Graphon Control Theory for Linear Systems on Complex Networks and Related Topics

by

Shuang Gao

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

November 2018

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

© November 2018 Shuang Gao
Abstract

To achieve general control objectives for extremely complex and very large scale networks using standard methods is a challenging task. This work provides a theoretical framework to analyze and control linear systems on networks with arbitrary or growing sizes and complex connections by the use of graphon theory and the theory of infinite dimensional systems.

The characterization of the underlying networks is by graphons, which allows the comparison among networks with differences in both sizes and connections. Since the spectral properties of graphons are important to the control design, the spectral decomposition of graphon operators is investigated, a systematic procedure to identify eigenfunctions of graphon operators is developed and spectral approximations of graphon operators are studied.

Then graphon differential equations are formulated in an appropriate infinite dimensional space. Based on this, the graphon dynamical system models are formulated in order to represent arbitrary-size networks of linear dynamical systems, and to define the convergence of sequences of network systems to limit systems. The exact controllability and the approximate controllability of graphon dynamical systems are then investigated.

The minimum energy state-to-state control problem and the linear quadratic regulator (LQR) problem are formulated for such graphon dynamical systems. The control problem for the graphon limit system is solved in each case using the infinite dimensional system theory and the respective control laws are then approximated to obtain control laws for the finite network system; furthermore, the convergence properties of the approximation schemes are established.

Moreover, to achieve decentralized control laws for LQR problems for systems on complex networks, team optimal solutions are developed, where each agent computes its control actions based on its own state observation and certain global information related to spectral properties of the network and initial conditions. The team solutions ensure that when all agents optimally solve their local LQR problems they collectively solve the global optimal control problem.

For multi-agent systems on a network with decentralized or distributed control strategies, agents may leave the network intentionally or due to failures and this would influence the performance of the control. In order to quantify this influence, the consensus-induced centrality measure for networks of systems is proposed, for which lower bounds and upper bounds are established.
Résumé

Atteindre des objectifs de contrôle pour des réseaux à grande complexité ou à grande échelle avec des méthodes standard est une tâche difficile. Basé sur la théorie des graphons et la théorie des systèmes à dimension infinie, ce travail de thèse définit un cadre théorique permettant d’analyser et de contrôler des systèmes linéaires sur des réseaux de tailles arbitraires ou croissantes.

La caractérisation des réseaux sous-jacents par des graphons permet de comparer des réseaux de tailles et de connections différentes. Puisque les propriétés spectrales des graphons sont importantes pour la conception du contrôle, la décomposition spectrale des graphons est étudiée, une procédure systématique pour identifier les fonctions propres des opérateurs graphiques est mise au point et les approximations spectrales des opérateurs de graphons sont analysées.

Ensuite, les équations différentielles des graphons sont formulées dans un espace de dimension infinie dûment choisi. Grâce à cela, les modèles de systèmes dynamiques de graphons sont formulés pour permettre de représenter des réseaux de systèmes linéaires dynamiques de tailles arbitraires. Ils permettent aussi d’obtenir la limite vers laquelle des séquences de systèmes de réseaux peuvent converger. Suite à cela, une étude est conduite sur les contrôlabilités exactes et approximatives des systèmes dynamiques de graphons.

Le contrôle entre deux états à énergie minimale et le régulateur quadratique linéaire (LQR) sont formulés pour de tels systèmes dynamiques à graphons. Le contrôle en régime limite de graphons est résolu dans chaque cas grâce à la théorie des systèmes à dimensions infinies. Les lois de contrôle respectives sont alors approximées pour obtenir les lois de contrôle pour les systèmes de réseaux à taille finie. En outre, les propriétés de convergence des méthodes d’approximation sont établies.

De plus, afin d’obtenir les lois de contrôle décentralisées pour les problèmes LQR de systèmes à réseaux complexes, des solutions optimales par équipe sont développées en utilisant des méthodes à découplage où chaque agent calcule ses actions de contrôle sur la base de ses propres observations d’état, de certaines informations globales relatives aux spectres et des conditions initiales du réseau. Les solutions par équipe garantissent que lorsque tous les agents résolvent de manière optimale leurs problèmes LQR locaux, ils résolvent collectivement le problème de contrôle optimal global.

Pour les systèmes à plusieurs agents sur des réseaux avec des stratégies de contrôle décentralisées ou distribuées, les agents peuvent rester dans le réseau ou le quitter, soit intentionnellement soit en raison de défaillances. Ceci influence la performance du contrôle. Afin de quantifier cette in-
cidence, la mesure de centralité induite par consensus pour les réseaux de systèmes est proposée, pour laquelle des bornes inférieures et supérieures sont établies.
Claims of Originality and Published Work

Claims of Originality

The main contributions presented in the thesis are as follows:

Chapter 2

- Formulation of graphon differential equations and the graphon unitary operator algebra
- Establishment of a systematic procedure to identify eigenfunctions of graphons.
- Development of spectral approximations of graphons and investigation of spectral decompositions of step function graphons
- Presentation of monotonically increasing graphon sequences
- Presentation of cosinusoidal graphons
  Related publications: [J1, P3, C1, C2, C3]

Chapter 3

- Formulation of the graphon dynamical system model and establishment of its relation with finite dimensional networks of linear systems
- Development of the convergence definition for sequences of graphon dynamical systems
- Proof of existence and uniqueness of solutions for graphon dynamical system equations
- Proof of the sufficient condition for exact controllability
  Related publications: [J1, C1, C2, C3]

Chapter 4

- Development of the graphon-state-to-state control strategy
- Proof of the approximation theorems for graphon state-to-state control
- Explicit solutions to the minimum energy graphon state-to-state control problems
  Related publications: [J1, C1]

Chapter 5

- Formulation of graphon LQR problem and establishment of graphon-network LQR strategy
• Establishment of the approximation of Riccati equation solutions for graphon dynamical systems
• Proofs of the convergence properties of the approximations of Riccati equation solutions
Related publications: [J1, C2, C3]

Chapter 6

• Formulation of graphon LQR problems with non-compact system operator
• Proof of existence and uniqueness of solutions
• Establishment of the team optimal control law via decoupling methods
Related publications: [P1]

Chapter 7

• Formulation of LQR problems on finite dimensional networks with multidimensional nodal states
• Proof of existence and uniqueness of solutions
• Establishment of the team optimal control law via decoupling methods
Related publications: [P2]

Chapter 8

• Formulation of the consensus-induced centrality measure for dynamical systems on positively weighted networks
• Establishment of lower bounds and upper bounds for the change of network eigenvalues with respect to the removal of node sets
• Establishment of lower bounds and upper bounds for the consensus-induced centrality measure
Related publications: [C4]

Publications

Journal Papers

2018.

**Conference Papers**


**Papers in Preparation**


**Presentations at Meetings with Abstract Volumes**


**Contribution of Co-authors**

Professor Peter E. Caines contributed 35% of the work in the papers cited above and their corresponding sections in this thesis.
Acknowledgement

I want to express my deep and sincere gratitude to my supervisor Professor Peter E. Caines for his guidance, continuous encouragement, solid support and his kindness. He provided lots of original inspirations and technical guidance for this thesis. He is a superb teacher, scholar and advisor with contagious enthusiasm, keen scientific insights and restless creativity. I feel privileged to have had the chance to work with and learn from him. I would also like to thank him for directing and encouraging me to work on the completely new “graphon control”.

I would like to take this opportunity to thank Professor Aditya Mahajan for many interesting discussions and inspirations. I thank him for sharing with me lots of important advice, research insights, and constructive suggestions for the work on team problems on networks in Chapter 7.

I would like to thank Professor Roland Malhamé and Professor Michael Rabbat for being my committee members and for their valuable advice and help. I am very grateful for Professor Odile Liboiron-Ladouceur for her support and her trust in me while I was working as a teaching assistant with her, without which I would not have received the McGill Teaching Assistant Award.

I would like to thank Chelsea Rogers, Nicholas Wilson, and Jan Binder at Centre for Intelligent Machines for their help and support all the time. I would also like to thank Jeniffer Zhao and Tara Mawhinney for showing me all the resources and wonderful services in McGill library. I am very grateful for both McGill library and Peter’s library where I can always find the mathematic books I wanted.

Many thanks go to my friends and colleagues in Montreal, Fei Zhang, Can Cui, Bozhong Lu, Jun Chen, Yuhan Ma, Di Wu, Dena Firoozi, Jhelum Chakravorty, Prokopis Prokopiou, Jayakumar Subramanian, Mehnaz Mannan, Hamed Layeghi, Ali Pakniyat, Jalal Arabneydi and Mohammad Afshari. Many thanks also to all my friends in McGill Chinese Graduate Student Association. Special thanks go to Thibaud Tonnellier for helping me translating the abstract into French.

Last but not least, I want to thank my parents for their endless love, patience and support. I want to thank all my grandparents for their love and their faith in me. I want to thank my girlfriend Mi Lin for shining her love into my life, and making my days wonderful and colourful.

Shuang Gao
McGill University,
Montréal, Québec, Canada
August 15, 2018
Contents

1 Introduction 1

I Graphon Control 6

2 Graphons 7
  2.1 Graphons as Graph Limits 7
     2.1.1 Graphs, Adjacency Matrices and Pixel Pictures 7
     2.1.2 Graphon Space 8
     2.1.3 Compactness of the Graphon Space 10
     2.1.4 Different Metrics 10
  2.2 Graphon Operators 11
     2.2.1 Graphon Operators 11
     2.2.2 Spectral Representations 12
     2.2.3 Convergence of Eigenvalues 15
     2.2.4 Identification of Eigenfunctions 15
  2.3 Approximation of Graphons 16
     2.3.1 Approximation by Spectral Sum 16
     2.3.2 Spectral Sum with Eigenfunctions Approximated by Fourier Functions 16
  2.4 Step Function Graphons 18
     2.4.1 Spectral Decomposition 19
     2.4.2 Kroneker Product 20
  2.5 The Graphon Unitary Operator Algebra 22
  2.6 Graphon Differential Equations 23
  2.7 Monotonically Increasing Graphon Sequences 24
## Contents

2.8 Cosinusoidal Graphons ........................................... 26
2.8.1 Cosinusoidal Graphons ................................ 26
2.8.2 Generalized Cosinusoidal Graphon ......................... 29
2.9 Appendices ......................................................... 30
2.9.1 Cauchy Graphon Sequences ................................ 30
2.9.2 Spectral Theorems .............................................. 31

3 Graphon Dynamical System Representations of Network Systems 33
3.1 Network Systems and Their Limit Systems ..................... 33
3.1.1 Network System Model ...................................... 33
3.1.2 Network Systems Described by Step Functions ............. 34
3.1.3 Limits of Sequences of Network Systems ................... 36
3.1.4 Limit Graphon Systems ...................................... 36
3.2 Graphon System Properties ...................................... 37
3.2.1 Uniqueness of the Solution .................................. 37
3.2.2 Controllability ................................................. 38

4 Graphon State-to-state Control of Large-scale Networks 41
4.1 Graphon State-to-state Control Strategy (GSCS) ............... 41
4.2 Approximation Theorems ........................................ 42
4.2.1 Approximation of $L^2[0, 1]$ Input Functions via Piece-wise Constant Functions ......................... 42
4.2.2 Limit Control for Network Systems ......................... 42
4.3 Minimum Energy State-to-state Control for Graphon Systems ........................................... 47
4.3.1 Minimum Energy Control of Infinite Dimensional Systems .......... 48
4.3.2 Minimum Energy Control Law ................................ 48
4.4 Simulation Example ............................................... 51
4.4.1 Network Systems with Sampled Weightings .................. 51
4.4.2 Minimum Energy Graphon State-to-state Control ................ 52
4.5 Appendix ......................................................... 54

5 Graphon Linear Quadratic Regulation (LQR) of Network Systems 58
5.1 Graphon-Network LQR (GLQR) Strategy ......................... 58
5.2 LQR Problems for Graphon Dynamical Systems ................ 59
5.3 Approximation Theorems ................................................................. 61
  5.3.1 Approximation of the Riccati Equation Solution and Its Convergence to the Optimal Riccati Equation Solution ........................................ 61
  5.3.2 Continuous Dependence of Riccati Equation Solution with Respect to the Data ................................................................. 63
  5.3.3 Convergence of States and Convergence of Costs ......................... 64
5.4 Simulation Example ......................................................................... 66

II Graphon Team Optimal Control 69

6 Team Optimal Graphon-LQR Control ................................................. 70
  6.1 Introduction .................................................................................. 70
  6.2 System Models with Non-compact Graphon System Operators .......... 71
    6.2.1 System Model ........................................................................ 71
    6.2.2 Cost Function ........................................................................ 72
    6.2.3 Existence and Uniqueness of Solutions to LQR Problems .......... 72
  6.3 Team Optimal Solutions via Decoupling ........................................ 73
    6.3.1 Decoupling Method ................................................................. 74
    6.3.2 Team Optimal Solutions .......................................................... 78
    6.3.3 Discussion ............................................................................. 79

7 Team Optimal Linear Quadratic Regulation (LQR) on Networks .......... 80
  7.1 Introduction .................................................................................. 80
  7.2 System Model and LQR Problems on Networks .............................. 82
    7.2.1 System Model ........................................................................ 82
    7.2.2 LQR Problems on Networks ................................................... 84
    7.2.3 Basic Assumptions ................................................................. 84
  7.3 Team Optimal Solutions via Decoupling ........................................ 85
    7.3.1 Decoupling in Dynamics ........................................................ 85
    7.3.2 Decoupling in the Cost Function ............................................. 86
    7.3.3 Team Optimal Solutions .......................................................... 93
  7.4 Numerical Examples ..................................................................... 95
    7.4.1 Complete Bipartite Networks .................................................. 95
### Contents

7.4.2 Cosinusoidal Networks ............................................. 96
7.5 Discussion ............................................................... 97

### III Centrality for Networks of Dynamical Systems 99

#### 8 Consensus-induced Centrality for Networks of Dynamical Systems 100

8.1 Introduction ............................................................ 100
8.2 Consensus and Synchronization on Networks .......................... 102
8.2.1 Consensus on Networks ........................................... 102
8.2.2 Synchronization on Networks ..................................... 103
8.2.3 Consensibility and Relative Consensibility of Network Systems ...... 104
8.2.4 Properties of Algebraic Connectivity ............................... 104
8.3 Consensus-induced Centrality Measure .................................. 105
8.3.1 Absolute Consensus-induced Centrality Measure ................. 105
8.3.2 Relative Consensus-induced Centrality Measure .................. 106
8.4 Basic Results on Consensus-induced Centrality Measure .............. 106
8.4.1 Upper Bound of ACIC ............................................. 106
8.4.2 Lower Bound of ACIC ............................................. 109
8.5 Generalization to Centrality Measure for Node Sets ................... 111
8.5.1 Consensus-induced Centrality Measures for Node Sets .......... 111
8.5.2 Spectrum Change with the Removal of Node Sets ............... 111
8.5.3 Upper and Lower Bounds of ACIC ................................ 114
8.6 Examples of Consensus-induced Centrality Measure for Networks .... 114
8.7 Discussions ............................................................. 119
8.8 Appendix ................................................................. 119
8.8.1 The Cauchy Interlace Theorem ................................... 119
8.8.2 The Courant-Fischer Theorem .................................... 120

#### 9 Summary and Future Research Directions 121

9.1 Summary ............................................................... 121
9.2 Future Research Directions .......................................... 123
9.2.1 Graphon Linear Quadratic Gaussian Problems .................... 123
9.2.2 Graphon Mean Field Games ..................................... 124
<table>
<thead>
<tr>
<th>9.2.3</th>
<th>Non-linear Local Dynamics</th>
<th>124</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.2.4</td>
<td>Control of Time Varying Graphon Dynamical Systems</td>
<td>125</td>
</tr>
<tr>
<td>9.2.5</td>
<td>Graphon as Non-parametric Models for Control Design</td>
<td>125</td>
</tr>
<tr>
<td>9.2.6</td>
<td>Centrality Measure of Dynamical Systems on Graphons (or Networks)</td>
<td>125</td>
</tr>
<tr>
<td>9.2.7</td>
<td>Graphon Control Applications</td>
<td>126</td>
</tr>
<tr>
<td>9.2.8</td>
<td>Other Future Directions</td>
<td>126</td>
</tr>
</tbody>
</table>

**References**  
128
## List of Figures

2.1 Dürer Graph, Adjacency Matrix, Pixel Diagram .................................. 7
2.2 Graph Sequence Converging to Its Limit ........................................... 8
2.3 Monotonically Increasing Graphon Sequence Example ......................... 25

3.1 A Weighted Graph from a Sequence Converging to the Limit Graphon $W(x, y) = 1 - \max(x, y), 0 \leq x, y \leq 1$ with $x, y$ measured from the top left .................. 36

4.1 Minimum Energy Graphon State-to-state Control ................................. 54

5.1 Parameters for Riccati Equation Solution ......................................... 67
5.2 Simulation on a Network of 320 Nodes .............................................. 68

7.1 A complete bipartite network example ............................................. 96
7.2 A numerical example on a complete bipartite network (with 20 agents), where each agent is optimally solving its local LQR problems and they together achieve the global optimal performance ................................. 96
7.3 An example of a cosinusoidal graph ................................................... 97
7.4 A numerical example on a network generated from a sinusoidal graphon (with 20 agents), where each agent is optimally solving its local LQR problems and they together achieve the global optimal performance ................................. 98

8.1 A Simple Network with Negative ACIC ............................................. 116
8.2 The Karate Club Network [1] .......................................................... 117
8.3 Small World Network Example ....................................................... 118
8.4 The Power Network ................................................................. 118
8.5 The Southern Women Club Network ................................................. 119
9.1 Control Design Procedure for Network Systems via Graphon Limits . . . . . . . 122
List of Tables

7.1 Summary of Notation ........................................... 83

Chapter 1

Introduction

Complex networks such as the Internet of Things (IoT), electric grids, biological neural networks, food webs, epidemic networks, stock market and social networks, are ubiquitous. Underlying most of these complex networks there are intrinsic nodal states, depending on the problem of interest, which evolve due to the interactions among the agents on the networks. The states can be for instance the status of devices (or vehicles), energy consumptions or productions, neuronal states (resting or excitation), the size of population of species, status of health, stock prices, opinions on certain topics, etc. The states of the agents can be freely involving, locally or globally controlled. Each agent on the networks has the ability to receive, process or output information encoded in the states.

Researchers have been studying networks of interacting dynamical systems to learn which collective behaviours may emerge from system interactions over complex networks [3, 4, 5, 6]. Furthermore, in addition to the structures of networks [7, 8], system theoretic notions such as controllability, observability, consensus dynamics and synchronization have been widely applied [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23]. In particular, network controllability has been very popular among the network science community during the past decade years. Studies in this direction include, but are not limited to, the number of inputs (or driver nodes) needed to achieve controllability at a structural level [10, 24, 25], the energy required for state-to-state control problems for networks [15, 26], the controllability via pinning a fraction of nodes on the networks [20, 27], the influence of intrinsic dynamics on controllability [28], and controllability from the spectrum perspective [14]. Moreover, network controllability as an effective tool to study complex networks has been applied to neuronal networks to predict neuronal functions...
These research studies focus on (a) analysis problems such as controllability [10], observability [14] and control energy metric [15], etc., and (b) synthesis problems with simple objectives or simple control, such as consensus [16], synchronization [19], ensemble control [21], etc. In spite of these contributions, the understanding of complex networks of dynamical systems and furthermore the creation of a control theory for such systems are still extremely challenging tasks. One of the reasons that controlling these networks is so difficult or even intractable lies in their large size and the complexity of connections.

In this work we propose and develop the graphon control theory to approximately control extremely complex and arbitrary size network systems by the use of graphon theory and the theory of infinite dimensional systems. The preliminary results of this work have been reported in [33, 34, 35].

Graphs, as the mathematical model of networks, have a rich theory developed by mathematicians. To study the possible limits of graph sequences and models for generating random graphs, the theory of graphon was introduced and developed in recent years by Lovász, Szegedy, Borgs, Chayes, Sós, and Vesztergombi among others (see [36, 37, 38, 39, 40]). This work draws on graph theory, measure theory, probability, and functional analysis, and recently has been applied in different areas such as games [41, 42], signal processing [43] and crowd-sourcing [44]. Evidently, graphon theory provides a potentially powerful theoretical tool for the study large networks and graph limits.

Infinite dimensional control theory has been developed by control theorists [45, 46, 47, 48] to solve control problems for systems governed by partial differential equations and functional differential equations. Its applications cover distributed parameter systems such as chemical process control, control of elastic structures, and the challenging issues such as stabilizing plasma instabilities. This provides the fundamental control theoretic framework for solving control problems on infinite dimensional networks of systems.

Team optimal control problems [49] with mean field coupling are recently proposed and studied in [50], where a large number of agents collectively achieve a common team objective by locally calculated control actions with global information on the mean field coupling. In this thesis, the team optimal control with graphon couplings is explored in the implementation of graphon control in a decentralized manner. More specifically, the LQR problem with graphon (or network) coupling is formulated and the team optimal solution is established. The information structure is not required to be partially nested [49] and permits more variations than mean-field
sharing [51].

In general, for multi-agent systems on networks with decentralized or distributed control strategies, agents may leave or stay in the network intentionally or due to failures and this would influence the performance of the decentralized or distributed control. In order to quantify this influence, centrality (or importance) measures for networks of dynamical systems need to be investigated. Most graph theoretic centrality measures on networks tend to focus on the structural importance of nodes in the underlying graph [52]. However, for controlled networks of dynamical systems, the dynamics and the control strategies, as well as the network structures, determine the centrality measure. In this thesis the consensus-induced centrality measure for networks of systems following consensus-type dynamics is proposed, and lower bounds and upper bounds are established. The consensus-induced centrality notion is applicable to network problems such as the synchronization problem, the pinning control problem, the consensus problem, and has possible applications to network systems such as social networks, power grid stability and wireless sensor networks, etc.

Organization of the Thesis

As explained above, the goal of this thesis is to provide a theoretical framework to analyze and control linear systems on networks with arbitrary or growing sizes and complex connections. The characterization of the underlying networks is by graphons, which allows the comparison among networks with differences in both sizes and connections.

Part I of the thesis introduces and develops graphon control theory and is organized as follows:

Chapter 2 introduces the theory of graphons, graphon operators, the graphon unitary operator algebra, the graphon differential equations, and the spectral representations and approximations of graphons, which form the basics of characterizing graphon dynamical systems.

Chapter 3 formulates the graphon dynamical system model to represent linear systems on arbitrary-size networks. The relation between a finite dimensional network and its counterpart represented by an infinite dimensional graphon system is studied. Furthermore, properties of graphon dynamical systems such as the existence of a unique solution and exact controllability are studied.

Chapter 4 studies the state-to-state control problem for graphon dynamical systems and introduces the graphon state-to-state control (GSSC) strategy. The GSSC strategy obtains a control
law based on the graphon approximation of networks, solves the corresponding control problem for the graphon system, then approximates the generated control law and applies it on the finite network systems lying in a sequence that converges to the limit system. The minimum energy state-to-state control problems for graphon dynamical systems are solved explicitly. The approximation theorems are developed for the GSSC strategy.

Chapter 5 studies the graphon linear quadratic regulation problem over a finite time horizon and introduces the graphon-network LQR (GLQR) strategy. The existence and uniqueness of optimal solutions are established. The optimal solution is given by a linear feedback form with the feedback gain generated from the corresponding infinite dimensional Riccati equation. The approximations of Riccati equation solutions are developed and the convergence of the approximations in terms of feedback performance is proved in the approximation theorems for the GLQR strategy.

Part II of the thesis discusses the team optimal control for finite or infinite dimensional networks of linear systems. It is organized as follows:

Chapter 6 presents the graphon system model with non-compact system operator as an extension of the model in Chapter 3. The uniqueness and existence of solutions are presented. Finally, the team optimal control solutions are established. This part considers linear systems on networks of arbitrary sizes, where the networks represent couplings in dynamics and cost functions. Agents on the networks are assumed to share the same nodal dynamics, except from the localized coupling determined by the networks. The global LQR problem is decoupled into LQR problems independently solved by each agent. All the agents collectively solve the global LQR problem by solving its local LQR problems with some global information that contains the network structure and initial states.

Chapter 7 presents the team optimal solutions to the LQR problem for linear systems on finite dimensional networks, where each agent on the network has a multi-dimensional state and a multi-dimensional input. Due to the finite dimensionality of the networks, the coupling in dynamics is not limit to coupling with averaging effects, and can be extended to network coupling without scaling. The team optimal solutions are established.

For both team optimal solutions in Chapter 6 and Chapter 7, under the assumption that the spectral representation of the network has a significantly smaller dimension compared to the network size, the computational complexity can be significantly reduced.

Part III of the thesis contains only Chapter 8 and studies the centrality measure for networks
of dynamical systems with decentralized or distributed control strategies. The consensus-induced centrality measure is proposed to quantify the importance of agents on networks, and lower bounds and upper bounds for this centrality are obtained.

Future research directions are presented in Chapter 9.
Part I

Graphon Control
Chapter 2

Graphons

2.1 Graphons as Graph Limits

2.1.1 Graphs, Adjacency Matrices and Pixel Pictures

The underlying structure of a network can be described by a graph $G = (V, E)$ specified by a node set $V$ and an edge set $E$ which represents the connections between nodes. An equivalent representation of a graph $G = (V, E)$ by a matrix called an adjacency matrix is defined to be the square $|V| \times |V|$ matrix $A$ such that an element $A_{ij}$ is one when there is an edge from node $i$ to node $j$, and zero otherwise. If the graph is a weighted graph where edges are associated with weights, then the adjacency matrix has corresponding weighted elements.

Another representation of the adjacency matrix is given by a pixel diagram where the 0s are replaced by white squares and the 1s by black squares. The whole pixel diagram is presented in a unit square, so the square elements have sides of length $\frac{1}{n}$, where $n$ is the number of nodes.

Fig. 2.1 Dürer Graph, Adjacency Matrix, Pixel Diagram
2.1.2 Graphon Space

Graphon theory was introduced and developed in recent years by Lovász, Szegedy, Borgs, Chayes, Sós, and Vesztergombi among others in [36, 37, 38, 39, 40]. This work draws on graph theory, measure theory, probability, and functional analysis. In the literature (see e.g. [40]), a meaningful convergence with respect to the cut metric is defined for sequences of dense and finite graphs. Graphons are then the limit objects of converging graph sequences. This concept is illustrated by a sequence of half graphs ([40]) represented by a sequence of pixel diagrams on the unit square converging to its limit in Fig. 2.2.

The set of finite graphs endowed with the cut metric gives rise to a metric space, and the completion of this space is the space of graphons. Graphons are represented by bounded symmetric Lebesgue measurable functions \( W : [0, 1]^2 \to [0, 1] \), which can be interpreted as weighted graphs on the node set \([0, 1]\). We note that in some papers, for instance [53], the word “graphon” refers to symmetric, integrable functions from \([0, 1]\) to \(\mathbb{R}\). In this work, unless stated otherwise, the term “graphon” is used to refer to functions \( W : [0, 1]^2 \to [-1, 1] \) and \( \tilde{G}_1^{sp} \) denotes the space of graphons. Let \( \hat{G}_0^{sp} \) represent the space of all graphons satisfying \( W_0 : [0, 1]^2 \to [0, 1] \) and let \( \hat{G}^{sp} \) denote the space of all symmetric measurable functions \( W : [0, 1]^2 \to \mathcal{R} \).

The cut norm of a graphon \( W \in \hat{G}_1^{sp} \) is then defined as

\[
\|W\|_\Box = \sup_{M,T \subset [0,1]} \left| \int_{M \times T} W(x,y) dx dy \right| \tag{2.1}
\]

with the supremum taking over all measurable subsets \( M \) and \( T \) of \([0,1]\). Evidently the following inequalities hold between norms on a graphon \( W \):

\[
\|W\|_\Box \leq \|W\|_1 \leq \|W\|_2 \leq \|W\|_\infty \leq 1, \tag{2.2}
\]

where the second to the fourth norms are given by the corresponding \( L^p \) norms on \( \hat{G}_1^{sp} \). Denote the set of measure preserving bijections from \([0,1]\) to \([0,1]\) by \( S_{[0,1]} \). The cut metric between
two graphons $V$ and $W$ is then given by

$$d_\Box(W, V) = \inf_{\phi \in S_{[0,1]}} \| W^{\phi} - V \|_\Box,$$

(2.3)

where $W^{\phi}(x, y) = W(\phi(x), \phi(y))$. We see that the cut metric $d_\Box(\cdot, \cdot)$ is given by measuring the maximum discrepancy between the integrals of two graphons over measurable subsets of $[0,1]$, then minimizing the maximum discrepancy over all possible measure preserving bijections.

Since the cut metric of two different graphons can be 0, strictly speaking it is not a metric. See [38, 54] for various characterizations of when the cut distance is 0. By identifying functions $V$ and $W$ for which $d_\Box(V, W) = 0$, we can construct the metric space $G_{1}^{sp}$ which denotes the image of $\tilde{G}_{1}^{sp}$ under this identification. Similarly we construct $G_{0}^{sp}$ from $\tilde{G}_{0}^{sp}$ and $G_{1}^{sp}$ from $\tilde{G}_{1}^{sp}$.

We define the $L^2$ metric for any graphons $W$ and $V$ as

$$d_{L^2}(W, V) = \| W - V \|_2 = \left( \int_{[0,1]^2} |W(x, y) - V(x, y)|^2 \, dx \, dy \right)^{\frac{1}{2}}$$

(2.4)

and the $\delta_2$ metric as

$$\delta_2(W, V) = \inf_{\phi \in S_{[0,1]}} d_{L^2}(W^{\phi}, V) = \inf_{\phi \in S_{[0,1]}} \| W^{\phi} - V \|_2;$$

(2.5)

Similarly, we define the $L^1$ metric as

$$d_{L^1}(W, V) = \| W - V \|_1,$$

(2.6)

and the $\delta_1$ metric as

$$\delta_1(W, V) = \inf_{\phi \in S_{[0,1]}} d_{L^1}(W^{\phi}, V) = \inf_{\phi \in S_{[0,1]}} \| W^{\phi} - V \|_1.$$

(2.7)

For any two graphons $W$ and $V$ the following inequalities are immediate:

$$\delta_0(W, V) \leq \delta_1(W, V) \leq \delta_2(W, V) \leq d_{L^2}(W, V).$$

(2.8)

The $\delta_2$ (or $\delta_1$) metric and $\delta_0$ metric share the same equivalence classes under the measure preserving transformations [40, Corollary 8.14]. Clearly, the $\delta_2$ (or $\delta_1$) metric is also well defined on $G_{1}^{sp}$. 


2.1.3 Compactness of the Graphon Space

**Theorem 2.1** ([40]). The space \((G_0^{\text{sp}}, \delta_0)\) is compact.

This remains valid if \(G_0^{\text{sp}}\) is replaced by any uniformly bounded subset of \(G^{\text{sp}}\) closed in the cut metric \([40]\).

**Theorem 2.2** ([40]). The space \((G_1^{\text{sp}}, \delta_1)\) is compact.

Sets in \(G_1^{\text{sp}}\) (or \(G_0^{\text{sp}}\)) compact with respect to the \(\delta_1\) metric are compact with respect to the cut metric. It follows immediately from (2.8) and Theorem 2.2 (or Theorem 2.1), if a graphon sequence is Cauchy in the \(\delta_2\) metric then it is also a Cauchy sequence in the cut metric and under both metrics, the limits are identical in \(G_1^{\text{sp}}\) (or \(G_0^{\text{sp}}\)).

Define the \(L^p\) closed ball in \(G^{\text{sp}}\) with radius \(C > 0\) as \(B_{L^p}(C) := \{ W : \| W \|_p \leq C, W \in G^{\text{sp}} \} \).

**Theorem 2.3** ([53]). The space \((B_{L^p}(C), \delta_\Box)\) with \(1 < p \leq \infty\) is compact.

By compactness, infinite sequences of graphons will necessarily possess one or more sub-sequential limits under the cut metric.

2.1.4 Different Metrics

**Proposition 2.1.** Let \(\{ W_i \}_{i=1}^\infty\) be a \(\delta_\Box\)-convergent sequence in \(G_1^{\text{sp}}\) with limit \(W\). If \(\| W \|_2 = \lim_{i \to \infty} \| W_i \|_2\), then \(\{ W_i \}_{i=1}^\infty\) converges to \(W\) also in the \(\delta_2\) metric.

**Proof.** Based on [55], the \(\{ W_i \}_{i=1}^\infty\) converges to \(W\) in the \(\delta_1\) metric. Since for any \(W, V \in G_1^{\text{sp}}\),

\[
\frac{1}{2} \delta_2(W, V)^2 \leq \delta_1(W, V) \leq \delta_2(W, V),
\]

one immediately obtains that \(\{ W_i \}_{i=1}^\infty\) converges to \(W\) also in the \(\delta_2\) metric. \(\square\)

**Proposition 2.2.** The space \(\tilde{G}_1^{\text{sp}}\) under the \(d_{L^2}\) metric is complete.

Henceforth we only consider the \(L^2\) topology on the space of graphons. Consequently the convergence of a sequence of graphons will be interpreted as convergence in the complete space of graphons in the \(L^2\) metric. By the ordering of metrics given in (2.8) this further implies convergence in the compact space (of equivalence classes) of graphons under in the weaker cut metric topology.

Examples of sets of graphons which have common limits in the \(L^2\) metric and cut metric topologies are given by the so-called monotone families of graphons. These correspond to graphs to which nodes or edges are recursively added at each of an infinite set of discrete time instants.
2 Graphons

2.2 Graphon Operators

2.2.1 Graphon Operators

A graphon operator \( \mathbb{W} : L^2[0, 1] \to L^2[0, 1] \) is defined by a graphon \( \mathbb{W} \in \tilde{G}_1^{gp} \) as follows:

\[
[\mathbb{W}f](x) = \int_0^1 \mathbb{W}(x,y)f(y)dy.
\] (2.9)

Clearly, the operator \( \mathbb{W} \) is Hermitian (or self-adjoint), since for any \( x, y \) in \( L^2[0, 1] \), \( \langle x, \mathbb{W}y \rangle = \langle \mathbb{W}x, y \rangle \).

**Definition 2.1** ([56, p.103]). A linear map \( T : L^2[0, 1] \to L^2[0, 1] \) is said to be compact if \( T \) maps the open unit ball in \( L^2[0, 1] \) to a set in \( L^2[0, 1] \) that has compact closure.

**Proposition 2.3** ([57, Chapter 2, Proposition 4.7]). If \( \mathbb{W} \in L^2[0, 1]^2 \), then

\[
[\mathbb{W}f](x) = \int_0^1 \mathbb{W}(x,y)f(y)dy
\] (2.10)

is a compact operator.

**Corollary 2.1.** Consider \( \mathbb{W} \in B_{L^2}(C) \) with \( 0 < C < \infty \). Then the corresponding operator \( \mathbb{W} \) is a compact operator.

**Proposition 2.4.** The graphon operator \( \mathbb{W} \) is a linear operator that is bounded, (hence) continuous, and compact.

**Proof.** Note that the range of graphon is \([-1, 1]\) and it is defined on \([0, 1]^2\). Therefore the graphon operator is bounded and hence continuous. The compactness is an immediate result of Proposition 2.3.

For simplicity of notation, henceforth we use the bold face letter (e.g. \( \mathbb{W}, U \)) to represent both a graphon and its corresponding graphon operator.

The operator product is then defined by

\[
[UW](x,y) = \int_0^1 U(x,z)W(z,y)dz,
\] (2.11)
where $U, W \in \tilde{G}_{1}^{sp}$. See [40] for more details. Note that if $U \in \tilde{G}_{1}^{sp}$ and $W \in \tilde{G}_{1}^{sp}$, then $UW \in \tilde{G}_{1}^{sp}$, since for all $x, y \in [0, 1]$

$$[[UW](x, y)] = \left| \int_{0}^{1} U(x, z)W(z, y) dz \right| \leq \int_{0}^{1} |U(x, z)W(z, y)| dz \leq 1. \tag{2.12}$$

Consequently, the power $W^{n}$ of an operator $W \in \tilde{G}_{1}^{sp}$ is defined as

$$W^{n}(x, y) = \int_{[0,1]^{n}} W(x, \alpha_{1}) \cdots W(\alpha_{n-1}, y) d\alpha_{1} \cdots d\alpha_{n-1} \tag{2.13}$$

with $W^{n} \in \tilde{G}_{1}^{sp}$ ($n \geq 1$). $W^{0}$ is formally defined as the identity operator on functions in $L^{2}[0, 1]$, but we note that $W^{0}$ is not a graphon.

For simplicity of notation, $UW$ is used to denote the graphon given by the convolution in (2.11); similarly, $Wv$ denotes the function defined by (2.9).

### 2.2.2 Spectral Representations

Since spectral properties of graphons are important to the control design for graphon dynamical systems, we study and present the spectral representation of graphons.

Denote the operator norm for a linear operator $L$ on $L^{2}[0, 1]$ as

$$\|L\|_{op} = \sup_{f \in L^{2}[0, 1], \|f\|_{2} = 1} \|Lf\|_{2}. \tag{2.14}$$

Define the kernel space (or null space) for a linear operator $L$ on $L^{2}[0, 1]$ as:

$$\ker(L) := \{ x \in L^{2}[0, 1] : Lx = 0 \}. \tag{*}$$

The spectrum $\sigma(L)$ of a linear bounded operator $L$ on $L^{2}[0, 1]$ is the set of all (complex or real) scalars $\lambda$ such that $L - \lambda I$ is not invertible, where $I$ is the identity operator. Thus $\lambda \in \sigma(L)$ if and only if at least one of the following two statements is true:

(i) The range of $L - \lambda I$ is not all of $L^{2}[0, 1]$.

(ii) $L - \lambda I$ is not one-to-one.

If (ii) holds, $\lambda$ is said to be an eigenvalue of $L$; the corresponding eigenspace is $\ker(L - \lambda I)$; each $x \in \ker(L - \lambda I)$ (except $x = 0$) is an eigenvector of $L$; it satisfies the equation $Lx = \lambda x$. See [56].
Proposition 2.5. Consider the graphon operator $W$ corresponding to a graphon $W \in \tilde{G}_1^{sp}$.

1) If $\lambda \in \sigma(W)$ (i.e. $\lambda$ is in the spectrum of $W$) and $\lambda \neq 0$, then $\lambda$ is an eigenvalue of $W$.

2) $W$ has a countable number of distinct eigenvalues and all the eigenvalues are real. Furthermore, if we denote the distinct nonzero eigenvalues of $W$ as $\{\eta_1, \eta_2, \ldots\}$ and the projection of $L^2[0,1]$ onto $\ker(W - \eta_n I)$ as $P_n$, then $P_m P_n = P_n P_m = 0$ if $n \neq m$, and

$$W = \sum_{n=1}^{\infty} \eta_n P_n$$

where the series converges to $W$ in the norm $\| \cdot \|_{op}$.

3) There is a set $\{\varphi_i\}$ consisting of a countable number of orthonormal elements in $L^2[0,1]$ such that the elements $\varphi_i$ are eigenfunctions to the eigenvalues $\lambda_i \in \mathbb{R}$ ordered as follows:

$$\|W\|_{op} = |\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \ldots \geq 0,$$

If the set $\{\varphi_i\}$ is infinite, we have the asymptotic behavior

$$\lim_{i \to \infty} \lambda_i = 0.$$

Furthermore, for any $u \in L^2[0,1]$, $Wu$ has representations as:

$$Wu = \sum_{i=1}^{\infty} \lambda_i \langle \varphi_i, u \rangle \varphi_i \quad \text{and} \quad \langle u, Wu \rangle = \sum_{i=1}^{\infty} \lambda_i |\langle \varphi_i, u \rangle|^2.$$

4) $$\sum_{i=1}^{\infty} \lambda_i^2 = \int_{[0,1]^2} W(x,y)^2 dxdy = \|W\|_2^2 \leq \infty, \quad \text{and} \quad |\lambda_k| \leq \frac{\|W\|_2}{\sqrt{k}}$$

5) Denote $W_m(x,y) = \sum_{k=1}^{m} \lambda_k \varphi_k(x) \varphi_k(y)$. Then

$$\lim_{m \to \infty} \|W_m - W\|_2 := \lim_{m \to \infty} \int_0^1 \int_0^1 (W_m(x,y) - W(x,y))^2 dxdy = 0.$$

6) If $W$ nonnegative (that is, $\int_0^1 \int_0^1 f(x) W(x,y) f(y) dxdy \geq 0$ for any $f(\cdot) \in L^2[0,1]$), and $W$ is continuous from $[0,1]^2$ to $\mathbb{R}$, then $\{W_m\}$ converges to $W$ absolutely and uniformly as $m \to \infty$, and the eigenfunctions corresponding to nonzero eigenvalues are continuous.
Proof. 1) Since $W$ is a compact operator, this is a special case of the result in [56, Theorem 4.25].

2) Since $W$ as an operator is self-join and compact, applying Theorem 2.8 and selecting the Hilbert space to be $L^2[0, 1]$, we obtain the result.

3) Applying Theorem 2.9 and selecting the Hilbert space to be $L^2[0, 1]$, we obtain the result.

4) See also [40, p.124].

$$k\lambda_k^2 \leq \sum_{i=1}^k \lambda_i^2 \leq \sum_{i=1}^\infty \lambda_i^2 = \int_{[0,1]^2} W(x, y)^2 dx dy = \|W\|^2_2.$$  

5) See e.g. [58, Chapter 5, section 2.4].

6) This is obtained by applying the Mercer’s Theorem [59] to kernels in $L^2[0, 1]^2$.

$\square$

Lemma 2.1. For any graphon $W$ or any function $W$ in $L^2[0, 1]^2$,

$$\|W\|_{op} \leq \|W\|_2.$$

Proof.

$$\|W\|_{op} = \sup_{x \neq 0, x \in L^2[0, 1]} \frac{\|Wx\|_2}{\|x\|_2}$$

$$= \sup_{x \neq 0, x \in L^2[0, 1]} \frac{\sqrt{\int_0^1 \left[ \int_0^1 W(\alpha, \beta) x(\beta) d\beta \right]^2 d\alpha}}{\|x\|_2}$$

$$\leq \sup_{x \neq 0, x \in L^2[0, 1]} \frac{\sqrt{\int_0^1 \left[ \int_0^1 W^2(\alpha, \beta) d\beta \int_0^1 x^2(\beta) d\beta \right] d\alpha}}{\|x\|_2}$$

$$= \sup_{x \neq 0, x \in L^2[0, 1]} \frac{\|x\|_2 \sqrt{\int_0^1 \int_0^1 W^2(\alpha, \beta) d\beta d\alpha}}{\|x\|_2}$$

$$= \sqrt{\int_0^1 \int_0^1 W^2(\alpha, \beta) d\beta d\alpha}$$

$\|W\|_2$.

$\square$

Remark: A graphon sequence convergences under $\| \cdot \|_2$ implies it convergences under $\| \cdot \|_{op}$. 
2.2.3 Convergence of Eigenvalues

For a graphon $W$, the eigenvalues form two sequences $\mu_1(W) \geq \mu_2(W) \geq \ldots \geq 0$ and $\mu'_1(W) \leq \mu'_2(W) \leq \ldots \leq 0$ converging to zero, where $\mu_i(W)$ and $\mu'_i(W)$ denote respectively the $i^{th}$ non-negative eigenvalue and the $i^{th}$ non-positive eigenvalue.

**Theorem 2.4** ([39]). Let $\{W_i\}_{i=1}^{\infty}$ be a sequence of uniformly bounded graphons, converging in the cut metric to a graphon $W$. Then for every $i \geq 1$,

$$\mu_i(W_n) \to \mu_i(W) \quad \text{and} \quad \mu'_i(W_n) \to \mu'_i(W) \quad \text{as } n \to \infty.$$

2.2.4 Identification of Eigenfunctions

Assume $\{h_i \in L^2[0,1]\}$ is a set of linearly independent functions spanning the range of a graphon operator $W$. Further, assume that the range space dimension of $W$ is finite and given by $d$, (i.e., the space spanned by the eigenspace of non-zero eigenvalues of $W$ is $d$ dimensional). We specify the following procedure to identify the eigenfunctions of $W$.

**Step 1.** By Gram-Schmidt orthogonalization of $\{h_i \in L^2[0,1]\}$, obtain the orthonormal basis $\{g_i \in L^2[0,1]\}$ of the range of the operator $W$. This is also $d$ dimensional.

**Step 2.** Calculate the inner product for all $i, j \in 1, \ldots, d$,

$$q_{ij} = q_{ji} := \langle g_i, W g_j \rangle,$$

which then gives the symmetric transformation matrix $Q = [q_{ij}] \in \mathbb{R}^{d \times d}$.

**Step 3.** For the finite dimensional matrix $Q$, perform the eigenvalue decomposition, which necessarily yields

$$Q = V \Lambda_d V^T$$

where $V = (v_1, \ldots, v_d)$, $\Lambda_d = \text{diag}(\lambda_1, \ldots, \lambda_d)$, $v_i$ is the normalized eigenvector of $Q$ corresponding to the (necessarily real) eigenvalue $\lambda_i$ of $Q$, $i \in \{1, \ldots, d\}$.

**Step 4.** Then $\lambda_i$ is the eigenvalue of $W$ with eigenfunction given by $f_i = v_i^T \cdot g \in L^2[0,1]$ where $g = (g_1, \ldots, g_d)^T$ and $v_i^T \cdot g := \sum_{k=1}^d v_i(k) g_k$ with $v_i(k)$ denoting the $k^{th}$ element of vector $v_i$. Therefore

$$W(x, y) = \sum_{i=1}^d \lambda_i f_i(x) f_i(y).$$
If the range of the operator $W$ is infinite dimensional space, it is unrealistic to find the infinite dimensional kernel matrix given by $Q$, however, we can approximate the graphon by a finite dimensional kernel space that corresponds to the most significant eigenvalues of $W$.

In particular, if the eigenfunctions are unknown, we can start with the Fourier functions, since Fourier functions form a complete basis for $L^2[0, 1]$.

2.3 Approximation of Graphons

2.3.1 Approximation by Spectral Sum

If the eigenvalues and the corresponding eigenfunctions of a graphon are known, by Proposition 2.5. (4), one can approximate the graphon by a finite spectral sum. Consider the approximation of a graphon $W$ by $W_m(x, y) = \sum_{k=1}^{m} \lambda_k \phi_k(x) \phi_k(y)$. Then the mean square error is bounded as:

$$
\|W - W_{m}\|_2 \leq \|W - W_{\infty}\|_2 + \|W_{\infty} - W_{m}\|_2
$$

$$
= \|W_{\infty} - W_{m}\|_2
$$

$$
= \sqrt{\sum_{k=m+1}^{\infty} \lambda_k^2}
$$

$$
= \sqrt{\|W\|_2^2 - \sum_{k=1}^{m} \lambda_k^2}
$$

(2.15)

2.3.2 Spectral Sum with Eigenfunctions Approximated by Fourier Functions

Denote $p_k(\cdot)$ as a polynomial function and $p_k(e^{2\pi i \cdot})$ is used to approximate the $k^{th}$ eigenfunction $\phi_k(\cdot)$ of $W$ with $W(x, y) = \sum_{k=1}^{\infty} \lambda_k \phi_k(x) \phi_k(y)$. Note that if the polynomial $p_k(\cdot)$ permits terms up to infinite order, $p_k(e^{2\pi i \cdot})$ is simply the Fourier representation of $\phi_k(\cdot)$. Denote the spectral sum with Fourier approximated eigenfunctions as

$$
W_{pm}(\vartheta, \psi) = \sum_{k=1}^{m} \lambda_k p_k(e^{2\pi i \vartheta}) p_k(e^{2\pi i \psi})
$$

(2.16)

There are two levels of approximations: (a) Fourier approximation of eigenfunctions; (b) spectral decomposition approximation of the graphon operator. Hence the approximation error is
upper bounded as follows:

\[ \| W - W_{pm} \|_2 \leq \| W - W_{\varphi m} \|_2 + \| W_{\varphi m} - W_{pm} \|_2 \]

where \( W_{\varphi m} := \sum_{k=1}^m \lambda_k \varphi_k(x) \varphi_k(y) \).

**Theorem 2.5.** If there exists \( c > 0 \) such that \( \| W \|_2 \leq c \) and \( \| W_{pm} \|_2 \leq c \), then

\[ \| W^n - (W_{pm})^n \|_2 \leq nc^n \| W - W_{pm} \|_2, \quad \text{and} \quad \| e^W - e^{W_{pm}} \|_{op} \leq ce^c \| W - W_{pm} \|_2, \]

where the exponential function of a graphon operator \( W \) defined as a linear operator for \( L^2[0, 1] \) functions as follows: \( e^W := \sum_{k=0}^\infty \frac{1}{k!} W^k \) with \( W^0 = I \) representing the identity operator.

**Proof.** The proof is similar to the proof in graphon control [33]. The key is to use \( W^n - U^n = (W - U)P_{k-1}(W, U) \) where \( P_k(x, y) = \sum_{i=0}^k x^iy^{k-i} \) and that \( \| W_k \|_2 \leq \| W \|_2 \leq 1, k \geq 1 \). We obtain

\[ \| W^n - (W_{pm})^n \|_2 \leq \| P_{n-1}(W, W_{pm}) \|_2 \cdot \| W - W_{pm} \|_2 \leq nc^n \| W - W_{pm} \|_2. \] (2.17)

For any \( x \in L^2[0, 1] \)

\[ \| (e^W - e^{W_{pm}})x \|_2 \leq \sum_{k=0}^\infty \frac{1}{k!} \| (W^k - (W_{pm})^k)x \|_2 \]

\[ = \sum_{k=1}^\infty \frac{1}{k!} \| (W^k - (W_{pm})^k)x \|_2 \]

\[ \leq \sum_{k=1}^\infty \frac{1}{k!} kce^c \| (W - W_{pm}) \|_2 \| x \|_2 \]

\[ = ce^c \| (W_{pm}) \|_2 \| x \|_2. \] (2.18)

According to the definition of the operator norm \( \cdot \}_{op} \), we have

\[ \| e^W - e^{W_{pm}} \|_{op} \leq ce^c \| W - W_{pm} \|_2. \] (2.19)
Relation Between \( p_k(e^{2\pi i \cdot}) \) and the Fourier Basis

Consider a polynomial \( p_k(e^{2\pi i \cdot}) \) with highest order \( 2n \), then there exists a matrix \( T_k \in \mathbb{R}^{n \times n} \) such that

\[
p_k(e^{2\pi i \cdot}) = (e_1, e_2, \ldots, e_n)^T T_k (e_1, e_2, \ldots, e_n)
\]

with \( e_k = e^{2\pi ki} \).

Fourier Transformation of \( L^2[0, 1] \) Functions

For \( f \in L^2[0, 1] \), \( m^{th} \) partial sum of the Fourier series is given by

\[
S_f^m(x) = \sum_{k=-m}^m F_k e^{2\pi ikx}, \quad F_k = \int_0^1 f(x) e^{-2\pi ikx} \, dx
\]

By Riesz-Fischer theorem,

\[
\lim_{m \to \infty} \| S_f^m - f \|_2 = 0,
\]

where \( \| \cdot \|_2 \) is the \( L^2[0, 1] \) norm. An \( L^2[0, 1] \) function \( f \) can be approximated by finite Fourier series and hence can also be approximated by a polynomial function of \( e^{2\pi i \cdot} \) as \( p(e^{2\pi i \cdot}) \).

2.4 Step Function Graphons

Graphons generalize weighted graphs in the following sense. A function \( W \in \tilde{G}^{sp}_1 \) is called a step function if there is a partition \( Q = \{ Q_1, \ldots, Q_k \} \) of \([0, 1]\) into measurable sets such that \( W \) is constant on every product set \( Q_i \times Q_j \). The sets \( Q_i \) are the steps of \( W \). For every weighted graph \( G \) (on node set \( V(G) \)), a step function \( S_G \in \tilde{G}^{sp}_1 \) is given as follows: partition \([0, 1]\) into \( n \) measurable sets \( Q_1, \ldots, Q_n \) of measure \( \mu(Q_i) = \alpha_i \alpha_G \), then for \( x \in Q_i \) and \( y \in Q_j \), we let \( S_G(x, y) = \beta_{ij}(G) \), where \( \alpha_i \) denotes the node weight of \( i^{th} \) node, \( \alpha(G) = \sum_i \alpha_i \) and \( \beta_{ij}(G) \) denotes the weight of the edge from node \( i \) to node \( j \) (i.e., \( \beta_{ij} \) is the \( ij^{th} \) entry in the adjacency matrix of \( G \)). Evidently the function \( S_G \) depends on the labelling of the nodes of \( G \). We define the uniform partition \( P^N = \{ P_1, P_2, \ldots, P_N \} \) of \([0, 1]\) by setting \( P_k = [\frac{k-1}{N}, \frac{k}{N}) \), \( k \in \{1, N-1\} \) and \( P_N = [\frac{N-1}{N}, 1] \). Then \( \mu(P_i) = \frac{1}{N} \), \( i \in \{1, 2, \ldots, N\} \). Under the uniform partition, step functions can be represented by the pixel diagram on the unit square (see [40]).
2.4.1 Spectral Decomposition

Let \( \{P_1, \ldots, P_N\} \) form a partition of \([0, 1]\). Then the piece-wise constant function in \( L^2[0, 1] \) corresponding to \( u \) in \( \mathbb{R}^N \) can be represented by

\[
S_u(x) = \sum_{i=1}^{N} \chi_{P_i}(x)u_i,
\]

where \( \chi_A(\cdot) \) represents the indicator function. Define \( S_u \cdot S_v^T \) as the following

\[
[S_u \cdot S_v^T](x, y) := \sum_{i=1}^{N} \chi_{P_i}(x)u_i \sum_{j=1}^{N} \chi_{P_j}(y)v_j = \sum_{i=1}^{N} \sum_{j=1}^{N} \chi_{P_i}(x)\chi_{P_j}(y)u_iv_j.
\]

Note that this assumes \( S_u \) and \( S_v^T \) share the same partition \( \{P_1, \ldots, P_N\} \). More generally, we can define

\[
[S_u \cdot S_v^T](x, y) := \sum_{i=1}^{N} \chi_{P_i}(x)u_i \sum_{j=1}^{N} \chi_{Q_j}(y)v_j = \sum_{i=1}^{N} \sum_{j=1}^{N} \chi_{P_i}(x)\chi_{Q_j}(y)u_iv_j,
\]

where \( \{P_1, \ldots, P_N\} \) and \( \{Q_1, \ldots, Q_N\} \) can be different partitions of \([0, 1]\).

Let \( u_l \in \mathbb{R}^{N_l} \) for all \( l \) and \( N_l \in \{1, 2, \ldots\} \). If the set of orthogonal piece-wise constant functions \( \{S_{u_l} \in L^2[0, 1] : \langle S_{u_l}, S_{u_k} \rangle = 0, \text{ if } k \neq l \} \) spans the range of a graphon operator \( W \), then

\[
W(x, y) = \sum_{l=1}^{\infty} \lambda_l S_{u_l}(x)S_{u_l}(y).
\]

A step function graphon can be represented by

\[
W(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \chi_{P_i}(x)\chi_{P_j}(y)m_{ij}, \quad (x, y) \in [0, 1]^2
\tag{2.20}
\]

where \( M = [m_{ij}] \) is the corresponding matrix that defines the weight in each partition of \( W \).

**Proposition 2.6.** If the matrix \( M \) has a spectral decomposition \( M = V\Lambda_dV^T \), where \( \Lambda_d = \text{diag}(\lambda_1, \ldots, \lambda_d) \) and \( V = (v_1, \ldots, v_d) \) with \( v_i \) representing the normalized eigenvector of \( \lambda_i \), then the graphon \( W \) defined by

\[
W(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \chi_{P_i}(x)\chi_{P_j}(y)m_{ij}, \quad (x, y) \in [0, 1]^2
\tag{2.21}
\]
has a spectral representation given by

\[
W(x, y) = \sum_{l=1}^{d} \lambda_l [S_{v_l} \cdot S_{v_l}^T](x, y), \quad (x, y) \in [0, 1]^2.
\] (2.22)

Equivalently,

\[
W = \sum_{l=1}^{d} \lambda_l (S_{v_l} \cdot S_{v_l}^T).
\]

**Proof.** For all \((x, y) \in [0, 1]^2,\)

\[
W(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \chi_{P_i}(x) \chi_{P_j}(y) m_{ij}
= \sum_{i=1}^{N} \sum_{j=1}^{N} \chi_{P_i}(x) \chi_{P_j}(y) \sum_{l=1}^{d} \lambda_l v_l(i) v_l(j)
= \sum_{l=1}^{d} \lambda_l \sum_{i=1}^{N} \chi_{P_i}(x) v_l(i) \sum_{j=1}^{N} \chi_{P_j}(y) v_l(j)
= \sum_{l=1}^{d} \lambda_l [S_{v_l} \cdot S_{v_l}^T](x, y).
\] (2.23)

\[\square\]

Note that

\[\langle S_{v_l}, S_{v_k} \rangle = \int_{0}^{1} S_{v_l}(x) S_{v_k}(x) dx = 0, \quad \text{if} \ l \neq k\]

and

\[\langle S_{v_l}, S_{v_k} \rangle = \int_{0}^{1} S_{v_l}(x) S_{v_k}(x) dx = \frac{1}{N}, \quad \text{if} \ l = k\]

where we assume \(\{P_1, ..., P_N\}\) is \(N\) uniform partition of \([0, 1]\).

We see that piece-wise constant functions in \(L^2[0, 1]\) form eigenfunctions of step function graphons. As we know piece-wise constant functions in \(L^2[0, 1]\) form a dense subset of \(L^2[0, 1]\) space, they can be used to approximate eigenfunctions of general graphons.

### 2.4.2 Kronecker Product

**Kronecker Product of Graphs**

Define the Kroneker product of graphs as the graph generated by the Kroneker product of the corresponding adjacency matrices. The eigenvalues of the Kroneker product \((A \otimes B)\) are given
by \( \{ \lambda_i \mu_j, 1 \leq i \leq n, 1 \leq j \leq m \} \) where \( \lambda_i \) and \( \mu_j \) is the \( i \)th and \( j \)th eigenvalues of \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{m \times m} \) respectively.

**Kroneker Product of an \( L^2[0, 1] \) Piece-wise Constant Function and a Vector**

Define the Kroneker product of \( L^2[0, 1] \) piece-wise constant function \( S_u \) (with the partition \( \{ P_1, ..., P_N \} \)) and vector \( v \) as

\[
v \otimes S_u := \sum_{i=1}^{nN} \chi_{Q_i}(x)[v \otimes u]_i
\]

(2.24)

where \( \{ Q_{(i-1)n+1}, ..., Q_{in} \} \) is the uniform partition of \( P_i \), i.e., \( \{ Q_1, ..., Q_{nN} \} \) forms a refined partition of \([0, 1]\).

**Kroneker Product of a Step Function Graphon and a Matrix**

Define the Kroneker product of a step function graphon \( W \) (with the partition \( P = \{ P_1, ..., P_N \} \)) and a matrix \( A \in \mathbb{R}^{n \times n} \) as follows:

\[
A \otimes W := \sum_{i=1}^{nN} \sum_{j=1}^{nN} \chi_{Q_i}(x) \chi_{Q_j}(y) [A \otimes M]_{ij}
\]

(2.25)

where \( \{ Q_{(i-1)n+1}, ..., Q_{in} \} \) is the uniform partition of \( P_i \), i.e., \( \{ Q_1, ..., Q_{nN} \} \) forms a refined partition of \([0, 1]\) and \( M = [m_{ij}] \) is the corresponding matrix that defines the weight in each partition of \( W \). A special case would be

\[
\text{ones}_n \otimes W = \sum_{i=1}^{nN} \sum_{j=1}^{nN} \chi_{Q_i}(x) \chi_{Q_j}(y) [\text{ones} \otimes M]_{ij},
\]

(2.26)

where \( \text{ones}_n := 1^T_n 1_n \) with \( 1_n = [1, ..., 1]_n \) (i.e. \( \text{ones}_n \) is the \( n \times n \) matrix with all elements being 1).

The Kroneker product \( \text{ones}_n \otimes W \) has the same non-zero eigenvalues as \( W \). Furthermore, the corresponding eigenfunctions of \( \text{ones}_n \otimes W \) are simply given by \( \{ 1_n \otimes f_l : l \in \{ 1, 2, \ldots \} \} \) with \( \{ f_l : l \in \{ 1, 2, \ldots \} \} \) representing all the eigenfunctions of \( W \) that corresponds to the non-zero eigenvalues.

We note that \( \{ \text{ones}_n \otimes W : n \geq 1 \} \) forms an equivalent class of graphons under both the \( \delta_\square \) metric and the \( \delta_2 \) metric, that is, each member of \( \{ \text{ones}_n \otimes W : n \geq 1 \} \) is of zero cut distance and zero \( \delta_2 \) distance from others.
Kroneker Product of a General Graphon and a Matrix

Define the Kroneker product of a general graphon $W$ and a matrix $A \in \mathbb{R}^{n \times n}$ as follows:

$$A \otimes W := \sum_{i=1}^{n} \sum_{j=1}^{n} \chi_{Q_i}(x) \chi_{Q_j}(y) W(nx(\text{mod } 1), ny(\text{mod } 1)) a_{ij},$$

where $\{Q_1, Q_2, \ldots, Q_n\}$ is a uniform partition of $[0, 1]$. Note that this definition is compatible with the previous definition in (2.25). Similar claims hold:

1. The Kroneker product $\text{ones}_n \otimes W$ has the same non-zero eigenvalues as $W$. Furthermore, the corresponding eigenfunctions of $\text{ones}_n \otimes W$ are simply given by $\{1_{n} \otimes f_l : l \in \{1, 2, \ldots\}\}$ with $\{f_l : l \in \{1, 2, \ldots\}\}$ representing all the eigenfunctions of $W$ that corresponds to the non-zero eigenvalues.

2. $\{\text{ones}_n \otimes W ; n \geq 1\}$ forms an equivalent class of graphons under both the $\delta_1$ metric and the $\delta_2$ metric.

2.5 The Graphon Unitary Operator Algebra

It is evident that the operator composition defined in (2.11) above yields an operator algebra with a multiplicative binary operation possessing the associativity, left distributivity, right distributivity properties and compatibility with the scalar field $\mathcal{R}$, that is, for any $V, W, H$ in the vector space $L^2[0, 1]^2$ and $a, b \in \mathcal{R}$,

$$(VW)H = V(WH),$$

$$V(W + H) = VW + VH,$$

$$(W + H)V = WV + HV,$$

$$(aW)(bH) = (ab)WH.$$ 

Thus we have an operator algebra $\mathcal{G}_A$ over the field $\mathcal{R}$ acting on elements of $L^2[0, 1]$ with operator multiplication as given in (2.9). By adjoining the identity element $I$ to the algebra $\mathcal{G}_A$ (see e.g. [60]) we obtain a unitary algebra $\mathcal{G}_{AI}$. The identity element $I$ is defined as follows: for any $W \in L^2[0, 1]^2$

$$[IW](x, y) = \int_0^1 W(z, y) \delta(x, z) dz = W(x, y),$$

(2.28)
where $\delta(\cdot, z)dz$ is the measure satisfying $\int_0^1 u(z)\delta(x,z)dz = u(x)$ for all $u \in L^2[0,1]$, and in particular $\int_0^1 \delta(x,z)dz = 1$.

The graphon unitary operator algebra $G_{AI}$ will be used in the definition of the controllability Gramian and the input operator. More specifically, we use the subset $G^{1}_{AI} = \{(aI + A) : A \in G^{1}_A, a \in \mathcal{R}\}$ where $G^{1}_A$ is the subset of $G_A$ that corresponds to $\tilde{G}_{sp}^{1}$.  

### 2.6 Graphon Differential Equations

Let $X$ be a Banach space. A linear operator $A : D(A) \subset X \to X$ is closed if $\{(x, Ax) : x \in D(A)\}$ is closed in the product space $X \times X$(see [45]). $L(X)$ denotes the Banach algebra of all linear continuous mappings $T : X \to X$. $L^p(a,b;X)$ denotes the Banach space of equivalent classes of strongly measurable (in the B"ochner sense) mappings $[a,b] \to X$ that are $p$-integrable, $1 \leq p < \infty$, with norm $\|f\|_{L^p(a,b;X)} = \left[ \int_a^b |f(s)|^p ds \right]^{\frac{1}{p}}$. Let $A : [0,1]^2 \to [-1,1]$ be a graphon and hence a bounded and closed linear operator from $L^2[0,1]$ to $L^2[0,1]$. Following [61], $A$ is the infinitesimal generator of the uniformly (hence strongly) continuous semigroup

$$S_A(t) := e^{At} = \sum_{k=0}^\infty \frac{t^k A^k}{k!}.$$ (2.29)

Therefore, the initial value problem of the graphon differential equation

$$\dot{y}_t = Ay_t, \quad y_0 \in L^2[0,1]$$ (2.30)

has a solution given by $y_t = e^{At}y_0$.

**Theorem 2.6.** Let $\{A_N\}_{N=1}^\infty$ be a sequence of graphons such that $A_N \to A_*$ as $N \to \infty$ in the $L^2$ metric. Then for all $x \in L^2[0,1]$, $e^{A_N t}x \to e^{A_* t}x$ as $N \to \infty$ in the $L^2$ metric where the convergence is pointwise in time and uniform on any time interval $[0, T]$.

**Proof.** Let us define $P_k(x,y) = \sum_{i=0}^k x^{k-i}y^i$, then

$$x^k - y^k = (x-y)P_{k-1}(x,y),$$ (2.31)

and we see that for $k \geq 1$

$$A_N^k - A_*^k = P_{k-1}(A_N, A_*)(A_N - A_*).$$
Consider a sequence of graphons associated with a partition $Q$. Let $A_\Delta = A_N - A_*$. For an arbitrary $x \in L^2[0, 1]$ and finite $t$, $0 \leq t < \infty$,

$$
\| e^{A_N t} x - e^{A_* t} x \|_2 = \left\| \sum_{k=0}^{\infty} \frac{t^k}{k!} (A_N^k - A_*^k) x \right\|_2
\leq \sum_{k=1}^{\infty} \frac{t^k}{k!} \| (A_N^k - A_*^k) x \|_2
\leq \sum_{k=1}^{\infty} \frac{t^k}{k!} \| P_{k-1}(A_N, A_*) \|_2 \cdot \| A_\Delta \|_2 \| x \|_2
\leq \sum_{k=1}^{\infty} \frac{t^k}{k!} \cdot k \cdot \| A_\Delta \|_2 \| x \|_2
= t e^t \cdot \| A_\Delta \|_2 \| x \|_2.
$$

Therefore, the left hand side goes to zero as $\| A_\Delta \|_2 = \| A_N - A_* \|_2$ goes to zero. It follows that for $t \in [0, T]$,

$$
\| e^{A_N t} x - e^{A_* t} x \|_2 \leq \max_{t \in [0, T]} \| e^{A_N t} x - e^{A_* t} x \|_2
\leq \max_{t \in [0, T]} t e^t \cdot \| A_\Delta \|_2 \| x \|_2
= T e^T \cdot \| A_\Delta \|_2 \| x \|_2.
$$

Hence the convergence is uniform in $t$ when $t$ is in some finite interval $[0, T]$.

\[ \square \]

2.7 Monotonically Increasing Graphon Sequences

Consider a sequence of graphons associated with a partition $Q = \{Q_1, Q_2, \ldots\}$ of $[0, 1]$ satisfying $\sum_{i=1}^{\infty} |Q_i| = 1$. Define the $(n + 1)^{th}$ strip graphon $U_{Q_{n+1}}$ as follows:

$$
U_{Q_{n+1}}(x, y) = \begin{cases} 
  c_{n+1}(x, y), & (x, y) \in Q_{n+1} \times \cup_{i=1}^{n} Q_i \\
  c_{n+1}(y, x), & (x, y) \in \cup_{i=1}^{n} Q_i \times Q_{n+1} \\
  l_{n+1}(x, y), & (x, y) \in Q_{n+1} \times Q_{n+1} \\
  0, & \text{otherwise}
\end{cases}
$$

where $c_{n+1} : Q_{n+1} \times \cup_{i=1}^{n} Q_i \to [-1, 1]$ is a measurable function and $l_{n+1} : Q_{n+1} \times Q_{n+1} \to [-1, 1]$ is a symmetric measurable function. Define a monotonically increasing graphon sequence $\{W_n\}$
recursively via
\[ W_{n+1} = W_n + U_{Q_{n+1}}, n \geq 0, \quad W_0 = 0. \]

From the construction, the optimal measure preserving bijection \( \phi \) which achieves the \( \delta_2 \) distance yields
\[ \delta_2(W_{n+m}, W_n) = \| U_{Q_{n+1}} + \cdots + U_{Q_{n+m}} \|_2, \quad m \geq 1, n \geq 0, \]

in other words \( \phi \) is the identity mapping from \([0, 1]\) to \([0, 1]\).

The nodal weight of a node in a finite graph can be characterized by the size of its corresponding block in the step function graphon representation. Without losing generality, the weights of the nodes on a graph can be scaled to sum up to 1. A graph is said to have uniform nodal weights, if all the nodes in the graph have the same weight.

If \( \{c_n(\cdot, \cdot)\} \) and \( \{l_n(\cdot, \cdot)\} \) are two sequences of constant functions, \( \{W_n\} \) then corresponds to a sequence of graphs with non-uniform nodal weights. An example of the size of the partitions is given by \( |Q_n| = \frac{1}{4}(\frac{3}{4})^{n-1} \). Intuitively this means that nodes that come later have discounted nodal weights with respect to previous ones. Note that \( W_n \) has only zero values on the last partition \( \cup_{i=n+1}^\infty Q_i \).

If the partition of \( W_n \) is interpreted as \( \{Q_1, Q_2, \ldots, Q_n, \cup_{i=n+1}^\infty Q_i\} \), a further refinement of this partition will yield a graph with uniform nodal weights that corresponds to \( W_n \). This is illustrated in Figure 2.3, where a graph with non-uniform nodal weights in (b) has an equivalent representation by a graph with uniform nodal weights in (c) where yellow nodes are the nodes with uniform nodal weights.

Under the refinement of the partition, one optimal measure preserving bijection is still the
identity mapping from \([0, 1]\) to \([0, 1]\). The measure preserving bijection can be considered as
the permutation of the refined partitions which corresponds to the relabelings on nodes of the
 corresponding finite graph. With the refinement, one can apply the graphon dynamical system
model directly without further modifications.

**Proposition 2.7.** Consider a monotonically increasing graphon sequence \(\{W_n\}\) along with its
partition \(Q = \{Q_1, Q_2\ldots\}\), \(\sum_{i=1}^{\infty} |Q_i| = 1\). Then \(\{W_n\}\) converges to a graphon limit \(W\) in \(\tilde{G}_{1}^{sp}\)
under the \(L^2\) metric. Furthermore, \(\{W_n\}\) converges to (the equivalence class of) the graphon
limit \(W\) in \(G_{1}^{sp}\) under the \(\delta_2\) metric and under the \(\delta_{\square}\) metric.

**Proof.** \(\sum_{i=1}^{\infty} |Q_i| = 1\) implies \(\lim_{n \to \infty} |Q_n| = 0\). Then, by construction, \(\{W_n\}\) forms a Cauchy
sequence in \(\tilde{G}_{1}^{sp}\) under the \(L^2\) metric, and hence converges to a unique limit \(W\) in \(\tilde{G}_{1}^{sp}\). By the
ordering of metrics given in (2.8), \(\{W_n\}\) converges to (the equivalence class of) the graphon
limit \(W\) in \(G_{1}^{sp}\) under the \(\delta_2\) metric and under the \(\delta_{\square}\) metric. \(\square\)

### 2.8 Cosinusoidal Graphons

#### 2.8.1 Cosinusoidal Graphons

We define a cosinusoidal graphon as any graphon that can be represented by

\[
W(\varphi, \vartheta) = a_0 + \sum_{k=1}^{\infty} b_k \cos(2\pi k(\varphi - \vartheta)), \quad (\varphi, \vartheta) \in [0, 1]^2
\]  

(2.33)

or equivalently by

\[
W(\varphi, \vartheta) = a_0 + \sum_{k=1}^{\infty} b_k \frac{e^{2\pi i k(\varphi - \vartheta)} + e^{-2\pi i k(\varphi - \vartheta)}}{2}, \quad (\varphi, \vartheta) \in [0, 1]^2.
\]  

(2.34)

Apparently cosinusoidal graphons are symmetric and they contain structures that are suitable to
fit Toeplitz matrices.

**Eigenfunctions and Eigenvalues**

Cosinusoidal graphons have simple spectral characterizations. The eigenfunctions of the graphon
in (2.33) are \(L^2[0, 1]\) functions as follows:

\[1, \{\cos 2\pi k(\cdot) : 1 \leq k \leq \infty\}, \{\sin 2\pi k(\cdot) : 1 \leq k \leq \infty\}\].
The corresponding eigenvalues are: \(a_0, \{\frac{b_k}{2} : 1 \leq k \leq \infty\}, \{\frac{b_k}{2} : 1 \leq k \leq \infty\}\). Moreover, the eigenfunctions form a complete orthogonal basis in \(L^2[0, 1]\) (see e.g. [62]) and satisfy the following: for all \(k, l \geq 1\),

\[
\begin{align*}
\langle \cos 2\pi k(\cdot) \cos 2\pi l(\cdot) \rangle &= \begin{cases} 
\frac{1}{2}, & k = l \\
0, & k \neq l
\end{cases}, \\
\langle \sin 2\pi k(\cdot) \sin 2\pi l(\cdot) \rangle &= \begin{cases} 
\frac{1}{2}, & k = l \\
0, & k \neq l
\end{cases}; \\
\langle \sin 2\pi k(\cdot) \cos 2\pi l(\cdot) \rangle &= 0; \\
\langle 1, \cos 2\pi k(\cdot) \rangle &= 0; \\
\langle 1, \sin 2\pi k(\cdot) \rangle &= 0;
\end{align*}
\] (2.35)

where \(\langle \cdot, \cdot \rangle\) denotes the inner product in \(L^2[0, 1]\) (i.e. \(\langle u, v \rangle = \int_0^1 u(\vartheta)v(\vartheta)\,d\vartheta\)).

**Operation on \(L^2[0, 1]\) Functions**

Consider an \(L^2[0, 1]\) function \(x(\cdot)\) represented by Fourier series as

\[
x(\vartheta) = K_0 + \sum_{k=1}^{\infty} (\alpha_k \sin(\vartheta) + \beta_k \cos(\vartheta)), \quad \vartheta \in [0, 1].
\] (2.36)

Operating \(W\) on \(x(\cdot)\) yields

\[
[Wx](\vartheta) = a_0 K_0 + \sum_{k=1}^{\infty} \frac{b_k}{2} (\alpha_k \sin(\vartheta) + \beta_k \cos(\vartheta)), \quad \vartheta \in [0, 1].
\] (2.37)

Similarly,

\[
[W^m x](\vartheta) = a_0^m K_0 + \sum_{k=1}^{\infty} \left(\frac{b_k}{2}\right)^m (\alpha_k \sin(\vartheta) + \beta_k \cos(\vartheta)), \quad \vartheta \in [0, 1];
\]

\[
[e^W x](\vartheta) = e^{a_0} K_0 + \sum_{k=1}^{\infty} e^{\frac{b_k}{2}} (\alpha_k \sin(\vartheta) + \beta_k \cos(\vartheta)), \quad \vartheta \in [0, 1].
\] (2.38)

**Functions of Cosinusoidal Graphons**

Power functions of a cosinusoidal graphon \(W\) are given by

\[
W^m(\varphi, \vartheta) = a_0^m + \sum_{k=1}^{\infty} \left(\frac{b_k}{2}\right)^{m-1} b_k \cos(2\pi k(\varphi - \vartheta)), \quad (\varphi, \vartheta) \in [0, 1]^2, \quad m \geq 1.
\] (2.39)

The zero power of \(W\) is defined as \(W^0 = I\) where \(I\) is the identity operator. It can be represented via a singular measure \(\sigma\) on the set of diagonal elements of \([0, 1]^2\) denoted by \(\mathcal{D} =\)
\{(x, y) : x = y, (x, y) \in [0, 1]^2\}. The measure \(\sigma\) satisfies: \(\sigma(x, y) = 1\), if \((x, y) \in \mathcal{D}\), and \(\sigma(x, y) = 0\) otherwise.

The exponential function of a cosinusoidal graphon is given by

\[ e^{Wt} = I + U_t, \quad t \in \mathcal{R} \tag{2.40} \]

with \(U_t\) given by

\[ U_t(\varphi, \vartheta) = (e^{a_0t} - 1) + 2 \sum_{k=1}^{\infty} (e^{b_kt} - 1) \cos 2\pi k(\varphi - \vartheta), \quad (\varphi, \vartheta) \in [0, 1]^2. \tag{2.41} \]

We note that \(e^{Wt} (t > 0)\) is an element in the graphon unitary operator algebra \(\mathcal{G}_{AI}\).

Furthermore, we can explicitly obtain the following function of \(W\):

\[
CG(W, [0, T]) := \int_0^T e^{Wt} e^{W^T} dt = \int_0^T (I + U_t)(I + U_t) dt = T I + \int_0^T (2U_t + U_t^2) dt, \tag{2.42}
\]

where the second part \(\int_0^T (2U_t + U_t^2) dt\) is represented as follows: for \((\varphi, \vartheta) \in [0, 1]^2\),

\[
\left[ \int_0^T (2U_t + U_t^2) dt \right](\varphi, \vartheta) = \int_0^T (e^{2a_0t} - 1) dt + 2 \sum_{k=1}^{\infty} \cos 2\pi k(\varphi - \vartheta) \int_0^T (e^{b_kt} - 1) dt \tag{2.43}
\]

\[
= \left( \frac{e^{2a_0T} - 1}{2a_0} - T \right) + 2 \sum_{k=1}^{\infty} \cos 2\pi k(\varphi - \vartheta) \left( \frac{e^{b_kT} - 1}{b_k} - T \right) \quad \text{(if } a_0 \text{ and } \{b_k : k \geq 1\} \text{ are non-zero coefficients)}
\]

This function \(CG(W, [0, T])\) is called the controllability Gramian operator for the graphon system \((W; I)\) which will be discussed in details in Chapter 4. We note that \(CG(W, [0, T])\) is also an element in the graphon unitary operator algebra \(\mathcal{G}_{AI}\).
2.8.2 Generalized Cosinusoidal Graphon

We define a \textit{generalized cosinusoidal graphon} as any graphon that can be represented by the following:

\[
W(\vartheta, \varphi) = a_0 + \sum_{k=1}^{\infty} [b_k \cos(2\pi k(\varphi - \vartheta)) + c_k \cos(2\pi k(\varphi + \vartheta))], \quad (\varphi, \vartheta) \in [0, 1]^2. \tag{2.44}
\]

This enables the representation of graphons with \(\cos 2\pi k(\varphi + \vartheta)\) terms and can fit a more general class of matrices than Toeplitz matrices.

The two dimensional function \(\cos 2\pi k(\varphi + \vartheta)\) with \((\varphi, \vartheta) \in [0, 1]^2\) has some interesting properties when operating on \(L^2[0, 1]\) functions:

\[
\begin{align*}
\int_0^1 \cos 2\pi k(\vartheta + \varphi) \sin 2\pi l \vartheta d\vartheta &= \begin{cases} 
-\frac{1}{2} \sin 2\pi k\varphi & k = l \\
0 & k \neq l 
\end{cases}; \\
\int_0^1 \cos 2\pi k(\vartheta + \varphi) \cos 2\pi l \vartheta d\vartheta &= \begin{cases} 
\frac{1}{2} \cos 2\pi k\varphi & k = l \\
0 & k \neq l 
\end{cases}. \tag{2.45}
\end{align*}
\]

The functions \(\cos 2\pi k(\vartheta + \varphi)\) and \(\cos 2\pi k(\vartheta - \varphi)\) with \((\varphi, \vartheta) \in [0, 1]^2\) satisfy:

\[
\begin{align*}
\int_0^1 \cos 2\pi k(\vartheta + \beta) \cos 2\pi l(\beta - \varphi) d\beta &= \begin{cases} 
\frac{1}{2} \cos 2\pi k(\vartheta + \beta) & k = l \\
0 & k \neq l 
\end{cases}; \\
\int_0^1 \cos 2\pi k(\vartheta + \beta) \cos 2\pi k(\beta + \varphi) d\beta &= \begin{cases} 
\frac{1}{2} \cos 2\pi k(\vartheta - \beta) & k = l \\
0 & k \neq l 
\end{cases}; \\
\int_0^1 \cos 2\pi k(\vartheta - \beta) \cos 2\pi k(\beta - \varphi) d\beta &= \begin{cases} 
\frac{1}{2} \cos 2\pi k(\vartheta - \beta) & k = l \\
0 & k \neq l 
\end{cases}. \tag{2.46}
\end{align*}
\]
Operation on \(L^2[0, 1]\) Functions

Let \(x(\cdot) = K_0 + \sum_{k=1}^{\infty}(\alpha_k \sin 2\pi k(\cdot) + \beta_k \cos 2\pi k(\cdot))\) be an \(L^2[0, 1]\) function. Then

\[
[W x](\varphi) = a_0 K_0 + \sum_{k=1}^{\infty} \alpha_k \left(\frac{b_k - c_k}{2}\right) \sin 2\pi k \varphi + \beta_k \left(\frac{b_k + c_k}{2}\right) \cos 2\pi k \varphi, \quad \varphi \in [0, 1];
\]

\[
[W^m x](\varphi) = a_0^m K_0 + \sum_{k=1}^{\infty} \alpha_k \left(\frac{b_k - c_k}{2}\right)^m \sin 2\pi k \varphi + \beta_k \left(\frac{b_k + c_k}{2}\right)^m \cos 2\pi k \varphi, \quad \varphi \in [0, 1];
\]

\[
[e^W x](\varphi) = e^{a_0} K_0 + \sum_{k=1}^{\infty} \alpha_k e^{\left(\frac{b_k - c_k}{2}\right)} \sin 2\pi k \varphi + \beta_k e^{\left(\frac{b_k + c_k}{2}\right)} \cos 2\pi k \varphi, \quad \varphi \in [0, 1].
\]

Functions of Generalized Cosinusoidal Graphons

The square function of \(W\) is given by:

\[
[W^2](\varphi, \psi) = a_0^2 + \sum_{k=1}^{\infty} \frac{(b_k^2 + c_k^2)}{2} \cos 2\pi k (\psi - \varphi) + \sum_{k=1}^{\infty} b_k c_k \cos 2\pi k (\varphi + \psi), \quad (\varphi, \psi) \in [0, 1]^2.
\]

(2.47)

Higher order power functions of generalized cosinusoidal graphons do not have simple forms like cosinusoidal graphons as in (2.39).

2.9 Appendices

2.9.1 Cauchy Graphon Sequences

**Theorem 2.7.** If a graphon sequence is Cauchy in the \(L^2\) metric then it is also a Cauchy sequence in the cut metric and under both metrics, the limits are identical up to measure preserving transformations.

**Proof.** Let \(\{X_n\}_{n=1}^{\infty}\) be a Cauchy graphon sequence in the \(L^2\), that is, for every positive real number \(\varepsilon > 0\) there is a positive integer \(N\) such that for all positive integers \(m, n > N\), the distance

\[
d_{L^2}(X_m, X_n) < \varepsilon.
\]

Since the graphon space \((\mathcal{G}^{sp}_1, d_{L^2})\) is complete, the limit denoted by \(X_{L^2}^{\infty}\) exists in the graphon space with the \(L^2\) metric. Note that \(L_2(X_m, X_n) \leq d_{L^2}(X_m, X_n)\). Therefore

\[
\delta_{L_2}(X_m, X_n) < \varepsilon.
\]
This implies that the graphon \( \{X_n\}_{n=1}^\infty \) is Cauchy in the cut metric. Since the metric space \((G_1^{sp}, \delta_\square)\) is compact, it is complete. Therefore the sequence converge to a unique limit in the cut metric. Since
\[
\lim_{n \to \infty} d_{L^2}(X_n, X_\infty^{L^2}) = 0
\]
and \(\delta_\square(X_n, X_\infty^{L^2}) \leq d_{L^2}(X_n, X_\infty^{L^2})\),
\[
\lim_{n \to \infty} \delta_\square(X_n, X_\infty^{L^2}) = 0.
\]
The uniqueness of the graphon limit in the cut metric ensures that \(X_\infty^{L^2}\) is also the unique limit of \(\{X_n\}\) in the cut metric. Therefore, under both metrics, the limits are identical up to measure preserving transformations.

The compactness of a set in \(G_1^{sp}\) with respect to the \(\delta_2\) metric implies its compactness with respect to the cut metric, and the limit of a convergent sequence in the \(\delta_2\) metric is the same unique limit of the same sequence in cut metric.

### 2.9.2 Spectral Theorems

**Theorem 2.8** ([57, Chapter 2, Theorem 5.2]). If \(T\) is a compact self-adjoint operator on Hilbert space \(\mathcal{H}\), then \(T\) has only a countable number of distinct eigenvalues. If \(\{\lambda_1, \lambda_2, \ldots\}\) are the distinct nonzero eigenvalues of \(T\), and \(P_n\) is the projection of \(\mathcal{H}\) onto \(\ker(T - \lambda_n I)\), then \(P_m P_n = P_n P_m = 0\) if \(n \neq m\), each \(\lambda_n\) is real, and
\[
T = \sum_{n=1}^{\infty} \lambda_n P_n
\]
where the series converges to \(T\) in the metric defined the norm of \(\mathcal{B}(\mathcal{H})\) (i.e. the set of bounded linear operator on \(\mathcal{H}\)).

**Corollary 2.2** ([57, Chapter 2, Corollary 5.4]). If \(T\) is a compact self-adjoint operator, then there is a sequence \(\{\mu_n\}\) of real numbers and an orthonormal basis \(\{e_n\}\) for (\(\ker T\))\(^\perp\) such that for all \(h\),
\[
T h = \sum_{n=1}^{\infty} \mu_n \langle h, e_n \rangle e_n
\]
Theorem 2.9 ([63, Theorem 7.3 (Spectral theorem of F. Rellich)]). Let the compact Hermitian operator $A : \mathcal{H} \to \mathcal{H}$ be given on the Hilbert space $\mathcal{H}$ satisfying $A \neq 0$. Then we have a finite or countably infinite system of orthonormal elements $\{\varphi_i\}_{i=1,2,...}$ in $\mathcal{H}$ such that

a) The elements $\varphi_i$ are eigenfunctions to the eigenvalues $\lambda_i \in \mathbb{R}$ ordered as follows:

$$\|A\| = |\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq ... \geq 0,$$

more precisely,

$$A\varphi_i = \lambda_i \varphi_i, \quad i = 1, 2, ...$$

If the set $\{\varphi_i\}_i$ is infinite, we have the asymptotic behavior

$$\lim_{i \to \infty} \lambda_i = 0.$$

b) For all $x \in \mathcal{H}$ we have the representations

$$Ax = \sum_{i=1}^{\infty} \lambda_i (\varphi_i, x) \varphi_i \quad \text{and} \quad \langle x, Ax \rangle = \sum_{i=1}^{\infty} \lambda_i |(\varphi_i, x)|^2.$$
Chapter 3

Graphon Dynamical System
Representations of Network Systems

3.1 Network Systems and Their Limit Systems

3.1.1 Network System Model

Consider an interlinked network $S^N$ of linear (symmetric) dynamical subsystems $\{S^N_i; 1 \leq i \leq N\}$, each with an $n$ dimensional state space. The subsystem $S^N_i$ at the node $V_i$ in the network $G_N(V,E)$ has interactions with $S^N_j$, $1 \leq j \leq N$, specified as below:

$$
\dot{x}^N_i = \frac{1}{nN} \sum_{j=1}^{N} A_{ij} x^j_t + \frac{1}{nN} \sum_{j=1}^{N} B_{ij} u^j_t, \quad t \in [0, T]
$$

$$
S^N_i: \quad x^i_t, u^i_t \in \mathbb{R}^n, i \in \{1, ..., N\},
$$

with $A_N = [A_{ij}], B_N = [B_{ij}] \in \mathbb{R}^{nN \times nN}$, the (symmetric) block-wise adjacency matrices of $G_N(V,E)$ and of the input graph, where $A_{ij} = [0]$ if $S^N_i$ has no connection to $S^N_j$ and similarity for $B_{ij}$. Then the (symmetric) linear dynamics for the network system $S^N(A_N, B_N, G_N)$ can be represented by

$$
\dot{x} = A_N \circ x_t + B_N \circ u_t, \quad t \in [0, T]
$$

$$
S^N: \quad x_t, u_t \in \mathbb{R}^{nN}, \quad A_N, B_N \in \mathbb{R}^{nN \times nN}, \quad (3.1)
$$
where \( \circ \) denotes the so-called averaging operator given by

\[
A_N \circ x = \frac{1}{(nN)} A_N x.
\] (3.2)

Let \( S = \times_{N=1}^{\infty} S^N \) where \( S^N = \cup_{A_N, B_N, G_N} S^N (A_N, B_N, G_N) \). For simplicity, we require the elements of \( A_N \) and \( B_N \) to be in \([-1, 1]\) for each \( N \) (note that in general \( A_N \) and \( B_N \) have elements that are bounded real numbers for which case we would achieve similar results). In addition, we note that if we take the supremum norm on vectors in \( \mathbb{R}^{nN} \), i.e. \( \| x \|_{\infty} = \sup_{|x| \neq 0} \frac{|A_N x|}{\| x \|_{\infty}} \), then \( \| A \|_{\text{op}\infty} \leq 1 \).

### 3.1.2 Network Systems Described by Step Functions

Let \( \{ (A_N; B_N) \}_{N=1}^{\infty} \in S \) denote a sequence of systems with the node averaging dynamics each of which is described according to (3.1). Let \( |A_{Nij}| \leq 1 \) and \( |B_{Nij}| \leq 1 \) for all \( i, j \in \{1, \ldots, nN\} \).

Let \( A_{s}[N], B_{s}[N] \in \tilde{G}^{\text{SP}}_1 \) be the step functions corresponding one-to-one to \( A_N \) and \( B_N \); these are specified using the uniform partition \( P^{nN} = \{ P_1, \ldots, P_{nN} \} \) of \([0, 1]\) with \( P_i = \left[ \frac{i-1}{nN}, \frac{i}{nN} \right) \), \( 1 \leq i < nN \), and \( P_{n,N} = \left[ \frac{nN-1}{nN}, 1 \right] \), by the following matrix-to-step function mapping \( M_G \): for all \( i, j \in \{1, 2, \ldots, nN\} \),

\[
A_{s}[N](x, y) := A_{Nij}, \quad \forall (x, y) \in P_i \times P_j,
\] (3.3)

and similar for \( B_{s}[N] \).

Define a piece-wise constant (PWC) function on \( \mathcal{R} \) to be any function of the form \( \sum_{k=1}^{l} \alpha_k \psi_{I_k} \) where \( \alpha_1, \ldots, \alpha_l \) are complex numbers and each \( I_k \) is a bounded interval (open, closed, or half-open). Let \( L^2_{\text{pwc}}[0, 1] \) denote the space of piece-wise constant \( L^2[0, 1] \) functions under the uniform partition \( P^{nN} \).

Let \( u^t_s \in L^2_{\text{pwc}}[0, 1] \) correspond one-to-one to \( u_t \in \mathcal{R}^{nN} \) via the following vector to PWC function mapping also denoted by \( M_G \): for all \( i \in \{1, \ldots, nN\} \),

\[
u_t^s(\alpha) := u_t(i), \quad \forall \alpha \in P_i,
\] (3.4)

and \( x^s_t \in L^2_{\text{pwc}}[0, 1] \) similarly correspond one-to-one to \( x_t \in \mathcal{R}^{nN} \).

**Lemma 3.1.** The trajectories of the system in (3.1) correspond one-to-one under the mapping
$M_G$ to the trajectories of the system

$$\dot{x}_t^s = A_s^{[N]} x_t^s + B_s^{[N]} u_t^s, \quad t \in [0, T]$$

with graphon operations defined according to (2.9).

Proof. Let the pair $(A_N; B_N)$ represent the system defined in (3.1) and $(A_s^{[N]}; B_s^{[N]})$ represent the system defined in (3.5).

Consider any $x_s \in L^2_{pwc}[0, 1]$ and its corresponding vector $x \in \mathcal{R}^{nN}$ following the vector-to-PWC-function mapping $M_G$. Since

$$[A_s^{[N]} x_s](\alpha) = \int_0^1 A_s^{[N]}(\alpha, \beta) x_s(\beta) d\beta, \quad x_s \in L^2_{pwc}[0, 1],$$

it follows that for all $\alpha \in P_i$,

$$[A_s^{[N]} x_s](\alpha) = \int_0^1 A_s^{[N]}(\alpha, \beta) x_s(\beta) d\beta$$

$$= \sum_{j=1}^{nN} \int_{P_j} A_s^{[N]}(\alpha, \beta) x_s(\beta) d\beta$$

$$= \sum_{j=1}^{nN} \int_{P_j} A_{Nij} x_j d\beta = \sum_{j=1}^{nN} \frac{1}{nN} A_{Nij} x_j$$

$$= \frac{1}{nN} [A_N x]_i = [A_N \circ x]_i,$$

where $x_j$ denotes the $j^{th}$ element of $x \in \mathcal{R}^{nN}$ and $[A_N x]_i$ denotes the $i^{th}$ element of $A_N x \in \mathcal{R}^{nN}$.

This implies that the step function $A_s^{[N]}$ in the graphon space, considered as an operator, represents a mapping in $L^2[0, 1]$; this operator is equivalent to the matrix transformation $A_N$ with $\circ$ operation in $\mathcal{R}^{nN}$ and the corresponding mapping $M_G$. A similar conclusion holds for $B_s^{[N]}$ and $B_N$.

Hence the we conclude that the trajectory of the system $(A_N; B_N)$ corresponds one-to-one to that of $(A_s^{[N]}; B_s^{[N]})$ under the corresponding vector-to-PWC-function mapping $M_G$. □
3.1.3 Limits of Sequences of Network Systems

Now the sequence of network systems with the node averaging dynamics can be described by the sequence of step function operators as \( \{(A_s^{[N]}, B_s^{[N]})\}_{N=1}^{\infty} \). Let the graphon sequences \( \{A_s^{[N]}\} \) and \( \{B_s^{[N]}\} \) be Cauchy sequences of step functions in \( L^2[0,1]^2 \) (under the same measure preserving bijection). Due to the completeness of \( L^2[0,1]^2 \), the respective graphon limits \( A \) and \( B \) exist and these will then necessarily be the limits in the cut metric (see Section 2.1.4 and [40]).

In fact, we can generalized the control input operator \( B \) to \( G_1A_1I \), i.e., \( B \) can consists of the identity operator part and the graphon part as \( B = \beta I + B_s^{[N]} \).

Consider a sequence of systems \( \{(A_s^{[N]}, B_s^{[N]})\} \in \tilde{G}_1^{sp} \times \tilde{G}_1^{sp} \). Decompose the input operator into the identity part and the graphon part as \( B^{[N]} = \beta_s^{[N]} I + B_s^{[N]} \).

**Definition 3.1.** A sequence of systems \( \{(A_s^{[N]}, B_s^{[N]})\} \in \tilde{G}_1^{sp} \times \tilde{G}_1^{sp} \) is convergent if

1. there exist \( \beta \in \mathcal{R} \) such that
   \[
   \lim_{N \to \infty} \beta_s^{[N]} = \beta
   \]

2. there exist \( A, B \in \tilde{G}_1^{sp} \) such that \( \{(A_s^{[N]}, B_s^{[N]})\} \) converges to \( (A; B) \) in the \( L^2 \) metric, i.e. \( A_s^{[N]} \to A \) and \( B_s^{[N]} \to B \) under the same sequence of measure preserving bijections in the \( L^2 \) metric.

Then the limit system is represented by \( (A; B) \) where \( B = \beta I + B \). With an abuse of notation, in the following sections we use \( B \) and \( B^{[N]}_s \) to represent input operators in \( G_1^{1}_{A_1T} \).

![Fig. 3.1 A Weighted Graph from a Sequence Converging to the Limit Graphon](image)

**Fig. 3.1** A Weighted Graph from a Sequence Converging to the Limit Graphon \( W(x, y) = 1 - \max(x, y), 0 \leq x, y \leq 1 \) with \( x, y \) measured from the top left

3.1.4 Limit Graphon Systems

We follow [45] and specialize both the Hilbert space of states \( H \) and the Hilbert space of controls \( U \) appearing there to the space \( L^2(\mathcal{R}; L^2[0,1]) \). We formulate an infinite dimensional linear...
system as follows:

\[ LS^\infty : \quad \dot{x}_t = Ax_t + Bu_t, \quad x_0 \in L^2[0,1], t \in [0,T] \]  \hspace{1cm} (3.7)

where \( A \in \tilde{G}^{sp}_1 \), \( B \in G_{A^2}^1 \), and are hence bounded operators on \( L^2[0,1] \), \( x_t \in L^2[0,1] \) is the system state at time \( t \) and \( u_t \in L^2[0,1] \) is the control input at time \( t \).

### 3.2 Graphon System Properties

#### 3.2.1 Uniqueness of the Solution

A solution \( x(\cdot) \in L^2(\mathbb{R}; L^2[0,1]) \) is a (mild) solution of (3.7) if

\[ x_t = e^{(t-a)A}x_a + \int_0^t e^{(t-s)A}Bu_s ds \]  \hspace{1cm} (3.8)

for all \( a \) and \( t \) in \( \mathbb{R} \) such that \( a \leq t \). Following [45] the assumptions on the operators \( A \) and \( B \) are

\[ (H1) \quad \left\{ \begin{array}{l}
(i) \quad A \text{ generates a strongly continuous semigroup } e^{tA} \text{ on } L^2[0,1], \\
(ii) \quad B \in L(L^2[0,1]),
\end{array} \right. \]

where the Hilbert space \( U \) (control space) in the present case is \( L^2[0,1] \). Under assumption \((H1)\), the system (3.7) has a unique solution \( x \in C([0,T]; L^2[0,1]) \) for any \( x_0 \in L^2[0,1] \) and any \( u \in L^2([0,T]; L^2[0,1]) \).

**Theorem 3.1.** The graphon system \( LS^\infty \) in (3.7) has a unique solution \( x \in C([0,T]; L^2[0,1]) \) for any \( x_0 \in L^2[0,1] \) and any \( u \in L^2([0,T]; L^2[0,1]) \).

**Proof.** Since \( A \) as a graphon operator generates a uniformly continuous semigroup, \( H1(i) \) is satisfied. Moreover \( B \in G_{A^2}^1 \) as a linear operator is bounded and hence a continuous linear mapping from control space \( L^2[0,1] \) to the state space \( L^2[0,1] \) satisfying \( H1(ii) \). Therefore the system (3.7) has a unique solution \( x \in C([0,T]; L^2[0,1]) \) for any \( x_0 \in L^2[0,1] \) and any \( u \in L^2([0,T]; L^2[0,1]) \).
3.2.2 Controllability

A system \((A; B)\) is exactly controllable on \([0, T]\) if for any initial state \(x_0 \in L^2[0, 1]\) and any target state \(x_f \in L^2[0, 1]\), there exists a control \(u \in L^2(0, T; U)\) driving the system from \(x_0\) to \(x_f\), i.e. \(x_T = x_f\) with \(x_T = e^{AT}x_0 + \int_0^T e^{A(T-t)}Bu_\tau dt\).

A system \((A; B)\) is approximately controllable on \([0, T]\) if for any initial state \(x_0 \in L^2[0, 1]\), any target state \(x_f \in L^2[0, 1]\) and any \(\varepsilon > 0\), there exists a control \(u \in L^2(0, T; U)\) driving the system from \(x_0\) to points in the state space within a \(\varepsilon\)-distance from \(x_f\), i.e. \(\|x_T - x_f\|_2 \leq \varepsilon\).

The controllability Gramian operator \(W_t : L^2[0, 1] \rightarrow L^2[0, 1]\) is defined as

\[
W_t := \int_0^t e^{A(t-s)}BB^Te^{A^T(t-s)}ds, \quad t > 0.
\]

**Theorem 3.2.** (see e.g. [45, 46]) A necessary and sufficient condition for exact controllability on \([0, T]\) is the uniform positive definiteness of \(W_T\), i.e.,

\[
(W_Th, h) \geq c_T\|h\|^2
\]

for all \(h \in L^2[0, 1]\), where \(c_T > 0\) and \(\cdot\|\) is the \(L^2[0, 1]\) norm.

The positive definiteness of the controllability Gramian operator \(W_T\) as a kernel is equivalent to the approximate controllability of the corresponding system (see [45, 46]).

**Theorem 3.3.** Let \(A\) be a graphon in \(\hat{G}_{1}^{\text{sp}}\) and let \(B\) be a bounded linear operator on \(L^2[0, 1]\). The linear system \((A; B)\) is exactly controllable on a finite time horizon \([0, T]\) if all the values in the spectrum of \(BB^T\) are lower bounded by a strictly positive constant.

**Proof.** Since \(A \in \hat{G}_{1}^{\text{sp}}\), \(A\) as an operator is self-joint. Let \(\lambda_\alpha\) be any eigenvalue of \(A\), which is necessarily real, and \(\varphi_\alpha \in L^2[0, 1]\) be any corresponding normalized eigenfunction of \(A\). Then

\[
\|A\|_{op} = \sup_{x \in L^2[0, 1], \|x\|_2 = 1} \|Ax\|_2 \geq \|A\varphi_\alpha\|_2 = |\lambda_\alpha|.
\]

(3.9)

For any \(A \in \hat{G}_{1}^{\text{sp}}\), there exists some finite \(c_1 > 0\), such that \(\|A\|_{op} \leq c_1\), i.e., \(A\) is a bounded operator on \(L^2[0, 1]\). Therefore \(c_1 \geq \lambda_\alpha \geq -c_1\) and hence for \(t > 0\)

\[
\|e^{At}\|_{op} \geq \|e^{A\tau}\varphi_\alpha\|_2 = \|e^{\lambda_\alpha t}\varphi_\alpha\|_2 = e^{\lambda_\alpha t} \geq e^{-c_1 t} > 0.
\]

(3.10)
Hence we see that $e^{At}$ as an $L^2[0, 1]$ operator is uniformly positive definite. Since all the values in the spectrum of $BB^T$ are lower bounded by a strictly positive constant, there exists $c > 0$ such that

$$\forall x \in L^2[0, 1], \quad (BB^T x, x) \geq c\|x\|_2^2.$$  \hspace{1cm} (3.11)

See e.g. [56, Theorem 12.12]. Consider the time horizon $[0, T]$. For any $h \in L^2[0, 1],$

$$\langle W_T h, h \rangle = \int_0^T (BB^T e^{At} t h, e^{At} T h) dt$$

$$\geq c \int_0^T \|e^{At} T h\|_2^2 dt$$

$$\geq cT (e^{-c_1 T})^2 \|h\|_2^2,$$

and hence the system $(A; B)$ is exactly controllable.

**Proposition 3.1.** Let $A$ and $B$ be graphons in $\tilde{G}_{1}^{sp}$. Then $(A; B)$ is not exactly controllable on any finite time horizon $[0, T]$.

**Proof.** Since $A$ and $B$ are graphons in $\tilde{G}_{1}^{sp}$, there exists $c_1 \geq 0$ and $c_2 \geq 0$, such that

$$c_1 \geq \|A\|_2 \geq \|A\|_{op} \quad \text{and} \quad c_2 \geq \|B\|_2 \geq \|B\|_{op}.$$  \hspace{1cm} (3.13)

See Lemma 2.1 for the relation between $L^2[0, 1]^2$ norm and the operator norm. Hence

$$\|e^{At}\|_{op} = \sup_{x \in L^2[0, 1], \|x\|_2 = 1} \|e^{At} x\|_2$$

$$\leq \sup_{x \in L^2[0, 1], \|x\|_2 = 1} \frac{1}{k!} \sum_{k=0}^{\infty} \|A^t\|_{op}^k \|x\|_2$$

$$= e^{\|A\|_{op} t} \leq e^{c_1 t}, \quad t \in [0, T].$$  \hspace{1cm} (3.14)

Therefore

$$\|W_T\|_2 \leq \int_0^T \|e^{At} BB^T e^{At} T\|_2 dt$$

$$\leq \int_0^T \|e^{At} B\|_2^2 dt$$

$$\leq \int_0^T (\|e^{At}\|_{op} \|B\|_2)^2 dt$$

$$\leq T (e^{c_1 T} c_2)^2 < \infty,$$  \hspace{1cm} (3.15)
which implies $W_T \in L^2[0, 1]^2$ and hence $W_T$ is a compact (and self-join) operator on $L^2[0, 1]$ functions (see e.g. [57, Chapter 2, Proposition 4.7]).

Based on the spectral property of a compact operator (see e.g. [63, Theorem 7.3 (Spectral theorem of F. Rellich)] which we also include as Theorem 2.9 for convenience), we obtain that $W_T$ has a countable number of nonzero (necessarily real) eigenvalues $\{\lambda_1, \lambda_2, \ldots\}$ such that $\lambda_n \to 0$ and each eigenvalue has finite multiplicity. Therefore $W_T$ is not uniformly positive definite and hence the system $(A; B)$ is not exactly controllable.

\[ \square \]

The results in Theorem 3.3 and Proposition 3.1 generalize to the case where $A$ lies in any uniformly bounded subset of $\tilde{G}^{sp}$ (that is, any set of symmetric measurable functions $W : [0, 1]^2 \to I$, where $I$ is a bounded interval in $\mathcal{R}$).
Chapter 4

Graphon State-to-state Control of Large-scale Networks

4.1 Graphon State-to-state Control Strategy (GSCS)

Consider the control problem of steering the states of each member of \( \{(A_N; B_N)\}_{N=1}^{\infty} \in \mathcal{S} \) to each of a sequence of desired states \( \{x^N_T \in \mathcal{R}^{n_N}\}_{N=1}^{\infty} \). The Graphon State-to-state Control (GSSC) Strategy consists of four steps:

**S.1** Let \( \{(A^N_s; B^N_s) \in \tilde{G}_{sp}^{1} \times G_{1}^{1} AI\}_{N=1}^{\infty} \) be the sequence of graphon dynamical systems equivalent to \( \{(A_N; B_N)\}_{N=1}^{\infty} \in \mathcal{S} \) under the mapping \( M_G \) defined in (3.3) and assume that it converges to the graphon system \((A; B) \in \tilde{G}_{sp}^{1} \times G_{1}^{1} AI\). Let \( \{x^N_T \in L^2[0, 1]\}_{N=1}^{\infty} \) be the image of \( \{x^N_T \in \mathcal{R}^{n_N}\}_{N=1}^{\infty} \) under \( M_G \) defined in (3.4), which is assumed to converge to some \( x^\infty_T \in L^2[0, 1] \) in the \( L^2[0, 1] \) norm.

**S.2** Specify the corresponding state-to-state control problem \( CP^\infty \) for \((A; B) \in \tilde{G}_{sp}^{1} \times G_{1}^{1} AI\) with \( x^\infty_T \) as the target terminal state and choose a tolerance \( \varepsilon > 0 \).

**S.3** Find a control law \( u^\infty := \{u_\tau \in L^2[0, 1], \tau \in [0, T]\} \) solving \( CP^\infty \).

**S.4** Then generate the control law \( \{u^N\} \) according to Theorems 4.2, 4.3, for which the convergence of \( \{x^N_T(u^N)\} \) to \( x^\infty_T \) is guaranteed. Together with the assumed convergence of \( \{x^N_T \in L^2[0, 1]\}_{N=1}^{\infty} \) to \( x^\infty_T \), it yields \( N_\varepsilon \) such that \( x^N_T(u^N) \) is within \( \varepsilon \) of \( x^N_T \) for all \( N \geq N_\varepsilon \) under the \( L^2[0, 1] \) norm.
4.2 Approximation Theorems

4.2.1 Approximation of $L^2[0,1]$ Input Functions via Piece-wise Constant Functions

**Theorem 4.1** ([64, p.198]). Let $\lambda$ be any measure on $\mathbb{R}$ and $\mathcal{B}_\lambda$ be the $\sigma$-algebra of $\lambda$-measurable sets, and let $1 \leq p < \infty$. Then piece-wise constant functions on $\mathbb{R}$ form a dense subset of $L^p(\mathbb{R}, \mathcal{B}_\lambda, \lambda)$.

Therefore piece-wise constant functions can approximate $L^2[0,1]$ functions arbitrarily well.

In our case, we want to approximate the control input $u_t(\cdot) \in L^2[0,1], 0 \leq t \leq T$, through a piece-wise constant function in $L^2[0,1]$ denoted by $u^N_t(\cdot)$. Specifically, the approximation of input $u_t(\cdot)$ by $u^N_t(\cdot)$ with the partition $Q = \{Q_1, Q_2, \cdots, Q_{nN}\}$ of $[0,1]$ is given as follows: for all $Q_i, i \in \{1, 2, \ldots, nN\}$,

$$u^N_t(\alpha) = \frac{1}{\mu(Q_i)} \int_{Q_i} u_t(\beta) d\beta, \quad \forall \alpha \in Q_i,$$

where $\mu(Q_i)$ denotes the measure of $Q_i$.

4.2.2 Limit Control for Network Systems

**Limit Control for Network Systems with Graphon Input Mappings**

Consider a finite dimensional system $(A_N; B_N)$ with node averaging dynamics as in (3.1) and $(A^N_s; B^N_s)$ as its equivalent step function system according to (3.3).

**Theorem 4.2.** Consider a sequence of network systems $\{(A^N_s; B^N_s)\}$ converging to a graphon system $(A; B)$ in the following sense: $A^N_s \to A$ in the $L^2$ metric and $B^N_s \to B$ in the $L^2$ metric as $N \to \infty$. Consider the problem of driving the systems from the origin to some target state. Then for any $T > 0$:

1) there exists a control $v^N$ for $(A^N_s; B^N_s)$ approximating the control $v$ for $(A; B)$ such that

$$\|x_T(v) - x_T^N(v^N)\|_2 \leq \|A^N_\Delta\|_2 \|B\|_2 \int_0^T e^{T-\tau} (T - \tau) \cdot \|v_\tau\|_2 d\tau + \|B^N_\Delta\|_2 \int_0^T e^{(T-\tau)}\|A^N_{s}\|_2 \cdot \|v_\tau\|_2 d\tau,$$

(4.2)
2) furthermore, for any $\varepsilon > 0$ there exists $N_\varepsilon > 0$ such that each $N \geq N_\varepsilon$,

$$\|x_T(v) - x_T^N(v[N])\|_2 < \varepsilon,$$  \hspace{1cm} (4.3)

where $A^N = A - A^N_s$, $B^N = B - B^N_s$, $x_T(v)$ represents the terminal state of $(A; B)$ under control $v$, $x_T^N(v[N])$ represents the terminal state of $(A^N_s; B^N_s)$ under control $v[N]$, and the control approximation is given in the following:

$$v_t^N(\alpha) = nN \int_{P_t} v_t(\beta)d\beta,$$  \hspace{1cm} (4.4)

for all $\alpha \in P_t$, $t \in [0, T]$, with the uniform partition $P_{mN} = \{P_1, \ldots, P_{nN}\}$.

**Proof.** Denote the terminal state of $(A; B)$ under a control $v \in L^2([0, T]; L^2[0, 1])$ by $G(v)$ and similarly the terminal state of $(A^N_s; B^N_s)$ under control $v \in L^2([0, T]; L^2[0, 1])$ by $G_N(v)$. Then

$$x_T(v) - x_T^N(v[N]) = G(v) - G_N(v[N]).$$  \hspace{1cm} (4.5)

We note that $(A^N_s; B^N_s)$ and $u[N]$ are defined via the same uniform partition $P = \{P_1, \ldots, P_{nN}\}$. Then following Lemma 4.1, one can show that for $\tau \in [0, T]$

$$B^N_s v^N_{\tau} = B^N_s v_{\tau},$$

$$A^N_s B^N_s v^N_{\tau} = A^N_s B^N_s v_{\tau}.$$  \hspace{1cm} (4.6)

For simplicity of notation, without causing confusion, we now set $A_s = A^N_s$ and $B_s = B^N_s$, then

$$G_N(v) = \int_0^T \left[ I + \sum_{k=1}^{\infty} \frac{1}{k!} (T - \tau)^k A^k_s \right] B_s v_{\tau} d\tau$$

$$= \int_0^T \left[ B_s v_{\tau} + \sum_{k=1}^{\infty} \frac{1}{k!} (T - \tau)^k A^k_s B_s v_{\tau} \right] d\tau$$

$$= \int_0^T \left[ B_s v_{\tau} - B_s v^N_{\tau} + B_s v^N_{\tau} + \sum_{k=1}^{\infty} \frac{1}{k!} (T - \tau)^k A^k_s B_s v^N_{\tau} \right] d\tau$$

$$= \int_0^T B_s [v_{\tau} - v^N_{\tau}] d\tau + \int_0^T \sum_{k=0}^{\infty} \frac{1}{k!} (T - \tau)^k A^k_s B_s v^N_{\tau} d\tau$$

$$= \int_0^T e^{A_s(T-\tau)} B_s v^N_{\tau} d\tau$$

$$= G_N(v[N]).$$
Therefore

$$\|G(v) - G^N(v^{[N]})\|_2 = \|G(v) - G^N(v)\|_2$$

$$= \left\| \int_0^T e^{A(T - \tau)} Bv_\tau d\tau - \int_0^T e^{A_s(T - \tau)} B_s v_\tau d\tau \right\|_2$$

$$\leq \left\| \int_0^T [e^{A(T - \tau)} - e^{A_s(T - \tau)}] Bv_\tau d\tau \right\|_2 + \left\| \int_0^T e^{A_s(T - \tau)} [B - B_s] v_\tau d\tau \right\|_2.$$  (4.8)

Denote $u_\tau = B_s v_\tau \in L^2[0, 1]$. A further simplification on the first part of the terminal state difference in (4.8) is as follows:

$$\left\| \int_0^T [e^{A(T - \tau)} - e^{A_s(T - \tau)}] Bv_\tau d\tau \right\|_2 \leq \int_0^T \sum_{k=0}^\infty \frac{(T - \tau)^k}{k!} \| (A^k - A_s^k) u_\tau \|_2 d\tau \quad \text{(Minkowski)}$$

$$= \int_0^T \sum_{k=1}^\infty \frac{(T - \tau)^k}{k!} \| (A^k - A_s^k) u_\tau \|_2 d\tau$$

$$\leq \int_0^T \sum_{k=1}^\infty \frac{(T - \tau)^k}{k!} k \cdot \| A_\Delta \|_2 \cdot \| u_\tau \|_2 d\tau \quad \text{(by Lemma 4.3)}$$

$$= \| A_\Delta \|_2 \int_0^T e^{T-\tau} (T - \tau) \| u_\tau \|_2 d\tau$$

$$\leq \| A_\Delta \|_2 \cdot \| B \|_2 \int_0^T e^{T-\tau} (T - \tau) \| v_\tau \|_2 d\tau,$$  (4.9)

where $A_\Delta = A - A_s^{[N]}$. The second part of the terminal state difference in (4.8) can be estimated as follows:

$$\left\| \int_0^T e^{A_s(T - \tau)} [B - B_s] v_\tau d\tau \right\|_2 = \left\| \int_0^T \sum_{k=0}^\infty A_s^k \frac{(T - \tau)^k}{k!} [B - B_s] v_\tau d\tau \right\|_2$$

$$\leq \int_0^T \sum_{k=0}^\infty \frac{(T - \tau)^k}{k!} \| A_s^k [B - B_s] v_\tau \|_2 d\tau \quad \text{(by Lemma 4.2 and Lemma 4.4)}$$

$$\leq \int_0^T \sum_{k=0}^\infty \frac{(T - \tau)^k}{k!} \| A_s \|_2 \cdot \| [B - B_s] v_\tau \|_2 d\tau$$

$$\leq \| B_\Delta \|_2 \cdot \int_0^T e^{\| A_s \|_2 (T - \tau)} \cdot \| v_\tau \|_2 d\tau$$  (4.10)
with $B_\Delta = B - B_s$. Finally, we obtain
\[
\|x_T(v) - x_T^N(v^{[N]}N)\|_2 \leq \|A_\Delta\|_2 \|B\|_2 \int_0^T e^{T-\tau}(T-\tau) \cdot \|v_\tau\|_{2d\tau} + \|B_\Delta\|_2 \int_0^T e^{(T-\tau)}|A^{[N]}_{1,2}| \cdot \|v_\tau\|_{2d\tau}. \tag{4.11}
\]

Then according to the $M_G$ mapping, the control law $\nu^N(\cdot)$ for the finite network system $(A_N; B_N)$ is given by
\[
\nu_t^N(i) = \nu_t^N(\alpha), \quad \forall i \in \{1, \ldots, nN\}, \forall \alpha \in P_i; t \in [0, T]. \tag{4.12}
\]

**Limit Control for Network Systems with the Identity Input Mapping**

In general, the control input mapping $B$ is not limited to be a graphon mapping. As long as the control input map is a linear continuous mapping from $L^2[0, 1]$ to $L^2[0, 1]$, the existence and uniqueness of a solution are guaranteed. The identity operator $I$ is a continuous mapping from $L^2[0, 1]$ to $L^2[0, 1]$ and hence the system $(A; I)$ has a unique solution. We note that while the identity operator $I$ may be represented by a positive measure on the diagonal in $[0, 1]^2$ and it may be treated as an element of $L^1[0, 1]^2$, it is not an element of $L^2[0, 1]^2$ and hence does not lie in $\hat{C}^{sp}_{1}$.

Consider a finite dimensional system $(A_N; I_N)$ with node averaging dynamics and $(A^{[N]}_s; I)$ as its equivalent step function system according to (3.3).

**Theorem 4.3.** Suppose $A^{[N]}_s \to A$ in the $L^2$ metric as $N \to \infty$. Consider the problem of driving the systems from the origin to some target state. Then for each $N, 1 \leq N < \infty$, there exists a control $u^{[N]}$ for $(A^{[N]}_s; I)$ approximating the control $u$ for $(A; I)$ such that
\[
\|x_T(u) - x_T^N(u^{[N]}N)\|_2 \leq \|A^N_\Delta\|_2 \int_0^T e^{T-\tau}(T-\tau) \cdot \|u_\tau\|_{2d\tau} + \int_0^T \|u_\tau - u^{[N]}_\tau\|_{2d\tau}, \tag{4.13}
\]

where $A^N_\Delta = A - A^{[N]}_s$, $x_T(u)$ represents the terminal state of $(A; I)$ under control $u$, $x_T^N(u^{[N]}N)$ represents the terminal state of $(A^{[N]}_s; I)$ under control $u^{[N]}$, and the control approximation is given by
\[
u_t^N(\alpha) = nN \int_{P_i} u_t(\beta) d\beta, \tag{4.14}
\]
for all $\alpha \in P_i$, with the uniform partition $P_{i\cdot}^n = \{P_1, \cdots, P_n\}$. Furthermore, for any $\varepsilon > 0$, there exists $N_\varepsilon > 0$ such that for each $N \geq N_\varepsilon$,

$$\|x_T(u) - x_T^N(u^{[N]})\|_2 < \varepsilon. \tag{4.15}$$

Proof. Denote the terminal state of $(A; B)$ under a control $v \in L^2([0, T]; L^2[0, 1])$ by $G(v)$ and similarly the terminal state of $(A^{[N]}_s; B^{[N]}_s)$ under control $v \in L^2([0, T]; L^2[0, 1])$ by $G^N(v)$. Then the difference between the two terminal states is given by $x_T(u) - x_T^N(u^{[N]}) = G(u) - G^N(u^{[N]})$. We note that $(A^{[N]}_s; B^{[N]}_s)$ and $u^{[N]}$ are defined via the same uniform partition $P = \{P_1, \cdots, P_n\}$. By Lemma 4.1, $A^{[N]}_s u^{[N]}_r = A^{[N]}_s u_r$.

For simplicity of notation, without causing confusion, we now set $A_s = A^{[N]}_s$, $u^N = u^{[N]}$ and $u^N_\tau = u^{[N]}_\tau$. Since the systems under consideration start with zero initial condition, we obtain

$$G^N(u) = \int_0^T e^{A_s(T-\tau)} u_\tau d\tau$$

$$= \int_0^T \left[1 + \sum_{k=1}^{\infty} \frac{1}{k!} (T-\tau)^k A^k_s\right] u_\tau d\tau$$

$$= \int_0^T \left[u_\tau + \sum_{k=1}^{\infty} \frac{1}{k!} (T-\tau)^k A^k_s u_\tau\right] d\tau$$

$$= \int_0^T [u_\tau - u^N_\tau] d\tau + \int_0^T \sum_{k=0}^{\infty} \frac{1}{k!} (T-\tau)^k A^k_s u^N_\tau d\tau$$

$$= \int_0^T [u_\tau - u^N_\tau] d\tau + \int_0^T e^{A_s(T-\tau)} u^N_\tau d\tau$$

$$= \int_0^T [u_\tau - u^N_\tau] d\tau + G^N(u^N). \tag{4.16}$$

Therefore

$$G(u) - G^N(u^N) = (G(u) - G^N(u)) + (G^N(u) - G^N(u^N))$$

$$= G(u) - G^N(u) + \int_0^T [u_\tau - u^N_\tau] d\tau. \tag{4.17}$$

By Lemma 4.3, we have

$$\| (A^k - A^k_s) u_\tau \|_2 \leq k \| A \|_2 \cdot \| u_\tau \|_2. \tag{4.18}$$
where $A_\Delta$ denotes $(A - A_s)$. Hence we have the following

$$
\|G(u) - G^N(u)\|_2 = \left\| \int_0^T \left[ e^{A(T-\tau)} - e^{A_s(T-\tau)} \right] u_\tau d\tau \right\|_2
$$

$$
= \left\| \int_0^T \sum_{k=0}^{\infty} \frac{(T-\tau)^k}{k!} (A^k - A_s^k) u_\tau d\tau \right\|_2
$$

$$
\leq \int_0^T \int_0^\infty \frac{(T-\tau)^k}{k!} \|A^k - A_s^k\|_2 \|u_\tau\|_2 d\tau
$$

(4.19)

$$
= \int_0^T \sum_{k=1}^{\infty} \frac{(T-\tau)^k}{k!} k \cdot \|A_\Delta\|_2 \cdot \|u_\tau\|_2 d\tau
$$

Finally, the $L^2$ difference in terminal states is bounded by

$$
\|x_T(u) - x_T^N(u^{[N]})\|_2 = \left\| G(u) - G^N(u) + \int_0^T [u_\tau - u_\tau^N] d\tau \right\|_2
$$

$$
\leq \|G(u) - G^N(u)\|_2 + \left\| \int_0^T [u_\tau - u_\tau^N] d\tau \right\|_2
$$

(4.20)

$$
= \|A_\Delta\|_2 \int_0^T e^{T-\tau} (T-\tau) \|u_\tau\|_2 d\tau + \left\| \int_0^T [u_\tau - u_\tau^N] d\tau \right\|_2.
$$

Then the control law $u^N(\cdot)$ according to the $M_G$ mapping for the finite network system $(A_N; I_N)$ is given by

$$
u^N_i(i) = u^{[N]}_i(\alpha), \quad \forall i \in \{1, ..., nN\}, \forall \alpha \in P_i, t \in [0, T].
$$

(4.21)

Note that $u^N$ always exists by definition since the control approximation given by (4.1) uses the same uniform partition as the step function approximation in the graphon space.

### 4.3 Minimum Energy State-to-state Control for Graphon Systems

A specific control law used in §2 of the GSCS is described in this section.
4.3.1 Minimum Energy Control of Infinite Dimensional Systems

Define the energy cost by the control over the time horizon \([0, T]\) as

\[
J(u) = \int_0^T \|u_\tau\|^2 d\tau.
\] (4.22)

The objective is to drive the system from some initial state \(x_0 \in L^2[0, 1]\) to some target state \(x_T \in L^2[0, 1]\) using minimum control energy. A function \(u^* \in L^2([0, T]; L^2[0, 1])\) is called an optimal control if \(J(u^*) \leq J(u)\), for all \(u \in L^2([0, T]; L^2[0, 1])\) that drive the system from \(x_0\) to \(x_T\).

4.3.2 Minimum Energy Control Law

Let \(X\) represent a Hilbert space. Let \(T\) and \(S\) be linear bounded operators on \(X\).

**Proposition 4.1.** [45] Assume that \(T\) and \(S\) are symmetric and nonnegative; then \(I + TS\) is one-to-one and onto. Moreover

\[
\|S(I + TS)^{-1}\| \leq \|S\|, \tag{4.23}
\]

and

\[
\|(I + TS)^{-1}\| \leq 1 + \|T\|\|S\|. \tag{4.24}
\]

Following this proposition, we prove the result on the existence of the inverse mapping of graphon controllability Gramian operator when the system is exactly controllable.

**Theorem 4.4.** If the graphon system \((A; B)\) with \(W_T\) as its graphon controllability Gramian operator is exactly controllable, then the inverse operator \(W_T^{-1}\) exists and is a bounded operator.

**Proof.** If the graphon system \((A; B)\) is exactly controllable, then

\[
\forall h \in L^2[0, 1], \quad \exists c_T > 0, \quad \langle W_T h, h \rangle \geq c_T \|h\|^2.
\]

Let \(I\) denote the identity operator from \(L^2[0, 1]\) to \(L^2[0, 1]\). Let \(M = W_T - \frac{1}{2} c_T I\), then \(W_T = \frac{1}{2} c_T (I + \frac{2}{c_T} M)\). By definition, the operator \(M\) is nonnegative and symmetric and hence \(\frac{2}{c_T} M\) is nonnegative and symmetric. By Proposition (4.1), \((I + \frac{2}{c_T} M)\) is one-to-one and onto, hence the inverse operator \((I + \frac{2}{c_T} M)^{-1}\) exists and is bounded. Therefore, by a scaling factor \(\frac{1}{2} c_T\) which is strictly positive and finite, \(W_T = \frac{1}{2} c_T (I + \frac{2}{c_T} M)\) is also one-to-one and onto, and the inverse operator \(W_T^{-1}\) exists and is bounded. \(\square\)
Assume the system \((A; B)\) is exactly controllable, then \(W_T^{-1}\) exists and the optimal control law that achieves the minimum energy control is given by

\[
u^*_\tau = B^T e^{A^T(T-\tau)} W_T^{-1} (x_T - e^{A(T)}x_0), \quad \tau \in [0, T].\] (4.25)

The minimum energy for controlling the system in time horizon \([0, T]\) is

\[
\|u\|_2^2 = [x_T - e^{A(T)}x_0]^T W_T^{-1} [x_T - e^{A(T)}x_0].
\] (4.26)

**Inverse of the Controllability Gramian Operator**

Since graphon \(A\) is a compact operator, it has a discrete spectrum.

**Proposition 4.2.** Assume the spectral decomposition of \(A\) is as follows

\[
A(x, y) = \sum_{l=1}^{\infty} \lambda_l f_l(x) f_l(y),
\]

where \(f_l\) is the normalized eigenfunction corresponding to the non-zero eigenvalues \(\lambda_l\) and assume \(B = I\). Then

1. the controllability Gramian operator is explicitly given by

\[
W_T = TI + \sum_{l=1}^{\infty} \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) f_l f_l^T;
\]

2. the inverse of the controllability Gramian operator for \((A; B)\) is explicitly given by

\[
W_T^{-1} = \frac{1}{T} I - \frac{1}{T} \sum_{l=1}^{\infty} \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T f_l f_l^T.
\]
Proof. The controllability Gramian is given by

\[
W_T = \int_0^T e^{At} e^{AT} dt = \int_0^T e^{2At} dt
\]

\[
= \int_0^T \left( I + \sum_{i=1}^\infty (2At)^i \frac{1}{i!} \right) dt
\]

\[
= \int_0^T I dt + \int_0^T \sum_{i=1}^\infty \sum_{l=1}^\infty (2\lambda_l t)^i f_l f_l^T dt
\]

\[
= \int_0^T I dt + \sum_{i=1}^\infty \int_0^T (e^{2\lambda_l t} - 1) f_l f_l^T dt
\]

\[
= TI + \sum_{l=1}^\infty \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) f_l f_l^T
\]

(4.27)

Suppose \( u = W_T x \). We should find the operator that maps \( u \) back to \( x \).

\[
u = W_T x = T x + \sum_{l=1}^\infty \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) f_l f_l^T x.
\]

Therefore

\[
x = \frac{1}{T} u - \frac{1}{T} \sum_{l=1}^\infty \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) f_l f_l^T x.
\]

By the definition of \( u \), we obtain in the direction of \( f_l \)

\[
f_l^T u = T f_l^T x + \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) f_l^T x = \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] \right) f_l^T x.
\]

Therefore

\[
x = \frac{1}{T} u - \frac{1}{T} \sum_{l=1}^\infty \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] \right) f_l f_l^T u.
\]

Equivalently, we obtain

\[
W_T^{-1} = \frac{1}{T} I - \frac{1}{T} \sum_{l=1}^\infty \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] \right) f_l f_l^T.
\]

Note that

\[
\lim_{\lambda_l \to 0} \left( \frac{1}{2\lambda_l} [e^{2\lambda_l T} - 1] - T \right) = 0.
\]
To achieve state-to-state control of linear system in infinite dimensional state space requires the system \((A; B)\) to be exactly controllable. Approximate controllability is not sufficient to achieve state-to-state control since the inverse operator of \(W_T\) might not be bounded in certain subspaces in \(L^2[0, 1]\) and then the energy required is unbounded.

### 4.4 Simulation Example

#### 4.4.1 Network Systems with Sampled Weightings

The generation of a randomly sampled network of size \(N\) from a graphon \(U\) is specified as follows:

1) Sample \(N\) points from a uniform distribution in \([0, 1]\). Sort the sample points in the decreasing order of their values and label them from node 1 to node \(N\). Denote the node set by \(V_N\) and the value of node \(i \in V_N\) by \(v_i\).

2) Connect the nodes \(i, j \in V_N\) with edge weight \(U(v_i, v_j)\) to generate the network \(G_N\). Then \(A_{Nij} = U(v_i, v_j)\) is the \(ij^{th}\) element of the adjacency matrix of \(G_N\).

Consider a network system evolving according to node averaging dynamics with \(G_N\) describing the dynamic interactions. Suppose each node has an independent input channel. Denote the system by \((A_N; I_N)\), where \(A_N\) is the adjacency matrix of \(G_N\) and \(I_N\) is the identity input mapping. The network system \((A_N; I_N)\) with node averaging dynamics is therefore described by

\[
\dot{x}_t^i = \frac{1}{N} \sum_{j=1}^{N} A_{Nij} x_t^j + u_t^i, \quad x_t^i, u_t^i \in \mathcal{R}, i \in \{1, ..., N\},
\]

(4.28)

where \(A_{Nij}\) is sampled from the graphon.

If \(U\) is almost everywhere continuous, then the step function \(A_{s}^{[N]}\) of \(A_N = [A_{Nij}]\) converges to \(U\) in the \(\delta_1\) metric as \(N \to \infty\) (see e.g. [65]), that is,

\[
\delta_1(A_{s}^{[N]}, U) \to 0 \quad \text{as} \quad N \to \infty.
\]

(4.29)

Further if we assume that \(U\) is uniformly bounded, then (4.29) implies \(\delta_2(A_{s}^{[N]}, U) \to 0\), as \(N \to \infty\). By the generation procedure, we obtain the labeling that achieves the minimum distance between the network and the limit, and hence a sequence of networks converges in the \(L^2\) metric.
4 Graphon State-to-state Control of Large-scale Networks

to the limit $U$. It follows that if $U$ is almost everywhere continuous and uniformly bounded, then we can apply the graphon control strategy to the sampled network systems.

4.4.2 Minimum Energy Graphon State-to-state Control

As an example, we consider the case where

$$U(x, y) = \cos(2\pi(x - y)) + 0.5 \cos(4\pi(x - y)), \ x, y \in [0, 1]$$

and solve the minimum energy control problem of driving the states of the network system $(A_N; I_N)$ to a Gaussian terminal state distribution $x_T^N$ from the origin over the time horizon $[0, T]$ with $T = 2$. Consider the target terminal state

$$x_T(\alpha) = \frac{1}{\sqrt{2\pi}} e^{-50(\alpha - 0.5)^2}, \ \alpha \in [0, 1]$$

The normalized eigenfunctions of the graphon operator $U$ are given by

$$f_1 = \sqrt{2} \cos 2\pi(\cdot), \ f_2 = \sqrt{2} \cos 4\pi(\cdot)$$

with corresponding eigenvalues $\lambda_1 = \frac{1}{2}$ and $\lambda_2 = \frac{1}{4}$. The inner products can be calculated as

$$\langle x_T, f_1 \rangle = -0.116088, \ \langle x_T, f_2 \rangle = 0.064211.$$

The systems $(U; I)$ is exactly controllable and the controllability Gramian operator is given by

$$W_T = \int_0^T e^{U(T-s)} e^{U(T-s)} ds = \int_0^T e^{2U(T-s)} ds. \ \ (4.30)$$

The minimum energy control for $(U; I)$ is given by

$$u_\tau^* = e^{U(T-\tau)} W_T^{-1} x_T, \ \ \tau \in [0, T], \ \ (4.31)$$

Since

$$e^{U(T-t)} = I + \sum_{l=1}^2 e^{\lambda_l(T-t)} f_l f_l^T \ \ (4.32)$$
and
\[
W_T^{-1} = \frac{1}{T} I - \frac{1}{T} \sum_{i=1}^{2} \frac{\left( \frac{1}{2\lambda_i} \left[ e^{2\lambda_i T} - 1 \right] - T \right) \mathbf{f}_i \mathbf{f}_i^T}{\left( \frac{1}{2\lambda_i} \left[ e^{2\lambda_i T} - 1 \right] \right) ^2}, \tag{4.33}
\]
the control law can be obtained as (4.34).
\[
u_t = \frac{1}{T} \mathbf{x}_T + \frac{1}{T} \sum_{i=1}^{2} \mathbf{f}_i(\mathbf{x}_T, \mathbf{f}_t) \left( -\frac{2\lambda_i \left( \frac{e^{\lambda_i T} - 1}{2\lambda_i} - T \right) (e^{\lambda_i(T-t)} - 1) + \left( e^{\lambda_i(T-t)} - 1 \right)}{e^{2\lambda_i T} - 1} \right),
\]
\[t \in [0, T], x \in [0, 1].\]
(4.34)

Replacing the parameters and set \(T = 2\), we obtain the control function for the graphon limit system (4.35).
\[
u_t = 0.199471 e^{-50(x-0.5)^2} - 0.0820869 \left( -\frac{(e^2 - 3) \left( \frac{e^{2\tau}}{2(e-1)} - 1 \right)}{e^2 - 1} + \left( \frac{e^{2\tau}}{2(e-1)} - 1 \right) \right) \cos(2\pi x)
+ 0.045404 \left( -\frac{2(e-1) - 2 \left( e^{2\tau} - 1 \right)}{2(e-1)} + \left( \frac{e^{2\tau}}{2(e-1)} - 1 \right) \right) \cos(4\pi x),
\]
\[t \in [0, 2], x \in [0, 1].\]
(4.35)

Further by approximating the control function (4.35) of the graphon limit system, we obtain the control law for any finite network system along the converging sequence to the graphon limit system. Then the control law \(u_N^*(i)\) for a network system \((A_N; I_N)\) generated by \(U\) comes from the following approximation:
\[
u_N^*(i) = N \int_{P_i} u^*_N(\beta) d\beta, \tau \in [0, T],
\tag{4.36}
\]
where \(P_i\) is the \(i^{th}\) element of the uniform partition \(P^N\) of \([0, 1]\). The error \(\| \mathbf{x}_T(u) - \mathbf{x}_T^N(u[N]) \|_2\) is bounded as in (4.13) and converges to 0 as \(N \to \infty\). The result of a simulation with a network system with 100 nodes using the proposed approximate control is shown in Figure 4.1.
4 Graphon State-to-state Control of Large-scale Networks

4.5 Appendix

**Lemma 4.1.** Consider a step function \( A_s^{[N]} \in \mathcal{G}_1^{sp} \) defined via partition \( P = \{ P_1, ..., P_{nN} \} \) and \( u_{\tau}^{[N]} \in L^2_{\text{pwc}}[0, 1] \) defined via the same partition \( P \) by

\[
    u_{\tau}^{[N]}(\alpha) = nN \int_{P_i} u_\tau(\beta) d\beta, \quad \forall \alpha \in P_i,
\]

where \( u_\tau \in L^2[0, 1] \). Then

\[
    A_s^{[N]} u_{\tau}^{[N]} = A_s^{[N]} u_{\tau},
\]

and

\[
    (A_s^{[N]})^k u_{\tau}^{[N]} = (A_s^{[N]})^k u_{\tau}, \quad k \geq 1.
\]

**Proof.** First, for all \( x \in P_i \),

\[
    [A_s^{[N]} u_{\tau}](x) = \int_0^1 A_s^{[N]}(x, y) u_\tau(y) dy = \sum_j \int_{P_j} A_{ij}^{[N]} u_\tau(y) dy,
\]
where $A_{ij}^{[N]} = A_{s}^{[N]}(x, y)$, for all $(x, y) \in (P_i, P_j)$; then

\[
[A_{s}^{[N]}u_{\tau}](x) = \sum_j \int_{P_j} A_{ij}^{[N]}u_{\tau}(y)dy, \quad \forall x \in P_i,
\]

\[
= \sum_j A_{ij}^{[N]} \int_{P_j} u_{\tau}(y)dy
\]

\[
= \sum_j A_{ij}^{[N]} \cdot \mu(P_j) \cdot u_{\tau}^{[N]}(x)
\]

(by the definition of $u_{\tau}^{[N]}$),

\[
[A_{s}^{[N]}u_{\tau}^{[N]}](x) = \sum_j \int_{P_j} A_{ij}^{[N]}u_{\tau}^{[N]}(y)dy, \quad \forall x \in P_i
\]

\[
= \sum_j A_{ij}^{[N]} \int_{P_j} u_{\tau}^{[N]}(y)dy
\]

\[
= \sum_j A_{ij}^{[N]} \cdot \mu(P_j) \cdot u_{\tau}^{[N]}(x).
\]

Finally, (4.40) and (4.41) give the equality in (4.37). An immediate implication of (4.37) is that for $k \geq 1$,

\[
(A_{s}^{[N]})^{k}u_{\tau}^{[N]} = (A_{s}^{[N]})^{k}u_{\tau}. \tag{4.42}
\]

\[\square\]

**Lemma 4.2.** Consider $A, A_{s} \in \tilde{G}_{s}^{sp}$ and $u_{\tau} \in L^2[0, 1]$. Then the following inequality holds:

\[
\|Au_{\tau} - A_{s}u_{\tau}\|_2^2 \leq \|A_{\Delta}\|_2^2 \|u_{\tau}\|_2^2, \tag{4.43}
\]

where $A_{\Delta}$ denotes $(A - A_{s})$.

**Proof.** An application of the Cauchy-Schwartz Inequality gives

\[
|\langle A_{\Delta}(x, \cdot), u_{\tau}(\cdot) \rangle|^2 = \int_0^1 A_{\Delta}(x, y)u_{\tau}(y)dy \leq \int_0^1 A_{\Delta}^2(x, y)dy \cdot \int_0^1 u_{\tau}^2(y)dy,
\]

where $A_{\Delta} = A - A_{s}$. 

\[
(A_{s}^{[N]})^{k}u_{\tau}^{[N]} = (A_{s}^{[N]})^{k}u_{\tau}.
\]
where $A_\Delta$ denotes $(A - A_s)$. Hence

\[
\|Au_r - A_s u_r\|_2^2 = \|(A - A_s)u_r\|_2^2 = \int_0^1 \bigg| \int_0^1 [A(x,y) - A_s(x,y)] u_r(y) dy \bigg|^2 dx
\]

\[
\leq \int_0^1 \left( \int_0^1 A_\Delta^2(x,y) dy \int_0^1 u_r^2(y) dy \right) dx
\]

\[
\leq \left( \int_0^1 \int_0^1 A_\Delta^2(x,y) dy dx \right) \int_0^1 u_r^2(y) dy
\]

\[
= \|A_\Delta\|_2^2 \|u_r\|_2^2.
\]  

(4.44)

**Lemma 4.3.** Consider $A, A_s \in \tilde{G}_{1}^{sp}$ and $u_r \in L^2[0,1]$. Then for $k \geq 1$

\[
\|(A^k - A_s^k)u_r\|_2 \leq k \|A_\Delta\|_2 \|u_r\|_2
\]

(4.45)

where $A_\Delta$ denotes $(A - A_s)$.

**Proof.** Define $P_k(x,y) = \sum_{i=0}^k x^{k-i} y^i$. Then

\[
x^k - y^k = (x - y) P_{k-1}(x,y),
\]

(4.46)

and we see that for $k \geq 1$

\[
A^k - A_s^k = P_{k-1}(A, A_s)(A - A_s).
\]

(4.47)

Since $A^{(k-i-1)} A_s^i \in G_{1}^{sp}$, for all $i \in \{0, 1, ... k - 1 \}$, we know that $\|A^{(k-i-1)} A_s^i\|_2 \leq 1$. Therefore, by the Minkowski inequality,

\[
\|P_{k-1}(A, A_s)\|_2 \leq \sum_{i=0}^{k-1} \|A^{(k-i-1)} A_s^i\|_2 \leq k \cdot 1.
\]

(4.48)

Together with Lemma 4.2, it yields

\[
\|(A^k - A_s^k)u_r\|_2 \leq \|P_{k-1}(A, A_s)\|_2 \cdot \|A_\Delta\|_2 \cdot \|u_r\|_2
\]

\[
\leq k \|A_\Delta\|_2 \|u_r\|_2
\]

(4.49)

where $A_\Delta$ denotes $(A - A_s)$.

\[\square\]
Lemma 4.4. Consider $A_s \in \tilde{G}_{1}^{sp}$. Then for $k \geq 1$

$$\|A_s^k\|_2 \leq \|A_s\|_2^k. \quad (4.50)$$

Proof. Let $V$ and $W$ be any two graphons in $\tilde{G}_{1}^{sp}$. Then for $k \geq 1$

$$\|VW\|_2^2 = \int_0^1 \int_0^1 [(VW)(x,y)]^2 dxdy$$

$$= \int_0^1 \int_0^1 \left[ \int_0^1 V(x,\beta)W(\beta,y) d\beta \right]^2 dxdy$$

$$\leq \int_0^1 \int_0^1 \left( \int_0^1 V(x,\beta)^2 d\beta \right) \cdot \int_0^1 \left( W(\alpha, y)^2 d\alpha \right) dxdy \quad (4.51)$$

$$= \int_0^1 \int_0^1 (V(x,\beta))^2 dxd\beta \cdot \int_0^1 \int_0^1 (W(\alpha, y))^2 d\alpha dxdy$$

$$= \|V\|_2^2 \|W\|_2^2.$$

Applying (4.51) to $A_s^k$ multiple times we have $\|A_s^k\|_2 \leq \|A_s\|_2^k$. \hfill \Box$

Results in Lemma 4.1, Lemma 4.2 and Lemma 4.4 generalize to functions in any uniformly bounded subset of $\tilde{G}^{sp}$. 
Chapter 5

Graphon Linear Quadratic Regulation (LQR) of Network Systems

5.1 Graphon-Network LQR (GLQR) Strategy

Consider the control problem of regulating the states of each member of \( \{(A_N; B_N)\}_{N=1}^{\infty} \in S \).

The \textbf{Graphon-Network LQR (GLQR) Strategy} is as follows:

S.1 \textit{Let} \( \{(A_s^{[N]}, B_s^{[N]}) \in \tilde{G}^{sp}_1 \times \mathcal{G}_{AI}^1 \}_{N=1}^{\infty} \) \textit{be the sequence of step function systems equivalent to} \( \{(A_N; B_N)\}_{N=1}^{\infty} \in S \) \textit{under the mapping} \( M_G \) \textit{defined in (3.3) and assume that it converges to the graphon system} \( (A; B) \in \tilde{G}^{sp}_1 \times \mathcal{G}_{AI}^1 \).

S.2 \textit{Define the linear quadratic cost for} \( (A; B) \) \textit{as}

\[
J(u) = \int_0^T [\|Cx_t\|^2 + \|u_t\|^2]dt + (P_0x_T, x_T)
\]

\textit{and the linear quadratic cost for} \( (A_s^{[N]}, B_s^{[N]}) \) \textit{as}

\[
J(u^{[N]}) = \int_0^T [\|C_s^{[N]}x_t^{[N]}\|^2 + \|u_t^{[N]}\|^2]dt + (P_{s0}x_T^{[N]}, x_T^{[N]})
\]

\textit{where it is assumed that} \( C_s^{[N]} \rightarrow C \) \textit{and} \( P_{s0}^{[N]} \rightarrow P \) \textit{in the strong operator sense. Solve the infinite dimensional Riccati equation for} \( (A; B) \) \textit{to generate the solution} \( P \).

S.3 \textit{Approximate} \( P \) \textit{to generate} \( \tilde{P}_N \) \textit{and hence the control law} \( u_t^{[N]} = -B_s^{[N]^T}\tilde{P}_N(T-t)x_t^{[N]} \) \textit{for} \( (A_s^{[N]}, B_s^{[N]}) \).
In this control strategy, the basic assumption in the formulation of LQR problems for linear systems distributed on complex networks is that the regulation problem for the infinite dimensional graphon limit systems can be solved (e.g. by established approximation methods) while the finite dimensional LQR problems for the original complex network systems are intractable due to their cardinality.

### 5.2 LQR Problems for Graphon Dynamical Systems

Let $\| \cdot \|$ and $\langle \cdot , \cdot \rangle$ denote the norm and the inner product in $L^2[0,1]$. For finite $T > 0$, consider the problem of minimizing the cost given by

$$J(u) = \int_0^T \left[ \| Cx_\tau \|^2 + \| u_\tau \|^2 \right] d\tau + \langle P_0 x_T, x_T \rangle \quad (5.1)$$

over all controls $u \in L^2([0,T]; L^2[0,1])$ subject to the system model constrains in (3.7). The assumptions for $C$ and $P_0$ are:

$$(H2) \quad \begin{cases} 
(iii) & P_0 \in \mathcal{L}(L^2[0,1]) \text{ is Hermitian and non-negative,} \\
(iv) & C \in \mathcal{L}(L^2[0,1]; Y) 
\end{cases}$$

where $Y$ is the Hilbert space of observations, which in the current case is $L^2[0,1]$.

Finding the feedback control via dynamic programming consists of the two following steps:

**Step 1.** Solve the Riccati equation

$$\dot{P} = A^T P + PA - PBB^T P + C^T C, \quad P(0) = P_0 \quad (5.2)$$

**Step 2.** Given the solution $P$ to the Riccati equation, it can be proved that the optimal control $u^*$ is given by

$$u^*_t = -B^T P(T - t)x^*_t, \quad t \in [0,T] \quad (5.3)$$

and moreover that $x^*$ is the solution to the closed loop equation

$$\dot{x}_t = Ax_t - BB^T P(T - t)x_t, \quad t \in [0,T], x_0 \in L^2[0,1]. \quad (5.4)$$
5 Graphon Linear Quadratic Regulation (LQR) of Network Systems

Basic Notation

Let
\[ \Sigma(L^2[0,1]) = \{ T \in \mathcal{L}(L^2[0,1]) : T \text{ is Hermitian} \} \] (5.5)
and
\[ \Sigma^+(L^2[0,1]) = \{ T \in \Sigma(L^2[0,1]) : \langle Tx, x \rangle \geq 0, \forall x \in L^2[0,1] \} . \] (5.6)

Denote the topological space of all strongly continuous mappings \( F : I \rightarrow \Sigma(L^2[0,1]) \) (i.e., \( F(\cdot)x \) is continuous in \( I \) for all \( x \in L^2[0,1] \)) endowed with strong convergence by \( C_s(I; \Sigma(L^2[0,1])) \) (see [45]), where \( I \) represents a compact time interval. The strong convergence of a sequence of strongly continuous mappings \( \{F_n\} \) to \( F \) is defined as
\[ \forall x \in L^2[0,1], \lim_{n \to \infty} \sup_{t \in I} \|F_n(t)x - F(t)x\| = 0. \] (5.7)

Existence and Uniqueness of Solutions

Applying the results in [45] and specializing the Hilbert space there to be \( L^2[0,1] \) space, one can show, under the assumption \((H1)\) and \((H2)\), the existence and uniqueness of the solution to the Riccati equation (5.2) and the existence and uniqueness of optimal solution pair \((u^*, x^*)\) in (5.3) and (5.4).

Following [45], if \( A \in \mathcal{L}(L^2[0,1]) \), then problem (5.2) is equivalent to the following integral form
\[ P(t)x = e^{tA^T}P_0e^{tA}x + \int_0^t e^{sA^T}C^TCe^{sA}xds \\
- \int_0^t \int_0^s e^{sA^T}P(s)BB^TP(s)e^{sA}xds, \quad x \in L^2[0,1]. \] (5.8)

A mild solution to the problem (5.2) in the time horizon \([0,T]\) is defined as a function \( P \in C_s([0,T]; \Sigma+(L^2[0,1])) \) that satisfies the integral equation (5.8).

**Theorem 5.1.** [45, p.393] Assume that \((H1)\) and \((H2)\) are verified. Then problem (5.2) has a unique (mild) solution \( P \in C_s([0,T]; \Sigma+(L^2[0,1])) \).

**Theorem 5.2.** [45, p.409] Assume that \((H1)\) and \((H2)\) are verified, and let \( x_0 \in L^2[0,1] \). Then there exists a unique optimal pair \((u^*, x^*)\). Moreover, the following statements hold:

1. \( x^* \in C([0,T]; L^2[0,1]) \) is the (mild) solution to the closed loop equation (5.4),
5 Graphon Linear Quadratic Regulation (LQR) of Network Systems

2. \( u^* \in C([0, T]; L^2[0, 1]) \) is given by the feedback formula

\[
u^*(t) = -B^*P(T - t)x^*(t), \quad t \in [0, T], \tag{5.9}\]

3. The optimal cost \( J(u^*) \) is given by

\[
J(u^*) = \langle P(T)x_0, x_0 \rangle. \tag{5.10}\]

5.3 Approximation Theorems

By approximating the Riccati equation solution \( P \) for \((A; B)\) we can generate \( \tilde{P}_N \) that provides the control law for the finite dimensional network system:

\[
u_{[N]}(t) = -B_{[N]}^{T}\tilde{P}_N(T - t)x_{[N]}(t), \tag{5.12}\]

5.3.1 Approximation of the Riccati Equation Solution and Its Convergence to the Optimal Riccati Equation Solution

We need to extend the step function approximation to step function approximation by local integration against measures.

First, we construct the equivalent representation of the linear operator \( P \) in \( C_s([0, T]; \Sigma^*(L^2[0, 1])) \) by integration against measures, that is, we first represent \( P \) by

\[
P(\cdot)(x, y)d\sigma(x, y), \quad (x, y) \in [0, 1]^2, \tag{5.11}\]

where \( d\sigma(x, y) := \sigma(dx, dy) \) and \( \sigma(x, y) \) represents the measure (which can be a singular measure, a Lebesgue measure or a mixed measure).

Second, we introduce a method to approximate the operator \( P \) by local integration with respect to measures over partitions. The local step function approximation against measures of \( P \) is performed by integration against measures as follows:

\[
\tilde{P}_N(\cdot)(x, y) = \frac{\int_{S_i \times S_j} P(\cdot)(x, y)d\sigma(x, y)}{\mu(S_i) \times \mu(S_j)}, \forall (x, y) \in S_i \times S_j, \tag{5.12}\]

where \( S_i, S_j \subset [0, 1], \mu(S_i) \) represents the length of the interval \( S_i \) and \( \sigma(x, y) \) represents the
measure (which can be a singular measure, a Lebesgue measure or a mixed measure).

Since $\tilde{P}_N x$ is the step function approximation of $Px$ in $L^2[0,1]$ under the interpretation of integration against measures, for any $x \in L^2[0,1],$

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \| \tilde{P}_N(t)x - P(t)x \|_2 = 0. \quad (5.13)$$

Therefore we obtain the following lemma.

**Lemma 5.1.** Let $\tilde{P}_N$ be generated by step function approximation against measures from $P$ via $N \times N$ uniform partition of $[0,1]^2$. Then

$$\lim_{N \to \infty} \tilde{P}_N = P, \quad \text{in } C([0,T]; \Sigma(L^2[0,1])).$$

**Theorem 5.3.** Let $\tilde{P}_N$ be generated by step function approximation against measures from $P$ via $N \times N$ uniform partition of $[0,1]^2$. For any $x \in L^2[0,1],$ for any $t \in [0,T],$

$$\lim_{N \to \infty} \| \tilde{P}_N(t)x - P_s^{[N]}(t)x \|_2 = 0,$$

where $P_s^{[N]}$ is the solution to Riccati equation of $(A_s^{[N]}, B_s^{[N]})$ that converges strongly to the solution $P.$

**Proof.** Since

$$\lim_{N \to \infty} \tilde{P}_N = P, \quad \text{in } C([0,T]; \Sigma(L^2[0,1]))$$

and

$$\lim_{N \to \infty} P_s^{[N]} = P, \quad \text{in } C([0,T]; \Sigma(L^2[0,1])),$$

we obtain the following for any $x \in L^2[0,1]$ and for any $t \in [0,T]:$

$$\lim_{N \to \infty} \| \tilde{P}_N(t)x - P(t)x \|_2 = 0$$

and

$$\lim_{N \to \infty} \| P(t)x - P_s^{[N]}(t)x \|_2 = 0.$$

Since

$$\| \tilde{P}_N(t)x - P_s^{[N]}(t)x \|_2 \leq \| \tilde{P}_N(t)x - P(t)x \|_2 + \| P(t)x - P_s^{[N]}(t)x \|_2,$$
we obtain
\[
\lim_{N \to \infty} \parallel \tilde{P}_N(t)x - P_s^{[N]}(t)x \parallel_2 = 0.
\]

5.3.2 Continuous Dependence of Riccati Equation Solution with Respect to the Data

Let \( \text{Ricc}(A, B, C, P_0) \) denote the following Riccati equation
\[
\dot{P} = A^TP + PA - PBB^TP + C^TC,
\]
\[
P(0) = P_0.
\]
(5.14)

Assumption 5.1.

1. For any \( N \geq 1 \), \((A_s^{[N]}, B_s^{[N]}, C_s^{[N]}, P_s^{[N]}_0)\) satisfies \((H1)\) and \((H2)\);
2. The system sequence \((A_s^{[N]}, B_s^{[N]})\) converges to \((A; B)\) in \( L^2 \) metric;
3. The sequences \( \{C_s^{[N]}\} \) and \( \{P_s^{[N]}_0\} \) converges strongly to \( C \) and \( P_0 \), respectively, as \( N \to \infty \);
4. \( C \) and \( C_s^{[N]} \) are self-joint linear operators.

Theorem 5.4. Consider a sequence of network systems \((A_N; B_N)\) \( \{N \geq 1\} \) with \((A_s^{[N]}, B_s^{[N]}) \in \tilde{G}_1^{sp} \times \tilde{G}_1^{sp} \) \( \{N = 1\} \) as the step function representation. Let \( P \) and \( P_s^{[N]} \) be the solution to \( \text{Ricc}(A, B, C, P_0) \) and \( \text{Ricc}(A_s^{[N]}, B_s^{[N]}, C_s^{[N]}, P_s^{[N]}_0) \) respectively. If Assumption 5.1 holds, then for any finite horizon \([0, T]\),
\[
\lim_{N \to \infty} P_s^{[N]} = P, \quad \text{in } C_s([0, T]; \Sigma(L^2[0, 1])).
\]

Proof. From Theorem 2.6, we know for all \( T > 0 \) and all \( x \in L^2[0, 1] \),
\[
\lim_{N \to \infty} e^{tA_s^{[N]}} x = e^{tA} x
\]
uniformly in \([0, T]\). Since the system sequence \((A_s^{[N]}, B_s^{[N]})\) converges to \((A; B)\) as in Definition 3.1, \( \{B_s^{[N]}\} \) converges to \( B \) in the strong operator sense. We may now apply Theorem 2.2, Part IV, [45], specialized to the Hilbert space \( L^2[0, 1] \). Since its hypotheses are then satisfied in the present case, the desired result follows.
\[\square\]
5.3.3 Convergence of States and Convergence of Costs

Let \( P_s^{[N]} \) denote the solution to the Riccati equation for \((A_s^{[N]}; B_s^{[N]})\) that converges strongly to the solution \( P \) of the Riccati equation for \((A; B)\). Let \( \tilde{P}_N \) be the step function approximation against measures for \( P \) generated via the \( N \times N \) uniform partition of \([0,1]^2\).

**Theorem 5.5.** Consider the time horizon \([0,T]\). Let the optimal linear quadratic control law for \((A_s^{[N]}; B_s^{[N]})\) be generated by

\[
u_t^{N^*} = -B_s^{[N]^T} P_s^{[N]} (T-t) x_t^{N^*},
\]

where the optimal state trajectory is given by \( x_t^{N^*} \), and let the graphon approximate control law for \((A_s^{[N]}; B_s^{[N]})\) be

\[
u_t^{[N]} = -B_s^{[N]^T} \tilde{P}_N (T-t) x_t^{[N]},
\]

where the corresponding state trajectory is given by \( x_t^{[N]} \). Then

\[
\forall t \in [0,T], \quad \lim_{N \to \infty} \|x_t^{N^*} - x_t^{[N]}\|_2 = 0,
\]

and

\[
\lim_{N \to \infty} |J(u^{N^*}) - J(u^{[N]})| = 0.
\]

**Proof.** The closed loop system with the optimal control law is given by

\[
\dot{x}_t^{*} = \left( A_s^{[N]} - B_s^{[N]} B_s^{[N]^T} P_s^{[N]} (T-t) \right) x_t^{*}, \quad t \in [0,T], x_0 \in L^2[0,1].
\]

(5.15)

The closed loop system under the graphon approximate control law is given by

\[
\dot{x}_t^{[N]} = \left( A_s^{[N]} - B_s^{[N]} B_s^{[N]^T} \tilde{P}_N (T-t) \right) x_t^{[N]}, \quad t \in [0,T], x_0 \in L^2[0,1].
\]

(5.16)

Let \( x_t^e = x_t^{*} - x_t^{[N]} \). By (5.15) and (5.16), we obtain

\[
\dot{x}_t^e = F_N(t) x_t^e + \nabla_N(t),
\]

(5.17)

where

\[
F_N(t) = \left( A_s^{[N]} - B_s^{[N]} B_s^{[N]^T} P_s^{[N]} (T-t) \right)
\]
and
\[ \mathbb{V}_N(t) = B_s^{[N]}B_s^{[N]^T}(P_s^{[N]}(T-t) - \tilde{P}_N(T-t))x_t^\ast. \]

The integral representation of (5.17) is given by
\[
x_t^\ast e_t = \int_0^t \mathbb{V}_N(s)ds + \int_0^t F_N(\tau)x_t^\ast d\tau.
\]

Hence we obtain
\[
\|x_t^\ast\| \leq \psi_N(t) + \int_0^t \chi_N(\tau)\|x_t^\ast\|d\tau.
\]

where \(\psi_N(t) = \|\int_0^t \mathbb{V}_N(s)ds\|_2\) and \(\chi_N(t) = \|F_N(t)\|_{op}, t \in [0,T]\). Applying the Gronwall-Bellman inequality [66, p.7], we obtain
\[
\|x_t^\ast\|_2 \leq \psi_N(t) + \int_0^t \psi_N(s)\chi_N(s)e^{\int_s^t \chi_N(u)du}ds.
\]

On one hand, we note that for any \(t \in [0,T]\), \(F_N(t)\) is a bounded linear operator, i.e., \(\chi_N(t)\) is bounded and hence there exist \(C > 0\)
\[
\chi_N(s)e^{\int_s^t \chi_N(u)du} \leq C, \quad \forall t, s \in [0,T], t \geq s.
\]

On the other hand, based on the fact that \(B_s^{[N]}\) is a bounded operator, \(x_t^\ast \in L^2[0,1]\) and the result in Theorem 5.3, we obtain, for any \(t \in [0,T]\),
\[
\lim_{N \to \infty} \|B_s^{[N]}B_s^{[N]^T}(P_s^{[N]}(T-t) - \tilde{P}_N(T-t))x_t^\ast\|_2 = 0,
\]

that is,
\[
\forall t \in [0,T], \quad \lim_{N \to \infty} \psi_N(t) = 0.
\]

Therefore
\[
\|x_t^\ast\|_2 \leq \psi_N(t) + C \int_0^t \psi_N(s)ds,
\]

and
\[
\lim_{N \to \infty} \|x_t^\ast - x_t^{[N]}\|_2 = 0, \quad t \in [0,T].
\]

Next we prove the convergence in the cost function. By definition,
\[
J(u_t^{N*}) = \int_0^T [\|Cx_t^\ast\|^2 + \|u_t^{N*}\|^2]d\tau
\]
and
\[ J(u^{[N]}) = \int_0^T \left[ \| Cx^*_t \|^2 + \| u^*_t \|^2 \right] dt. \]

Hence
\[
J(u^{N*}) - J(u^{[N]})
= \int_0^T \left[ \| Cx^*_t \|^2 - \| Cx^*_t \|^2 + \| u^*_t \|^2 - \| u^*_t \|^2 \right] dt
= \int_0^T \left[ (Cx^*_t, x^*_t) - (Cx^*_t, x^*_t) + (u^*_t, u^*_t) - (u^*_t, u^*_t) \right] dt
= \int_0^T \left[ (C(x^*_t + x^*_t), x^*_t) + (u^*_t + u^*_t, u^*_t - u^*_t) \right] dt
\leq \int_0^T (\| C(x^*_t + x^*_t) \| \cdot \| x^*_t - x^*_t \| + \| u^*_t + u^*_t \| \cdot \| u^*_t - u^*_t \|) dt
\]

Since for all \( t \in [0, T] \)
\[ \lim_{N \to \infty} \| x^*_t - x^*_t \|_2 = 0, \]
together with the result in Theorem 5.3, we obtain
\[ \lim_{N \to \infty} \| u^*_t - u^*_t \|_2 = 0. \]

Considering the fact that \( Cx^*_t, Cx^*_t, u^*_t, \) and \( u^*_t \) are \( L^2[0, 1] \) functions for all \( t \in [0, T] \), we have
\[ \lim_{N \to \infty} |J(u^{N*}) - J(u^{[N]})| = 0. \] (5.24)

### 5.4 Simulation Example

Consider a network system evolving according to the node averaging dynamics with weighted graph \( G_N \) describing the dynamic interactions. Suppose each node has an independent input channel. Denote the system by \((A_N; I_N)\), where \( A_N \) is the adjacency matrix of \( G_N \) and \( I_N \) is the identity input mapping. The network system \((A_N; I_N)\) with (normalized) node dynamics is
therefore described by

$$\dot{x}_i^t = \frac{1}{N} \sum_{j=1}^{N} A_{Nij} x_j^t + u_i^t, \quad x_i^t, u_i^t \in \mathcal{R}, \ i \in \{1, \ldots, N\}. \quad (5.25)$$

where $A_{Nij}$

The objective is to regulate the network states around origin from random initial states with minimum quadratic cost.

As an example, we consider a sequence of networks converging to the graphon limit $U(x, y) = 4 \cos(2\pi(x - y))$ for all $x, y \in [0, 1]$ as in figure (h) and solve the LQR problem over the time horizon $[0, T]$ for the network sequence. See Section 4.4.1 for the detailed description of the generation of a convergent network sequence.

In this simulation, as shown in Figure 5.2, a network of size 320 along the sequence is considered. The system is represented by $(A_{320}, I_{320})$ with $A_{320}$ as the adjacency matrix of the weighted network and $I_{320}$ as the identity input matrix of size 320. $B = I_{320}, C = \sqrt{2}I_{320}, P_0 = I_{320}$. The infinite dimensional limit Riccati equation can be solved with the solution given by $P_t = \alpha_t I + \beta_t U$ where $\alpha_t$ and $\beta_t$ satisfy

$$\dot{\alpha}_t = 2 - \alpha_t^2, \quad \dot{\beta}_t = 2\alpha_t + 16\beta_t - 2\alpha_t\beta_t - 8\beta_t^2 \quad (5.26)$$

with $\alpha_0 = 1, \beta_0 = 0$. The finite dimensional control law is generated by approximating the Riccati equation solution as in (5.12). As the networks increase in size and converge to the limit graphon, the strong convergence of the approximated graphon Riccati equation solution to the finite dimensional Riccati equation solution is guaranteed by Theorem 5.3. Furthermore, the convergence in the state trajectory (and cost) to the optimal state trajectory (and the optimal cost) is guaranteed by Theorem 5.5.

Both the graphon-LQR control and the LQR optimal control regulate the system from the
same random initial states to the origin as shown in figures (a) and (b). From figures (e) and (f), we see that the graphon-LQR control achieves remarkably similar performance to the LQR optimal control. The maximum trajectory difference from the optimal control is less than 4% of the maximum initial states. With the graph interpreted as a $L^2[0, 1]^2$ function, the distance between the graph and the graphon limit in $L^2[0, 1]^2$ is 0.000813. The Graphon-LQR control cost is only 0.133% higher than the optimal LQR control cost.
Part II

Graphon Team Optimal Control
Chapter 6

Team Optimal Graphon-LQR Control

6.1 Introduction

One approach to achieve scalable control design for complex networks of agents (i.e. dynamical systems each of which has an independent control) is to design decentralized control strategies where the control is calculated locally at each agent and applied locally. In team optimal control solutions in [50, 51], each agent computes its control actions based on the observation of its own state and certain global information, which we note is closely analogous to the decentralized strategies appearing in Mean Field Game theory [67, 68, 69]. The global information in the team strategies appearing in this and the subsequent chapter is related to the spectral properties of the network and its initial states. Moreover, the team optimal solutions in this chapter permit decentralized control on networks with an infinite number of nodal agents. Specifically, here we study graphon LQR control problems with non-compact system operators and establish decentralized strategies via solutions to the corresponding team optimal graphon control problems. We prove that the team optimal control solutions are equivalent to the corresponding centralized optimal control solutions in the sense that they achieve the same performance. The advantage of team optimal control solutions is that they can be locally computed and implemented by the agents.

We assume the graphon can be exactly characterized by a finite spectral sum, although the previous results presented in approximation theorems [70] still hold for team optimal solutions with approximated graphons. Henceforth, no discussions on approximations will be included. Compared to the previous chapters on graphon control, the system operator is extended to any operator in the graphon unitary operator algebra, which is not necessarily compact. However, it is assumed that the input operators lie in the span of the identity input operator.
This framework permits the characterization of couplings in higher dimensional spaces compared to the weighted (or unweighted) mean-field coupling in [51, 71, 72, 73], and the coupling is represented as a graphon (or network) effect rather than a mean-field effect.

The organization of this chapter is as follows: we start with the centralized graphon-LQR control problem, then solve the problem via decoupling, and finally show that decentralized team optimal solutions with global information and local state observations optimally solve the original centralized graphon control problem.

### 6.2 System Models with Non-compact Graphon System Operators

#### 6.2.1 System Model

Consider the following graphon system model

\[
\dot{x}_t = \alpha x_t + Ax_t + \beta u_t, \\
x_t, u_t \in L^2[0, 1], \quad \alpha, \beta \in \mathbb{R}, \quad t \in [0, T],
\]

where \( A \in \tilde{G}_1^{sp} \) is the underlying graphon. Equivalently, the system model is represented by

\[
\dot{x}_t = (\alpha I + A)x_t + (\beta I)u_t, \\
x_t, u_t \in L^2[0, 1], \quad \alpha, \beta \in \mathbb{R}, \quad t \in [0, T],
\]

where \( I \) denotes the identity operator, \( (\alpha I + A) \) and \( (\beta I) \) are elements in the graphon unitary operator algebra \( G_{AI}^1 \).

It is obvious that \( (\alpha I + A) \) and \( (\beta I) \) are bounded linear operators on \( L^2[0, 1] \). Following [61], \( (\alpha I + A) \) is the infinitesimal generator of the uniformly (hence strongly) continuous semigroup

\[
S_A(t) := e^{(\alpha I + A)t} = \sum_{k=0}^{\infty} \frac{t^k(\alpha I + A)^k}{k!}.
\]

Therefore, the initial value problem of the graphon unitary algebra differential equation

\[
\dot{y}_t = (\alpha I + A)y_t, \quad y_0 \in L^2[0, 1]
\]

has a solution given by \( y_t = e^{(\alpha I + A)t}y_0 \).
Theorem 6.1. The system (6.2) has a unique mild solution $x \in C([0, T]; L^2[0, 1])$ for any $x_0 \in L^2[0, 1]$ and any $u \in L^2([0, T]; L^2[0, 1])$.

Proof. This comes from the fact that $(\alpha I + A)$ generates a strongly continuous semigroup and $(\beta I)$ is a (bounded) linear operator on $L^2[0, 1]$. Following the same lines as the proof in Theorem 3.1, we obtain the result. \qed

6.2.2 Cost Function

Assumption 6.1. $Q$ and $P_0$ are linear operators on $L^2[0, 1]$ that are Hermitian and non-negative (i.e., $Q, P_0 \succeq 0$).

The instantaneous cost is given by

$$c_t(u_t, x_t) = \langle x_t, Qx_t \rangle + \langle u_t, u_t \rangle,$$

and the terminal cost is given by

$$c_T(x_T) = \langle x_T, P_0x_T \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2[0, 1]$. The control objective is under Assumption 6.1 to obtain the control law $u$ that minimizes the quadratic cost

$$J(u, x) = \int_0^T c_t(x_t, u_t) dt + c_T(x_T),$$

subject to system constrains in (6.1) over the time horizon $[0, T]$.

6.2.3 Existence and Uniqueness of Solutions to LQR Problems

Denote $A = \alpha I + A$ and $B = \beta I$. Consider the following Riccati equation

$$\dot{P} = A^T P + PA - PBB^T P + Q, \quad P(0) = P_0.\quad (6.6)$$

Given the solution $P$ to the Riccati equation, the optimal control $u^*$ is given by

$$u^*_t = -B^T P(T - t)x^*_t, \quad t \in [0, T]$$

(6.7)
and moreover that $x^*$ is the solution to the closed loop equation

$$\dot{x}_t = Ax_t - BB^TP(T-t)x_t,$$

$$t \in [0, T], x_0 \in L^2[0, 1].$$

(6.8)

See e.g. [45].

Applying the results in [45] and specializing the Hilbert space there to be $L^2[0, 1]$ space, we can establish, under Assumption 6.1, the existence and uniqueness of the solution to the Riccati equation (6.6) and the existence and uniqueness of optimal solution pair $(u^*, x^*)$ in (6.7) and (6.8).

### 6.3 Team Optimal Solutions via Decoupling

According to [57, Chapter 2, Proposition 4.7], under the fact $\|A\|_2 < \infty$, the operator $A$ is a compact operator and hence it has a countable spectrum decomposition

$$A(x, y) = \sum_{i=1}^{\infty} \lambda_i f_i(x) f_i(y), \quad (x, y) \in [0, 1]^2,$$

(6.9)

where the convergence is in the $L^2[0, 1]^2$ sense, $\{\lambda_1, \lambda_2, \ldots\}$ is the set of eigenvalues (which are not necessarily distinct) and $\{f_1, f_2, \ldots\}$ denotes the set of the corresponding orthonormal eigenfunctions (i.e. $\|f_i\|_2 = 1$, and $\langle f_l, f_k \rangle = 0$ if $l \neq k$).

We know that the only accumulating point of the eigenvalues is zero, i.e., $\lim_{l \to \infty} \lambda_l = 0$ (see [40] and (2.5)). This implies that as a compact operator $A$ can be approximated by a finite truncation of spectrum decomposition with the most significant eigenvalues, which is also shown in (2.15). Therefore, it is reasonable to make the following assumption on the spectrum of the underlying graphon.

**Assumption 6.2.** The graphon $A$ as an operator has a finite number ($d$) eigenfunctions corresponding to the non-zero eigenvalues.

Furthermore, the cost structure is assumed to be a function of the underlying graphon.

**Assumption 6.3.** $Q = \text{poly}_Q(A) = q_0 I + \sum_{k=1}^{h} q_k A^k$, $P_0 = \text{poly}_{P_0}(A) = z_0 I + \sum_{k=1}^{r} z_k A^k$, $r > 0, h > 0.$
Notations
\[ \bar{x}_t^l = \langle x_t, f_l \rangle f_l \in L^2[0, 1] \text{ and } \bar{u}_t^l = \langle u_t, f_l \rangle f_l \in L^2[0, 1] \]
denote the projections of \( x_t \) and \( u_t \) onto the eigenspace spanned by \( f_l \in L^2[0, 1] \), respectively. \( \bar{x}_t^l = \langle x_t, f_l \rangle \) and \( \bar{u}_t^l = \langle u_t, f_l \rangle \) denote the real values for inner products.

6.3.1 Decoupling Method

An auxiliary state and an auxiliary control are introduced to decouple the optimal control problem into local parts and global parts. Define the auxiliary state and control as:
\[ \ddot{x}_t = x_t - \sum_{l=1}^d \bar{x}_t^l, \quad \ddot{u}_t = u_t - \sum_{l=1}^d \bar{u}_t^l \]  
(6.10)
where \( \bar{x}_t^l = \langle x_t, f_l \rangle f_l \) and \( \bar{u}_t^l = \langle u_t, f_l \rangle f_l \) denote the projections of \( x_t \) and \( u_t \) onto the eigenspace spanned by \( f_l \), respectively.

Decoupling in Dynamics

Applying the projection on (6.2), we obtain the following dynamics of the projections:
\[ \ddot{x}_t^l = (\alpha + \lambda_l)\bar{x}_t^l + \beta\bar{u}_t^l, \quad 1 \leq l \leq d. \]  
(6.11)
Furthermore,
\[ \dot{x}_t = \alpha x_t + Ax_t + \beta u_t \]
\[ - (\alpha + \lambda_1)\bar{x}_t^1 - \beta\bar{u}_t^1 \]
\[ ... \]
\[ - (\alpha + \lambda_d)\bar{x}_t^d - \beta\bar{u}_t^d. \]
(6.12)

Since
\[ Ax_t = \sum_{l=1}^d \lambda_l \langle x_t, f_l \rangle f_l = \sum_{l=1}^d \lambda_l \bar{x}_t^l, \]  
(6.13)
we obtain
\[ \dot{x}_t = \alpha \ddot{x}_t + \beta \ddot{u}_t. \]  
(6.14)
Decoupling in Costs

**Lemma 6.1.** Assume $Q = \text{poly}(A)$ with $\text{poly}(\cdot)$ representing polynomial functions. Then the following decoupling holds

$$
\langle x_t, Qx_t \rangle = \langle \bar{x}_t, Q\bar{x}_t \rangle + \sum_{l=1}^{d} \langle \bar{x}^l_t, Q\bar{x}^l \rangle.
$$

(6.15)

Furthermore, if $\text{poly}(s) = \sum_{k=0}^{h} q_k s^k$, i.e., $Q = q_0 I + \sum_{k=1}^{h} q_k A^k$, then

$$
\langle x_t, Qx_t \rangle = q_0 \|\bar{x}_t\|_2^2 + \sum_{l=1}^{d} \text{poly}(\lambda_l) \|\bar{x}_l^l\|_2^2.
$$

(6.16)

**Proof.** First, we show for any $k \geq 0$

$$
\langle x_t, A^k x_t \rangle = \langle \bar{x}_t, A^k \bar{x}_t \rangle + \sum_{l=1}^{d} \langle \bar{x}^l_t, A^k \bar{x}^l \rangle
$$

(6.17)

By decomposing the left hand side of (6.17), we have

$$
\langle x_t, A^k x_t \rangle = \langle \bar{x}_t, A^k \bar{x}_t \rangle + \sum_{l=1}^{d} \langle \bar{x}^l_t, A^k \bar{x}^l \rangle + 2 \langle \bar{x}_t, A^k \sum_{l=1}^{d} \bar{x}^l \rangle
$$

(6.18)

Further, the second term on the right hand side of (6.18) gives

$$
\sum_{l=1}^{d} \langle \bar{x}^l_t, A^k x^l_t \rangle = \sum_{l=1}^{d} \langle x_t, f_l \rangle f_l, A^k \sum_{l=1}^{d} \langle x_t, f_l \rangle f_l \rangle

= \sum_{l=1}^{d} \langle x_t, f_l \rangle f_l, \lambda^k_l \sum_{l=1}^{d} \langle x_t, f_l \rangle f_l \rangle

= \sum_{l=1}^{d} \langle x_t, f_l \rangle f_l, \lambda^k_l \langle x_t, f_l \rangle f_l \rangle

= \sum_{l=1}^{d} \langle x_t, f_l \rangle f_l, A^k \langle x_t, f_l \rangle f_l \rangle

= \sum_{l=1}^{d} \langle x^l_t, A^k x^l_t \rangle = \sum_{l=1}^{d} \lambda^k_l \|x^l_t\|_2^2
$$

(6.19)
and the last term on the right hand side of (6.18) gives

\[
\langle \ddot{x}_t, A^k \sum_{l=1}^d \dot{x}_l^i \rangle = \langle x_t - \sum_{l=1}^d x_l^i, A^k \sum_{l=1}^d x_l^i \rangle \\
= \langle x_t - \sum_{l=1}^d (x_t, f_l) f_l, A^k \sum_{l=1}^d (x_t, f_l) f_l \rangle \\
= \langle x_t, \sum_{l=1}^d \lambda_l^k (x_t, f_l) f_l \rangle - \langle \sum_{l=1}^d (x_t, f_l) f_l, \sum_{l=1}^d \lambda_l^k (x_t, f_l) f_l \rangle \\
= \sum_{l=1}^d \lambda_l^k (x_t, f_l)^2 - \sum_{l=1}^d \lambda_l^k (x_t, f_l)^2 \| f_l \|_2^2 \\
= (\text{since } \| f_l \|_2 = 1) \\
= 0.
\]

Hence, we prove (6.17). Since this separation result holds for all powers of $A$, together with the linearity of inner product, we have

\[
\langle x_t, \text{poly}(A)x_t \rangle = \langle \ddot{x}_t, \text{poly}(A)x_t \rangle + \sum_{l=1}^d \langle \dot{x}_l^i, \text{poly}(A)x_l^i \rangle \\
= \sum_{l=1}^d \langle \dot{x}_l^i, \text{poly}(A)x_l^i \rangle.
\]  

(6.21)

With $Q = \text{poly}(A)$, we prove the result in (6.15). Furthermore,

\[
\langle \ddot{x}_t, \text{poly}(A)x_l^i \rangle = \langle (x_t, f_l) f_l, \text{poly}(A)(x_t, f_l) f_l \rangle \\
= \langle (x_t, f_l) f_l, \text{poly}(\lambda_l)(x_t, f_l) f_l \rangle \\
= \text{poly}(\lambda_l) \| \ddot{x}_l^i \|_2
\]

(6.22)

and

\[
\langle \ddot{x}_t, Qx_t \rangle = \langle \ddot{x}_t, q_0 \ddot{x}_t \rangle + \langle \ddot{x}_t, \sum_{k=1}^h q_k A^k \ddot{x}_t \rangle \\
= \langle \ddot{x}_t, q_0 \ddot{x}_t \rangle = q_0 \| \ddot{x}_t \|_2
\]  

(6.23)

we have the result in (6.16). □

**Lemma 6.2.** If **Assumption 6.3** holds, i.e., $Q$ and $P_0$ are polynomials of $A$, then

\[
c_t(u_t, x_t) = \ddot{c}(\ddot{u}_t, \ddot{x}_t) + \sum_{l=1}^d c_l^T(\ddot{u}_l^i, \ddot{x}_l^i) \quad \text{and} \quad c_T(x_t) = \ddot{c}_T(\ddot{x}_t) + \sum_{l=1}^d \ddot{c}_T(\ddot{x}_l^i),
\]
By applying the result in Lemma 6.1 to the cost functions, we obtain the result.

**Proof.** By applying the result in Lemma 6.1 to the cost functions, we obtain the result.

Therefore, we can separate the LQR problem into \((d + 1)\) decoupled LQR problems:

\[
\begin{align*}
\dot{x}_l^i &= (\alpha + \lambda_l)x_l^i + \beta \tilde{u}_l^i, \\
\tilde{J}(\tilde{u}_l, \tilde{x}_l) &= \int_0^T \tilde{c}_l(\tilde{u}_l^i, \tilde{x}_l^i) dt + \tilde{c}_T(\tilde{x}_T), \quad 1 \leq l \leq d
\end{align*}
\]  

(6.24)

where \(\tilde{c}_l(\tilde{u}_l^i, \tilde{x}_l^i) = poly_Q(\lambda_l)\|\tilde{x}_l^i\|_2^2 + \|\tilde{u}_l^i\|_2^2\) and \(\tilde{c}_T(\tilde{x}_T) = poly_{P_0}(\lambda_l)\|\tilde{x}_T\|_2^2\);

\[
\begin{align*}
\dot{x}_l &= \alpha \tilde{x}_l + \beta \tilde{u}_l, \\
\tilde{J}(\tilde{u}_l, \tilde{x}_l) &= \int_0^T \tilde{c}_l(\tilde{u}_l, \tilde{x}_l) dt + \tilde{c}_T(\tilde{x}_T),
\end{align*}
\]  

(6.25)

where \(\tilde{c}_l(\tilde{u}_l, \tilde{x}_l) = q_0\|\tilde{x}_l\|_2^2 + \|\tilde{u}_l\|_2^2\) and \(\tilde{c}_T(\tilde{x}_T) = z_0\|\tilde{x}_T\|_2^2\).

**Lemma 6.3.** If Assumptions 6.1-6.3 are satisfied, and \(poly_Q(\cdot)\) and \(poly_{P_0}(\cdot)\) are nonnegative for any eigenvalue \(\lambda_l(1 \leq l \leq d)\) of \(A\), then solving the optimal control problems (6.24) and (6.25) is equivalent to solving the original optimal control problem defined by (6.2) and (6.5). The optimal control solution exists and is unique.

**Proof.** The cost defined by (6.5) can be decoupled as

\[
J(u, x) = \tilde{J}(\tilde{u}_t, \tilde{x}_t) + \sum_{l=1}^d \tilde{J}(\tilde{u}_l^i, \tilde{x}_l^i),
\]  

(6.26)

with the summation of non-negative terms on the right hand side. Therefore \(J(u, x)\) is minimized if and only if \(\tilde{J}(\tilde{u}_t, \tilde{x}_t)\) and \(\tilde{J}(\tilde{u}_l^i, \tilde{x}_l^i), (1 \leq l \leq d)\), are minimized. Furthermore, the original dynamics defined by (6.2) are decoupled into dynamics of auxiliary system and that in each eigenfunction direction (i.e. the space spanned by the corresponding eigenfunction). Hence, solving the optimal control problems (6.24) and (6.25) is equivalent to solving the original optimal control problem defined by (6.2) and (6.5).
6.3.2 Team Optimal Solutions

We use “eigenfunction direction” to represent the space spanned by the corresponding eigenfunction. To solve the problem in a decentralized manner, each agent should solve the following optimal control problems in all eigenfunction directions:

\[
\begin{aligned}
\dot{x}_l^i &= (\alpha + \lambda_l)x_l^i + \beta \bar{u}_l^i, \\
J_l^i(u_l^i, x_l^i) &= \int_0^T c_l^i(u_l^i, x_l^i) dt + \bar{c}_T(x_l^i), 1 \leq l \leq d,
\end{aligned}
\]  

(6.27)

where \(c_l^i(u_l^i, x_l^i) = \text{poly}_Q(\lambda_l)(x_l^i)^2 + (u_l^i)^2\), \(\bar{c}_T(x_l^i) = \text{poly}_{P_0}(\lambda_l)(x_l^i)^2\), \(x_l^i = \langle x_t, f_l \rangle\), and \(\bar{u}_l^i = \langle u_t, f_l \rangle\); in addition, for the agent with the index \(\gamma \in [0, 1]\), it should solve the following optimal control problem of the auxiliary system:

\[
\begin{aligned}
\dot{x}_l(\gamma) &= \alpha \dot{x}_l(\gamma) + \beta \ddot{u}_l(\gamma), \\
J_l(\dot{u}_l(\gamma), \ddot{x}_l(\gamma)) &= \int_0^T \ddot{c}_l(\dot{u}_l(\gamma), \ddot{x}_l(\gamma)) dt + \ddot{c}_T(\ddot{x}_l(\gamma)),
\end{aligned}
\]  

(6.28)

where \(\ddot{c}_l(\dot{u}_l(\gamma), \ddot{x}_l(\gamma)) = q_0(\ddot{x}_l(\gamma))^2 + (\dot{u}_l(\gamma))^2\) and \(\ddot{c}_T(\ddot{x}_l(\gamma)) = z_0(\ddot{x}_l(\gamma))^2\).

**Theorem 6.2.** If Assumptions 6.1-6.3 are satisfied, and \(\text{poly}_Q(\cdot)\) and \(\text{poly}_{P_0}(\cdot)\) are nonnegative for any eigenvalue \(\lambda_l (1 \leq l \leq d)\) of \(A\), then solving the optimal control problems defined by (6.27) and (6.28) locally is equivalent to solving the original optimal control problem defined by (6.2) and (6.5). Moreover, the team optimal control law for the \(\gamma^{th}\) agent is given by

\[
u_l(\gamma) = -\beta \left( L_{T-t} \ddot{x}_l(\gamma) + \sum_{l=1}^d M_{T-t}^{(l)} \bar{x}_l^i(\gamma) \right),
\]  

(6.29)

where \(L := \{L_t : t \in [0, T]\}\) is the solution to the Riccati equation

\[
\dot{L}_t = 2\alpha L_t - \beta^2 L_t^2 + q_0, \quad L_0 = z_0,
\]  

(6.30)

and \(M^{(l)} := \{M_t^{(l)} : t \in [0, T]\}\) is the solution to the Riccati equation

\[
\begin{aligned}
\dot{M}_t^{(l)} &= 2(\alpha + \lambda_l)M_t^{(l)} - (M_t^{(l)})^2 + \text{poly}_Q(\lambda_l), \\
M_0^{(l)} &= \text{poly}_{P_0}(\lambda_l), 1 \leq l \leq d.
\end{aligned}
\]  

(6.31)

**Proof.** Firstly, since \(\bar{x}_l^i = x_l^i f_i\), \(\bar{u}_l^i = \bar{u}_l^i f_i\) and \(\|f_i\|_2 = 1\) for \(1 \leq l \leq d\), the optimal control
problem in (6.24) can be equivalently solved by solving (6.27) and then recovering the optimal pair \((\bar{x}_l, \bar{u}_l)\) in the eigenfunction direction \(f_l, \ 1 \leq l \leq d\). Secondly, notice that \(J(\bar{u}_l, \bar{x}_l) = \int_0^1 J(\bar{u}_l(\gamma), \bar{x}_l(\gamma))d\gamma\) and \(J(\bar{u}_l(\gamma), \bar{x}_l(\gamma))\) is non-negative for any \(\gamma \in [0, 1]\). Therefore the optimal control problems (6.25) and (6.28) are equivalent. Finally, together with the result in Lemma 6.3, we obtain that solving the optimal control problems defined by (6.27) and (6.28) locally is equivalent to solving the original optimal control problem defined by (6.2) and (6.5).

It is obvious that (6.30) and (6.31) are the Riccati equations for the LQR problems (6.28) and (6.27), respectively. By the standard LQR theory, the optimal control laws are given by

\[
\bar{u}_l(\gamma) = -\beta L_T^{-1} \bar{x}_l(\gamma) \quad \text{and} \quad \bar{u}_l = -\beta M_T^{(l)} \bar{x}_l f_l(\gamma),
\]

respectively.

Since \(u_t(\gamma) = \bar{u}_t(\gamma) + \sum_{l=1}^d \bar{u}_l(\gamma) = \bar{u}_t(\gamma) + \sum_{l=1}^d \bar{u}_l f_l(\gamma)\), we obtain the team optimal control law for the original problem defined by (6.2) and (6.5):

\[
u_t(\gamma) = -\beta \left( L_{T-t} \bar{x}_t(\gamma) + \sum_{l=1}^d M_T^{(l)} \bar{x}_l f_l(\gamma) \right).
\] (6.32)

\[\square\]

### 6.3.3 Discussion

The following the global information is required by each agent to generate the team optimal solution:

1. all the eigenvalues of \(A\) and the value of the respective eigenfunctions at its index location (i.e. \(\lambda_l, f_l(\gamma)\) for all \(1 \leq l \leq d\))

2. the projections of the initial state \(x_0\) onto each eigenfunction direction (i.e. \(\bar{x}_0^l\) for all \(1 \leq l \leq d\)).

The computation complexity for each agent involves

1. solving the Riccati equation corresponding to auxiliary state dynamics, which is one-dimensional;

2. solving \(d\) one-dimensional Riccati equations corresponding to \(d\) eigenfunction directions.

Note that if the underlying graphon is an uniform graphon \(A(x, y) = 1\) for all \(x, y \in [0, 1]\), which gives \(d = 1, f_1 = 1 \in L^2[0, 1]\) and \(\lambda_1 = 1\), then the LQR problem with graphon coupling reduces to the LQR problem with mean-field coupling.
Chapter 7

Team Optimal Linear Quadratic Regulation (LQR) on Networks

7.1 Introduction

In this chapter, we study LQR optimal control problems for finite dimensional networks of linear systems, where agents (or subsystems) on a network have the same local dynamics and are coupled through the underlying network. Depending on the locations in the network, each agent receives different network influences on its dynamics. The total cost function also depends on the network structure. It turns out that, if the network can be characterized by a spectral decomposition that has a smaller dimension compared to the network size, the complexity of the control problem can be reduced. Moreover, an equivalent team optimal LQR problem is formulated by decoupling the original LQR problem for systems on networks. Finally, the team optimal LQR control law for systems on networks is established.

The coupling for the multi-agent systems is represented as a network effect rather than a mean field effect. This framework permits charactering couplings in a higher dimensional space compared to the mean field coupling in [51, 71].

Compared to Chapter 6, due to the finite dimensionality of the networks, the couplings in the dynamics are not limited to couplings with averaging dynamics, and can be extended to network couplings without scaling.
Comparisons with Existing Literatures

The idea of extending the mean field coupling to the network coupling for multi-agent systems is original which generalizes the coupling to higher dimensions. The coupling is represented as a network effect rather than a mean field effect.

Solving a linear quadratic optimal control problem as decoupled problems has been used in the previous work such as [50, 51, 71, 72, 73]. The present work is different from the previous in the following:

- **Difference in dynamics couplings:** The coupling in the present work is the coupling over networks, which allows a higher dimensional characterization of the coupling than the mean field coupling (which is one dimensional) in [50, 51, 71, 72, 73]. If the network permits a simple spectrum characterization by only one nonzero eigenvalue, the problem then reduces to the (weighted) mean field coupling case in [51, 71]. The work in [51] considers the case with heterogeneous populations in the sense that each population can have different dynamics. All the above mentioned work other than [51] assumes homogenous populations. In this work it is assumed that all the agents have the same local dynamics but they can receive different network influences from the network coupling.

- **Difference in cost structure:** In this work the cost can be any positive polynomial structures of the underlying network. This has a simpler representation to show the possible cost structures with which we can still achieve the decoupling of the LQR problems on networks. The cost structure in [51] is limited to functions of complete graphs or the Kronecker product of such a complete graph with finite dimensional graphs. In the weighted mean field case in [51, 71], the underlying cost structure can be represented as the linear combination of an identity and the underlying network (which has only one nonzero eigenvalue).

- **Difference in solutions:** In this work every agent with same local dynamics can receive a different network coupling and the problem can still be solved by each agent solving \((d + 1)\) Riccati equations where \(d\) is the number of non-zero eigenvalues of the underlying networks. To solve this same problem using the solutions in [51], the number of Riccati equations each agent needs to solve is \(K + 1\) where \(K\) is the number of different classes. Note that for the problem with homogenous local dynamics and network couplings the
following relation holds: $d \leq K \leq N$ where $N$ denotes the size of the population. This difference can be illustrated by the numerical example on the cosinusoidal network.

- **Difference in characterizing exchangeability:** The notion of exchangeability (in the sense that exchanging two agents does not affect the cost and dynamics) is proposed in [51]. The problem with multiple subpopulations with exchangeability in each subpopulation [51] boils down to constraining the underlying networks to be a complete network, or the Kronecker product of a complete network with finite networks. In this work, no exchangeability assumption is posed. However, we require the cost structure to be a function of the underlying network. This has a connection to the notion of exchangeability in the sense that relabelings on the nodes in the network do not change the system dynamics and costs. Nodes (or agents) that are exchangeable are simply twin nodes (i.e. nodes that share the same connection structure in the network).

**Notation**

In this chapter, $n$ and $p$ denote respectively the dimension of state and that of control for each agent, $N$ denotes the size of the networks, and $d$ denotes the total number of orthonormal eigenvectors of the network corresponding to non-zero eigenvalues.

### 7.2 System Model and LQR Problems on Networks

#### 7.2.1 System Model

Consider LQR problem on systems coupled over an undirected network with the following agent dynamics

$$
\dot{x}_i^t = A_t x_i^t + B_t u_i^t + D_t z_i^t + E_t \gamma_i^t, \quad x_i^t, z_i^t \in \mathcal{R}^n, u_i^t, \gamma_i^t \in \mathcal{R}^p
$$

where $A_t$, $B_t$, $D_t$ and $E_t$ are matrices with appropriate dimensions, and

$$
z_t^i = \sum_{j=1}^{N} x_j^t m_{ji} \quad \text{and} \quad \gamma_t^i = \sum_{j=1}^{N} u_j^t m_{ji}
$$

are respectively the locally perceived network coupling of state and that of control actions with $m_{ij}$ representing the $ij^{th}$ element of the network adjacency matrix $M \in \mathcal{R}^{N \times N}$. Equivalently, the
Table 7.1  Summary of Notation

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t^i \in \mathcal{R}^n$</td>
<td>state of agent $i$</td>
</tr>
<tr>
<td>$u_t^i \in \mathcal{R}^p$</td>
<td>control of agent $i$</td>
</tr>
<tr>
<td>$z_t^i \in \mathcal{R}^n$</td>
<td>network field for agent $i$</td>
</tr>
<tr>
<td>$x_t = (x_t^1, x_t^2, ..., x_t^N) \in \mathcal{R}^{n \times N}$</td>
<td>state matrix</td>
</tr>
<tr>
<td>$u_t = (u_t^1, u_t^2, ..., u_t^N) \in \mathcal{R}^{p \times N}$</td>
<td>control input matrix</td>
</tr>
<tr>
<td>$z_t = (z_t^1, z_t^2, ..., z_t^N) \in \mathcal{R}^{n \times N}$</td>
<td>network field matrix</td>
</tr>
<tr>
<td>$v_l \in \mathcal{R}^n, 1 \leq l \leq d$</td>
<td>the $l^{th}$ eigenvector of the network adjacency matrix</td>
</tr>
<tr>
<td>$v_l(j) \in \mathcal{R}, 1 \leq j \leq n$</td>
<td>the $j^{th}$ element of $v_l \in \mathcal{R}^n$</td>
</tr>
<tr>
<td>$\bar{x}_t^{(l)} = (x_t^1, ..., x_t^N)v_l v_l^T \in \mathcal{R}^{n \times N}$</td>
<td>the projection of $x_t$ into $v_l$</td>
</tr>
<tr>
<td>$\bar{x}_t^{(l)}j = (x_t^1, ..., x_t^N)v_l v_l(j) \in \mathcal{R}^{n \times 1}$</td>
<td>the $j^{th}$ column of $\bar{x}_t^{(l)}$</td>
</tr>
<tr>
<td>$\bar{u}_t^{(l)} = (u_t^1, ..., u_t^N)v_l v_l^T \in \mathcal{R}^{p \times N}$</td>
<td>the projection of $u_t$ into $v_l$</td>
</tr>
<tr>
<td>$\bar{u}_t^{(l)}j = (u_t^1, ..., u_t^N)v_l v_l(j) \in \mathcal{R}^{p \times 1}$</td>
<td>the $j^{th}$ column of $\bar{u}_t^{(l)}$</td>
</tr>
<tr>
<td>$\bar{x}<em>t^i = x_t^i - \sum</em>{l=1}^{d} \bar{x}_t^{(l)i}$</td>
<td>auxiliary state of agent $i$</td>
</tr>
<tr>
<td>$\bar{u}<em>t^i = u_t^i - \sum</em>{l=1}^{d} \bar{u}_t^{(l)i}$</td>
<td>auxiliary control of agent $i$</td>
</tr>
</tbody>
</table>

compact form of system dynamics is given by

$$\dot{x}_t = A_t x_t + B_t u_t + D_t z_t + E_t \gamma_t,$$  \hspace{1cm} (7.2)

where $x_t = (x_t^1, x_t^2, ..., x_t^N) \in \mathcal{R}^{n \times N}$, $z_t = (z_t^1, ..., z_t^N) = x_t M \in \mathcal{R}^{n \times N}$, $u_t = (u_t^1, ..., u_t^N) \in \mathcal{R}^{p \times N}$ and $\gamma_t = (\gamma_t^1, ..., \gamma_t^N) = u_t M \in \mathcal{R}^{p \times N}$. As an analogy to the “mean field”, we call $z_t^i$ (and $\gamma_t^i$) the “network field” of the state (and control actions) at node $i$. 
### 7.2.2 LQR Problems on Networks

Let \( \langle \cdot, \cdot \rangle \) denote the vector inner product and for simplicity let \( \|x\|_Q^2 := \langle x, Qx \rangle \) for any \( x \in \mathbb{R}^N, Q \in \mathbb{R}^{N \times N}, Q \geq 0 \). The instantaneous cost is given by

\[
c_t(u_t, x_t, z_t) = \sum_{i=1}^{N} \sum_{j=1}^{N} g_{ij}(x^i_t, Q_t x^j_t) + \sum_{i=1}^{N} \sum_{j=1}^{N} o_{ij}(u^i_t, R_t u^j_t) + \sum_{i=1}^{N} \|z^i_t\|_{H_t}^2 \tag{7.3}
\]

where \( [g_{ij}] = G \) and \( G := \text{poly}_q(M) \) is a polynomial of the network adjacency matrix \( M \) with \( \text{poly}_q(x) := q_0 + \sum_{i=1}^{c_q} q_i x^i; \) similarly, \( [o_{ij}] = O \) and \( O := \text{poly}_r(M) \) with \( \text{poly}_r(x) := r_0 + \sum_{i=1}^{c_r} r_i x^i. \) Here \( c_q \) and \( c_r \) denotes the orders of the respective polynomial functions \( \text{poly}_q(\cdot) \) and \( \text{poly}_r(\cdot). \)

Note that if \( M \) is symmetric, then \( G \) and \( O \) are both symmetric. The terminal cost is given by

\[
c_T(x_T, z_T) = \sum_{i=1}^{N} \sum_{j=1}^{N} g_{ij}(x^i_T, Q_T x^j_T) + \sum_{i=1}^{N} \|z^i_T\|_{H_T}^2. \tag{7.4}
\]

This formulation for cost functions applies to cases where the cost functions contain cross-coupling terms for \( z_t \) since it can be absorbed into the cross-coupling terms for \( x_t \). (See Lemma 7.7). For simplicity, we do not include such cross-coupling terms for \( z_t \) in (7.3) and (7.4).

The control objective is to minimize the quadratic cost

\[
J(u, x) = \int_0^T c_t(u_t, x_t, z_t) dt + c_T(x_T, z_T) \tag{7.5}
\]

over the time horizon \([0, T]\) subject to the system dynamics contrains in (7.1).

### 7.2.3 Basic Assumptions

Since the underlying network we consider is undirected, the underlying adjacency matrix \( M \) is symmetric. Therefore, there exists a spectral decomposition for \( M \) such that

\[
M = [v_1, v_2, \ldots, v_d] \Lambda_d [v_1, v_2, \ldots v_d]^T
\]

where \( v_1, \ldots, v_d \) denote the orthonormal eigenvectors with non-zero eigenvalues, \( \Lambda_d = \text{diag}(\lambda_1, \ldots, \lambda_d) \) with \( \lambda_1, \ldots, \lambda_d \) as the corresponding non-zero eigenvalues.

**Assumption 7.1.** \( \text{poly}_q(s) \geq 0 \) and \( \text{poly}_r(s) > 0 \) for all \( s \in \{0, \lambda_1, \ldots, \lambda_d\}. \)
Assumption 7.2. \( Q_t \geq 0 \) and \( H_t \geq 0 \) for all \( t \in [0, T] \); \( R_t > 0 \) for all \( t \in [0, T] \).

Lemma 7.1. If Assumption 7.1 holds, then \( \text{poly}_q(M) \geq 0 \) and \( \text{poly}_r(M) > 0 \).

Proof. It is obvious that the eigenvalues of \( \text{poly}_q(M) \) are given by

\[
\{ \text{poly}_q(s) : s \text{ is an eigenvalue of } M \}.
\]

Hence Assumption 7.1 implies all the eigenvalues of \( \text{poly}_q(M) \) are nonnegative and therefore \( \text{poly}_q(M) \geq 0 \). Similarly, we obtain \( \text{poly}_r(M) > 0 \).

Denote \( V = (v_1, v_2, \ldots, v_d) \in \mathbb{R}^{N \times d}, \bar{x}^{(l)}_t = (x^1_t, \ldots, x^N_t) v_l v_l^T \in \mathbb{R}^{n \times N} \), and \( \bar{u}^{(l)}_j = (x^1_t, \ldots, x^N_t) v_l v_l(j) \in \mathbb{R}^{n \times 1} \) with \( v_l(j) \) representing the \( j \)-th element of the vector \( v_l \in \mathbb{R}^n \). Similarly, denote \( \bar{u}^{(l)}_t = (u^1_t, \ldots, u^N_t) v_l v_l^T \in \mathbb{R}^{p \times N} \) and \( \bar{u}^{(l)}_t = (u^1_t, \ldots, u^N_t) v_l v_l(j) \in \mathbb{R}^{p \times 1} \).

### 7.3 Team Optimal Solutions via Decoupling

#### 7.3.1 Decoupling in Dynamics

Multiplying \( v_l v_l^T \) from the right on (7.2), we obtain

\[
\dot{x}^i_t v_l v_l^T = A_t x^j_t v_l v_l^T + B_t u^j_t v_l v_l^T + D_t z_t v_l v_l^T. \tag{7.6}
\]

With a further simplification, the dynamics in \( l \)-th eigendirection is given by

\[
\begin{cases}
\dot{x}^{(l)}_t = (A_t + D_t \lambda_l) \bar{x}^{(l)}_t + B_t \bar{u}^{(l)}_t, \\
\bar{x}^{(l)}_t \in \mathbb{R}^{n \times N}, \bar{u}^{(l)}_t \in \mathbb{R}^{p \times N}, 1 \leq l \leq d
\end{cases} \tag{7.7}
\]

and equivalently by the nodal form

\[
\begin{cases}
\dot{x}^{(l)i}_t = (A_t + D_t \lambda_l) \bar{x}^{(l)i}_t + (B_t + E_t \lambda_l) \bar{u}^{(l)i}_t, \\
\bar{x}^{(l)i}_t \in \mathbb{R}^{n \times 1}, \bar{u}^{(l)i}_t \in \mathbb{R}^{p \times 1}, 1 \leq l \leq d, 1 \leq i \leq N.
\end{cases} \tag{7.8}
\]

Denote the auxiliary state and control as

\[
\dot{x}_t^i = x_t^i - \sum_{l=1}^d \bar{x}^{(l)i}_t, \quad \dot{u}_t^i = u_t^i - \sum_{l=1}^d \bar{u}^{(l)i}_t.
\]
Lemma 7.2. The auxiliary system dynamics is given by

\[ \dot{x}_t = A_t \dot{x}_t + B_t \dot{u}_t \]  

(7.9)

and equivalently by

\[ \dot{x}_i^t = A_t \dot{x}_i^t + B_t \dot{u}_i^t, \quad 1 \leq i \leq N, \]  

(7.10)

where \( \dot{x}_t = (\dot{x}_1^t, \ldots, \dot{x}_d^t) \) and \( \dot{u}_t = (\dot{u}_1^t, \ldots, \dot{u}_d^t) \).

Proof. From (7.2) and (7.7), we obtain

\[ \dot{x}_t = A_t \dot{x}_t + B_t \dot{u}_t + D_t \left( z_t - \sum_{l=1}^{d} \lambda_l \bar{x}_t^{(l)} \right) + E_t \left( \gamma_t - \sum_{l=1}^{d} \lambda_l \bar{u}_t^{(l)} \right) \]

\[ = A_t \dot{x}_t + B_t \dot{u}_t + D_t \left( x_t M - \sum_{l=1}^{d} \lambda_l x_t v_l v_l^T \right) + E_t \left( u_t M - \sum_{l=1}^{d} \lambda_l u_t v_l v_l^T \right) \]

\[ = A_t \dot{x}_t + B_t \dot{u}_t + D_t \left( \sum_{l=1}^{d} \lambda_l x_t v_l v_l^T - \sum_{l=1}^{d} \lambda_l x_t v_l v_l^T \right) + E_t \left( \sum_{l=1}^{d} \lambda_l u_t v_l v_l^T - \sum_{l=1}^{d} \lambda_l u_t v_l v_l^T \right) \]

\[ = A_t \dot{x}_t + B_t \dot{u}_t. \]

\[ \square \]

7.3.2 Decoupling in the Cost Function

Denote

\[ (x_t, Q_t x_t)_G := \sum_{i=1}^{N} \sum_{j=1}^{N} g_{ij} (x_i^t, Q_t x_j^t), \]

where \( G = [g_{ij}] \).

Lemma 7.3. Let \( G = M^k, \quad (k > 1) \). Assume \( Q_t \) and \( M = [m_{ij}] \) are symmetric. Then

\[ (x_t, Q_t x_t)_G = \sum_{l=1}^{d} \lambda_l^k (\bar{x}_t^{(l)}, Q_t \bar{x}_t^{(l)})_I \]

where \( I \in \mathcal{R}^{n \times n} \) is the identity matrix.

Proof. The symmetry of \( M \) implies the symmetry of \( G \) and hence matrices \( G \) and \( M \) are diagonalizable with real eigenvalues. It is obvious that \( G \) and \( M \) share the same eigenvectors and
$G v_l = \lambda^k v_l$. Replacing the states by auxiliary states and states in each eigenvector direction yields

$$(x_t, Q_t x_t)_G = \left( \bar{x}_t + \sum_{l=1}^{d} \bar{x}^{(l)}_t, Q_t (\bar{x}_t + \sum_{l=1}^{d} \bar{x}^{(l)}_t) \right)_G.$$ 

By the linearity of inner product and summation, we obtain

$$(x_t, Q_t x_t)_G = (\bar{x}_t, Q_t \bar{x}_t)_G + \left( \sum_{l=1}^{d} \bar{x}^{(l)}_t, Q_t \sum_{l=1}^{d} \bar{x}^{(l)}_t \right)_G \tag{7.11}$$

To show the first term in (7.11) is zero, we start with

$$(\bar{x}_t, Q_t \bar{x}_t)_M = \sum_{i=1}^{N} \sum_{j=1}^{N} m_{ij} (\bar{x}^i_t, Q_t \bar{x}^j_t)$$

$$= \sum_{j=1}^{N} \left( \sum_{i=1}^{N} m_{ij} (x^i_t - \sum_{l=1}^{d} \bar{x}^{(l)i}_t), Q_t \bar{x}^j_t \right).$$

Note that

$$\sum_{i=1}^{N} m_{ij} (x^i_t - \sum_{l=1}^{d} \bar{x}^{(l)i}_t) = \sum_{i=1}^{N} m_{ij} (x^i_t - \sum_{l=1}^{d} x_t v_l v_l(i))$$

$$= \sum_{i=1}^{N} m_{ij} x^i_t - \sum_{l=1}^{d} x_t v_l (\sum_{i=1}^{N} m_{ij} v_l(i))$$

$$= \sum_{i=1}^{N} m_{ij} x^i_t - \sum_{l=1}^{d} \lambda_l x_t v_l v_l(j)$$

(since both terms represent the $j^{th}$ column of $x_t M$)

$$= 0.$$ 

Therefore $(\bar{x}_t, Q_t \bar{x}_t)_M = 0$ and similarly $(\bar{x}_t, Q_t \bar{x}_t)_M^k = 0$. The second term in (7.11) can be
simplified as follows:

\[
\begin{align*}
\left( \sum_{l=1}^{d} \bar{x}_t^{(l)}, Q_t \sum_{l=1}^{d} \bar{x}_t^{(l)} \right)_G = & \sum_{i=1}^{N} \sum_{j=1}^{N} g_{ij} \left( \sum_{l=1}^{d} \bar{x}_t^{(l)i}, \sum_{h=1}^{d} \bar{x}_t^{(h)j} \right) \\
= & \sum_{i=1}^{N} \sum_{j=1}^{N} \left( g_{ij} \sum_{l=1}^{d} x_t v_l v_l(i), Q_t \sum_{h=1}^{d} \bar{x}_t^{(h)j} \right) \\
= & \sum_{j=1}^{N} \left( \sum_{l=1}^{d} x_t v_l \sum_{i=1}^{N} g_{ij} v_l(i), Q_t \sum_{h=1}^{d} \bar{x}_t^{(h)j} \right) \\
= & \left( \text{since } G v_l = \lambda_t^k v_l \right) \\
= & \sum_{j=1}^{N} \left( \sum_{l=1}^{d} x_t v_l \lambda_t^k v_l(j), Q_t \sum_{h=1}^{d} \bar{x}_t^{(h)j} \right) \\
= & \sum_{j=1}^{N} \left( \sum_{l=1}^{d} \bar{x}_t^{(l)j} \lambda_t^k, Q_t \sum_{h=1}^{d} \bar{x}_t^{(h)j} \right) \\
= & \sum_{l=1}^{d} \sum_{h=1}^{d} \lambda_t^k \left( \sum_{j=1}^{N} (x_t v_l v_l(j), Q_t \bar{x}_t^{(h)j}) \right) \\
= & \sum_{l=1}^{d} \sum_{h=1}^{d} \lambda_t^k \left( \sum_{j=1}^{N} v_l(j) v_h(j) \right) \left( x_t v_l, Q_t x_t v_h \right) \\
= & \left( \text{since for all } h \neq l, v_l^T v_h = 0 \right) \\
= & \sum_{l=1}^{d} \lambda_t^k \left( \sum_{j=1}^{N} v_l(j) v_l(j) \right) \left( x_t v_l, Q_t x_t v_l \right) \\
= & \sum_{l=1}^{d} \lambda_t^k \left( \sum_{j=1}^{N} \bar{x}_t^{(l)j} \right) \left( x_t v_l, Q_t \bar{x}_t^{(l)j} \right) \\
= & \sum_{l=1}^{d} \lambda_t^k \left( \bar{x}_t^{(l)}, Q_t \bar{x}_t^{(l)} \right)_l.
\end{align*}
\]
The third term in (7.11) can be simplified as follows:

\[
\begin{aligned}
\left(\bar{x}_t, Q_t \sum_{l=1}^{d} \bar{x}_t^{(l)} \right)_G &= \left( x_t - \sum_{l=1}^{d} \bar{x}_t^{(l)}, Q_t \sum_{l=1}^{d} \bar{x}_t^{(l)} \right)_G \\
&= \left( x_t, Q_t \sum_{l=1}^{d} \bar{x}_t^{(l)} \right)_G - \left( \sum_{l=1}^{d} \bar{x}_t^{(l)}, Q_t \sum_{l=1}^{d} \bar{x}_t^{(l)} \right)_G \\
&= \left( x_t, Q_t \sum_{l=1}^{d} \bar{x}_t^{(l)} \right)_G - \sum_{l=1}^{d} \lambda^k_l \left( \sum_{j=1}^{N} v_l(j) v_l(j) \right) \left( x_t v_l, Q_t x_t v_l \right) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} g_{ij} \left( x_t^i, Q_t \sum_{h=1}^{d} x_t v_h v_h(j) \right) - \sum_{l=1}^{d} \lambda^k_l \left( x_t v_l, Q_t x_t v_l \right) \\
&= \sum_{i=1}^{N} \sum_{h=1}^{d} \left( \sum_{j=1}^{N} g_{ij} v_h(j) \right) \left( x_t^i, Q_t x_t v_h \right) - \sum_{l=1}^{d} \lambda^k_l \left( x_t v_l, Q_t x_t v_l \right) \\
&= \sum_{i=1}^{N} \sum_{h=1}^{d} \lambda^k_h v_h(i) \left( x_t^i, Q_t x_t v_h \right) - \sum_{l=1}^{d} \lambda^k_l \left( x_t v_l, Q_t x_t v_l \right) \\
&= \sum_{h=1}^{d} \lambda^k_h \left( \sum_{i=1}^{N} v_h(i) x_t^i, Q_t x_t v_h \right) - \sum_{l=1}^{d} \lambda^k_l \left( x_t v_l, Q_t x_t v_l \right) \\
&= \sum_{h=1}^{d} \lambda^k_h \left( x_t v_h, Q_t x_t v_h \right) - \sum_{l=1}^{d} \lambda^k_l \left( x_t v_l, Q_t x_t v_l \right) \\
&= 0.
\end{aligned}
\]

Similarly, the last term in (7.11)

\[
\left( \sum_{l=1}^{d} \bar{x}_t^{(l)}, Q_t \bar{x}_t \right)_G = 0. \tag{7.12}
\]

\[\square\]

**Lemma 7.4.** Assume \( Q_t \) is symmetric, and denote

\[
(x_t, Q_t x_t)_I = \sum_{i=1}^{N} (x_t^i, Q_t x_t^i). \tag{7.13}
\]
Then
\[
(x_t, Q_t x_t)_I = (\ddot{x}_t, Q_t \ddot{x}_t)_I + \sum_{l=1}^{d} (\ddot{x}_t^{(l)}, Q_t \ddot{x}_t^{(l)})_I. \tag{7.14}
\]

**Proof.** An expansion of \((x_t, Q_t x_t)_I\) yields
\[
(x_t, Q_t x_t)_I = \sum_{i=1}^{N} (x_t^{i}, Q_t x_t^{i})
= \sum_{i=1}^{N} (\ddot{x}_t^{i} + \sum_{l=1}^{d} \ddot{x}_t^{(l)i}), Q_t (\ddot{x}_t^{i} + \sum_{l=1}^{d} \ddot{x}_t^{(l)i}))
= \sum_{i=1}^{N} (\ddot{x}_t^{i}, Q_t \ddot{x}_t^{i}) + \sum_{i=1}^{N} (\sum_{l=1}^{d} \ddot{x}_t^{(l)i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i}) + 2 \sum_{i=1}^{N} (\ddot{x}_t^{i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i}). \tag{7.15}
\]

We notice the last term is zero as show below:
\[
\sum_{i=1}^{N} (\ddot{x}_t^{i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i}) = \sum_{i=1}^{N} (\ddot{x}_t^{i} - \sum_{l=1}^{d} \ddot{x}_t^{(l)i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i})
= \sum_{i=1}^{N} (\ddot{x}_t^{i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i}) - \sum_{i=1}^{N} (\sum_{l=1}^{d} \ddot{x}_t^{(l)i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i})
= \sum_{i=1}^{N} (\ddot{x}_t^{i} x_t v_l (i) Q_t x_t v_l) - \sum_{i=1}^{N} (\sum_{l=1}^{d} x_t v_l v_l (i), Q_t \sum_{l=1}^{d} x_t v_l v_l (i))
= \sum_{i=1}^{d} (\sum_{l=1}^{d} v_l (i) x_t^{i}, Q_t x_t v_l) - \sum_{i=1}^{d} (\sum_{l=1}^{d} v_l (i) v_l (i)) (x_t v_l, Q_t x_t v_l)
= \sum_{i=1}^{d} (x_t v_l, Q_t x_t v_l) - \sum_{l=1}^{d} (x_t v_l, Q_t x_t v_l) = 0. \tag{7.16}
\]

Also we notice that
\[
\sum_{i=1}^{N} (\sum_{l=1}^{d} \ddot{x}_t^{(l)i}, Q_t \sum_{l=1}^{d} \ddot{x}_t^{(l)i}) = \sum_{i=1}^{d} \sum_{l=1}^{d} (\sum_{i=1}^{N} v_l (i) v_l (i)) (x_t v_l, Q_t x_t v_l)
= \sum_{i=1}^{d} (\sum_{l=1}^{d} v_l (i) v_l (i)) (x_t v_l, Q_t x_t v_l)
= \sum_{i=1}^{d} (\ddot{x}_t^{i}, Q_t \ddot{x}_t^{i}) \tag{7.17}
= \sum_{i=1}^{d} (\ddot{x}_t^{i}, Q_t \ddot{x}_t^{i})_I.
\]
Therefore we have
\[(x_t, Q_t x_t)_I = (\ddot{x}_t, Q_t \ddot{x}_t)_I + \sum_{l=1}^{d} (\dddot{x}_t^{(l)}, Q_l \dddot{x}_t^{(l)})_I.\]

\[\square\]

**Lemma 7.5.** Let \(G = \text{poly}_q(M)\) where \(\text{poly}_q(x) = q_0 + \sum_{i=1}^{d} q_i x^i\). Assume \(Q_t\) and \(M = [m_{ij}]\) are symmetric. Then

\[\begin{align*}
(x_t, Q_t x_t)_G &= q_0(\ddot{x}_t, Q_t \ddot{x}_t)_I + \sum_{l=1}^{d} \text{poly}_q(\lambda_l)(\dddot{x}_t^{(l)}, Q_l \dddot{x}_t^{(l)})_I \\
&= q_0 \sum_{i=1}^{N} (\dddot{x}_t^i, Q_t \dddot{x}_t^i) + \sum_{i=1}^{N} \sum_{l=1}^{d} \text{poly}_q(\lambda_l)(\dddot{x}_t^{(l)i}, Q_l \dddot{x}_t^{(l)i})
\end{align*}\]

(7.18)

where \(I \in \mathbb{R}^{n \times n}\) is the identity matrix.

**Proof.** The symmetry of \(M\) implies the symmetry of \(G\) and hence matrices \(G\) and \(M\) are diagonalizable with real eigenvalues. It is obvious that \(G\) and \(M\) share the same eigenvectors and \(G v_l = \lambda_l^k v_l\).

\[(x_t, Q_t x_t)_G = (x_t, Q_t x_t)_{q_0 I} + (x_t, Q_t x_t)_{\sum_{i=1}^{d} q_i M^i}
\]

(by Lemma 7.3 and Lemma 7.4)

\[\begin{align*}
&= (\ddot{x}_t, Q_t \ddot{x}_t)_{q_0 I} + \sum_{l=1}^{d} (\dddot{x}_t^{(l)}, Q_t \dddot{x}_t^{(l)})_{q_0 I} \\
&\quad + \sum_{i=1}^{N} \sum_{l=1}^{d} q_i \lambda_l^j (\dddot{x}_t^{(l)i}, Q_l \dddot{x}_t^{(l)i})_I \\
&= q_0 (\ddot{x}_t, Q_t \ddot{x}_t)_I + \sum_{l=1}^{d} q_0 (\dddot{x}_t^{(l)}, Q_t \dddot{x}_t^{(l)})_I \\
&\quad + \sum_{l=1}^{d} \sum_{i=1}^{d} q_i \lambda_l^j (\dddot{x}_t^{(l)i}, Q_l \dddot{x}_t^{(l)i})_I \\
&= q_0 (\ddot{x}_t, Q_t \ddot{x}_t)_I + \sum_{l=1}^{d} \text{poly}_q(\lambda_l)(\dddot{x}_t^{(l)}, Q_l \dddot{x}_t^{(l)})_I
\end{align*}\]

(by definition)

\[\begin{align*}
&= q_0 \sum_{i=1}^{N} (\dddot{x}_t^i, Q_t \dddot{x}_t^i) + \sum_{i=1}^{N} \sum_{l=1}^{d} \text{poly}_q(\lambda_l)(\dddot{x}_t^{(l)i}, Q_l \dddot{x}_t^{(l)i}).
\end{align*}\]

\[\square\]
Lemma 7.6. Let \( O = \text{poly}_r(M) \) where \( \text{poly}_r(x) = r_0 + \sum_{i=1}^{c_r} r_i x^i \). Assume \( R_t \) and \( M = [m_{ij}] \) are symmetric. Then

\[
(u_t, R_t u_t)_O = r_0 (\bar{u}_t, R_t \bar{u}_t)_I + \sum_{i=1}^{d} \text{poly}_r(\lambda_l)(\bar{u}_t^{(l)}(i), R_t \bar{u}_t^{(l)}(i))
\]

(7.19)

where \( I \in \mathbb{R}^{n \times n} \) is the identity matrix.

Proof. The proof follows the same lines of the proof for Lemma 7.5.

\[ \square \]

Lemma 7.7. Denote \( (z_t, H_t z_t)_I = \sum_{i=1}^{N} \|z_i^t\|_{H_t}^2 \). Then

\[
(z_t, H_t z_t)_I = (x_t, H_t x_t)_{M^2},
\]

(7.20)

where \( z_t = M x_t \) is the network field state matrix as \( z_t = (z_1^t, \ldots, z_N^t) \).

Proof.

\[
(z_t, H_t z_t)_I = \sum_{i=1}^{N} (z_i^t, H_t z_i^t)
\]

\[
= \sum_{i=1}^{N} (\sum_{j=1}^{N} x_j^t m_{ij}, H_t \sum_{k=1}^{N} x_k^t m_{ik})
\]

\[
= \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{i=1}^{N} m_{ij} m_{ik} (x_j^t, H_t x_k^t)
\]

(7.21)

(since \( M \) is symmetric)

\[
= \sum_{j=1}^{N} \sum_{k=1}^{N} [M^2]_{jk} (x_j^t, H_t x_k^t)
\]

\[
= (x_t, H_t x_t)_{M^2}.
\]

\[ \square \]

Base on Lemma 7.3 and Lemma 7.7, we know

\[
(z_t, H_t z_t)_I = \sum_{i=1}^{d} \lambda_i^2 (\bar{x}_t^{(l)}(i), H_t \bar{x}_t^{(l)}(i))_I = \sum_{i=1}^{N} \sum_{l=1}^{d} \lambda_i^2 (\bar{x}_t^{(l)}(i), H_t \bar{x}_t^{(l)}(i)).
\]

(7.22)
7 Team Optimal Linear Quadratic Regulation (LQR) on Networks

7.3.3 Team Optimal Solutions

**Proposition 7.1.** Under Assumptions 7.1 and 7.2, the original network LQR problem defined by (7.1) and (7.5) can be solved by solving the following LQR problems locally for all agents: 1) the LQR problem for the auxiliary system

\[
\begin{aligned}
\dot{x}_t^i &= A_t x_t^i + B_t \bar{u}_t^i, \\
\dot{J}(\bar{u}_t^i, x_t^i) &= \int_0^T \dot{c}_t(\bar{u}_t^i, x_t^i) dt + \dot{c}_T(x_T^i), \\
&= q_0 \| \bar{x}_t^i \|_{Q_t}^2 + r_0 \| \bar{u}_t^i \|_{R_t}^2, \quad \text{and} \quad \dot{c}_T(x_T^i) = q_0 \| x_T^i \|_{Q_T}^2;
\end{aligned}
\]  

(7.23)

where

\[
\dot{c}_t(\bar{u}_t^i, x_t^i) = q_0 \| \bar{x}_t^i \|_{Q_t}^2 + r_0 \| \bar{u}_t^i \|_{R_t}^2,
\]

and 2) the LQR problems in the eigendirections

\[
\begin{aligned}
\dot{x}_t^{(l)i} &= (A_t + D_t \lambda_t) x_t^{(l)i} + (B_t + E_t \lambda_t) \bar{u}_t^{(l)i}, \\
\dot{J}(\bar{u}_t^{(l)i}, \bar{x}_t^{(l)i}) &= \int_0^T \dot{c}_t^{(l)}(\bar{u}_t^{(l)i}, \bar{x}_t^{(l)i}) dt + \dot{c}_T^{(l)}(\bar{x}_T^{(l)i}), \\
&= \left( \text{poly}_q(\lambda_t) \| x_t^{(l)i} \|_{Q_t}^2 + \lambda_t^2 \| \bar{x}_t^{(l)i} \|_{H_t}^2 \right) + \text{poly}_r(\lambda_t) \| \bar{u}_t^{(l)i} \|_{R_t}^2,
\end{aligned}
\]  

(7.24)

where

\[
\begin{aligned}
\dot{c}_t^{(l)}(\bar{u}_t^{(l)i}, \bar{x}_t^{(l)i}) &= \left( \text{poly}_q(\lambda_t) \| x_t^{(l)i} \|_{Q_t}^2 + \lambda_t^2 \| \bar{x}_t^{(l)i} \|_{H_t}^2 \right), \\
\dot{c}_T^{(l)}(\bar{x}_T^{(l)i}) &= \left( \text{poly}_q(\lambda_t) \| x_T^{(l)i} \|_{Q_T}^2 + \lambda_t^2 \| \bar{x}_T^{(l)i} \|_{H_T}^2 \right).
\end{aligned}
\]

**Proof.** The instantaneous cost can be expanded as

\[
\begin{aligned}
c_t(u_t, x_t, z_t) &= q_0 \sum_{i=1}^N \left( \bar{x}_t^i, Q_t \bar{x}_t^i \right) + \sum_{i=1}^N \sum_{l=1}^d \text{poly}_q(\lambda_t) \left( \bar{x}_t^{(l)i}, Q_t \bar{x}_t^{(l)i} \right) \\
&\quad + r_0 \sum_{i=1}^N \left( \bar{u}_t^i, R_t \bar{u}_t^i \right) + \sum_{i=1}^N \sum_{l=1}^d \text{poly}_r(\lambda_t) \left( \bar{u}_t^{(l)i}, R_t \bar{u}_t^{(l)i} \right) \\
&\quad + \sum_{i=1}^N \sum_{l=1}^d \lambda_t^2 \left( \bar{x}_t^{(l)i}, H_t \bar{x}_t^{(l)i} \right) \\
&= q_0 \sum_{i=1}^N \| \bar{x}_t^i \|_{Q_t}^2 + r_0 \sum_{i=1}^N \| \bar{u}_t^i \|_{R_t}^2 \\
&\quad + \sum_{i=1}^N \sum_{l=1}^d \left( \text{poly}_q(\lambda_t) \| \bar{x}_t^{(l)i} \|_{Q_t}^2 + \lambda_t^2 \| \bar{x}_t^{(l)i} \|_{H_t}^2 \right) \\
&\quad + \sum_{i=1}^N \sum_{l=1}^d \text{poly}_r(\lambda_t) \| \bar{u}_t^{(l)i} \|_{R_t}^2.
\end{aligned}
\]  

(7.25)
Similarly the terminal cost can be expanded as
\[
  c_T(x_T, z_T) = q_0 \sum_{i=1}^{N} \|\dddot{x}_i^T\|^2_{Q_T} + \sum_{i=1}^{N} \sum_{l=1}^{d} \left( \text{poly}_q(\lambda_l) \|\dddot{x}_T^{(l)i}\|^2_{Q_T} + \lambda_l^2 \|\dddot{x}_T^{(l)i}\|^2_{H_T} \right). \tag{7.26}
\]

Therefore, we can separate the cost as follows:
\[
  c_t(u_t, x_t, z_t) = \sum_{i=1}^{N} \left( \hat{c}_t(\dddot{u}_t^i, \dddot{x}_t^i) + \sum_{l=1}^{d} \hat{c}_t^{(l)i}(\dddot{u}_t^{(l)i}, \dddot{x}_t^{(l)i}) \right) \tag{7.27}
\]
and
\[
  c_T(x_T, z_T) = \sum_{i=1}^{N} \left( \hat{c}_T(\dddot{x}_T^i) + \sum_{l=1}^{d} \hat{c}_T^{(l)i}(\dddot{x}_T^{(l)i}) \right). \tag{7.28}
\]

Hence
\[
  \tilde{J}(u, x) = \sum_{i=1}^{N} \left( \tilde{J}(\dddot{u}^i, \dddot{x}^i) + \sum_{l=1}^{d} \tilde{J}^{(l)i}(\dddot{u}^{(l)i}, \dddot{x}^{(l)i}) \right). \tag{7.29}
\]

Since \( \tilde{J}(\dddot{u}^i, \dddot{x}^i) \) and \( \tilde{J}^{(l)i}(\dddot{u}^{(l)i}, \dddot{x}^{(l)i}) \) are non-negative, minimizing the costs \( \tilde{J}(\dddot{u}^i, \dddot{x}^i) \) and \( \tilde{J}^{(l)i}(\dddot{u}^{(l)i}, \dddot{x}^{(l)i}) \) for all \( i \in \{1, \ldots, N\} \) and all \( l \in \{1, \ldots, d\} \) is equivalently minimizing the cost \( \tilde{J}(u, x) \). Moreover, for each agent, the dynamics is decoupled into the auxiliary system dynamics (7.23) and dynamics in the eigenvector directions (7.24). Therefore, we can optimally solve the original LQR problem as decoupled LQR problems with decoupled dynamics and decoupled costs in the decentralized way.

Let \( \text{Ricc}(A_t, B_t, Q_t, Q_T, R_t) \) denote the continuous time Riccati equation
\[
  \dot{P}_t = A_t^T P_t + P_t A_t - P_t B_t R_t^{-1} B_t^T P_t + Q_t, \quad P_0 = Q_T \tag{7.30}
\]
over the time horizon \([0, T]\).

**Theorem 7.1.** Under Assumptions 7.1 and 7.2, the team optimal control law for each agent is given by
\[
  u_t^i = -R_t^{-1} \left( \frac{B_t^T}{r_0} \dddot{P}_t^i + \sum_{l=1}^{d} \frac{(B_t + E_t \lambda_l) T}{\text{poly}_r(\lambda_l)} \frac{P^{(l)i} \dddot{x}_t^{(l)i}}{T_{T-t} \dddot{x}_t^i} \right) \tag{7.31}
\]
with \( \dddot{P}_t^i \) as the solution to the Riccati equation
\[
  \text{Ricc}(A_t, B_t, q_0 Q_t, q_0 Q_T, r_0 R_t), \tag{7.32}
\]
and $P^{(l)i}$ as the solution to the Riccati equation

$$Ricc([A_t + D_t \lambda_t], [B_t + E_t \lambda_t], [poly_q(\lambda_t)Q_t + \lambda_t^2 H_t], [poly_q(\lambda_t)Q_T + \lambda_T^2 H_T], [poly_r(\lambda_t)R_t])$$.

(7.33)

Furthermore, there exists a unique team optimal control law that solves the original LQR problem defined by (7.2) and (7.5).

**Proof.** Following the classic LQR theory [74], the optimal control solution for the problem in (7.23) is given by

$$\ddot{u}^i_t = -(r_0 R_t)^{-1} B^T P^{(l)i}_T \ddot{x}^i_t$$;

the optimal solution for the problem in (7.24) is given by

$$\ddot{u}^{(l)i}_t = -(poly_r(\lambda_t) R_t)^{-1} (B_t + E_t \lambda_t)^T P^{(l)i}_T \ddot{x}^{(l)i}_t$$

with $\dot{P}^i$ as the solution to the continuous time Riccati equation (7.32), and $P^{(l)i}$ as the solution to the continuous time Riccati equation (7.33). Together with the following relation

$$u^i_t = \ddot{u}^i_t + \sum_{l=1}^d \ddot{u}^{(l)i}_t$$;

we obtain the result on the optimal control law.

Under the **Assumptions 7.1** and **7.2**, the original LQR problem defined by (7.2) and (7.5) has a unique solution and each of the LQR problems in (7.23) and (7.24) has a unique solution. Therefore, the team optimal solution exists and is unique for the original LQR problem.  

7.4 Numerical Examples

The parameters and functions in the numerical examples are as follows: $A_t = 2; B_t = 4; D_t = 3; E_t = 0; Q_t = 7; Q_T = 5; H_t = 6; H_T = 2; R_t = 3; poly_r(s) = 1 + 2s + s^2; poly_q(s) = 1 + 2s + s^2$.

7.4.1 Complete Bipartite Networks

Many of the demand-supply networks, for instance, Uber or Didi networks, can be represented by bipartite networks. In this example, we use a complete bipartite network of 20 nodes with the
adjacency matrix given as in Figure 7.1(b). It is known that any complete bipartite graph has only two non-zero eigenvalues and in this example the two non-zero eigenvalues are $-30$ and $30$.

Fig. 7.1 A complete bipartite network example

![Fig. 7.1](image)

**Fig. 7.2** A numerical example on a complete bipartite network (with 20 agents), where each agent is optimally solving its local LQR problems and they together achieve the global optimal performance

7.4.2 Cosinusoidal Networks

Cosinusoidal networks are generated from a cosinusoidal graphon which can be represented by two non-zero eigenvalues and the corresponding cosinusoidal eigenfunctions in $L^2[0, 1]$. In this example, the cosinusoidal network can be represented by a spectral representation with two non-zero eigenvalues (i.e. $9.0317, 10.9683$). See Figure 7.3(b) for the adjacency matrix.
7 Team Optimal Linear Quadratic Regulation (LQR) on Networks

7.5 Discussion

To solve the team problem, each agent needs to solve:

a) The LQR problem for the auxiliary system, which is \( n \) dimensional;

b) The LQR problems in each of the \( d \) eigenvector directions, each of which is \( n \) dimensional.

The computational complexity of the decentralized team optimal control law at each agent depends on the dimension of the network eigenspace corresponding to all the non-zero eigenvalues, and it does not depend directly on the size of the network (or equivalently the size of the agent population).

In the mean field team LQR case, the only non-zero eigenvalue is 1 and the corresponding eigenvector is \( 1 \in \mathbb{R}^N \). The problem reduces to the problem of solving two decoupled \( n \)-dimensional LQR problems.

Information required by each agent to generate the control action is:

1) All non-zero eigenvalues of the underlying network to obtain the Riccati equations for the decoupled LQR problems.

2) Eigenvectors corresponding to each non-zero eigenvalue and initial states of all agents, to determine the initial states of the decoupled systems.

3) Its own state observation

Alternatively 2) can be replaced by the information on the projections of initial states of all agents onto all different eigenvector directions and the value of each eigenvector evaluated at the agent’s index.
We note that this work studies the finite horizon problem and the infinite horizon problem can be solve when the global system is stabilizable. In that case, each agent only need to solve the algebraic Riccati equation for the auxiliary system and the corresponding algebraic Riccati equation in each eigendirection.

This work can be directly generalized to solve the LQG problems for finite dimensional networks of stochastic systems with additive noise and exact state observation, based on the certainty equivalence principle [75].

Compared to the case with weighted mean field couplings in [51], this allows more variations in the cost coupling and the coupling of the dynamics. Moreover, we consider the weighted mean field as a whole network effect rather than an agent-to-agent effect. This network effect is then projected into orthogonal eigendirections of the network and enables the characterization of the weighted mean field coupling in the eigendirections, which is intuitively simpler and leads to simpler characterization of the optimal solutions.

This work illustrate that the fundamental property of decoupling an LQR problem on networks into an auxiliary LQR problem and LQR problems relying on global information is the network structure preserved in both the dynamics and the cost functions. However, we note that the limitation of this work lies in the assumption that agents have similar dynamics, apart from the difference in the network coupling.
Part III

Centrality for Networks of Dynamical Systems
Chapter 8

Consensus-induced Centrality for Networks of Dynamical Systems

In multi-agent systems where agents are coupled through the underlying networks and follow the consensus dynamics, the ability for agents on networks to reach consensus depends on the network structure, and more specifically depends on the spectra of the Laplacian of the networks. Each agent on the network plays a role in aggregating the information for the agent population to reach consensus. To evaluate the influence that one agent (or one group of agents) would have on the network to reach consensus, we propose the consensus-induced centrality measure. Examples of consensus-induced centrality analysis for different types of networks are given. The consensus-induced centrality has possible applications in many network systems, such as social networks, power grid stability and wireless sensor networks, etc.

8.1 Introduction

Networks are everywhere ranging from social networks, biological networks to engineering networks. Underlying most networks, there are dynamical processes. To fully understand and control these networks, it is important to study both the network structure and the underlying dynamics. For instance, in the study of consensus and synchronization on networks, the dynamics is specified locally following consensus protocol [76, 77, 11] and the key structure property is represented by the spectrum of the graph Laplacian.

Most graph theoretic measures for node importance (centrality) on networks tend to focus
on the structural importance of nodes in the underlying graph [52]. However, to study network dynamical systems, it is more meaningful to represent the node importance considering also the underlying dynamical process on the networks.

This work provides a measure for nodes representing their importance in a network under a dynamical setting. The underlying methodology is to evaluate the impact of the removal of an agent on the ability of the network to reach consensus or synchronization. The importance (centrality) values of the agents are ranked according to their impact factors. Further, since the ability to reach consensus is an invariant of the network system represented by the spectrum of the underlying graph, the change in the graph spectrum is used to evaluate the change of the ability to reach consensus. We define the consensus-induced centrality for a network agent as the change of the underlying graph spectrum with the removal of that agent. See [52] for general induced centrality measures on graphs.

For unweighted graphs, the centrality based on the change of algebraic connectivity with removal of nodes is called the vertex-deleted centrality measure and is studied in [78] where tight lower bound and upper bound are established. The results in our paper apply to general weighted graphs, and lower bounds and upper bounds based on local connection weights are proven. These results are straightforward to generalize to centrality measures for groups of nodes. Most importantly, we give a clear interpretation of the induced centrality measure in the context of multi-agent network systems.

In this work, the centrality measure for connected networks is considered, however, it is straightforward to generalize the results to disconnected networks.

**Basic notations**

Let $G_n = (V_n, E, W)$ denote an undirected positively weighted graph with the node set $V_n$ where $n$ is the size of the graph, i.e. $|V_n| = n$, the edge set $E$ and the positive edge weight set $W$. Simple graphs (i.e. unweighted graphs with no self-loops and no multiple edges) are a special case where $W$ only takes 0 and 1 as its elements. Let $A(G_n)$ represent the symmetric adjacency matrix of the graph $G_n$ where the $ij^{th}$ element of $A(G_n)$ is $a_{ij}$ if there is an undirected connection between the $i^{th}$ node and the $j^{th}$ node with edge weight $a_{ij}$ in the graph $G_n$ and zero otherwise. Let $1 = [1, 1, ..., 1]^T \in \mathbb{R}^n$. Let $L(G_n)$ be the Laplacian of graph $G_n$, i.e. $L(G_n) = \text{diag}(A(G_n)1) - D$.
A(G_n). Let the eigenvalues of \( L(G_n) \) be ordered as

\[
\lambda_1(G_n) \leq \lambda_2(G_n) \leq \cdots \leq \lambda_n(G_n).
\]

We use the word “network” to refer to a network system with an underlying graph and an underlying dynamics which in the present case is the consensus or synchronization dynamics.

### 8.2 Consensus and Synchronization on Networks

#### 8.2.1 Consensus on Networks

Consider a network with \( n \) agents each of which follows the local dynamics and protocol

\[
\dot{x}_i(t) = u_i(t), \quad (8.1)
\]

\[
u_i(t) = -\sum_{j \in N_i} a_{ij}(x_i(t) - x_j(t)), \quad (8.2)
\]

\[x_i(t), u_i(t) \in \mathbb{R}, \quad i \in \{1, \ldots, n\}, \quad t \in [0, \infty),
\]

where \( x_i, u_i \) and \( N_i \) represent respectively the state, the control input and the neighbourhood set of agent \( i \) on the network, and \( a_{ij} \) denotes the edge weight connecting agent \( i \) and agent \( j \). Denote the weighted adjacency matrix by \( A = [a_{ij}] \) and denote the Laplacian matrix by \( L \). The system is said to reach consensus at time \( t \) if \( x_1(t) = x_2(t) = \cdots = x_n(t) \) where \( t \) could be either finite or infinite.

The system dynamics can be written in the following compact form

\[
\dot{X}(t) = -LX(t), \quad (8.3)
\]

where \( X(t) = [x_1(t), \ldots, x_n(t)]^T \). The solution of the system is given by

\[
X(t) = e^{-Lt}X_0 = Ve^{-\Lambda t}V^TX_0 \\
= e^{-\lambda_1 t}(v_1^TX_0)v_1 + e^{-\lambda_2 t}(v_2^TX_0)v_2 + \cdots + e^{-\lambda_n t}(v_n^TX_0)v_n \quad (8.4)
\]

where \( X(0) = X_0 \) and \( L = V\Lambda V^T = [v_1, \ldots, v_n] \text{diag}(\lambda_1, \ldots, \lambda_n)[v_1, \ldots, v_n]^T \), with \( v_1, \ldots, v_n \) as the normalized eigenvectors of \( L \) that correspond to the eigenvalues \( \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \) respectively. Based on the definition of \( L \), it is obvious that \( \lambda_1 = 0 \) always holds. Under the proposed local
control protocol (8.2), the system (8.1) can reach consensus for any initial condition as $t \to \infty$ if and only if $\lambda_2 > 0$ (which also means the underlying graph is connected).

From (8.4) we know that in the direction of $v_1$ the system trajectory will always have a length $v_1^T X_0$ since $\lambda_1 = 0$ always holds for $L$, and the convergence rate of the consensus problem is determined by the second smallest eigenvalue $\lambda_2$ of the Laplacian. See [79] for more details.

### 8.2.2 Synchronization on Networks

A typical dynamical network model [80, 81] for synchronization analysis is as follows. The dynamical network contains $N$ identical nodes which are diffusively and linearly coupled. On each node there exist an $n$-dimensional dynamical system.

\[
\dot{x}_i = f(x_i) - c \sum_{j=1}^{N} L_{ij} \Gamma x_j, \quad i = 1, 2, \ldots, N \tag{8.5}
\]

with $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T \in \mathbb{R}^n$ as the state vector of node $i$, constant $c > 0$ as the coupling strength, and $\Gamma \in \mathbb{R}^{n \times n}$ a constant matrix with only 0-1 elements as variables’ coupling structure. $L_{ij} = L_{ji} = -1(i \neq j)$ if node $i$ and node $j$ are connected, otherwise $L_{ij} = L_{ji} = 0(i \neq j)$. $L_{ii} = d_i$ where $d_i$ is the connection degree of node $i$. Note that $L$ is the Laplacian of the underlying graph representing connection between agents. The objective is to (asymptotically) synchronize the networks, i.e. to achieve

\[
\lim_{t \to \infty} \|x_i(t) - s(t)\| = 0, \quad i = 1, 2, \ldots, N, \tag{8.6}
\]

where $s(t)$ is the solution to $\dot{s}(t) = f(s(t))$. Assume the underlying network is connected, that is, the adjacency matrix $[a_{ij}]$ is irreducible. This implies $0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_N$. The synchronization state is exponentially stable if

\[
c \geq T \lambda_2^{-1} \tag{8.7}
\]

where $T > 0$ is a constant determined by the dynamics of an isolated node and the state mapping structure $\Gamma$ and $c > 0$ is the coupling strength [19, 81]. This shows that the second smallest eigenvalue of the Laplacian of the underlying graph determines the synchronizability of the dynamical network (8.5).

There are other variations of network consensus and synchronization problems. The key index to these problems is the spectrum of the Laplacian of the underlying networks. The first
eigenvalue of the graph Laplacian is always zero. The second smallest eigenvalue of the graph Laplacian (also called algebraic connectivity), determines the rate of convergence of consensus and synchronization capability of the network.

8.2.3 Consensibility and Relative Consensibility of Network Systems

Consider a network system with an underlying connected graph $G_n$. The consensibility (i.e. the ability to reach consensus) of the network system is defined as the second smallest eigenvalue of the underlying graph Laplacian, i.e. the algebraic connectivity of the underlying graph, denoted by $\lambda_2(G_n)$. The relative consensibility of the network system is defined as $\frac{\lambda_2(G_n)}{n}$ representing the ability of the network reaching consensus compared to the fully connected network on the same node set.

8.2.4 Properties of Algebraic Connectivity

As is known, the edge connectivity $e(G)$ of an undirected simple graph $G = (V, E)$ is the smallest cardinality of a subset $E_1 \subset E$ satisfying the property that $G_1 = (V, E \setminus E_1)$ is not connected. Similarly, the node connectivity $\nu(G)$ of $G = (V, E)$ is the smallest cardinality of a subset $V_1 \subset V$ having the property that the subgraph $G_1$ generated by removing $V_1$ from $G$ is not connected. The minimum degree $\delta(G_n)$ of graph $G_n$ is the minimum degree of its nodes.

**Theorem 8.1** ([82]). Let $G_n(n \geq 2)$ be a non-complete simple graph with node connectivity $\nu(G_n)$, edge connectivity $e(G_n)$ and the minimum degree $\delta(G_n)$. Then

$$\lambda_2(G_n) \leq \nu(G_n) \leq e(G_n) \leq \delta(G_n).$$

**Theorem 8.2** ([83]). Let $G_n(n \geq 2)$ be a connected simple graph with $n$ nodes. Then

$$2(1 - \cos \frac{\pi}{n}) \leq \lambda_2(G_n) \leq n$$

with the left equality if and only if $G_n$ is a path.

**Theorem 8.3** ([82]). Let $G_n(n \geq 2)$ be a simple graph. Then $\lambda_2(G_n) > 0$ if and only if $G_n$ is connected.

This result generalizes to undirected positively weighted graphs.
Theorem 8.4 (see e.g. [84]). For an undirected and positively weighted graph $G_n = (V, E, W)$, $n \geq 2$, $\lambda_2(G_n) > 0$, if and only if $G_n$ is connected.

For the purpose of convenience, we attach a proof here.

Proof. Since the eigenvalues of $L(G_n)$ are non-negative, we only need to prove that $\lambda_2(G_n) = 0$ if and only if the graph $G_n$ is not connected.

First, we prove that $G_n$ is not connected implies $\lambda_2(G_n) = 0$. If $G_n$ is not connected, it consists at least two separated subgraph, denoted by $G^1$ and $G^2$. Then the Laplacian $L(G_n)$ consists of two blocks $L(G^1)$ and $L(G^2)$, each of which has one zero eigenvalue. Therefore $L(G_n)$ has at least two zero eigenvalues and $\lambda_2(G_n) = 0$.

Second, we prove that $\lambda_2(G_n) = 0$ implies $G_n$ is not connected. The algebraic connectivity of a weighted graph can be written as follows [84] :

$$\lambda_2(G_n) = \min_{x \neq 0, x \perp 1} \frac{\sum_{(i,j) \in E} a_{ij} (x_i - x_j)^2}{\sum_{i=1}^n x_i^2},$$  

(8.8)

where $a_{ij}$ is the positive edge weight between node $i$ and node $j$. $\lambda_2(G_n) = 0$ implies that there exist $x \neq 0$ such that $\sum_{i=1}^n x_i = 0$ and $\sum_{(i,j) \in E} a_{ij} (x_i - x_j)^2 = 0$. If $G_n$ is connected, $\sum_{(i,j) \in E} a_{ij} (x_i - x_j)^2 = 0$ implies $x_1 = x_2 = \ldots = x_n$. Together with $\sum_{i=1}^n x_i = 0$, this implies $x_1 = x_2 = \ldots = x_n = 0$, which contradicts the fact that $x \neq 0$. Therefore the underlying graph $G_n$ is not connected.

\[
\square
\]

8.3 Consensus-induced Centrality Measure

The importance of each node of the network in the process to reach consensus and synchronization can be reflected by the change of the graph algebraic connectivity if that particular node is removed. Based on this intuition, we define the consensus-induced centrality measure for networks. There can be two definitions: absolute consensus-induced centrality (ACIC) measure and relative consensus-induced centrality (RCIC) measure.

8.3.1 Absolute Consensus-induced Centrality Measure

We define the absolute consensus-induced centrality measure in the following. The subgraph of $G_n = (V_n, E, W)$ generated by removing node $v \in V_n$ of $G_n$ is denoted by $G^{(v)}_{n-1}$. For a node $v$ in
a graph $G_n$, the absolute consensus-induced centrality (ACIC) value is given by

$$ACIC_v = \lambda_2(G_n) - \lambda_2(G_{n-1}^{(v)}).$$

This value is the value drop of the second smallest eigenvalue of the graph Laplacian after removing node $v$ in the $G_n$. It reflects the drop of the rate of reaching consensus on $G_n$.

### 8.3.2 Relative Consensus-induced Centrality Measure

As we know that for a complete graph of size $n > 1$, the eigenvalues of its Laplacian are \{0, n, \cdots n\}. Since the two graphs $G_n$ and $G_{n-1}^{(v)}$ are of different sizes and the size of the network determines the maximum value of eigenvalue that graph achieve, it is meaningful to compare the relative change of the eigenvalues. Therefore we define the relative consensus-induced centrality (RCIC) value for node $v$ of graph $G_n$ as follows:

$$RCIC_v = \frac{\lambda_2(G_n)}{n} - \frac{\lambda_2(G_{n-1}^{(v)})}{n-1}.$$

This value measures the change of the ratio between the algebraic connectivity and the maximum algebraic connectivity of graphs on the same node set by removing node $v$ in the graph $G_n$.

According to the definition of RCIC, if the network is sparse and network size is large, RCIC can be very small. RCIC and ACIC would represent the same property if the network is very large. But for small networks, RCIC and ACIC represent relatively different properties.

### 8.4 Basic Results on Consensus-induced Centrality Measure

#### 8.4.1 Upper Bound of ACIC

Based on the Cauchy Interlace Theorem and the Courant-Fischer Theorem, we prove the following results on the change of eigenvalues of the graph Laplacian when removing a node from the graph.

**Lemma 8.1.** Consider a graph $G_n = (V_n, E, W)$ and its subgraph $G_{n-1}^{(v)}$ generated by removing node $v$ from $G_n$. Then

$$\lambda_k(G_{n-1}^{(v)}) \geq \lambda_k(G_n) - \alpha_{\text{max}}^v, \quad 1 \leq k \leq n - 1,$$
where $\alpha^v_{\text{max}}$ is maximum of the connection weights between $v$ and all other nodes in $G_n$.

The proof follows the proof idea of [85] and extends the result to weighted graphs.

**Proof.** Denote the Laplacian of $G_n$ by $L_n$ and denote the Laplacian of the graph $G^{(v)}_{n-1}$ by $L^{(v)}_{n-1}$. Removing the row and column of $L_n$ that correspond to node $v$, we have the principle sub-matrix denoted by $P^{(v)}_{n-1}$. Label the nodes of $G_n$ by $\{1, 2, ..., n\}$. Suppose the label of node $v$ is $i$. We note that $P^{(v)}_{n-1}$ and $L^{(v)}_{n-1}$ are different only in the diagonal terms as follows:

$$L^{(v)}_{n-1} = P^{(v)}_{n-1} - \text{diag}(\alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_n),$$  \hspace{1cm} (8.11)

where $\alpha_k$ is the connection weight between the $i^{th}$ (i.e. node $v$) and $k^{th}$ node. Then

$$\alpha^v_{\text{max}} = \max\{\alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_n\}.$$  

For simplicity, let $\Lambda_v = \text{diag}(\alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_n)$. Then

$$L^{(v)}_{n-1} = P^{(v)}_{n-1} - \Lambda_v.$$  \hspace{1cm} (8.12)

Fix $k \in \{1, 2, ..., n-1\}$. We let $U_k$ be the set of subspaces $\{U \subset \mathbb{R}^n \text{ with } \dim(U) = n - k + 1\}$. It follows from the Courant-Fisher Theorem that,

$$\lambda_k(P^{(v)}_{n-1}) = \max_{U \in U_k} \min_{x \in U \setminus \{0\}} \left\{ \frac{x^T P^{(v)}_{n-1} x}{x^T x} \right\}. $$

Substituting $P^{(v)}_{n-1}$ by $L^{(v)}_{n-1} + \Lambda_v$, we yield

$$\lambda_k(P^{(v)}_{n-1}) = \max_{U \in U_k} \min_{x \in U \setminus \{0\}} \left\{ \frac{x^T (L^{(v)}_{n-1} + \Lambda_v) x}{x^T x} \right\}. $$
Further simplifying the right hand side, we obtain

\[
\lambda_k(P_{n-1}^{(v)}) \leq \max \min_{U \in \mathcal{U}_k, x \in U} \left\{ \frac{x^T (L_{n-1}^{(v)} + \alpha_{\max}^v) x}{x^T x}; x \neq 0 \right\}
\]

\[
= \max \min_{U \in \mathcal{U}_k, x \in U} \left\{ \frac{x^T L_{n-1}^{(v)} x}{x^T x} + \alpha_{\max}^v; x \neq 0 \right\}
\]

\[
= \max \min_{U \in \mathcal{U}_k, x \in U} \left\{ \frac{x^T L_{n-1}^{(v)} x}{x^T x}; x \neq 0 \right\} + \alpha_{\max}^v
\]

\[
= \lambda_k(L_{n-1}^{(v)}) + \alpha_{\max}^v.
\]

By the Cauchy Interlace Theorem, we have

\[
\lambda_k(L_n) \leq \lambda_k(P_{n-1}^{(v)}).
\]

Therefore we obtain the following inequality

\[
\lambda_k(L_{n-1}^{(v)}) \geq \lambda_k(L_n) - \alpha_{\max}^v.
\]

Hence

\[
\lambda_k(G_{n-1}^{(v)}) \geq \lambda_k(G_n) - \alpha_{\max}^v, \quad 1 \leq k \leq n - 1.
\]

We can see that the maximum drop in the relative eigenvalues of the Laplacians when removing a node \(v\) from graph \(G_n\) is upper bounded by \(\alpha_{\max}^v\), i.e. the maximum of the connection weights between \(v\) and all other nodes in \(G_n\).

Based on this result and the definition of ACIC, we obtain an upper bound for the ACIC of the nodes.

**Theorem 8.5.** Consider a network \(G_n = (V_n, E, W)\). For any \(v \in V_n\)

\[
ACIC_v \leq \alpha_{\max}^v,
\]

where \(\alpha_{\max}^v\) is maximum of the connection weights between \(v\) and all other nodes in \(G_n\).
Corollary 8.1. Consider a network $G_n = (V_n, E, W)$. Then

$$\forall v \in V_n, \quad ACIC_v \leq \alpha_{\text{max}},$$

where $\alpha_{\text{max}}$ is the maximum edge weight in $G_n$.

$\alpha_{\text{max}}$ provides the upper bound for $ACIC$ and hence a robustness estimate for synchronization when the removal of one node is possible. Specifically, if $\lambda_2 - \alpha_{\text{max}} > 0$, then the coupling strength

$$c \geq T(\lambda_2 - \alpha_{\text{max}})^{-1},$$

based on (8.7), ensures synchronization under the removal of any agent in the network.

8.4.2 Lower Bound of ACIC

Lemma 8.2. Consider a graph $G_n = (V_n, E, W)$ and its subgraph $G_{n-1}^{(v)}$ generated by removing node $v$ from $G_n$. Then

$$\lambda_k(G_{n-1}^{(v)}) \leq \lambda_{k+1}(G_n) - \alpha^v_{\text{min}} \quad 1 \leq k \leq n - 1,$$

where $\alpha^v_{\text{min}}$ is minimum of the connection weights (including zero weights) between $v$ and all other nodes in $G_n$.

Proof. Following the same set up in the proof of Lemma 8.1, we have

$$\lambda_k(P_{n-1}^{(v)}) = \max_{U \in \mathcal{U}_k} \min_{x \in \mathcal{U}} \left\{ \frac{x^T(L_{n-1}^{(v)} + \Lambda_v)x}{x^T x}; x \neq 0 \right\}.$$
Further simplifying the right hand side, we obtain

\[
\lambda_k(P_{n-1}^{(v)}) \geq \max_{U \in \mathcal{U}} \min_{x \in U} \left\{ \frac{x^T (L_{n-1}^{(v)} + \alpha_{\min}^v) x}{x^T x}; x \neq 0 \right\}
\]

\[
= \max_{U \in \mathcal{U}} \min_{x \in U} \left\{ \frac{x^T L_{n-1}^{(v)} x}{x^T x} + \alpha_{\min}^v; x \neq 0 \right\}
\]

\[
= \max_{U \in \mathcal{U}} \min_{x \in U} \left\{ \frac{x^T L_{n-1}^{(v)} x}{x^T x}; x \neq 0 \right\} + \alpha_{\min}^v
\]

\[
= \lambda_k(L_{n-1}^{(v)}) + \alpha_{\min}^v.
\]

By the Cauchy Interlace Theorem, we have

\[
\lambda_{k+1}(L_n) \geq \lambda_k(P_{n-1}^{(v)}).
\]

Therefore we obtain the following inequality

\[
\lambda_{k+1}(L_n) \geq \lambda_k(L_{n-1}^{(v)}) + \alpha_{\min}^v.
\]

Hence

\[
\lambda_k(G_{n-1}^{(v)}) \leq \lambda_{k+1}(G_n) - \alpha_{\min}^v, \quad 1 \leq k \leq n - 1.
\]

**Theorem 8.6.** Consider a network \(G_n = (V_n, E, W)\). Then, for any \(v \in V_n\)

\[
ACIC_v \geq \alpha_{\min}^v + \lambda_2(G_n) - \lambda_3(G_n),
\]

where \(\alpha_{\min}^v\) is minimum of the connection weights between \(v\) and all other nodes in \(G_n\).

**Corollary 8.2.** Consider a network \(G_n = (V_n, E, W)\). Then

\[
\forall v \in V_n, \quad ACIC_v \geq \alpha_{\min} + \lambda_2(G_n) - \lambda_3(G_n),
\]

where \(\alpha_{\min}\) is the minimum connection weight (including zero weights) over all nodes.
Corollary 8.3. Consider a network \( G_n = (V_n, E, W) \). Then
\[
\forall v \in V_n, \quad ACIC_v \geq \lambda_2(G_n) - \lambda_3(G_n).
\]

8.5 Generalization to Centrality Measure for Node Sets

8.5.1 Consensus-induced Centrality Measures for Node Sets

Denote the subgraph of \( G_n = (V_n, E, W) \) generated by removing node set \( S \subset V_n \) with \( |S| = m \) of \( G_n \) by \( G_n^S \). Similar to (8.9), for a node set \( S \) in a graph \( G_n \), the absolute consensus-induced centrality (ACIC) value is given by
\[
ACIC_S = \lambda_2(G_n) - \lambda_2(G_n^S).
\]

(8.13)

Similar to (8.10), we define the relative consensus-induced centrality (RCIC) value for a node set \( S \) of graph \( G_n \) as follows:
\[
RCIC_S = \frac{\lambda_2(G_n)}{n} - \frac{\lambda_2(G_n^S)}{n-1}.
\]

(8.14)

8.5.2 Spectrum Change with the Removal of Node Sets

To prove bounds for the spectrum change with the removal of a set of nodes, we use the following extension of the Cauchy Interlace Theorem.

An \((n - m)^{th}\) order principle sub-matrix of \( A \in \mathcal{R}^{n \times n} \) is defined as the \( m \times m \) matrix formed by deleting \( n - m \) rows of \( A \), and the same \( n - m \) columns of \( A \).

Lemma 8.3. Let \( A \) be a Hermitian matrix of order \( n \), and let \( B \) be the \((n - m)^{th}\) order principle sub-matrix of \( A \) with \( 1 \leq m < n \). If \( \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n-1} \leq \lambda_n \) lists the eigenvalues of \( A \) and \( \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{n-2} \leq \mu_{n-m} \) the eigenvalues of \( B \), then for \( k \in \{1, 2, \ldots, n - m\} \)
\[
\lambda_k \leq \mu_k \leq \lambda_{k+m}.
\]

(8.15)

Proof. The proof idea is simply to apply the Cauchy Interlace Theorem multiple times. \(\square\)

By the use of the Courant-Fisher Theorem, we can prove the following theorem.
Theorem 8.7. Consider a network \( G_n = (V_n, E, W) \). Removing a node set \( S \subset V \) with \(|S| = m \) from \( G_n \) generates a network \( G_{n-m}^S \). Then for \( i \leq n-m \),

\[
\lambda_i(G_n) - \alpha_{\text{max}}^S \leq \lambda_i(G_{n-m}^S) \leq \lambda_{i+m}(G_n) - \alpha_{\text{min}}^S,
\]

where \( \alpha_{\text{min}}^S \) and \( \alpha_{\text{max}}^S \) denote respectively the minimum and maximum connection weights between \( S \) and each node in the complement node set \( S^c = V \setminus S \), i.e.,

\[
\alpha_{\text{max}}^S = \max_{i \in S^c} \sum_{j \in S} w_{ij}, \quad \alpha_{\text{min}}^S = \min_{i \in S^c} \sum_{j \in S} w_{ij}.
\]

The proof follows the same proof ideas in Lemma 8.1 and Lemma 8.2.

Proof. Denote the Laplacian of \( G_n \) by \( L_n \) and denote the Laplacian of the graph \( G_{n-m}^S \) by \( L_{n-m}^S \). Removing the rows and columns of \( L_n \) that correspond to the node set \( S \), we have the \( m \)th order principle sub-matrix of \( L_n \) denoted by \( P_{n-m}^S \). Label the nodes of \( G_n \) by \( \{1, 2, ..., n\} \). Since we can relabel the graph, without loss of generality, we can denote the labels for nodes in \( S \) as \( \{ (n-m+1), ..., n \} \). We note that \( P_{n-m}^S \) and \( L_{n-m}^S \) are different only in the diagonal terms as follows:

\[
L_{n-m}^S = P_{n-m}^S - \text{diag}(\alpha_1, \alpha_2, ..., \alpha_{n-m}),
\]

where \( \alpha_k (1 \leq k \leq n-m) \) is the connection weight between node \( k \) in \( S^c \) and the node set \( S \), i.e.,

\[
\alpha_k = \sum_{j \in S} w_{kj}.
\]

Furthermore, let

\[
\alpha_{\text{max}}^S = \max_{i \in S^c} \sum_{j \in S} w_{ij}, \quad \alpha_{\text{min}}^S = \min_{i \in S^c} \sum_{j \in S} w_{ij}.
\]

For simplicity, let \( \Lambda_S = \text{diag}(\alpha_1, \alpha_2, ..., \alpha_{n-m}) \). Then

\[
L_{n-1}^S = P_{n-1}^S - \Lambda_S.
\]

Fix \( k \in \{1, 2, ..., n-m\} \). We denote \( U_k \) as the set of subspaces \( \{ U \subset \mathbb{R}^n \text{ with } \dim(U) = n-k+1 \} \).

It follows from Courant-Fisher Theorem that,

\[
\lambda_k(P_{n-m}^S) = \max_{U \in U_k} \min_{x \neq 0} \left\{ \frac{x^T P_{n-m}^S x}{x^T x} \right\}.
\]
Substituting $P_{n-m}^S$ by $L_{n-m}^S + \Lambda_S$ we have

$$\lambda_k(P_{n-m}^S) = \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T (L_{n-m}^S + \Lambda_S) x}{x^T x}; x \neq 0 \right\}.$$ 

Further simplifying the right hand side, we obtain

$$\lambda_k(P_{n-m}^S) \leq \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T (L_{n-m}^S + \alpha_{\max}^S) x}{x^T x}; x \neq 0 \right\}$$

$$= \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T L_{n-m}^S x}{x^T x} + \alpha_{\max}^S; x \neq 0 \right\}$$

$$= \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T L_{n-m}^S x}{x^T x}; x \neq 0 \right\} + \alpha_{\max}^S$$

$$= \lambda_k(L_{n-m}^S) + \alpha_{\max}^S.$$ 

By Lemma 8.3, we have $\lambda_k(L_n) \leq \lambda_k(P_{n-m}^S)$. Therefore we obtain the following inequality

$$\lambda_k(L_{n-m}^S) \geq \lambda_k(L_n) - \alpha_{\max}^S,$$

that is,

$$\lambda_k(G_{n-m}^S) \geq \lambda_k(G_n) - \alpha_{\max}^S, \quad 1 \leq k \leq n - m.$$ 

Similarly, we can prove the upper bound for $\lambda_k(G_{n-m}^S)$. Fix $k \in \{1, 2, ..., n - m\}$,

$$\lambda_k(P_{n-m}^S) \geq \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T (L_{n-m}^S + \alpha_{\min}^S I) x}{x^T x}; x \neq 0 \right\}$$

$$= \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T L_{n-m}^S x}{x^T x} + \alpha_{\min}^S; x \neq 0 \right\}$$

$$= \max_{U \in U_k} \min_{x \in U} \left\{ \frac{x^T L_{n-m}^S x}{x^T x}; x \neq 0 \right\} + \alpha_{\min}^S$$

$$= \lambda_k(L_{n-m}^S) + \alpha_{\min}^S.$$ 

By Lemma 8.3, we have $\lambda_k(P_{n-m}^S) \leq \lambda_{k+m}(L_n)$. Therefore we obtain the following inequality

$$\lambda_k(G_{n-m}^S) \leq \lambda_{k+m}(G_n) - \alpha_{\min}^S, \quad 1 \leq k \leq n - m.$$
8 Consensus-induced Centrality for Networks of Dynamical Systems

8.5.3 Upper and Lower Bounds of ACIC

Based on Theorem 8.7, we have the following result on upper and lower bounds of ACIC for node sets.

**Proposition 8.1.** Consider a network $G_n = (V_n, E, W)$. Removing a node set $S \subset V$ with $|S| = m$ from $G_n$ generates a network $G_{n-m}^S$ with $n > m$. Then

$$\lambda_2(G_n) - \lambda_{2+m}(G_n) + \alpha_{\text{min}}^S \leq ACIC_S \leq \alpha_{\text{max}}^S$$  \hspace{1cm} (8.19)

where $\alpha_{\text{max}}^S = \max_{i \in S^c} \sum_{j \in S} w_{ij}$, $\alpha_{\text{min}}^S = \min_{i \in S^c} \sum_{j \in S} w_{ij}$ with $S^c = V \setminus S$.

8.6 Examples of Consensus-induced Centrality Measure for Networks

We calculate the ACIC and RCIC for some simple networks in Table 8.1 and real world networks. For simple networks, since the edge weight is either 1 or 0, $\alpha_{\text{max}} = 1$ (if the graph is non-empty) and therefore $ACIC \leq 1$.

**Table 8.1** Algebraic Connectivity for Some Simple Networks [2]

<table>
<thead>
<tr>
<th>Graph G</th>
<th>Algebraic Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Graph $K_n$</td>
<td>$\lambda_2(K_n) = n$</td>
</tr>
<tr>
<td>Path $P_n$, ($n &gt; 1$)</td>
<td>$\lambda_2(P_n) = 2(1 - \cos \frac{\pi}{n})$</td>
</tr>
<tr>
<td>Cycle $C_n$, ($n &gt; 2$)</td>
<td>$\lambda_2(C_n) = 2(1 - \cos \frac{2\pi}{n})$</td>
</tr>
<tr>
<td>Bipartite complete graph $K_{p,q}$, ($p \geq 1, q &gt; 1$)</td>
<td>$\lambda_2(K_{p,q}) = \min{q, p}$</td>
</tr>
<tr>
<td>Star $K_{1,q}$, ($q &gt; 1$)</td>
<td>$\lambda_2(K_{1,q}) = 1$</td>
</tr>
</tbody>
</table>

**Complete Networks**

The nodes on a complete graph (i.e. fully connected graph) are indifferent from one another, and they all have the same importance. For any node $v$ in an complete graph $C_n$:

$$ACIC_v = n - (n - 1) = 1$$
$$RCIC_v = \frac{n}{n} - \frac{n - 1}{n - 1} = 0$$
ACIC is 1 means that the algebraic connectivity decreases by 1 if any node is removed. RCIC is 0 implies that compared to the full capacity of the network, the relative consensibility of the network does not change when any node is removed.

**Complete Bipartite Networks**

For a complete bipartite network $K_{|P|,|Q|}$ with $n$ nodes where $P$ and $Q$ are the two complete cliques. Suppose $|P| < |Q|$. For a node $v_p$ in $P$,

$$ACIC_{v_p} = |P| - (|P| - 1) = 1,$$
$$RCIC_{v_p} = \frac{|P|}{n} - \frac{|P| - 1}{n - 1}.$$

For a node $v_q$ in $Q$

$$ACIC_{v_q} = |P| - |P| = 0,$$
$$RCIC_{v_q} = \frac{|P|}{n} - \frac{|P|}{n - 1} = \frac{-|P|}{n(n - 1)}.$$

This implies that nodes in the smaller clique $P$ are more important than nodes in the larger clique $Q$ in terms of consensibility. The removal of $(|Q| - |P|)$ nodes in the larger clique $Q$ does not affect the consensibility of the network.

Star graphs are special cases of complete bipartite graphs where the smaller complete clique $P$ has only one node. For the center node $v_c$ in the star network $K_{1,n-1}, (n > 3),$

$$ACIC_{v_c} = 1, \quad RCIC_{v_c} = \frac{1}{n},$$

For any leaf node $v_l$ in a star network $K_{1,n-1}$

$$ACIC_{v_l} = 0, \quad RCIC_{v_l} = -\frac{1}{n(n - 1)}.$$

This implies that the center node is more important than the leaf nodes in terms of consensability on a star network. $RCIC$ is negative for leaf nodes implies that the remove a leaf node can increase the network relative consensibility.
Path Networks

For a node $v_b$ on the boundary of a path network

\[
ACIC_{v_b} = 2\left(\cos\frac{\pi}{n-1} - \cos\frac{\pi}{n}\right),
\]
\[
RCIC_{v_b} = \frac{2\left(1 - \cos\frac{\pi}{n}\right)}{n} - \frac{2\left(1 - \cos\frac{\pi}{n-1}\right)}{n-1}.
\]

For an internal node $v_i$ on a path network $P_n$

\[
ACIC_{v_i} = 2\left(1 - \cos\frac{\pi}{n}\right) - 0 = 2\left(1 - \cos\frac{\pi}{n}\right),
\]
\[
RCIC_{v_i} = \frac{2\left(1 - \cos\frac{\pi}{n}\right)}{n} - 0 = \frac{2\left(1 - \cos\frac{\pi}{n}\right)}{n}.
\]

This suggests that on path networks, nodes in the middle is more important than nodes on the boundary in terms of the influence to consensibility.

Cycle Networks

For any node $v$ on a cycle network $C_n$ ($n > 2$)

\[
ACIC_v = 2\left(\cos\frac{\pi}{n} - \cos\frac{2\pi}{n-1}\right)
\]
\[
RCIC_v = \frac{2\left(1 - \cos\frac{2\pi}{n}\right)}{n} - \frac{2\left(1 - \cos\frac{\pi}{n-1}\right)}{n-1}
\]

A Simple Network with Negative ACIC

Here is a network example on which some nodes have negative ACIC and RCIC.

![Fig. 8.1 A Simple Network with Negative ACIC](image-url)
Negative ACIC and negative RCIC of node 1 respectively imply that its removal would increase the consensibility and relative consensibility of the network. Note that node 1 is weakly connected to the close community formed by nodes 2, 3 and 4. Node 2 which is located at the center of the network has the largest ACIC and the largest RCIC. Zero ACICs of nodes 3 and 4 suggest the removal of them does not change the absolute consensibility. Negative RCIC of nodes 3 and 4 implies that each removal would result in a network that has better relative consensibility.

### The Karate Club Network

We examine the centrality measure on the Karate Club Network. Node 1 has the largest ACIC and RCIC and the removal of which will disconnect the network. Node 34 has the second largest ACIC and RCIC. Nodes 2, 3, 32 and 33 has relatively large ACIC and RCIC compared to other nodes. Node 17 has the smallest negative ACIC and RCIC, which means its removal would increase network consensibility the most. Nodes that have larger ACIC and RCIC tend to be the hubs of a densely connected community and nodes that have small ACIC and RCIC tend to be nodes that are weakly connected to or distant from communities.

![The Karate Club Network](image)

**Fig. 8.2** The Karate Club Network [1]

### Small World Networks

We use the small world network model [8] to generate the following network, by starting with 20 nodes, connecting each node to 4 nearest neighbours in ring topology and for each edge the rewiring probability is 0.5. The node has the highest ACIC and RCIC is node 8, and it has a degree of 4 which is not the highest node degree. Node 1 has the highest node degree which is 6. Node 9 has the lowest ACIC and RCIC and it has a node degree of 3.
8 Consensus-induced Centrality for Networks of Dynamical Systems

Power Networks

We examine the centrality measure for the undirected, unweighted network with 4941 nodes representing the topology of the Western States Power Grid of the United States [8]. The power grid stability problem is modelled as the synchronization of coupled oscillators [77]. We note that the algebraic connectivity provides a conservative estimate for achieving synchronization [77].

The algebraic connectivity: $7.5921 \times 10^{-4}$. The max ACIC of all nodes is $7.5921 \times 10^{-4}$. There are 1229 nodes achieved the maximum ACIC under error bound $10^{-8}$. The minimum ACIC of all nodes is $-5.9562 \times 10^{-7}$. Under error bound $10^{-8}$, there are 27 nodes that achieve the minimum ACIC.

Since the network size is large and the network is sparse, the RCIC would have very small values. The maximum RCIC is $1.5366 \times 10^{-7}$. The minimum RCIC is $-1.5167 \times 10^{-10}$.

The Southern Women Club Network

The Southern Women Club network [86] is an undirected weighted network that contains the observed attendance at 14 social events by 18 southern women. The edge weight is the number of co-attended events. The maximum edge weight in the network is 7.

The network consensibility (algebraic connectivity) is 12.7572. Node 5 achieves the minimum
ACIC (-0.2870) and minimum RCIC (-0.0586). Node 14 achieves the maximum ACIC (1.9528) and the maximum RCIC (0.0732).

8.7 Discussions

An application is on power grid stability since the grid stability is modelled as the synchronization of coupled oscillators [77], although the algebraic connectivity provides a conservative condition for achieving synchronization. If one node on the grid needs to be removed, this work answers the question that which node can be remove first to ensure that the resulting network structure have at least the same performance in reaching synchronization.

In the case of wireless sensor networks where information needs to be processed through aggregation, this can be used to reduce the use the redundant sensors to improve the consensus rate, by careful choice of nodes according to ACIC.

Future directions of this work include: (1) consensus-induced measures for edges, (2) centrality measures for groups of nodes, (3) centrality measures for directed network systems, (4) induced centrality measures for dynamical systems on networks controlled via other decentralized strategies.

8.8 Appendix

8.8.1 The Cauchy Interlace Theorem

Theorem 8.8 (Cauchy Interlace Theorem [87]). Let \( A \) be a Hermitian matrix of order \( n \), and let \( B \) be a principal sub-matrix of \( A \) of order \( n - 1 \). If \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n \) lists the eigenvalues of \( A \) and \( \mu_1 \leq \mu_2 \leq \ldots \leq \mu_{n-2} \leq \mu_{n-1} \) the eigenvalues of \( B \), then \( \lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \ldots \leq \lambda_{n-1} \leq \mu_{n-1} \leq \lambda_n \).
8.8.2 The Courant-Fischer Theorem

The minimax and maximin characterization of eigenvalues of Hermitian matrices, known as the Courant-Fischer Theorem, is represented in the following.

**Theorem 8.9** (Courant-Fischer Theorem). *Let $M$ be an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_k \leq \cdots \leq \lambda_n$ then*

\[
\lambda_k = \min_U \max_x \left\{ \frac{x^T M x}{x^T x} : U \subset \mathbb{R}^n, \dim(U) = k, x \in U = \text{span}(U), x \neq 0 \right\}
\]

*and*

\[
\lambda_k = \max_U \min_x \left\{ \frac{x^T M x}{x^T x} : U \subset \mathbb{R}^n, \dim(U) = n - k + 1, x \in U = \text{span}(U), x \neq 0 \right\}.
\]
Chapter 9

Summary and Future Research Directions

9.1 Summary

This work studies the analysis and the control design for large-scale and complex networks of linear systems.

We employ the graphon theoretical framework to model the problem for arbitrary-size networks. More specifically, we formulate the graphon differential equations, graphon dynamical systems and graphon optimal control problems. A sufficient condition for exact controllability is obtained. The optimal graphon state-to-state control problem and graphon LQR control problem are solved using infinite dimensional system theory. Under the assumption that the graphon is the limit of the underlying sequence of networks under $L^2$ metric, we develop the approximation theorem and show the convergence property of the terminal state for the state-to-state control problem, and that of the state trajectory and regulation cost for the LQR problem.

Based on these results, we propose the graphon based control methodology (which includes GSSC and GLQR strategies) for controlling complex network systems. The general graphon control strategy consists of the following steps:

1) Identify the graphon limit of the sequence $\tilde{S}$ of networks as the number of nodes goes to infinity.

2) Solve the corresponding control problem for the limit graphon dynamical system.

3) Approximate the control law for the limit system so as to generate approximate control laws for finite network systems.
4) Apply the resulting control laws to the networks of systems along the sequence $\tilde{S}$.

Both the GSSC and GLQR strategies are illustrated by Figure 9.1. They are developed to solve the otherwise intractable or computationally heavy optimal control problems for complex networks of linear systems.

Specifically, in Part I of this thesis, the minimum energy state-to-state control problem and the linear quadratic regulator problem are solved for complex network systems using this graphon control strategy. The main contributions of Part I of the thesis include:

- the formulation of graphon differential equations and graphon dynamical systems, which form fundamental components of this study of arbitrary-size networks of linear systems.
- the development of graphon state-to-state control methodology to solve state-to-state control problem on complex networks.
- the proposed graphon linear quadratic regulation methodology to solve linear quadratic regulator problems on complex networks.

Within this problem formulation, it turns out that the spectral property of the graphon (or network) is important to the control synthesis. Therefore, we develop a standard procedure to
identify the eigenfunctions of a graphon. Meanwhile, the spectral approximation of graphons are studied. The approximation consists of two levels: the spectral approximation by the most significant eigenvalues, and the approximation of the corresponding eigenfunctions via Fourier functions.

To achieve the control synthesis in a decentralized manner, we obtain the team optimal solutions, where each agent is minimizing a local cost based on its own state observation and some global information related to network spectra and initial conditions (more specifically, the sizes of the projections of initial states in all eigenfunction directions, the evaluations of eigenfunctions at its local index, and the eigenvalues). This team solution ensures that, when all agents are optimally solving their local LQR problem, they together solve the global optimal control problem.

For multi-agent systems on networks with decentralized or distributed control strategies, agents may leave or stay in the network intentionally or due to failures and this would influence the performance of the decentralized or distributed control. To quantify the influence, the consensus-induced centrality measure for networks of systems is proposed, with lower bounds and upper bounds established. This centrality measure is illustrated in real world network examples.

9.2 Future Research Directions

9.2.1 Graphon Linear Quadratic Gaussian Problems

In this thesis, the problem formulation is for deterministic systems distributed on networks. A straightforward extension of this work should be for stochastic systems on networks, especially, the linear quadratic Gaussian problem on large networks. In this direction, the formulation for $L^2$ white noise is required for the graphon linear quadratic Gaussian model.

In the finite dimensional network case, this work can be directly generalized to the problem for finite dimensional networks of stochastic systems with additive noise and exact state observation, based on certainty equivalence principle [75]. However, if the additive noise also exists in local state observations, one needs to design a decentralized or distributed Kalman filter to achieve the local state estimate. The design of a decentralized or distributed Kalman filter based on global information of network spectra and local state observation needs to be further explored.
9.2.2 Graphon Mean Field Games

Mean Field Game [67, 68] theory solves the problem with completely connected networks or networks with finite completely connected subclasses. Examples of applications can be financial markets, electric water heating [88], CDMA power allocation, social optima [89], etc. The development of Mean Field Game theory has provided a decentralized optimal control solution concept for large dimensional multi-agent systems. The limitation is the restriction on the underlying networks, i.e., completely connected networks or finitely many completely connected subclasses. However, many real world networks do not exhibit such properties, instead, they exhibit heterogeneity and bounded local connections. Mean Field Games on networks characterized by graphons inherently permit the study of networks with both an infinite number and a finite number of agents.

It should be mentioned that the work in this thesis on graphon control theory has fundamentally influenced the current research effort in Graphon Mean Field Games (GMFG) and the GMFG equations [41, 90].

9.2.3 Non-linear Local Dynamics

Many real world systems on networks have non-linear nodal dynamics. For instance, in a neuronal network, each neurone exhibits nonlinear nodal dynamics with action potential thresholds and non-linear excitation responses. To understand and control the behaviour of the interconnected nonlinear dynamical systems as such are the most challenging and important problems. A framework for non-linear dynamical systems distributed on graphons need to be further explored. Nonlinear optimal control tools are required to solve the optimal control problem.

Networks of non-linear systems are well studied in synchronization and pining control, where the network coupling strength and input nodes are the control variables. One common key feature of these network systems is that agents share similar or same nodal dynamics with differences appear only in network coupling, which helps significantly simplify the analysis. The algebraic connectivity of the underlying networks, together with nodal nonlinear dynamics, determines the coupling strength required to achieve synchronization.

The synchronization problems and pinning control problems on networks of growing or infinite sizes can be posed and studied via graphon representations of the underlying networks. This requires the formulation of the algebraic connectivity for graphons and that of graphon Laplacians, which have been seen in [91]. Topics of pinning a subpopulation to control graphon
systems need to be explored.

9.2.4 Control of Time Varying Graphon Dynamical Systems

Many real world networks, such as, social, technological and biological networks, may grow both in the connections and the cardinality. The advantage of using the graphon control framework in time varying network systems is that graphons can characterize networks that evolve not only in their structure but also in their cardinality. In this thesis, the underlying graphon is assumed to be constant over time. This work has a straightforward and meaningful generalization to the time varying case, which requires more knowledge on the characterization of time varying infinite dimensional systems. Furthermore, under the graphon dynamical system framework, the control problems for time-varying graphon systems with the underlying graphons converging over time to a certain limit can be posed.

9.2.5 Graphon as Non-parametric Models for Control Design

In this thesis, graphon theory is employed to characterize the limit of graph sequences. Another interpretation of graphons is to consider graphons as generative models for random graphs. The generative models can be non-parametric models derived from data. A formulation of network systems with randomness in network structures can be considered, that is, the dynamics of the network systems contain randomness from the underlying network. This randomness can be statistically estimated from observed network data and represented in terms of graphon models.

9.2.6 Centrality Measure of Dynamical Systems on Graphons (or Networks)

For multi-agent systems on networks with decentralized or distributed control strategies, agents may leave the network intentionally or due to network failures and this would influence the performance of the decentralized or distributed control. The sensitivity of the control law to the removal of an agent (or a group of agents) should be considered in future works. Along this direction, we have studied the centrality measure for agents on networks following the consensus type control protocol in [92], which immediately applies to synchronization problems on networks. For networks of infinite agents with averaging dynamics, the removal of one agent has a negligible effect but the removal a proportion of the agent population may have a significant effect. For networks of finite dimension with strong couplings rather than averaging couplings, the
removal of an agent may have significant effect on the network spectrum and hence the control performance.

Future research studies should include consensus (or synchronization) problems for random networks, where the underlying graphs are characterized by graphons and each agent follows the consensus protocol (or the synchronization protocol) similar to the finite network case. This would immediately generalize problems of consensus over random networks [93, 94] to the cases where the underlying networks are not necessarily of the same size.

Furthermore, issues concerning the consensus-induced centrality measures for dynamical systems on graphons should be studied.

9.2.7 Graphon Control Applications

The work in the thesis has possible applications on many real world networks such as electrical networks, market networks, social networks, sensor networks, and DiDi (or Uber) networks, etc. Graphon control allows to compare networks of different sizes, which enables us to design control laws regardless of the network size. The computation is less complex if the underlying network has simple spectral representations. With spectral characterization of the networks, each agent on the network by optimally solving their local optimization problem and collectively solves the global optimal control problem. Examples of such networks include complete bipartite networks \( \left\{ K_{m,n} \right\} \) which contain three eigenvalues \( \left\{ \pm \sqrt{mn}, 0, \sqrt{m+n} \right\} \). Bipartite networks inherently can represent networks with service providers and consumers such as DiDi (or Uber) networks, electric market networks, social networks, etc. There are many other examples of networks that have simple spectral representations. In fact, in engineering networks such as IoT, power grids, sensor networks, etc., during the process of designing network structures, it is even possible to construct networks with small number of non-zero eigenvalues.

The control problems on these networks with arbitrary large sizes can be solved using graphon control and graphon team optimal control with lower complexity.

9.2.8 Other Future Directions

Extension to Sparse Networks: Graphons are designed for graphs with dense connections, that is, graphs \( \left\{ G_n = (V_n, E_n) \right\} \) such that \( \frac{|E_n|}{|V_n|^2} \) does not shrink to zero as \( n \to \infty \). Graphings [40] or graphexes [95, 96] as the counterparts of graphons for sparse networks shall be explored for systems on networks with sparse connections.
**Observability for Graphon Dynamical Systems:** Approximate observability and exact observability as the dual concepts of approximate controllability and exact controllability should be explored for graphon dynamical systems.

**Dynamical Systems on Directed Networks:** In this thesis, the networks under consideration are undirected and the graphons under consideration are symmetric objects. The extension of the work in this thesis to directed networks should be further investigated.
References


References


[34] ——, “Graphon linear quadratic regulation of large-scale networks of linear systems,” Accepted for the *57th IEEE Conference on Decision and Control (CDC)*, Miami Beach, FL, USA, December 2018.


[41] P. E. Caines and M. Huang, “Graphon mean field games and the gmfg equations,” Accepted for the 57th *IEEE Conference on Decision and Control*, Miami Beach, FL, USA, December 2018.


[90] P. E. Caines and M. Huang, “Mean field games on infinite networks and the graphon mean field game equations,” Presented at the Conference on Stochastic Control, Computational Methods and Applications IMA, University of Minneapolis, Minneapolis, MN, USA, May 2018.


