \begin{bmatrix} p \\ v \end{bmatrix} \;=\; -\,\begin{bmatrix} F_f & F_b & F_v \end{bmatrix} \begin{bmatrix} C_f & O \\ C_f^T & O \\ O & I \end{bmatrix} \begin{bmatrix} p \\ v \end{bmatrix}$$

where p, v, and u denote the position error, the velocity error, and the control input vectors, e.g., $p = [\,p_1 \;\cdots\; p_N\,]^T$. Furthermore, the $N \times N$ diagonal feedback gains are determined by

$$F_f \;:=\; \mathrm{diag}\,\{f_n\}, \quad F_b \;:=\; \mathrm{diag}\,\{b_n\}, \quad F_v \;:=\; \mathrm{diag}\,\{g_n\}$$

and C_f is a sparse Toeplitz matrix with 1 on the main diagonal and -1 on the first

lower sub-diagonal. For example, for $N = 4$,

$$F_f = \begin{bmatrix} f_1 & 0 & 0 & 0 \\ 0 & f_2 & 0 & 0 \\ 0 & 0 & f_3 & 0 \\ 0 & 0 & 0 & f_4 \end{bmatrix}, \quad C_f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}. \tag{9.3}$$

Thus, $C_f\, p$ determines the vector of the relative position errors $p_n - p_{n-1}$ between each vehicle and the one in front of it; similarly, $C_f^T\, p$ determines the vector of the relative position errors $p_n - p_{n+1}$ between each vehicle and the one behind it.

We will also consider formations with no fictitious followers. In this case, the Nth vehicle only uses relative position error with respect to the $(N-1)$th vehicle, i.e., $b_N = 0$ implying that $u_N = -f_N\,(p_N - p_{N-1})$ for the single-integrator model and $u_N = -f_N\,(p_N - p_{N-1}) - g_N\, v_N$ for the double-integrator model.

9.1.3 Closed-loop stability: the role of fictitious vehicles

We next show that at least one fictitious vehicle is needed in order to achieve closed-loop stability. This is because the absence of GPS devices in the formation prevents vehicles from tracking their absolute desired trajectories.

For the single-integrator model, the state-feedback gain $K_p = F_f\, C_f + F_b\, C_f^T$ is a structured tridiagonal matrix satisfying $K_p\, \mathbb{1} = \begin{bmatrix} f_1 & 0 & \cdots & 0 & b_N \end{bmatrix}^T$ where $\mathbb{1}$ is the vector of all ones. If neither the 1st nor the Nth vehicle has access to its own global position, i.e., $f_1 = b_N = 0$, then K_p has a zero eigenvalue with corresponding eigenvector $\mathbb{1}$. Hence, the closed-loop system is not asymptotically stable regardless of the choice of the feedback gains $\{f_n\}_{n=2}^{N}$ and $\{b_n\}_{n=1}^{N-1}$.

For the double-integrator model, the action of $A_{\text{cl}} = A - B_2 FC$ on $\begin{bmatrix} \mathbb{1}^T & \mathbf{0}^T \end{bmatrix}^T$ is

given by

$$\begin{bmatrix} O & I \\ -K_p & -F_v \end{bmatrix} \begin{bmatrix} \mathbb{1} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -K_p \mathbb{1} \end{bmatrix}$$

where $\mathbf{0}$ is the N-vector of all zeros. Thus, if $f_1 = b_N = 0$ then A_{cl} has a zero eigenvalue with corresponding eigenvector $[\,\mathbf{1}^T \quad \mathbf{0}^T\,]^T$. Therefore, for both the single- and the double-integrator models, we need at least one vehicle with access to its global position in order to achieve closed-loop stability.

9.1.4 Structured $\mathcal{H}_2$ problem

The state-space representation of the vehicular formation is given by

$$\begin{aligned} \dot{x} &= A\,x + B_1\,d + B_2\,u \\ y &= C\,x, \qquad u = -F\,y. \end{aligned}$$

For the single-integrator model (9.1), the state vector is $x = p$, the measured output y is given by the relative position errors between the neighboring vehicles, and

$$A = O, \qquad B_1 = B_2 = I, \qquad C = \begin{bmatrix} C_f \\ C_f^T \end{bmatrix}, \qquad F = \begin{bmatrix} F_f & F_b \end{bmatrix}. \tag{VP1}$$

For the double-integrator model (9.2), the state vector is $x = [\,p^T \; v^T\,]^T$, the measured output y is given by the relative position errors between the neighboring vehicles and the absolute velocity errors, and

$$A = \begin{bmatrix} O & I \\ O & O \end{bmatrix}, \quad B_1 = B_2 = \begin{bmatrix} O \\ I \end{bmatrix}, \quad C = \begin{bmatrix} C_f & O \\ C_f^T & O \\ O & I \end{bmatrix}, \quad F = \begin{bmatrix} F_f & F_b & F_v \end{bmatrix}. \tag{VP2}$$

Here, O and I denote the zero and identity matrices, and $\{F_f,\, F_b,\, F_v,\, C_f\}$ are defined in (9.3).

Upon closing the loop, we have

$$\begin{aligned} \dot{x} &= (A - B_2FC)\,x + B_1 d \\ z &= \begin{bmatrix} Q^{1/2}\,x \\ r^{1/2}\,u \end{bmatrix} = \begin{bmatrix} Q^{1/2} \\ -r^{1/2}FC \end{bmatrix} x \end{aligned}$$

where z encompasses the penalty on both the state and the control. Here, Q is a symmetric positive semi-definite matrix and r is a positive scalar. The objective is to design the *structured feedback gain* F such that the influence of the white stochastic disturbance d, with zero mean and unit variance, on the performance output z is minimized (in the $\mathcal{H}_2$ sense). This control problem can be formulated as

$$\begin{aligned} &\text{minimize} \quad J = \operatorname{trace}\left(PB_1B_1^T\right) \\ &\text{subject to} \;\; (A - B_2FC)^T P + P(A - B_2FC) = -(Q + r\,C^TF^TFC), \quad F \in \mathcal{S} \end{aligned} \tag{9.4}$$

where $\mathcal{S}$ denotes the structural subspace that F belongs to.

The necessary conditions for optimality of (9.4) are given by the set of coupled matrix equations in F, P, and L

$$(A - B_2FC)^T\,P + P\,(A - B_2FC) = -\left(Q + r\,C^TF^TFC\right) \tag{NC-1}$$

$$(A - B_2FC)\,L + L\,(A - B_2FC)^T = -B_1B_1^T \tag{NC-2}$$

$$(rFCLC^T) \circ I_{\mathcal{S}} = (B_2^TPLC^T) \circ I_{\mathcal{S}}. \tag{NC-3}$$

Here, P and L are the closed-loop observability and controllability Gramians, $\circ$ denotes the elementwise multiplication of two matrices, and the matrix $I_{\mathcal{S}}$ in (NC-3) denotes the

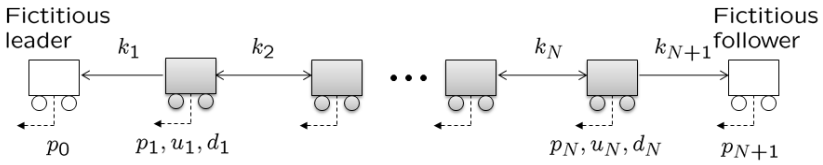


Figure 9.3: Formation of vehicles with localized symmetric gains.

structural identity of the subspace $\mathcal{S}$ under the elementwise multiplication $F \circ I_{\mathcal{S}} = F$, with $I_{\mathcal{S}} = [\, I \quad I \,]$ for the single-integrator model and $I_{\mathcal{S}} = [\, I \quad I \quad I \,]$ for the double-integrator model. For example, $[\, F_f \quad F_b \,] \circ [\, I \quad I \,] = [\, F_f \quad F_b \,]$. In the absence of the fictitious follower, an additional constraint $b_N = 0$ is imposed in (9.4) and thus, the structural identity for the single- and the double-integrator models are given by $[\, I \;\; I_z \,]$ and $[\, I \;\; I_z \;\; I \,]$, respectively. Here, I_z is a diagonal matrix with its main diagonal given by $[\, 1 \cdots 1 \; 0 \,]$.

9.2 Design of symmetric gains for the single-integrator model: A convex problem

In this section, we design the optimal symmetric feedback gains for the single-integrator model; see Fig. 9.3. This is a special case of the localized design, obtained by restricting the forward and the backward gains between the neighboring vehicles to be equal to each other $f_n = b_{n-1}$ for $n \in \{2, \ldots, N\}$. Under this assumption, we show that (9.4) is a convex optimization problem for the single-integrator model. This implies that the *global minimum* can be computed efficiently. Furthermore, in the absence of the fictitious follower, we provide *analytical expressions* for the optimal feedback gains.

Let us denote $k_1 = f_1$ and $k_{N+1} = b_N$ and let

$$k_n \;=\; f_n \;=\; b_{n-1}, \qquad n \in \{2, \ldots, N\}. \tag{9.5}$$

For the single-integrator model, the structured gain becomes a *symmetric* tridiagonal matrix

$$\begin{aligned} K \;&=\; F_f\, C_f \;+\; F_b\, C_f^T \\ &= \begin{bmatrix} k_1 + k_2 & -k_2 & & & \\ -k_2 & k_2 + k_3 & \ddots & & \\ & \ddots & \ddots & -k_N & \\ & & -k_N & k_N + k_{N+1} \end{bmatrix}. \end{aligned} \tag{9.6}$$

Consequently, $A_{\text{cl}} = -K$ is Hurwitz if and only if K is positive definite, in which case the Lyapunov equation in (9.4) simplifies to

$$KP \;+\; PK \;=\; Q \;+\; rKK.$$

The application of [70, Lemma 1] transforms the problem (9.4) of optimal symmetric design for the single-integrator model to

$$\begin{aligned} &\underset{K}{\text{minimize}} \quad J(K) \;=\; \frac{1}{2}\,\text{trace}\left(QK^{-1} \;+\; rK\right) \\ &\text{subject to} \;\; K \;\succ\; 0 \;\text{and}\; K \in \mathcal{S}_K \end{aligned} \tag{9.7}$$

where $K \in \mathcal{S}_K$ is a linear structural constraint given by (9.6). Specifically, $K = F_f C_f + F_b C_f^T$ is a symmetric tridiagonal matrix with the linear constraint (9.5). By introducing an auxiliary variable $X = X^T \geq Q^{1/2} K^{-1} Q^{1/2}$, we can formulate (9.7) as an SDP in X and K

$$\begin{aligned} &\underset{X,\,K}{\text{minimize}} \quad \frac{1}{2}\,\text{trace}\,(X \;+\; rK) \\ &\text{subject to} \quad K \;\succ\; 0, \quad K \in \mathcal{S}_K, \quad \begin{bmatrix} K & Q^{1/2} \\ Q^{1/2} & X \end{bmatrix} \;\succeq\; 0 \end{aligned}$$

which can be solved using available SDP solvers. Here, we have used the Schur complement in conjunction with $K > 0$ to express $X \geq Q^{1/2} K^{-1} Q^{1/2}$ as an LMI.

Next, we exploit the structure of K to express J in (9.7) with $Q = I$ in terms of the feedback gains $\{k_n\}_{n=1}^{N+1}$ between the neighboring vehicles. Since the inverse of the symmetric tridiagonal matrix K can be determined analytically [71, Theorem 2.3], the ijth entry of K^{-1} is given by

$$(K^{-1})_{ij} \;=\; \frac{\gamma_i\,(\gamma_{N+1}-\gamma_j)}{\gamma_{N+1}}, \quad j \geq i, \quad \text{where} \quad \gamma_i \;=\; \sum_{n=1}^{i} \frac{1}{k_n} \tag{9.8}$$

yielding the following expression for J

$$\begin{aligned} J \;&=\; \frac{1}{2}\,\mathrm{trace}\left(K^{-1} \,+\, rK\right) \\ &=\; \frac{1}{2}\sum_{n=1}^{N} \frac{\gamma_n(\gamma_{N+1}-\gamma_n)}{\gamma_{N+1}} \,+\, r\left(\frac{k_1+k_{N+1}}{2} \,+\, \sum_{n=2}^{N} k_n\right). \end{aligned}$$

The above expression for J is well-defined for $\{k_n\}_{n=1}^{N+1}$ that guarantee positive definiteness of K in (9.6); this is because the closed-loop A-matrix is determined by $A_{\mathrm{cl}} = -K$. The global minimizer of J can be computed using the gradient method [43].

For the formations without the fictitious follower, we next derive *explicit analytical expression* for the global symmetric minimizer $K = K^T > 0$ of (9.7) with $Q = I$. In this case $k_{N+1} = 0$ and the ijth entry of K^{-1} in (9.8) simplifies to $(K^{-1})_{ij} = \gamma_i$ for $j \geq i$. Consequently, the unique minimum of

$$\begin{aligned} J \;&=\; \frac{1}{2}\sum_{n=1}^{N} \gamma_n \,+\, r\left(\frac{k_1}{2} \,+\, \sum_{n=2}^{N} k_n\right) \\ &=\; \frac{1}{2}\sum_{n=1}^{N} \frac{N+1-n}{k_n} \,+\, r\left(\frac{k_1}{2} \,+\, \sum_{n=2}^{N} k_n\right) \end{aligned}$$

is attained for

$$k_1 \;=\; \sqrt{N/r}, \qquad k_n \;=\; \sqrt{(N+1-n)/(2r)}, \qquad n \in \{2,\ldots,N\}. \tag{9.9}$$

We also note that

$$\text{trace}\left(K^{-1}\right) \;=\; \sum_{n=1}^{N} \gamma_n \;=\; \sum_{n=1}^{N} \frac{N+1-n}{k_n} \;=\; r\left(k_1 \,+\, 2\sum_{n=2}^{N} k_n\right) \;=\; r\,\text{trace}\,(K) \tag{9.10}$$

where the third equality follows from (9.9).

Figure 9.4 shows the optimal symmetric gains for a formation with $N = 50$ vehicles, $Q = I$, and $r = 1$. Since the fictitious leader and the follower always move along their desired trajectories, the vehicles that are close to them have larger gains than the other vehicles. When no fictitious follower is present, the gains decrease monotonically from the first to the last vehicle; see ($\times$) in Fig. 9.4. In other words, the farther away the vehicle is from the fictitious leader the less weight it places on the information coming from its neighbors. When both the fictitious leader and the follower are present, the gains decrease as one moves from the boundary to the center of the formation; see ($\circ$) in Fig. 9.4.

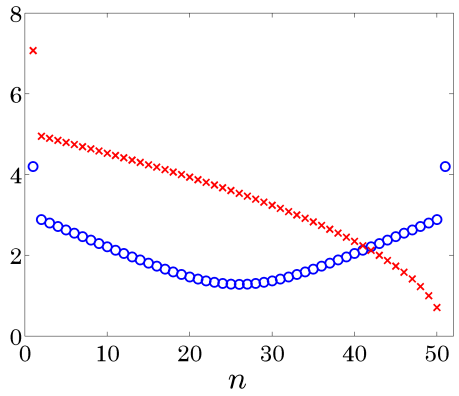


Figure 9.4: Optimal symmetric gains for formations with follower ($\circ$) and without follower ($\times$) for $N = 50$, $Q = I$, and $r = 1$. ($\times$) are obtained by evaluating formula (9.9) and ($\circ$) are computed using the gradient method.

9.3 Homotopy-based Newton's method

In this section, we remove the symmetric feedback gain restriction and utilize a homotopy-based Newton's method to solve (9.4). We solve a parameterized family of problems that ranges between an easily solvable problem and the problem of interest. In particular, we consider

$$Q(\varepsilon) \;=\; Q_0 \;+\; \varepsilon\,(Q_d \;-\; Q_0) \tag{9.11}$$

where Q_0 is the initial weight to be selected, Q_d is the *desired* weight, and $\varepsilon \in [0,1]$ is the homotopy parameter. Note that $Q = Q_0$ for $\varepsilon = 0$, and $Q = Q_d$ for $\varepsilon = 1$. The homotopy-based Newton's method consists of three steps: (i) For $\varepsilon = 0$, we find the initial weight Q_0 with respect to which a spatially uniform gain F_0 is *inversely optimal*. This is equivalent to solving problem (9.4) *analytically* with the performance weight Q_0. (ii) For $0 < \varepsilon \ll 1$, we employ perturbation analysis to determine the first few terms in the expansion $F(\varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n F_n$. (iii) For larger values of ε, we use Newton's method to solve (9.4). We gradually increase ε and use the structured optimal gain obtained for the previous value of ε to initialize the next round of iterations. This process is repeated until the desired value $\varepsilon = 1$ is reached.

In the remainder of this section, we focus on the single-integrator model. In Section 9.4, we solve problem (9.4) for the double-integrator model.

9.3.1 Spatially uniform symmetric gain: Inverse optimality for $\varepsilon = 0$

One of the simplest localized strategies is to use *spatially uniform gain*, where F_f and F_b are diagonal matrices with $f_n = f$ and $b_n = b$ for all n and some positive f and b. In particular, for $F_f = F_b = I$ it is easy to show closed-loop stability and to find the

performance weight Q_0 with respect to which the spatially uniform symmetric gain

$$K_0 \,=\, F_0\, C \,=\, \left[\begin{array}{cc} I & I \end{array}\right] \left[\begin{array}{c} C_f \\ C_f^T \end{array}\right] \,=\, T$$

is inversely optimal. The problem of inverse optimality amounts to finding the performance weight Q_0 for which an *a priori* specified K_0 is the corresponding optimal state-feedback gain [72, 73]. From linear quadratic regulator theory, the optimal state-feedback gain is given by $K_0 \,=\, R^{-1}B_2^T P_0$ where P_0 is the positive definite solution of

$$A^T P_0 \,+\, P_0 A \,+\, Q_0 \,-\, P_0 B_2 R^{-1} B_2^T P_0 \,=\, 0.$$

For the kinematic model (VP1), $A = O$ and $B_2 = I$, with $R = rI$, we have $K_0 = r^{-1}P_0$ and $Q_0 - r^{-1}P_0P_0 \,=\, 0$. Therefore, the state penalty $Q_0 \,=\, rK_0^2 \,=\, rT^2$ guarantees inverse optimality of the spatially uniform symmetric gain K_0. The above procedure of finding Q_0 can be applied to any structured gain F_0 that yields a symmetric positive definite K_0, e.g., the optimal symmetric gain of Section 9.2.

9.3.2 Perturbation analysis for $\varepsilon \ll 1$

We next utilize perturbation analysis to solve (9.4) with $Q(\varepsilon)$ given by (9.11) for $\varepsilon \ll 1$. For small ε, by representing P, L, and F as

$$P \,=\, \sum_{n\,=\,0}^{\infty} \varepsilon^n P_n, \qquad L \,=\, \sum_{n\,=\,0}^{\infty} \varepsilon^n L_n, \qquad F \,=\, \sum_{n\,=\,0}^{\infty} \varepsilon^n F_n$$

substituting in (NC-1)-(NC-3), and collecting same-order terms in ε, we obtain the set of equations (PA) with $A_0 := A - B_2F_0C$. Note that these equations are *conveniently coupled* in one direction, in the sense that for any $n \geq 1$, $O(\varepsilon^n)$ equations depend only on the solutions of $O(\varepsilon^m)$ equations for $m \leq n$. In particular, it is easy to verify that

$$
\begin{aligned}
O(1):&\begin{cases} A_0^T P_0 + P_0 A_0 &= -(Q_0 + r\, C^T F_0^T F_0 C) \\ A_0 L_0 + L_0 A_0^T &= -B_1 B_1^T \\ (r F_0 C L_0 C^T) \circ I_{\mathcal{S}} &= (B_2^T P_0 L_0 C^T) \circ I_{\mathcal{S}} \end{cases} \\
O(\varepsilon):&\begin{cases} A_0^T P_1 + P_1 A_0 &= -(Q_d - Q_0) \\ A_0 L_1 + L_1 A_0^T &= (B_2 F_1 C) L_0 + L_0 (B_2 F_1 C)^T \\ (r F_1 C L_0 C^T) \circ I_{\mathcal{S}} &= (B_2^T P_1 L_0 C^T) \circ I_{\mathcal{S}} \end{cases} \\
O(\varepsilon^2):&\begin{cases} A_0^T P_2 + P_2 A_0 &= (B_2 F_1 C)^T P_1 + P_1 (B_2 F_1 C) - r C^T F_1^T F_1 C \\ A_0 L_2 + L_2 A_0^T &= (B_2 F_1 C) L_1 + L_1 (B_2 F_1 C)^T + \\ & \quad (B_2 F_2 C) L_0 + L_0 (B_2 F_2 C)^T \\ (r F_2 C L_0 C^T) \circ I_{\mathcal{S}} &= (B_2^T P_1 L_1 C^T + B_2^T P_2 L_0 C^T - r F_1 C L_1 C^T) \circ I_{\mathcal{S}} \end{cases} \\
&\vdots \qquad\qquad\qquad\qquad \vdots
\end{aligned}
\tag{PA}
$$

---

the first and the third equations of $O(1)$ are satisfied with $K_0 = F_0 C = r^{-1} B_2^T P_0$ and with $Q_0 = r K_0^2$ identified in Section 9.3.1. Thus, the matrix L_0 can be obtained by solving the second equation of $O(1)$, and the matrices P_1, F_1, and L_1 can be obtained by solving the first, the third, and the second equations of $O(\varepsilon)$, respectively. The higher order terms F_n, P_n, and L_n can be determined in a similar fashion. The matrix F found by this procedure is the *unique optimal* solution of the control problem (9.4) for $\varepsilon \ll 1$. This is because the equations (PA), under the assumption of convergence for small ε, give a unique matrix $F(\varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n F_n$.

We next provide analytical expressions for $F_1 = [\, F_f^{(1)} \;\; F_b^{(1)} \,]$ obtained by solving the $O(\varepsilon)$ equations in (PA) with $r = 1$, $Q_0 = T^2$, and $Q_d = I$. When a fictitious follower is present, we have

$$
\begin{aligned}
f_n^{(1)} &= \frac{n(n-N-1)(4n(N+1) - N(2N+7) + 1)}{12\,(N^2-1)} - \frac{1}{2} \\
b_n^{(1)} &= \frac{n(N+1-n)(4n(N+1) - N(2N+1) - 5)}{12\,(N^2-1)} - \frac{1}{2}
\end{aligned}
\tag{9.12}
$$

where $f_n^{(1)}$ and $b_n^{(1)}$ denote the nth diagonal entries of $F_f^{(1)}$ and $F_b^{(1)}$. From (9.12) it follows that $f_n^{(1)} = b_{N+1-n}^{(1)}$ for $n \in \{1, \ldots, N\}$. When a fictitious follower is not present, we have

$$\begin{aligned} f_n^{(1)} &= \frac{1}{2}(-n^2 + (N+1)n - 1), \qquad n \in \{1, \ldots, N-1\} \\ f_N^{(1)} &= \frac{1}{2}(N-1), \\ b_n^{(1)} &= \frac{1}{2}(n^2 - Nn - 1), \qquad n \in \{1, \ldots, N-1\} \\ b_N^{(1)} &= 0. \end{aligned}$$

To compute the optimal structured feedback gain for larger values of ε, we use $F(\varepsilon)$ obtained from perturbation analysis to initialize Newton's method, as described in Section 9.3.3.

9.3.3 Newton's method for larger values of ε

In this section, we employ Newton's method to solve the necessary conditions for optimality (NC-1)-(NC-3) as ε is gradually increased to 1. Newton's method is an iterative descent algorithm for finding local minima in optimization problems [43]. Specifically, given an initial stabilizing structured gain F^0, a decreasing sequence of the objective function $\{J(F^i)\}$ is generated by updating F according to $F^{i+1} = F^i + s^i \tilde{F}^i$. Here, $\tilde{F}^i$ is the Newton direction that satisfies the structural constraint and s^i is the step-size.

For small ε, we initialize Newton's method using $F(\varepsilon)$ obtained from the perturbation expansion up to the first order in ε, $F(\varepsilon) = F_0 + \varepsilon F_1$. We then increase ε slightly and use the optimal structured gain resulting from Newton's method at the previous ε to initialize the next round of iterations. We continue increasing ε gradually until desired value $\varepsilon = 1$ is reached, that is, until the optimal structured gain F for the desired Q_d is obtained.

Since the homotopy-based Newton's method solves a family of optimization problems

parameterized by ε, the optimal feedback gain is a function of $\varepsilon \in [0,1]$. To see the incremental change relative to the spatially uniform gain F_0, we consider the difference between the optimal forward gain $f_n(\varepsilon)$ and the uniform gain $f_n(0) = 1$,

$$\tilde{f}_n(\varepsilon) := f_n(\varepsilon) - f_n(0) = f_n(\varepsilon) - 1.$$

Figure 9.5a shows the normalized profile $\tilde{f}(\varepsilon)/\|\tilde{f}(\varepsilon)\|$ for a formation with fictitious follower, $N = 50$, $r = 1$, $Q_0 = T^2$, and $Q_d = I$. The values of ε are determined by 20 logarithmically spaced points between 10^{-4} and 1. As ε increases, the normalized profile changes from an almost sinusoidal shape (cf. analytical expression in (9.12)) at $\varepsilon = 10^{-4}$ to an almost piecewise linear shape at $\varepsilon = 1$. Note that the homotopy-based Newton's method converges to the same feedback gains at $\varepsilon = 1$ when it is initialized by the optimal symmetric controller obtained in Section 9.2.

Since the underlying path-graph exhibits symmetry between the edge pairs associated with f_n and b_{N+1-n}, the optimal forward and backward gains satisfy a *central symmetry* property,

$$f_n = b_{N+1-n}, \qquad n \in \{1, \ldots, N\}$$

for all $\varepsilon \in [0,1]$; see Fig. 9.5b for $\varepsilon = 1$. We note that the first vehicle has a larger forward gain than other vehicles; this is because it neighbors the fictitious leader. The forward gains decrease as one moves away from the fictitious leader; this is because information about the absolute desired trajectory of the fictitious leader becomes less accurate as it propagates down the formation. Similar interpretation can be given to the optimal backward gains, which monotonically increase as one moves towards the fictitious follower.

Since the 1st vehicle has a *negative* backward gain (see Fig. 9.5b), if the distance between the 1st and the 2nd vehicles is greater than the desired value δ, then the 1st

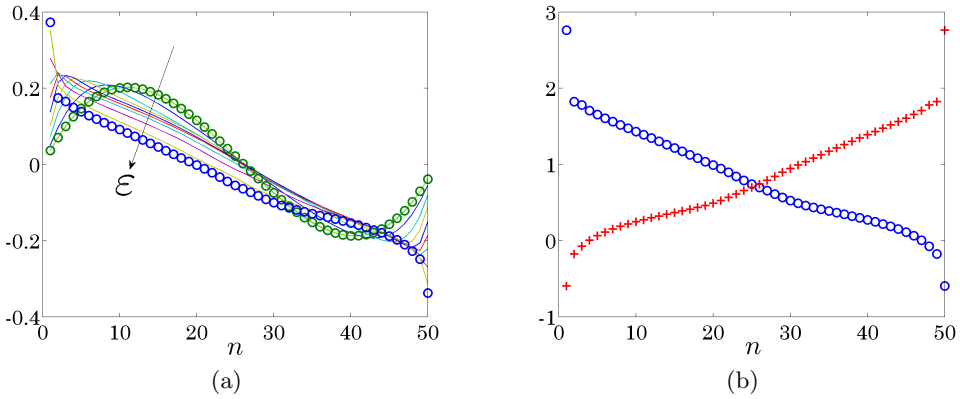


Figure 9.5: Formation with fictitious follower, $N = 50$, $r = 1$, $Q_0 = T^2$, and $Q_d = I$. (a) Normalized optimal forward gain $\tilde{f}(\varepsilon)/\|\tilde{f}(\varepsilon)\|$ changes from an almost sinusoidal shape (cf. analytical expression in (9.12)) at $\varepsilon = 10^{-4}$ to an almost piecewise linear shape at $\varepsilon = 1$. (b) Optimal forward ($\circ$) and backward ($+$) gains at $\varepsilon = 1$.

vehicle distances itself even further from the 2nd vehicle. On the other hand, if the distance is less than δ, then the 1st vehicle pulls itself even closer to the 2nd vehicle. This negative backward gain of the 1st vehicle can be interpreted as follows: Since the 1st vehicle has access to its global position, it aims to correct the absolute positions of other vehicles in order to enhance formation coherence. If the 2nd vehicle is too close to the 1st vehicle, then the 1st vehicle moves towards the 2nd vehicle to push it back; this in turn pushes other vehicles back. If the 2nd vehicles is too far from the 1st vehicle, then the 1st vehicle moves away from the 2nd vehicle to pull it forward; this in turn pulls other vehicles forward. Similar interpretation can be given to the negative forward gain of the Nth vehicle that neighbors the fictitious follower. Also note that the forward gain of the Nth vehicle becomes *positive* when the fictitious follower is *removed* from the formation; see Fig. 9.6c. This perhaps suggests that negative feedback gains of the 1st and the Nth vehicles are a consequence of the fact that both of them have access to their own global positions.

As shown in Figs. 9.6a and 9.6b, the normalized optimal gains for the formation

without the fictitious follower also change continuously as ε increases to 1. In this case, however, the optimal forward and backward gains do not satisfy the central symmetry; see Fig. 9.6c. Since the optimal controller puts more emphasis on the vehicles ahead when the fictitious follower is *not* present, the forward gains have larger magnitudes than the backward gains. As in the formations with the fictitious follower, the optimal forward gains decrease monotonically as one moves away from the fictitious leader. On the other hand, the optimal backward gains at first increase as one moves away from the 1st vehicle and then decrease as one approaches the Nth vehicle in order to satisfy the constraint $b_N = 0$.

9.4 Optimal localized design for the double-integrator model

In this section, we solve (9.4) for the double-integrator model using the homotopy-based Newton's method. For a formation in which each vehicle – in addition to relative positions with respect to its immediate neighbors – has access to *its own velocity*, our results highlight similarity between optimal forward and backward position gains for the single- and the double-integrator models. We further show that the performance measures exhibit similar scaling properties to those found in single-integrators. We also establish convexity of (9.4) for the double-integrator model by restricting the controller to symmetric position and uniform diagonal velocity gains.

The perturbation analysis and the homotopy-based Newton's method closely follow the procedure described in Sections 9.3.2 and 9.3.3, respectively. In particular, $F_0 = [\,\alpha I \quad \alpha I \quad \beta I\,]$ yields $K_0 = F_0\, C = [\,\alpha T \quad \beta I\,]$. It is readily shown that, for positive α and β with $\beta^2 > 8\alpha$, this spatially uniform structured feedback gain is stabilizing and

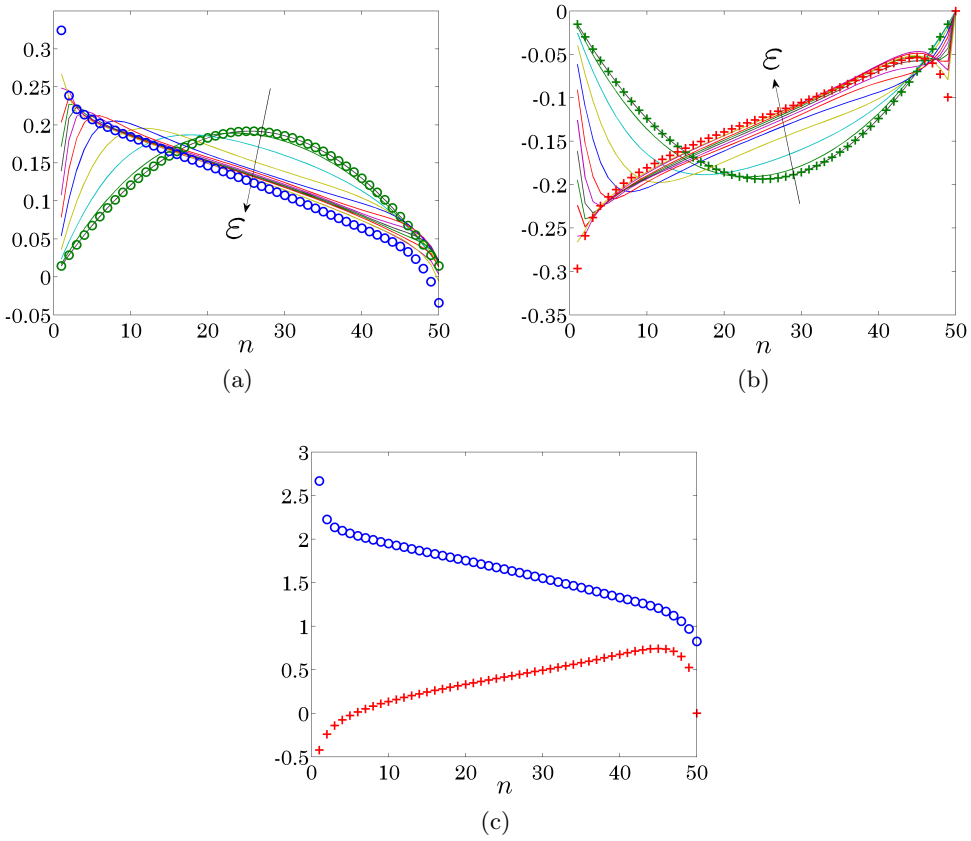


Figure 9.6: Formation without fictitious follower, $N = 50$, $r = 1$, $Q_0 = T^2$, and $Q_d = I$. Normalized optimal (a) forward and (b) backward gains. (c) Optimal forward ($\circ$) and backward ($+$) gains at $\varepsilon = 1$.

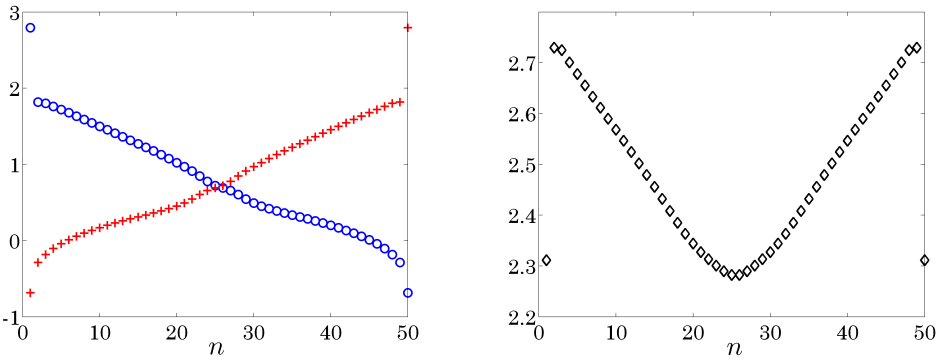


Figure 9.7: Double-integrator model with fictitious follower, $N = 50$, $Q = I$ and $r = 1$. (a) The optimal forward ($\circ$) and backward gains ($+$); (b) the optimal velocity gains ($\diamond$).

inversely optimal with respect to

$$Q_0 \;=\; \begin{bmatrix} Q_p & O \\ O & Q_v \end{bmatrix}, \qquad Q_p \;=\; r\alpha^2\, T^2, \qquad Q_v \;=\; r(\beta^2\, I - 2\,\alpha\, T), \qquad r \;>\; 0.$$

In what follows, we choose $\alpha = 1$ and $\beta = 3$ and employ the homotopy-based Newton's method to solve (9.4) for the double-integrator model. For a formation with fictitious follower, $N = 50$, $Q = I$, and $r = 1$ the optimal forward and backward position gains are shown in Fig. 9.7a and the optimal velocity gains are shown in Fig. 9.7b. We note remarkable similarity between the optimal position gains for the single- and the double-integrator models; cf. Fig. 9.7a and Fig. 9.5b. For a formation without fictitious follower, the close resemblance between the optimal position gains for both models is also observed.

For the double-integrator model with $K = FC = [\, K_p \;\; \beta I \,]$ and fixed $\beta > 0$ we next show convexity of (9.4) with respect to $K_p = K_p^T > 0$. The Lyapunov equation in (9.4), for the block diagonal state weight Q with components Q_1 and Q_2, can be rewritten in

terms of the components of P with

$$P \;=\; \begin{bmatrix} P_1 & P_0 \\ P_0^T & P_2 \end{bmatrix}$$

that is,

$$K_p P_0^T \;+\; P_0 K_p \;=\; Q_1 \;+\; K_p K_p \tag{9.13a}$$

$$K_p P_2 \;-\; P_1 \;+\; \beta P_0 \;=\; \beta K_p \tag{9.13b}$$

$$2\beta P_2 \;=\; P_0 \;+\; P_0^T \;+\; Q_2 \;+\; \beta^2 I. \tag{9.13c}$$

Linearity of the trace operator in conjunction with $B_1 = [\,O \;\; I\,]^T$ and (9.13c) yields

$$\begin{aligned} J \;&=\; \text{trace}\,(P_2) \;=\; \frac{1}{2\beta}\,\text{trace}\,\big(2P_0 + Q_2 + \beta^2 I\big) \\ &=\; \frac{1}{2\beta}\,\text{trace}\,\big(K_p^{-1} Q_1 + K_p + Q_2 + \beta^2 I\big) \end{aligned}$$

where the last equation is obtained by multiplying (9.13a) from the left with K_p^{-1} and using $\text{trace}\,(K_p^{-1} P_0 K_p) = \text{trace}\,(P_0)$. For $Q_1 \geq 0$, similar argument as in Section 9.2 can be used to conclude convexity of J with respect to $K_p = K_p^T > 0$.

Chapter 10

Performance of vehicular formations with nearest neighbor interactions

We consider how the performance of the optimally-controlled vehicular formation scales with the number of vehicles. We consider both macroscopic and microscopic performance measures based on whether attention is paid to the absolute position error of each vehicle or the relative position error between neighboring vehicles. The macroscopic performance measure quantifies the resemblance of the formation to a rigid lattice. This performance measure is also referred to as the *coherence* of the formation [37]. It was shown in [37] that, for *localized symmetric feedback* with uniform bounds on control energy at each vehicle, it is impossible to have large coherent formations that behave like rigid lattice. This was established by exhibiting linear scaling $O(N)$ of macroscopic performance measure with the number of vehicles N. It is thus of interest to examine the coherence of vehicular formations with the optimal localized feedback design introduced in Chapter 9.

When the control penalty in the quadratic performance objective is formation-size-independent, it turns out that the optimal symmetric and non-symmetric controllers asymptotically provide $O(\sqrt{N})$ and $O(\sqrt[4]{N})$ scalings of formation coherence. However, these controllers require similar growth with N of the control energy per vehicle. We show that bounded control energy can be obtained by judicious selection of an N-dependent control penalty, leading to $O(N)$ and $O(\sqrt{N})$ scalings of formation coherence for the optimal symmetric and non-symmetric controllers, respectively. These results are summarized in Table 10.1 and they hold for both single- and double-integrators for formations in which each vehicle has access to *its own velocity*; see Sections 10.2 and 10.3 for details.

In addition, we provide a *spatially uniform non-symmetric* controller that outperforms the *optimal spatially varying symmetric* controller in the scaling trend of macroscopic performance measure. This indicates that departure from symmetric gains can improve coherence of large-scale formations and that the controller structure may play a more important role than the optimal selection of feedback gains. Our results show that the localized controller that achieves the best performance is both *non-symmetric and spatially-varying*.

This chapter is organized as follows. In Section 10.1, we introduce the microscopic and macroscopic performance measures that quantify the local and global position errors of vehicular formations. In Section 10.2, we examine the performance of spatially uniform symmetric and non-symmetric controllers, and optimal symmetric and non-symmetric controllers for the single-integrator model. In Section 10.3, we consider the performance of localized controllers for the double-integrator model.

Table 10.1: Summary of asymptotic scalings with the number of vehicles N for the optimal symmetric and non-symmetric position gains. The N-independent control penalty, $R = r_0 I$, in the quadratic performance objective leads to similar growth with N of formation coherence and control energy (per vehicle). On the other hand, the N-dependent control penalty that provides bounded control energy yields less favorable coherence.

Optimal position gains	**Control penalty** $R = rI$	**Control energy (per vehicle)**	**Formation coherence**
symmetric	$r(N) = r_0$	$O(\sqrt{N})$	$O(\sqrt{N})$
non-symmetric	$= \text{const.}$	$O(\sqrt[4]{N})$	$O(\sqrt[4]{N})$
symmetric	$r(N) \sim N$	$O(1)$	$O(N)$
non-symmetric	$r(N) \sim \sqrt{N}$	$O(1)$	$O(\sqrt{N})$

10.1 Performance of optimal localized controller

To evaluate the performance of the optimal localized controller F, obtained by solving (9.4) with $Q = I$, we consider the closed-loop system

$$\begin{aligned} \dot{x} &= (A - B_2 F C)\, x + B_1 d \\ \zeta &= \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} Q_s^{1/2} \\ -FC \end{bmatrix} x, \qquad s = g \;\text{ or }\; s = l, \end{aligned} \tag{10.1}$$

where ζ_1 is the global or local performance output and ζ_2 is the control input. Motivated by [37], we examine two state performance weights for the single-integrator model

- Macroscopic (global): $Q_g = I$,
- Microscopic (local): $Q_l = T$,

where T is an $N \times N$ symmetric Toeplitz matrix with its first row given by

$$[\, 2 \;\; -1 \;\; 0 \;\; \cdots \;\; 0 \,] \in \mathbb{R}^N.$$

For example, for $N = 4$,

$$T = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}. \tag{10.2}$$

The macroscopic performance weight $Q_g = I$ penalizes the *global* (absolute) position errors,

$$\zeta_1^T \zeta_1 = p^T Q_g \, p = \sum_{n=1}^{N} p_n^2$$

and the microscopic performance weight $Q_l = T$ penalizes the *local* (relative) position errors,

$$\zeta_1^T \zeta_1 = p^T Q_l \, p = \sum_{n=0}^{N} (p_n - p_{n+1})^2$$

with $p_0 = p_{N+1} = 0$. These state weights induce the macroscopic and microscopic performance measures determined by the *formation-size-normalized* $\mathcal{H}_2$ norm

$$\Pi_s(N) = \frac{1}{N} \|G_1\|_2^2, \qquad s = g \ \text{ or } \ s = l,$$

where G_1 is the transfer function of (10.1) from d to ζ_1. On the other hand, the microscopic performance measure Π_l quantifies how well regulated the distances between the neighboring vehicles are. We will also examine the formation-size-normalized control energy (variance) of the closed-loop system (10.1),

$$\Pi_{\text{ctr}}(N) = \frac{1}{N} \|G_2\|_2^2,$$

which is determined by the $\mathcal{H}_2$ norm of the transfer function G_2 from d to $\zeta_2 = u$.

Similarly, for the double-integrator model, we use the following performance weights

- Macroscopic (global), $Q_g = \begin{bmatrix} I & O \\ O & I \end{bmatrix}$,
- Microscopic (local), $Q_l = \begin{bmatrix} T & O \\ O & I \end{bmatrix}$.

10.2 Performance vs. size for the single-integrator model

In this section, we study the performance of the optimal symmetric and non-symmetric gains obtained in Sections 9.2 and 9.3.3. This is done by examining the dependence on the formation size of performance measures Π_g, Π_l, and Π_{ctr} introduced in Section 10.1. Our results highlight the role of non-symmetry and spatial variations on the scaling trends in large-scale formations. They also illustrate performance improvement achieved by the optimal controllers relative to spatially uniform symmetric and non-symmetric feedback gains.

For the spatially uniform *symmetric* gain with $\{f_n, b_n\}$ in (9.1b) being $f_n = b_n = \alpha > 0$, we show *analytically* that Π_g is an *affine* function of N. This implies that the formation coherence scales linearly with N irrespective of the value of α. We also *analytically* establish that the spatially uniform *non-symmetric* gain with $\{f_n = \alpha > 0, b_n = 0\}$ (look-ahead strategy) provides a *square-root* asymptotic dependence of Π_g on N. Thus, symmetry breaking between the forward and backward gains may improve coherence of large-scale formations. We then investigate how *spatially varying* optimal feedback gains, introduced in Sections 9.2 and 9.3.3, influence coherence of the formation. We show that the optimal symmetric gain provides a *square-root* dependence of Π_g on N and that the optimal non-symmetric gain provides a *fourth-root* dependence of Π_g on N.

Even though we are primarily interested in asymptotic scaling of the global performance measure Π_g, we also examine the local performance measure Π_l and the control energy Π_{ctr}. For the single-integrator model (VP1) in Chapter 9, the closed-loop system (10.1) simplifies to

$$\begin{aligned} \dot{x} &= -FC\,x + d \\ \zeta &= \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} Q_s^{1/2} \\ -FC \end{bmatrix} x, \qquad Q_s = \begin{cases} I, & s = g \\ T, & s = l. \end{cases} \end{aligned}$$

Then

$$\begin{aligned} \Pi_s &= \frac{1}{N}\,\text{trace}\,(L\,Q_s) \\ \Pi_{\text{ctr}} &= \frac{1}{N}\,\text{trace}\,\big(L\,C^T F^T FC\big) \end{aligned} \tag{10.3}$$

where L denotes the closed-loop controllability Gramian,

$$(-FC)\,L + L\,(-FC)^T = -I. \tag{10.4}$$

The asymptotic scaling properties of Π_g, Π_l, and Π_{ctr}, for the above mentioned spatially uniform controllers and the spatially varying optimal controllers, obtained by solving (9.4) with $Q = I$ and $r = 1$, are summarized in Table 10.2. For both spatially uniform symmetric and look-ahead strategies, we analytically determine the dependence of these performance measures on the formation size in Sections 10.2.1 and 10.2.2. Furthermore, for the formation without the fictitious follower subject to the optimal symmetric gains, we provide analytical results in Section 10.2.3. For the optimal symmetric and non-symmetric gains in the presence of fictitious followers, the scaling trends are obtained with the aid of numerical computations in Section 10.2.3.

Several comments about the results in Table 10.2 are given next. First, in contrast to

Table 10.2: Asymptotic dependence of Π_g, Π_l, and Π_{ctr} on the formation size N for uniform symmetric, uniform non-symmetric (look-ahead strategy), and optimal symmetric and non-symmetric gains of Sections 9.2 and 9.3.3 with $Q = I$ and $r = 1$. The results in the first three rows are determined *analytically*; the scalings in the last two rows are estimated based on numerical computations.

Controller	Π_g	Π_l	Π_{ctr}
uniform symmetric with/without follower	$O(N)$	$O(1)$	$O(1)$
uniform non-symmetric	$O(\sqrt{N})$	$O(1)$	$O(1)$
optimal symmetric without follower	$O(\sqrt{N})$	$O(1/\sqrt{N})$	$O(\sqrt{N})$
optimal symmetric with follower	$O(\sqrt{N})$	$O(1/\sqrt{N})$	$O(\sqrt{N})$
optimal non-symmetric with/without follower	$O(\sqrt[4]{N})$	$O(1/\sqrt[4]{N})$	$O(\sqrt[4]{N})$

the spatially uniform controllers, the optimal symmetric and non-symmetric gains, resulting from an N-independent control penalty r in (9.4), do not provide uniform bounds on the control energy per vehicle, Π_{ctr}. This implies the trade-off between the formation coherence Π_g and control energy Π_{ctr} in the design of the optimal controllers. It is thus of interest to examine formation coherence for optimal controllers with bounded control energy per vehicle (see Remark 4). Second, the controller structure (e.g., symmetric or non-symmetric gains) plays an important role in the formation coherence. In particular, departure from symmetry in localized feedback gains can significantly improve coherence of large-scale formations (see Remark 5).

10.2.1 Spatially uniform symmetric gain

For the spatially uniform symmetric controller with $f_n = b_n = \alpha > 0$, we next show that Π_g is an affine function of N and that, in the limit of an infinite number of vehicles, both Π_l and Π_{ctr} become formation-size-independent. These results hold irrespective of the presence of the fictitious follower.

For the single-integrator model with the fictitious follower we have $K = FC = \alpha T$

(see (10.2) for the definition of T), and $L = T^{-1}/(2\alpha)$ solves the Lyapunov equation (10.4) [70, Lemma 1]. Since the nth diagonal entry of T^{-1} is determined by

$$(T^{-1})_{nn} = \frac{n\,(N+1-n)}{N+1}$$

from (10.3) we conclude that the global performance measure Π_g is an affine function of N, and that both Π_l and Π_{ctr} are formation-size-independent,

$$\begin{aligned}
\Pi_g &= \frac{1}{2\alpha N}\,\text{trace}\left(T^{-1}\right) \\
&= \frac{1}{2\alpha N}\sum_{n=1}^{N} n \;-\; \frac{1}{2\alpha N(N+1)}\sum_{n=1}^{N} n^2 \;=\; \frac{N+2}{12\alpha} \\
\Pi_l &= \frac{1}{2\alpha N}\,\text{trace}\left(T\,T^{-1}\right) \;=\; \frac{1}{2\alpha} \\
\Pi_{\text{ctr}} &= \frac{1}{2\alpha N}\,\text{trace}\left(\alpha^2\,T\,T\,T^{-1}\right) \;=\; \alpha.
\end{aligned}$$

For the formation without the fictitious follower, the following expressions

$$\Pi_g = \frac{N+1}{4\alpha}, \qquad \Pi_l = \frac{1}{\alpha}, \qquad \Pi_{\text{ctr}} = \frac{\alpha(3N+1)}{2N}$$

imply that, for the spatially uniform symmetric controller, the asymptotic scaling trends do not depend on the presence of the fictitious follower.

10.2.2 Spatially uniform non-symmetric gain (look-ahead strategy)

We next examine the asymptotic scaling of the performance measures for the spatially uniform non-symmetric gain with $\{f_n = \alpha > 0,\, b_n = 0\}$. We establish the *square-root* scaling of Π_g with N and the formation-size-independent scaling of Π_l. Furthermore, in the limit of an infinite number of vehicles, we show that Π_{ctr} becomes N-independent.

For the single-integrator model with $K = FC = \alpha C_f$ (see (9.3) for the definition of

C_f), the solution of the Lyapunov equation (10.4) is given by

$$L \;=\; \int_0^\infty \mathrm{e}^{-\,\alpha\, C_f\, t}\, \mathrm{e}^{-\,\alpha\, C_f^T\, t}\, \mathrm{d}t. \tag{10.5}$$

As shown below, the inverse Laplace transform of $(sI + \alpha C_f)^{-1}$ can be used to determine the analytical expression for $\mathrm{e}^{-\,\alpha\, C_f\, t}$, yielding the following formulae,

$$\begin{aligned} \Pi_g(N) \;&=\; \frac{1}{N}\sum_{n\,=\,1}^{N} L_{nn} \;=\; \frac{1}{N}\sum_{n\,=\,1}^{N} \frac{\alpha\,\Gamma(n+1/2)}{\sqrt{\pi}\,\Gamma(n)} \;=\; \frac{2\,\alpha\,\Gamma(N+3/2)}{3\,\sqrt{\pi}\,\Gamma(N+1)} \\ \Pi_l \;&=\; \alpha \\ \Pi_{\mathrm{ctr}} \;&=\; \alpha \;-\; \frac{1}{N}\, L_{NN} \end{aligned}$$

with $\Gamma(\cdot)$ denoting the Gamma function.

We next show that, in the limit of an infinite number of vehicles, a look-ahead strategy for the single-integrator model provides the square-root dependence of Π_g on N and the formation-size-independent Π_l and Π_{ctr}. The solution of the Lyapunov equation (10.4) with $FC = \alpha C_f$ is determined by (10.5). Since the ith entry of the first column of the lower triangular Toeplitz matrix $(sI + \alpha C_f)^{-1}$ is $\alpha^i/(s+\alpha)^i$, the corresponding entry of the matrix exponential in (10.5) is determined by the inverse Laplace transform of $\alpha^i/(s+\alpha)^i$, that is,

$$\frac{\alpha\,(\alpha t)^{i-1}\,\mathrm{e}^{-\alpha t}}{(i-1)!}.$$

Thus, the nth element on the main diagonal of the matrix L in (10.5) is given by

$$L_{nn} \;=\; \int_0^\infty \sum_{i\,=\,1}^{n} \left(\alpha \mathrm{e}^{-\alpha t}\, \frac{(\alpha t)^{i-1}}{(i-1)!}\right)^2 \mathrm{d}t \;=\; \frac{\alpha\,\Gamma(n+1/2)}{\sqrt{\pi}\,\Gamma(n)} \;=\; \frac{\alpha\,(2n)!}{2^{2n}(n-1)!\,n!} \tag{10.6}$$

thereby yielding

$$\Pi_g \;=\; \sum_{n=1}^{N} \frac{L_{nn}}{N} \;=\; \frac{2\,\alpha\,\Gamma(N+3/2)}{3\,\sqrt{\pi}\,\Gamma(N+1)} \;=\; \frac{2}{3}\frac{\alpha\,(2N+2)!}{2^{2N+2}N!(N+1)!}. \tag{10.7}$$

A similar procedure can be used to show that the $n(n+1)$th entry of L is determined

$$L_{n(n+1)} \;=\; L_{(n+1)(n+1)} \;-\; \frac{\alpha}{2}, \qquad n \;=\; 1,\ldots,N-1. \tag{10.8}$$

Now, from (10.8) and the fact that $L_{11} = \alpha/2$, we obtain

$$\Pi_l \;=\; \frac{1}{N}\,\mathrm{trace}\,(TL) \;=\; \frac{2}{N}\left(\sum_{n=1}^{N} L_{nn} \;-\; \sum_{n=1}^{N-1} L_{n(n+1)}\right) \;=\; \alpha.$$

Similarly,

$$\begin{aligned}\Pi_{\mathrm{ctr}} \;&=\; \frac{1}{N}\,\mathrm{trace}\left(LC_f^T C_f\right)\\ &=\; \frac{2}{N}\left(\sum_{n=1}^{N} L_{nn} \;-\; \sum_{n=1}^{N-1} L_{n(n+1)}\right) \;-\; \frac{1}{N}\,L_{NN}\\ &=\; \alpha \;-\; \frac{1}{N}\,L_{NN}.\end{aligned}$$

Using Stirling's approximation $n! \approx \sqrt{2\pi n}\,(n/\mathrm{e})^n$ for large n, we have

$$\lim_{n\to\infty} \frac{L_{nn}}{\sqrt{n}} \;=\; \lim_{n\to\infty} \frac{\alpha}{\sqrt{\pi}}\sqrt{\frac{n}{n-1}}\left(\frac{n}{n-1}\right)^{n-1}\frac{1}{\mathrm{e}} \;=\; \frac{\alpha}{\sqrt{\pi}}$$

where we used the fact that

$$\lim_{n\to\infty}\left(\frac{n}{n-1}\right)^{n-1} \;=\; \mathrm{e}.$$

Consequently,

$$\lim_{N\to\infty} \Pi_{\mathrm{ctr}}(N) \;=\; \alpha.$$

From (10.6) and (10.7), it follows that $\Pi_g = (2/3)L_{(N+1)(N+1)}$ and thus,

$$\lim_{N\to\infty} \frac{\Pi_g(N)}{\sqrt{N}} = \frac{2\alpha}{3\sqrt{\pi}}.$$

We conclude that Π_g asymptotically scales as a square-root function of N and that Π_{ctr} is formation-size-independent as N increases to infinity.

10.2.3 Optimal symmetric and non-symmetric controllers

We next examine the asymptotic scaling of the performance measures for the optimal symmetric and non-symmetric gains of Sections 9.2 and 9.3.3. For the formation without the fictitious follower, we analytically establish that the optimal symmetric gains asymptotically provide $O(\sqrt{N})$, $O(1/\sqrt{N})$, and $O(\sqrt{N})$ scalings of Π_g, Π_l, and Π_{ctr}, respectively. We then use numerical computations to (i) confirm these scaling trends for the optimal symmetric gains in the presence of the fictitious follower; and to (ii) show a fourth-root dependence of Π_g and Π_{ctr} on N and an $O(1/\sqrt[4]{N})$ dependence of Π_l for the optimal non-symmetric gains. All these scalings are obtained by solving (9.4) with the formation-size-independent control penalty r and $Q = I$. We also demonstrate that uniform control variance (per vehicle) can be obtained by judicious selection of an N-dependent r. For the optimal symmetric and non-symmetric gains, this constraint on control energy (variance) increases the asymptotic dependence of Π_g on N to linear and square-root, respectively.

For the formation without the fictitious follower, the optimal symmetric gains are given by (9.9). As shown in (9.10), $\text{trace}\,(K^{-1}) = \text{trace}\,(rK)$, thereby yielding

$$\Pi_g \;=\; r\,\Pi_{\text{ctr}} \;=\; \frac{1}{2N}\,\text{trace}\left(K^{-1}\right) \;=\; \frac{\sqrt{r}}{2N}\left(\sqrt{N} + \sum_{n=1}^{N-1}\sqrt{2n}\right). \tag{10.9}$$

In the limit of an infinite number of vehicles,

$$\lim_{N\to\infty}\frac{\Pi_g(N)}{\sqrt{N}} \;=\; \lim_{N\to\infty}\sum_{n=1}^{N-1}\sqrt{\frac{rn}{2N}}\,\frac{1}{N} \;=\; \int_0^1\sqrt{\frac{rx}{2}}\,\mathrm{d}x \;=\; \sqrt{\frac{2r}{9}}$$

which, for an N-independent r, leads to an asymptotic square-root dependence of Π_g and Π_{ctr} on N,

$$\begin{aligned}\Pi_g(N) \;&=\; \sqrt{\frac{2\,rN}{9}}+\sqrt{\frac{r}{4N}},\quad N\gg 1\\ \Pi_{\text{ctr}}(N) \;&=\; \sqrt{\frac{2N}{9r}}+\frac{1}{\sqrt{4\,rN}},\quad N\gg 1.\end{aligned}\tag{10.10}$$

Similar calculation can be used to obtain $O(1/\sqrt{N})$ asymptotic scaling of Π_l.

We next use numerical computations to study the scaling trends for the optimal symmetric and non-symmetric gains in the presence of fictitious followers. The optimal symmetric gain (cf. ($\circ$) in Fig. 9.4) provides a *square-root* scaling of Π_g with N; see Fig. 10.1a. On the other hand, the optimal non-symmetric gain (cf. Fig. 9.5b) leads to a *fourth-root* scaling of Π_g with N; see Fig. 10.1b. The local performance measure Π_l decreases monotonically with N for both controllers, with Π_l scaling as $O(1/\sqrt{N})$ for the optimal symmetric gain and as $O(1/\sqrt[4]{N})$ for the optimal non-symmetric gain; see Fig. 10.2. For both the optimal symmetric and non-symmetric controllers, our computations indicate equivalence between the control energy and the global performance measure when $r = 1$. (For the optimal symmetric gain without the fictitious follower and $r = 1$, we have analytically shown that $\Pi_{\text{ctr}} = \Pi_g$; see formula (10.9).) Therefore, the asymptotic scaling of the formation-size-normalized control energy is $O(\sqrt{N})$ for the optimal symmetric gain and $O(\sqrt[4]{N})$ for the optimal non-symmetric gain. Finally, for the formations without the fictitious follower, our computations indicate that the optimal non-symmetric gains also asymptotically provide $O(\sqrt[4]{N})$, $O(1/\sqrt[4]{N})$, and $O(\sqrt[4]{N})$ scalings of Π_g, Π_l, and Π_{ctr}, respectively.

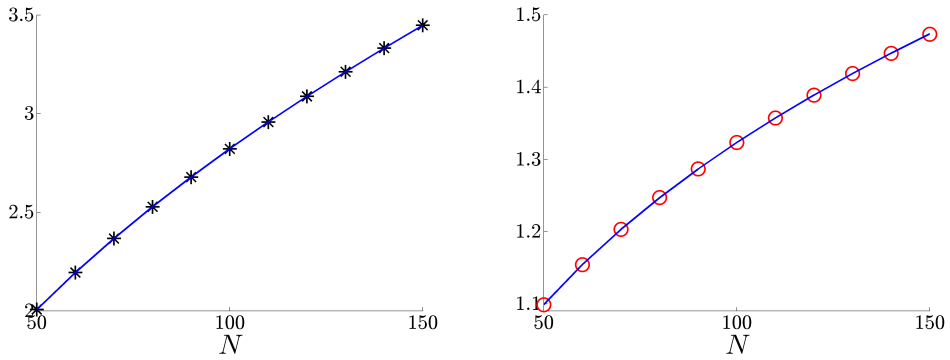


Figure 10.1: (a) Square-root scaling of Π_g $(*)$ using optimal symmetric gain of Section 9.2, $0.2784\sqrt{N} + 0.0375$ (curve); and (b) Fourth-root scaling of Π_g $(\circ)$ using optimal non-symmetric gain of Section 9.3.3, $0.4459\sqrt[4]{N} - 0.0866$ (curve). The optimal controllers are obtained by solving (9.4) with $Q = I$ and $r = 1$ for the formation with the fictitious follower.

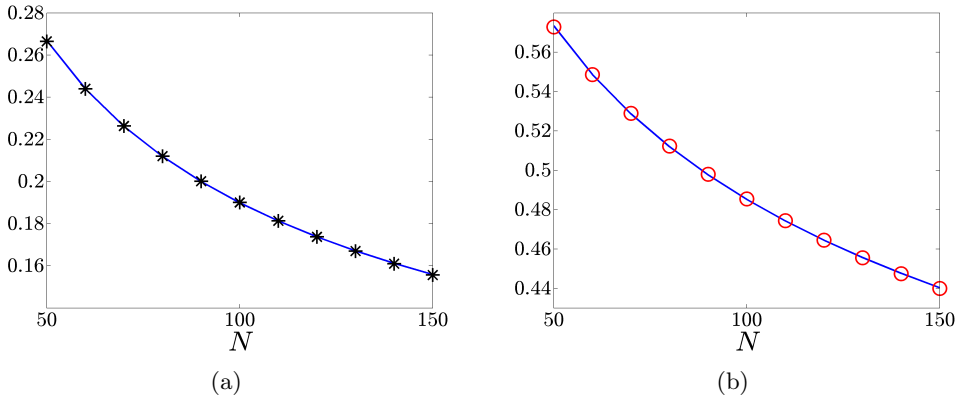


Figure 10.2: (a) Π_l $(*)$ using the optimal symmetric gain of Section 9.2, $1.8570/\sqrt{N} + 0.0042$ (curve); and (b) Π_l $(\circ)$ using the optimal non-symmetric gain of Section 9.3.3, $1.4738/\sqrt[4]{N} + 0.0191$ (curve). The optimal controllers are obtained by solving (9.4) with $Q = I$ and $r = 1$ for the formation with the fictitious follower.

Remark 4. *In contrast to the spatially uniform controllers, the optimal structured controllers of Sections 9.2 and 9.3.3, resulting from an N-independent control penalty r in (9.4), do not provide uniform bounds on the formation-size-normalized control energy. These controllers are obtained using $\mathcal{H}_2$ framework in which control effort represents a 'soft constraint'. It is thus of interest to examine formation coherence for optimal controllers with bounded control energy per vehicle. For formations without the fictitious follower, from (10.10) we see that the optimal symmetric controller with $r(N) = 2N/9$ asymptotically yields $\Pi_{\rm ctr} \approx 1$ and $\Pi_g \approx 2N/9 + 1/(3\sqrt{2})$. Similarly, for formations with followers, the optimal gains that result in $\Pi_{\rm ctr} \approx 1$ for large N can be obtained by changing control penalty from $r = 1$ to $r(N) = 0.08N$ for the optimal symmetric gain and to $r(N) = 0.175\sqrt{N}$ for the optimal non-symmetric gain*[1]*. These N-dependent control penalties provide an* affine *scaling of Π_g with N for the optimal symmetric gain and a* square-root *scaling of Π_g with N for the optimal non-symmetric gain; see Fig. 10.3. The asymptotic scalings for formations without followers subject to the optimal symmetric gains are obtained analytically (cf. (10.10)); all other scalings are obtained with the aid of computations.*

Remark 5. *Figure 10.3 illustrates the global performance measure Π_g obtained with four aforementioned structured controllers that asymptotically yield $\Pi_{\rm ctr} \approx 1$ for formations with fictitious follower. Note that the simple look-ahead strategy outperforms the optimal symmetric gain; $O(\sqrt{N})$ vs. $O(N)$ scaling. Thus, departure from symmetry in localized feedback gains can significantly improve coherence of large-scale formations. In particular, we have provided an example of a spatially uniform non-symmetric controller that yields better scaling trends than the optimal spatially varying controller obtained by restricting design to symmetric gains. Given the extra degrees of freedom in the optimal symmetric gain this is perhaps a surprising observation, indicating that the network*

[1]Both spatially uniform symmetric and look-ahead strategies with $\alpha = 1$ yield $\Pi_{\rm ctr} = 1$ in the limit of an infinite number of vehicles.

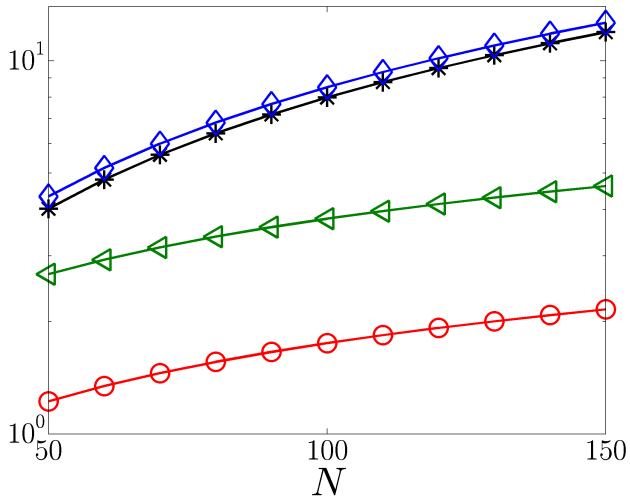


Figure 10.3: Π_g using four structured gains with $\Pi_{\text{ctr}} \approx 1$ for formations with fictitious follower: spatially uniform symmetric ($\diamond$), $N/12 + 1/6$ (blue curve), spatially uniform non-symmetric ($\triangleleft$), $2\sqrt{N}/(3\sqrt{\pi})$ (green curve), optimal symmetric ($*$), $0.0793N + 0.0493$ (black curve), and optimal non-symmetric ($\circ$), $0.1807\sqrt{N} - 0.0556$ (red curve).

topology may play a more important role than the optimal selection of the feedback gains in performance of large-scale interconnected systems. On the other hand, our results show that the optimal localized controller that achieves the best performance is both non-symmetric and spatially-varying.

10.3 Performance vs. size for the double-integrator model

For the double-integrator model with each vehicle having access to *its own velocity*, we show that the performance measures exhibit similar scaling properties to those found in single-integrators. Specifically, the optimal localized controller obtained by solving (9.4) with $Q = I$ and $r = 1$ provides a fourth-root dependence of the macroscopic performance measure Π_g on N; see Fig. 10.4a. Furthermore, the microscopic performance measure and control energy asymptotically scale as $O(1/\sqrt[4]{N})$ and $O(\sqrt[4]{N})$, respectively; see

Fig. 10.4b and Fig. 10.4c.

For comparison, we next provide the scaling trends of the performance measures for both the spatially uniform symmetric and look-ahead controllers. As in the single-integrator model, the spatially uniform symmetric gain $F_0 = [\,\alpha I \quad \alpha I \quad \beta I\,]$ provides linear scaling of Π_g with N and the formation-size-independent Π_l and $\Pi_{\rm ctr}$,

$$\begin{aligned} \Pi_g(N) &= \frac{N+2}{12\alpha\beta} + \frac{1}{2\beta} \\ \Pi_l &= \frac{1}{2\alpha\beta} + \frac{1}{2\beta} \\ \Pi_{\rm ctr} &= \frac{\alpha}{\beta} + \frac{\beta}{2}. \end{aligned}$$

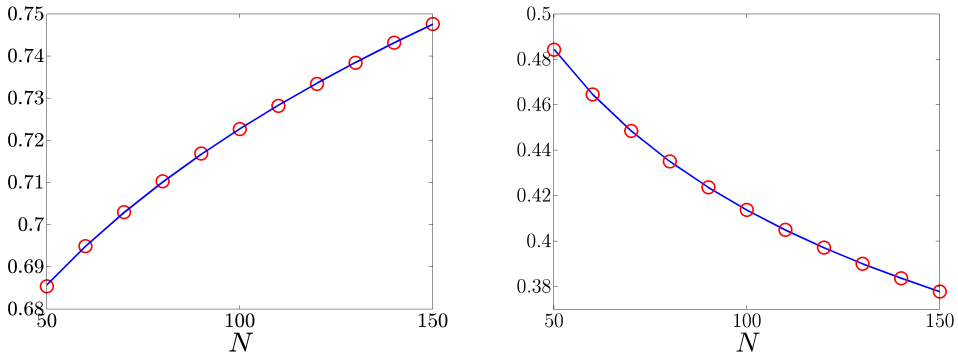
On the other hand, for the double-integrator model the performance of the look-ahead strategy $K = FC = [\,\alpha C_f \quad \beta I\,]$ heavily depends on the choices of α and β. In particular, for $\alpha = 1/4$ and $\beta = 1$, using similar techniques as in Section 10.2.2, we obtain

$$\Pi_g(N) = \frac{1}{\sqrt{\pi}} \sum_{n=1}^{N} \frac{(N-n+1)}{2N\,\Gamma(2n)} \Big(8\,\Gamma(2n-\frac{1}{2}) + \Gamma(2n-\frac{3}{2}) \Big)$$

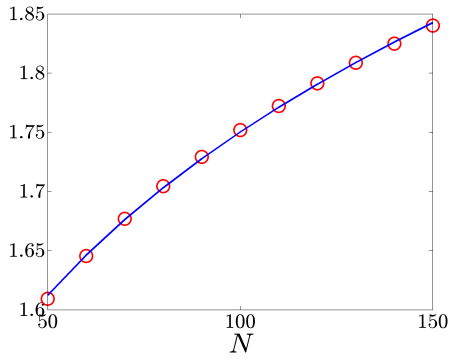
which asymptotically leads to the formation-size-independent scaling of $\Pi_{\rm ctr}$ and the square-root scaling of Π_g with N, i.e.,

$$\lim_{N\to\infty} \frac{\Pi_g(N)}{\sqrt{N}} = \frac{16}{3\sqrt{2\pi}};$$

also see Fig. 10.5a. This is in sharp contrast to $\alpha = \beta = 1$ which leads to an *exponential* dependence of Π_g on N; see Fig. 10.5b. Therefore, the design of the look-ahead strategy is much more subtle for double-integrators than for single-integrators.

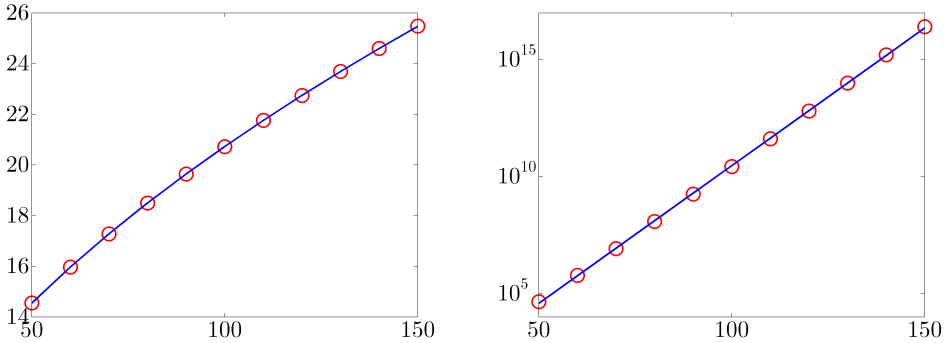


(a) Π_g (b) Π_l



(c) Π_{ctr}

Figure 10.4: Double-integrator model with the optimal non-symmetric gain obtained by solving (9.4) with $Q = I$ and $r = 1$ for formations with the fictitious follower: (a) Π_g ($\circ$), $0.0736\sqrt[4]{N} + 0.4900$ (curve) (b) Π_l ($\circ$), $1.1793/\sqrt[4]{N} + 0.0408$ (curve); (c) Π_{ctr} ($\circ$), $0.2742\sqrt[4]{N} + 0.8830$ (curve).



(a) (b)

Figure 10.5: Double-integrator model with look-ahead strategies $K = [\,\alpha C_f \quad \beta I\,]$. (a) For $\alpha = 1/4$ and $\beta = 1$, square-root scaling of Π_g ($\circ$), $2.1113\sqrt{N} - 0.3856$ (curve); (b) for $\alpha = \beta = 1$, exponential scaling of Π_g ($\circ$), $10^{0.1177N - 1.3058}$ (line).

Chapter 11

Conclusions and future directions

Conclusions

We consider the optimal control of one-dimensional formations with nearest neighbor interactions between vehicles. We formulate a structured optimal control problem in which local information exchange of relative positions between immediate neighbors imposes structural constraints on the feedback gains. We study the design problem for both the single- and the double-integrator models and employ a homotopy-based Newton's method to compute the optimal structured gains. We show that design of symmetric gains for the single-integrator model is a convex optimization problem, which we solve analytically for formations with no fictitious followers. For double-integrators, we identify a class of convex problems by restricting the controller to symmetric position and uniform diagonal velocity gains.

As in [68], we employ perturbation analysis to determine the departure from a stabilizing spatially uniform profile that yields nominal diffusion dynamics on a one-dimensional lattice; in contrast to [68], we find the 'mistuning' profile by optimizing a performance index rather than by performing spectral analysis. We also show how a homotopy-based Newton's method can be employed to obtain non-infinitesimal variation

in feedback gains that minimizes the desired objective function.

We establish several explicit scaling relationships and identify a spatially uniform non-symmetric controller that performs better than the optimal symmetric spatially varying controller ($O(\sqrt{N})$ vs. $O(N)$ scaling of coherence with $O(1)$ control energy per vehicle). This suggests that departure from symmetry can improve coherence of large-scale formations and that the controller structure may play a more important role than the optimal feedback gain design. On the other hand, our results demonstrate that the best performance is achieved with the optimal localized controller that is both non-symmetric and spatially-varying.

Extensions and future directions

Double-integrator model with relative position and velocity. The developed tools can be readily extended to the design of optimal structured controllers for double-integrators with relative position and velocity measurements; this is a topic of our ongoing research.

Central symmetry property. As noted in Section 9.3.3, the optimal forward and backward gains satisfy the *central symmetry* property, i.e., $f_n = b_{N+1-n}$ for $n \in \{1, \ldots, N\}$. This symmetry property implies that there may be hidden convexity of the non-symmetric feedback design. This is because a necessary condition for convexity of an optimization problem on graphs is that the solution is invariant under graph automorphism [74]. In other words, central symmetry is a necessary condition for the non-symmetric feedback design to be a convex problem. Furthermore, this symmetry can be exploited to reduce the number of variables and improve the efficiency of design algorithms.

Phase transition in scaling trends. For the double-integrator model with look-ahead strategy, the scaling trend of the macroscopic measure switches from a square-root

function of N (for $\alpha = 1/4$ and $\beta = 1$) to an exponential function of N (for $\alpha = 1$ and $\beta = 1$). It is thus of interest to study this scaling trend as a function of α with fixed $\beta = 1$ for the double-integrator model.

Part IV

Algorithms for leader selection in consensus networks

Chapter 12

Noise-corrupted leader selection problem

We consider consensus networks with two groups of nodes. Ordinary nodes, the so-called *followers*, form their control actions using relative information exchange with their neighbors; while special nodes, the so-called *leaders*, in addition to relative information from their neighbors also have access to their own states. This setting arises, for example, in the control of vehicular formations where all vehicles are equipped with ranging devices that provide relative distances from their neighbors, and the leaders additionally have GPS devices that provide their global position information.

Suppose that we want to equip a number of vehicles with GPS devices to keep all vehicles in the formation in their desired positions under the influence of noise. More precisely, we are interested in assigning an *a priori* specified number of nodes as leaders to minimize the steady-state variance of the deviation from consensus of the network. Related sensor selection problems arise in parameter estimation [75] and in localization of sensor networks [76, 77].

For undirected networks[1] in which *all nodes* including leaders are subject to stochastic disturbances, we show that the *noise-corrupted leader selection* problem has a convex objective function. In spite of this, the combinatorial nature of Boolean constraints (a node is either a leader or it is not) makes determination of the global minimum challenging for large networks. Instead, we focus on developing efficient algorithms to compute lower and upper bounds on the global optimal value.

This chapter is organized as follows. In Section 12.1, we formulate the noise-corrupted leader selection problem and discuss its connection to the sensor selection problem. In Section 12.2, we consider the convex relaxation of Boolean constraints and obtain a lower bound on the global optimal value using a customized interior point method. In Section 12.3, we obtain upper bounds on the global optimal value using two different algorithms. The first algorithm utilizes the one-leader-at-a-time greedy approach followed by a swap procedure that improves performance by checking possible swaps between leaders and followers. In both steps, algorithmic complexity is significantly reduced by exploiting structure of low-rank modifications to Laplacian matrices. The second algorithm utilizes the alternating direction method of multipliers (ADMM), which is capable of handling the nonconvex Boolean constraints by a simple projection. In Section 12.4, we use two examples to illustrate that the gap between the lower and upper bounds is small and sometimes even tight bounds can be attained.

12.1 Noise-corrupted leader selection

Consider n nodes with integrator dynamics

$$\dot{\psi}_i = u_i + w_i, \qquad i = 1, \ldots, n$$

[1] See Section 6.1 for the definition of undirected networks.

where ψ_i is the scalar state of node i, u_i is the control input, and w_i is the white stochastic disturbance with zero-mean and unit-variance. A node is a *follower* if it uses *only* relative information exchange with its neighbors to form its control action,

$$u_i \;=\; -\sum_{j\,\in\,\mathcal{N}_i} (\psi_i - \psi_j).$$

On the other hand, a node is a *leader* if, in addition to relative information exchange with its neighbors, it also has access to its own state ψ_i

$$u_i \;=\; -\sum_{j\,\in\,\mathcal{N}_i} (\psi_i - \psi_j) \;-\; \kappa_i\,\psi_i.$$

Here, κ_i is a positive number and $\mathcal{N}_i$ is the set neighboring nodes of node i.

The communication network is modeled by a connected, undirected graph; thus, the graph Laplacian F is a symmetric positive semidefinite matrix with a single eigenvalue at zero and the corresponding eigenvector $\mathbb{1}$ of all ones. A state-space representation of the leader-follower consensus network is given by

$$\dot{\psi} \;=\; -\,(F \;+\; D_\kappa D_x)\,\psi \;+\; w, \tag{12.1}$$

where

$$D_\kappa \;:=\; \mathrm{diag}\,(\kappa)\,, \quad D_x \;:=\; \mathrm{diag}\,(x)$$

are diagonal matrices formed from the vectors $\kappa = [\,\kappa_1 \;\cdots\; \kappa_n\,]^T$ and $x = [\,x_1 \;\cdots\; x_n\,]^T$. Here, x is a Boolean-valued vector with its ith entry $x_i \in \{0,\,1\}$, indicating that node i is a leader if $x_i = 1$ and that node i is a follower if $x_i = 0$. In connected networks with at least one leader $F + D_\kappa D_x$ is a positive definite matrix and the steady-state

covariance of ψ is determined by

$$\begin{aligned}\Sigma &:= \lim_{t\to\infty} \mathbf{E}\left(\psi(t)\,\psi^T(t)\right)\\ &= \int_0^\infty \mathrm{e}^{-(F+D_\kappa D_x)t}\,\mathrm{e}^{-(F+D_\kappa D_x)^T t}\,\mathrm{d}t\\ &= \frac{1}{2}\,(F+D_\kappa D_x)^{-1}.\end{aligned}$$

We use the total steady-state variance

$$\mathrm{trace}\,(\Sigma) \;=\; \frac{1}{2}\,\mathrm{trace}\left((F+D_\kappa D_x)^{-1}\right) \tag{12.2}$$

to quantify the performance of consensus networks subject to stochastic disturbances.

We are interested in identifying N_l leaders that are most effective in reducing the steady-state variance (12.2). For an *a priori* specified number of leaders $N_l < n$, the leader selection problem can thus be formulated as

$$\begin{array}{lrcl}\underset{x}{\text{minimize}} & J(x) & = & \mathrm{trace}\left((F\,+\,D_\kappa D_x)^{-1}\right)\\ \text{subject to} & x_i & \in & \{0,1\}, \qquad i\,=\,1,\ldots,n\\ & \mathbb{1}^T x & = & N_l.\end{array} \tag{LS1}$$

In (LS1), the number of leaders N_l as well as the matrices F and D_κ are the problem data, and the vector x is the optimization variable. As shown in Section 12.2, for a positive definite matrix $F + D_\kappa D_x$, the objective function J in (LS1) is a convex function of x. The challenging aspect of (LS1) comes from the nonconvex Boolean constraints $x_i \in \{0,1\}$; in general, finding the solution to (LS1) requires an intractable combinatorial search.

12.1.1 Connections to the sensor selection problem

The problem of estimating a vector $\psi \in \mathbb{R}^n$ from m relative measurements corrupted by additive white noise

$$y_{ij} \;=\; \psi_i \;-\; \psi_j \;+\; w_{ij}$$

arises in distributed localization in sensor networks. We consider the simplest scenario in which all ψ_i's are scalar-valued, with ψ_i denoting the position of sensor i; see [76, 77] for vector-valued localization problems. Let $\mathcal{I}_r$ denote the index set of the m pairs of distinct nodes between which the relative measurements are taken and let e_{ij} belong to $\mathbb{R}^n$ with 1 and -1 at its ith and jth elements, respectively, and zero everywhere else. Then,

$$y_{ij} \;=\; e_{ij}^T\,\psi \;+\; w_{ij}, \quad (i,j) \;\in\; \mathcal{I}_r$$

or, equivalently in the matrix form,

$$y_r \;=\; E_r^T\psi \;+\; w_r \tag{12.3}$$

where y_r is the vector of relative measurements and $E_r \in \mathbb{R}^{n\times m}$ is the matrix whose columns are determined by e_{ij} for $(i,j) \in \mathcal{I}_r$. Since $\psi + a\mathbb{1}$ for any scalar a results in the same y_r, with relative measurements the position vector ψ can be determined only up to an additive constant. This can also be verified by noting that $E_r^T\mathbb{1} = 0$.

Suppose that N_l sensors can be equipped with GPS devices that allow them to measure their absolute positions

$$y_a \;=\; E_a^T\psi \;+\; E_a^T w_a$$

where $E_a \in \mathbb{R}^{n\times N_l}$ is the matrix whose columns are determined by the ith unit vector e_i in $\mathbb{R}^n$ and the index i belongs to the index set of absolute measurements $i \in \mathcal{I}_a$. Then

the vector of all measurements is given by

$$\begin{bmatrix} y_r \\ y_a \end{bmatrix} = \begin{bmatrix} E_r^T \\ E_a^T \end{bmatrix} \psi + \begin{bmatrix} I & 0 \\ 0 & E_a^T \end{bmatrix} \begin{bmatrix} w_r \\ w_a \end{bmatrix} \tag{12.4}$$

where w_r and w_a are zero-mean white stochastic disturbances with

$$\mathbf{E}(w_r w_r^T) = W_r, \quad \mathbf{E}(w_a w_a^T) = W_a, \quad \mathbf{E}(w_r w_a^T) = 0.$$

Given the measurement vector y in (12.4), the linear minimum variance unbiased estimate of ψ is determined by [78, Chapter 4.4]

$$\hat{\psi} = (E_r W_r^{-1} E_r^T + E_a (E_a^T W_a E_a)^{-1} E_a^T)^{-1} (E_r W_r^{-1} y_r + E_a (E_a^T W_a E_a)^{-1} y_a)$$

with the covariance of the estimation error

$$\Sigma = \mathbf{E}((\psi - \hat{\psi})(\psi - \hat{\psi})^T) = (E_r W_r^{-1} E_r^T + E_a (E_a^T W_a E_a)^{-1} E_a^T)^{-1}.$$

Furthermore, let us assume that

$$W_r = I, \quad W_a = D_\kappa^{-1}.$$

Then

$$\Sigma = (E_r E_r^T + E_a E_a^T D_\kappa E_a E_a^T)^{-1},$$

where we use the fact that the inverse of a diagonal matrix is determined by the inverse of individual elements on the diagonal to obtain

$$(E_a^T D_\kappa^{-1} E_a)^{-1} = E_a^T D_\kappa E_a.$$

Since $E_a E_a^T$ is a diagonal matrix with its ith diagonal element being 1 for $i \in \mathcal{I}_a$ and $E_r E_r^T$ is the Laplacian matrix of the relative measurement graph, it follows that

$$D_x = E_a E_a^T, \quad F = E_r E_r^T, \quad \Sigma = (F + D_x D_\kappa D_x)^{-1} = (F + D_\kappa D_x)^{-1}$$

where $D_x D_\kappa D_x = D_\kappa D_x$ because D_x and D_κ commute and $D_x D_x = D_x$. Therefore, we have established the equivalence between the noise-corrupted leader selection problem (LS1) and the problem of choosing N_l sensors with absolute position measurements such that the variance of the estimation error is minimized.

12.2 Lower bound on global performance

Since the objective function J in (LS1) is the composition of a convex function trace $(\bar{F}^{-1})$ of a positive definite matrix $\bar{F} \succ 0$ with an affine function $\bar{F} = F + D_\kappa D_x$, it follows that J is a convex function of x. By enlarging the Boolean constraint set $x_i \in \{0, 1\}$ to its convex hull $x_i \in [0, 1]$ (i.e., the smallest convex set that contains the Boolean constraint set), we obtain a convex relaxation of (LS1)

$$\begin{array}{ll} \underset{x}{\text{minimize}} & J(x) = \text{trace}\left((F + D_\kappa D_x)^{-1}\right) \\ \text{subject to} & \mathbb{1}^T x = N_l, \quad 0 \leq x_i \leq 1, \qquad i = 1, \ldots, n. \end{array} \tag{CR}$$

Since we have enlarged the constraint set, the solution x^* of the relaxed problem (CR) provides a lower bound on J_{opt}. However, x^* may not provide a selection of N_l leaders, as it may turn out not to be Boolean-valued. If x^* is Boolean-valued, then it is the global solution of (LS1).

The convex optimization problem (CR) can be formulated as an SDP

$$\begin{array}{ll} \underset{X,\,x}{\text{minimize}} & \text{trace}\,(X) \\ \text{subject to} & \left[\begin{array}{cc} X & I \\ I & F + D_\kappa D_x \end{array}\right] \succeq 0 \\ & \mathbb{1}^T x \,=\, N_l, \quad 0 \,\leq\, x_i \,\leq\, 1, \quad i \,=\, 1, \ldots, n. \end{array}$$

For small networks (e.g., $n \leq 30$), this problem can be solved efficiently using standard SDP solvers. For large networks, we develop a customized interior point method in Section 12.2.1.

12.2.1 Customized interior point method for (CR)

We begin by augmenting the objective function in (CR) with log-barrier functions associated with the inequality constraints on x_i

$$\begin{array}{ll} \underset{x}{\text{minimize}} & q(x) \,=\, \tau\,\text{trace}\left((F \,+\, D_\kappa D_x)^{-1}\right) \,+\, \displaystyle\sum_{i\,=\,1}^{n}\big(-\,\log(x_i) \,-\, \log(1-x_i)\big) \\ \text{subject to} & \mathbb{1}^T x \,=\, N_l. \end{array} \tag{12.5}$$

The solution of the approximate problems (12.5) converges to the solution of the convex relaxation (CR) as the positive scalar τ increases to infinity [43, Section 11.2]. We solve a sequence of problem (12.5) by gradually increasing τ, and by starting each minimization using the solution from the previous value of τ. We use Newton's method to solve (12.5) for a fixed τ, and the Newton direction for problems with linear constraints is given by (e.g., see [43, Section 10.2])

$$x_{\text{nt}} \,=\, -\,(\nabla^2 q)^{-1}\nabla q \,-\, \delta(\nabla^2 q)^{-1}\mathbb{1}$$

where

$$\delta \;=\; -\frac{\mathbb{1}^T(\nabla^2 q)^{-1}\nabla q}{\mathbb{1}^T(\nabla^2 q)^{-1}\mathbb{1}}.$$

Here, the expressions for the ith entry of the gradient direction ∇q and for the Hessian matrix are given by

$$\begin{aligned}(\nabla q)_i \;&=\; -\;\tau\,\kappa_i\,((F+D_\kappa D_x)^{-2})_{ii}\;-\;x_i^{-1}\;-\;(x_i-1)^{-1}\\ \nabla^2 q\;&=2\tau\,(D_\kappa(F+D_\kappa D_x)^{-2}D_\kappa)\circ(F+D_\kappa D_x)^{-1}\;+\;\mathrm{diag}\left(x_i^{-2}\;+\;(1-x_i)^{-2}\right),\end{aligned}$$

where $\circ$ denotes the elementwise multiplication of matrices and $\mathrm{diag}(x)$ denotes a diagonal matrix with the main diagonal determined by x.

We next examine complexity of computing the Newton direction x_{nt}. The major cost of computing $\nabla^2 q$ is to form $(F+D_\kappa D_x)^{-2}$, which takes $(7/3)n^3$ operations to form $(F+D_\kappa D_x)^{-1}$ and n^3 operations to form $(F+D_\kappa D_x)^{-2}$. Computing x_{nt} requires solving two linear equations,

$$(\nabla^2 q)\,y \;=\; -\nabla q, \quad (\nabla^2 q)\,z \;=\; -\mathbb{1}$$

which takes $(1/3)n^3$ operations using Cholesky factorization. Thus, computation of each Newton step requires $(7/3+1+1/3)n^3 = (11/3)n^3$ operations.

12.3 Upper bounds on global performance

We next show that upper bounds can be obtained using an efficient greedy algorithm and the alternating direction method of multipliers (ADMM). Greedy algorithm selects one leader at a time, which introduces low-rank modifications to the Laplacian matrix. We exploit this feature in conjunction with the matrix inversion lemma to gain computational efficiency. On the other hand, the ADMM algorithm handles the Boolean

constraints explicitly by a simple projection onto a discrete nonconvex set.

12.3.1 Greedy algorithm to obtain an upper bound

With the lower bound on the optimal value J_{opt} resulting from the convex relaxation (CR) in Section 12.2, we next use a greedy algorithm to compute an upper bound on J_{opt}. This algorithm selects one leader at a time by assigning the node that provides the largest performance improvement as the leader. Once this is done, an attempt to improve a selection of N_l leaders is made by checking possible swaps between the leaders and the followers. We show that substantial improvement in algorithmic complexity can be achieved by exploiting structure of the low-rank modifications to the Laplacian matrix.

One-leader-at-a-time algorithm

As the name suggests, we select one leader at a time by assigning the node that results in the largest performance improvement as the leader. To select the first leader, we compute

$$J_1^i \;=\; \text{trace}\left((F \,+\, \kappa_i e_i e_i^T)^{-1}\right)$$

for $i = 1, \ldots, n$, and assign the node, say v_1, that achieves the minimum value of $\{J_1^i\}$. If two or more nodes provide the largest performance improvement, we select one of these nodes as a leader. After choosing s leaders, $v_1, \ldots, v_s$, we compute

$$\begin{aligned} J_{s+1}^i &= \text{trace}\left((F_s \,+\, \kappa_i e_i e_i^T)^{-1}\right) \\ F_s &= F \,+\, \kappa_{v_1} e_{v_1} e_{v_1}^T \,+\, \cdots \,+\, \kappa_{v_s} e_{v_s} e_{v_s}^T \end{aligned}$$

for $i \notin \{v_1, \ldots, v_s\}$, and select node v_{s+1} that yields the minimum value of $\{J_{s+1}^i\}$. This procedure is repeated until all N_l leaders are selected.

Without exploiting structure, the above procedure requires $O(n^4 N_l)$ operations. On the other hand, the rank-1 update formula obtained from matrix inversion lemma

$$(F_s + \kappa_i e_i e_i^T)^{-1} = F_s^{-1} - \frac{F_s^{-1} \kappa_i e_i e_i^T F_s^{-1}}{1 + \kappa_i e_i^T F_s^{-1} e_i} \tag{12.6}$$

yields

$$J_{s+1}^i = \text{trace}\,(F_s^{-1}) - \frac{\kappa_i \|(F_s^{-1})_i\|_2^2}{1 + \kappa_i (F_s^{-1})_{ii}},$$

where $(F_s^{-1})_i$ is the ith column of F_s^{-1} and $(F_s^{-1})_{ii}$ is the iith entry of F_s^{-1}. To initiate the algorithm, we use the generalized rank-1 update [79],

$$F_1^{-1} = F^\dagger - (F^\dagger e_i)\mathbb{1}^T - \mathbb{1}(F^\dagger e_i)^T + ((1/\kappa_i) + e_i^T F^\dagger e_i)\mathbb{1}\mathbb{1}^T$$

and thus,

$$J_1^i = \text{trace}\,(F^\dagger) + n\,((1/\kappa_i) + e_i^T F^\dagger e_i),$$

where $F^\dagger$ denotes the pseudo-inverse of F

$$F^\dagger = (F + \mathbb{1}\mathbb{1}^T/n)^{-1} - \mathbb{1}\mathbb{1}^T/n.$$

Therefore, once F_s^{-1} is determined, the inverse of the matrix on the left-hand-side of (12.6) can be computed using $O(n^2)$ operations and J_{s+1}^i can be evaluated using $O(n)$ operations. Overall, N_l rank-1 updates, $nN_l/2$ objective function evaluations, and one full matrix inverse (for computing F_s^{-1}) require $O(n^2 N_l + n^3)$ operations as opposed to $O(n^4 N_l)$ operations without exploiting the low-rank structure. In large-scale networks, further computational advantage can be gained by exploiting structure of the underlying Laplacian matrices; see [80].

Swap algorithm

Having determined a selection of leaders using one-leader-at-a-time algorithm, we swap one of the N_l leaders with one of the $n-N_l$ followers, and check if such a swap leads to a decrease in J. If no decrease occurs for all $(n-N_l)N_l$ swaps, the algorithm terminates; if a decrease in J occurs, we update the leader and then restart checking the possible $(n-N_l)N_l$ swaps for the new leader selection. This swap procedure has been used as an effective means for improving performance of combinatorial algorithms encountered in graph partitioning [81], sensor selection [75], and community detection problems [82].

Since a swap between a leader i and a follower j leads to a rank-2 modification (12.7) to the matrix $\bar{F} = F + D_\kappa D_x$, we can exploit this low-rank structure to gain computational efficiency. Using the matrix inversion lemma, we have

$$\left(\bar{F} \,-\, \kappa_i e_i e_i^T \,+\, \kappa_j e_j e_j^T\right)^{-1} = \bar{F}^{-1} \,-\, \bar{F}^{-1}\,\bar{E}_{ij}\,(I_2 \,+\, E_{ij}^T \bar{F}^{-1} \bar{E}_{ij})^{-1}\,E_{ij}^T\,\bar{F}^{-1}, \quad (12.7)$$

where $E_{ij} = [\, e_i \;\; e_j \,]$, $\bar{E}_{ij} = [\, -\kappa_i e_i \;\; \kappa_j e_j \,]$, and I_2 is the 2×2 identity matrix. Thus, the objective function after the swap between leader i and follower j is given by

$$J_{ij} \;=\; J \;-\; \mathrm{trace}\left((I_2 + E_{ij}^T \bar{F}^{-1} \bar{E}_{ij})^{-1} E_{ij}^T\,\bar{F}^{-2}\bar{E}_{ij}\right). \quad (12.8)$$

Here, we do not need to form the full matrix $\bar{F}^{-2}$, since

$$E_{ij}^T\,\bar{F}^{-2}\bar{E}_{ij} \;=\; \begin{bmatrix} -\kappa_i(\bar{F}^{-2})_{ii} & \kappa_j(\bar{F}^{-2})_{ij} \\ -\kappa_i(\bar{F}^{-2})_{ji} & \kappa_j(\bar{F}^{-2})_{jj} \end{bmatrix}$$

and the ijth entry of $\bar{F}^{-2}$ can be computed by multiplying the ith row of $\bar{F}^{-1}$ with the jth column of $\bar{F}^{-1}$. Thus, evaluation of J_{ij} takes $O(n)$ operations and computation of the matrix inverse in (12.7) requires $O(n^2)$ operations.

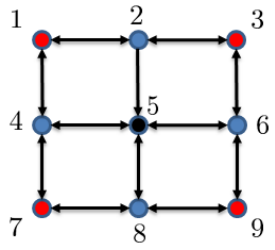


Figure 12.1: A 3×3 grid.

Remark 6. *Since the total number of swaps can be large, we limit the maximum number of swaps with a linear function of the number of nodes n. On the other hand, particular structure of networks can be exploited to reduce the required number of swaps. To illustrate this, let us consider the problem of selecting one leader in a network with* 9 *nodes shown in Fig. 12.1. Suppose that nodes in set $S_1 := \{1, 3, 7, 9\}$ have the same feedback gain κ_1 and that nodes in set $S_2 := \{2, 4, 6, 8\}$ have the same feedback gain κ_2. In addition, suppose that node* 5 *is chosen as a leader. Owing to symmetry, to check if selecting other nodes as a leader can improve performance we only need to swap node* 5 *with one node in each set S_1 and S_2. We note that more sophisticated symmetry exploitation techniques can be found in [74, 83].*

12.3.2 Alternating direction method of multipliers

Since the previously introduced greedy algorithm may not yield an optimal selection of leaders, we next employ the ADMM algorithm [50] as an alternative approach to a selection of N_l leaders for problem (LS1). Although the convergence of this method depends on the initial conditions and on the algorithmic parameters, ADMM is capable of handling the nonconvex Boolean constraints explicitly by a simple projection onto a discrete nonconvex set

$$\mathcal{B} := \left\{ x \,|\, \mathbb{1}^T x = N_l,\ x_i \in \{0, 1\},\ i = 1, \ldots, n \right\}. \tag{12.9}$$

We can rewrite (LS1) as an unconstrained optimization problem

$$\underset{X,\,x}{\text{minimize}} \quad J(x) + \mathcal{I}(x), \tag{12.10}$$

where $\mathcal{I}(x)$ is indicator function associated with set $\mathcal{B}$

$$\mathcal{I}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{B} \\ +\infty & \text{if } x \notin \mathcal{B}. \end{cases}$$

Now, (12.10) can be put into the following equivalent form suitable for the application of ADMM

$$\begin{aligned} \underset{x,\,z}{\text{minimize}} \quad & J(x) + \mathcal{I}(z) \\ \text{subject to} \quad & x - z = 0 \end{aligned} \tag{12.11}$$

and the augmented Lagrangian associated with (12.11) is given by

$$\mathcal{L}_\rho(x, z, \lambda) = J(x) + \mathcal{I}(z) + \lambda^T(x - z) + \frac{\rho}{2}\|x - z\|_2^2,$$

where $\lambda \in \mathbb{R}^n$ is the dual variable and ρ is a positive number. The ADMM algorithm updates x, z, and λ in an iterative fashion

$$x^{k+1} := \arg\min_x \mathcal{L}_\rho(x, z^k, \lambda^k) \tag{12.12a}$$

$$z^{k+1} := \arg\min_z \mathcal{L}_\rho(x^{k+1}, z, \lambda^k) \tag{12.12b}$$

$$\lambda^{k+1} := \lambda^k + \rho(x^{k+1} - z^{k+1}), \tag{12.12c}$$

for $k = 0, 1, \ldots$ until $\|x^{k+1} - z^{k+1}\|_2 \leq \epsilon$ and $\|z^{k+1} - z^k\|_2 \leq \epsilon$.

Splitting the optimization variables into two copies $\{x,\, z\}$ and updating them in an alternating fashion yields the minimization problems (12.12a) and (12.12b) that are

easy to solve.

x-minimization step

By completion of squares in $\mathcal{L}_\rho$ with respect to x, problem (12.12a) can be expressed as

$$\underset{x}{\text{minimize}} \quad \text{trace}\left((F + D_\kappa D_x)^{-1}\right) + \frac{\rho}{2}\|x - u^k\|_2^2 \tag{12.13}$$

where

$$u^k := z^k - (1/\rho)\lambda^k.$$

Since (12.13) is equivalent to the following problem,

$$\begin{array}{ll} \underset{x,\,\mu}{\text{minimize}} & \text{trace}\left((F + D_\kappa D_x)^{-1}\right) + \mu \\ \text{subject to} & \dfrac{\rho}{2}\|x - u^k\|_2^2 \le \mu \end{array}$$

it can be expressed as an SDP

$$\begin{array}{ll} \underset{X,\,x,\,\mu}{\text{minimize}} & \text{trace}\,(X) + \mu \\ \text{subject to} & \begin{bmatrix} X & I \\ I & F + D_\kappa D_x \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} I & x - u^k \\ (x - u^k)^T & 2\mu/\rho \end{bmatrix} \succeq 0 \end{array}$$

where the second LMI is an alternative way of writing the quadratic constraint

$$2\mu/\rho - (x - u^k)^T(x - u^k) \ge 0$$

using Schur complement. Thus, for small networks, problem (12.13) can be solved efficiently using standard SDP solvers. For large networks, we use descent methods [43]

(e.g., Newton's method) with the gradient and Hessian of $\mathcal{L}_\rho$ with respect to x being given by

$$\begin{aligned} \nabla \mathcal{L}_\rho &= -\kappa \circ \operatorname{diag}\left((F + D_\kappa D_x)^{-2}\right) + \rho\,(x - u^k) \\ \nabla^2 \mathcal{L}_\rho &= 2\,(D_\kappa (F + D_\kappa D_x)^{-2} D_\kappa) \circ (F + D_\kappa D_x)^{-1} + \rho I, \end{aligned}$$

where $\operatorname{diag}(M)$ denotes the vector determined by the main diagonal of a matrix M.

z-minimization step

Using similar argument as in [50, Section 9.1], the z-minimization problem (12.12b) can be solved explicitly using a simple projection onto the set $\mathcal{B}$

$$z_i = \begin{cases} 1 & \text{if } v_i^k \geq [v^k]_{N_l} \\ 0 & \text{if } v_i^k < [v^k]_{N_l}, \end{cases} \tag{12.14}$$

where

$$v^k := x^{k+1} + (1/\rho)\lambda^k$$

and $[v^k]_{N_l}$ denotes the (N_l)th largest entry of v^k.

We next provide detailed derivation for (12.14). We use completion of squares to obtain the following problem which is equivalent to (12.12b)

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & (\rho/2)\|z - v^k\|_2^2 \\ \text{subject to} \quad & z \in \mathcal{B} \end{aligned}$$

whose solution is given by (12.14). To see this, consider $\bar{z} \in \mathcal{B}$, i.e., $\mathbb{1}^T \bar{z} = N_l$ and $\bar{z}_i \in \{0, 1\}$, but $\bar{z}$ is not the projection determined by (12.14). Thus, there exists at least one entry of $\bar{z}$, say the rth entry, such that $\bar{z}_r = 1$ for $v_r^k < [v^k]_{N_l}$, and at least one

entry, say the jth entry, such that $\bar{z}_j = 0$ for $v_j^k \geq [v^k]_{N_l}$. Consider

$$\delta_{rj} \;=\; (\bar{z}_r - v_r^k)^2 \;+\; (\bar{z}_j - v_j^k)^2 \;=\; (1 - v_r^k)^2 \;+\; (v_j^k)^2$$

and $\delta_{jr} = (v_r^k)^2 + (1 - v_j^k)^2$. Since $\delta_{rj} - \delta_{jr} = 2(v_j^k - v_r^k) > 0$, it follows that the objective function $(\rho/2)\|z - v^k\|_2^2$ will decrease if we choose $\{\bar{z}_r = 0,\ \bar{z}_j = 1\}$ instead of $\{\bar{z}_r = 1,\ \bar{z}_j = 0\}$. Therefore, we can reduce the objective function by exchanging the values of two entries $\bar{z}_r = 1$ (with $v_r^k < [v^k]_{N_l}$) and $\bar{z}_j = 0$ (with $v_j^k \geq [v^k]_{N_l}$) until (12.14) is satisfied for all $i = 1, \ldots, n$.

12.4 Examples

We next provide two examples, a small network from [41] and a 2D lattice, to illustrate the performance of the developed methods. In both examples we set κ_i to be the degree of node i. We set the initial conditions of the ADMM algorithm to $\{z^0 = 0,\ \lambda^0 = 0\}$ and the penalty weight to $\rho = 10^3$.

12.4.1 A small network

For the network shown in Fig. 12.2 with $N_l \leq 5$, we determine the global minima to the noise-corrupted leader selection problem (LS1) by exhaustive search. It turns out that the one-leader-at-a-time algorithm followed by the swap algorithm actually finds the global minima. As shown in Table 12.1, ADMM provides the global minima for the problems with 4 and 5 leaders.

Figure 12.3a shows lower bounds resulting from convex relaxation and upper bounds resulting from ADMM and from greedy algorithm. As the number of leaders N_l increases, the gap between the lower and upper bounds from greedy algorithm decreases; see Fig. 12.3b.

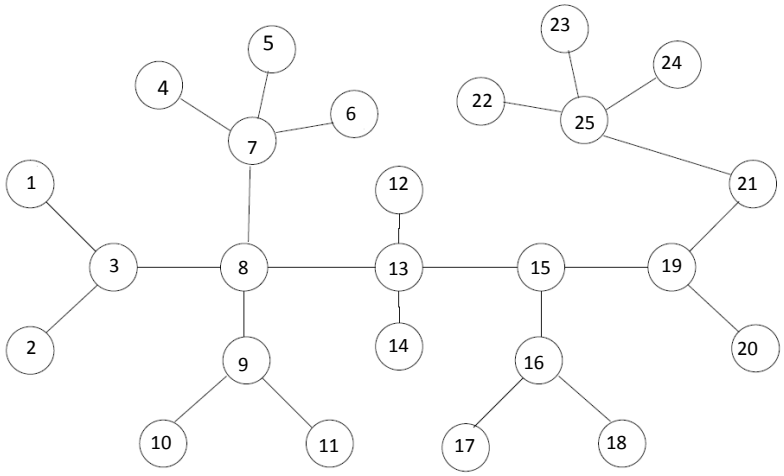
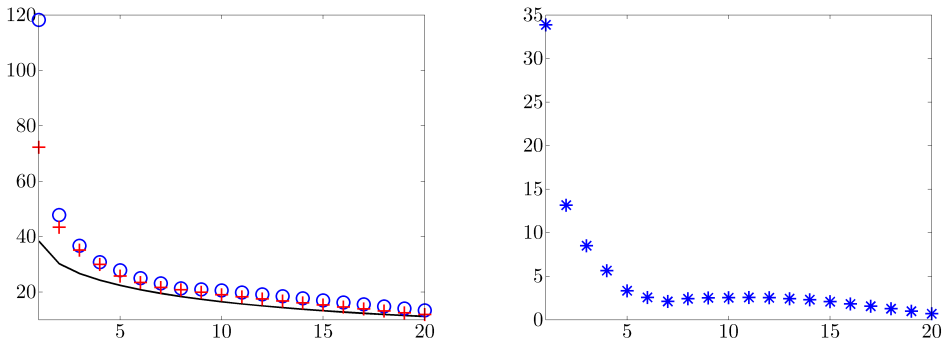


Figure 12.2: A small network with 25 nodes.

Table 12.1: Lower and upper bounds on the noise-corrupted leader selection problem (LS1) for the example shown in Fig. 12.2. Lower bounds J_{lb} are obtained by solving the convex relaxation (CR); upper bounds J_{ub} from greedy algorithm – the one-leader-at-a-time algorithm followed by the swap algorithm – are actually tight, i.e., $J_{\text{ub}} = J_{\text{opt}}$; upper bounds J_{ub} from ADMM are tight for $N_l = 4, 5$.

		greedy algorithm		ADMM	
N_l	J_{lb}	J_{ub}	leaders	J_{ub}	leaders
1	38.4	72.3	13	118.3	25
2	30.3	43.4	8, 25	47.9	7, 25
3	26.7	35.2	8, 16, 25	36.7	7, 16, 25
4	24.3	30.0	3, 7, 16, 25	30.0	3, 7, 16, 25
5	22.4	25.8	3, 7, 9, 16, 25	25.8	3, 7, 9, 16, 25



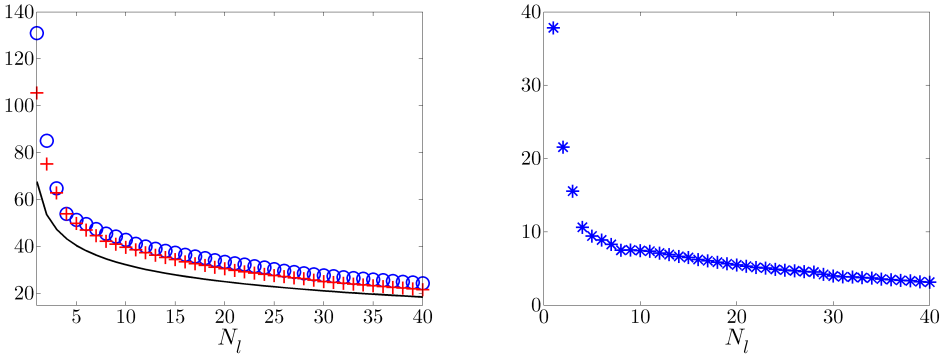
(a) (b)

Figure 12.3: Computational results for the network with 25 nodes shown in Fig. 12.2: (a) lower bounds ($-$) resulting from convex relaxation and upper bounds resulting from greedy algorithm (i.e., one-leader-at-a-time algorithm followed by swap algorithm) ($+$) and from ADMM ($\circ$); (b) the gap between lower bounds and upper bounds resulting from greedy algorithm.

12.4.2 A 2D lattice

We next consider the leader selection problem for a 9×9 regular lattice. Figure 12.4a shows lower bounds resulting from convex relaxation and upper bounds resulting from ADMM and from greedy algorithm, i.e., the one-leader-at-a-time algorithm followed by the swap algorithm. As the number of leaders N_l increases, the gap between the lower and upper bounds from greedy algorithm decreases; see Fig. 12.4b. For $N_l = 1, \ldots, 40$, the number of swap updates ranges between 1 and 19 and the average number of swaps is 10.

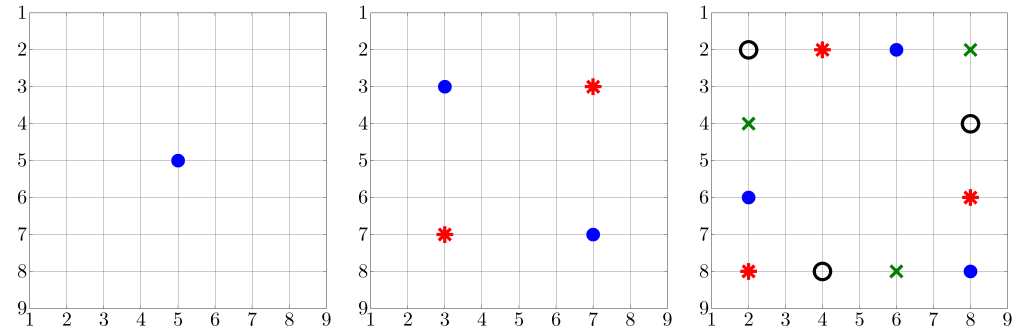
Figure 12.5 shows selection of leaders resulting from the greedy algorithm for different choices of N_l. For $N_l = 1$, the *center* node $(5, 5)$ provides the optimal selection of a single leader. As N_l increases, nodes away from the center node $(5, 5)$ are selected; for example, for $N_l = 2$, nodes $\{(3, 3), (7, 7)\}$ are selected and for $N_l = 3$, nodes $\{(2, 6), (6, 2), (8, 8)\}$ are selected. Selection of nodes farther away from the center becomes more significant for $N_l = 4$ and $N_l = 8$. The selection of leaders exhibits symmetry shown in



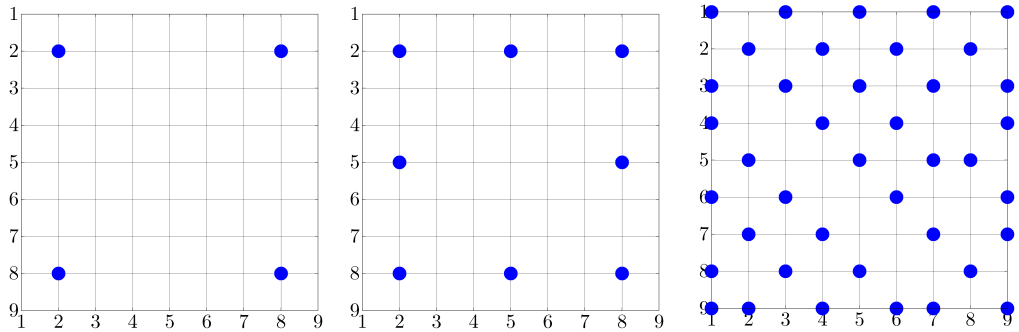
(a) (b)

Figure 12.4: The computational results for the 2D lattice example: (a) lower bounds ($-$) resulting from convex relaxation and upper bounds resulting from greedy algorithm (i.e., one-leader-at-a-time algorithm followed by swap algorithm) ($+$) and from ADMM ($\circ$); (b) the gap between lower bounds and upper bounds resulting from greedy algorithm.

Fig. 12.5. In particular, when N_l is large, almost uniform spacing between the leaders is observed; see Fig. 12.5f for $N_l = 40$.



(a) $N_l = 1$ (b) $N_l = 2$ (c) $N_l = 3$



(d) $N_l = 4$ (e) $N_l = 8$ (f) $N_l = 40$

Figure 12.5: Selections of leaders (•) obtained using the one-at-a-time algorithm followed by the swap algorithm for a 2D lattice. The two selections of two leaders denoted by (•) and (∗) in (b) provide the same objective function J. The four selections of three leaders denoted by (•), (∗), (×), and (○) in (c) provide the same J.

Chapter 13

Noise-free leader selection problem

In this chapter, we consider the noise-free leader selection problem in which leaders are assumed to be immune to noise and they follow their desired trajectories at all times. This idealized setting has recently be studied by several authors; e.g., see [41, 42]. For connected networks with at least one leader, it was shown in [41] that adding leaders always improves performance. In view of this, the one-at-a-time greedy algorithm was proposed in [41] for the noise-free leader selection problem. It was shown in [42] that the variance of deviation from consensus is a supermodular function of the set of noise-free leaders. This implies that the performance improvement by adding additional leaders is *diminishing* as the number of leaders increases. Furthermore, the supermodular optimization framework is then employed to show that the greedy algorithm provides selection of leaders that is within a provable bound from globally optimal solution [42].

In contrast to the above references, we use convex optimization to select noise-free leaders. We first provide an explicit expression for the objective function to identify the source of nonconvexity and to suggest an LMI-based convex relaxation. We then

relax the *hard* Boolean constraint on the number of leaders with a *soft-constraint* by augmenting the objective function with the ℓ_1 norm of the optimization variables. The ℓ_1 norm provides a means for obtaining a sparse solution whose nonzero elements identify the leaders. The developed algorithm produces a trade-off curve between the number of noise-free leaders and the variance of the deviation from consensus by solving a parameterized family of convex optimization problems.

This chapter is organized as follows. In Section 13.1, we formulate the noise-free leader selection problem and discuss its connections to the sensor selection problem. In Section 13.2, we provide an explicit expression for the objective function and we propose the soft-constraint method. In Section 13.3, we use two examples to illustrate that the performance of the developed method is close to that of the greedy algorithm, and both approaches outperform the simple degree-heuristics-based method.

13.1 Noise-free leader selection

We consider the selection of *noise-free* leaders which follow their desired trajectories at all times. Equivalently, in coordinates that determine deviation from the desired trajectory, the state of every leader is identically equal to zero. Therefore, the dynamics of the followers are given by

$$\dot{\psi}_f \;=\; -\,F_f\,\psi_f \;+\; w_f.$$

Here, F_f is obtained from the Laplacian matrix F of the network by eliminating all rows and columns associated with the leaders. Thus, the problem of selecting leaders

that minimize the steady-state variance of ψ_f amounts to

$$\begin{array}{lrcl} \underset{x}{\text{minimize}} & J_f(x) & = & \text{trace}\,(F_f^{-1}) \\ \text{subject to} & x_i & \in & \{0,1\}, \qquad i \,=\, 1,\ldots,n \\ & \mathbb{1}^T x & = & N_l. \end{array} \tag{LS2}$$

As in (LS1), the Boolean constraints $x_i \in \{0,1\}$ are nonconvex. Furthermore, as we demonstrate in Section 13.2, the objective function J_f in (LS2) is a nonconvex function of x.

In what follows, we establish the equivalence between the noise-corrupted and the noise-free leader selection problems (LS1) and (LS2) when all leaders use arbitrarily large feedback gains on their own states. Partitioning ψ into the state of the leader nodes ψ_l and the state of the follower nodes ψ_f brings system (12.1) to the following form[1]

$$\begin{bmatrix} \dot{\psi}_l \\ \dot{\psi}_f \end{bmatrix} = - \begin{bmatrix} F_l + D_{\kappa_l} & F_0 \\ F_0^T & F_f \end{bmatrix} \begin{bmatrix} \psi_l \\ \psi_f \end{bmatrix} + \begin{bmatrix} w_l \\ w_f \end{bmatrix}. \tag{13.1}$$

Here, $D_{\kappa_l} := \text{diag}\,(\kappa_l)$ with $\kappa_l \in \mathbb{R}^{N_l}$ being the vector of feedback gains associated with the leaders. Taking the trace of the inverse of the 2×2 block matrix in (13.1) yields

$$J \,=\, \text{trace}\left(F_f^{-1} \,+\, F_f^{-1}\, F_0^T\, S_{\kappa_l}^{-1}\, F_0\, F_f^{-1} \,+\, S_{\kappa_l}^{-1}\right)$$

where

$$S_{\kappa_l} \,=\, F_l \,+\, D_{\kappa_l} \,-\, F_0\, F_f^{-1}\, F_0^T$$

is the Schur complement of F_f. Since $S_{\kappa_l}^{-1}$ vanishes as each component of the vector κ_l goes to infinity, the variance of the network is solely determined by the variance of

[1]Note that D_x does not show in (13.1) since the partition is performed with respect to the indices of the 0 and 1 diagonal elements of D_x.

the followers, $J_f = \text{trace}\,(F_f^{-1})$, where F_f is the reduced Laplacian matrix obtained by removing all columns and rows corresponding to the leaders from F.

13.1.1 Connections to the sensor selection problem

Recall that in Section 12.1.1, we show that the noise-corrupted leader selection problem (LS1) is equivalent to the sensor selection problem in which the selected sensors have access to their absolute measurements. To formulate a sensor selection problem that is equivalent to the noise-free leader selection problem (LS2), we assume that the exact positions of N_l sensors are known *a priori*. Let ψ_l denote the positions of these *reference sensors* and let ψ_f denote the positions of the other sensors. We can thus write the relative measurement equation (12.3) as

$$y_r \;=\; E_r^T\psi \;+\; w_r \;=\; E_l^T\psi_l \;+\; E_f^T\psi_f \;+\; w_r$$

and the linear minimum variance unbiased estimate of ψ_f is given by

$$\hat{\psi}_f \;=\; (E_f E_f^T)^{-1} E_f W_r^{-1}\,(y_r \;-\; E_l^T\psi_l)$$

with the covariance of the estimation error being

$$\Sigma_f \;=\; (E_f E_f^T)^{-1}.$$

Identifying $E_f E_f^T$ with F_f in the Laplacian matrix F

$$F \;=\; E_r E_r^T \;=\; \begin{bmatrix} E_l E_l^T & E_l E_f^T \\ E_f E_l^T & E_f E_f^T \end{bmatrix} \;=\; \begin{bmatrix} F_l & F_0 \\ F_0^T & F_f \end{bmatrix}$$

establishes the equivalence between problem (LS2) and the problem of assigning N_l sensors with known reference positions to minimize the variance of the estimation error of sensor network.

13.2 Linear approximation and soft-constraint method

In this section, we provide an alternative expression for the objective function J_f in the noise-free leader selection problem (LS2). We use this explicit expression to identify the source of nonconvexity and to suggest an LMI-based convex approximation. We then relax the hard constraint of having exactly N_l leaders in (LS2) by augmenting the objective function J_f with the ℓ_1 norm of the optimization variable x. This *soft-constraint* approach yields a parameterized family of optimization problems whose solution provides a trade-off between the ℓ_1 norm of x and the convex approximation of the variance amplification of the network.

13.2.1 Explicit expression for the objective function

Since the objective function J_f in (LS2) is not expressed explicitly in terms of the optimization variable x, it is difficult to examine its basic properties such as convexity. We next provide an alternative expression for J_f that allows us to establish the lack of convexity and to suggest an LMI-based convex approximation of J_f.

Proposition 7. *For networks with at least one leader, the objective function J_f in the noise-free leader selection problem (LS2) can be written as*

$$J_f \;=\; \mathrm{trace}\,(F_f^{-1}) \;=\; \mathrm{trace}\,\big((I \,-\, D_x)(G \,+\, D_x \circ F)^{-1}(I \,-\, D_x)\big) \tag{13.2}$$

where $\circ$ denotes the elementwise multiplication of matrices, and

$$G = (I - D_x)\, F\, (I - D_x), \quad D_x = \operatorname{diag}(x), \quad x_i \in \{0,1\}, \quad i = 1, \ldots, n.$$

Furthermore, J_f is a nonconvex function of x over the smallest convex set $x_i \in [0,1]$ that contains feasible points $x_i \in \{0,1\}$ for $i = 1, \ldots, n$.

Proof. After an appropriate relabeling of the nodes as done in (13.1), F and D_x can be partitioned conformably into 2×2 block matrices,

$$F = \begin{bmatrix} F_l & F_0 \\ F_0^T & F_f \end{bmatrix}, \quad D_x = \begin{bmatrix} I_{N_l \times N_l} & O_{N_l \times p} \\ O_{p \times N_l} & O_{p \times p} \end{bmatrix}, \quad p := n - N_l$$

which leads to

$$G = \begin{bmatrix} O_{N_l \times N_l} & O_{N_l \times p} \\ O_{p \times N_l} & F_f \end{bmatrix}, \quad D_x \circ F = \begin{bmatrix} I_{N_l \times N_l} \circ F_l & O_{N_l \times p} \\ O_{p \times N_l} & O_{p \times p} \end{bmatrix}$$

$$G + D_x \circ F = \begin{bmatrix} I_{N_l \times N_l} \circ F_l & O_{N_l \times p} \\ O_{p \times N_l} & F_f \end{bmatrix}.$$

Since $I_{N_l \times N_l} \circ F_l$ is a diagonal matrix with positive diagonal elements and since the principal submatrix F_f of the Laplacian F is positive definite for connected graphs [4, Lemma 10.36], we have

$$G + D_x \circ F \succ 0. \tag{13.3}$$

Consequently,

$$\operatorname{trace}\left((I - D_x)(G + D_x \circ F)^{-1}(I - D_x)\right) = \operatorname{trace}(F_f^{-1})$$

which yields the desired result (13.2).

We next use a simple example to illustrate the lack of convexity of J_f over $x_i \in [0,1]$. Let

$$F = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad D_x = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}$$

with $x_1 \in [0,1]$ and $x_2 = 1$. From

$$G + F \circ D_x = \begin{bmatrix} (1-x_1)^2 + x_1 & 0 \\ 0 & 1 \end{bmatrix} \succ 0 \quad \text{and} \quad J_f = \frac{(1-x_1)^2}{(1-x_1)^2 + x_1}$$

it can be verified that, for $x_1 \in [0, 1/3]$ the second derivative of J_f with respect to x_1 is negative; therefore, J_f is not convex. □

Explicit expression (13.2) in conjunction with Schur complement can be used to convert the minimization of J_f into the following problem

$$\begin{aligned} \underset{X,\,x}{\text{minimize}} \quad & \text{trace}\,(X) \\ \text{subject to} \quad & \begin{bmatrix} X & I - D_x \\ I - D_x & G + D_x \circ F \end{bmatrix} \succeq 0 \end{aligned} \tag{13.4}$$

where $X \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix. To see this, note that since $G + D_x \circ F \succ 0$, we have

$$\begin{bmatrix} X & I - D_x \\ I - D_x & G + D_x \circ F \end{bmatrix} \succeq 0 \quad \Leftrightarrow \quad X \succeq (I - D_x)(G + D_x \circ F)^{-1}(I - D_x).$$

Thus, to minimize trace (X) subject to the inequality constraint, we take

$$X = (I - D_x)(G + D_x \circ F)^{-1}(I - D_x),$$

which shows the equivalence between the objective functions in (13.4) and in (13.2).

Thus, the noise-free leader selection problem (LS2) can be formulated as

$$\begin{aligned} \underset{X,\,x}{\text{minimize}} \quad & \operatorname{trace}(X) \\ \text{subject to} \quad & \begin{bmatrix} X & I - D_x \\ I - D_x & G + D_x \circ F \end{bmatrix} \succeq 0 \\ & G = (I - D_x)\, F\, (I - D_x) \\ & D_x = \operatorname{diag}(x), \quad \mathbb{1}^T x = N_l, \quad x_i \in \{0,1\}, \quad i = 1, \ldots, n. \end{aligned} \tag{13.5}$$

In addition to the Boolean constraints, the quadratic dependence of G on D_x provides another source of nonconvexity in (13.5). Thus, in contrast to (LS1), relaxation of the Boolean constraints to $x_i \in [0,1]$ for $i = 1, \ldots, n$ is not enough to guarantee convexity of the optimization problem (13.5).

13.2.2 Linear approximation of G

As established in Section 13.2.1, the alternative formulation (13.5) of the noise-free leader selection problem (LS2) identifies two sources of nonconvexity: the quadratic matrix inequality and the Boolean constraints. In view of this, we use linearization of the matrix G to approximate the quadratic matrix inequality in (13.5) with an LMI. Furthermore, instead of imposing Boolean constraints, we augment the objective function with the ℓ_1 norm of x. This choice is used as a proxy for obtaining a sparse solution x whose nonzero elements identify the leaders.

The idea of using linearization comes from [22], where a linear approximation of the objective function $\operatorname{trace}(YZ)$ at the point (Y_0, Z_0) was considered

$$\frac{1}{2} \operatorname{trace}(Y_0 Z + Y Z_0).$$

To design fixed-order output feedback controllers, the authors of [22] minimize trace $(Y_0 Z + Y Z_0)$ with respect to Y and Z, set $Y_0 \leftarrow Y$, $Z_0 \leftarrow Z$, and repeat. Motivated by this iterative scheme, we consider the following linear approximation of G

$$G_0 := \frac{1}{2}(I - D_x) F (I - D_{x_0}) + \frac{1}{2}(I - D_{x_0}) F (I - D_x) \tag{13.6}$$

where D_{x_0} is our current-best-estimate of D_x. Replacing G with G_0 leads to an LMI approximation of the quadratic matrix inequality in (13.5).

In addition to the linearization, we relax the *hard* constraint $\mathbb{1}^T x = N_l$ for Boolean-valued x with a *soft* one. This is achieved by augmenting the objective function with the ℓ_1 norm of x,

$$\text{trace}\,(X) + \gamma \|x\|_{\ell_1}$$

where, as discussed in Section 2.1, the positive number γ characterizes our emphasis on the sparsity of the vector x. Putting this soft-constraint approach and linearization (13.6) together, we obtain a *convex* optimization problem

$$\begin{array}{ll} \underset{X,\,x}{\text{minimize}} & \text{trace}\,(X) + \gamma \|x\|_{\ell_1} \\ \text{subject to} & \begin{bmatrix} X & I - D_x \\ I - D_x & G_0 + D_x \circ F \end{bmatrix} \succeq 0 \\ & G_0 = \frac{1}{2}(I - D_x) F (I - D_{x_0}) + \frac{1}{2}(I - D_{x_0}) F (I - D_x) \\ & D_x = \text{diag}\,(x) \end{array} \tag{13.7}$$

which can be solved efficiently for small size problems (e.g., $n \leq 30$) using standard SDP solvers. For large problems, we develop a customized algorithm in Section 13.2.3.

For a fixed value of γ, we start with $D_{x_0} = 0$ and solve problem (13.7) as part of an iterative loop; the solution $D_x = \text{diag}\,(x)$ at every iteration is treated as the

current-best-estimate $D_{x_0} = \mathrm{diag}\,(x_0)$ for the linearization in the next iteration until $\|x - x_0\|_2 \leq \epsilon$. Ranging γ from small to large values, the solution to the γ-parameterized family of problems (13.7) provides a trade-off between minimization of $\mathrm{trace}\,(X)$ and minimization of $\|x\|_{\ell_1}$. Larger values of γ promote smaller $\|x\|_{\ell_1}$ and *typically* lead to fewer nonzero elements in x. Depending on the structure of the network, there may not exist values of γ that lead to a vector x with exactly N_l nonzero elements. In this case, we find the solution x^* that has the least number of nonzero elements N^* with $N^* > N_l$, and use the indices of the N_l largest entries of x^* to determine the leaders.

13.2.3 ADMM for the soft-constraint method

We next employ ADMM for the soft-constraint method developed in Section 13.2.2. We consider the following minimization problem

$$\underset{x}{\text{minimize}} \quad f(x) + \gamma \|x\|_{\ell_1}$$

where f is the convex approximation of (13.2)

$$f(x) = \mathrm{trace}\left((I - D_x)(G_0 + D_x \circ F)^{-1}(I - D_x)\right)$$

and G_0 is the linear approximation of G given by (13.6). This problem is equivalent to the constrained problem

$$\begin{aligned} \underset{x,\,z}{\text{minimize}} \quad & f(x) + \gamma \|z\|_{\ell_1} \\ \text{subject to} \quad & x - z = 0 \end{aligned}$$

and the associated augmented Lagrangian function is given by

$$\mathcal{L}_\rho(x, z, \lambda) = f(x) + \gamma \|z\|_{\ell_1} + \lambda^T(x - z) + \frac{\rho}{2}\|x - z\|_2^2.$$

By completion of squares in $\mathcal{L}_\rho$ with respect to z, the z-minimization problem (12.12b) can be expressed as

$$\underset{z}{\text{minimize}} \;\; \gamma \|z\|_{\ell_1} + \frac{\rho}{2} \|z - v^k\|_2^2$$

where $v^k = x^{k+1} + (1/\rho)\lambda^k$. The solution is given by the soft thresholding operator (e.g., see [50, Section 4.4.3])

$$z_i^* = \mathcal{S}_{\gamma/\rho}(v_i^k) = \begin{cases} \left(1 - \dfrac{\gamma/\rho}{|v_i^k|}\right) v_i^k, & |v_i^k| > \gamma/\rho \\ 0, & |v_i^k| \leq \gamma/\rho \end{cases} \tag{13.8}$$

for $i = 1, \ldots, n$. On the other hand, by completing squares in $\mathcal{L}_\rho$ with respect to x, we obtain

$$\underset{x}{\text{minimize}} \;\; \phi(x) = f(x) + \frac{\rho}{2} \|x - u^k\|_2^2$$

where $u^k = z^k - (1/\rho)\lambda^k$. This problem can be solved using descent methods (e.g., gradient method [43]). Here, we provide the expression for the gradient of ϕ

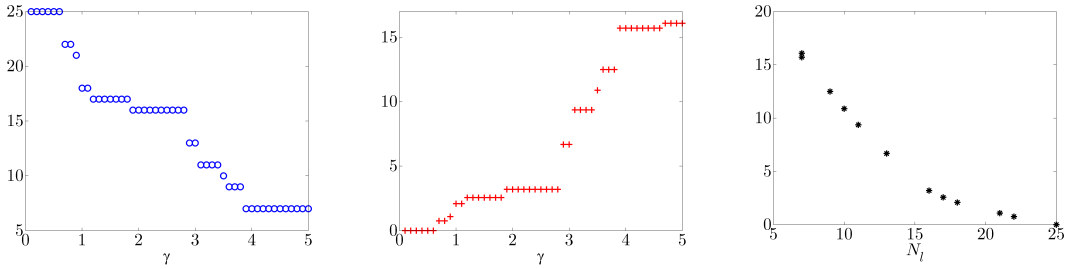
$$\begin{aligned} \nabla\phi(x) \;\; = \;\; & -2\,\text{diag}\left((I - D_x)M^{-1}\right) + \text{diag}\left(F(I - D_{x_0})M^{-1}(I - D_x)^2 M^{-1}\right) \\ & - \text{diag}\left(M^{-1}(I - D_x)^2 M^{-1}\right) \circ \text{diag}\,(F) + \rho(x - u^k) \end{aligned}$$

where $M = G_0 + D_x \circ F$.

13.3 Examples

13.3.1 A small network

We next use the soft-constraint method of Section 13.2.2 to select leaders for a small network with 25 nodes shown in Fig. 12.2. As shown in Figs. 13.1a and 13.1b, the number of leaders N_l decreases and the variance J_f of the followers increases as γ increases. The trade-off between the number of leaders and the variance of followers is



(a) Number of leaders N_l (b) Variance of the network J_f (c) Trade-off between N_l and J_f

Figure 13.1: Performance of the soft-constraint method for the network shown in Fig. 12.2: (a) the number of leaders N_l decreases as γ increases; (b) the variance of the followers J_f increases as γ increases; and (c) the trade-off between N_l and J_f.

illustrated in Fig. 13.1c.

Figure 13.2 compares performance of the soft-constraint method to performance of the greedy algorithm [41,42,84], which chooses one leader at a time by assigning the node that provides the largest performance improvement as a leader. Using a supermodular optimization framework, it was shown in [42] that the greedy algorithm selects noise-free leaders that are within a provable performance bound from the global solution to (LS2). This motivates us to use greedy algorithm as a benchmark for performance of the soft-constraint method. As shown in Fig. 13.2a, for a small number of leaders (e.g., $N_l \leq 5$), the greedy algorithm outperforms the soft-constraint method; the only exception happens for $N_l = 3$. A more detailed comparison is reported in Table 13.1, with the global solution to (LS2) for $N_l \leq 5$ obtained using the exhaustive search.

When the number of leaders is large (e.g., $N_l \geq 9$), the soft-constraint method outperforms the greedy algorithm; see Fig. 13.2b. The heuristics of assigning nodes with large degrees (i.e., large number of neighbors) as leaders is outperformed by both greedy and soft-constraint methods. The poor performance of the simple degree-heuristics-based-selection was also noted in [41, 42, 84].

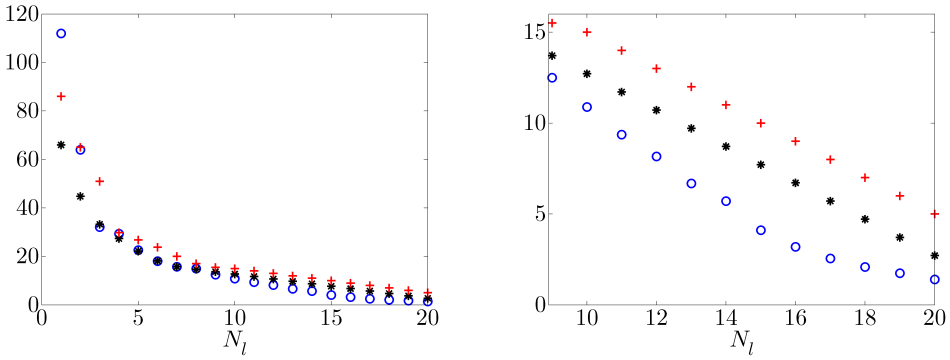


Figure 13.2: (a) The variance of the followers J_f obtained using the soft-constraint method ($\circ$), the greedy algorithm ($*$), and the degree heuristics ($+$) for the network shown in Fig. 12.2. (b) Comparison of three algorithms for $N_l \geq 9$.

Table 13.1: Performance comparison of greedy algorithm and soft-constraint method with the global solution to the noise-free leader selection problem (LS2) for the network shown in Fig. 12.2.

	global solution		greedy algorithm		soft-constraint	
N_l	J_f	leaders	J_f	leaders	J_f	leaders
1	66.0	13	66.0	13	112.0	25
2	38.4	8, 25	44.8	13, 25	64.0	16, 25
3	30.0	8, 16, 25	33.3	7, 13, 25	32.1	7, 16, 25
4	25.3	7, 9, 16, 25	27.4	7, 13, 16, 25	29.4	7, 16, 20, 25
5	20.7	3, 7, 9, 16, 25	22.2	3, 7, 13, 16, 25	22.6	3, 7, 16, 20, 25

13.3.2 A random network

We next consider the selection of noise-free leaders in a network with 100 randomly distributed nodes in a unit square. A pair of nodes can communicate with each other if their distance is not greater than 0.2. This scenario arises in sensor networks with prescribed omnidirectional (i.e., disk shape) sensing range [4]. As shown in Figs. 13.3a and 13.3b, the number of leaders N_l decreases and the variance J_f of followers increases with γ; also see the trade-off curve between N_l and J_f in Fig. 13.3c.

For this random network example, we observe similar selection of leaders and similar performance of the soft-constraint and greedy algorithms. Furthermore, for $N_l > 1$, both these algorithms significantly outperform the degree-heuristics-based-selection; see Fig. 13.5. To gain some insight into the selection of leaders, we compare the results obtained using soft-constraint method and the degree heuristics. As shown in Fig. 13.4b, the degree heuristics chooses nodes that turn out to be in the proximity of each other. In contrast, the soft-constraint method select leaders that, in addition to having large degrees, are far from each other; see Fig. 13.4a. As a result, the selected leaders can influence more followers and thus more effectively improve the performance of the network.

The contrast between the degree-heuristics-based method and the soft-constraint method becomes even more dramatic for large number of leaders. As shown in Figs. 13.4c and 13.4d, the leader sets obtained using the soft-constraint method and degree heuristics are almost *complements* of each other. While the degree heuristics clusters the leaders around the center of the network, the soft-constraint method distributes the leaders around the boundary of the network.

Figures 13.6a and 13.6b show the degree distribution of all the nodes in the random network and of the 41 nodes that are selected as leaders (see Fig. 13.4c). In contrast to the degree heuristics, the soft-constraint method chooses nodes with both large- and

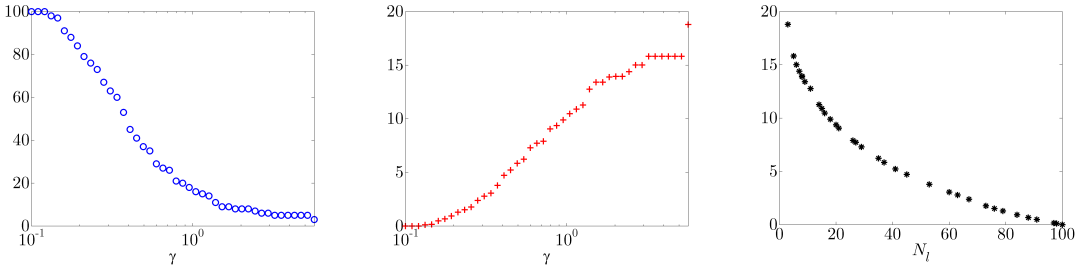
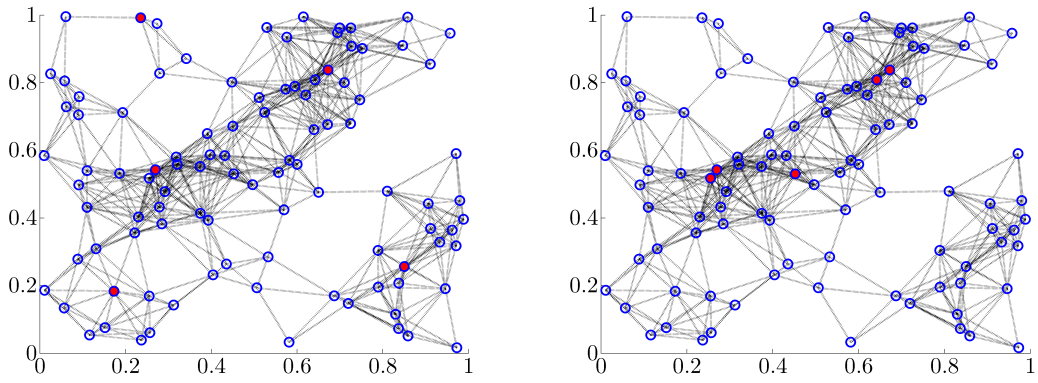


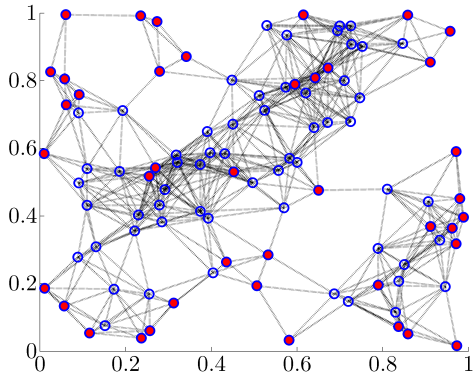
Figure 13.3: Computational results of a random network with 100 nodes: (a) the number of leaders N_l decreases as γ increases; (b) the variance of the followers J_f increases as γ increases; and (c) the trade-off curve between N_l and J_f.

small-degrees as leaders; in particular, all nodes with degree less than 8 and all nodes with degree greater than 18 are selected.

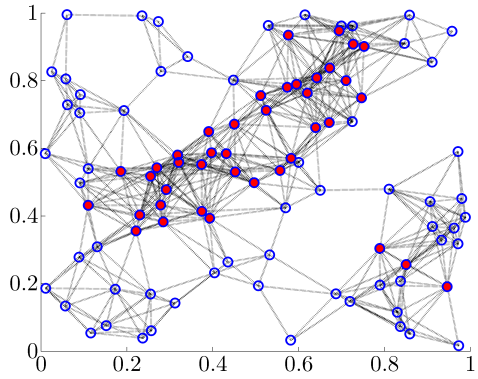


(a) $N_l = 5$

(b) $N_l = 5$



(c) $N_l = 41$



(d) $N_l = 40$

Figure 13.4: Selection of leaders (●) for the random network example using soft-constraint method in (a) and (c) and using degree-heuristics-based method in (b) and (d).

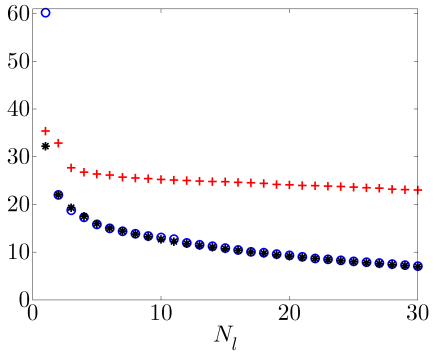
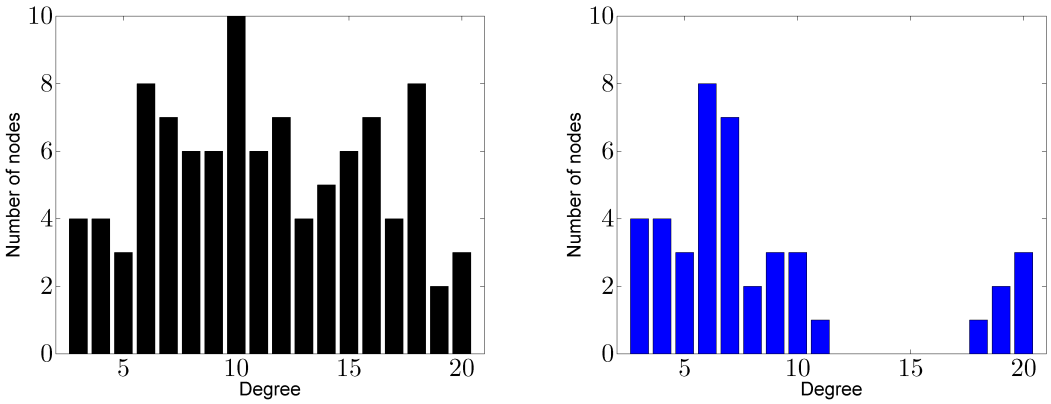


Figure 13.5: The objective function J_f obtained using the soft-constraint method ($\circ$), the greedy algorithm ($*$), and the degree heuristics ($+$) for the random network.



(a) (b)

Figure 13.6: The degree distribution of (a) the random network of Section 13.3.2 and of (b) 41 leaders selected using soft-constraint method. Note that the soft-constraint method chooses all nodes with degree less than 8 and all nodes with degree greater than 18.

Chapter 14

Conclusions and future directions

Conclusions

We develop efficient algorithms that facilitate selection of leaders in large stochastically forced consensus networks. For the noise-corrupted leader selection problem (LS1), we focus on computing lower and upper bounds on the global optimal value. A lower bound is obtained by solving a convex relaxation, and upper bounds result from a simple but efficient greedy algorithm and the alternating direction method of multipliers. For the noise-free leader selection problem (LS2), we provide an explicit expression for the variance amplification of the network. This allows us to identify sources of nonconvexity and to propose a convex relaxation of the objective function in (LS2). Furthermore, we use augmentation of the objective function with the ℓ_1 norm of the vector of optimization variables as a surrogate for obtaining a sparse solution whose nonzero elements identify the leaders.

Ongoing research and future directions

Leader selection in real-world networks. We are currently applying the developed algorithms for leader selection problems in different types of real-world networks,

such as collaboration of scientists [82] and power grid networks [85]. Our preliminary results show interesting connections between the selection of leaders and the influential nodes identified using metrics from social networks [86]. Studies along these lines (e.g., [87]) will bridge the gap between notions in network science and tools from control theory and optimization in understanding complex dynamical networks.

Leader selection in directed networks. Recently, controllability of directed networks has received considerable attention in network science [88]. Inspired by this effort, we also study the effect of leaders in directed trees and directed lattices. In [89], we consider the steady-state variance distribution in directed lattices with leaders being placed along the boundary of the lattice. We show that the variance of followers along the diagonal of the 2D lattice scales as a logarithmic function of node indices

Phase transition phenomenon. As observed in the random network example in Section 13.3.2, the selection of leaders switches from nodes with large degrees (when the number of leaders is small) to nodes with small degrees (when the number of leaders is large enough). It is of interest to study under what conditions this phase transition of leader selection from large degree nodes to small degree nodes takes place. Tools from random graphs [90, 91] will play an important role in this effort.

References

[1] R. Albert and A. L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97, 2002.

[2] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[3] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks.* Princeton University Press, 2009.

[4] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks.* Princeton University Press, 2010.

[5] B. Bamieh, F. Paganini, and M. A. Dahleh. Distributed control of spatially invariant systems. *IEEE Trans. Automat. Control*, 47(7):1091–1107, 2002.

[6] G. A. de Castro and F. Paganini. Convex synthesis of localized controllers for spatially invariant system. *Automatica*, 38:445–456, 2002.

[7] P. G. Voulgaris, G. Bianchini, and B. Bamieh. Optimal H^2 controllers for spatially invariant systems with delayed communication requirements. *Syst. Control Lett.*, 50:347–361, 2003.

[8] R. D'Andrea and G. E. Dullerud. Distributed control design for spatially interconnected systems. *IEEE Trans. Automat. Control*, 48(9):1478–1495, 2003.

[9] C. Langbort, R. S. Chandra, and R. D'Andrea. Distributed control design for systems interconnected over an arbitrary graph. *IEEE Trans. Automat. Control*, 49(9):1502–1519, 2004.

[10] X. Qi, M. Salapaka, P. Voulgaris, and M. Khammash. Structured optimal and robust control with multiple criteria: A convex solution. *IEEE Trans. Automat. Control*, 49(10):1623–1640, 2004.

[11] B. Bamieh and P. G. Voulgaris. A convex characterization of distributed control problems in spatially invariant systems with communication constraints. *Syst. and Control Lett.*, 54:575–583, 2005.

[12] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Trans. Automat. Control*, 51(2):274–286, 2006.

[13] N. Motee and A. Jadbabaie. Optimal control of spatially distributed systems. *IEEE Trans. Automat. Control*, 53(7):1616–1629, 2008.

[14] F. Borrelli and T. Keviczky. Distributed LQR design for identical dynamically decoupled systems. *IEEE Trans. on Automat. Control*, 53(8):1901–1912, 2008.

[15] J. Swigart and S. Lall. An explicit state-space solution for a decentralized two-player optimal linear-quadratic regulator. In *Proceedings of the 2010 American Control Conference*, pages 6385–6390, 2010.

[16] M. Fardad and M. R. Jovanović. Design of optimal controllers for spatially invariant systems with finite communication speed. *Automatica*, 47(5):880–889, May 2011.

[17] P. Shah and P. A. Parrilo. H_2-optimal decentralized control over posets: A state-space solution for state-feedback. *IEEE Trans. Automat. Control*, 2011. submitted; also arXiv:1111.1498v1.

[18] L. Lessard and S. Lall. Optimal controller synthesis for the decentralized two-player problem with output feedback. In *Proceedings of the 2012 American Control Conference*, pages 6314–6321, 2012.

[19] G. E. Dullerud and F. Paganini. *A course in robust control theory*. Springer, 2000.

[20] A. Zečević and D. D. Šiljak. *Control of Complex Systems, Structural Constraints and Uncertainty*. Springer Verlag, 2010.

[21] P. L. D. Peres and J. C. Geromel. An alternate numerical solution to the linear quadratic problem. *IEEE Trans. Automat. Control*, 39:198–202, 1994.

[22] L. El Ghaoui, F. Oustry, , and M. AitRami. A cone complementarity linearization algorithm for static output-feedback and related problems. *IEEE Trans. Automat. Control*, 42(8):1171–1176, 1997.

[23] V. L. Syrmos, C. T. Abdallah, P. Dorato, and K. Grigoriadis. Static output feedback-a survey. *Automatica*, 33(2):125–137, 1997.

[24] M. C. de Oliveira and J. C. Geromel. Numerical comparison of output feedback design methods. In *Proceeding of the 1997 American Control Conference*, pages 72–76, 1997.

[25] J. C. Geromel, C. C. de Souza, and R. E. Skelton. Static output feedback controllers: stability and convexity. *IEEE Trans. Automat. Control*, 43(1):120–125, 1998.

[26] V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 2000.

[27] K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice-Hall, 1996.

[28] M. H. DeGroot. Reaching a consensus. *J. Amer. Statist. Assoc.*, 69(345):118–121, 1974.

[29] B. Golub and M. Jackson. Naive learning social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, 2010.

[30] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *J. Parallel Distrib. Comput.*, 7(2):279–301, 1989.

[31] J. E. Boillat. Load balancing and Poisson equation in a graph. *Concurrency: Practice and Experience*, 2(4):289–313, 1990.

[32] 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.

[33] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Automat. Control*, 48(6):988–1001, 2003.

[34] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automat. Control*, 49(9):1520–1533, 2004.

[35] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Control*, 50(2):169–182, 2005.

[36] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri. Communication constraints in the average consensus problem. *Automatica*, 44(3):671–684, 2007.

[37] B. Bamieh, M. R. Jovanović, P. Mitra, and S. Patterson. Coherence in large-scale networks: dimension dependent limitations of local feedback. *IEEE Trans. Automat. Control*, 57(9):2235–2249, September 2012.

[38] P. Seiler, A. Pant, and K. Hedrick. Disturbance propagation in vehicle strings. *IEEE Trans. Automat. Control*, 49(10):1835–1841, 2004.

[39] M. R. Jovanović and B. Bamieh. On the ill-posedness of certain vehicular platoon control problems. *IEEE Trans. Automat. Control*, 50(9):1307–1321, September 2005.

[40] M. R. Jovanović, J. M. Fowler, B. Bamieh, and R. D'Andrea. On the peaking phenomenon in the control of vehicular platoons. *Syst. Control Lett.*, 57(7):528–537, July 2008.

[41] S. Patterson and B. Bamieh. Leader selection for optimal network coherence. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 2692–2697, 2010.

[42] A. Clark, L. Bushnell, and R. Poovendran. A supermodular optimization framework for leader selection under link noise in linear multi-agent systems. *IEEE Trans. Automat. Control*, 2012. submitted; also arXiv:1208.0946v1.

[43] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[44] E. J. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.*, 59(8):1207–1223, 2006.

[45] R. Tibshirani. Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc B.*, 58(1):267–288, 1996.

[46] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM J. on Sci. Comp.*, 20(1):33–61, 1998.

[47] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algoirthms. *Phys. D*, 60(1-4):259–268, 1992.

[48] E. J. Candès, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl*, 14:877–905, 2008.

[49] M. R. Jovanović, M. Arcak, and E. D. Sontag. A passivity-based approach to stability of spatially distributed systems with a cyclic interconnection structure. *IEEE Trans. Automat. Control: Special Issue on Systems Biology*, 53:75–86, January 2008.

[50] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–124, 2011.

[51] B. D. O. Anderson and J. B. Moore. *Linear optimal control.* Prentice-Hall, 1971.

[52] P. M. Mäkilä and H. T. Toivonen. Computational methods for parametric LQ problems – a survey. *IEEE Trans. Automat. Control*, 32(8):658–671, 1987.

[53] T. Rautert and E. W. Sachs. Computational design of optimal output feedback controllers. *SIAM J. Optim.*, 7(3):837–852, 1997.

[54] D. P. Bertsekas. *Nonlinear programming.* Athena Scientific, 1999.

[55] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer, 1999.

[56] D. G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Springer, 2008.

[57] A. Nedić and A. Ozdaglar. Distributed subgradient methods for multiagent optimization. *IEEE Trans. on Automat. Control*, 54(1):48–61, 2009.

[58] E. Wei, A. Ozdaglar, and A. Jadbabaie. A distributed newton method for network utility maximization. *arXiv:1005.2633 [math.OC]*, 2011.

[59] G. F. Young, L. Scardovi, and N. E. Leonard. Robustness of noisy consensus dynamics with directed communication. In *Proceedings of the 2010 American Control Conference*, pages 6312–6317, 2010.

[60] D. Zelazo and M. Mesbahi. Edge agreement: Graph-theoretic performance bounds and passivity analysis. *IEEE Trans. Automat. Control*, 56(3):544–555, 2011.

[61] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[62] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[63] W. S. Levine and M. Athans. On the optimal error regulation of a string of moving vehicles. *IEEE Trans. Automat. Control*, 11(3):355–361, 1966.

[64] 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.

[65] D. Swaroop and J. Hedrick. String stability of interconnected systems. *IEEE Trans. Automat. Control*, 41(3):349–356, 1996.

[66] J. A. Fax and R. M. Murray. Information flow and cooperative control of vehicle formations. *IEEE Trans. Automat. Control*, 49(9):1465–1476, 2004.

[67] G. Lafferriere, A. Williams, J. Caughman, and J. J. P. Veerman. Decentralized control of vehicle formations. *Syst. Control Lett.*, 54:899–910, 2005.

[68] P. Barooah, P. G. Mehta, and J. P. Hespanha. Mistuning-based control design to improve closed-loop stability margin of vehicular platoons. *IEEE Trans. Automat. Control*, 54(9):2100–2113, 2009.

[69] R. H. Middleton and J. H. Braslavsky. String instability in classes of linear time invariant formation control with limited communication range. *IEEE Trans. Automat. Control*, 55(7):1519–1530, 2010.

[70] B. Bamieh and M. Dahleh. Exact computation of traces and H^2 norms for a class of infinite dimensional problems. *IEEE Trans. Automat. Control*, 48(4):646–649, 2003.

[71] G. Meurant. A review of the inverse of symmetric tridiagonal and block tridiagonal matrices. *SIAM J. Matrix Anal. Appl.*, 13(3):707–728, 1992.

[72] R. E. Kalman. When is a linear control system optimal? *J. Basic Eng.*, 86:51–60, 1964.

[73] M. R. Jovanović. On the optimality of localized distributed controllers. *Int. J. Systems, Control and Communications*, 2(1/2/3):82–99, 2010. Special issue on Information Processing and Decision Making in Distributed Control Systems.

[74] S. Boyd, P. Diaconis, P. Parrilo, and L. Xiao. Fastest mixing Markov chain on graphs with symmetries. *SIAM J. Optim.*, 20(2):792–819, 2009.

[75] S. Joshi and S. Boyd. Sensor selection via convex optimization. *IEEE Trans. Signal Process.*, 57(2):451–462, 2009.

[76] P. Barooah and J. P. Hespanha. Estimation on graphs from relative measurements: Distributed algorithms and fundamental limits. *IEEE Control Systems Magazine*, 27(4):57–74, 2007.

[77] P. Barooah and J. P. Hespanha. Estimation from relative measurements: Electrical analogy and large graphs. *IEEE Trans. Signal Process.*, 56(6):2181–2193, 2008.

[78] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1968.

[79] C. D. Meyer. Generalized inversion of modified matrices. *SIAM Journal of Applied Mathematics*, 24(3):315–323, 1973.

[80] D. A. Spielman. Algorithms, graph theory, and linear equations in Laplacian matrices. *Proceedings of the International Congress of Mathematicians*, IV:2698–2722, 2010.

[81] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *Bell System Technical Journal*, 49:291–307, 1970.

[82] M. E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 74:036104, 2006.

[83] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt. Controllability of multi-agent systems from a graph theoretic perspective. *SIAM J. Control Optim.*, 48(1):162–186, 2009.

[84] A. Clark and R. Poovendran. A submodular optimization framework for leader selection in linear multi-agent systems. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 3614–3621, 2011.

[85] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[86] S. P. Borgatti and M. G. Everett. A graph-theoretic perspective on centrality. *Social networks*, 28(4):466–484, 2006.

[87] I. Poulakakis, L. Scardovi, and N. E. Leonard. Node classification in networks of stochastic evidence accumulators. *arXiv:1210.4235 [cs.SY]*, 2012.

[88] Y. Y. Liu, J. J. Slotine, and A. L. Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

[89] F. Lin, M. Fardad, and M. R. Jovanović. Performance of leader-follower consensus algorithms in directed trees and lattices. In *Proceedings of the 51th IEEE Conference on Decision and Control*, Maui, HI, 2012. to appear.

[90] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.

[91] R. Durrett. *Random Graph Dynamics*. Cambridge University Press, 2007.

[92] S. Haykin. *Nonlinear Methods of Spectral Analysis*. Springer-Verlag, 1979.

[93] P. Stoica and R. Moses. *Introduction to Spectral Analysis*. Prentice Hall, 1997.

[94] T. T. Georgiou. Spectral estimation via selective harmonic amplification. *IEEE Trans. Automat. Control*, 46(1):29–42, 2001.

[95] C. I. Byrnes, T. T. Georgiou, and A. Lindquist. A new approach to spectral estimation: a tunable high-resolution spectral estimator. *IEEE Trans. Signal Process.*, 48(11):3189–3205, 2000.

[96] C. I. Byrnes, T. T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Trans. Automat. Control*, 45(6):822–839, 2001.

[97] T. T. Georgiou. Structured covariances and related approximation questions. In *Directions in Mathematical Systems Theory and Optimization*, pages 135–140. Springer-Verlag, 2003.

[98] T. T. Georgiou. The structure of state covariances and its relation to the power spectrum of the input. *IEEE Trans. Automat. Control*, 47(7):1056–1066, 2002.

[99] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[100] L. J. Goldstein. *Abstract Algebra; A First Course.* Prentice Hall, 1973.

[101] N. J. Higham. Computing the nearest correlation matrix - a problem from finance. *IMA J. Numer. Analysis*, 22:329–343, 2002.

[102] J. Malick. A dual approach to semidefinite least-squares problems. *SIAM J. Matrix Anal. Appl.*, 26(1):272–284, 2004.

[103] S. Boyd and L. Xiao. Least-squares covariance matrix adjustment. *SIAM J. Matrix Anal. Appl.*, 27(2):532–546, 2005.

[104] H. Qi and D. Sun. A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM J. Matrix Anal. Appl.*, 26:360–385, 2006.

[105] L. Qi and J. Sun. A nonsmooth version of Newton's method. *Math. Programming*, 58:353–367, 1993.

[106] T. Kato. *A Short Introduction to Perturbation Theory for Linear Operators*. Springer-Verlag, 1982.

[107] J. Löberg. Yalmip: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

Appendix A

Least-squares approximation of structured covariances

The use of second order statistics has been extensively studied in spectral estimation [92–94]. Recently, there has been renewed interest [95, 96] in utilizing state covariances of linear filters to extract information about the power spectra of the input processes. To qualify as a valid state covariance, a positive semi-definite matrix has to satisfy a certain linear constraint imposed by the underlying dynamics. However, the sample covariances, computed from a finite measurement record, almost always fail to have the required structure. Most methods in spectral estimation [93] take sample covariances even though the effect of inaccuracy is not well understood nor analyzed in any detail [97].

In view of the above, it is pertinent to find a nonnegative definite matrix with required structure to approximate the given sample covariance. The natural Euclidean distance gives a least-squares problem which can be solved by standard semi-definite programming (SDP) solvers. For the $n \times n$ covariance matrix, however, the number of optimization variables is of $O(n^2)$, which implies numerical difficulty (computational effort of $O(n^6)$) of the interior-point methods employed in available SDP solvers. In this

appendix, we develop an alternative approach to this optimization problem.

Our presentation is organized as follows: we set up the problem and give an equivalent formulation in Section A.1. We derive the dual problem and present the unconstrained optimization methods in Section A.2. Then, a numerical example is provided with the computational results presented in Section A.3. The appendix is concluded with a brief summary in Section A.4.

A.1 Problem formulation

Let a finite dimensional linear system be given by its state equation

$$\dot{x} \;=\; Ax \;+\; Bd$$

where $d \in \mathbb{C}^m$ is a stationary, zero-mean stochastic process and $x \in \mathbb{C}^n$ is the state vector. The system is characterized by the controllable pair (A, B), where $A \in \mathbb{C}^{n\times n}$ is Hurwitz, and $B \in \mathbb{C}^{n\times m}$ is full column rank. Under these assumptions, the steady state covariance $X := \lim_{t\to\infty} \mathcal{E}\{x(t)x^*(t)\}$ satisfies the following linear constraint (cf. [98])

$$AX \;+\; XA^* \;=\; -(BH \;+\; H^*B^*) \tag{LC}$$

where $\mathcal{E}(\cdot)$ is the expectation operator and $(\cdot)^*$ is the complex conjugate transpose. The matrix $H \in \mathbb{C}^{m\times n}$ depends on the input power spectrum and the pair (A, B). It was also established in [98] that the condition for a positive semi-definite matrix X to be the state covariance of a linear system (A, B) for some stationary, zero-mean, stochastic input d, is equivalent to the solvability of (LC) in terms of H. However, the sample covariance

$$\Sigma \;:=\; \frac{1}{k}\sum_{i=1}^{k} x_i x_i^*$$

computed from k samples almost always fails to satisfy (LC) [97]. In view of this, we formulate the following approximation problem:

- *Given a positive semi-definite matrix $\Sigma = \Sigma^* \succeq 0$ and a controllable pair (A, B) with A Hurwitz and B full column rank, find $X = X^* \succeq 0$ that is closest to Σ in the least-squares sense and satisfies (LC) for some $H \in \mathbb{C}^{m\times n}$.*

This optimization problem can be formulated as follows:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|X - \Sigma\|_F^2 \\ \text{subject to} \quad & X = X^* \succeq 0 \\ & AX + XA^* = -(BH + H^*B^*) \end{aligned} \tag{P1}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, and X and H are the optimization variables.

A.1.1 Standard SDP formulation

The primal problem (P1) is a convex optimization problem with a norm objective function and a linear constraint in the positive semi-definite cone $\mathcal{S}_n^+$. By introducing an auxiliary variable κ, (P1) can be cast into an SDP problem [99],

$$\begin{aligned} \text{minimize} \quad & \kappa \\ \text{subject to} \quad & \frac{1}{2}\|X - \Sigma\|_F^2 \le \kappa \\ & X = X^* \succeq 0 \\ & AX + XA^* = -(BH + H^*B^*) \end{aligned}$$

which can be solved by standard primal-dual interior-point methods. However, the number of optimization variables is of $O(n^2)$, which implies the computational complexity $O(n^6)$ of these standard methods.

A.1.2 Equivalent constraints

For a given positive semi-definite matrix X, the solvability of (LC) in terms of H qualifies X to be a valid steady state covariance. However, having H as an optimization variable increases the problem size by $m \times n$, and computations become more expensive as the number of inputs m increases. We note that the Lyapunov-type constraint (LC) implies that X must lie in the range of a certain linear operator $\mathcal{L}$, i.e., $X \in \mathcal{R}(\mathcal{L})$. Namely, the constraint (LC) can be equivalently represented as

$$X \;=\; \int_0^\infty \mathrm{e}^{At}(BH \,+\, H^*B^*)\mathrm{e}^{A^*t}\,\mathrm{d}t \;=:\; \mathcal{L}(H)$$

where $\mathcal{L}$ maps H into X. Equivalently, X must be orthogonal to the null space of the adjoint of $\mathcal{L}$, i.e., $X \perp \mathcal{N}(\mathcal{L}^{ad})$. Next, we determine the basis of $\mathcal{N}(\mathcal{L}^{ad})$.

Real field case

Let us first consider the linear constraint over the field of real numbers. The complex conjugate transpose in (LC) is then replaced by transpose, $AX + XA^T = -(BH + H^TB^T)$. The linear operator $\mathcal{L}(H) := \int_0^\infty \mathrm{e}^{At}(BH + H^TB^T)\mathrm{e}^{A^Tt}\,\mathrm{d}t$ maps $\mathbb{R}^{m\times n}$ to $\mathbb{R}^{n\times n}$. Let symmetric matrix $G \in \mathbb{R}^{n\times n}$ be in the range of $\mathcal{L}$. The unique linear operator $\mathcal{L}^{ad}$ exists and satisfies $\langle G, \mathcal{L}(H)\rangle = \langle \mathcal{L}^{ad}(G), H\rangle$ where the inner product is defined as $\langle M, N\rangle := \mathrm{trace}\,(M^TN)$. Hence,

$$\begin{aligned}
\langle G, \mathcal{L}(H)\rangle \;&=\; \mathrm{trace}\,(G^T\int_0^\infty \mathrm{e}^{At}(BH + H^TB^T)\mathrm{e}^{A^Tt}\,\mathrm{d}t)\\
&=\; \int_0^\infty \mathrm{trace}\,(\mathrm{e}^{A^Tt}G\mathrm{e}^{At}(BH + H^TB^T)\,\mathrm{d}t)\\
&=\; 2\;\mathrm{trace}\,(\int_0^\infty \mathrm{e}^{A^Tt}G\mathrm{e}^{At}\,\mathrm{d}t\,BH).
\end{aligned}$$

Thus,

$$\mathcal{L}^{ad}(G) \;=\; 2B^T(\int_0^\infty \mathrm{e}^{A^T t} G \mathrm{e}^{At} \mathrm{d}t) \;=:\; 2B^T Z$$

where Z represents the solution to the Lyapunov equation $A^T Z + ZA = -G$. To construct the basis of $\mathcal{N}(\mathcal{L}^{ad})$, we introduce the change of coordinates $\{\tilde{B} = PB,\ \tilde{A} = PAP^{-1}\}$, such that $\tilde{B} = [I_{m\times m} \;\; O_{(n-m)\times m}]^T$. Under this coordinate transformation, the basis $\tilde{G}$ satisfies

$$\begin{aligned}\mathcal{L}^{ad}(\tilde{G}) \;&=\; \tilde{B}^T \tilde{Z} \;=\; \begin{bmatrix} I & O \end{bmatrix} \begin{bmatrix} \tilde{Z}_1 & \tilde{Z}_2^T \\ \tilde{Z}_2 & \tilde{Z}_3 \end{bmatrix} \\ &=\; \begin{bmatrix} \tilde{Z}_1 & \tilde{Z}_2^T \end{bmatrix} \;=\; \begin{bmatrix} O & O \end{bmatrix}.\end{aligned}$$

Therefore, any symmetric matrix $\tilde{Z}$ of the form

$$\tilde{Z} \;=\; \begin{bmatrix} O & O \\ O & \tilde{Z}_3 \end{bmatrix}$$

gives a member of the $\mathcal{N}(\mathcal{L}^{ad})$ by substituting $\tilde{Z}$ into $\tilde{A}^T\tilde{Z} + \tilde{Z}\tilde{A} = -\tilde{G}$. Matrices $\tilde{G}_i$'s are determined by substituting the basis elements of $\mathcal{S}_{(n-m)\times(n-m)}$ for $\tilde{Z}$. Thus, G_i's in the original coordinates $\{A,\ B\}$ are recovered by $G_i = P^T \tilde{G}_i P$. Finally, the Gram-Schmidt procedure is employed to orthonormalize G_i's. The number of basis elements, r, is easily determined by the size of $\tilde{Z}_3$, $r = 0.5(n-m)(n-m+1)$.

Complex field case

When the matrices are defined over the field of complex numbers, the previous inner product procedure fails to give a linear operator $\mathcal{L}^{ad}$, because H^* is not linear with respect to H. To circumvent this difficulty, we note that the bijection between a complex

matrix $X = X_r + jX_i$ and a real matrix $\bar{X}$ of the form

$$\bar{X} \;=\; \begin{bmatrix} X_r & -X_i \\ X_i & X_r \end{bmatrix}$$

is a ring isomorphism [100]. By mapping $\{A, B, H\}$ into $\{\bar{A}, \bar{B}, \bar{H}\}$, the constraint (LC) transforms to $\bar{A}\bar{X} + \bar{X}\bar{A}^T = -(\bar{B}\bar{H} + \bar{H}^T\bar{B}^T)$. This can be also verified by expanding (LC) and equating the real and imaginary parts on both sides of the resulting equation. Thus, the procedure of Section A.1.2 can be employed to construct the basis $\bar{G}_i$. Correspondingly, the number of basis elements, r, for the complex case is $r = (n-m)(2n-2m+1)$.

Thus, (LC) is now transformed into the following equivalent set of constraints

$$\text{trace}\,(G_i X) \;=\; 0, \quad i = 1, 2, \ldots, r. \tag{TC}$$

As already mentioned, the number of G_i's to span $\mathcal{N}(\mathcal{L}^{ad})$ is $r = 0.5(n-m)(n-m+1)$ (in the real case) and $r = (n-m)(2n-2m+1)$ (in the complex case). If $m \geq n$, then (LC) is always satisfied for some H; thus, we assume $m < n$ in the sequel. The transformation of (LC) to (TC) is advantageous for optimization because:

- it eliminates H, which contains $m \times n$ optimization variables;
- the number of corresponding trace constraints decreases as the number of inputs increases.

Examples where the number of inputs is close to the number of states, i.e. $m \approx n$, can be found in spatially distributed systems. A particular example, which is the main motivation for current developments, is encountered in wall-bounded shear flows of incompressible fluids (e.g., boundary layers subject to spatio-temporal excitation in

the form of a free-stream turbulence). The algorithms developed in this appendix are expected to be useful in the study of these problems.

The primal problem is now cast into

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|X - \Sigma\|_F^2 \\ \text{subject to} \quad & X = X^* \succeq 0 \\ & \text{trace}(G_i X) = 0, \quad i = 1, 2, \ldots, r \end{aligned} \tag{P2}$$

where G_i's form the orthonormal basis of $\mathcal{N}(\mathcal{L}^{ad})$. After solving (P2), the least-squares solution H can be found by a very cheap computation as follows. Let $D := AX^\star + X^\star A^*$, where $X^\star$ is the unique solution of (P2). Left-multiplication of (LC) with B^* and right-multiplication with B gives

$$B^*B(HB) + (HB)^*B^*B = -B^*DB. \tag{A.1}$$

With $X \in \mathcal{R}(\mathcal{L})$, there exists H such that HB is Hermitian (see Section 3, Remark 1 in [98]). Therefore, H can be computed by

$$H = -(B^*B)^{-1}(B^*D + MB^*)$$

where matrix $M := HB$ is obtained as the solution to the Lyapunov equation (A.1).

In the sequel, we study optimization problem (P2). The covariance matrix approximation problems have been recently studied by several research groups. Higham first introduced the *nearest correlation matrix* problem and proposed an alternating projection method [101]. Malick studied the *semi-definite least-squares* (SDLS) problem [102], which generalized $\mathcal{S}_n^+$ to any closed convex cone. He proposed a quasi-Newton algorithm and gave a dual interpretation for the alternating projection method as the standard

gradient algorithm (see Section 5.2 in [102]). Boyd and Xiao studied the *least-squares covariance adjustment problem* (LSCAP) [103], which is an extension of the SDLS with trace inequality constraints. They proposed a projected gradient method and exploited structure (such as sparsity) to reduce computational expense. The objective functions of the dual problems in both SDLS and LSCAP are not twice continuously differentiable. This implies that the convergence rates of the proposed methods in [102] and [103] are at best linear [104]. Utilizing recent results for strongly semi-smooth functions [105], Qi and Sun developed a generalized Newton method with quadratic convergence rate, which is highly efficient and outperforms quasi-Newton and projected gradient methods (as reported in [104]).

Although fruitful results have been developed for general problems (as evident from the above references), for our specific problem (P2) we derive and cast its dual problem into an unconstrained problem via standard optimization theory. We then implement unconstrained maximization methods proposed in [102–104].

A.2 Dual problem

In this section, the primal problem (P2) is cast into its dual problem via standard Lagrange multipliers method. It is then converted into an unconstrained maximization problem by projection on $\mathcal{S}_n^+$. To begin with, the Lagrangian [54] is formed by introducing the Lagrange multipliers $\nu_i \in \mathbb{C}$ and $Z \in \mathbb{C}^{n\times n}$

$$L(\nu, Z, X) := \frac{1}{2}\|\Sigma - X\|_F^2 - \text{trace}(ZX) + \sum_{i=1}^{r} \nu_i \text{trace}(G_i X)$$

with $Z = Z^* \succeq 0$ corresponding to the inequality constraint. The minimizer of $L(\nu, Z, X)$ over X satisfies $\partial L(\nu, Z, X)/\partial X = 0$, which gives $X_{\min} = \Sigma + Z - \sum_{i=1}^{r} \nu_i G_i$. By choosing $X = X_{\min}$ and denoting $G_\nu := \sum_{i=1}^{r} \nu_i G_i$, we have the

dual objective function

$$\begin{aligned} g(\nu, Z) &= -0.5\|Z - G_\nu\|_F^2 - \text{trace}((Z - G_\nu)\Sigma) \\ &= -0.5\|\Sigma + Z - G_\nu\|_F^2 + 0.5\|\Sigma\|_F^2. \end{aligned}$$

Thus, the dual problem is given by

$$\begin{aligned} &\text{maximize} \quad g(\nu, Z) = -0.5\|\Sigma + Z - G_\nu\|_F^2 + 0.5\|\Sigma\|_F^2 \\ &\text{subject to} \quad Z = Z^* \succeq 0. \end{aligned}$$

Note that any Hermitian matrix can be decomposed as $X = X_+ + X_-$ with X_+ and X_-, respectively, being the positive and negative semi-definite parts of X, i.e., $X_+ = \sum_{\bar{\lambda}_i > 0} \bar{\lambda}_i \bar{u}_i \bar{u}_i^*$ and $X_- = \sum_{\bar{\lambda}_i < 0} \bar{\lambda}_i \bar{u}_i \bar{u}_i^*$. Here $\bar{u}_1, \ldots, \bar{u}_n$ denote a set of orthonormal eigenvectors of X with the corresponding eigenvalues $\{\bar{\lambda}_1, \ldots, \bar{\lambda}_n\}$. Hence, the eigenvalue decomposition of X gives

$$\|X\|_F^2 = \text{trace}(\bar{U}\bar{\Lambda}\bar{U}^*\bar{U}\bar{\Lambda}\bar{U}^*) = \sum_{i=1}^{n} \bar{\lambda}_i^2.$$

Thus, for given vector ν, the choice $Z^\star = -(\Sigma - G_\nu)_-$ eliminates the negative eigenvalues of $\Sigma - G_\nu$. It follows that $Z^\star$ is the maximizer over $\mathcal{S}_n^+$. Hence, the dual problem is finally cast into the following *unconstrained maximization* problem

$$\text{maximize} \quad g(\nu) = -\frac{1}{2}\|(\Sigma - G_\nu)_+\|_F^2 + \frac{1}{2}\|\Sigma\|_F^2 \tag{D}$$

where $(\Sigma - G_\nu)_+$ is the corresponding positive semi-definite part of $\Sigma - G_\nu$. The operation $(\cdot)_+$ on a Hermitian matrix can be interpreted as the projection on the positive semi-definite cone $\mathcal{S}_n^+$. Specifically, one can compute the eigenvalue decomposition and replace the negative eigenvalues by zero to obtain the nonnegative definite part. The

dual variables are scalars ν_i's.

For the convex primal problem with *linear constraints*, feasibility is sufficient to guarantee strong duality [54, p. 504]. The primal problem constraints are easily seen feasible (for instance, $X = 0$). Therefore, there is no duality gap and the optimal can be obtained by solving the unconstrained dual problem (D). With the solution from dual problem given by $\nu_i^\star$'s, the optimal solution of the primal problem is determined by $X^\star = (\Sigma - G_{\nu^\star})_+$. One advantage of the dual formulation is that the number of the variables $r = 0.5(n-m)(n-m+1)$ is only a fraction of the size of the primal problem (P2), $0.5(n^2+n)$, when $m \approx n$. Furthermore, due to the absence of constraints any method for unconstrained maximization problem can be utilized.

A.2.1 Unconstrained maximization methods

We derive the gradient of the objective function using standard perturbation theory and borrow available results from [104] to obtain generalized second order derivative. Let λ_j be the eigenvalue of $\Sigma - G_\nu$ with the associated orthonormal eigenvector u_j. Define $\phi(\lambda) := \max(0, \lambda)$. Then, the objective function is given by

$$g(\nu) \;=\; -\frac{1}{2}\sum_{j=1}^{n} \phi^2(\lambda_j) \;+\; \frac{1}{2}\|\Sigma\|_F^2.$$

We now employ results from standard perturbation analysis [106] to determine the gradient of $g(\nu)$. For simplicity, consider a matrix F perturbed by νG, where G and F are Hermitian matrices and ν is a scalar. The derivative of an isolated eigenvalue λ of the resulting matrix $F - \nu G$ with respect to ν is given by $\partial\lambda/\partial\nu = -u^* G u$, where u is the unit eigenvector associated to λ. Therefore, the ith entry of ∇g is determined by

$\partial g(\nu)/\partial \nu_i = \sum_{\lambda_j > 0} \lambda_j u_j^* G_i u_j$, and it can be rewritten compactly as

$$\begin{aligned}
\frac{\partial g(\nu)}{\partial \nu_i} &= \sum_{\lambda_j > 0} \lambda_j u_j^* G_i u_j = \sum_{\lambda_j > 0} u_j^* (\Sigma - G_\nu)_+ G_i u_j \\
&= \text{trace}\Big(\sum_{\lambda_j > 0} u_j u_j^* (\Sigma - G_\nu)_+ G_i \Big) \\
&= \text{trace}(UU^* (\Sigma - G_\nu)_+ G_i) \\
&= \text{trace}((\Sigma - G_\nu)_+ G_i).
\end{aligned}$$

The quasi-Newton method utilizes the gradients of two consecutive steps to construct the approximation of the second order derivative (Hessian). The Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme is employed due to its efficiency for general problems. Starting with negative identity, the negative definite matrix H_s in our maximization problem is updated using the following BFGS scheme (see [54, p. 150])

$$H_s^+ = H_s + \frac{yy^T}{y^T s} - \frac{H_s s s^T H_s}{s^T H_s s}$$

where $y := \nabla g(\nu^+) - \nabla g(\nu)$ and $s := \nu^+ - \nu$ with ν^+ and ν denoting the current and previous step variables, respectively. The BFGS direction is determined by $v_{BFGS} = -H_s^{-1} \nabla g(\nu)$. Given ν_0, the first step ν_1 can be obtained by a gradient method which is necessary for the BFGS iterations to proceed.

The objective function is not twice continuously differentiable when $\Sigma - G_\nu$ has zero eigenvalues [104]. This implies that the classical Hessian needs to be generalized. The detailed discussion about the smoothness property of the objective function and the derivation of generalized Newton direction can be found in [104]. To compute

generalized $\nabla^2 g(\nu)$, one constructs the symmetric matrix Ω as follows,

$$\Omega_{ij} = \begin{cases} (\phi(\lambda_i) - \phi(\lambda_j))/(\lambda_i - \lambda_j) & \text{if } \lambda_i \neq \lambda_j \\ 1 & \text{if } \lambda_i = \lambda_j > 0 \\ 0 & \text{if } \lambda_i = \lambda_j \leq 0. \end{cases}$$

Then for any $\Delta\nu \in \mathbb{R}^r$, the generalized Hessian acting on $\Delta\nu$ is given by

$$\nabla^2 g(\nu)(\Delta\nu) = \begin{pmatrix} \operatorname{trace}(-U(\Omega \circ (U^* G_{\Delta\nu} U))U^* G_1) \\ \vdots \\ \operatorname{trace}(-U(\Omega \circ (U^* G_{\Delta\nu} U))U^* G_r) \end{pmatrix}$$

where $\circ$ denotes the Hadamard product, i.e. entrywise multiplication. By solving $\nabla^2 g(\nu)(v_{NT}) = -\nabla g(\nu)$ using conjugate gradient (CG) method [54, 104], we obtain the generalized Newton direction v_{NT}.

A.2.2 Implementation

With the three ascending directions $\nabla g, v_{BFGS}$ and v_{NT} determined above, the algorithm for the unconstrained problem is given next. We choose the standard backtracking line search as the step size method, with parameters $\alpha = 0.3$, $\beta = 0.5$ (see [43, p. 464]).

Algorithm

Start with initial point ν_0 (e.g., $\nu_0 = 0$) and at each step k

repeat:

1. project $\Sigma - G_{\nu_k}$ onto $\mathcal{S}_n^+$, then compute ascending direction v_k,
2. use the backtracking line search to determine step size t, then update $\nu_{k+1} = \nu_k + t v_k$.

until: stopping criterion $\|\nabla g(\nu)\|_2 < \epsilon$ is reached.

The convergence to the global optimal is guaranteed by the convexity of the problem.

A.2.3 Complexity analysis

The computational effort of each algorithm is studied in this section. At each step of the gradient method, the computational effort is $O(\max(n^3, rn^2))$ operations, where $O(n^3)$ operations are required for the eigenvalue decomposition of $\Sigma - G_\nu$ and $O(n^2)$ operations are required for computation of the matrix inner product $\text{trace}((\Sigma - G_\nu)_+ G_i)$ for each $i = 1, 2, \ldots, r$. As BFGS method uses gradient directions to form H_s, it requires the same amount of operations as the gradient method. The extra effort comes from computing the inverse of H_s, which requires $O(r^3)$ operations. Thus, each BFGS step costs $O(\max(n^3, rn^2, r^3))$ operations. To compute the generalized Newton direction, it takes $O(\max(n^3, rn^2))$ operations to determine $U(\Omega \circ (U^* G_{\Delta\nu} U))U^*$ in each CG step, where $O(rn^2)$ operations are required to form the sum $G_{\Delta\nu}$ and $O(n^3)$ operations are required for matrix multiplications. Thus, the cost of each CG step is $O(\max(n^3, rn^2))$ operations. The number of CG steps is $O(r)$ provided that the problem is well preconditioned [54, 104]. Therefore, each Newton step costs $O(\max(rn^3, r^2n^2))$ operations assuming $O(r)$ CG steps.

For a given problem with n states, the cost of the above algorithms relies heavily on the number of dual variables. From the construction of the basis in Section A.1.2, the number of inputs m dictates the size of $r = 0.5(n-m)(n-m+1)$. Thus, if the number of inputs is close to the number of states, i.e. $m \approx n$, presented algorithms are expected to be computationally efficient.

A.3 Numerical experiments

We present a mass-spring-damper example and compare the computation results for different formulations of the problem. As discussed in Section A.1, the original primal

problem (P1) is equivalent to primal problem (P2), which is then cast into the dual problem (D) in Section A.2. The first two formulations can be solved by available SDP solvers such as SeDuMi. Following [103, 104], we develop unconstrained maximization methods for the dual formulation. The following experiments are performed in MATLAB on a personal computer with 3.2GHz CPU and 2.5GB RAM. Primal problems (P1) and (P2) are solved using YALMIP [107] with SeDuMi as its SDP solver.

A.3.1 Mass-spring-damper example

We consider a mass-spring-damper system consisting of l masses and $l+1$ springs and dampers on a line as in Fig. A.1. (The dampers are not shown in the figure.) The dynamics of the ith mass m_i are given by

$$\begin{aligned} m_i\ddot{p}_i \;+\; b_{i-1}(\dot{p}_i \;-\; \dot{p}_{i-1}) \;+\; b_i(\dot{p}_i \;-\; \dot{p}_{i+1}) \quad + \\ k_{i-1}(p_i - p_{i-1}) \;+\; k_i(p_i - p_{i+1}) \quad = \quad d_i \end{aligned}$$

where p_i represents the displacement from a reference position of the ith mass, b_i is the damping coefficient of the ith damper, and k_i is the spring constant of the ith spring. We assign unit values to $\{m_i, b_i, k_i\}$, and assume that a stationary Gaussian white stochastic process, d_i, with zero-mean and unit variance is introduced to the ith mass. The first and the last masses are connected to fixed boundaries; hence, $\dot{p}_0 \equiv 0$, $p_0 \equiv 0$, $\dot{p}_{l+1} \equiv 0$, $p_{l+1} \equiv 0$. By selecting the state variables $x_1 := \mathrm{col}\,\{p_i\}$ and $x_2 := \mathrm{col}\,\{\dot{p}_i\}$, the state-space representation is determined by

$$A \;=\; \begin{bmatrix} O & I \\ T & T \end{bmatrix}, \quad B \;=\; \begin{bmatrix} O \\ I \end{bmatrix}$$

where $T := \mathrm{toeplitz}([\,-2\ 1\ 0\ \cdots\ 0\,])$, I is $l \times l$ identity matrix, O is $l \times l$ zero matrix, and input $d := \mathrm{col}\,\{d_i\}$. To demonstrate the performance with respect to different number

of inputs, we assume the first $l-m$ components in d to be identically equal to zero; the definition of input matrix B should be changed correspondingly in this case.

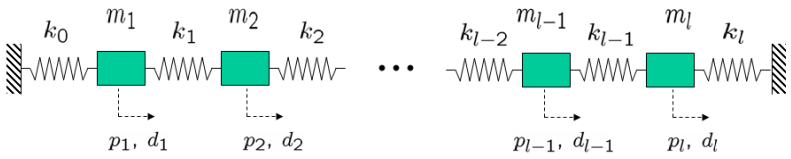


Figure A.1: Mass-spring-damper system.

Knowledge of the dynamics of the linear system and covariance Q of d can be used to obtain the steady state covariance X by solving the Lyapunov equation

$$AX \ + \ XA^T \ = \ -BQB^T.$$

However, we consider a situation in which only limited observed sample data is available to estimate state covariance and infer the second order statistics of the inputs. In the numerical experiments, we take 1000 state samples (uniformly sampled in time from 0 to 10 seconds) and compute the sample covariance Σ. Invariably, Σ fails to satisfy (LC) and we utilize approximation algorithms to find least-squares estimate of Σ.

A.3.2 Performance comparison of three formulations

As discussed in Section A.2.3, the computational effort of all unconstrained maximization algorithms depends heavily on the number of dual variables determined by $r = 0.5(n-m)(n-m+1)$ (in this example, $n = 2l$). Hence, two sets of optimization experiments are carried out with difference in the number of inputs m relative to the number of states n ($m = 0.1n$ in the first set and $m = 0.5n$ in the second set with the results shown in Tables A.1 and A.2, respectively). The time for all computations is

given in seconds. We only present results of BFGS method for the dual formulation because it generally outperforms gradient and generalized Newton methods. It is observed in this example that the conjugate gradient method usually runs into difficulty when $r \geq 100$. However, in numerical examples reported in [104], the generalized Newton method outperformed BFGS method significantly. The purpose of our experiments is not to provide comparison between different unconstrained maximization methods but rather to compare the different formulations of the problem in terms of their computational efficiency. For simplicity, the initial condition for BFGS method is set to be a zero vector and the stopping criterion is $\|\nabla g(\nu)\|_2 \leq 10^{-5}$.

As evident from Tables A.1 and A.2, three different formulations give very close optimal solutions. For $m = 0.1n$, the primal formulation (P1) can be solved more efficiently than formulation (P2) by standard SDP solver. The BFGS (as well as gradient and generalized Newton methods) has difficulty in solving the dual formulation. For $m = 0.5n$, however, the unconstrained formulation can be solved very efficiently and the BFGS method significantly outperforms standard SDP solvers. Also, note that even for the same SDP solver, formulation (P2) is much easier to handle than formulation (P1). Another aspect of the dual formulation is the construction of the basis. From results listed in Table A.3, the time required to construct the basis is actually comparable to the optimization time. However, the basis can be computed off-line and stored for future computations.

In the following tables, the time is given in seconds. In Table A.1 and A.2, the optimal value $\|X^\star - \Sigma\|_F^2$ and the ratio $\|X^\star - \Sigma\|_F^2/\|\Sigma\|_F^2$ are reported. The time for (P1) and (P2) is the time required to run the SDP solver SeDuMi. We note that the interface Yalmip requires more time than solver SeDuMi. The stopping criterion for BFGS method is $\|\nabla g(\nu)\|_2 \leq 10^{-5}$. The number of dual variables is given by $r = 0.5(n-m)(n-m+1)$.

		Time(s)	It. No.	$\|X^\star - \Sigma\|_F^2$	$\frac{\|X^\star-\Sigma\|_F^2}{\|\Sigma\|_F^2}$
$r = 378$	BFGS	25.1	238	3.6280e-4	13.5%
$n = 30$	P1	7.7	21	3.6288e-4	13.5%
	P2	13.0	21	3.6288e-4	13.5%
$r = 666$	BFGS	108.3	298	4.4958e-4	13.5%
$n = 40$	P1	34.6	23	4.4965e-4	13.5%
	P2	64.1	22	4.4965e-4	13.5%
$r = 1035$	BFGS	1182.8	1104	3.2386e-3	12.7%
$n = 50$	P1	127.1	28	3.2388e-3	12.7%
	P2	281.6	28	3.2388e-3	12.7%

Table A.1: Performance comparison for $m = 0.1n$.

		Time(s)	It. No.	$\|X^\star - \Sigma\|_F^2$	$\frac{\|X^\star-\Sigma\|_F^2}{\|\Sigma\|_F^2}$
$r = 120$	BFGS	0.1	5	6.3172e-3	5.3%
$n = 30$	P1	24.7	24	6.3172e-3	5.3%
	P2	4.5	24	6.3172e-3	5.3%
$r = 210$	BFGS	0.2	5	7.2892e-3	5.9%
$n = 40$	P1	117.3	24	7.2893e-3	5.9%
	P2	17.9	24	7.2893e-3	5.9%
$r = 325$	BFGS	0.5	5	6.7127e-2	7.4%
$n = 50$	P1	413.5	23	6.7127e-2	7.4%
	P2	62.4	24	6.7127e-2	7.4%

Table A.2: Performance comparison for $m = 0.5n$.

	$n = 30$	$n = 40$	$n = 50$
	$r = 378$	$r = 666$	$r = 1035$
Basis time(s)	9.3	34.6	219.2
BFGS time(s)	25.1	108.3	1182.8

Table A.3: BFGS and basis time for $m = 0.1n$.

A.4 Summary

The state sample covariances almost always fail to satisfy linear constraint imposed by the underlying dynamics. The consistency with such dynamics is crucial in addressing the problem of characterizing the input power spectra. We formulate the structured covariance least-squares problem and convert the linear matrix constraint into an equivalent set of trace constraints. The corresponding dual problem can be solved efficiently by unconstrained maximization methods when the number of inputs is close to the number of states.