

© 2009 by Azeem Sarwar. All rights reserved.

SPATIOTEMPORAL SYSTEMS: GRADUAL VARIATIONS,
IDENTIFICATION, ADAPTATION AND ROBUSTNESS

BY

AZEEM SARWAR

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mechanical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Doctoral Committee:

Professor Srinivasa M Salapaka, Chair
Professor Petros G Voulgaris, Director of Dissertation Research
Professor Carolyn Beck
Professor Prashant G. Mehta

ABSTRACT

Motivated by the increasing size of complex systems by mere interconnection of simple units, this dissertation considers a set of important open research problems related to the stability, identification, adaptation and robustness of spatiotemporal systems. First, we consider the l_∞ stability of linear spatiotemporally varying (LSTV) systems when the underlying controllers are designed based on local linear spatiotemporally invariant (LSTI) approximations. We show that the l_∞ to l_∞ performance of global LSTV systems cannot be much worse than the worst frozen spatially and temporally l_∞ to l_∞ performance, given that the rates of variation of the plant and the controller are sufficiently small. Next, we consider the problem of system identification of LSTI systems where the subsystems cooperatively attempt to identify the dynamics common to every one. We propose a distributed projection algorithm that guarantees to bring the local estimates arbitrarily close to each other for large enough time, hence resulting in a slowly varying spatiotemporal system. Coupled with the results on the stability of LSTV systems, we next propose an indirect adaptive control scheme based on certainty equivalence. Last, we look at the robust l_∞ and l_2 stability of LSTI systems and address the necessary and sufficient conditions for robust stability in the presence of LSTV perturbations. We also investigate the robust stability of these systems with the underlying perturbations being nonlinear spatiotemporally invariant. We show that the robustness conditions are analogous to the scaled small gain condition (which is equivalent to a spectral radius condition and a linear matrix inequality for the l_∞ and l_2 case respectively) derived for standard linear time invariant models subject to linear time varying or nonlinear perturbations. Future research directions are also provided.

To my Parents

میرے ماں باپ سلامت میری جنت آباد
اک دعا ان کے لیے اٹھتے ہی سب سے پہلے

ACKNOWLEDGMENTS

“He who does not thank people does not thank Allah (God)”.
(Prophet Muhammed (SAW), Abu Daud No. 4177, Tirmidhee No. 1877)

I would like to thank everyone, from the depth of my heart, who has supported and encouraged me throughout my years of study. I am greatly indebted to my advisor, Professor Petros G. Voulgaris, for his phenomenal support, guidance, and supervision. His knowledge, dedication, enthusiasm, patience, calm demeanor, and personality have made my graduate experience at University of Illinois extremely rewarding and indeed unforgettable. I am also indebted to my co-advisor Prof. Srinivasa M. Salapaka, for his support, guidance, encouragement, and ever availability to discuss various research ideas. Given his friendly and welcoming attitude, I never felt shy in stopping by his office for discussions concerning my research and other related issues. I am also grateful to my committee members: Professor Carolyn Beck, and Professor Prashant G. Mehta, for their kind service in my dissertation committee, invaluable comments on my dissertation, and edificatory discussions about my research.

I would like to thank Dr. Naeem Shiekh for his help with various mathematical problems I encountered throughout my stay at the University of Illinois at Urbana Champaign (UIUC). I am also grateful to Prof. Zhong-Jin Ruan, Prof. I. D. Berg, and Prof. Angelia Nedich, for frequently sparing the time to discuss problems related to my research. The encouragement, and motivation I received from Prof. Geir E. Dullerud was extremely helpful to me in coming to (UIUC), after having to wait for fifteen months in visa processing. I am also appreciative of the cooperation of Celia Snyder in this regard. I would like to thank Prof. Geir E. Dullerud for also providing me guidance and lending me patient ears in my first year at UIUC. I would also like to thank Prof. Franceso Bullo for providing me support, and guidance in my first year at UIUC.

I would like to extend special thanks to Dr. Shengxiang Jiang for pro-

viding not only a great company while he was at Coordinates Sciences Lab. (CSL), but also helping me out even after his graduation. I am also thankful to Raja Katta for providing me study material for the qualifying exam. I must also thank Frances Bridges, Rebecca (Becky) Lonberger, Ronda Rigdon, Jana Lenz, John Bambenek, Jennifer Rideout, Pamela S. VanEtta, and Kathy Smith for their cooperation at CSL and Mechanical Sciences and Engineering (MechSe) building.

It has been a great experience to be a student at UIUC and a member of the CSL. Not only I have gained knowledge from the research community, but have also been inspired by their dedication to carry out research and the willingness to help and share with others. My friends at University of Illinois have made the graduate study enjoyable and memorable. Thanks to Rouzbeh, Orhan, Serdar, Aneel, Nick, Hemant, Anand, Shreyas, Wei, Sun, Robert McCabe, and all other people I have come to know in Urbana-Champaign through CSL.

Dr. Sorin Benghea, who supervised me during my summer internships at EATON Corporation has had a great influence on my professional development. I am grateful to him for being a great mentor and a source of inspiration in striving for nothing but excellence. His professional, friendly and calm demeanor made my stay at EATON Corporation very enjoyable.

I am thankful to Dr. Irfan Ahmed and Dr. Naeem Shiekh for being a source of inspiration, motivation, and guidance with regards to undertaking various activities for community development. I would also like to thank Dr. Rizwan Uddin, Dr. Zeeshan Ahmed, Dr. Arif Masud, Dr. Safdar Ali Khan, Ammar Bhutta, Ali Tahir Kazmi, Faisal Saeed, and their families for providing me with a homely atmosphere in Urbana Champaign.

Special thanks to Dr. Jalal Almuhtadi, Dr. Nahil Sobh, Muhammed Aaqib Dada, Shakil Mustafa, Nouredine Tayebi, Ibrahim Zubair Mohammed, Adeel Dosani, Salman Syed, Ahmed Habib, Mustafa Kherbane, Hilmi Wan, Mohammed Abusaad, Naseer Yahya, Omar Galaria, Youness Ktiri, Humad Choudhry, Atif Saeed, Dr. Zayd Leseman, Rafay Zahid Ansari, Shakil Mustafa, Naeem Shams, Khawar Abbas, Khurram Farouque, Zahid Anwar, Fouzan Ansari, Noor Razvan, Abrar Manzar, Shahzad Javed, Kazim Ali Malik, Khurram Farouque, Imran Khan, Hassan Khan, Imaad Khan, Moaz Khan, Ahsan Mirza, Shadi Hussain, Ahsan Alvi, Jehanzeb Abbas, Jehanzeb Hameed, Haris Anwar, Yerkin Abdildin, Osman Sarood, Shehla

Saleem, Asma Faiz, Zara Fahim, Shaista Babar, Kashif Altaf, Adeel Ahmed, Qazi Aurangzeb, Asfand Waqar, Numair Ahmed, Ahmed Qadir, Salman Noshear, Abid Ameen, Ibrahim Pasha, Zakir Hussain, Manzoor Hussain, Aamir Wali, Ahmad Ajmal Cheema, Fawad Hassan Ismail, Adeel Zafar, Hamood Arham, Kamran Akhtar, Salman Qureshi, Umar Mahmud, Usman Tariq, Muhammad Yasar, Aizaz Syed, Kashif Manzoor, Kashif Mohammad, Syed Imad, Rakesh Kumar, Tanweer Alam, Ahmed Ibrahim Abd El-Rahman, and others for their companionship.

It would be unfair not to make a special note for the support I received from one of my best friends, Mohamed Yousef Mohamed. I, for all times to come, shall remain indebted to him for lending a helping hand, patient ears, and words of encouragement and wisdom whenever I was in need.

I would like to acknowledge the financial support by National Science Foundation under NSF Awards No CCR 03-25716 ITR, CMS 0301516, and AFOSR grant FA9950-06-1-0252.

Last, but for sure not the least in the list of people, is my family. I am forever grateful to them for supporting and helping me to get through all these years. My late mother Iffat Sarwar remains an inspiration for me to strive for excellence. I am forever indebt to her life long sacrifices for my wellbeing of all sorts. The tremendous support and encouragement I received from my father Dr. Mohammad Sarwar Kamran during my stay at University of Illinois was invaluable and vital for me to carry out my work with composure and tranquility. I am highly appreciative of his guidance and priceless words of wisdom throughout my lifetime, and making every possible effort for my comfort and success. I am extremely grateful to his ever-presence during my dark days. My sister, Dr. Zoonah Sarwar, my brother Dr. Fartash Sarwar and their families have all been ever ready to provide advice, encouragement, and support. I would also like to thank my wife Dr. Ayesha Sarfaraz, and her family for their support. I am grateful to my wife for allowing me to spend her share of my time in preparing and writing this dissertation over the last two years.

While I have tried my best to acknowledge the support and favors of others who have contributed in preparation of this dissertation, I am sure that the forgetful nature of human beings is deemed to exercise itself. I, therefore, present my sincere apologies to those whom I may have missed out above.

At the end, I humbly acknowledge the favors of Allah (the Lord of heavens

and earth and the one true God) for giving me the opportunity to undertake graduate studies at a school as prestigious as UIUC, and blessing me with the company of all those whom I have mentioned above. I am grateful to Him for providing me the means that I had never imagined, nor deserved. I do not have the words, nor do I even know how to thank the One who has been so kind for as long as I can remember and beyond. Indeed He is the Best Lord one can have. Alhamd-o-Lillah-e-Rab-il-Aa'lameen (All praises to Allah, the Lord of all the worlds)!

TABLE OF CONTENTS

LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xii
CHAPTER 1 INTRODUCTION	1
1.1 Literature Review	4
1.1.1 Distributed Estimation/Identification	5
1.1.2 Distributed Adaptive Control	7
1.1.3 System Robustness	8
1.2 Contributions and Organization	9
1.3 General Framework	11
1.4 Notation and Symbols	12
CHAPTER 2 STABILITY AND PERFORMANCE OF GRADUALLY VARYING SPATIOTEMPORAL SYSTEMS	13
2.1 Basic Setup	13
2.1.1 Notations	13
2.1.2 Spatiotemporal Varying Systems	14
2.1.3 Spatially Invariant Systems	15
2.1.4 Frozen Spatiotemporal Systems	16
2.1.5 Support of m	17
2.1.6 Recursively Computable Spatiotemporal Systems	17
2.1.7 Gradually Varying Spatiotemporal System	17
2.1.8 Integral Time and Space Absolute Error	18
2.1.9 z, λ Transform	18
2.2 Frozen Space-Time Control	19
2.3 Stability Analysis	23
2.4 Performance Analysis	32
2.5 Remarks	35
2.6 Conclusion	35
CHAPTER 3 SYSTEM IDENTIFICATION OF SPATIOTEMPORALLY INVARIANT SYSTEMS	37
3.1 Distributed Projection Algorithm	37
3.1.1 Main Idea	38

3.1.2	Cumulative Improvement Index $I_i(\cdot)$	39
3.1.3	The ϵ -Rule:	41
3.1.4	Properties of Distributed Projection Algorithm	42
3.1.5	Information Exchange	45
3.1.6	Simulation	48
3.2	Parameter Estimation With Bounded Noise	52
3.2.1	Simulation	57
3.3	Conclusions	59
CHAPTER 4 ADAPTIVE CONTROLLERS FOR SPATIOTEMPORALLY INVARIANT SYSTEMS		60
4.1	Basic Setup	60
4.2	Characterization of a Class of Gradually Varying Spatiotemporal Controllers	62
4.2.1	Discussion	67
4.3	Convergence of Adaptive Scheme	68
4.4	Conclusion	69
CHAPTER 5 l_∞ AND l_2 ROBUSTNESS OF SPATIALLY INVARIANT SYSTEMS		71
5.1	Basic Setup	72
5.1.1	Spatially Invariant Systems	72
5.1.2	Perturbation Models	74
5.2	l_∞ Stability Robustness	76
5.2.1	LSTV Unstructured Perturbations	76
5.2.2	LSTV Structured Perturbations	79
5.2.3	NLSTI Structured Perturbations	84
5.2.4	Numerical Example	85
5.3	l_2 Stability Robustness	86
5.3.1	LSTV Unstructured Perturbations	86
5.3.2	LSTV Structured Perturbations	91
5.3.3	NLSTI Structured Perturbations	93
5.3.4	Remark	94
5.3.5	Numerical Example	94
5.4	Conclusion	95
CHAPTER 6 CONCLUSIONS		96
6.1	Optimality of Frozen Space Time Control	97
6.2	Excitation Conditions for Distributed Projection Algorithm	97
6.3	Robustness Analysis of LSTV Systems	98
6.4	Robust Adaptive Control	98
6.5	l_1 Optimal Control Problem	99
REFERENCES		103

APPENDIX A DETAILS FOR NUMERICAL EXAMPLE GIVEN IN CHAPTER 5	111
AUTHOR’S BIOGRAPHY	113

LIST OF FIGURES

1.1	General Framework	12
2.1	Finite Area Region of Support With Lattice Sector Having Angle Less Than 180°	18
2.2	General Form of Closed Loop	20
2.3	Main Idea for Proof of Theorem 2.2.1	28
3.1	Flowchart Depicting the Distributed Projection Algorithm . .	39
3.2	Figure Showing ϵ -Rule	42
3.3	Calculation of Upper Bound on the Distance of Local Esti- mates	46
3.4	Circulant System	48
3.5	Comparison of Distributed Projection Algorithm (a) with Stan- dard Projection Algorithm (b)	50
3.6	Figure Showing the Effect of Choice of ϵ for the ϵ -rule in Dis- tributed Projection Algorithm	51
3.7	Parameter Estimation with Bounded Noise: Comparison of Distributed Projection Algorithm with (a) with Standard Pro- jection Algorithm (b)	58
5.1	Stability Robustness Problem.	75
5.2	Signal Construction for Unstructured Uncertainty.	77
5.3	Signal Construction.	82
5.4	Schematic showing the layout of the infinite dimensional mi- crocantilever array with mechanical and electrostatic coupling. .	85
6.1	Setup for Robust Adaptive Control for LSTI Systems	99

LIST OF ABBREVIATIONS

AFM	atomic force microscope
FIR	finite impulse response
LMI	linear matrix inequality
LPV	linear parameter-varying
LQG	linear quadratic Gaussian
LSTI	linear spatiotemporally invariant
LSTV	linear spatiotemporally varying
LTl	linear time-invariant
LTV	linear time-varying
MIMO	multi-input multi-output
NP	nondeterministic polynomial-time
NLSTI	nonlinear spatiotemporally invariant
SISO	single-input single-output
YJB	Youla-Jabr-Bongiorno

CHAPTER 1

INTRODUCTION

With the advancement in sensing and actuating techniques coupled with the incessant increase in computational power, the idea of developing more and more complex systems by putting together simpler smaller units is turning into a reality. Examples of such systems can now be cited from various areas such as: satellite constellations [1], cross-directional control in paper processing applications [2], airplane formation flight [3], [4], automated highway systems [5] and very recently, microcantilever array control for various nano-robotic applications [6]. Lumped approximations of partial differential equations (PDEs) can also be considered in this regard-examples include the deflection of beams, plates, and membranes, and the temperature distribution of thermally conductive materials [14].

Most of the examples cited above have an inherent distributed structure associated with them. For example, many of these systems have sensing and actuation capabilities at every unit (or subsystem). This can be seen clearly in the case for automated highway systems, airplane formation flight, satellite constellations, cross-directional control systems and microcantilever arrays. The rapid advancements in micro-electromechanical actuators and sensors are now enabling deployment of distributed sensors and actuators for systems governed by partial differential equations only to validate the lumped approximations of such systems.

The history of distributed control design can be traced back to 1960s under the title of so-called *Team Theory* [15]. Although distributed control design is still a daunting task in general, several results on distributed control using recently developed techniques are now available [16,17] for spatially invariant systems. Spatial invariance is a strong property of a given system, which means that the dynamics of the system do not vary as we translate along some spatial axis.

Many systems, however, may not be spatially invariant in general or even

time invariant. Distributed systems with time-dependent spatial domains arise naturally in many physical situations. Consider, for example, the problem of controlled annealing of a solid by dipping it in a fluid medium [18], or shape stabilization of flexible structures [19], where the spatial domain of the underlying distributed system is time dependent. The evolution of the spatial domains with time is described by a finite-dimensional system of ordinary differential equations, while the distributed systems are described by first-order or second-order linear evolution equations defined on appropriate spaces.

Spatiotemporal varying systems also arise frequently in the process industry where system dynamics are parameter dependent, and change with e.g. change in temperature, pressure, concentration of chemicals etc. There is a large number of industrial control problems which involve transport-reaction processes with moving boundaries such as: crystal growth, metal casting, and gas-solid reaction systems. The motion of boundaries is usually a result of a phase change, such as a chemical reaction, mass and heat transfer, and melting or solidification. The mathematical models of transport-reaction processes with moving boundaries are usually obtained from the dynamic conservation equations and consist of parabolic partial differential equations (PDEs) with time-dependent spatial domains [20]. A two dimensional linear PDE model is used to describe the spatiotemporal evolution of the thin film surface coating process [72, 73]. The control problem is formulated as regulating the thin film thickness and surface roughness by manipulating the substrate temperature and adsorption rate. Spatiotemporal varying systems also arise in distributed parameter estimation schemes. Stability of spatiotemporally varying systems is one of the aspects that we shall address in this dissertation.

For any given process, the modeling equations can usually be specified from basic conservation laws along with constitutive relations. However, parameters remain to be determined. Much of classical and modern science and engineering has been concerned with this problem. Laboratory and experimental determination of chemical rate constants, heat transfer coefficients, specific heats, electromagnetic properties, gas properties, diffusion constants, elastic moduli, strain properties, etc is an ongoing effort throughout the scientific world. Traditionally, the measurement of such parameters has been under rigidly controlled experimental conditions. The problem of identifying

parameters in systems from dynamic (time series) data can be traced back to early celestial observation problems.

Examples of applications areas abound. They include structural analysis and design where vibrations and dynamic behavior are central, acoustics problems, design in the basic process industries where heat and mass transfer and chemical reactions are important, geophysical analysis including underground water and oil exploration, earthquake study, meteorological predictions, agricultural productivity assessments, and demographic analysis [7, 8].

A major body of work exists which treats problems of identification and estimation for lumped systems described by ordinary differential equations (ODE's) (see e.g. [9–11]). Whenever a set of ODE's is used to describe a process, there is an implicit assumption that a distributed phenomena can be approximated by a lumped model. Briefly, a lumped representation of a physical system implies that local spatial variations are ignored and that the media may be considered homogeneous throughout the volume under consideration. The answer to the question of whether a lumped representation is valid is not simple. Generally, if the response shows significant instantaneous differences along a spatial direction, a distributed representation is preferable.

Since all processes are by nature distributed, and since our ability to simulate complex processes is continually increasing, parameter identification has become a basic tool in model building, and the distributed parameter identification problem has, of necessity, received great attention in recent times. We shall be looking at the problem of distributed parameter identification of spatially invariant systems in this dissertation from a control theoretic perspective.

We note here that control design of any system is as good as the system model. When the system model is not available upfront, system identification and control action have to be implemented in parallel. As the system model gets updated, the control law needs to adapt in order to guarantee stability/performance. In large distributed systems where centralized identification and control is not an option, one must resort to a distributed setting to carry out the task at hand. Distributed adaptation is another aspect that we shall touch upon in this dissertation.

Control design of distributed systems, on the other hand, for which a priori model are available, has been under the spot light of the research

community for quite some time now. The relationship between models and the reality they represent is, however, subtle and complex. A mathematical model provides a map from inputs to responses. The quality of a model depends on how closely its responses match those of the true plant. Delays, nonlinearities, time variations, and other kind of dynamics that are difficult to model in a lumped system become even harder to capture in a distributed setting, and are frequently ignored.

Since no single fixed model can respond exactly like the true plant, we need, at the very least, a set of maps. This mismatch between the model and the actual process should be represented in a certain way to avoid exciting unmodeled dynamics that could cause serious deterioration in the stability and performance of the system. Consider, for example, an array of (about 4,000) closely packed identical microcantilevers to be employed in an Atomic Force Microscope (AFM) application [6]. In reality, however, it is impossible to fabricate such a system where all the microcantilevers are identical owing to the imperfections in the fabrication process. Each microcantilever will be having slightly different length, thickness, mass and hence the associated spring constant with regards to the nominal design. Moreover, the actual array is finite and hence the spatial invariant approximation model with an infinite number of identical elements entails additional errors. It is, therefore, imperative to analyze the behavior of these ideal models in the presence of perturbations. These perturbations may not, in general, be spatially invariant or even linear. Robustness analysis of spatially invariant systems is another aspect that shall be covered in this dissertation.

1.1 Literature Review

This section reviews briefly the existing literature on distributed estimation in networked/distributed systems, adaptive control of distributed systems and system robustness of spatially invariant systems. These topics strongly relate to the main thrust of this dissertation.

1.1.1 Distributed Estimation/Identification

Distributed estimation/identification is an area that finds applications in distributed optimization, network consensus, sensor fusion, dynamic systems characterized by PDEs, and wireless networks to name a few. Each of the aforementioned areas brings its own flavor to the quest of distributed estimation/identification. We note here that distributed estimation/identification is not the same as the estimation/identification of distributed system. While the former concerns the problem of parameter estimation (identification) in (of) distributed systems in a distributed setting, the latter may only address the problem assuming a centralized setting. The literature on system identification of distributed systems (assuming a centralized setting) is abundant, with the early attempts geared towards investigations dealing with the ‘Inverse Problem’ in Heat Transfer ([21]- [25]). For a thorough historic development in this regard see [7], [8], [26], and the references therein. We shall limit ourselves in providing a literature review that is most pertinent to this dissertation from the view point of distributed estimation/identification.

One of the early works in the area of distributed estimation and consensus can be attributed to Aumann [27] who showed that if two people have the same prior, and their posteriors for a given event are common knowledge, then these posteriors must be equal. Addressing the open problems left by Aumann, Geanakoplos and Polemarchakis [28] showed that for an arbitrary event, under the assumption of common priors, if the information partitions of both agents are finite, then simply by communicating their posteriors back and forth the agents will be led to make revisions that in finitely many steps converge to a common equilibrium posterior. Borkar and Varaiya [29] extended the work of Aumann, Geanakoplos, and Polemarchakis to a system consisting of several (finite) agents. Each agent updates its estimate of the same random variable whenever he makes a new observation or receives the estimate made by another agent. In turn each agent transmits his estimate to a randomly chosen subset of the other agents. If the agents communicate with each other infinitely often, then the estimates of all agents asymptotically agree. Willsky et al. [30] considered the problem of combining and updating estimates that may have been generated in a distributed fashion or may represent estimates, generated at different times, of the same process sample path. In [31], Bertsekas et al, consider a model of

distributed iterative algorithms where several processors participate in the computation, while collecting possibly stochastic information from the environment or other processors via communication links. Issues of asymptotic convergence and agreement are explored under very weak assumptions on the ordering of computations and the timing of information reception. Recent attempts to tackle the network consensus problem under various network topologies are summarized in the survey paper by Ren et al in [32].

More recent relevant literature can be found in [80] where incremental adaptive strategies for distributed parameter estimation have been proposed by Lopes and Syed that require a cyclic pattern of collaboration among the subsystems (nodes). Each subsystem updates the estimate that it receives from its predecessor, based on the information available to it and passes it to its successor to do the same. Employing the incremental adaptive strategies, Rabbat and Nowak address the question that ‘when, in fact, does in-network processing use less energy, and how much energy is saved’ in [33]?. In the same spirit, Li, and Chambers [34] propose an incremental affine projection-based adaptive (APA) algorithm for distributed networks. Lopes and Syed present diffusion techniques in [81], that require more network connectivity. Each subsystem combines its current estimate with the estimate of its neighbor, based on some performance criterion, to come up with an aggregate. This aggregate is then used for carrying out the next estimation update. A similar space-time diffusion approach had earlier been proposed in [82]. An iterative optimization algorithm for a networked system is considered in [83]. Each subsystem (agent) obtains a weighted average of its own iterate with the iterates of its neighbors, and updates the average using the subgradient of its local function to generate the new iterate. Analysis of the estimate difference between the iterates across subsystems (agents) is also presented by considering the difference of each estimate with respect to the running system average of the estimates. Identification of circulant systems is considered in [84] by employing a spatial Fourier transform. The identified data available to each subsystem, however, should be processed centrally in order to construct the global system matrices.

It is interesting to note that the literature cited above concerns only systems with finitely many subsystems. Our interest in this dissertation is to see how ‘local’ consensus might be achieved for identification of unknown common parameter with infinitely many subsystems while at the same time

guaranteeing continuous reduction in the difference of observed and predicted system output.

1.1.2 Distributed Adaptive Control

Several attempts have been made in the last three decades or so to address the problem of decentralized adaptive control of interconnected systems employing various approaches. The most notable early work can be attributed to Ioannou and Kokotovic [35] in this regard, where weakly interconnected subsystems were studied. It was shown that no matter how weak the interconnections may be, a decentralized adaptive control scheme (assuming that the system is decoupled) can become unstable. Singular perturbation theory was used to show boundedness of closed loop signals and convergence of tracking and parameter errors to small residual sets. Later, Ioannou established global stability for a decentralized model reference adaptive controller for interconnected systems in [36] while assuming the relative degree to be less than or equal to two. Taking the M-Matrix approach Ossman showed in [37] that stability of a decentralized system is ensured if there exists a positive definite M-matrix, which is related to the bound of the interconnections. This approach focused on linear subsystems with possibly nonlinear interconnections. An alternative decentralized adaptive control method using the high gain approach was developed in [38], where a standard strict matching condition is assumed on the disturbances. A methodology for handling higher-order interconnections in a decentralized adaptive control framework was developed in [43].

One of the key challenges in decentralized control is the issue of dealing with uncertainty, both in the nonlinearities of the local subsystems as well as in the interconnections. A recent approach for dealing with uncertainty is based on the use of neural networks to approximate the unknown interconnections. This approach is used in [44], and [45], where a decentralized control design scheme was developed for systems with interconnections that are bounded by first-order polynomials. A composite Lyapunov function is employed in [39] for handling both unknown nonlinear model dynamics and interconnections. The interconnections are assumed to be bounded by unknown smooth functions, which are indirectly approximated by neural net-

works. In [46–48] it is assumed that the decentralized controllers share prior information about their reference models, enabling the subsystems to asymptotically track their desired outputs.

No literature exists, to our knowledge, that addresses the problem of adaptive control of LSTI systems. We shall attempt to address this issue as highlighted in the next section.

1.1.3 System Robustness

The contemporary methods of robustness analysis of lumped systems date back to 1960s on feedback stability for nonlinear systems from an input-output perspective [40]. In this category fall the small gain theorem [49, 50] and methods based on passivity, particularly emphasized by the Russian school, where they were termed “absolute stability theory” [51, 52]. The observation that these methods could be used to study uncertain systems was not highlighted until the late 1970s [53] and the connection to \mathcal{H}_∞ norms was pointed out [54]. In particular, the formalization of uncertainty in terms of blocks appeared in [42, 55] and [42] introduced the matrix structured singular value for studying time invariant uncertainty. Subsequently, a great deal of research activity was devoted to its study and to its extension to mixed real parametric and linear time invariant (LTI) uncertainty [56]. Much of this research is summarized in [57–59].

Computing the structured singular value for an arbitrary number of blocks turned out to be NP-hard [60, 61]. In fact [61] also showed that the same holds for multidimensional systems. In [62] it was shown that the gap between the structured singular value and its upper bound is not uniformly bounded with respect to the number of blocks. It was shown in [68] that scaled small gain conditions in the L_∞ -induced norm were exact for the analysis of arbitrary time varying operators. These results led to the study of analogous questions for the case of L_2 norms, leading to the results of [69].

While the field of system robustness with regards to lumped (one dimensional) systems has come a long way, robustness analysis of distributed (multidimensional) systems, on the other hand, has not attracted much attention thus far. The notions of structured uncertainty for linear spatiotemporal invariant (LSTI) systems appeared in [65] for l_∞ signals and in [66] for l_2 signals. In these references, interactions between individual components were

treated as disturbances that satisfy certain magnitude bounds and it was shown that this approach was equivalent to constructing a feedback that is robust stabilizing with structured uncertainties. Perturbation of spatially invariant systems has been considered in [12,13] where symmetry breaking has been introduced for stabilizing the underlying systems. Robust l_2 stability analysis for LSTI systems was carried out for LSTI \mathcal{H}_∞ stable perturbations in [67] and μ -like conditions were established generalizing the results for standard LTI systems.

We shall look at the robust l_∞ and l_2 stability of LSTI systems for other kind of perturbations as described in the next section.

1.2 Contributions and Organization

This dissertation focuses on stability, identification, adaptation, and robustness of spatiotemporal systems. The main contributions are enumerated below:

1. Stability of spatiotemporally varying systems is addressed first when the underlying controllers are designed based on the local frozen linear spatiotemporally invariant approximants. It is shown that the stability can be guaranteed provided the rates of dynamic variations (in space as well as in time) are small enough. Performance is also shown not to be worse than that of the worst frozen LSTI closed loop.
2. System identification of spatiotemporally invariant systems is next addressed in a fashion that the estimated system is amenable to indirect adaptive control. In particular, we propose an identification algorithm that cooperatively establishes a ‘local’ consensus with regards to the estimated system.
3. Based on the first two results, i.e. stability of spatiotemporally varying systems and identification of spatially invariant systems, an indirect adaptive control scheme is presented under certainty equivalence considerations.
4. Necessary and sufficient conditions required to ascertain the robust stability and performance of spatiotemporally invariant systems, in the

presence of various spatiotemporal perturbations, are presented at the end.

The organization of this thesis is presented in the following. Chapter 2 contributes towards providing sufficient conditions for stability and performance of spatiotemporally varying systems. The advancement in sensing and actuating techniques can now allow better lumped approximations of spatiotemporal systems by having densely populated sensors that have fast sampling rates and actuators that have high bandwidths. As the discretization steps in time and space decreases, the dynamics tend to be rather ‘quasi-invariant’ from one step to another (in space as well as time). This motivates us to look at the task of controlling spatiotemporal varying systems in an abstract setting, and as such we seek to answer the following question: under what conditions can a spatiotemporal varying system be stabilized by controllers that are designed based on the local LSTI approximations? In addition, we look into the aspect of not adjusting the controllers for every instance of space and time but use them for some fixed window in time and space before implementing new controllers. The lengths of these windows then enter as additional parameters in the stability analysis. Another motivation for looking at the stability of such systems also comes from the indirect distributed adaptive control of spatially invariant systems, where such systems arise naturally.

In Chapter 3 we present a distributed projection algorithm for identification of spatiotemporally invariant systems. From a control theoretic perspective, we are interested in ascertaining if the following is achieved for a distributed system:

1. Estimation error (difference between actual and predicted output) goes to zero, regardless of the convergence of estimates to the true value
2. Estimates get close to each other arbitrarily as time increases (at least locally)

While the necessity of having 1) is quite clear, the requirement of having 2) is motivated from the possibility of generating an estimated system that is amenable to adaptation using the results of [41], for example. Most of the work cited in the literature review above (and the references therein) concerns with the convergence of the estimates to the true parameter and seldom

addresses the above mentioned objectives. While [83] does address the difference of estimates across the iterates, the analysis cannot be extended to a system having infinitely many subsystems. The objective function, moreover, is not dynamic and the bound provided on the difference of estimates is fairly large, unless the constant step size chosen is very small. Our motivation is exactly summarized in the above mentioned objectives, with the underlying system being spatiotemporally invariant. We propose a distributed projection algorithm for system identification of spatiotemporally invariant systems that builds on the standard projection algorithm [9]. The main idea, however, can be applied to other gradient based schemes such as least squares. Our goal is to achieve the above mentioned objectives together with providing better performance than the standard projection algorithm if run independently on every subsystem.

Chapter 4 presents an indirect adaptive control scheme for spatiotemporally invariant systems and generalizes the results presented in [86]. We base the controller design on certainty-equivalence approach, where at each step system parameters are estimated and the controller is implemented using the estimated parameters. At each estimation stage a modeling error is committed which affects the output of the plant. We show that under suitable assumptions drawn along the lines of the Chapter 2, coupled with the results presented in Chapter 3, yield a globally stable adaptive scheme.

In Chapter 5 we provide robust l_∞ and l_2 stability analysis for spatiotemporally invariant systems. In particular, we provide necessary and sufficient conditions for robust stability of spatiotemporally invariant stable systems in the presence of linear spatiotemporal varying perturbations. The investigation on robust stability of spatiotemporally invariant stable systems with the underlying perturbations being nonlinear spatiotemporal invariant is also presented.

We conclude our discussion in Chapter 6 along with some remarks on future research work.

1.3 General Framework

This dissertation deals with various aspects of spatiotemporal systems. The general framework is shown in Figure 1.1. Throughout this dissertation, we

shall be concerned with an infinite array of interconnected systems as shown in Figure 1.1, where P_i refers to the i_{th} subsystem, and u_i, y_i refer to the input and output respectively. The subsystem P_i is taken as single-input single-output. In the case of spatiotemporally varying systems (presented in Chapter 2), P_i shall be a linear time varying system with the possibility that $P_{i-1} \neq P_i \neq P_{i+1}$. In the case of spatiotemporally invariant systems P_i shall be a linear time invariant system, with $P_i = P_j \forall i, j$. The definitions shall be made precise in their respective chapters.

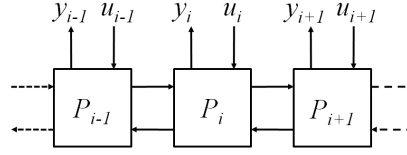


Figure 1.1 General Framework

1.4 Notation and Symbols

The notations common to all chapters are presented here. The set of reals is denoted by \mathbb{R} and the set of integers is denoted by \mathbb{Z} . The set of non-negative integers is denoted by \mathbb{Z}^+ . We use l_∞^e to denote the set of all real double sequences $f = \{f_i(t)\}_{i=-\infty, t=0}^{i=\infty, t=\infty}$. These sequences correspond to spatiotemporal signals with a 2-sided spatial support ($-\infty \leq i \leq \infty$) and one sided temporal ($0 \leq t \leq \infty$). We use l_∞ to denote the set of such sequences with $\|f\|_\infty := \sup_{i,t} |f_i(t)| < \infty$. Similarly, l_2 denotes the set of (double) sequences $f = \{f_i(t)\}_{i=-\infty, t=0}^{i=\infty, t=\infty}$ with $\|f\|_2 := (\sum_{i,t} |f_i(t)|^2)^{1/2} < \infty$. Note that for $f \in l_\infty^e$, we can represent it as a one-sided (causal) temporal sequence as $f = \{f(0), f(1), \dots\}$, where

$$f(t) = \left(\dots, f_{-1}(t), f_0(t), f_{+1}(t) \dots \right)', \quad t \in \mathbb{Z}^+$$

and each $f_j(t) \in \mathbb{R}$, with $j \in \mathbb{Z}$. $\mathcal{B}(x, r)$ denotes a ball in l_2 space with center x and radius r .

CHAPTER 2

STABILITY AND PERFORMANCE OF GRADUALLY VARYING SPATIOTEMPORAL SYSTEMS

This chapter considers the stability of gradually varying spatiotemporal systems. In particular, we restrict our focus to a certain class of discrete distributed systems that have gradually varying dynamics in time as well as in space. We focus on the recursively computable spatiotemporal systems. Recursibility is a property of certain difference equations which allows one to iterate the equation by choosing an indexing scheme so that every output sample can be computed from outputs that have already been found from initial conditions and from samples of the input sequence [95]. The notion of recursibility shall be formalized in the sequel. We show that these gradually varying spatiotemporal systems can be stabilized using the frozen linear spatiotemporally invariant (LSTI) controllers provided the rates of the variations in the spatiotemporal dynamics are sufficiently small. Our result is a generalization of the results on slowly time-varying systems presented in [74] and [76]. We also show that the worst l_∞ to l_∞ performance of global spatiotemporally varying system can be brought arbitrarily close to the worst frozen spatially and temporally l_∞ to l_∞ performance given that the rates of variation of the plant and the controller are sufficiently small. This is a generalization of the result presented in [75].

2.1 Basic Setup

2.1.1 Notations

Systems described by capital letter without subscripts will represent LSTV systems (e.g. A , B). Systems described by capital letters with subscripts

will represent LSTI systems that are obtained by freezing the LSTV systems that the alphabet alone (without any subscripts) would represent. For example $A_{i,t}$ represents an LSTI system obtained by freezing the LSTV system A at time index (t) , and space index (i) . Given a family of spatially invariant operators $\{X_{i,t}\}$ that are indexed in time t and space i , the associated spatiotemporally varying operator X is defined as: $(Xu)_i(t) := (X_{i,t}u)_i(t)$.

2.1.2 Spatiotemporal Varying Systems

Linear spatiotemporally varying systems (LSTV) are systems $M : u \rightarrow y$ on l_∞^e given by the convolution

$$y_i(t) = \sum_{\tau=0}^{\tau=t} \sum_{j=-\infty}^{j=\infty} m_{i,i-j}(t, t-\tau) u_j(\tau) \quad (2.1)$$

where $\{m_{i,j}(t, \tau)\}$ is the kernel representation of M . These systems can be viewed as an infinite interconnection of different linear time varying systems. For simplicity, we assume that each of these subsystems is single-input-single-output (SISO). Let $y_i = (y_i(0), y_i(1), y_i(2), \dots)'$, then the corresponding input-output relationship of the i_{th} block can be given as follows:

$$\begin{aligned} \begin{pmatrix} y_i(0) \\ y_i(1) \\ y_i(2) \\ \vdots \end{pmatrix} &= \begin{pmatrix} m_{i,0}(0,0) & 0 & 0 & \cdots \\ m_{i,0}(1,0) & m_{i,0}(1,1) & 0 & \cdots \\ m_{i,0}(2,0) & m_{i,0}(2,1) & m_{i,0}(2,2) & \cdots \\ \cdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_i(0) \\ u_i(1) \\ u_i(2) \\ \vdots \end{pmatrix} \\ &+ \begin{pmatrix} m_{i,-1}(0,0) & 0 & 0 & \cdots \\ m_{i,-1}(1,0) & m_{i,-1}(1,1) & 0 & \cdots \\ m_{i,-1}(2,0) & m_{i,-1}(2,1) & m_{i,-1}(2,2) & \cdots \\ \cdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_{i-1}(0) \\ u_{i-1}(1) \\ u_{i-1}(2) \\ \vdots \end{pmatrix} \\ &+ \begin{pmatrix} m_{i,1}(0,0) & 0 & 0 & \cdots \\ m_{i,1}(1,0) & m_{i,1}(1,1) & 0 & \cdots \\ m_{i,1}(2,0) & m_{i,1}(2,1) & m_{i,1}(2,2) & \cdots \\ \cdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_{i+1}(0) \\ u_{i+1}(1) \\ u_{i+1}(2) \\ \vdots \end{pmatrix} + \cdots \end{aligned} \quad (2.2)$$

where $\{u_i(t)\}$ is the input applied at the i_{th} block with $u_i(t) \in \mathbb{R}$ and $t \in \mathbb{Z}^+$ is the time index, and $\{m_{i,j}(t, \tau)\}$ is the kernel representation of M . Also, $\{y_i(t)\}$ is the output sequence of the i_{th} block, with $y_i(\cdot) \in \mathbb{R}$. We can write the overall input-output relationship for a LSTV system as follows:

$$\begin{pmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ \vdots \end{pmatrix} = \begin{pmatrix} M^{0,0} & & & & \\ M^{1,0} & M^{1,1} & & & \\ M^{2,0} & M^{2,1} & M^{2,2} & & \\ M^{3,0} & M^{3,1} & M^{3,2} & M^{3,3} & \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \\ u(2) \\ u(3) \\ \vdots \end{pmatrix} \quad (2.3)$$

Where, $u(t) = (\cdots, u_{-1}(t), u_0(t), u_{+1}(t), \cdots)'$ and

$$M^{t,\tau} = \begin{pmatrix} \cdots & \vdots & & \vdots & & \vdots & \cdots \\ \cdots & m_{i-1,0}(t, \tau) & m_{i-1,1}(t, \tau) & m_{i-1,2}(t, \tau) & \cdots \\ \cdots & m_{i,-1}(t, \tau) & m_{i,0}(t, \tau) & m_{i,1}(t, \tau) & \cdots \\ \cdots & m_{i+1,-2}(t, \tau) & m_{i+1,-1}(t, \tau) & m_{i+1,0}(t, \tau) & \cdots \\ \cdots & \vdots & & \vdots & & \vdots & \cdots \end{pmatrix} \quad (2.4)$$

where $t, \tau \in \mathbb{Z}^+$. The l_∞ induced operator norm on M in this case is given as

$$\|M\| = \sup_{i,t} \sum_{\tau=0}^t \sum_{j=-\infty}^{i=\infty} |m_{i,j}(t, \tau)| \quad (2.5)$$

The space of l_∞ bounded LSTV systems will be denoted as \mathcal{L}_{STV}

2.1.3 Spatially Invariant Systems

Linear spatially invariant systems are spatiotemporal systems $M : u \rightarrow y$ on l_∞^e given by the convolution

$$y_i(t) = \sum_{\tau=0}^{\tau=t} \sum_{j=-\infty}^{j=\infty} m_{i-j}(t - \tau) u_j(\tau) \quad (2.6)$$

where $\{m_i(t)\}$ is the impulse response of M . These systems can be viewed as an infinite array of interconnected linear time invariant (LTI) systems. The subspace of \mathcal{L}_{STV} that contains the stable LSTI systems will be denoted as

\mathcal{L}_{STI} . The induced l_∞ operator norm on M in this case is given as

$$\|M\| = \sum_{t=0}^{\infty} \sum_{i=-\infty}^{i=\infty} |m_i(t)| \quad (2.7)$$

2.1.4 Frozen Spatiotemporal Systems

Given a LSTV system M , consider the lower triangular representation of M as shown in (5.1). For any given pair (i, t) (where $i \in \mathbb{Z}$ represents a spatial coordinate, and $t \in \mathbb{Z}^+$ represents time), define a LSTI system $M_{i,t}$ from a pulse response based on the row of $y(t)$ as follows

$$M_{i,t} = \{M^{t,t}, \dots, M^{t,0}, M^{t+1,0}, M^{t+2,0}, \dots\} \quad (2.8)$$

where each $M^{t,\tau}$ is frozen at the i_{th} spatial coordinate, i.e. the i_{th} row is picked in (2.4) and the matrix representation of $M^{t,\tau}$ is assumed to have a toeplitz structure with respect to this row. This ensures that $(M_{i,t}u)_i(t) = (Mu)_i(t)$. We will refer to $M_{i,t}$ as the local or frozen system corresponding to the pair (i, t) . As an example, let $i = 1$, and $t = 3$, then the input output relationship of $M_{1,3}$ can be given as:

$$\begin{pmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ \vdots \end{pmatrix} = \begin{pmatrix} M^{3,3} & & & & \\ M^{3,2} & M^{3,3} & & & \\ M^{3,1} & M^{3,2} & M^{3,3} & & \\ M^{3,0} & M^{3,1} & M^{3,2} & M^{3,3} & \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \\ u(2) \\ u(3) \\ \vdots \end{pmatrix} \quad (2.9)$$

Where, $u(t) = (\dots, u_{-1}(t), u_0(t), u_{+1}(t), \dots)'$ and

$$M^{t,\tau} = \begin{pmatrix} \dots & \vdots & \vdots & \vdots & \dots \\ \dots & m_{0,0}(t, \tau) & m_{0,1}(t, \tau) & m_{0,2}(t, \tau) & \dots \\ \dots & m_{0,-1}(t, \tau) & m_{0,0}(t, \tau) & m_{0,1}(t, \tau) & \dots \\ \dots & m_{0,-2}(t, \tau) & m_{0,-1}(t, \tau) & m_{0,0}(t, \tau) & \dots \\ \dots & \vdots & \vdots & \vdots & \dots \end{pmatrix} \quad (2.10)$$

Let A and B be two LSTV systems. We define a global product AB to mean

the usual composition of these two systems. Given a pair (i, t) , the local product of LSTV systems A and B corresponds to the product (composition) of the frozen LSTI systems $A_{i,t}$ and $B_{i,t}$, i.e. $A_{i,t}B_{i,t}$.

2.1.5 Support of m

We define the support of a sequence $\{m_i(t)\}$ by $Supp(m)$, i.e.

$$Supp(m) = \{[i, t] \in \mathbb{Z}^2 : m_i(t) \neq 0\} \quad (2.11)$$

2.1.6 Recursively Computable Spatiotemporal Systems

A LSTV system P is called recursively computable spatiotemporal system with the input-output relationship defined by an equation of the form

$$(A_{i,t}y)_i(t) = (B_{i,t}u)_i(t) \quad (2.12)$$

where y is the output and u is the input. With $\{a_{i,j}(t, \tau)\}, \{b_{i,j}(t, \tau)\}$ being the kernel representations of the operators $A_{i,t}, B_{i,t}$ in \mathcal{L}_{STI} respectively, we can write the above equation explicitly as follows;

$$\sum_{\substack{j \\ (j,\tau) \in I_{a(i,t)}}} \sum_{\tau} a_{i,j}(t, \tau) y_{i-j}(t - \tau) = \sum_{\substack{j \\ (j,\tau) \in I_{b(i,t)}}} \sum_{\tau} b_{i,j}(t, \tau) u_{i-j}(t - \tau) \quad (2.13)$$

where $I_{a(i,t)}$ (output mask) and $I_{b(i,t)}$ (input mask) denote, respectively, the finite area region of support for $\{a_{i,j}(t, \tau)\}$ and $\{b_{i,j}(t, \tau)\}$. The system in (2.13) is well defined if $\{a_{i,0}(t, 0)\} \neq 0$, and $\{a_{i,j}(t, \tau)\} \neq 0$ for some (j, τ) , and $Supp(\{a_{i,j}(t, \tau)\})$ is a subset of the lattice sector with vertex $(0, 0)$ of angle less than 180° , for every pair (i, t) [95] (a general schematic is shown in Figure 2.1). We will assume that all the spatiotemporal systems under consideration are well defined.

2.1.7 Gradually Varying Spatiotemporal System

A LSTV system A is said to be gradually space-time varying if given two pairs (i, t) , and (i, τ) , the following holds for the corresponding frozen systems

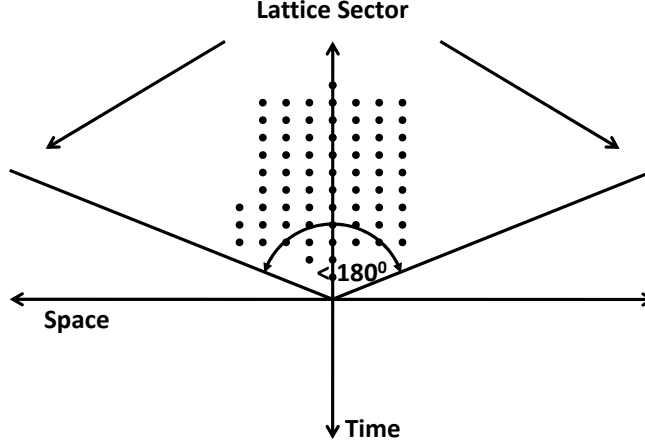


Figure 2.1 Finite Area Region of Support With Lattice Sector Having Angle Less Than 180°

$A_{i,t}, A_{\mathbf{i},\tau};$

$$\|A_{i,t} - A_{\mathbf{i},\tau}\| \leq \gamma(|i - \mathbf{i}| + |t - \tau|) \quad (2.14)$$

where $\gamma \in \mathbb{Z}^+$ is a constant. Such systems will be denoted by $\text{GSTV}(\gamma)$

2.1.8 Integral Time and Space Absolute Error

Given a LSTI system M , the integral time and space absolute error (ITSAE) is defined as

$$ITSAE(M) = \sum_{t=0}^{\infty} \sum_{i=-\infty}^{\infty} (|i| + |t|) |m_i(t)| \quad (2.15)$$

2.1.9 z, λ Transform

We define the z, λ transform for a LSTI SISO system M as

$$\hat{M}(z, \lambda) = \sum_{t=0}^{\infty} \sum_{k=-\infty}^{\infty} (m_k(t) z^k) \lambda^t \quad (2.16)$$

with the associated spectral or \mathcal{H}_∞ norm

$$\|\hat{M}\|_\infty := \sup_{\theta, \omega} |\hat{M}(e^{i\theta}, e^{j\omega})| \quad (2.17)$$

It is well known (see e.g. [95]) that for a system $M \in \mathcal{L}_{STI}$, M^{-1} is in \mathcal{L}_{STI} if and only if

$$\inf_{|z|=1, |\lambda| \leq 1} |\hat{M}(z, \lambda)| > 0 \quad (2.18)$$

2.2 Frozen Space-Time Control

Consider the general form of a closed loop system given in Figure 2.2. The plant P is a LSTV recursively computable spatiotemporal system with the input-output relationship defined by an equation of the form

$$(A_{i,t}y_1)_i(t) = (B_{i,t}y_4)_i(t) \quad (2.19)$$

with $\{a_{i,j}(t, \tau)\}$, $\{b_{i,j}(t, \tau)\}$, being the kernel representations of the operators $A_{i,t}$, $B_{i,t}$ in \mathcal{L}_{STI} respectively, we can write the above equation explicitly as follows:

$$\sum_{\substack{j \\ (j,\tau) \in I_{a(i,t)}}} \sum_{\tau} a_{i,j}(t, \tau) y_{1,i-j}(t - \tau) = \sum_{\substack{j \\ (j,\tau) \in I_{b(i,t)}}} \sum_{\tau} b_{i,j}(t, \tau) y_{4,i-j}(t - \tau) \quad (2.20)$$

The control law is designed on the basis of frozen-time and frozen-space plants. Given an instance in space and time, the plant is thought of as a LSTI system, with the defining operators fixed at that time and space. The controllers are designed for the corresponding frozen LSTI system. Allowing ourselves the flexibility of using a designed controller for several instances in time and space, we will consider the controller design every T steps in time and every S steps in space. Define $n_t = nT$ and $k_i = kS$, where n and k are smallest integers such that t and i lie in the interval $[nT, (n+1)T]$ and $[kS, (k+1)S]$ respectively. The controller is designed at intervals of nT , and kS in time and space respectively. We note here that this abstract setting also addresses the control problem of distributed systems where each subsystem has knowledge of its own dynamics, but otherwise only knows that

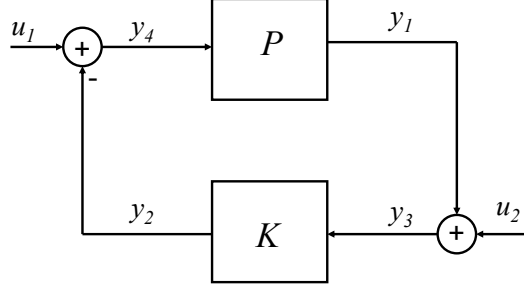


Figure 2.2 General Form of Closed Loop

the rest of the subsystems are ‘similar’. The closed loop is stable if the map from u_1, u_2 to y_1, y_2 is bounded. The dynamics of the control law K are given by

$$(L_{k_i, n_t} y_2)_i(t) = (M_{k_i, n_t} y_3)_i(t) \quad (2.21)$$

where $L_{k_i, n_t}, M_{k_i, n_t} \in \mathcal{L}_{STI}$ for each pair of indices (k_i, n_t) . The evolution of these operators is given by

$$(L_{k_i, n_t} y_2)_i(t) = \sum_j \sum_{\substack{\tau \\ (j, \tau) \in I_l(k_i, n_t)}} l_{k_i, i-j}(n_t, t - \tau) y_{2,j}(\tau) \quad (2.22)$$

$$(M_{k_i, n_t} y_3)_i(t) = \sum_j \sum_{\substack{\tau \\ (j, \tau) \in I_m(k_i, n_t)}} m_{k_i, i-j}(n_t, t - \tau) y_{3,j}(\tau) \quad (2.23)$$

where $I_l(k_i, n_t)$ and $I_m(k_i, n_t)$ denote, respectively, the finite area region of support for $\{l_{k_i, j}(n_t, \tau)\}$ and $\{m_{k_i, j}(n_t, \tau)\}$. The frozen space and time operator that defines the above control law stabilizes the following Bezout identity

$$L_{k_i, n_t} A_{k_i, n_t} + M_{k_i, n_t} B_{k_i, n_t} = G_{k_i, n_t} \quad (2.24)$$

where $G_{k_i, n_t}^{-1} \in \mathcal{L}_{STI}$ for each fixed pair (k_i, n_t) . That is, for every frozen plant given by $A_{k_i, n_t}, B_{k_i, n_t}$, the control generated by $L_{k_i, n_t}, M_{k_i, n_t}$ is such that the “frozen” closed loop map G_{k_i, n_t}^{-1} is stable. Note that the frozen plant

is LSTI, and hence a frozen LSTI controller that satisfies the frozen closed loop can be obtained using various methods, e.g. [16], [17]. Here, we are not interested in any specific method. We only require that K operates as described above and provides frozen stability.

We investigate the interaction of the gradual variations of the plant and the controller in time as well as in space. In particular, we show how these variations play their role with regards to stability of the closed loop. The fact that the controller is updated only every T steps in time and after every S number of plants in space introduces a new parameter in the stability analysis. In the sequel, we show as to how large T and S can be without endangering the stability of the closed loop system. Intuitively, the larger the T and S are, the slower the plant variations should be in the respective time and spatial domains. On the other hand, for the extreme case such that $T \rightarrow \infty$, the system should be time invariant, and for $S \rightarrow \infty$, the system should be spatially invariant. Equivalently, if we require $S, T \rightarrow \infty$ simultaneously, the system should be LSTI. From Figure 2.2, we can write the closed loop equations for the controlled system as follows:

$$(A_{i,t}y_1)_i(t) = (B_{i,t}(u_1 - y_2))_i(t) \quad (2.25)$$

$$(L_{k_i,n_t}y_2)_i(t) = (M_{k_i,n_t}(u_2 + y_1))_i(t) \quad (2.26)$$

$$L_{k_i,n_t}A_{k_i,n_t} + M_{k_i,n_t}B_{k_i,n_t} = G_{k_i,n_t} \quad (2.27)$$

In the following we obtain a relation that connects the input sequences $\{u_{1,i}(t)\}$, $\{u_{2,i}(t)\}$ to the outputs $\{y_{1,i}(t)\}$ and $\{y_{2,i}(t)\}$. Operating on equation (2.25) by L_{k_i,n_t} , we get

$$(LAy_1)_i(t) = (LBu_1)_i(t) - (LBy_2)_i(t) \quad (2.28)$$

Adding, subtracting, and grouping certain terms we get:

$$\begin{aligned} & \{(L_{k_i,n_t}A_{k_i,n_t} + B_{k_i,n_t}M_{k_i,n_t})y_1 + (L_{k_i,n_t}\nabla A_{i,t} + (L_{k_i,n_t}A_{i,t} - L_{k_i,n_t}A_{k_i,n_t}) \\ & + B_{i,t}\nabla M_{k_i,n_t} + (B_{i,t}M_{k_i,n_t} - B_{k_i,n_t}M_{k_i,n_t}))y_1 \\ & + (L_{k_i,n_t}\nabla B_{i,t} - B_{i,t}\nabla L_{k_i,n_t})y_2\}_i(t) = (LBu_1)_i(t) - (BMu_2)_i(t) \end{aligned} \quad (2.29)$$

where we have used the notation; $A_{i,t} \nabla B_{i,t} = AB - A_{i,t} B_{i,t}$, i.e., $A_{i,t} \nabla B_{i,t}$ is the difference between the global and local product of operators given a pair (i, t) . To obtain a second closed loop equation, we operate on equation (2.25) by M_{k_i, n_t} :

$$(MAy_1)_i(t) = (MBu_1)_i(t) - (MB y_2)_i(t) \quad (2.30)$$

Again adding, subtracting, and grouping certain terms we get:

$$\begin{aligned} & \{(M_{k_i, n_t} B_{k_i, n_t} + A_{k_i, n_t} L_{k_i, n_t}) y_2 + (M_{k_i, n_t} \nabla B_{i,t} + (M_{k_i, n_t} B_{i,t} - M_{k_i, n_t} B_{k_i, n_t}) \\ & + A_{i,t} \nabla L_{k_i, n_t} + (A_{i,t} L_{k_i, n_t} - A_{k_i, n_t} L_{k_i, n_t})) y_2 \\ & + (A_{i,t} \nabla M_{k_i, n_t} - M_{k_i, n_t} \nabla A_{i,t}) y_1\}_i(t) = (MBu_1)_i(t) + (AMu_2)_i(t) \end{aligned} \quad (2.31)$$

For $t \in \mathbb{Z}^+$, $i \in \mathbb{Z}$, define the following

$$\begin{aligned} X_{i,t} &= L_{k_i, n_t} \nabla A_{i,t} + (L_{k_i, n_t} A_{i,t} \\ & - L_{k_i, n_t} A_{k_i, n_t}) + B_{i,t} \nabla M_{k_i, n_t} + (B_{i,t} M_{k_i, n_t} - B_{k_i, n_t} M_{k_i, n_t}) \end{aligned} \quad (2.32)$$

$$Y_{i,t} = L_{k_i, n_t} \nabla B_{i,t} - B_{i,t} \nabla L_{k_i, n_t} \quad (2.33)$$

$$Z_{i,t} = M_{k_i, n_t} \nabla A_{i,t} - A_{i,t} \nabla M_{k_i, n_t} \quad (2.34)$$

$$\begin{aligned} W_{i,t} &= M_{k_i, n_t} \nabla B_{i,t} + (M_{k_i, n_t} B_{i,t} - M_{k_i, n_t} B_{k_i, n_t}) + A_{i,t} \nabla L_{k_i, n_t} \\ &+ (A_{i,t} L_{k_i, n_t} - A_{k_i, n_t} L_{k_i, n_t}) \end{aligned} \quad (2.35)$$

Denote by X, Y, Z, W, G the spatiotemporally varying operators associated with the families $X_{i,t}, Y_{i,t}, Z_{i,t}, W_{i,t}, G_{k_i, n_t}$ respectively. Using (2.27) we can write the closed loop equation as follows:

$$\left(\begin{bmatrix} G + X & Y \\ -Z & G + W \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right)_i(t) = \left(\begin{bmatrix} LB & -BM \\ MB & AM \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right)_i(t) \quad (2.36)$$

The idea is to analyze the above system by considering the operators X, Y, Z, W as perturbations. We state below the main result of this chapter regarding stability of the system given in (2.36). We prove this result in the next section.

Theorem 2.2.1. *Assume the following for system (2.36):*

1. *The operators defining the plant are gradually time and space varying*

with rates γ_A and γ_B , i.e. $A_{i,t} \in GSTV(\gamma_A)$, and $B_{i,t} \in GSTV(\gamma_B)$.

2. The sequence of controllers are gradually time and space varying, i.e. $M_{k_i,n_t} \in GSTV(\gamma_M)$ and $L_{k_i,n_t} \in GSTV(\gamma_L)$.
3. The l_∞ induced norms and the ITSAE of the operators $A_{i,t}$, $B_{i,t}$, L_{k_i,n_t} , M_{k_i,n_t} are uniformly bounded in i , and t . From this and 1, 2, and the Bezout identity it follows that the operator G_{k_i,n_t} will also be gradually varying in space and time and, hence, we can write $G_{k_i,n_t} \in GSTV(\gamma_G)$
4. The l_∞ to l_∞ norms and the ITSAE of the LSTI operators G_{k_i,n_t}^{-1} are bounded uniformly in i , and t .

Then there exists a non-zero constant γ such that if γ_A , γ_B , γ_M , γ_L , $\gamma_G \leq \gamma$, the closed loop system is internally stable.

2.3 Stability Analysis

In this section we study the stability of the closed loop system arising from the frozen time and space control design. From equation (2.36) we see that the map G_{k_i,n_t} is perturbed by a few operators, each of which falls in one of the two categories:

1. $A_{i,t} \nabla L_{k_i,n_t}$
2. $L_{k_i,n_t} (A_{i,t} - A_{k_i,n_t})$

In the following lemmas we show how the l_∞ induced norms of these operators can be made small by controlling the rates of spatiotemporal variations involved in the problem at hand.

Lemma 2.3.1. *Let $L_{k_i,n_t} \in GSTV(\gamma_L)$, and $A_{i,t} \in GSTV(\gamma_A)$ then*

$$A_{i,t} \nabla L_{k_i,n_t} \in \mathcal{L}_{STV}$$

and its induced norm satisfies

$$\begin{aligned}
\|A_{i,t}\nabla L_{k_i,n_t}\| &= \sup_{i,t} \|A_{i,t}\nabla L_{k_i,n_t}\| \\
&\leq \gamma_L \left(2(S+T) \sup_{i,t} \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,j}(t,\tau)| \right. \\
&\quad \left. + \sup_{i,t} \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |j| |a_{i,j}(t,\tau)| + \sup_{i,t} \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} \tau |a_{i,j}(t,\tau)| \right)
\end{aligned} \tag{2.37}$$

Proof. Let $u \in l_\infty$, then the operator $A_{i,t}\nabla L_{k_i,n_t}$ acts on u as follows

$$\begin{aligned}
&(A_{i,t}\nabla L_{k_i,n_t}u)_i(t) \\
&= \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} a_{i,i-j}(t,t-\tau) \left(\sum_{r=0}^{\tau} \sum_{s=-\infty}^{\infty} (l_{k_j,j-s}(n_\tau,\tau-r) - l_{k_i,i-s}(n_t,\tau-r)) u_s(r) \right)
\end{aligned} \tag{2.38}$$

Taking the absolute value of the above equation we get:

$$\begin{aligned}
&|A_{i,t}\nabla L_{k_i,n_t}u_i(t)| \\
&\leq \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,i-j}(t,t-\tau)| \left(\sum_{r=0}^{\tau} \sum_{s=-\infty}^{\infty} |l_{k_j,j-s}(n_\tau,\tau-r) - l_{k_i,i-s}(n_t,\tau-r)| \|u\|_\infty \right)
\end{aligned} \tag{2.39}$$

$$= \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,i-j}(t,t-\tau)| \left(\sum_{r=0}^{\tau} \sum_{s=-\infty}^{\infty} |l_{k_j,j-s}(n_\tau,r) - l_{k_i,i-s}(n_t,r)| \|u\|_\infty \right) \tag{2.40}$$

$$\leq \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,i-j}(t,t-\tau)| \|L_{k_j,n_\tau} - L_{k_i,n_t}\| \|u\|_\infty \tag{2.41}$$

$$\leq \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,i-j}(t,t-\tau)| \gamma_L (|k_j - k_i| + |n_\tau - n_t|) \|u\|_\infty \tag{2.42}$$

Now,

$$\begin{aligned}
|k_j - k_i| + |n_\tau - n_t| &= |k_j - j + j - i + i - k_i| + |n_\tau - \tau + \tau - t + t - n_t| \\
&\leq 2S + 2T + |j - i| + |\tau - t|
\end{aligned} \tag{2.43}$$

since, $|k_j - j| \leq S$, and $|n_\tau - \tau| \leq T$. The above inequality can now be written as:

$$\begin{aligned}
& |A_{i,t} \nabla L_{k_i,n_t} u_i(t)| \\
& \leq \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,i-j}(t, t-\tau)| \gamma_L (2S + 2T + |j-i| + |\tau-t|) \|u\|_{\infty} \quad (2.44) \\
& \leq \gamma_L \left(2(T+S) \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |a_{i,j}(t, \tau)| + \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |j| |a_{i,j}(t, \tau)| \right. \\
& \quad \left. + \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} \tau |a_{i,j}(t, \tau)| \right) \|u\|_{\infty} \quad (2.45)
\end{aligned}$$

□

Lemma 2.3.2. *Let the assumptions in Lemma 2.3.1 hold. $L_{k_i,n_t}(A_{i,t} - A_{k_i,n_t}) \in LSTV$ and its induced norm satisfies*

$$\begin{aligned}
\|L_{k_i,n_t}(A_{i,t} - A_{k_i,n_t})\| &= \sup_{i,t} \|L_{k_i,n_t}(A_{i,t} - A_{k_i,n_t})\| \\
&\leq \gamma_A(T+S) \sup_{i,t} \sum_{\tau=0}^t \sum_{j=-\infty}^{\infty} |l_{k_i,j}(n_t, \tau)| \quad (2.46)
\end{aligned}$$

Proof. The proof follows in a similar fashion as above and is hence omitted. □

We now proceed to present the proof of Theorem 2.2.1.

Proof of Theorem 2.2.1: Consider the first equation in (2.36), expressed in operator form,

$$Gy_1 + Xy_1 + Yy_2 = v \quad (2.47)$$

where $v_i(t) = (LBu_1)_i(t) - (BMu_2)_i(t)$. Let (i, τ) be a fixed instance in space and time, we can write

$$G_{k_i,n_\tau} y_1 + (G - G_{k_i,n_\tau}) y_1 + Xy_1 + Yy_2 = v \quad (2.48)$$

where $G_{k_i, \tau} \in \mathcal{L}_{STI}$. Denote by $H_{k_i, \tau}$ the inverse of $G_{k_i, \tau}$. By assumption (4), $H_{k_i, \tau} \in \mathcal{L}_{STI}$. The above equation can, therefore, be written as

$$y_1 + H_{k_i, n_\tau}(G - G_{k_i, n_\tau})y_1 + H_{k_i, n_\tau}Xy_1 + H_{k_i, n_\tau}Yy_2 = H_{k_i, n_\tau}v \quad (2.49)$$

Evaluating the above operator equation at (i, τ) we obtain

$$\begin{aligned} y_{1,i}(\tau) &+ (H_{k_i, n_\tau}(G - G_{k_i, n_\tau})y_1)_i(\tau) + (H_{k_i, n_\tau}Xy_1)_i(\tau) + (H_{k_i, n_\tau}Yy_2)_i(\tau) \\ &= (H_{k_i, n_\tau}v)_i(\tau). \end{aligned} \quad (2.50)$$

Similarly we can write

$$\begin{aligned} -(H_{k_i, n_\tau}Zy_1)_i(\tau) &+ y_{2,i}(\tau) + (H_{k_i, n_\tau}(G - G_{k_i, n_\tau})y_2)_i(\tau) + (H_{k_i, n_\tau}Wy_2)_i(\tau) \\ &= (H_{k_i, n_\tau}w)_i(\tau) \end{aligned} \quad (2.51)$$

where $w_i(t) = (MBu_1)_i(t) + (AMu_2)_i(t)$. Combining the above equations, we get the following closed loop system:

$$\left(\begin{pmatrix} I + F \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right)_i(\tau) = \begin{pmatrix} H_{k_i, n_\tau}v \\ H_{k_i, n_\tau}w \end{pmatrix}_i(\tau) \quad (2.52)$$

where

$$F = \begin{pmatrix} H_{k_i, n_\tau}(G - G_{k_i, n_\tau}) + H_{k_i, n_\tau}X & H_{k_i, n_\tau}Y \\ -H_{k_i, n_\tau}Z & H_{k_i, n_\tau}(G - G_{k_i, n_\tau}) + H_{k_i, n_\tau}W \end{pmatrix} \quad (2.53)$$

The idea is to show that the induced norm of the spatiotemporally varying perturbing operator F can be made less than one by choosing the rates of variations sufficiently small. From the previous lemmas, and the fact that H_{k_i, n_τ} is uniformly bounded, it is clear that each of the spatiotemporally varying operators generated from each family of operators $H_{k_i, n_\tau}X$, $H_{k_i, n_\tau}Y$, $H_{k_i, n_\tau}Z$, $H_{k_i, n_\tau}W$, have induced norms that are controlled by the rates of variation γ_A , γ_B , γ_L , γ_M , γ_G . The internal stability will follow from the small gain theorem if we show that the induced norm of the operator $H_{k_i, n_\tau}(G - G_{k_i, n_\tau})$ can be analogously controlled. We present in the following a calculation of an upper bound of the norm of the operator $H_{k_i, n_\tau}(G - G_{k_i, n_\tau})$. Let $y \in l_\infty$

and the output of the system be x , then

$$\begin{aligned} x_{\mathbf{i}}(\tau) &= \sum_{m=0}^{\tau} \sum_{j=-\infty}^{\infty} h_{k_{\mathbf{i}}, \mathbf{i}-j}(n_{\tau}, \tau - m) \\ &\quad \times \sum_{r=0}^m \sum_{s=-\infty}^{\infty} (g_{k_j, j-s}(n_m, m - r) - g_{k_{\mathbf{i}}, \mathbf{i}-s}(n_{\tau}, m - r)) y_s(r) \end{aligned} \quad (2.54)$$

Taking absolute values we get,

$$\begin{aligned} |x_{\mathbf{i}}(\tau)| &\leq \sum_{m=0}^{\tau} \sum_{j=-\infty}^{\infty} |h_{k_{\mathbf{i}}, \mathbf{i}-j}(n_{\tau}, \tau - m)| \\ &\quad \times \sum_{r=0}^m \sum_{s=-\infty}^{\infty} |g_{k_j, j-s}(n_m, m - r) - g_{k_{\mathbf{i}}, \mathbf{i}-s}(n_{\tau}, m - r)| \|y\|_{\infty} \end{aligned} \quad (2.55)$$

By an argument similar to one given in the proof of Lemma 1, it follows that:

$$\begin{aligned} \|x\|_{\infty} &\leq \gamma_G \left(2(T + S) \sup_{\mathbf{i}, \tau} \sum_{m=0}^{\tau} \sum_{j=-\infty}^{\infty} |h_{k_{\mathbf{i}}, j}(n_{\tau}, m)| \right. \\ &\quad \left. + \sup_{\mathbf{i}, \tau} \sum_{m=0}^{\tau} \sum_{j=-\infty}^{\infty} m |h_{k_{\mathbf{i}}, j}(n_{\tau}, m)| + \sup_{\mathbf{i}, \tau} \sum_{m=0}^{\tau} \sum_{j=-\infty}^{\infty} |j| |h_{k_{\mathbf{i}}, j}(n_{\tau}, m)| \right) \|y\|_{\infty} \end{aligned} \quad (2.56)$$

It follows by assumption (4) that there exist constants $C_1, C_2 \geq 0$ such that

$$\|x\|_{\infty} \leq \gamma_G (2(S + T)C_1 + C_2 + C_3) \|y\|_{\infty} \quad (2.57)$$

We have, hence, shown that the induced norms of all the perturbing operators that comprise F can be made small by choosing the rates of variations small enough. Internal stability now follows by an application of the small gain theorem. This concludes the proof of Theorem 2.2.1.

The main idea behind the proof of Theorem 2.2.1 is captured in Figure 2.3. The Δ 's on the top capture the time and space varying mismatch between the actual system dynamics and the one assumed by the controller. We note here that a closed form expression of these Δ 's cannot be given as it would depend on the specific system dynamics. Controllers are shown to be updated after every three subsystems (this may not be the case in reality as

it would depend on the specific design). We emphasize here that Figure 2.3 is only presented to convey the main idea from an intuitive standpoint, and an exact correspondence with the proof should not be made.

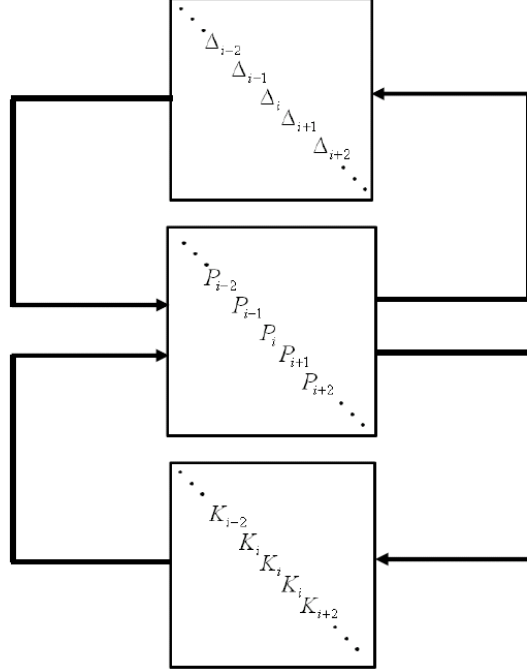


Figure 2.3 Main Idea for Proof of Theorem 2.2.1

Theorem 2.2.1 shows that if the assumptions (1-4) are satisfied and if the variations are small enough, then the closed loop system will be l_∞ stable. The assumptions (1-2) are quite reasonable and are typically satisfied for the recursively computable spatiotemporal system that we focus on. The first part of assumption (3), requiring uniform bounds on the operators, is also quite reasonable. Intuitively the second part of assumption (3), that requires uniform bounds on the ITSAE of operators, implies that the LTV building blocks of the LSTV system have decaying memory (temporal), and decaying spatial dependence on the neighbors (as one goes away from the reference in space). Assumption (4), however, is harder to satisfy. This assumption implies that the zeros in λ of $\hat{G}_{i,t}(e^{j\theta}, \lambda)$, lie outside a disc of radius $1 + \epsilon$, for some $\epsilon > 0$ and for all θ . Precisely, this assumption implies:

$$\inf_{|z|=1, |\lambda| \leq 1} |\hat{G}_{i,t}(z, \lambda)| \geq \delta > 0, \quad \forall i, \forall t \quad (2.58)$$

Satisfying this condition only, however, does not imply that the l_∞ induced norms and the ITSAE of $G_{i,t}^{-1}$ are uniformly bounded, i.e., assumption (4) is not satisfied in general. The above spectral condition does imply that the \mathcal{H}_∞ norm of $G_{i,t}^{-1}$ is uniformly bounded in i , and t , since:

$$\left\| \hat{G}_{i,t}^{-1} \right\|_\infty = \sup_{|z|=1, |\lambda| \leq 1} \frac{1}{|\hat{G}_{i,t}(z, \lambda)|} = \frac{1}{\inf_{|z|=1, |\lambda| \leq 1} |\hat{G}_{i,t}(z, \lambda)|} \leq \frac{1}{\delta}, \quad \forall i, \forall t \quad (2.59)$$

In the following we show that with some additional mild assumptions on the \mathcal{H}_∞ norm of $\hat{G}_{i,t}(z, \lambda)$ and its partial derivatives, the spectral condition is enough to verify the uniform bounds on the l_∞ induced norms and the ITSAE of $G_{i,t}^{-1}$. For partial derivatives, we define the notation $\hat{G}_{i,t,(\xi)} = \frac{\partial \hat{G}_{i,t}}{\partial \xi}$, $\hat{G}_{i,t,(\xi\xi)} = \frac{\partial^2 \hat{G}_{i,t}}{\partial \xi \partial \xi}$, $\hat{G}_{i,t,(\xi\xi\xi)} = \frac{\partial^3 \hat{G}_{i,t}}{\partial \xi \partial \xi \partial \xi}$, where ξ, ζ can be z or λ .

Theorem 2.3.1. *Given the following conditions:*

1. $\left\| \hat{G}_{i,t} \right\|_\infty \leq M_1, \forall i, t$
2. $\left\| \hat{G}_{i,t,(z)} \right\|_\infty \leq M_2, \forall i, t$
3. $\left\| \hat{G}_{i,t,(\lambda)} \right\|_\infty \leq M_3, \forall i, t$
4. $\left\| \hat{G}_{i,t,(zz)} \right\|_\infty \leq M_4, \forall i, t$
5. $\left\| \hat{G}_{i,t,(\lambda\lambda)} \right\|_\infty \leq M_5, \forall i, t$
6. $\left\| \hat{G}_{i,t,(z\lambda)} \right\|_\infty \leq M_6, \forall i, t$
7. $\left\| \hat{G}_{i,t,(zz\lambda)} \right\|_\infty \leq M_7, \forall i, t$
8. $\left\| \hat{G}_{i,t,(\lambda zz)} \right\|_\infty \leq M_8, \forall i, t$

9. $\inf_{|z|=1, \lambda \leq 1} |\hat{G}_{i,t}(z, \lambda)| \geq \delta > 0 \quad \forall i, t \text{ (spectral condition)}$

Then the l_∞ induced norm and ITSAE of $G_{i,t}^{-1}$ are uniformly bounded in i , and t .

Proof. The proof is based on Hardy's theorem [78], and its extension for two variables. We present Hardy's theorem in one variable along with its extension form in two variables without proof in the following. Given a function $\hat{R} \in \mathcal{H}_\infty$ with

$$\hat{R}(\lambda) = \sum_{t=0}^{\infty} m(t) \lambda^t \quad (2.60)$$

there exists a constant $0 < C < +\infty$ such that the coefficients $m(t)$ satisfy:

$$\sum_{t=1}^{\infty} \frac{1}{t} |m(t)| \leq C \|\hat{R}\|_\infty \quad (2.61)$$

Similarly, if we have a function $\hat{R} \in \mathcal{H}_\infty$ with

$$\hat{R}(z, \lambda) = \sum_{t=0}^{\infty} \sum_{k=-\infty}^{\infty} (m_k(t) z^k) \lambda^t \quad (2.62)$$

then there exists a constant $0 < C < +\infty$ such that the coefficients $m_k(t)$ satisfy:

$$\sum_{t=1}^{\infty} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \frac{1}{t|k|} |m_k(t)| \leq C \|\hat{R}\|_\infty \quad (2.63)$$

To show that $\|G_{i,t}^{-1}\|$ is uniformly bounded, we apply Hardy's theorem on $\hat{R}(z, \lambda) = \hat{G}_{i,t,(z\lambda)}^{-1}$. Note that

$$\frac{\partial^2 \hat{G}_{i,t}^{-1}}{\partial z \partial \lambda} = \frac{-\hat{G}_{i,t,(z\lambda)} \hat{G}_{i,t} + 2\hat{G}_{i,t,(z)} G_{i,t,(\lambda)}}{\hat{G}_{i,t}^3} \quad (2.64)$$

Hence,

$$\|\hat{R}\|_\infty \leq \frac{\|\hat{G}_{i,t,(z\lambda)}\|_\infty \|\hat{G}_{i,t}\|_\infty + 2\|\hat{G}_{i,t,(z)}\|_\infty \|G_{i,t,(\lambda)}\|_\infty}{\delta^3} \leq \frac{M_9}{\delta^3} \quad (2.65)$$

Where $M_9 := M_6M_1 + 2M_2M_3$. Let $\hat{G}_{i,t}^{-1}$ be given by:

$$\hat{G}_{i,t}^{-1} = \sum_{\tau=0}^{\infty} \sum_{k=-\infty}^{\infty} (h_{i,k}(t, \tau) z^k) \lambda^\tau \quad (2.66)$$

Then,

$$\hat{R}_{i,t} = \sum_{\tau=0}^{\infty} \sum_{k=-\infty}^{\infty} \tau k (h_{i,k}(t, \tau) z^{k-1}) \lambda^{\tau-1} \quad (2.67)$$

Applying Hardy's theorem on $z\lambda\hat{R}_{i,t}$, we get

$$\sum_{\tau=1}^{\infty} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} |h_{i,k}(t, \tau)| \leq C \frac{M_9}{\delta} \quad (2.68)$$

Note that in the last expression above, we are missing the terms

$$\sum_{k=-\infty}^{\infty} |h_{i,k}(t, 0)|, \text{ and } \sum_{\tau=0}^{\infty} |h_{i,0}(t, \tau)|$$

in order to establish a bound on $\|G_{i,t}^{-1}\|$. Let

$$\hat{h}1 = \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} h_{i,k}(t, 0) z^k; \quad \hat{h}2 = \sum_{\tau=0}^{\infty} h_{i,0}(t, \tau) \lambda^\tau \quad (2.69)$$

The \mathcal{H}_∞ norms of the operators $\hat{h}1$, $\hat{h}2$ are finite, since $\|\hat{G}_{i,t}^{-1}\|_\infty$ is finite. Also the \mathcal{H}_∞ norms of $\hat{h}1_{(z)}$, and $\hat{h}2_{(\lambda)}$ are finite since $\|\hat{R}_{i,t}\|_\infty$ is finite. Using the Hardy's theorem again and reasoning in a similar fashion as above, we can establish bounds on, $\sum_{k=-\infty, k \neq 0}^{\infty} |h_{i,k}(t, 0)|$, and $\sum_{\tau=1}^{\infty} |h_{i,0}(t, \tau)|$. Let the sum of their bounds be denoted by C_h . We now have the following bound for $\|G_{i,t}^{-1}\|$:

$$\|G_{i,t}^{-1}\| \leq C \frac{M_9}{\delta} + C_h + |h_{i,0}(t, 0)| \leq C \frac{M_9}{\delta} + C_h + \delta^{-1} \quad (2.70)$$

A similar argument works for $\text{ITSAE}(G_{i,t}^{-1})$ by considering $G_{i,t,(zz\lambda)}^{-1}$, and $G_{i,t,(\lambda\lambda z)}^{-1}$. We omit the details as they follow the above argument. \square

The benefit of the above theorem is that one can check if assumption (4) is satisfied by checking easily computable \mathcal{H}_∞ norms.

Theorem 2.3.2. *Let the assumptions (1-3), and the spectral condition (9, Theorem 2.3.1) hold along with the uniform boundedness of the following quantities*

- $\left\| \hat{A}_{i,t,(zz)} \right\|_\infty, \left\| \hat{A}_{i,t,(zz\lambda)} \right\|_\infty, \left\| \hat{A}_{i,t,(\lambda\lambda z)} \right\|_\infty$
- $\left\| \hat{B}_{i,t,(zz)} \right\|_\infty, \left\| \hat{B}_{i,t,(zz\lambda)} \right\|_\infty, \left\| \hat{B}_{i,t,(\lambda\lambda z)} \right\|_\infty$
- $\left\| \hat{L}_{i,t,(zz)} \right\|_\infty, \left\| \hat{L}_{i,t,(zz\lambda)} \right\|_\infty, \left\| \hat{L}_{i,t,(\lambda\lambda z)} \right\|_\infty$
- $\left\| \hat{M}_{i,t,(zz)} \right\|_\infty, \left\| \hat{M}_{i,t,(zz\lambda)} \right\|_\infty, \left\| \hat{M}_{i,t,(\lambda\lambda z)} \right\|_\infty$

Then the closed loop system (2.36) is stable.

Proof. The proof is straight forward since the conditions in Theorem 2.2.1 will be satisfied from the Bezout identity relating $\hat{G}_{i,t}$ to the above quantities. \square

2.4 Performance Analysis

In this section we seek a relationship between the performance of the frozen-time pair $(P_{i,t}, K_{k_i,n_t})$ and the actual time-varying feedback pair (P, K) . This is addressed in the following theorem.

Theorem 2.4.1. *Let S^{kl} ($k = 1, 2, l = 1, 2, 3, 4$) represent the map from u_k to y_l in the system of Figure 2.2 and $S_{i,t}^{kl}$ the LSTI map from u_k to y_l for the frozen system $(P_{i,t}, K_{k_i,n_t})$. Now, let the assumptions of Theorem 2.2.1 hold. Given $\epsilon > 0$, there exists a nonzero constant γ_p with $\gamma \leq \gamma_p$ such that*

$$(1 - \epsilon) \|S^{kl}\| \leq \sup_{i,t} \|S_{i,t}^{kl}\| + \epsilon$$

Proof. Let $u_1 = 0$ and $\|u_2\| \leq 1$. From the system equations we get

$$\begin{aligned} y_{1,i}(t) &= -(H_{k_i,n_t}(G - G_{k_i,n_t})y_1)_i(t) - (H_{k_i,n_t}Xy_1)_i(t) - (H_{k_i,n_t}Yy_2)_i(t) \\ &\quad - (H_{k_i,n_t}(BMu_2))_i(t) \end{aligned} \quad (2.71)$$

Consider now the frozen LSTI feedback system given a pair (i, t) and subjected to the same input u_2 . Let \hat{y}_1 denote the output that corresponds to y_1 in the time varying loop. Evaluating \hat{y}_1 at (i, t) we have

$$\hat{y}_{1,i}(t) = -(H_{k_i,n_t} B_{i,t} M_{k_i,n_t} u_2)_i(t) \quad (2.72)$$

Subtracting (2.71) from (2.72) we obtain

$$\begin{aligned} \hat{y}_{1,i}(t) - y_{1,i}(t) &= (H_{k_i,n_t} (G - G_{k_i,n_t}) y_1)_i(t) + (H_{k_i,n_t} X y_1)_i(t) \\ &+ (H_{k_i,n_t} Y y_2)_i(t) + (H_{k_i,n_t} (BM - B_{i,t} M_{k_i,n_t}) u_2)_i(t) \end{aligned} \quad (2.73)$$

The idea is to bound $|(H_{k_i,n_t} (BM - B_{i,t} M_{k_i,n_t}) u_2)_i(t)|$ by some constant. For this purpose, define the operator $Q \in \mathcal{L}_{STV}$ as

$$(Qz)_i(\tau) = (B_{i,\tau} M_{k_i,n_\tau} z)_i(\tau), \quad i \in \mathbb{Z}, \tau \in \mathbb{Z}^+ \quad (2.74)$$

then

$$\begin{aligned} (H_{k_i,n_t} (BM - B_{i,t} M_{k_i,n_t}) u_2)_i(t) &= (H_{k_i,n_t} (BM - Q) u_2)_i(t) \\ &+ (H_{k_i,n_t} (Q - B_{i,t} M_{k_i,n_t}) u_2)_i(t) \end{aligned} \quad (2.75)$$

By Lemma 2.3.1, and the fact that H_{k_i,n_t} has uniformly bounded norm, it follows that

$$|(H_{k_i,n_t} (BM - Q) u_2)_i(t)| \leq \gamma c_1 \quad (2.76)$$

where c_1 is a positive constant. We have the following for the term $(H_{k_i,n_t} (Q - B_{i,t} M_{k_i,n_t}) u_2)_i(t)$:

$$\begin{aligned} \|B_{i,\tau} M_{k_i,n_\tau} - B_{i,t} M_{k_i,n_t}\| &\leq \|B_{i,\tau}\| \|M_{k_i,n_\tau} - M_{k_i,n_t}\| \\ &+ \|M_{k_i,n_t}\| \|B_{i,t} - B_{i,\tau}\| \end{aligned} \quad (2.77)$$

$$\begin{aligned} &\leq \|B_{i,\tau}\| \gamma_M (|i - i| + |t - \tau|) \\ &+ \|M_{k_i,n_t}\| \gamma_B (|i - i| + |t - \tau|) \end{aligned} \quad (2.78)$$

$$\leq \gamma c_2 (|i - i| + |t - \tau|) \quad (2.79)$$

Hence if $z_i(\tau) = ((Q - B_{i,t}M_{k_i,n_t})u_2)_i(\tau)$, then

$$|z_i(\tau)| \leq \gamma c_2(|i - i| + |t - \tau|), \quad i \in \mathbb{Z}, \quad \tau \in \mathbb{Z}^+, \quad \text{with } c_2 > 0$$

However, from the fact that H_{k_i,n_t} has bounded (uniformly in t , and i) IT-SAE, it follows that

$$|(H_{k_i,n_t}(Q - B_{i,t}M_{k_i,n_t})u_2)_i(t)| = \left| \sum_{\tau=0}^t \sum_{i=-\infty}^{\infty} (h_{k_i,i-\mathbf{i}}(n_t, t - \tau)) z_i(\tau) \right| \quad (2.80)$$

$$\leq \gamma c_2 \sum_{\tau=0}^t \sum_{i=-\infty}^{\infty} (h_{k_i,i}(n_t, \tau)) (|i| + |\tau|) \quad (2.81)$$

$$\leq \gamma c_3, \quad c_3 > 0 \quad (2.82)$$

Looking at the rest of the terms, and since $\|u_2\| \leq 1$, we have

$$|(H_{k_i,n_t}Xy_1)_i(t)| \leq \gamma c_4 \|S^{12}\| \quad (2.83)$$

$$|(H_{k_i,n_t}Yy_2)_i(t)| \leq \gamma c_5 \|S^{22}\| \quad \text{and} \quad (2.84)$$

$$|(H_{k_i,n_t}(G - G_{k_i,n_t})y_1)_i(t)| \leq \gamma c_6 \|S^{12}\| \quad (2.85)$$

Putting every thing together, it follows that there are constants $c, c_{12}, c_{22} > 0$ such that

$$|\hat{y}_{1,i}(t) - y_{1,i}(t)| \leq \gamma c + \gamma c_{12} \|S^{12}\| + \gamma c_{22} \|S^{22}\| \quad (2.86)$$

Since $\|u_2\| \leq 1$, we have $|\hat{y}_{1,i}(t)| \leq \|S_{i,t}^{12}\|$, and therefore

$$\sup_{i,t} |y_{1,i}(t)| \leq \sup_{i,t} \|S_{i,t}^{12}\| + \gamma c + \gamma c_{12} \|S^{12}\| + \gamma c_{22} \|S^{22}\| \quad (2.87)$$

Since u_2 is arbitrary

$$\|S^{12}\| \leq \sup_{i,t} \|S_{i,t}^{12}\| + \gamma c + \gamma c_{12} \|S^{12}\| + \gamma c_{22} \|S^{22}\| \quad (2.88)$$

Similarly working for $\|S^{22}\|$ we get

$$\|S^{22}\| \leq \sup_{i,t} \|S_{i,t}^{22}\| + \gamma k + \gamma k_{22} \|S^{22}\| + \gamma k_{12} \|S^{12}\| \quad (2.89)$$

Noting that $\|H_{k_i, n_t}\|$ is uniformly bounded, we have

$$\sup_{i,t} \|S_{i,t}^{12}\|, \sup_{i,t} \|S_{i,t}^{22}\| < \infty$$

and hence by assuming γ_p sufficiently small, the proof of the theorem is complete. \square

2.5 Remarks

Before closing, some final remarks are in order. First, we note that although the previous stability and performance analysis relies on a sufficiently small rate of change in time and space of local approximants, this does not mean that all the local approximants are close to each other. If the spatiotemporal distance is large, the local systems may be very different. Secondly, the specific overall controller structure can be viewed as a collection of local controllers distributed in space and time. Each local station has to know the measurements and controls of the others, in fact only a limited number that depends on the specific (polynomial) order of L_{k_i, n_t} , M_{k_i, n_t} that we choose for the local, frozen design. Each local station though, does not have to know the dynamics of the others as it assumes they are the same as its own. Therefore, only their own, local models need to be available to the local stations.

2.6 Conclusion

In this chapter we have considered the stability analysis of systems that have gradually varying dynamics in time as well as in space. In particular, we have looked at the case where the controllers were not necessarily adjusted for every instance in space and time, and hence were used for some fixed spatiotemporal window before new controllers were implemented. We showed how the length of these windows entered in the stability analysis. It was shown that the actual time varying system can be stabilized using the frozen space-time controllers provided the variations in the spatiotemporal

dynamics are sufficiently small. We have also shown that the worst l_∞ to l_∞ performance of global spatiotemporally varying system can be brought arbitrarily close to the worst frozen spatially and temporally l_∞ to l_∞ performance given that the rates of variation of the plant and the controller are sufficiently small.

In the next chapter we shall present a distributed projection algorithm for system identification of LSTI systems and it shall be shown that the proposed algorithm eventually results in an estimated system that is slowly varying in time as well as in space.

CHAPTER 3

SYSTEM IDENTIFICATION OF SPATIOTEMPORALLY INVARIANT SYSTEMS

In Chapter 2 we presented results regarding the stability and performance of gradually varying spatiotemporal systems. In this chapter we look at the problem of system identification of spatiotemporally invariant systems from a control theoretic perspective. In particular, we present a distributed projection algorithm in which each subsystem coordinates with other subsystems in the neighborhood to achieve the following:

1. estimation error (difference between actual and predicted system output) goes to zero, regardless of the convergence of estimates to the true value
2. estimates get arbitrarily close to each other as time increases (at least locally)
3. better performance (in terms of getting closer to the true parameter) than the standard projection algorithm if run independently on every subsystem

Our proposed algorithm builds on the standard projection algorithm. The main idea, however, can be applied to other gradient-based schemes such as least squares. We shall see in Chapter 4 that the above mentioned properties of our proposed algorithm will allow us to apply the results of Chapter 2 to prove stability of an indirect adaptive control scheme.

3.1 Distributed Projection Algorithm

We assume that the deterministic dynamical system can be described by a model that may be expressed succinctly in the following simple form:

$$y_i(t) = \phi_i(t-1)^T \theta_0 \tag{3.1}$$

where $y_i(t)$ denotes the (scalar) system output of subsystem i at time ‘ t ’, $\phi_i(t-1)$ denotes a vector that is a linear function of $Y(t)$ and $U(t)$ where:

$$\begin{aligned} Y(t) &:= [\{y_i(t-1), y_i(t-2), \dots\}, \{y_{i-1}(t-1), y_{i-1}(t-2), \dots\}, \\ &\quad \{y_{i+1}(t-1), y_{i+1}(t-2), \dots\}, \dots] \\ U(t) &:= [\{u_i(t-1), u_i(t-2), \dots\}, \{u_{i-1}(t-1), u_{i-1}(t-2), \dots\}, \\ &\quad \{u_{i+1}(t-1), u_{i+1}(t-2), \dots\}, \dots] \end{aligned}$$

θ_0 denotes a parameter vector that is unknown. We introduce the following notations that shall be used frequently in the sequel: $\hat{\theta}_i(t-1)$ is the estimate of the true parameter vector θ_0 at time ‘ t ’ available at the i_{th} subsystem. $\hat{\theta}(0)$ is the initial estimate/guess which is given at each subsystem.

3.1.1 Main Idea

The main idea of the algorithm is depicted in Figure 3.1. We assume that all the subsystems start with the same initial guess $\hat{\theta}_i(0) =: \hat{\theta}(0)$. Each subsystem runs the same algorithm, which is described in this section. At time ‘ $t+1$ ’, subsystem i has access to its own estimate $\hat{\theta}_i(t)$, and the estimate of its immediate neighbors $\hat{\theta}_{i-1}(t)$, and $\hat{\theta}_{i+1}(t)$. Before updating its current estimate based on the information available at time ‘ $t+1$ ’, subsystem i resets its current estimate to the best available estimate from the set $\{\hat{\theta}_i(t), \hat{\theta}_{i-1}(t), \hat{\theta}_{i+1}(t)\}$. This estimate, call it $\hat{\theta}^i(t)$, is then taken as the current estimate by the subsystem i in order to establish the next estimate. If there is a tie, i.e., there is more than one estimate that qualifies as the best, an average is taken. In the sequel, the best estimate available to the i_{th} subsystem for next iterate shall be referred to as $\hat{\theta}^i(\cdot)$. The update equation of the algorithm is given in Equation (3.2). The cumulative improvement index helps in identifying the best estimate available.

$$\hat{\theta}_i(t+1) = \hat{\theta}^i(t) + \frac{a(t)\phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \left[y_i(t+1) - \phi_i(t)^T \hat{\theta}^i(t) \right] \quad (3.2)$$

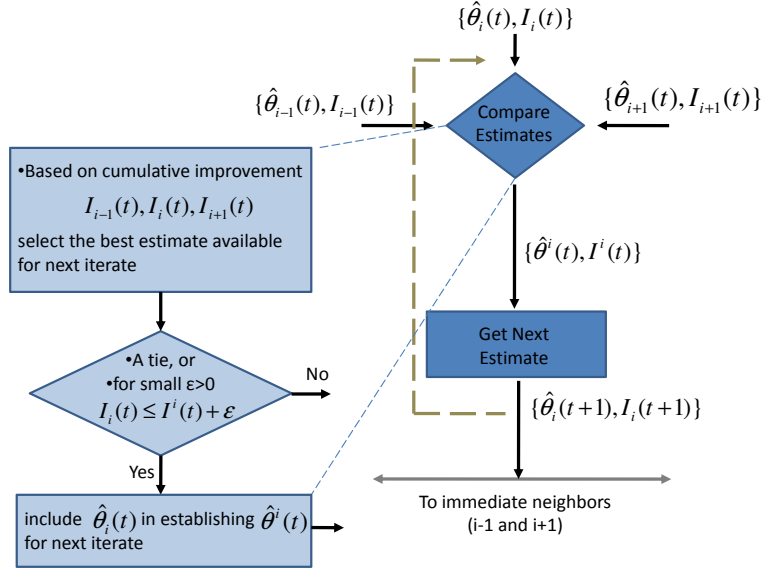


Figure 3.1 Flowchart Depicting the Distributed Projection Algorithm

where

$$\theta^i(t) \in \left\{ \begin{array}{l} \hat{\theta}_i(t), \hat{\theta}_{i-1}(t), \hat{\theta}_{i+1}(t), \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2}, \frac{\hat{\theta}_i(t) + \hat{\theta}_{i+1}(t)}{2}, \\ \frac{\hat{\theta}_{i+1}(t) + \hat{\theta}_{i-1}(t)}{2}, \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t) + \hat{\theta}_{i+1}(t)}{3} \end{array} \right\}$$

where $c > 0$, and $0 < a(t) < 2$.

3.1.2 Cumulative Improvement Index $I_i(\cdot)$

Each subsystem shares its current estimate along with the associated cumulative improvement index with its immediate neighbors. The cumulative improvement index $I_i(\cdot)$ is a measure of how much an estimate has improved from the original guess $\hat{\theta}(0)$. In the case of a system with no noise, one can exactly establish $I_i(\cdot)$ for each estimate at every time step ' t '. In the case of a system with bounded measurement noise, however, one can establish only a lower bound on the magnitude of $I_i(\cdot)$, and then proceed as discussed here. In order to explain the cumulative improvement index $I_i(\cdot)$ mathematically,

let us first introduce some notations:

$$\tilde{\theta}^i(t) := \hat{\theta}^i(t) - \theta_0 \quad (3.3)$$

$$\tilde{\theta}_i(t) := \hat{\theta}_i(t) - \theta_0 \quad (3.4)$$

$$\begin{aligned} e_i(t) &:= y_i(t) - \phi_i(t-1)^T \hat{\theta}^i(t-1) \\ &= -\phi_i(t-1)^T \tilde{\theta}^i(t-1) \end{aligned} \quad (3.5)$$

We shall refer to $e_i(t)$ as the estimation error. Consider Equation (3.2). Subtracting θ_0 from both sides of the (3.2), and using (3.1) along with (3.3), we obtain

$$\tilde{\theta}_i(t+1) = \tilde{\theta}^i(t) - \frac{a(t)\phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \phi_i(t)^T \tilde{\theta}^i(t) \quad (3.6)$$

Using (3.5), we have

$$\left\| \tilde{\theta}_i(t+1) \right\|^2 - \left\| \tilde{\theta}^i(t) \right\|^2 = a(t) \left[-2 + \frac{\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \quad (3.7)$$

Note that

$$a(t) \left[-2 + \frac{\phi_i(t)^T \phi_i(t)}{1 + \phi_i(t)^T \phi_i(t)} \right] < 0 \quad (3.8)$$

Expanding $\left\| \tilde{\theta}^i(t) \right\|$, and its successive terms, we can write

$$\begin{aligned} \left\| \tilde{\theta}_i(t+1) \right\|^2 &= \left\| \tilde{\theta}(0) \right\|^2 + \\ &\sum_{j=1}^t a(j) \left[-2 + \frac{\phi_k(j-1)^T \phi_k(j-1)}{c + \phi_k(j-1)^T \phi_k(j-1)} \right] \frac{e_k(j)^2}{c + \phi_k(j-1)^T \phi_k(j-1)} \\ &\quad + a(t) \left[-2 + \frac{\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \end{aligned} \quad (3.9)$$

where the subscript k captures the evolution path of $\hat{\theta}^i(t)$, i.e. the sequence of the subsystems that were involved in establishing $\hat{\theta}^i(t)$. We now define the cumulative improvement index in the following way:

$$I^i(t) := \sum_{j=1}^t a(t) \left[-2 + \frac{\phi_k(j-1)^T \phi_k(j-1)}{c + \phi_k(j-1)^T \phi_k(j-1)} \right] \frac{e_k(j)^2}{c + \phi_k(j-1)^T \phi_k(j-1)} \quad (3.10)$$

From (3.10) we have

$$I_i(t+1) = I^i(t) + a(t) \left[-2 + \frac{\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \quad (3.11)$$

It is clear from Equation (3.10), that $I_i(t+1) \leq 0$. Its magnitude is exactly the square of the distance $\hat{\theta}_i(t+1)$ has traveled from $\hat{\theta}(0)$ towards θ_0 . Also note that $I_i(t+1) \leq I_i(t)$. As is evident from Equation (3.9), the cumulative improvement index is just a scalar that can be updated iteratively at each time step. The bigger the magnitude of the cumulative improvement index, the better the estimate is.

3.1.3 The ϵ -Rule:

The key idea of the algorithm we propose is that we can write any estimate as the sum of the initial guess and some improvement terms. The cumulative improvement index is a sum of nonpositive real numbers. At the same time, $\|\tilde{\theta}_i(t)\|$ is a bounded function (bounded below by 0, and above by $\|\tilde{\theta}(0)\|$). As $t \rightarrow \infty$, the cumulative improvement index $I_i(t)$ will contain infinitely many nonpositive terms (this is obvious by observing the expression of Equation 3.10). Since, however, $\|\tilde{\theta}_i(t)\|$ is bounded, $\lim_{t \rightarrow \infty} I_i(t)$ should exist. This implies that the improvement in estimates will slow down as time increases. We note here that it is possible, for example, to have $\|\tilde{\theta}_i(t)\| \approx \|\tilde{\theta}_{i-1}(t)\|$, but $\hat{\theta}_i(t)$ and $\hat{\theta}_{i-1}(t)$ still be far apart. This is depicted in Figure 3.2. For such a situation, when the neighboring estimates (in space as well as time) are at almost ‘equal distance’ from the true value but still far from each other, we introduce the ϵ rule, which is as follows:

- There is a small number $\epsilon > 0$ specified a priori and known to each subsystem.
- At a given time t , if $\|\tilde{\theta}_i(t)\| \leq \|\tilde{\theta}^i(t)\| + \epsilon$, redefine $\hat{\theta}^i(t)$ to include $\hat{\theta}_i(t)$.

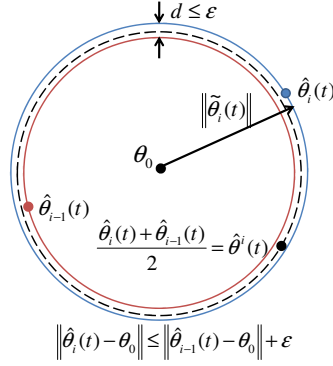


Figure 3.2 Figure Showing ϵ -Rule

For example, if $\hat{\theta}^i(t) = \hat{\theta}_{i-1}(t)$, redefine as $\hat{\theta}^i(t) = \frac{\hat{\theta}_{i-1}(t) + \hat{\theta}_i(t)}{2}$ etc.

The purpose of the ϵ -rule is to bring the estimates close to each other locally when their respective rate of improvement slows down. At the same time, this rule ensures that the next estimate will be as good as the current, if not better, in terms of its distance from the true value. The algorithm, therefore, guarantees improvement throughout its execution making sure that the identification error at each subsystem will eventually become small. The proposed estimation scheme also serves to bring the estimates ‘close’ to one another asymptotically as shown below.

3.1.4 Properties of Distributed Projection Algorithm

The properties of distributed projection algorithm are summarized below.

Lemma 3.1.1. *For the algorithm (3.2) and subject to (3.1) it follows that*

1. $\|\hat{\theta}_i(t) - \theta_0\| \leq \|\hat{\theta}^i(t-1) - \theta_0\| \leq \|\hat{\theta}(0) - \theta_0\| \quad \text{for } t \geq 1$
2. $\lim_{t \rightarrow \infty} \frac{e_i(t)}{[c + \phi_i(t-1)^T \phi_i(t-1)]^{1/2}} = 0$
3. $\lim_{t \rightarrow \infty} \|\hat{\theta}_i(t) - \hat{\theta}_k(t)\| = 0 \text{ for } k \in \{i-1, i+1\}$
4. $\lim_{t \rightarrow \infty} \|\hat{\theta}_i(t) - \hat{\theta}_i(t-1)\| = 0 \quad \forall i. \text{ This together with 3) implies}$

$$\lim_{t \rightarrow \infty} \|\hat{\theta}_i(t) - \hat{\theta}_k(t+l)\| = 0 \text{ for } k \in \{i, i-1, i+1\} \text{ and for finite } l$$

Proof. The proof of 1) is evident from the discussion provided above.

For 2), we observe that $\left\|\tilde{\theta}_i(t)\right\|^2$ is a bounded nonincreasing function, and by summing (3.7) we get

$$\begin{aligned} \left\|\tilde{\theta}_i(t)\right\|^2 = & \left\|\tilde{\theta}(0)\right\|^2 + \sum_{j=1}^{t-1} a(t) \left[-2 + \frac{\phi_k(j-1)^T \phi_k(j-1)}{c + \phi_k(j-1)^T \phi_k(j-1)} \right] \frac{e_k(j)^2}{c + \phi_k(j-1)^T \phi_k(j-1)} \\ & + a(t) \left[-2 + \frac{\phi_i(t-1)^T \phi_i(t-1)}{c + \phi_i(t-1)^T \phi_i(t-1)} \right] \frac{e_i(t)^2}{c + \phi_i(t-1)^T \phi_i(t-1)} \end{aligned} \quad (3.12)$$

Since $\left\|\tilde{\theta}_i(t)\right\|^2$ is nonnegative and (3.8) holds, we conclude that 2) holds.

In order to prove 3), first note the following:

1. Given $\epsilon_0 > 0 \exists T_{\epsilon_0} \mid \forall t \geq T_{\epsilon_0}$, and $\forall i$ we have

$$\left| \frac{e_i(t)}{[c + \phi_i(t-1)^T \phi_i(t-1)]^{1/2}} \right| \leq \epsilon_0 \quad (3.13)$$

i.e. the (normalized) estimation error will be small (and shall remain small) for all subsystems after time T_{ϵ_0} .

2. Given $\epsilon_1 > 0 \exists T_{\epsilon_1} \mid \forall t \geq T_{\epsilon_1} \forall i$, we have

$$\left\|\tilde{\theta}_i(t-1)\right\| - \left\|\tilde{\theta}_i(t)\right\| \leq \epsilon_1 \quad (3.14)$$

i.e. the improvement in the parameter estimates will be small (and shall remain small) for all subsystems after time T_{ϵ_1} . Also note that if $\epsilon_1 \leq \epsilon$, where ϵ is the constant used for ϵ -rule, then ϵ -rule will stay in place $\forall t \geq T_{\epsilon_1}$.

3. If $\left\|\tilde{\theta}_i(t)\right\| = \left\|\tilde{\theta}_{i-1}(t)\right\|$, and $\hat{\theta}_i(t) \neq \hat{\theta}_{i-1}(t)$, then:

$$\begin{aligned} \left\| \theta_0 - \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2} \right\| & < \frac{1}{2} \|\theta_0 - \theta_i(t)\| + \frac{1}{2} \|\theta_0 - \theta_{i-1}(t)\| \\ & \text{(triangle inequality is strict since the estimates are not aligned)} \\ & = \|\theta_0 - \theta_i(t)\| \\ & = \|\theta_0 - \theta_{i-1}(t)\| \end{aligned} \quad (3.15)$$

To see the main idea, let $\hat{\theta}^i(t) = \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2}$, and $\hat{\theta}_i(t) \neq \hat{\theta}_{i-1}(t)$ then

$$\begin{aligned} \|\tilde{\theta}_i(t+1)\|^2 &= \left\| \frac{\tilde{\theta}_i(t) + \tilde{\theta}_{i-1}(t)}{2} \right\|^2 \\ &+ a(t+1) \left[-2 + \frac{\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \underbrace{\frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)}}_{\leq \epsilon_0^2} \end{aligned} \quad (3.16)$$

$$\|\tilde{\theta}_i(t+1)\|^2 \leq \left\| \frac{\tilde{\theta}_i(t) + \tilde{\theta}_{i-1}(t)}{2} \right\|^2 < \|\tilde{\theta}_i(t)\|^2 \quad (3.17)$$

Since $\|\tilde{\theta}_i(t)\| - \|\tilde{\theta}_i(t+1)\| \leq \epsilon_1$, the improvement resulting from averaging should be less than ϵ_1 . We show below that if the estimates are far apart then the improvement caused by averaging will exceed ϵ_1 .

For $t \geq \max\{T_{\epsilon_1}, T_{\epsilon_0}\}$ and $\epsilon_1 \leq \epsilon$, we demonstrate below that:

- either the estimate(s) from the immediate neighbor(s) ($\hat{\theta}_{i+1}(t)$, or $\hat{\theta}_{i-1}(t)$, or both) shall be used to establish the estimate at the i_{th} subsystem ($\hat{\theta}_i(t+1)$), or
- the estimate of the i_{th} subsystem ($\hat{\theta}_i(t)$) is used in establishing the estimate(s) of the immediate neighbor(s) ($\hat{\theta}_{i+1}(t+1)$, or $\hat{\theta}_{i-1}(t+1)$, or both), or
- both of the above (in case the estimates are equally good)

In other words, the estimation algorithm has strong interdependencies in a given local neighborhood. At any given time instance t , we have one of the following possibilities for the subsystem i :

1. $\hat{\theta}^i(t) = \hat{\theta}_i(t)$ in which case we will have
 $\hat{\theta}^{i+1}(t) \in \left\{ \frac{\hat{\theta}_{i+1}(t) + \hat{\theta}_i(t)}{2}, \frac{\hat{\theta}_{i+1}(t) + \hat{\theta}_i(t) + \hat{\theta}_{i+2}(t)}{3} \right\}$, and
 $\hat{\theta}^{i-1}(t) \in \left\{ \frac{\hat{\theta}_{i-1}(t) + \hat{\theta}_i(t)}{2}, \frac{\hat{\theta}_{i-1}(t) + \hat{\theta}_i(t) + \hat{\theta}_{i-2}(t)}{3} \right\}$
2. $\hat{\theta}^i(t) = \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2}$ in which case we will have
 $\hat{\theta}^{i+1}(t) \in \left\{ \frac{\hat{\theta}_{i+1}(t) + \hat{\theta}_i(t)}{2}, \frac{\hat{\theta}_{i+1}(t) + \hat{\theta}_i(t) + \hat{\theta}_{i+2}(t)}{3} \right\}$
3. $\hat{\theta}^i(t) = \frac{\hat{\theta}_i(t) + \hat{\theta}_{i+1}(t)}{2}$ in which case we will have
 $\hat{\theta}^{i-1}(t) \in \left\{ \frac{\hat{\theta}_{i-1}(t) + \hat{\theta}_{i-2}(t)}{2}, \frac{\hat{\theta}_{i-1}(t) + \hat{\theta}_i(t) + \hat{\theta}_{i-2}(t)}{3} \right\}$,
4. $\hat{\theta}^i(t) = \frac{\hat{\theta}_{i+1}(t) + \hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{3}$

We will provide an upper bound on $\|\hat{\theta}_i(t) - \hat{\theta}_k(t)\|$, where $k \in \{i-1, i+1\}$. Note that the farther apart the estimates are, the closer the resulting averaged estimate shall be to the true parameter. We shall assume that we have situation 4) as this covers the rest (1-3). This becomes obvious if we further assume that $\hat{\theta}_i(t) = \hat{\theta}_{i+1}(t)$ and let $e_i(t) = 0$. This ensures that the improvement can be attributed to the process of averaging alone. We also assume that $\|\tilde{\theta}_i(t)\| = \|\tilde{\theta}_{i-1}(t)\|$. Our goal is to see how far $\hat{\theta}_{i-1}(t)$ can be from $\hat{\theta}_i(t)$ so that the resulting improvement does not exceed ϵ_1 . The calculation is presented in Figure 3.3. As shown in Figure 3.3(a), we are interested in establishing the distance c . It is, however, easier to calculate the distance \bar{c} where $\bar{c} > c$ (see Figure 3.3(b)). Solving the right triangle $AB\theta_0$ we see that $\bar{c} = 2\sqrt{3a\epsilon_1 - 9/4\epsilon_1^2}$. Clearly, this distance is controlled by ϵ_1 . We can, therefore, conclude that $\|\hat{\theta}_i(t) - \hat{\theta}_k(t)\| \leq 2\sqrt{3a\epsilon_1 - 9/4\epsilon_1^2}$, where $k \in \{i-1, i+1\}$.

To prove 4) we can bound the successive iterates at subsystem i as follows:

$$\begin{aligned} & \|\hat{\theta}_i(t+1) - \hat{\theta}_i(t)\| \\ & \leq \left\| \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t) + \hat{\theta}_{i+1}(t)}{3} + \frac{a(t+1)\phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} e_i(t+1) - \hat{\theta}_i(t) \right\| \end{aligned} \quad (3.18)$$

$$\leq \frac{1}{3} \|\hat{\theta}_{i-1}(t) - \hat{\theta}_i(t)\| + \frac{1}{3} \|\hat{\theta}_{i+1}(t) - \hat{\theta}_i(t)\| + \epsilon_0 \quad (3.19)$$

$$\leq \frac{4}{3} \sqrt{3a\epsilon_1 - 9/4\epsilon_1^2} + \epsilon_0 \quad (3.20)$$

Similarly we can calculate an upper bound on $\|\hat{\theta}_i(t+1) - \hat{\theta}_k(t)\|$ for $k \in \{i-1, i+1\}$ controlled by ϵ_1 , and ϵ_0 . Clearly as ϵ_0 and ϵ_1 approach zero with $t \rightarrow \infty$, the difference between the estimates also approach zero. This completes the proof. \square

3.1.5 Information Exchange

In this section we provide details about the information exchange necessary to take place between the subsystems to execute the proposed algorithm.

1. $\{\hat{\theta}_i(t), I_i(t)\}$:

After each iterate or update, each subsystem must provide its estimate and the associated cumulative improvement index to its immediate

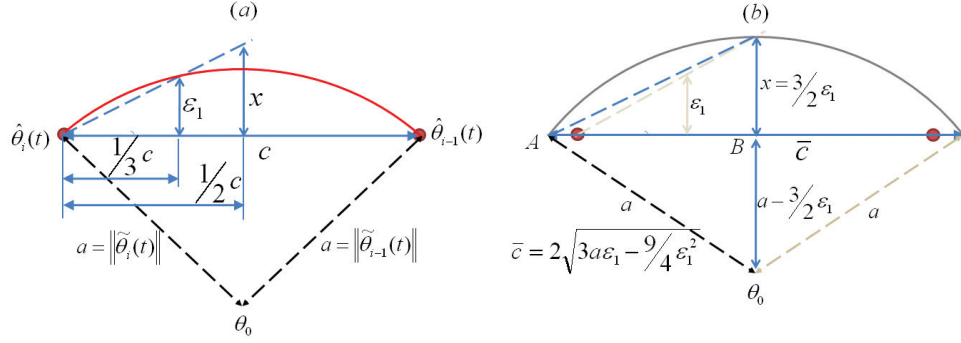


Figure 3.3 Calculation of Upper Bound on the Distance of Local Estimates

neighbors. The cumulative improvement index $I_i(t)$ can be computed iteratively, e.g. at each step, all one needs to calculate is the term

$$a(t) \left[-2 + \frac{\phi_i(t-1)^T \phi_i(t-1)}{c + \phi_i(t-1)^T \phi_i(t-1)} \right] \frac{e_i(t)^2}{c + \phi_i(t-1)^T \phi_i(t-1)}$$

and add it to $I_j(t-1)$ to obtain, $I_i(t)$

2. $[y_{i-1}(t-1), y_{i+1}(t-1), \dots, y_{i-1}(t-2), y_{i+1}(t-2), \dots], [u_{i-1}(t-1), u_{i+1}(t-1), \dots, u_{i-1}(t-2), u_{i+1}(t-2), \dots]$:

Depending on the structure of the model of the system, each system must receive information about the outputs and inputs from other sub-systems that affect it in order to construct the regression vector $\phi_i(t-1)$.

3. More information exchange is required whenever the ϵ -rule comes into play or averaging takes place in order to keep track of the cumulative improvement index. Define:

$$E_i(t) := \hat{\theta}_i(t) - \hat{\theta}(0) \quad (3.21)$$

Note that we can equivalently write,

$$E_i(t) = \sum_{\tau=1}^t \left[\frac{\phi_k(\tau-1)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) \quad (3.22)$$

To demonstrate the requirement of this additional information ex-

change, assume that at time t , we need to take $\hat{\theta}^i(t) = \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2}$ for the subsystem i . The calculation of $I^i(t)$ is given as follows:

$$\begin{aligned}\hat{\theta}^i(t) &= \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2} = \frac{\hat{\theta}(0) + E_i(t) + \hat{\theta}(0) + E_{i-1}(t)}{2} \\ &= \hat{\theta}(0) + \frac{E_i(t) + E_{i-1}(t)}{2}\end{aligned}\quad (3.23)$$

subtracting θ_0 from both sides and taking the inner product of $\hat{\theta}^i(t)$ with itself, we get

$$\begin{aligned}\|\tilde{\theta}^i(t)\|^2 &= \|\tilde{\theta}(0)\|^2 + \tilde{\theta}(0)^T [E_i(t) + E_{i-1}(t)] \\ &\quad + \frac{1}{4} [E_i(t) + E_{i-1}(t)]^T [E_i(t) + E_{i-1}(t)]\end{aligned}\quad (3.24)$$

From the above expression, we can write $I^i(t)$ as follows

$$I^i(t) = \tilde{\theta}(0)^T [E_i(t) + E_{i-1}(t)] + \frac{1}{4} [E_i(t) + E_{i-1}(t)]^T [E_i(t) + E_{i-1}(t)] \quad (3.25)$$

The expression $\tilde{\theta}(0)^T E_i(t)$ can be iteratively computed as follows:

$$\begin{aligned}\tilde{\theta}(0)^T E_i(t) &= \tilde{\theta}(0)^T \left(\sum_{\tau=1}^t \left[\frac{\phi_k(\tau-1)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) \right) \\ &= \sum_{\tau=1}^t \left[\frac{\hat{\theta}(0)^T \phi_k(\tau-1) - y_k(\tau)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) \\ &= \tilde{\theta}(0)^T E_j(t-1) + \left[\frac{\hat{\theta}(0)^T \phi_i(t-1) - y_i(t)}{c + \phi_i(t-1)^T \phi_i(t-1)} \right] e_i(t)\end{aligned}$$

We summarize the required exchange of information below:

(a) $E_i(t)$

The vector $E_i(t)$ should be shared amongst the immediate neighbors in order to evaluate the expression $E_i(t)^T E_{i-1}(t)$. The record of this vector can be kept iteratively.

(b) $\tilde{\theta}(0)^T E_i(t)$

Each subsystem should also share the scalar $\tilde{\theta}(0)^T E_i(t)$ with its

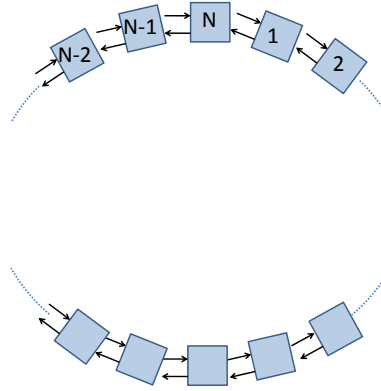


Figure 3.4 Circulant System

immediate neighbors. The record of this scalar can be kept iteratively.

3.1.6 Simulation

In this section we present an implementation of our algorithm on a circulant system. The basic structure of a circulant system is presented in Figure 3.4. Each subsystem has the same dynamics. For $i = N$, $i + 1 = 1$. Likewise, for $i = 1$, $i - 1 = N$. The dynamics of the i_{th} subsystem are given as follows:

$$y_i(t + 1) = -0.6y_i(t) + 0.1y_{i-1}(t) + 0.1y_{i+1}(t) + u_i(t) \quad (3.26)$$

The initial guess for all the subsystems was taken as $\hat{\theta}(0) = [1, 1, 1, 1]^T$. The input for the subsystems was chosen as given below:

$$u_i(t) = \cos(t + i) + \sin(t - i/2) \quad (3.27)$$

A system comprised of $N = 100$ subsystems was simulated. The standard projection algorithm (where there is no coordination amongst the subsystems) was simulated along with the proposed distributed projection algorithm. For the implementation of the distributed projection algorithm, a constant value of $\epsilon = 10^{-5}$ was chosen. The results are captured in Figure (3.5). The top plots in Figure (3.5) present the distance of the estimates of subsystem 1 from its neighbors, four to its right and four to its left. The

bottom plots, on the other hand, present the distance of the subsystems $(N - 3, \dots, N, 1, \dots, 5)$ from the true parameter θ_0 . A clear difference can be seen amongst the two plots. For the standard projection algorithm, the parameter estimates cease to change after about $20(\times 10)$ time steps. Also, the parameter estimates of the neighbors for subsystem 1 are quite far apart from it. The performance of the distributed projection algorithm, on the other hand, is quite outstanding. Not only do the parameters for the subsystems converge to the true value, their distance from subsystem 1 can also be seen to go to zero. It is quite interesting to note that the input signal was not ‘exciting’ enough for the standard projection algorithm, whereas convergence to the true parameter was seen for the distributed projection algorithm. Establishing excitation conditions necessary for true system identification is part of future research work.

Choice of ϵ

Simulations were also carried out to demonstrate the choice of ϵ for the ϵ -rule. A system of 10 subsystems, having the same dynamics as given in (3.26), was simulated with the initial guess for all the subsystems taken as $\hat{\theta}(0) = [1, 1, 1, 1]^T$. The input for the subsystems was chosen as follows:

$$u_i(t) = \cos(t + i) \quad (3.28)$$

Figure (3.6) presents a comparison of choosing $\epsilon = 10^{-5}$ against choosing $\epsilon = 10^{-1}$. It is interesting to note that a smaller value of ϵ results in bringing the estimates closer at a faster rate than a bigger value that enforces averaging. This is due to the fact that a smaller value of ϵ emphasizes improving the estimates rather than bringing the estimates closer. In doing so, the estimates get closer to the true parameter and hence to each other as a result.

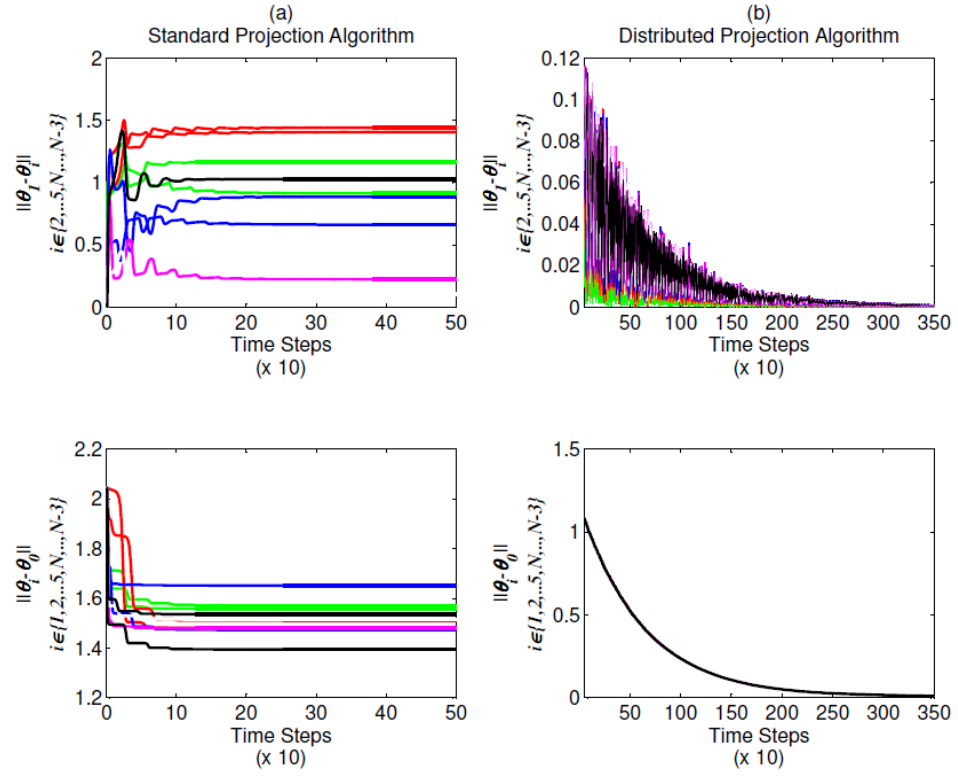


Figure 3.5 Comparison of Distributed Projection Algorithm (a) with Standard Projection Algorithm (b)

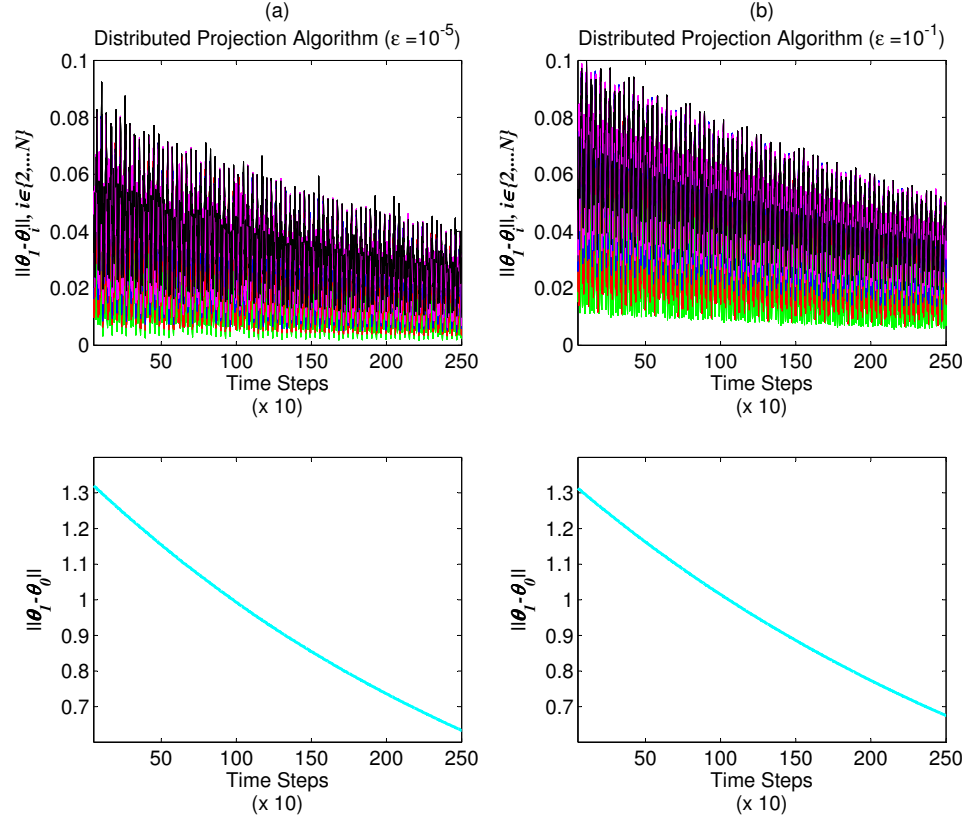


Figure 3.6 Figure Showing the Effect of Choice of ϵ for the ϵ -rule in Distributed Projection Algorithm

3.2 Parameter Estimation With Bounded Noise

We now consider the following system model

$$y_i(t) = \phi_i(t-1)^T \theta_0 + w_i(t) \quad (3.29)$$

where $w_i(t)$ denotes a bounded ‘noise’ term that can account for measurement noise, inaccurate modeling, computer round-off error etc. We, however, know that $w_i(t)$ is bounded, i.e. $\sup_{i,t} |w_i(t)| \leq \Delta$. We modify the distributed projection algorithm presented above as follows to account for the measurement noise.

$$\hat{\theta}_i(t+1) = \hat{\theta}^i(t) + \frac{a(t)\phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \left[y_i(t+1) - \phi_i(t)^T \hat{\theta}^i(t) \right] \quad (3.30)$$

where $\theta(0)$ is given, $c > 0$, and

$$a(t) = \begin{cases} 1 & \text{if } \left| y_i(t+1) - \phi_i(t)^T \hat{\theta}^i(t) \right| > 2\Delta \\ 0 & \text{otherwise} \end{cases} \quad (3.31)$$

The motivation above is to turn off the algorithm when the prediction error is small compared to the size of noise. We note here that it no longer remains possible to exactly establish the cumulative improvement index. We can, however, calculate a lower bound on its magnitude. This lower bound can be used to identify the better estimate as we explain in the sequel. Let us examine the update of the system i from time $t = 0$ to time $t = 1$.

$$\hat{\theta}_i(1) = \hat{\theta}(0) + \frac{a(0)\phi_i(0)}{c + \phi_i(0)^T \phi_i(0)} \left[y_i(1) - \phi_i(0)^T \hat{\theta}(0) \right] \quad (3.32)$$

Subtracting θ_0 from both sides and taking the inner product with itself, we get

$$\begin{aligned} \left\| \tilde{\theta}_i(1) \right\|^2 &= \left\| \tilde{\theta}(0) \right\|^2 + \frac{a(0)^2 \phi_i(0)^T \phi_i(0)}{[c + \phi_i(0)^T \phi_i(0)]^2} e_i(1)^2 \\ &\quad + \frac{2a(0)\tilde{\theta}(0)^T \phi_i(0)}{c + \phi_i(0)^T \phi_i(0)} e_i(1) \end{aligned} \quad (3.33)$$

Using the fact that

$$\hat{\theta}(0)^T \phi_i(0) = w_i(1) - e_i(1) \quad (3.34)$$

we can rewrite (3.33) as follows:

$$\begin{aligned} \|\tilde{\theta}_i(1)\|^2 &= \|\tilde{\theta}(0)\|^2 + a(0) \left[-2 + \frac{a(0)\phi_i(0)^T \phi_i(0)}{c + \phi_i(0)^T \phi_i(0)} \right] \frac{e_i(1)^2}{c + \phi_i(0)^T \phi_i(0)} \\ &\quad + \frac{a(0)}{c + \phi_i(0)^T \phi_i(0)} [2e_i(1)w_i(1)] \end{aligned} \quad (3.35)$$

$$\begin{aligned} \|\tilde{\theta}_i(1)\|^2 &\leq \|\tilde{\theta}(0)\|^2 \\ &\quad + a(0) \left[-2 + \frac{a(0)\phi_i(0)^T \phi_i(0)}{c + \phi_i(0)^T \phi_i(0)} \right] \frac{e_i(1)^2}{c + \phi_i(0)^T \phi_i(0)} + \frac{a(0)}{c + \phi_i(0)^T \phi_i(0)} [2|e_i(1)|\Delta] \end{aligned} \quad (3.36)$$

We define $\hat{I}_i(1)$ as follows:

$$\begin{aligned} \hat{I}_i(1) &:= a(0) \left[-2 + \frac{a(0)\phi_i(0)^T \phi_i(0)}{c + \phi_i(0)^T \phi_i(0)} \right] \frac{e_i(1)^2}{c + \phi_i(0)^T \phi_i(0)} \\ &\quad + \frac{a(0)}{c + \phi_i(0)^T \phi_i(0)} [2|e_i(1)|\Delta] \end{aligned} \quad (3.37)$$

$\hat{I}_i(1)$ is an overestimate (or its magnitude serves as a lower bound) for the actual cumulative improvement index $I_i(1)$. Note that the use of a dead-zone in the algorithm above makes sure that $\hat{I}_i(1)$ remains non-positive. To see this note that

$$a(0) \left[-2 + \frac{a(0)\phi_i(0)^T \phi_i(0)}{c + \phi_i(0)^T \phi_i(0)} \right] \leq -1 \text{ if } |y_i(t+1) - \phi_i(t)^T \hat{\theta}_j(t)| > 2\Delta \quad (3.38)$$

and

$$\frac{e_i(1)^2}{c + \phi_i(0)^T \phi_i(0)} \geq \frac{2|e_i(1)|\Delta}{c + \phi_i(0)^T \phi_i(0)} \quad (3.39)$$

One can write any update in the similar fashion as described above and the upper bound on the cumulative improvement index can be established iteratively. For example, from the update relationship of $\hat{\theta}_i(t+1)$, we can establish $\hat{I}_i(t+1)$ as follows:

$$\begin{aligned}
\left\| \tilde{\theta}_i(t+1) \right\|^2 &= \left\| \tilde{\theta}^i(t) \right\|^2 + a(t) \left[-2 + \frac{a(t)\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \\
&\quad + \frac{a(t)}{c + \phi_i(t)^T \phi_i(t)} [2e_i(t+1)w_i(t+1)]
\end{aligned} \tag{3.40}$$

$$\begin{aligned}
\left\| \tilde{\theta}_i(t+1) \right\|^2 &\leq \left\| \tilde{\theta}(0) \right\|^2 + \hat{I}^i(t) + a(t) \left[-2 + \frac{a(t)\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \\
&\quad + \frac{a(t)}{c + \phi_i(t)^T \phi_i(t)} [2|e_i(t+1)|\Delta]
\end{aligned} \tag{3.41}$$

and $\hat{I}_i(t+1)$ is given as:

$$\begin{aligned}
\hat{I}_i(t+1) &= \hat{I}^i(t) + a(t) \left[-2 + \frac{a(t)\phi_i(t)^T \phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} \right] \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \\
&\quad + \frac{a(t)}{c + \phi_i(t)^T \phi_i(t)} [2|e_i(t+1)|\Delta]
\end{aligned} \tag{3.42}$$

The upper bound on cumulative improvement index, whenever averaging needs to take place, can be calculated as explained in the following. Assume that at time ‘ t ’ we have $\hat{\theta}^i(t) = \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2}$ for the subsystem i . The calculation of $I^i(t)$ is given as:

$$\begin{aligned}
\hat{\theta}^i(t) &= \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t)}{2} = \frac{\hat{\theta}(0) + E_i(t) + \hat{\theta}(0) + E_{i-1}(t)}{2} \\
&= \hat{\theta}(0) + \frac{E_i(t) + E_{i-1}(t)}{2}
\end{aligned} \tag{3.43}$$

subtracting θ_0 from both sides and taking inner product with itself, we get

$$\begin{aligned}
\left\| \tilde{\theta}^i(t) \right\|^2 &= \left\| \tilde{\theta}(0) \right\|^2 + \tilde{\theta}(0)^T [E_i(t) + E_{i-1}(t)] \\
&\quad + \frac{1}{4} [E_i(t) + E_{i-1}(t)]^T [E_i(t) + E_{i-1}(t)]
\end{aligned} \tag{3.44}$$

From the above expression, we can write $I^i(t)$ as follows

$$I^i(t) = \tilde{\theta}(0)^T [E_i(t) + E_{i-1}(t)] + \frac{1}{4} [E_i(t) + E_{i-1}(t)]^T [E_i(t) + E_{i-1}(t)] \tag{3.45}$$

The upper bound on the expression $\tilde{\theta}(0)^T E_i(t)$ can be iteratively computed as follows:

$$\begin{aligned}
\tilde{\theta}(0)^T E_i(t) &= \tilde{\theta}(0)^T \left(\sum_{\tau=1}^t \left[\frac{\phi_k(\tau-1)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) \right) \\
&= \sum_{\tau=1}^t \left[\frac{\tilde{\theta}(0)^T \phi_k(\tau-1) - y_k(\tau) + w_k(\tau)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) \\
&\leq \sum_{\tau=1}^t \left[\frac{\tilde{\theta}(0)^T \phi_k(\tau-1) - y_k(\tau)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) + \sum_{\tau=1}^t \left[\frac{\Delta}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] |e_k(\tau)| \\
&= \underbrace{\sum_{\tau=1}^t \left(\left[\frac{\tilde{\theta}(0)^T \phi_k(\tau-1) - y_k(\tau)}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] e_k(\tau) + \left[\frac{\Delta}{c + \phi_k(\tau-1)^T \phi_k(\tau-1)} \right] |e_k(\tau)| \right)}_{=:\tilde{\theta}(0)^T \hat{E}_i(t)} \\
&= \tilde{\theta}(0)^T \hat{E}_i(t-1) + \left[\frac{\tilde{\theta}(0)^T \phi_i(t-1) - y_i(t)}{c + \phi_i(t-1)^T \phi_i(t-1)} \right] e_i(t) + \left[\frac{\Delta}{c + \phi_i(t-1)^T \phi_i(t-1)} \right] |e_i(t)| \quad (3.46)
\end{aligned}$$

We can, therefore, write the expression for $\hat{I}^i(t)$ as follows:

$$\hat{I}^i(t) = \tilde{\theta}(0)^T [\hat{E}_i(t) + \hat{E}_{i-1}(t)] + \frac{1}{4} [E_i(t) + E_{i-1}(t)]^T [E_i(t) + E_{i-1}(t)] \quad (3.47)$$

It is, however, not clear at this time that the upper bound $\hat{I}^i(t)$ calculated in this way is at least as good as $\hat{I}_i(t)$. This is required for correct execution of the algorithm and to make sure that the algorithm guarantees continuous improvement, at least in terms of establishing bounds on the successive terms. To see that $\hat{I}^i(t)$ is indeed as good as $\hat{I}_i(t)$, note that we can write $\hat{I}^i(t)$ as follows:

$$\begin{aligned}
\hat{I}^i(t) &= \underbrace{\tilde{\theta}(0)^T [\hat{E}_i(t) + \hat{E}_{i-1}(t)] + \frac{1}{2} E_i(t)^T E_i(t) + \frac{1}{2} E_{i-1}(t)^T E_{i-1}(t)}_{= \frac{\hat{I}_i(t) + \hat{I}_{i-1}(t)}{2}} \\
&\quad + \frac{1}{2} E_i(t)^T E_{i-1}(t) - \frac{1}{4} (E_i(t)^T E_i(t) + E_{i-1}(t)^T E_{i-1}(t)) \\
\hat{I}^i(t) &= \frac{\hat{I}_i(t) + \hat{I}_{i-1}(t)}{2} - \frac{1}{4} [E_i(t) - E_{i-1}(t)]^T [E_i(t) - E_{i-1}(t)] \\
\hat{I}^i(t) &\leq \frac{\hat{I}_i(t) + \hat{I}_{i-1}(t)}{2} - \frac{1}{4} [E_i(t) - E_{i-1}(t)]^T [E_i(t) - E_{i-1}(t)] \\
&\quad \text{Since } \hat{I}_{i-1}(t) \leq \hat{I}_i(t) \\
\hat{I}^i(t) &\leq \hat{I}_i(t) \quad (3.48)
\end{aligned}$$

Note that for $\Delta = 0$, the above modified algorithm coincides with the one described for the deterministic system. We now sum up the properties of this algorithm in the following Lemma.

Lemma 3.2.1. *For the algorithm (3.30,3.31) and subject to (3.29) with the upper bound $\hat{I}_j(t)$ given as in (3.41), it follows that*

1. $\left\| \hat{\theta}_i(t) - \theta_0 \right\| \leq \left\| \hat{\theta}^i(t-1) - \theta_0 \right\| \leq \left\| \hat{\theta}(0) - \theta_0 \right\| \quad t \geq 1$
2. $\lim_{t \rightarrow \infty} a(t-1) \frac{e_i(t)^2 - 4\Delta^2}{[c + \phi_i(t-1)^T \phi_i(t-1)]} \leq 0$
3. $\lim_{t \rightarrow \infty} \left\| \hat{\theta}_i(t) - \hat{\theta}_k(t) \right\| = 0$ with $k \in \{i-1, i+1\}$.
4. $\lim_{t \rightarrow \infty} \left\| \hat{\theta}_i(t) - \hat{\theta}_i(t-1) \right\| \leq \frac{2\Delta}{\sqrt{c}}$

Proof. 1) The proof is obvious from the discussion above.

2) From (3.40), we have

$$\begin{aligned}
\left\| \tilde{\theta}_i(t+1) \right\|^2 &\leq \left\| \tilde{\theta}(0) \right\|^2 + \hat{I}^i(t) - a(t) \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} + \frac{a(t) 2e_i(t+1)w_i(t+1)}{c + \phi_i(t)^T \phi_i(t)} \\
&\leq \left\| \tilde{\theta}(0) \right\|^2 + \hat{I}^i(t) - a(t) \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \\
&\quad + \frac{a(t)}{c + \phi_i(t)^T \phi_i(t)} \left[\frac{e_i(t+1)^2}{2} + 2w_i(t+1)^2 \right] \\
&\quad \text{Since } 2ab \leq ka^2 + \frac{b^2}{k} \text{ for any } k \\
&\leq \left\| \tilde{\theta}(0) \right\|^2 + \hat{I}^i(t) - a(t) \frac{e_i(t+1)^2}{c + \phi_i(t)^T \phi_i(t)} \\
&\quad + \frac{a(t)}{c + \phi_i(t)^T \phi_i(t)} \left[\frac{e_i(t+1)^2}{2} + 2\Delta^2 \right] \\
&\leq \left\| \tilde{\theta}(0) \right\|^2 + \hat{I}^i(t) - \frac{a(t)}{2} \left[\frac{e_i(t+1)^2 - 4\Delta^2}{c + \phi_i(t)^T \phi_i(t)} \right]
\end{aligned} \tag{3.49}$$

Since the left hand side above is bounded below by zero, the result follows.

The proof of 3) follows exactly as presented in Section II.

To prove 4) note that given $\epsilon_0 > 0 \exists T_{\epsilon_0} \mid \forall t \geq T_{\epsilon_0}$, and $\forall i$ we have

$$\left| a(t-1) \frac{e_i(t)}{[c + \phi_i(t-1)^T \phi_i(t-1)]^{1/2}} \right| \leq a(t-1) \frac{2\Delta}{\sqrt{c}} + \epsilon_0 \tag{3.50}$$

i.e. the (normalized) estimation error will be close to 2Δ (and shall remain as such) for all subsystems after time T_{ϵ_0} . We have

$$\begin{aligned} & \left\| \hat{\theta}_i(t+1) - \hat{\theta}_i(t) \right\| \\ & \leq \left\| \frac{\hat{\theta}_i(t) + \hat{\theta}_{i-1}(t) + \hat{\theta}_{i+1}(t)}{3} + \frac{a(t)\phi_i(t)}{c + \phi_i(t)^T \phi_i(t)} e_i(t+1) - \hat{\theta}_i(t) \right\| \end{aligned} \quad (3.51)$$

$$\leq \frac{1}{3} \left\| \hat{\theta}_{i-1}(t) - \hat{\theta}_i(t) \right\| + \frac{1}{3} \left\| \hat{\theta}_{i+1}(t) - \hat{\theta}_i(t) \right\| + a(t) \frac{2\Delta}{\sqrt{c}} + \epsilon_0 \quad (3.52)$$

$$\leq \frac{4}{3} \sqrt{3a\epsilon_1 - 9/4\epsilon_1^2} + a(t) \frac{2\Delta}{\sqrt{c}} + \epsilon_0 \quad (3.53)$$

$$\leq \frac{4}{3} \sqrt{3a\epsilon_1 - 9/4\epsilon_1^2} + \frac{2\Delta}{\sqrt{c}} + \epsilon_0 \quad (3.54)$$

The result follows as $t \rightarrow \infty$. \square

3.2.1 Simulation

In this section we present an implementation of our algorithm on the same circulant system that was presented in Section II with noise added to the output. The dynamics of the i_{th} subsystem are given as follows:

$$y_i(t+1) = -0.6y_i(t) + 0.1y_{i-1}(t) + 0.1y_{i+1}(t) + u_i(t) + w_i(t) \quad (3.55)$$

where $|w_i(t)| \leq \Delta = 0.4$. The initial guess for all the subsystems was taken as $\hat{\theta}(0) = [1, 1, 1, 1]^T$. The input for the subsystems was chosen as follows:

$$u_i(t) = \cos(t + i) + \sin(t - i/2) \quad (3.56)$$

A system, comprising of $N = 100$ subsystems, was simulated. The standard projection algorithm (where there is no coordination amongst the subsystems) was simulated along with the proposed distributed projection algorithm. For the implementation of distributed projection algorithm, $\epsilon = 10^{-5}$ was chosen. The results are captured in Figure (3.7). The top plots for Figure (3.7) present the distance of the estimates of the subsystem 1 from its neighbors, four to its right and four to its left. The bottom plots, on the other hand, present the distance of the subsystems ($N - 3, \dots, N, 1, \dots, 5$) from the true parameter θ_0 . A clear difference can be seen between the standard projection and distributed projection algorithm. For the standard projection algorithm, the parameter estimates cease to change at about $250(\times 10)$ time

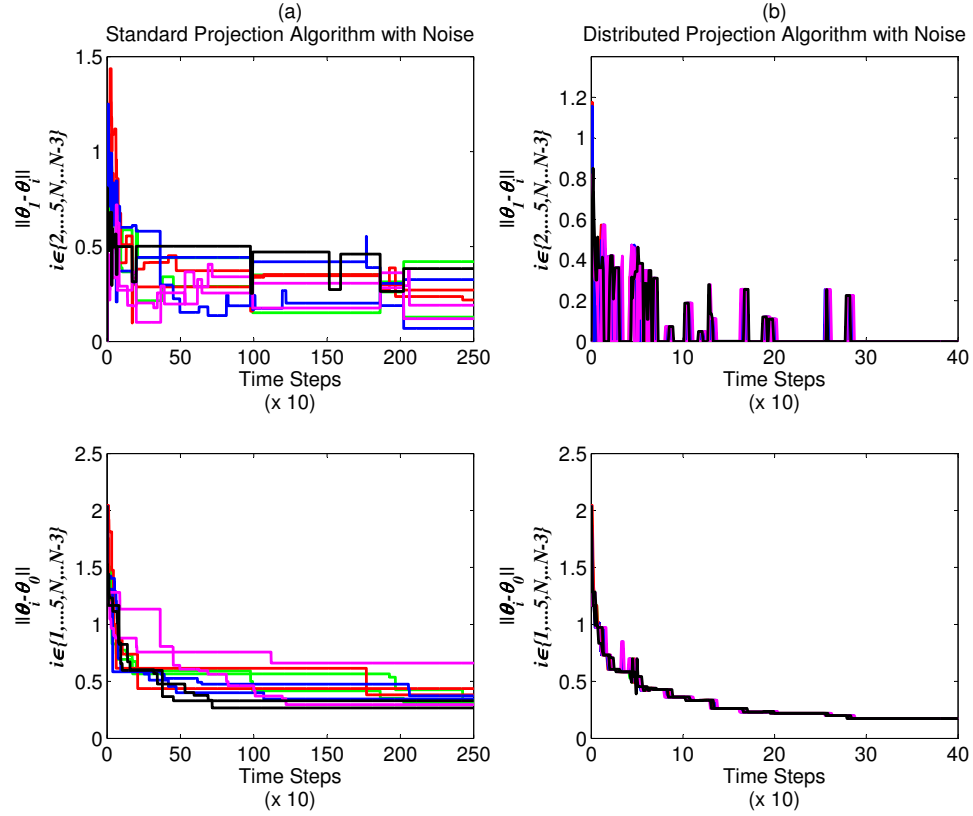


Figure 3.7 Parameter Estimation with Bounded Noise: Comparison of Distributed Projection Algorithm with (a) with Standard Projection Algorithm (b)

steps. Also, the parameter estimates of the neighbors for subsystem 1 are quite far apart from it. While the distributed projection algorithm shows no significant improvement in terms of getting closer to the true estimate, it settles to the final estimate a lot faster. The parameter estimation in this case ceases to improve after about $30(\times 10)$ time steps. It should be noted that the convergence of the parameter estimates in both the algorithms is enforced by the defined dead-zone, which is the same in the implementation of both algorithms. For the projection algorithm, it can be seen that the distance between the subsystem 1 and its neighbors vanishes completely at around $30(\times 10)$ time steps.

3.3 Conclusions

We have presented a distributed projection algorithm for system identification of spatiotemporally invariant systems in this chapter. Each subsystem receives information from all of its neighbors affecting it in order to construct the regression vector. Each subsystem, however, communicates only with its immediate neighbor to share its current estimate along with the related information (cumulative improvement index etc.). The best estimate available is picked in order to carry out the next iterate. For small estimation error, the scheme switches over to a smart-averaging routine. The scheme ensures continuous decay of estimation error and it serves to bring the local estimates arbitrarily close to one another. For a system with bounded noise added to the output, it was shown that for a given time step, the dead-zone algorithm serves to bring the local estimates arbitrarily close to one another. The scheme was seen to operate significantly better than the standard projection algorithm, even in the presence of bounded noise. It was also seen that the parameter estimates converge to the true value even when the standard projection algorithm fails to do so. This calls for an investigation into the excitation conditions necessary for true parameter identification for the distributed projection algorithm. This work is part of our future research work.

We remark at the end that the idea presented in this chapter can be extended easily to other identification schemes such as least squares. This is obvious since the key point is the ability to write the given estimate as the sum of the original guess and some improvement terms.

In the next chapter we combine the results of this chapter and the previous one in order to propose an indirect adaptive control scheme for LSTI systems based on certainty equivalence.

CHAPTER 4

ADAPTIVE CONTROLLERS FOR SPATIOTEMPORALLY INVARIANT SYSTEMS

In Chapter 2 we presented stability analysis of gradually varying spatiotemporal systems where the over all controller was generated by a collection of controllers indexed in space and time based on frozen spatially and temporally invariant descriptions of the plant. It was shown that the actual spatiotemporally varying system can be stabilized using frozen in space and time controllers, provided the rate of variations in the spatiotemporal dynamics are small enough. In Chapter 3 we presented a system identification scheme for spatiotemporally varying systems. It was seen that the local estimates get arbitrary close to one another and the normalized estimation error approaches zero as time $\rightarrow \infty$. In other words, the estimated plant qualifies to be gradually spatiotemporally varying (based on the definition presented in Chapter 2) for large enough time. We combine the results presented in Chapter 2 and 3 to present an indirect adaptive control scheme for spatiotemporally invariant systems. This approach generalizes the results presented in [86]. We base the controller design on certainty-equivalence approach, where at each step system parameters are estimated and the controller is implemented using the estimated parameters. At each estimation stage a modeling error is committed which affects the output of the plant. We show that under suitable assumptions drawn along the lines of the Chapter 2, coupled with the results presented in Chapter 3, a globally (weakly) stable adaptive scheme can be guaranteed.

4.1 Basic Setup

We will focus on SISO discrete-time spatiotemporally invariant systems that are recursively computable. Such plants can be represented by the following

transfer function for the i_{th} subsystem.

$$\hat{P} = \frac{\hat{B}(z, \lambda)}{\hat{A}(z, \lambda)} \quad (4.1)$$

Where \hat{B} , and \hat{A} are polynomials in z , and λ given by

$$\hat{A}(z, \lambda) = 1 + \sum_{t=0}^{m_1} \sum_{k=-n_1, k \neq 0}^{n_1} (a_k(t) z^k) \lambda^t \quad (4.2)$$

$$\hat{B}(z, \lambda) = \sum_{t=0}^{m_2} \sum_{k=-n_2, k \neq 0}^{n_2} (b_k(t) z^k) \lambda^t \quad (4.3)$$

The coefficients $\{a_k(t)\}$, and $\{b_k(t)\}$ are not known a priori. However, we will assume knowledge of the bound on the degrees of \hat{A} , and \hat{B} . We mark this down as an assumption in the following.

AS-1: The integers $m = \max(m_1, m_2)$, and $n = \max(n_1, n_2)$ are known a priori.

The above model can be written as:

$$y_i(t) = \phi_i(t-1)^T \theta_0 \quad (4.4)$$

where $y_i(t)$ denotes the (scalar) system output of subsystem i at time ' t ', $\phi_i(t-1)$ denotes a vector that is a linear function of

$$\begin{aligned} Y(t) &:= [\{y_i(t-1), y_i(t-2), \dots\}, \{y_{i-1}(t-1), y_{i-1}(t-2), \dots\}, \\ &\quad \{y_{i+1}(t-1), y_{i+1}(t-2), \dots\}, \dots] \\ U(t) &:= [\{u_i(t-1), u_i(t-2), \dots\}, \{u_{i-1}(t-1), u_{i-1}(t-2), \dots\}, \\ &\quad \{u_{i+1}(t-1), u_{i+1}(t-2), \dots\}, \dots] \end{aligned}$$

θ_0 is a vector that is formed from the coefficients $\{a_k(t)\}$, and $\{b_k(t)\}$. We shall employ the distributed projection algorithm from Chapter 3 for recursive estimation part of the adaptive scheme. At each instant of time the estimation algorithm supplies an estimate $\hat{\theta}_i(t)$ at the i_{th} subsystem from which we obtain the estimates $\hat{A}_{i,t}$, and $\hat{B}_{i,t}$. For the sake of completion, we list below the properties of the distributed projection algorithm from Chapter 3.

1. $\left\| \hat{\theta}_i(t) - \theta_0 \right\| \leq \left\| \hat{\theta}^i(t-1) - \theta_0 \right\| \leq \left\| \hat{\theta}(0) - \theta_0 \right\| \quad t \geq 1$
2. $\lim_{t \rightarrow \infty} \frac{e_i(t)}{[c + \phi_i(t-1)^T \phi_i(t-1)]^{1/2}} = 0$
3. $\lim_{t \rightarrow \infty} \left\| \hat{\theta}_i(t) - \hat{\theta}_k(t) \right\| = 0$ for $k \in \{i-1, i+1\}$
4. $\lim_{t \rightarrow \infty} \left\| \hat{\theta}_i(t) - \hat{\theta}_i(t-1) \right\| = 0 \quad \forall i$. This together with 3 implies

$$\lim_{t \rightarrow \infty} \left\| \hat{\theta}_i(t) - \hat{\theta}_k(t+l) \right\| = 0 \text{ for } k \in \{i, i-1, i+1\} \text{ and for finite } l$$

The estimation algorithm implies that the estimates remain bounded, their variation slows down locally, and the normalized estimation error gets small as time progresses.

4.2 Characterization of a Class of Gradually Varying Spatiotemporal Controllers

As discussed earlier, the analysis approach of Chapter 2 is extended to an indirect adaptive scheme where the plant is estimated recursively via the distributed projection algorithm of Chapter 3. The sequence of estimated plants is viewed as a gradually varying spatiotemporal system. The notions of gradual spatiotemporal variations is borrowed verbatim from Chapter 2 and the definitions shall not be repeated here. The local control law $\hat{K}_{i,t} = \frac{\hat{M}_{i,t}}{\hat{L}_{i,t}}$ is designed on the basis of frozen time and frozen space plants. Given an instance in space and time, the plant is thought of as a LSTI system, with the defining operators fixed at that time and space. The controllers are designed for the corresponding frozen LSTI system. The overall controller, thus, forms a sequence that can be regarded as a spatiotemporally varying controller. Since the frozen plant is LSTI, a frozen LSTI controller can be obtained using various methods, e.g. [16], [17] with different design objectives. Our approach covers these cases with the advantage of being applicable to more elaborate control techniques.

The frozen space and time operator that defines the control law satisfies the following Bezout identity

$$L_{i,t}A_{i,t} + M_{i,t}B_{i,t} = G_{i,t} \quad (4.5)$$

where $G_{i,t}^{-1} \in \mathcal{L}_{STI}$ for each fixed pair (i, t) , is the closed loop polynomial. The following result gives sufficient conditions for the l_∞ stability of a class of adaptive controllers.

Theorem 4.2.1. *Given $\hat{P} = \frac{\hat{B}}{\hat{A}}$ a LSTI plant, and N an integer such that the degrees of \hat{A} , and \hat{B} are bounded by N . Let $A_{i,t}$, and $B_{i,t}$ be the estimates of A , and B at the i_{th} subsystem at time t . Assume that a spatiotemporal varying controller K is implemented as follows.*

$$L_{i,t}u_i(t) = M_{i,t}(r_i(t) - y_i(t)) \quad (4.6)$$

$$L_{i,t}A_{i,t} + M_{i,t}B_{i,t} = G_{i,t} \quad (4.7)$$

where $L_{i,t}$, $M_{i,t}$, $G_{i,t} \in \mathcal{L}_{STI}$, and $\{r_i(t)\}$ is a bounded reference input. Let the following conditions hold:

1. The operators defining estimates of the plant are gradually time and space varying after time $T_p < \infty$ with rates γ_A and γ_B , i.e. $A_{i,t} \in GSTV(\gamma_A)$, and $B_{i,t} \in GSTV(\gamma_B)$.
2. The sequence of controllers are gradually time and space varying after time $T_k < \infty$, i.e. $M_{i,t} \in GSTV(\gamma_M)$, $L_{i,t} \in GSTV(\gamma_L)$, and $G_{i,t} \in GSTV(\gamma_G)$.
3. There exists an integer N_2 such that the degrees of $L_{i,t}$, $M_{i,t}$ are bounded by N_2 for all (i, t)
4. The zeros in λ of $\hat{G}_{i,t}(e^{j\theta}, \lambda)$, lie outside a disc of radius $1 + \epsilon$, for some $\epsilon > 0$ and for all θ .
5. The l_∞ to l_∞ norms of the LSTI operators $G_{i,t}^{-1} L_{i,t}$, $M_{i,t}$ are bounded uniformly in i , and t .

Then there exists a non-zero constant γ such that if γ_A , γ_B , γ_M , γ_L , $\gamma_G \leq \gamma$, the spatiotemporally varying controller will result in stable adaptive scheme.

Proof. The proof is along the lines of the proof of Theorem 2.2.1 as presented in Chapter 2. We highlight the difference in the sequel. The error signal $e_i(t)$ will appear as a disturbance of the above system, and hence the

operator mapping it to the signals $u_i(t)$, $y_i(t)$ is stable. Property 2 of the estimation scheme guarantees the boundedness of the error signal and consequently $u_i(t)$, $y_i(t)$, resulting in a stable adaptive system.

Spatiotemporally varying polynomials $A_{i,t}$, and $B_{i,t}$ are obtained from the distributed projection algorithm, driven by the error term $e_i(t) = y_i(t) - \phi_i(t)\hat{\theta}^i(t-1)$. The following equations are the basic components of the adaptive scheme:

$$e_i(t) = A_{i,t-1}y_i(t) - B_{i,t-1}u_i(t) \quad (4.8)$$

$$L_{i,t}u_i(t) = M_{i,t}(-y_i(t) + r_i(t)) \quad (4.9)$$

$$G_{i,t} = L_{i,t}A_{i,t} + M_{i,t}B_{i,t} \quad (4.10)$$

The basic idea is to relate the sequences $\{u_i(t)\}$ and $\{y_i(t)\}$ to the sequence $\{e_i(t)\}$ and $\{r_i(t)\}$, and show that the resulting operator is l_∞ stable. Using the Equations (4.8)-(4.10), this can be easily done and the resulting equations can be written as :

$$\begin{bmatrix} G_{i,t} + X_{i,t} & Y_{i,t} \\ -Z_{i,t} & G_{i,t} + W_{i,t} \end{bmatrix} \begin{bmatrix} u_i(t) \\ y_i(t) \end{bmatrix} = \begin{bmatrix} w_i(t) - M_{i,t}e_i(t) \\ z_i(t) + L_{i,t}e_i(t) \end{bmatrix} \quad (4.11)$$

where

$$X_{i,t} = A_{i,t}\nabla L_{i,t} + M_{i,t}\nabla B_{i,t-1} + M_{i,t}(B_{i,t} - B_{i,t-1}) \quad (4.12)$$

$$Y_{i,t} = A_{i,t}\nabla M_{i,t} - M_{i,t}\nabla A_{i,t-1} + M_{i,t}(A_{i,t} - A_{i,t-1}) \quad (4.13)$$

$$Z_{i,t} = B_{i,t}\nabla L_{i,t} - L_{i,t}\nabla B_{i,t-1} + L_{i,t}(B_{i,t} - B_{i,t-1}) \quad (4.14)$$

$$W_{i,t} = B_{i,t}\nabla M_{i,t} + L_{i,t}\nabla A_{i,t-1} - L_{i,t}(A_{i,t} - A_{i,t-1}) \quad (4.15)$$

$$w_i(t) = (AMr)_i(t) \quad (4.16)$$

$$z_i(t) = (BMr)_i(t) \quad (4.17)$$

For any pair (i, τ) , we factor $G_{i,\tau}$, evaluate the equations at $i = i, t = \tau$, and consider the evolution of the system as a function of i, τ . The equations can be written as

$$\left(\begin{pmatrix} I + F \end{pmatrix} \begin{pmatrix} u \\ y \end{pmatrix} \right)_i(\tau) = \begin{pmatrix} H_{i,\tau}w_i(\tau) - H_{i,\tau}M_{i,t}e_i(\tau) \\ H_{i,\tau}z_i(\tau) + H_{i,\tau}L_{i,t}e_i(\tau) \end{pmatrix} \quad (4.18)$$

where

$$F = \begin{pmatrix} H_{i,\tau}(G_{i,t} - G_{i,\tau}) + H_{i,\tau}X_{i,t} & H_{i,\tau}Y_{i,t} \\ H_{i,\tau}Z_{i,t} & H_{i,\tau}(G_{i,t} - G_{i,\tau}) + H_{i,\tau}W_{i,t} \end{pmatrix} \quad (4.19)$$

$H_{i,\tau}$ is the inverse of $G_{i,\tau}$. By assumption, $H_{i,\tau} \in \mathcal{L}_{STI}$. From the proof of Theorem 2.2.1 (Chapter 2), we know that there exists an integer T , such that $\|(I - P_T)F\| < 1$ where P_T is a temporal truncation operator. Note that the invertibility of $I + F$ is in essence concerned with the solvability of the equation

$$(\tilde{y} + F\tilde{y})_i(t) = (\tilde{e})_i(t) \quad (4.20)$$

for $(\tilde{e})_i(t) \in l_\infty$. Let $f_{i,i}(t, \tau)$ be the kernel representing the operator F . As argued above, there exists an integer T such that

$$C_1 = \sup_{i,t>T} \sum_{k=-\infty}^{\infty} \sum_{j=0}^t \|f_{i,k}(t, j)\| < 1 \quad (4.21)$$

Let us investigate the operator $I + F$ on the time segment $[0, T]$. On this time segment the operator $I + F$ is finite dimensional (temporally) and is given by.

$$\begin{pmatrix} I + F^{00} & 0 & \cdots & 0 \\ F^{10} & I + F^{11} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ F^{T0} & F^{T1} & \cdots & I + F^{TT} \end{pmatrix} \begin{pmatrix} \tilde{y}(0) \\ \tilde{y}(1) \\ \vdots \\ \tilde{y}(T) \end{pmatrix} = \begin{pmatrix} \tilde{e}(0) \\ \tilde{e}(1) \\ \vdots \\ \tilde{e}(T) \end{pmatrix} \quad (4.22)$$

Where, $\tilde{y}(t) = (\cdots, \tilde{y}_{i-1}(t), \tilde{y}_i(t), \tilde{y}_{i+1}(t), \cdots)'$ and

$$F^{t\tau} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \cdots \\ \cdots & f_{i-1,i-1}(t, \tau) & f_{i-1,i}(t, \tau) & f_{i-1,i+1}(t, \tau) & \cdots \\ \cdots & f_{i,i-1}(t, \tau) & f_{i,i}(t, \tau) & f_{i,i+1}(t, \tau) & \cdots \\ \cdots & f_{i+1,i-1}(t, \tau) & f_{i+1,i}(t, \tau) & f_{i+1,i+1}(t, \tau) & \cdots \\ \cdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (4.23)$$

The operator $P_T(I + F)P_T$ maps $P_T(l_\infty)$ into $P_T(l_\infty)$. Note that we have

$F^{tt} = 0$, and hence $P_T(I + F)P_T$ is invertible and the inverse is algebraic (does not require inversion of operators). Therefore, there exists a constant C , such that

$$\|P_T \tilde{y}\|_\infty = C \|\tilde{e}\|_\infty \quad (4.24)$$

Since the ‘tail’ of the operator is small, we should be able to bound the term $(I - P_T)\tilde{y}_i(t)$ in terms of $\tilde{e}(t)$, arising from the solution of $((I + F)\tilde{y})_i(t) = (\tilde{e})_i(t)$. We have

$$(I - P_T)\tilde{y}_i(t) + (I - P_T)F\tilde{y}_i(t) = (I - P_T)\tilde{e}_i(t) \quad (4.25)$$

which implies that

$$\|(I - P_T)\tilde{y}\|_\infty - \|(I - P_T)F\tilde{y}\|_\infty \leq \|(I - P_T)\tilde{e}\|_\infty \leq \|\tilde{e}\|_\infty \quad (4.26)$$

Investigating the term

$$(I - P_T)F\tilde{y}_i(t) = \begin{cases} \sum_{k=-\infty}^{\infty} \sum_{j=0}^t f_{i,k}(t, j) & \text{if } t > T \\ 0 & t \leq T \end{cases}$$

This implies that

$$\|(I - P_T)F\tilde{y}\| \leq \sum_{k=-\infty}^{\infty} \sum_{j=0}^t \|f_{i,k}(t, j)\| \|\tilde{y}_i(t)\| \quad \text{for } t > T \quad (4.27)$$

$$\begin{aligned} &= \sum_{k=-\infty}^{\infty} \sum_{j=0}^T \|f_{i,k}(t, j)\| \|\tilde{y}_i(t)\| \\ &+ \sum_{k=-\infty}^{\infty} \sum_{j=T+1}^t \|f_{i,k}(t, j)\| \|\tilde{y}_i(t)\| \end{aligned} \quad (4.28)$$

$$\begin{aligned} &\leq \|P_T \tilde{y}\|_\infty \sup_{i, t > T} \sum_{k=-\infty}^{\infty} \sum_{j=0}^T \|f_{i,k}(t, j)\| \\ &+ \|(I - P_T)\tilde{y}\|_\infty \sup_{i, t > T} \sum_{k=-\infty}^{\infty} \sum_{j=T+1}^t \|f_{i,k}(t, j)\| \end{aligned} \quad (4.29)$$

From Equation (4.26) we have

$$\|(I - P_T)\tilde{y}\|_\infty \leq M_1 \|P_T \tilde{y}\|_\infty + M_2 \|(I - P_T)\tilde{y}\|_\infty + \|\tilde{e}\|_\infty \quad (4.30)$$

which implies that

$$\|(I - P_T)\tilde{y}\|_\infty \leq \frac{M_2 C}{1 - M_1} \|\tilde{e}\|_\infty + \frac{1}{1 - M_1} \|\tilde{e}\|_\infty \quad (4.31)$$

Combining the results from (4.24), (4.26), and (4.31) we get the following bound,

$$\|\tilde{y}\|_\infty \leq \|P_T \tilde{y}\|_\infty + \|(I - P_T)\tilde{y}\|_\infty \leq k_1 \|\tilde{e}\|_\infty \quad (4.32)$$

for some positive constant k_1 . We have, therefore, established that in 4.11, the sequences $\{u_i(t), y_i(t)\}$ are bounded by the sequences $\{e_i(t), w_i(t), z_i(t)\}$. Equivalently,

$$\|\phi_i(t)\| \leq K_1 + K_2 \max_{i, \tau \leq t} \|e_i(\tau)\| \quad (4.33)$$

In order to complete the proof, we now introduce the following technical lemma [87].

Lemma 4.2.1. *If*

$$\lim_{t \rightarrow \infty} \frac{e_i(t)}{[c + \phi_i(t-1)^T \phi_i(t-1)]^{1/2}} = 0 \quad (4.34)$$

and (4.33) holds, then it follows that

$$\lim_{t \rightarrow \infty} e_i(t) = 0 \quad (4.35)$$

and $\{\|\phi_i(t)\|\}$ is bounded.

It now follows that e , u , and y are bounded.

□

4.2.1 Discussion

We consider two cases here.

System With No Noise ($\Delta = 0$)

Condition 1 in Theorem 4.2.1 is immediately satisfied from property 2 of the estimation scheme. In fact γ_A , γ_B are smaller than any positive γ for T large

enough. If one designs a compensator uniformly continuous with respect to the coefficients in $A_{i,t}$ and $B_{i,t}$, with the stability region in the complement of the disc of radius $1 + \epsilon$ for all θ , conditions 2, 4 will be satisfied. The boundedness conditions 3, 5 are generally satisfied when $\|G_{i,t}\|$ does not approach zero. Hence, any frozen space-time control design methodology that stabilizes the estimates and at the same time is continuous with respect to these estimates will result in stabilizing the unknown system.

System With Noise ($\Delta \neq 0$)

While the estimates in this case may be slowly varying in space, they may not be slowly varying in time (see Chapter 3), hence not guaranteeing condition 1 of Theorem 4.2.1. This means that the speed of estimation has to be controlled after some finite time T . This can be done by choosing the parameter $a(t)$ to be small enough (instead of being constantly 1 outside the dead-zone). Also, it is worth noting that the speed of the estimation scheme need not be controlled for all time but it has to be controlled for large enough time. The question of how small $a(t)$ have to be is difficult to answer a priori. The estimates derived in Theorem 4.2.1 give a very clear idea about the tradeoffs involved, but the issue remains dependent on the control scheme employed. We note that the results hold without any assumptions of persistence of excitation to force the parameters to converge, and their value is obvious in showing the limitation of the adaptive control in the presence of noise. Also, this characterization has the advantage of providing us with large class of stabilizing adaptive controllers, which makes it possible to satisfy performance specifications by choosing an appropriate one.

4.3 Convergence of Adaptive Scheme

We have already established the following,

1. The estimates $A_{i,t}$ and $B_{i,t}$ remain bounded, and are gradually varying in space and time
2. $\{u_i(t)\}$ is a bounded sequence
3. $\{y_i(t)\}$ is a bounded sequence

Theorem 4.3.1. *Subject to assumption Given in Theorem 4.2.1 (1-5), the following holds for the closed loop polynomial*

$$\lim_{t \rightarrow \infty} [G_{i,t}y_i(t) - B_{i,t}M_{i,t}r(t+1)] = 0 \quad (4.36)$$

Proof. We have already concluded in the proof of Theorem 4.2.1 that

$$\lim_{t \rightarrow \infty} e_i(t) = 0$$

Since

$$e_i(t) = A_{i,t-1}y_i(t) - B_{i,t-1}u_i(t)$$

we can write

$$\begin{aligned} L_{i,t}e_i(t+1) &= (LAy)_i(t+1) - (LBu)_i(t+1) \\ &= ([LA - L_{i,t}A_{i,t}]y)_i(t+1) - ([LB - L_{i,t}B_{i,t}]u)_i(t+1) \\ &+ L_{i,t}A_{i,t}y_i(t+1) - (B_{i,t}M_{i,t}r - B_{i,t}M_{i,t}y)_i(t) \\ &= ([LA - L_{i,t}A_{i,t}]y)_i(t+1) - ([LB - L_{i,t}B_{i,t}]u)_i(t+1) \\ &+ G_{i,t}y_i(t+1) - B_{i,t}M_{i,t}r_i(t+1) \end{aligned} \quad (4.37)$$

where

$$G_{i,t} = M_{i,t}B_{i,t} + A_{i,t}L_{i,t} \quad (4.38)$$

Taking limit as $t \rightarrow \infty$ of both sides of the above expression, and using the boundedness of $A_{i,t}$, $B_{i,t}$, $L_{i,t}$, $\{y_i(t)\}$, $\{u_i(t)\}$, $\{r_i(t)\}$ we get

$$\lim_{t \rightarrow \infty} [G_{i,t}y_i(t) - B_{i,t}M_{i,t}r(t+1)] = 0 \quad (4.39)$$

□

4.4 Conclusion

Based on the results presented in Chapter 2, and 3, we have presented an indirect adaptive control scheme for LSTI systems that is independent of the underlying control design methodology. We employ certainty-equivalence

approach, where at each step system parameters are estimated and the controller is implemented using the estimated parameters. We showed that under suitable assumptions drawn along the lines of the Chapter 2, a globally stable adaptive scheme can be guaranteed.

In Chapter 5 we shall turn our attention towards establishing conditions necessary and sufficient for robust stability of LSTI system under various kinds of perturbations.

CHAPTER 5

l_∞ AND l_2 ROBUSTNESS OF SPATIALLY INVARIANT SYSTEMS

In the previous chapters we analyzed parametric uncertainty in LSTI systems. Herein we consider nonparametric perturbations of various sorts in LSTI systems. In particular, we analyze the robustness of LSTI systems in the presence of spatiotemporal perturbations, that are not necessarily parametric, to derive necessary and sufficient conditions for robust stability.

Robust l_2 stability analysis for linear spatiotemporal invariant (LSTI) systems has been carried out for LSTI \mathcal{H}_∞ stable perturbations in [67] and μ -like conditions were established. The focus of this chapter is the robust l_∞ and l_2 stability analysis for other types of perturbations. In particular, this chapter aims to address the necessary and sufficient conditions for robust stability of LSTI stable systems in the presence of LSTV perturbations. We also investigate the robust stability of LSTI stable systems with the underlying perturbations being nonlinear spatiotemporal invariant (NLSTI). Inline with the approach taken in the previous chapters, we capitalize on the time domain representation of these spatiotemporal systems to show that the robustness conditions are analogous to the scaled small gain condition (which is equivalent to a spectral radius condition and a linear matrix inequality (LMI) for the l_∞ and l_2 case respectively) derived for standard linear time invariant (LTI) models subject to linear time varying or nonlinear perturbations (see [68], [69], or [70]).

5.1 Basic Setup

5.1.1 Spatially Invariant Systems

We consider spatiotemporal systems $M : u \rightarrow y$ on l_∞^e given by the convolution

$$y_i(t) = \sum_{\tau=0}^{\tau=t} \sum_{j=-\infty}^{j=\infty} m_{i-j}(t-\tau)u_j(\tau)$$

These systems can be viewed as an infinite array of interconnected LTI systems. These form identical building blocks in the system and the corresponding input-output relationship of the i_{th} block can be given as follows:

$$\begin{aligned} \begin{pmatrix} y_i(0) \\ y_i(1) \\ y_i(2) \\ \vdots \end{pmatrix} &= \begin{pmatrix} m_0(0) & 0 & 0 & \cdots \\ m_0(1) & m_0(0) & 0 & \cdots \\ m_0(2) & m_0(1) & m_0(0) & \cdots \\ \cdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_i(0) \\ u_i(1) \\ u_i(2) \\ \vdots \end{pmatrix} \\ &+ \begin{pmatrix} m_{-1}(0) & 0 & 0 & \cdots \\ m_{-1}(1) & m_{-1}(0) & 0 & \cdots \\ m_{-1}(2) & m_{-1}(1) & m_{-1}(0) & \cdots \\ \cdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_{i-1}(0) \\ u_{i-1}(1) \\ u_{i-1}(2) \\ \vdots \end{pmatrix} \\ &+ \begin{pmatrix} m_1(0) & 0 & 0 & \cdots \\ m_1(1) & m_1(0) & 0 & \cdots \\ m_1(2) & m_1(1) & m_1(0) & \cdots \\ \cdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_{i+1}(0) \\ u_{i+1}(1) \\ u_{i+1}(2) \\ \vdots \end{pmatrix} + \cdots \end{aligned}$$

where $\{u_i(t)\}$ is the input applied at the i_{th} block with $u_i(t) \in \mathbb{R}$ and $t \in \mathbb{Z}^+$ is the time index, and $\{m_i(t)\}$ is the pulse response corresponding to the i_{th} input, with $m_i(\cdot) \in \mathbb{R}$. Also, $\{y_i(t)\}$ is the output sequence of the i_{th} block, with $y_i(t) \in \mathbb{R}$. We can write the overall input-output relationship for a LSTI system as follows:

$$\begin{pmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} M_0 & & & & \\ M_1 & M_0 & & & \\ M_2 & M_1 & M_0 & & \\ M_3 & M_2 & M_1 & M_0 & \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u(0) \\ u(1) \\ u(2) \\ u(3) \\ \vdots \end{pmatrix} \quad (5.1)$$

where, $u(t) = (\cdots, u_{-1}(t), u_0(t), u_{+1}(t), \cdots)'$ and

$$M_t = \begin{pmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdots & m_{-1}(t) & m_0(t) & m_1(t) & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \quad t \in \mathbb{Z}^+$$

l_∞ Stability

We say that a LSTI SISO system M is l_∞ stable if its l_∞ induced norm is finite. It is a straightforward exercise to show that this condition requires the pulse response of the LSTI system to be absolutely summable. With regards to the system representation of (5.1), this condition reduces to the requirement that the l_1 norm of M satisfies

$$\|M\|_1 := \sum_{t=0}^{\infty} \sum_{i=-\infty}^{\infty} |m_i(t)| < \infty$$

Note that $\|M\|_1 = \sum_{t=0}^{\infty} \|M_t\|_1$.

l_2 Stability

We first define the z, λ transform for a LSTI system M as

$$\hat{M}(z, \lambda) = \sum_{t=0}^{\infty} \sum_{k=-\infty}^{\infty} (m_k(t) z^k) \lambda^t \quad (5.2)$$

It is known [16] that the l_2 induced norm of a LSTI system is equal to the \mathcal{H}_∞ norm of $\hat{M}(z, \lambda)$

$$\|M\|_{l_2\text{-ind}} = \left\| \hat{M} \right\|_{\mathcal{H}_\infty} := \sup_{\theta, \omega} | \hat{M}(e^{i\theta}, e^{j\omega}) | \quad (5.3)$$

We say that a LSTI system M is l_2 stable if $\left\| \hat{M} \right\|_{\mathcal{H}_\infty} < \infty$.

Remark

For the case when M is a n by p LSTI MIMO system, i.e., when $\{m_i(t)\}$ are n by p (real) matrices for every i and t , and hence M can be represented as a matrix of n by p LSTI SISO systems $\{M_{kj}\}$, the previous stability and

induced norm definitions have the usual generalizations as in the standard LTI systems [70].

5.1.2 Perturbation Models

We will consider various forms of temporaly causal perturbations. As usual, by a temporaly causal (proper) system T on l_∞^e we mean that $P_t T = P_t T P_t$ for all $t \in \mathbb{Z}^+$, where P_t is t steps truncation defined as $P_t(x) = P_t(x(0), x(1), \dots) = (x(0), x(1), \dots, x(t), 0, 0, \dots)$ for any $x \in l_\infty^e$. T is strictly temporaly causal (strictly proper) if $P_t T = P_t T P_{t-1}$. In the sequel we will use the terms causal (proper) to mean temporaly causal (proper).

LSTV Perturbations

The space of LSTV temporaly causal and stable perturbations $\Delta : y \rightarrow u$ are given by the convolution

$$u_i(t) = \sum_{\tau=0}^{\tau=t} \sum_{j=-\infty}^{j=\infty} \delta_{i,j}(t, \tau) y_j(\tau)$$

These perturbations can also be represented as a temporaly causal system

$$\Delta = \begin{pmatrix} \Delta(0,0) & & & & \\ \Delta(1,0) & \Delta(1,1) & & & \\ \Delta(2,0) & \Delta(2,1) & \Delta(2,2) & & \\ \Delta(3,0) & \Delta(3,1) & \Delta(3,2) & \Delta(3,3) & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \quad (5.4)$$

The various blocks $\Delta(i, j)$ in the perturbations are infinite matrices, the elements of which are obtained from the spatiotemporal pulse response $\{\delta_{i,j}(t, \tau)\}$.

We define the set $\mathbf{B}_{\Delta_{LSTV,p}}$ as

$$\mathbf{B}_{\Delta_{LSTV,p}} = \{\Delta \text{ causal, LSTV, with } \|\Delta\|_{l_p\text{-ind}} < 1\} \quad (5.5)$$

for $p = 2, \infty$

NLSTI Perturbations

We will also consider NLSTI temporally causal and stable perturbations. For $p = 2, \infty$, the set $\mathbf{B}_{\Delta_{NLSTI,p}}$ is defined as

$$\mathbf{B}_{\Delta_{NLSTI,p}} = \{\Delta \text{ causal, NLSTI, with } \|\Delta\|_{l_p\text{-ind}} < 1\} \quad (5.6)$$

System Interconnection

Throughout this chapter, we will be interested in the stability of the interconnected system shown in Figure 5.1, with $\Delta \in \mathbf{B}_{\Delta_{LSTV,p}}$ or $\Delta \in \mathbf{B}_{\Delta_{NLSTI,p}}$ strictly causal, and M a LSTI system which is l_p stable. We will investigate

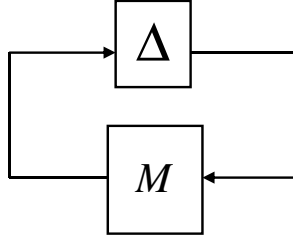


Figure 5.1 Stability Robustness Problem.

the stability in the cases 1) when Δ is unstructured and 2) Δ is structured. By structured, we mean Δ to be of the form $\Delta = \text{diag}(\Delta_1, \Delta_2, \dots, \Delta_n)$, where Δ_i is LSTV or NLSTI l_p -stable perturbation for all $1 \leq i \leq n$.

Structured Norm

Along the lines of [70], which we base our work on, we define the Structured Norm as follows. The Structured Norm, SN, is a map from the space of stable systems to the nonnegative reals defined as

$$SN_{\Delta,p}(M) = \frac{1}{\inf_{\Delta} \{ \|\Delta\|_{l_p\text{-ind}} \mid (I - \Delta M)^{-1} \text{ is not } l_p \text{ stable} \}}$$

where M is a LSTI l_p stable system and Δ in a given class i.e., LSTV or NLSTI. It is straightforward to verify directly from the definition that

$$(I - \Delta M)^{-1} \text{ is } l_p\text{-stable for all } \Delta \Leftrightarrow SN_{\Delta,p}(M) \leq 1$$

5.2 l_∞ Stability Robustness

In this section we present the robustness analysis of l_∞ stable LSTI systems with regards to unstructured and structured perturbations that are LSTV. We also present an investigation when the underlying perturbations are structured NLSTI.

5.2.1 LSTV Unstructured Perturbations

Consider the interconnection of l_∞ stable LSTI system M , with a non-structured perturbation $\mathbf{B}_{\Delta_{LSTV,\infty}}$ as shown in Figure 5.1. The following theorem presents necessary and sufficient conditions for the stability robustness of such a closed-loop system.

Theorem 1. *The closed loop system of Figure 5.1, with $\Delta \in \mathbf{B}_{\Delta_{LSTV,\infty}}$ and strictly proper, is robustly stable if and only if $\|M\|_1 \leq 1$.*

Proof. The sufficiency follows directly from the small gain theorem and the sub-multiplicative property of the norm, i.e.,

$$\|\Delta M\|_{l_\infty\text{-ind}} \leq \|\Delta\|_{l_\infty\text{-ind}} \|M\|_1 < 1 \quad (5.7)$$

Strict properness guarantees the well posedness of the closed loop system. For necessity we will show that if $\|M\|_1 \geq \gamma > 1$, then there exists a destabilizing $\Delta \in \mathbf{B}_{\Delta_{LSTV,\infty}}$. For simplicity, we will consider the case when M and Δ are SISO and divide the proof in two steps 1) we construct an unbounded signal and 2) use this unbounded signal for the construction of a destabilizing perturbation.

Construction of Unbounded Signal

M is shown in Figure 5.2 with $\xi \in l_\infty^e$ as its input and $z \in l_\infty^e$ as its output. The signal $y \in l_\infty^e$ is made up of the output z after a bounded signal, the output of a sign function (the operation of which is interpreted component wise so that the summation $\gamma \text{sgn}(z) + z$ makes sense), has been added to it. We interpret this bounded signal as an external signal injected for stability analysis. We aim to construct ξ satisfying the following:

1. ξ is unbounded.

2. ξ results in a signal y , such that $\| P_k \xi \|_\infty \leq \frac{1}{\gamma} \| P_k y \|_\infty$ where P_k is the truncation operator.

If y and ξ satisfy the second condition, then it is always possible to find Δ so that Δ is causal, has induced norm less than one and satisfies $\Delta y = \xi$. If the first requirement is also met then this Δ is also a destabilizing perturbation.

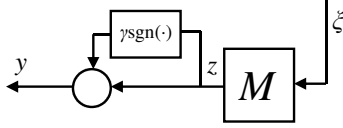


Figure 5.2 Signal Construction for Unstructured Uncertainty.

For simplicity of exposition, we assume that M has finite temporal impulse response of length N . While keeping $|\xi(k)| \leq 1$ for $k = 0, \dots, N-1$, the first N components can be constructed so that $\| M \|_1$ is achieved, where here, with some abuse of notation, we use $|\xi(k)|$ to indicate $\sup_i |\xi_i(k)|$. This implies that $\| P_{N-1} z \|_\infty \geq \gamma$, since $\| M \|_1 \geq \gamma$. This in turn implies that $\| P_{N-1} y \|_\infty \geq 2\gamma$. Note that the way ξ has been constructed, we have

$$\| P_{N-1} y \|_\infty \geq \gamma \| P_{N-1} \xi \|_\infty + \gamma \quad (5.8)$$

This relationship allows us to increase the magnitude of $|\xi(k)|$ for $k > N-1$ without violating the second requirement on ξ . In particular we let $|\xi(k)|$ to be as large as 2 for $k = N, \dots, 2N-1$. Again, we can pick $|\xi(k)|$ for this range of k as proceeded before so as to satisfy

$$\| P_{2N-1} y \|_\infty \geq \gamma \| P_{2N-1} \xi \|_\infty + \gamma \quad (5.9)$$

which allows us to further increase $|\xi(k)|$ by 1 for the next N components of ξ , and follow the entire procedure again. It is evident from this construction that when ξ is completely specified it will be unbounded, hence, meeting the first requirement. The second requirement is also met since we kept choosing $\xi(k)$ in a way that it does not grow too fast.

Construction of Destabilizing Perturbation

Given $\xi = \{\xi_i(t)\}_{i=-\infty, t=0}^{i=\infty, t=\infty}$ and $y = \{y_i(t)\}_{i=-\infty, t=0}^{i=\infty, t=\infty}$ from the previous section, we construct a destabilizing perturbation as follows. The construction of $\Delta(0, 0)$ is trivial if $y(0) = 0$, so we assume that there is at least one $i \in \mathbb{Z}$ such that $y_i(0) \neq 0$. Without loss of generality, we assume $i = 0$. We can now specify $\Delta(0, 0)$ as follows

$$\Delta(0, 0) = \begin{pmatrix} \vdots & & & & \\ \dots & 0 & \vdots & 0 & \dots \\ \dots & 0 & \frac{\xi_{-1}(0)}{y_0(0)} & 0 & \dots \\ \dots & 0 & \frac{\xi_0(0)}{y_0(0)} & 0 & \dots \\ \dots & 0 & \frac{\xi_{+1}(0)}{y_0(0)} & 0 & \dots \\ \dots & 0 & \vdots & 0 & \dots \\ \vdots & & & & \end{pmatrix} \quad (5.10)$$

Clearly, $\xi(0) = \Delta(0, 0)y(0)$. Also note that $\Delta(0, 0)$ has a spatially varying structure, and $\|\Delta(0, 0)\|_{l_\infty\text{-ind}} \leq \frac{1}{\gamma} < 1$ by construction, since each row has exactly one element and the magnitude of each element is strictly less than 1. Next we pick $\Delta(1, 0) = 0$, and specify $\Delta(1, 1)$ as follows. If $y_i(1) = 0$ for all $i \in \mathbb{Z}$, we simply pick $\Delta(1, 1) = 0$. We, hence, assume that there is at least one i such that $y_i(1) \neq 0$. Again, without loss of generality, we assume $i = 0$. The construction of $\Delta(1, 1)$ is given as follows:

$$\Delta(1, 1) = \begin{pmatrix} \vdots & & & & \\ \dots & 0 & \vdots & 0 & \dots \\ \dots & 0 & \frac{\xi_{-1}(1)}{y_0(1)} & 0 & \dots \\ \dots & 0 & \frac{\xi_0(1)}{y_0(1)} & 0 & \dots \\ \dots & 0 & \frac{\xi_{+1}(1)}{y_0(1)} & 0 & \dots \\ \dots & 0 & \vdots & 0 & \dots \\ \vdots & & & & \end{pmatrix} \quad (5.11)$$

Clearly, $\xi(1) = \Delta(1, 1)y(1)$. Also, note that $\Delta(1, 1)$ has a spatially varying structure and $\|\Delta(1, 1)\|_{l_\infty\text{-ind}} \leq \frac{1}{\gamma} < 1$ by construction. Next we pick $\Delta(2, 0) = \Delta(2, 1) = 0$ and specify $\Delta(2, 2)$ in the same way as above. It is

clear that when Δ is completely specified, it will have a diagonal structure as shown in (5.12) with only one element in any given row guaranteeing that $\|\Delta\|_{l_\infty\text{-ind}} \leq \frac{1}{\gamma} < 1$, and satisfying $\Delta y = \xi$. Moreover, Δ is causal and the construction above can be repeated by introducing delay in the construction of ξ so that Δ is strictly causal. This will guarantee the well posedness of the closed-loop system, implying that $(I - \Delta M)^{-1}$ exists, and is unstable by construction.

$$\Delta = \begin{pmatrix} \Delta(0,0) & & & & \\ 0 & \Delta(1,1) & & & \\ 0 & 0 & \Delta(2,2) & & \\ 0 & 0 & 0 & \Delta(3,3) & \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \quad (5.12)$$

□

5.2.2 LSTV Structured Perturbations

For simplicity we will show our result only for two SISO perturbation blocks. The entire result can be generalized for n perturbation blocks in a straight forward manner. It is easy to show that, since $\|D\Delta D^{-1}\|_{l_\infty\text{-ind}} = \|\Delta\|_{l_\infty\text{-ind}}$, the structured norm satisfies [70]

$$SN_{\Delta_{LSTV,\infty}}(M) \leq \inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 \quad (5.13)$$

where $\mathbf{D} = \{\text{diag}(d_1, d_2), d_i \in \mathbb{R} d_i > 0\}$.¹ Corresponding to the two perturbation blocks i.e $\Delta = \text{diag}(\Delta_1, \Delta_2)$, we can partition M as follows:

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad (5.14)$$

¹An element $D \in \mathbf{D}$ is a spatially and temporally constant operator and its z, λ transform is the matrix $\{\text{diag}(d_1, d_2)\}$.

where each M_{ij} is a temporally causal LSTI stable system. We now introduce the following non-negative matrix associated with (5.14).

$$\widetilde{M} = \begin{pmatrix} \|M_{11}\|_1 & \|M_{12}\|_1 \\ \|M_{21}\|_1 & \|M_{22}\|_1 \end{pmatrix} \quad (5.15)$$

Proposition 1. *The following conditions are equivalent:*

1. $\rho(\widetilde{M}) \leq 1$, where $\rho(\cdot)$ denotes the spectral radius.
2. The system of inequalities $x < \widetilde{M}x$ and $x \geq 0$ has no solutions, where the vector inequalities are to be interpreted componentwise.
3. $\inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 \leq 1$.

Proof. 1 \Leftrightarrow 2 follows exactly along the lines of [70]. We will only show 1 \Leftrightarrow 3 by showing that $\rho(\widetilde{M}) = \inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1$. By definition,

$$\begin{aligned} \|D^{-1}MD\|_1 &= \max_i \sum_{j=1}^2 \|(d_i^{-1}M_{ij}d_j)\|_1 = \max_i \sum_{j=1}^2 \left(\sum_{t=0}^{\infty} \sum_{k=-\infty}^{\infty} \left| \frac{d_j}{d_i} M_{ij,k}(t) \right| \right) \\ &= \max_i \sum_{j=1}^2 \frac{d_j}{d_i} \left(\sum_{t=0}^{\infty} \sum_{k=-\infty}^{\infty} |M_{ij,k}(t)| \right) = \max_i \sum_{j=1}^2 \frac{d_j}{d_i} \|M_{ij}\|_1 \end{aligned} \quad (5.16)$$

The expression on the right is also equal to the standard 1-norm of the matrix $D^{-1}\widetilde{M}D$. Denoting this norm by $|\cdot|_1$, we have $\|D^{-1}MD\|_1 = |D^{-1}\widetilde{M}D|_1$. Since the spectral radius of a matrix is bounded from above by any matrix norm of that matrix, we have

$$\inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 = \inf_{D \in \mathbf{D}} |D^{-1}\widetilde{M}D|_1 \geq \inf_{D \in \mathbf{D}} \rho(D^{-1}\widetilde{M}D) = \rho(\widetilde{M})$$

Choosing $D = \text{diag}(d_1, d_2)$ where $(d_1, d_2)'$ is the positive eigenvector corresponding to the eigenvalue $\rho(\widetilde{M})$, the inequality becomes an equality hence establishing an equivalence between 1 and 3. \square

The fact that for the optimum scaling $D = \text{diag}(d_1, d_2)$ all the rows $D^{-1}MD$ have the same norm will be exploited next in showing the necessity of condition 3 in the above proposition.

Theorem 2. *The system in Figure 5.1 achieves robust stability for all $\Delta \in \mathbf{B}_{\Delta_{LSTV}, \infty}$ if and only if*

$$\min_{D \in \mathbf{D}} \|D^{-1}MD\|_1 \leq 1$$

Equivalently, the Structured Norm can be computed exactly and is given by

$$SN_{\Delta_{LSTV}, \infty}(M) = \min_{D \in \mathbf{D}} \|D^{-1}MD\|_1 = \rho(\widetilde{M})$$

Proof. The sufficiency of this condition follows from (5.13). We now demonstrate that $\inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 \leq 1$ is necessary for robust stability. The approach is to show how to construct a destabilizing perturbation whenever $\inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 > 1$. Suppose that $\inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 \geq \gamma > 1$. Given that this infimum is in fact a minimum, and the fact that the rows of $D^{-1}MD$ will have equal norms, we have the following relationship for $n = 2$

$$\|(D^{-1}MD)_1\|_1 = \|(D^{-1}MD)_2\|_1 \quad (5.17)$$

where $(D^{-1}MD)_i$ denotes the i_{th} row of $D^{-1}MD$. The proof follows along the footsteps of the previous section, with the first step being a construction of an unbounded signal that gets amplified componentwise by $\|D^{-1}MD\|_1$ at the optimum D , and the second being the construction of a destabilizing perturbation using this signal.

Construction of Unbounded Signals

$D^{-1}MD$ shown in Figure 5.3 has $\xi = (\xi_1, \xi_2)$ as its input, where each $\xi_i \in l_\infty^e$, and $z = (z_1, z_2)$ as the output, with each $z_i \in l_\infty^e$. $y = (y_1, y_2)$ is made up of the output $z = (z_1, z_2)$ after a bounded signal, the output of a sign function, has been added to it, where the sign function operates on z_i component wise. Again, we interpret this bounded signal as an external signal injected for stability analysis. We aim to construct ξ satisfying the following:

1. ξ is unbounded.
2. ξ results in a signal y , satisfying $\|P_k \xi_i\|_\infty \leq \frac{1}{\gamma} \|P_k y_i\|_\infty$ for $i = 1, 2$, with P_k a temporal truncation operator.

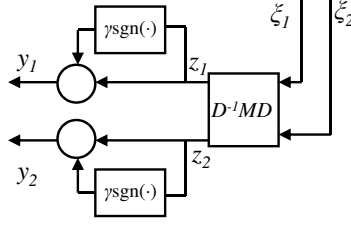


Figure 5.3 Signal Construction.

Simplifying the exposition, we assume that all M_{ij} 's have finite temporal pulse response of length N . While keeping $|\xi(k)| \leq 1$ for $k = 0, \dots, N-1$, the first N temporal components of ξ can be constructed so as to achieve $\|(D^{-1}MD)_1\|_1$. Since $\|(D^{-1}MD)_1\|_1 \geq \gamma$, this implies that $\|P_{N-1}z_1\|_\infty \geq \gamma$, which in turn implies that $\|P_{N-1}y_1\|_\infty \geq 2\gamma$. Next, while still keeping $|\xi(k)| \leq 1$, we pick the next N temporal components of ξ so as to achieve the second row norm $\|(D^{-1}MD)_2\|_1$. As a result we have $\|P_{2N-1}z_1\|_\infty \geq \gamma$, which implies that $\|P_{2N-1}y_1\|_\infty \geq 2\gamma$. Note that the second requirement mentioned above is met for $k = 0, \dots, 2N-1$. In addition, the way the first $2N$ terms of ξ have been constructed, we have

$$\|P_{2N-1}y_i\|_\infty \geq \gamma \|P_{2N-1}\xi_i\|_\infty + \gamma \quad i = 1, 2 \quad (5.18)$$

This relationship allows us to increase the magnitude of $|\xi(k)|$ to 2 for $k = 2N, \dots, 4N-1$ while satisfying the following relationship:

$$\|P_{4N-1}y_i\|_\infty \geq \gamma \|P_{4N-1}\xi_{i,j}\|_\infty + \gamma \quad i = 1, 2 \quad (5.19)$$

This allows us to increase $|\xi(\cdot)|$ by 1 for the next $2N$ temporal components, and repeat the whole procedure again. Hence, by the time when ξ is completely specified it will be unbounded, meeting the first requirement. The second requirement is also met since ξ was not allowed to grow too fast.

Construction of Destabilizing Perturbation

Given $\xi_1 = \{\xi_{1,i}(t)\}_{i=-\infty, t=0}^{i=\infty, t=\infty}$ and $y_1 = \{y_{1,i}(t)\}_{i=-\infty, t=0}^{i=\infty, t=\infty}$ from the previous section, we proceed as follows for the construction of destabilizing perturbation. The construction of $\Delta_1(0, 0)$ is trivial if $y_1(0) = 0$, so we assume that

there is at least one i such that $y_{1,i}(0) \neq 0$. Without loss of generality, we assume $i = 0$. We can specify $\Delta_1(0,0)$ as follows

$$\Delta_1(0,0) = \begin{pmatrix} \vdots & & & & \\ \dots & 0 & \vdots & 0 & \dots \\ \dots & 0 & \frac{\xi_{1,-1}(0)}{y_{1,0}(0)} & 0 & \dots \\ \dots & 0 & \frac{\xi_{1,0}(0)}{y_{1,0}(0)} & 0 & \dots \\ \dots & 0 & \frac{\xi_{1,+1}(0)}{y_{1,0}(0)} & 0 & \dots \\ \vdots & & & & \end{pmatrix} \quad (5.20)$$

Clearly, $\xi_1(0) = \Delta_1(0,0)y_1(0)$. Also, note that $\Delta_1(0,0)$ has a spatially varying structure and $\|\Delta_1(0,0)\|_{l_\infty\text{-ind}} \leq \frac{1}{\gamma} < 1$ by construction. Next we pick $\Delta_1(1,0) = 0$ and specify $\Delta_1(1,1)$ as follows. If $y_{1,i}(1) = 0$ for all i , we simply pick $\Delta_1(1,1) = 0$. We, hence, assume that there is at least one i such that $y_{1,i}(1) \neq 0$. Again, without loss of generality, we assume $i = 0$. The construction of $\Delta_1(1,1)$ is given as follows:

$$\Delta_1(1,1) = \begin{pmatrix} \vdots & & & & \\ \dots & 0 & \vdots & 0 & \dots \\ \dots & 0 & \frac{\xi_{1,-1}(1)}{y_{1,0}(1)} & 0 & \dots \\ \dots & 0 & \frac{\xi_{1,0}(1)}{y_{1,0}(1)} & 0 & \dots \\ \dots & 0 & \frac{\xi_{1,+1}(1)}{y_{1,0}(1)} & 0 & \dots \\ \vdots & & & & \end{pmatrix} \quad (5.21)$$

Clearly, $\xi_1(1) = \Delta_1(1,1)y_1(1)$. Also, note that $\Delta_1(1,1)$ has a spatially varying structure and $\|\Delta_1(1,1)\|_{l_\infty\text{-ind}} \leq \frac{1}{\gamma} < 1$ by construction. Next we pick $\Delta_1(2,0) = \Delta_1(2,1) = 0$ and specify $\Delta_1(2,2)$ in the same way as above. It is clear that when Δ is completely specified, it will have a diagonal structure as shown in (5.22) with only one element in any given row guaranteeing that $\|\Delta_1\|_{l_\infty\text{-ind}} \leq \frac{1}{\gamma} < 1$, and satisfying $\Delta_1 y_1 = \xi_1$. Moreover, Δ_1 is causal and the construction above can be repeated so that Δ_1 is strictly causal. This will guarantee the well posedness of the closed-loop system implying that

$(I - \Delta M)^{-1}$ exists and is unstable by construction.

$$\Delta_1 = \begin{pmatrix} \Delta_1(0,0) & & & & \\ 0 & \Delta_1(1,1) & & & \\ 0 & 0 & \Delta_1(2,2) & & \\ 0 & 0 & 0 & \Delta_1(3,3) & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5.22)$$

□

5.2.3 NLSTI Structured Perturbations

We will only discuss the case of structured NLSTI perturbations as the case of unstructured NLSTI perturbations will become obvious from it. From the definition of structured norm, it follows that

$$SN_{\Delta_{NLSTI},\infty}(M) \leq \inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1 \quad (5.23)$$

In the following, we show that the equality also holds in (5.23).

Theorem 3. *The structured norm for NLSTI perturbations satisfies*

$$SN_{\Delta_{NLSTI},\infty}(M) = SN_{\Delta_{LSTV},\infty}(M) = \inf_{D \in \mathbf{D}} \|D^{-1}MD\|_1$$

Proof. The proof of this fact follows exactly as the proof of Theorem 2, except for the construction of the destabilizing perturbation. Given the signals y and ξ , we show that a NLSTI perturbation can be constructed to destabilize the closed-loop. Let the signals y_i and ξ_i be given as before ($i = 1, 2$). Δ_i must be such that $\|\Delta_i\|_{l_\infty-\text{ind}} < 1$ and $\Delta_i y_i = \xi_i$. We define Δ_i as follows:

$$(\Delta_i f)(k) = \begin{cases} \xi_{i,l-m}(k-j) & \text{if for some integers } m \text{ and } j, \text{ with } j \geq 0, \\ P_k f = P_k S_{m,j} y_i & \\ 0 & \text{otherwise} \end{cases}$$

where $S_{m,j}$ is the shift operator by m spatial and j temporal steps.² It can be seen that Δ_i is a causal, nonlinear spatiotemporal invariant system. It

²For example, if $x \in l_\infty^e$, then

$$S_{1,1}((\cdots, x_{-1}(0), x_0(0), x_1(0), \cdots)', (\cdots, x_{-1}(1), x_0(1), x_1(1), \cdots)', (\cdots, x_{-1}(2), x_0(2), x_1(2), \cdots)' \cdots)$$

has a norm less than one and maps y_i to ξ_i . \square

We note here that if, instead of NLSTI, the underlying perturbations are causal l_∞ -stable nonlinear spatiotemporal varying (NLSTV), the structured norm is obviously the same as in the above theorem since LSTV perturbations are a subset of NLSTV.

5.2.4 Numerical Example

As an example to calculate the sub-optimal scaling D , we consider a LSTI approximation that models the microcantilever array presented in [6]. The system consists of infinitely many microcantilevers connected to a base, each forming a micro-capacitor, with the second rigid plate located underneath the microcantilever. The vertical displacement of each microcantilever can be controlled by applying a voltage across the plates. Although each microcantilever is independently actuated, its dynamics are influenced by the presence of other microcantilevers. As elaborated in Figure 5.4, this coupling has two sources of origin: 1) mechanical, since the microcantilevers are attached to the same base and 2) electrical, due to the fringing fields generated by the micro-capacitors nearby. The dynamics for the i_{th} microcantilever can

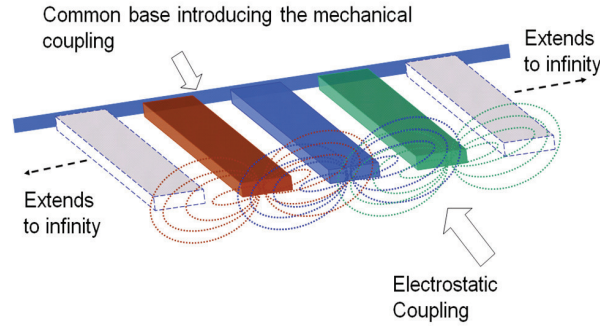


Figure 5.4 Schematic showing the layout of the infinite dimensional microcantilever array with mechanical and electrostatic coupling.

$$= (0, (\cdots, x_{-2}(0), x_{-1}(0), x_0(0), \cdots)', (\cdots, x_{-2}(1), x_{-1}(1), x_0(1), \cdots)', \cdots)$$

be given as follows:

$$\begin{aligned}\dot{x}_i &= Ax_i + Bu_i + \sum_{j=-\infty, j \neq i}^{\infty} G_{i,j}x_j \\ y_i &= Cx_i + Du_i\end{aligned}$$

where $G_{i,j}$ captures the mechanical and electrostatic coupling effects from the neighboring microcantilevers. A distributed controller was designed to decouple the dynamics of this system allowing independent actuation of each microcantilever. We are interested in assessing the robust stability of the (nominal) closed loop system M (system formed by closing the loop of the plant and the designed controller). The (nominal) closed loop M is a 2×2 LSTI system. The (nominal) closed loop system satisfies.

$$\|M\|_1 = \gamma \approx 2.14$$

Considering the perturbation block of $\Delta = \text{diag}\{\Delta_1, \Delta_2\}$ to assess robust stability of the system, we calculate the corresponding matrix \widetilde{M} (necessary details are given in the Appendix A). The spectral radius $\rho(\widetilde{M})$ comes out to be ≈ 0.0011 . This implies that the system can tolerate any structured spatiotemporal varying perturbation $\Delta = \text{diag}\{\Delta_1, \Delta_2\}$ with $\|\Delta_1\|_1 < 874$ and $\|\Delta_2\|_1 < 874$. Note that if the diagonal structure of Δ is ignored and we use the l_1 norm criterion for robustness, the size of allowable perturbations reduces dramatically to $1/\gamma = 0.467$ as opposed to the SN value achieved by solution of $\inf_{d \in \mathbf{D}} \|D^{-1}MD\|_1 = \hat{\gamma} \approx 0.0011$.

5.3 l_2 Stability Robustness

In this section we present the robustness analysis of l_2 stable LSTI systems with regards to unstructured and structured perturbations that are LSTV. We also present an investigation when the underlying perturbations are structured NLSTI.

5.3.1 LSTV Unstructured Perturbations

In [71], l_2 stability analysis of multidimensional systems subject to specific types of structured perturbations was carried under an LMI framework to

obtain robustness conditions equivalent to a scaled small gain condition. Our approach here produces the same outcome and, hence, in a sense is equivalent to [71], although the overall development is different as it relies on the same ideas presented in the previous section for l_∞ robustness and generalizes the one in [70]. In the sequel we elaborate in some detail on these developments. Let M be a LSTI l_2 stable system. It follows from the small gain theorem argument that

$$SN_{\Delta_{LSTV,2}}(M) \leq \left\| \hat{M} \right\|_{\mathcal{H}_\infty} \quad (5.24)$$

As in the l_∞ case, the upper bound is equal to the structured norm as stated in the following theorem.

Theorem 4. *The system in Figure 5.1 achieves robust stability for all $\Delta \in \mathbf{B}_{\Delta_{LSTV,2}}$ if and only if $\left\| \hat{M} \right\|_{\mathcal{H}_\infty} \leq 1$. Equivalently, the structured norm is given as*

$$SN_{\Delta_{LSTV,2}}(M) = \left\| \hat{M} \right\|_{\mathcal{H}_\infty} \quad (5.25)$$

Proof. We will show the result for the case of SISO block only. To establish this result, we first show the following lemmas. For simplicity, we assume that M is temporally FIR of length N_1 . The result generalizes in a straightforward way. Define:

$$k(f) := \|Mf\|_2^2 - \|f\|_2^2 \quad \text{where } f \in l_2 \quad (5.26)$$

Lemma 1. *If $\left\| \hat{M} \right\|_{\mathcal{H}_\infty} > 1$, then there exists $f \in l_2$ such that $k(f) > 0$*

Proof. Suppose on the contrary, for every $f \in l_2$, we have $k(f) \leq 0$, then

$$k(f) = \|Mf\|_2^2 - \|f\|_2^2 \leq 0 \quad (5.27)$$

$$\implies \sup \frac{\|Mf\|_2}{\|f\|_2} = \left\| \hat{M} \right\|_{\mathcal{H}_\infty} \leq 1 \quad (5.28)$$

which is a contradiction. \square

Lemma 2. *There exists a destabilizing perturbation $\Delta \in \mathbf{B}_{\Delta_{LSTV,2}}$ of the system in Figure 5.1, if there exists an $f \in l_2$ such that $k(f) > 0$.*

Proof. Since $k(f) > 0$, then there exists an $N_2 \geq N_1$ and a $\gamma^2 \geq 1$ such that

$$\|P_{N_2-1}Mf\|_2^2 \geq \gamma^2 \|P_{N_1-1}f\|_2^2 \quad (5.29)$$

where P_k is the temporal truncation operator. Without loss of generality, we assume that f has a finite temporal length N_2 , i.e. $f(k) = 0$ for all $k \geq N_2$. The proof is divided into two steps. The first step is the construction of a signal $\xi \in l_\infty^e \setminus l_2$, such that the output is amplified by γ^2 . The next step is to use this signal to construct a destabilizing perturbation.

Construction of Unbounded Signal

Define the signal ξ as follows:

$$\xi = \sum_{k=0}^{\infty} S_{k(N_1+N_2)}f \quad (5.30)$$

where $S_{k(N_1+N_2)}$ is the temporal shift operator by $k(N_1 + N_2)$ temporal steps. We remark here that this operator is same as the spatiotemporal shift operator $S_{m,j}$ presented in Section III (C), with $m = 0$. For simplicity, we omit the subscript identifying the spatial shift. The signal ξ can be visualized as a signal made up from the nonzero components of f (which we denote by f) by shifting it, and adding zeros in between, i.e.,

$$\xi = \{ \underbrace{f}_{N_2}, \overbrace{0}^{N_1}, f, 0, \dots \} \quad (5.31)$$

The action of M on ξ can be decomposed as follows:

$$\begin{aligned} y = M\xi &= \sum_{k=0}^{\infty} S_{k(N_1+N_2)}Mf \\ &= \sum_{k=0}^{\infty} S_{k(N_1+N_2)}(P_{N_2-1}Mf + (P_{N_1+N_2-1} - P_{N_2-1})Mf) \end{aligned}$$

Define M_0 and M_1 as follows:

$$M_0 := P_{N_2-1}MP_{N_2-1}, \quad M_1 := S_{-N_2}(P_{N_1+N_2-1} - P_{N_2-1})MP_{N_2-1}$$

Then y can be written as

$$y = \{M_0 f, M_1 f, M_0 f, M_1 f \dots\} \quad (5.32)$$

Defining $\hat{y} = M_0 f$, and $\tilde{y} = M_1 f$, we can write y as follows:

$$y = \{\hat{y}, \tilde{y}, \hat{y}, \tilde{y} \dots\} \quad (5.33)$$

It follows from (5.29) and (5.32) that for any $k \geq 0$

$$\|P_{k(N_1+N_2-1)} y\|_2^2 \geq \gamma^2 \|P_{k(N_1+N_2-1)} \xi\|_2^2 \quad (5.34)$$

Construction of Destabilizing Perturbation

We construct $\hat{\Delta}$ satisfying the following:

$$f = \hat{\Delta} \hat{y}, \quad \|\hat{\Delta}\|_{l_2\text{-ind}} \leq \frac{1}{\gamma} \quad (5.35)$$

The construction of such a $\hat{\Delta}$ can be given as:

$$\hat{\Delta} = \begin{pmatrix} \hat{\Delta}_{0,0} & \cdots & \hat{\Delta}_{0,N_2-1} \\ \vdots & & \vdots \\ \hat{\Delta}_{N_2-1,0} & \cdots & \hat{\Delta}_{N_2-1,N_2-1} \end{pmatrix}$$

where

$$\begin{aligned} \hat{\Delta}_{i,j} &= \frac{1}{\|\hat{y}\|_2^2} \cdot \begin{pmatrix} \vdots \\ f_{-1}(i) \\ f_0(i) \\ f_{+1}(i) \\ \vdots \end{pmatrix} (\cdots y_{-1}(j) \ y_0(j) \ y_{+1}(j) \ \cdots) \\ &= \frac{1}{\|\hat{y}\|_2^2} \cdot f(i) y(j)' \end{aligned}$$

Equivalently, we can write

$$\hat{\Delta} = \frac{1}{\|\hat{y}\|_2^2} \cdot \begin{pmatrix} f(0)y'(0) & \cdots & f(0)y'(N_2 - 1) \\ \vdots & & \vdots \\ f(N_2 - 1)y'(0) & \cdots & f(N_2 - 1)y'(N_2 - 1) \end{pmatrix}$$

or succinctly

$$\hat{\Delta} = \frac{f\hat{y}'}{\|\hat{y}\|_2^2}$$

and $\|\hat{\Delta}\|_{l_2\text{-ind}}$ can be evaluated as follows:

$$\begin{aligned} \|\hat{\Delta}\|_{l_2\text{-ind}} &:= \sup \frac{\|\hat{\Delta}g\|_2}{\|g\|_2} = \frac{\sqrt{g'\hat{\Delta}'\hat{\Delta}g}}{\|g\|_2} = \sqrt{\frac{g'\hat{y}f'f\hat{y}'g}{\|\hat{y}\|_2^4 \|g\|_2^2}} \\ &= \sqrt{\frac{\|f\|_2^2 g'\hat{y}\hat{y}'g}{\|\hat{y}\|_2^4 \|g\|_2^2}} \\ &\leq \sqrt{\frac{\|f\|_2^2 (\|g\|_2 \|y\|_2)(\|y\|_2 \|g\|_2)}{\|\hat{y}\|_2^4 \|g\|_2^2}} \leq \frac{1}{\gamma} \end{aligned}$$

Now, define the perturbation:

$$\Delta = \begin{pmatrix} 0 & \cdots & & & \\ 0 & 0 & \cdots & & \\ \hat{\Delta} & 0 & 0 & \cdots & \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \hat{\Delta} & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

With this perturbation, the output in each channel for the input ξ is given by

$$(f, 0, f - \hat{\Delta}\hat{y}, 0, f - \hat{\Delta}\hat{y}, \cdots) = (f, 0, 0, \cdots) \in l_2$$

This immediately implies that $(I - \Delta M)^{-1}$ is not l_2 -stable since it maps a signal in l_2 to a signal $l_\infty^e \setminus l_2$. Notice that $(I - \Delta M)^{-1}$ is well defined since Δ is strictly proper. \square

If $\|\hat{M}\|_{\mathcal{H}_\infty} \leq 1$ then the system is stable. Suppose that $\|\hat{M}\|_{\mathcal{H}_\infty} > 1$. By Lemma 1, there exists a function $f \in l_2$ such that $k(f) > 0$. It follows by Lemma 2 that there exists a destabilizing perturbation of the system in Figure 5.1. \square

5.3.2 LSTV Structured Perturbations

For simplicity, we will show our result only for two SISO perturbation blocks. The entire result can be generalized in a straight forward manner. Let M be a LSTI l_2 stable system. It follows from the small gain theorem argument that

$$SN_{\Delta_{LSTV,2}}(M) \leq \inf_{D \in \mathbf{D}} \|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty} \quad (5.36)$$

Theorem 5. *The system in Figure 5.1 achieves robust stability for all $\Delta \in \mathbf{B}_{\Delta_{LSTV,2}}$ if and only if $\inf_{D \in \mathbf{D}} \|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty} \leq 1$. Equivalently, the structured norm is given as*

$$SN_{\Delta_{LSTV,2}}(M) = \inf_{D \in \mathbf{D}} \|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty} \quad (5.37)$$

Proof. For simplicity, we assume that M is temporally FIR of length N_1 . The generalization is straightforward. Define:

$$k_i(f) := \| (Mf)_i \|_2^2 - \| f_i \|_2^2 \quad \text{for } i = 1, 2 \text{ and } f \in l_2 \quad (5.38)$$

$(Mf)_i$ denotes the i_{th} row of Mf . In order to establish the result of Theorem 5, we first invoke Lemma 7.3.2 in [70], that states the following:

Lemma 3. *Suppose that \hat{M} is such that $\inf_{D \in \mathbf{D}} \|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty}$ has a finite non zero minimizer. If*

$$\inf_{D \in \mathbf{D}} \|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty} > 1 \quad (5.39)$$

then there exists a function $f \in l_2$ such that

$$k_i(f) > 0 \quad \text{for } i = 1, 2 \quad (5.40)$$

We now show the following lemma.

Lemma 4. *There exists a destabilizing perturbation $\Delta \in \mathbf{B}_{\Delta_{LSTV,2}}$ of the system in Figure 5.1, if there exists an $f \in l_2$ such that $k_i(f) > 0$ for $i = 1, 2$.*

Proof. The proof follows in the footsteps of the proof of Lemma 2. Since $k_i(f) > 0$, then there exists an $N_2 \geq N_1$ and a $\gamma^2 \geq 1$ such that

$$\| P_{N_2-1}(Mf)_i \|_2^2 \geq \gamma^2 \| P_{N_1-1}f_i \|_2^2 \quad \text{for } i = 1, 2 \quad (5.41)$$

where P_k is the temporal truncation operator. Without loss of generality, we assume that f has a finite temporal length N_2 , i.e. $f(k) = 0$ for all $k \geq N_2$.

Construction of Unbounded Signal

With the definitions of ξ , M_0 , and M_1 presented in the previous section, we can directly write

$$y = \{M_0f, M_1f, M_0f, M_1f \dots\} = \{\hat{y}, \tilde{y}, \hat{y}, \tilde{y} \dots\} \quad (5.42)$$

It follows from (5.41), and (5.42) that for any $k \geq 0$

$$\| P_{k(N_1+N_2-1)}y_i \|_2^2 \geq \gamma \| P_{k(N_1+N_2-1)}\xi_i \|_2^2 \quad \text{for } i = 1, 2 \quad (5.43)$$

Construction of Destabilizing Perturbation

We construct $\hat{\Delta}_i$ as follows

$$\hat{\Delta}_i = \frac{f_i \hat{y}'_i}{\| \hat{y}_i \|_2^2}; \quad \left\| \hat{\Delta}_i \right\|_{l_2\text{-ind}} \leq \frac{1}{\gamma} \quad \text{for } i = 1, 2$$

Now, define the perturbation:

$$\Delta = \begin{pmatrix} 0 & \dots & & & \\ 0 & 0 & \dots & & \\ \hat{\Delta} & 0 & 0 & \dots & \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \hat{\Delta} & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{with } \hat{\Delta} = \begin{pmatrix} \hat{\Delta}_1 & 0 \\ 0 & \hat{\Delta}_2 \end{pmatrix}$$

With this perturbation, the output in each channel for the input ξ is given

by

$$(f_i, 0, f_i - \hat{\Delta}_i \hat{y}_i, 0, f_i - \hat{\Delta}_i \hat{y}_i, \dots) = (f_i, 0, 0, \dots) \in l_2$$

This immediately implies that $(I - \Delta M)^{-1}$ is not l_2 -stable since it maps a signal in l_2 to a signal $l_\infty^e \setminus l_2$. Notice that $(I - \Delta M)^{-1}$ is well defined since Δ is strictly proper. \square

If $\left\| D^{-1} \hat{M} D \right\|_{\mathcal{H}_\infty} \leq 1$ then the system is stable. Suppose that

$$\left\| D^{-1} \hat{M} D \right\|_{\mathcal{H}_\infty} > 1$$

By Lemma 3, there exists a function $f \in l_2$ such that $k(f) > 0$. It follows by Lemma 4 that there exists a destabilizing perturbation of the system in Figure 5.1. \square

5.3.3 NLSTI Structured Perturbations

Here we present the case of structured NLSTI perturbations only, since the case of unstructured NLSTI perturbations will become obvious from it. From the definition of structured norm, it follows that

$$SN_{\Delta_{NLSTI,2}}(M) \leq \inf_{D \in \mathbf{D}} \left\| D^{-1} \hat{M} D \right\|_{\mathcal{H}_\infty} \quad (5.44)$$

We remark here that the above inequality also holds if the underlying perturbations are causal l_2 -stable nonlinear spatiotemporal varying. However, since this set contains LSTV perturbations, it follows that equality holds. Similar to the l_∞ case, the equality in (5.44) also holds as shown in the following theorem.

Theorem 6. *The structured norm for NLSTI perturbations satisfies*

$$SN_{\Delta_{NLSTI,2}}(M) = SN_{\Delta_{LSTV,2}}(M) = \inf_{D \in \mathbf{D}} \left\| D^{-1} \hat{M} D \right\|_{\mathcal{H}_\infty}$$

Proof. The proof follows exactly as in the LSTV case if we can show that a NLSTI perturbation can be constructed such that

$$\Delta = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix}, \quad \Delta_i y_i = (0, 0, f_i, 0, f_i, \dots) = S_{N_1+N_2} \xi_i, \quad i = 1, 2$$

with $\|\Delta\|_{l_2\text{-ind}} < 1$. Consider the following perturbations

$$(\Delta_i g)(k) = \begin{cases} 0 & \text{if } k < N_1 + N_2 \\ \xi_{i,l-m}(k - j - N_1 - N_2) & \text{if for some integers } m \text{ and } j, \\ & \text{with } j \geq 0, P_k g = P_k S_{m,j} y_i \\ 0 & \text{otherwise} \end{cases}$$

It can be verified that Δ is a causal NLSTI perturbation. It satisfies

$$\|\Delta\|_{l_2\text{-ind}} < 1 \text{ with } \Delta_i y_i = S_{N_1+N_2} \xi_i.$$

□

5.3.4 Remark

The condition requiring the existence of a scaling matrix D can be readily cast into a family of LMIs over the spatial Fourier frequency parameter θ . For $\{\hat{A}(z), \hat{B}(z), \hat{C}(z), \hat{D}(z)\}$ a state space realization of \hat{M} parameterized by spatial Fourier transform (see [16] for details), it is a straight forward exercise to show (using the KYP lemma for discrete systems) that the condition $\|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty} < 1$ is equivalent to the feasibility condition of the following LMI over the Fourier frequency parameter θ :

$$\begin{aligned} & \begin{bmatrix} \hat{A}(e^{i\theta}) & \hat{B}(e^{i\theta}) \\ \hat{C}(e^{i\theta}) & \hat{D}(e^{i\theta}) \end{bmatrix}^* \begin{bmatrix} X(e^{i\theta}) & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} \hat{A}(e^{i\theta}) & \hat{B}(e^{i\theta}) \\ \hat{C}(e^{i\theta}) & \hat{D}(e^{i\theta}) \end{bmatrix} \\ & \quad - \begin{bmatrix} X(e^{i\theta}) & 0 \\ 0 & D \end{bmatrix} < 0 \\ & \quad X(e^{i\theta}) > 0 \\ & \quad D = \text{diag}(d_1, \dots, d_n) > 0 \end{aligned} \quad (5.45)$$

5.3.5 Numerical Example

As an example to calculate the SN sub-optimal scaling D , we consider the same model as in the previous example. The system satisfies

$$\|M(e^{i\theta})\|_{\mathcal{H}_\infty} = \left\| \left[\begin{array}{c|c} \hat{A}(e^{i\theta}) & \hat{B}(e^{i\theta}) \\ \hline \hat{C}(e^{i\theta}) & \hat{D}(e^{i\theta}) \end{array} \right] \right\|_{\mathcal{H}_\infty} = \gamma \approx 0.58 \text{ for all } \theta \in [0, 2\pi]$$

With a structured perturbation block of size two i.e., $\Delta = \text{diag}\{\Delta_1, \Delta_2\}$ we grid θ over the interval $[0 \ 2\pi]$, stack the resulting LMI's in (5.45) together with constraints for $X(e^{i\theta})$ and D to form a single LMI. We then check the feasibility of the resulting LMI over various values of $\hat{\gamma}$, where $\hat{\gamma}$ is the upper bound on $\|D^{-1}\hat{M}D\|_{\mathcal{H}_\infty}$. An optimal (almost) value of $\hat{\gamma} = 0.0015$ was found. This implies that the system can tolerate any structured spatiotemporal varying perturbation $\Delta = \text{diag}\{\Delta_1, \Delta_2\}$ with $\|\Delta_1\|_{l_2\text{-ind}} < 666$ and $\|\Delta_2\|_{l_2\text{-ind}} < 666$. Note that if the diagonal structure of Δ is ignored and we use the \mathcal{H}_∞ norm criterion for robustness, the size of allowable perturbations reduces dramatically to $1/\gamma = 1.72$ as opposed to the optimal (almost) value ($\hat{\gamma} = 0.0015$) achieved by the optimal (almost) scaling $D = \text{diag}\{4.38 \times 10^6, 1.034\}$.

5.4 Conclusion

In this chapter we have presented necessary and sufficient conditions for robust stability for LSTI stable systems with respect to l_∞ induced norm and l_2 induced norm when the underlying perturbations are LSTV and NLSTI stable (in the sense of corresponding induced norms). We have shown that the structured norm has the same value for two classes of perturbations 1) nonlinear spatiotemporal invariant perturbations 2) linear spatiotemporal varying perturbations. These conditions turn out to be analogous to the robustness conditions of standard LTI stable (l_∞ , and l_2) systems.

CHAPTER 6

CONCLUSIONS

In this dissertation, we considered the stability and performance aspects of gradually varying spatiotemporal systems. We showed that these systems can be stabilized with controllers designed based on frozen in time and space LSTI approximants of the original system provided that the spatiotemporal variations are slow enough. We emphasize here that the stability of the spatiotemporal varying systems was established in a distributed or modular way. The controllers were designed based on only the knowledge of local dynamics, measurements and control actions of the neighbors affecting a given subsystem. No knowledge about the dynamics of other subsystems (including the immediate neighbors) was assumed. It was also shown that the performance of the overall system cannot be much worse than that of worst case frozen in time and space performance. We remark here that although our result is developed for a setting that involves infinitely many subsystems, it can be easily extended/applied to a system with finitely many subsystems. In this case the underlying controllers can be designed based on the local frozen in time and space circulant approximants. The boundedness requirement of *ITSAE* will reduce to simply boundedness of the integral time absolute error (*ITAE*).

We then looked at the problem of distributed system identification of spatiotemporally invariant systems that performed better than the standard projection algorithm if run independently on each subsystem. The proposed algorithm guaranteed to bring the local estimates arbitrarily close to each other (in the absence of additive noise in the output), hence making the resulting estimated system a suitable candidate for the application of results on the gradually varying spatiotemporal systems. This resulted in an indirect adaptive control scheme based on certainty equivalence and we showed that it is stable. Having analyzed parametric uncertainty in spatiotemporally invariant systems, we turned our attention to consider nonparametric

perturbations of various sorts in spatiotemporally invariant systems and provided necessary and sufficient conditions for robust l_∞ and l_2 stability. Built on the results presented in previous chapters, there are several relevant extensions/future topics that can be researched into.

6.1 Optimality of Frozen Space Time Control

In Chapter 2 we looked at the relationship between the overall and frozen performance. A natural question that arises is if the optimal controller (over all spatiotemporal varying controllers) can be approximated by the frozen space-time control law that is optimum at each step in time and space. It seems intuitive to conjecture that the answer could be affirmative. For the standard case, Wang [91] argued that if the unstable zeros are not allowed to exit the unit disk then the frozen-time optimal \mathcal{H}_∞ design yields arbitrarily close (depending on the rate of variation) optimal $l_2 - l_2$ global behavior. Voulgaris and Dahleh presented a counter example in [75] showing that in the presence of discontinuities in the \mathcal{H}_∞ design, the optimal frozen-time design may not be optimal or near optimal regardless of the rate of variations. Establishing such properties for LSTV systems is an aspect that can be looked into for future research.

6.2 Excitation Conditions for Distributed Projection Algorithm

We saw in Chapter 3 that the distributed projection algorithm converged to the true parameter while the standard projection algorithm did not. It is already argued by Keum and Bentsman [88] that the concept of persistency of excitation, which guarantees the parameter error convergence to zero in finite-dimensional adaptive systems, in infinite-dimensional adaptive systems should be investigated in relation to time variable, spatial variable, and also boundary conditions. They also showed that unlike the finite-dimensional case, in infinite dimensional adaptive systems even a constant input can be persistently exciting in the sense that it guarantees the convergence of

parameter errors to zero. Further investigations would be of interest to ascertain if some generalizations can be made with respect to the structure of a given LSTI system about the excitations conditions necessary for system identification.

6.3 Robustness Analysis of LSTV Systems

In Chapter 2 we considered the stability of spatiotemporally varying systems by employing spatiotemporally varying controllers designed based on the frozen in time and space approximants of the original system. In Chapter 5 we have already provided necessary and sufficient conditions for robust stability of spatiotemporally invariant systems. It is only natural to carry out similar investigation for spatiotemporally varying systems. For the standard case, this has been carried out by Khammash in [89]. Generalization of this result is thought to be possible with some effort.

6.4 Robust Adaptive Control

Based on the results on stability and performance of spatiotemporally varying systems, a characterization of a class of indirect adaptive controllers can be worked out that can stabilize LSTI systems subjected to both parametric and $l_\infty - l_\infty$ bounded unstructured uncertainty. In particular, consider the setup shown in Figure 6.1. We are given the SISO LSTI discrete system

$$((A_0 - \Delta_A)y)(t) = ((B_0 + \Delta_B)u)(t) + d(t)$$

where A_0, B_0 are operators in \mathcal{L}_{STI} , with a polynomial z, λ transform representation given as follows;

$$\begin{aligned}\hat{A}_0(z, \lambda) &= 1 + \sum_{t=0}^{m_1} \sum_{k=-n_1, k \neq 0}^{n_1} (a_k^0(t) z^k) \lambda^t \\ \hat{B}_0(z, \lambda) &= \sum_{t=0}^{m_2} \sum_{k=-n_2, k \neq 0}^{n_2} (b_k^0(t) z^k) \lambda^t\end{aligned}$$

with the coefficients of $\{a_k^0(t)\}$ and $\{b_k^0(t)\}$ are not known a priori, Δ_A , and Δ_B are unknown, possibly spatiotemporally varying, operators in \mathcal{L}_{STV} , and

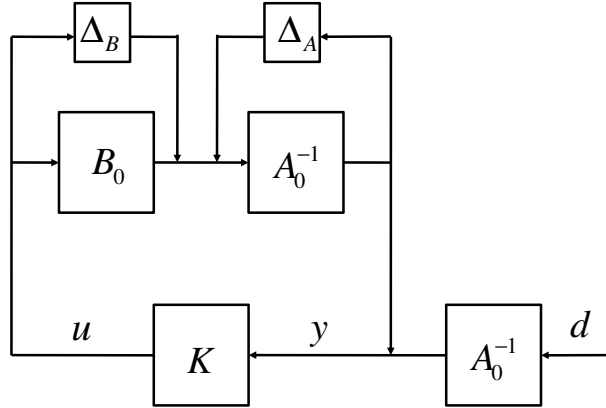


Figure 6.1 Setup for Robust Adaptive Control for LSTI Systems

d is a bounded disturbance, i.e. $\|d\| \leq D$, for some $D > 0$. It is assumed, however, that we have the knowledge of the upper bounds on the degrees of the polynomials $\hat{A}_0(z, \lambda)$, and $\hat{B}_0(z, \lambda)$. It can be further assumed that some bound D_Δ such that $\|(\Delta_A \Delta_B)\| < D_\Delta$ is known. With the given plant $\hat{P} = \frac{\hat{B}_0(z, \lambda)}{\hat{A}_0(z, \lambda)}$, the quest is to find a controller K that stabilizes the feedback loop shown in Figure 6.1. For the standard case this has been carried out by Voulgaris et al. in [90] employing the results of [86], and [75]. The generalization of this result for LSTI systems is expected to fall out with little effort.

6.5 l_1 Optimal Control Problem

The l_1 optimal control problem arises naturally within the context of adaptive control. The idea is to design the controller to minimize the effect of disturbances and error signal on the output. Intuitively, this means that the estimation algorithm is less affected by the disturbance, and hence produces more realistic estimates. Equivalently, minimizing the effect of the error on the output suggests that the graphs of the plant and the frozen-time estimates, restricted on the domain of all possible signals generated by the control scheme, are close. Since the bounds on the disturbances and error signals are l_∞ bounds, the l_1 methodology is most adequate in this set-

ting. Roughly, one can explain the procedure as follows: The input/output sequences are related through the true model by

$$Ay_i(t) = Bu_i(t)$$

Assume at time τ and at subsystem i , the available estimates of A and B are $A_{i,\tau}$, $B_{i,\tau}$. Using this estimate we can write the input/output relation as

$$A_{i,\tau}y_i(t) = B_{i,\tau}u_i(t) + e_i(t)$$

where $e_i(t)$ is an error signal given by

$$e_i(t) = (A_{i,\tau} - A)y_i(t) + (B - B_{i,\tau})u_i(t)$$

The control law $u_i(t)$ is given by

$$u_i(t) = K_{i,\tau}(-y_i(t) + r_i(t))$$

where $K_{i,\tau}$ is constructed to stabilize the plant $P_{i,\tau} = \frac{B_{i,\tau}}{A_{i,\tau}}$ and $r_i(t)$ is a bounded reference signal. With this control law the output of the plant is given by

$$y_i(t) = P_{i,\tau}(1 + P_{i,\tau}K_{i,\tau})^{-1}r_i(t) + \frac{1}{A_{i,\tau}}(1 + P_{i,\tau}K_{i,\tau})^{-1}e_i(t)$$

The first goal is to find a controller $K_{i,\tau}$, that internally stabilizes $P_{i,\tau}$. Internal stability means that every element of the following matrix is a stable transfer function

$$\hat{H}(P_{i,\tau}, K_{i,\tau}) = \begin{bmatrix} \frac{1}{1 + \hat{P}_{i,\tau}\hat{K}_{i,\tau}} & \frac{\hat{K}_{i,\tau}}{1 + \hat{P}_{i,\tau}\hat{K}_{i,\tau}} \\ \frac{\hat{P}_{i,\tau}}{1 + \hat{P}_{i,\tau}\hat{K}_{i,\tau}} & \frac{1}{1 + \hat{P}_{i,\tau}\hat{K}_{i,\tau}} \end{bmatrix}$$

The family of all compensators that stabilize $P_{i,\tau}$, denoted by $S(P_{i,\tau})$, is appropriately parameterized via the YJB parametrization [93]. If we think of the error signal $e_i(t)$ as a bounded l_∞ sequence, then it is reasonable to choose a compensator $K_{i,\tau}$ from $S(P_{i,\tau})$ that minimizes the induced operator

norm. This is exactly an l_1 optimal control problem defined as follows

$$\inf_{K_{i,\tau} \in S(P_{i,\tau})} \left\| \frac{1}{A_{i,\tau}} (1 + P_{i,\tau} K_{i,\tau})^{-1} \right\|_1$$

where

$$S(P_{i,\tau}) = \{K_{i,\tau} \mid K_{i,\tau} \text{ internally stabilizes } P_{i,\tau}\}$$

The idea is that at each stage the controller is chosen for its ability to reject the modeling error, as well as for its stabilizing properties.

For the standard case, when the unstable zeros are strictly within the unit disc, the l_1 optimal control problem reduces to a finite dimensional linear program. Also, the closed loop is known to be FIR in this case. No such result exists at this time for spatiotemporal systems. Over the last two years or so we have managed to establish [92] that the optimal l_1 closed loop for LSTI systems is not FIR in general (along the spatial domain in particular) if the temporal zeros are restricted to be strictly inside the unit disc (spatial zeros lie on the boundary in this case). Generalization of l_1 optimal control problem for LSTI systems along the lines of the standard case is not straight forward. The unstable zeros no longer remain discrete as they constitute a region of continuum. For a thorough treatment on discrete multidimensional systems see [94], and [95]. How these zeros may affect the l_1 optimal control design and in turn performance as the estimates vary in space as well as in time is an issue that has not been researched into.

Parallel to its counterpart for the standard case, there are various aspects within the quest for the solution of l_1 optimal control problem for LSTI systems that are of critical importance. We list a couple of them in the following:

1. can the l_1 optimal control problem be posed in to a dual space that is helpful in characterizing the closed loop, or the solution to the original problem ?
2. can the l_1 optimal control problem be cast into a finite dimensional optimization scheme (e.g. a linear program), resulting in an FIR closed loop along the temporal domain ?

We remark here that while we have looked at systems that have infinite spatial extent, this is never the case in reality. The infinite spatial dimensionality lends itself to more elegant, unified, and easier mathematical treatment. For the real systems, boundary effects are always present and do need special considerations. Presently the boundary effects are treated in an ad hoc fashion, see e.g. [6], which lack rigor and may even lead to instability. More elaborated approaches are required to incorporate the boundary effects in a unified manner to guarantee system stability and performance. In this direction Dullerud and D'Andrea [96] have worked on the control design for distributed systems where they do not require the underlying system dynamics to be spatiotemporally invariant and when quadratic criteria are of interest.

We end this thesis by making the note that the area of spatiotemporal systems is still open and full of challenges with respect to control system synthesis and analysis and requires more effort from the research community to provide answers to various pending questions.

REFERENCES

- [1] G. B. Shaw, D. W. Miller, and D. E. Hastings, “The generalized information network analysis methodology for distributed satellite systems”, *Ph.D. dissertation, Mass. Inst. Technol.*, Cambridge, MA, 1998.
- [2] G. E. Stewart, “Two-dimensional loop shaping controller design for paper machine cross-directional processes”, *Ph.D. dissertation, Univ. British Columbia*, Vancouver, BC, Canada, 2000.
- [3] J. D. Wolfe, D. F. Chichka, and J. L. Speyer, “Decentralized controllers for unmaned aerial vehicle formation flight”, *Amer. Inst. Aeronautics Astronautics*, 1996.
- [4] D. F. Chichka and J. L. Speyer, “Solar-powered, formation-enhanced aerial vehicle system for sustained endurance”, in *Proc. Amer. Control Conf.*, 1998, pp. 684688.
- [5] H. Raza and P. Ioannou, “Vehicle following control design for automated highway systems”, *IEEE Control System Magazine.*, vol. 16, pp. 4360, June 1996.
- [6] A. Sarwar, P.G. Voulgaris and S.M. Salapaka. “Modeling and Distributed Control of Electrostatically Actuated Microcantilever Arrays”, *In Proc. American Control Conference 2007.* (New York)
- [7] C.S. Kubrusly, “Distributed parameter system identification: A survey”, *International Journal of Control*, vol. 26, no. 4, pp. 509–535, Oct. 1977
- [8] M.P. Polis, R.E. Goodson, “Parameter identification in distributed systems: A synthesizing overview”, *Proceedings of the IEEE*, Jan. Vol. 64, Issue: 1, 1976, pp. 45- 61.
- [9] Graham C. Goodwin and Kwai Sang Sin, “Adaptive Filtering Prediction and Control”, *Prentice Hall*, 1984
- [10] N.K. Sinha and B. Kuszta, “Modeling and identification of dynamic systems”, *New York : Van Nostrand Reinhold*, 1983
- [11] L. Ljung, “System Identification: Theory for the User”, *Prentice Hall PTR*, December 1998

- [12] Barooah, P., P. G. Mehta and J. Hespanha, "Mistuning-based Control Design to Improve Closed-loop Stability of Vehicular Platoons", *IEEE Transactions on Automatic Control* 54:9, 2100-2113, September 2009
- [13] Mehta, P. G., G. Hagen, and A. Banaszuk, "Symmetry and Symmetry-breaking for a Wave Equation with Feedback", *SIAM Journal of Applied Dynamical Systems*, 6:3, 549-575, July 20
- [14] M. E. Taylor, "Partial Differential Equations I: Basic Theory", *New York: Springer-Verlag*, 1996
- [15] R. Radner, "Team decision problems", *The Annals of Mathematical Statistics*, Vol. 33, No. 3, pp. 857-991, 1962.
- [16] Bamieh, B.F. Paganini and M. A. Dahleh, "Distributed Control of Spatially Invariant Systems", *IEEE Transactions on Automatic Control*, Vol. 47, No. 7, pp.1091-1107. 2002.
- [17] Raffaello D'Andrea and Geir E. Dullerud, "Distributed Control Design for Spatially Interconnected Systems", *IEEE Transaction on Automatic Control*, Vol. 48, No. 9, 2003.
- [18] P. K. C. Wang, "Stabilization and control of distributed systems with time-dependent spatial domains", *Journal of Optimization Theory and Applications*, Volume 65, Number 2, pp. 331-362 May, 1990
- [19] Zolesio, J. P., "Shape Stabilization of Flexible Structure, Distributed Parameter Systems, Edited by F. Kappel, K. Kunisch, and W. Schappacher", *Springer- Verlag*, Berlin, Germany, pp. 446-460, 1985
- [20] Dubljevic, S., P. D. Christofides and I. G. Kevrekidis, "Distributed Non-linear Control of Diffusion-Reaction Processes", in *Proceedings of American Control Conference*, pp. 1341-1348, Denver, Colorado, 2003.
- [21] Stolz, G. Jr. "Numerical Solution to Inverse Problems of Heat conduction for Simple Shapes", *Trans. ASME J. Heat Transfer*, Vol 82-C, No.1, pp. 20-26, Feb 1960
- [22] Frank J. "An Application of Least Squares Method to the Solution of the Inverse Problem of Heat Conduction", *Trans. ASME J. Heat Transfer* Vol. 85-C, No. 4, pp. 378-379, Nov 1963.
- [23] Burggraf, O. R. "An Exact Solution to Inverse Problem in Heat Conduction Theory and Applications", *Trans. ASME J. Heat Transfer* Vol. 86-C, No. 3, pp. 373-382, Aug 1964.

- [24] Beck, J.V. “ The Optimum Analytical Design of Transient Experiments for simultaneous Determinations of Thermal Conductivity and Specific Heat”, Ph.D Thesis, Dept. of Mech. Eng., Michigan State University, 1964.
- [25] Beck, J.V. “ Analytical Determination of Optimim Transient Experiments for the Measurement of Thermal Properties”, *3rd Int. Heat transfer Conf.*, Vol. IV, pp. 301-329, 1966.
- [26] H.T. Banks, K. Kunish, “Estimation techniques for distributed parameter systems”, *Systems & Control: Foundation & Applications*, Vol. 1, Note(s): XIII-315, 1989.
- [27] R. J. Aumann. “Agreeing to disagree”, *Annals of Statistics.*, vol. 4, no. 6, pp. 1236-1239, 1976.
- [28] J. D. Geanakoplos and H. M. Polemarchakis. “We can’t disagree Forever”, *Inst. Math. Studies in Social Sci., Stanford Univ.*, Stanford, CA. Tech. Rep. 277, 1978.
- [29] Borkar, V. and P. Varaiya, “Asymptotic Agreement in Distributed Estimation”, *IEEE Trans. on Automatic Control*, vol. AC-27, No. 3, 1982, pp. 650-655.
- [30] Willsky, A.S., M. Bello, D.A. Castanon, B.C. Levy and G. Verghese, “Combining and Updating of Local Estimates and Regional Maps along Sets of One-Dimensional Tracks”, *IEEE Transaction on Automatic Control*, Vol. AC-27, No. 4, 1982, pp. 799-813.
- [31] Dimitri P. Bertsekas, John N. Tsitsiklis and Michael Athans “Convergence theories of distributed iterative process : a survey”, *Lecture Notes in Control and Information Sciences*, Springer Berlin/Heidelberg, pages, 107-139, 1985
- [32] Wei Ren, Randal W. Beard, and Ella M. Atkins, “A Survey of Consensus Problems in Multi-agent Coordination”, *In Proc. American Control Conference* 2005, June 8-10, 2005. Portland, OR, USA
- [33] Michael G. Rabbat and Robert D. Nowak “Quantized Incremental Algorithms for Distributed Optimization”, *IEEE Journal on Selected Areas in Communications*, Vol 23, No. 4, April 2005
- [34] Leilei Li, and Jonathon A. Chambers “A new incremental affine projection-based adaptive algorithm for distributed networks”, *Signal Processing*, 88 (2008) 2599-2603
- [35] P. Ioannou and P. Kokotovic “ Decentralized adaptive control of interconnected systems with reduced-order models”, *Automatica*, Vol 21(4) pp. 401-412, 1985.

- [36] P.A. Ioannou. “Decentralized adaptive control of interconnected systems”, *IEEE Transactions on Automatic Control*, Vol 31 pp. 291298, 1986.
- [37] K.A. Ossman “Indirect adaptive control for interconnected systems”, *IEEE Transactions on Automatic Control*, Vol 34(8) pp. 908911, 1989.
- [38] D.T. Gavel and D.D. Siljak. “Decentralized adaptive control: structural conditions for stability”, *IEEE Transactions on Automatic Control*, Vol 34(4) pp. 413426, 1989.
- [39] S. Huang, K.K. Tan, and T.H. Lee “Decentralized control design for large-scale systems with strong interconnections using neural networks”, *IEEE Transactions on Automatic Control*, Vol 48(5) pp. 805810, 2003.
- [40] G.,E.Dullerud and F.Paganini, “A course in robust control theory: a convex approach”, *Springer-Verlag* New York, 2000.
- [41] Azeem Sarwar, Petros G. Voulgaris, and Srinivasa M. Salapaka. “Stability Of Slowly Varying Spatiotemporal Systems”, *In Proc. IEEE Conference on Decision and Control*, Cancun, Mexico, Dec, 2008
- [42] John C. Doyle, “Analysis of feedback systems with structured uncertainty”, *IEEE Proceedings* 129, 1982
- [43] L. Shi and SK Singh “Decentralized adaptive controller design for large-scale systems with higher order interconnections”, *IEEE Transactions on Automatic Control*, Vol 37(8) pp. :11061118, 1992.
- [44] J.T. Spooner and K.M. Passino “Adaptive control of a class of decentralized nonlinear systems”, *IEEE Transactions on Automatic Control*, Vol 41(2) pp. 280284, 1996
- [45] J.T. Spooner and K.M.Passino “Decentralized adaptive control of nonlinear systems using radialbasis neural networks”, *IEEE Transactions on Automatic Control*, Vol 44(11) pp. 20502057, 1999.
- [46] K.S. Narendra and N.O. Oleng “Exact output tracking in decentralized adaptive control systems”, *IEEE Transactions on Automatic Control*, Vol 47(2) pp. 390395, 2002.
- [47] B.-J. Yang N. Hovakimyan, E. Lavretsky and A.J. Calise “Coordinated decentralized adaptive output feedback control of interconnected systems”, *IEEE Transactions on Neural Networks*, Vol 16 pp. 10459227, 2005.
- [48] B.M. Mirkin and P.O. Gutman “Decentralized Output-Feedback MRAC of Linear State Delay Systems”, *IEEE Transactions on Automatic Control*, Vol 48(9) pp. :1613, 2003.

- [49] I.W. Sandberg. "An observation concerning the application of the contraction mapping fixed-point theorem and a result concerning the norm boundedness of solutions of nonlinear functional equations", *Bell Systems Technical Journal* Vol 44 pp. 809-1812, 1965
- [50] G.Zames "On the input-output stability of time varying nonlinear feedback system. Part I: Conditions derived using concepts of loop gain, conicity, and positivity", *IEEE Transactions on Automatic Control* Vol 11. pp. 228-238 1966
- [51] V.M. Popov "On the absolute stability of nonlinear systems of automatic control", *Automation and Remote Control* Vol 22. pp. 857-875 1961
- [52] V.A. Yakubovich "Frequency conditions for the absolute stability of control systems with several nonlinear or linear nonstationary units", *Automat. Telemekh* pp. 5-30 1967
- [53] M.G.Safonov "Stability and Robustness of Multivariable Feedback Systems", *MIT Press* 1980
- [54] G.Zames "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses", *IEEE Transactions on Automatic Control* Vol 26. pp. 301-320 1981
- [55] M.G.Safonov "Tight bounds on the response of multivariable systems with component uncertainty", in *Proc. Allerton Conference* 1978
- [56] M.Fan, A. Tits, and J.C. Doyle "Robustness in the presence of mixed parametric uncertainty and unmodled dynamics", *IEEE Transactions on Automatic Control* Vol 36. pp. 25-38, 1991
- [57] C.L. Beck and J.C. Doyle "A necessary and sufficient minimality condition for uncertain systems", *IEEE Transactions on Automatic Control* Vol 44. pp. 1802-1813, 1999
- [58] A. Packard and J.C. Doyle "The complex structured singular value", *Automatica* Vol 29. pp. 71-109, 1993
- [59] P.M. Young "Robustness with parametric and dynamic uncertainty", *PhD Thesis* California Institute of Technology 1993
- [60] R. Braatz, P.M. Young, J. C. Doyle, and M. Morari "Computation complexity of μ calculation", *IEEE Transactions on Automatic Control* Vol 39. pp. 1000-1002, 1994
- [61] O. Toker and H. Ozbay "On the NP-hardness of the purely complex μ computation, analysis/synthesis, and some related problems in multidimensional systems", in *Proc. ACC* 1995

- [62] S. Treil “The gap between the complex structured singular value and its upper bound is infinite”, *Available online at Brown University, www.math.brown.edu/faculty/treil.html*, 1999.
- [63] M.Khammash and J.B. Pearson. “Performance Robustness of Discrete-Time Systems with Structured Uncertainty”, *IEEE Transactions on Automatic Control*. (Vol. 36, No. 4, April 1991.)
- [64] J. S. Shamma. “Robust Stability with Time-Varying Structured Uncertainty”, *IEEE Transactions on Automatic Control*. (Vol. 39, No. 4, April 1994.)
- [65] Jeff S. Shamma “A Connection between Structured Uncertainty and Decentralized Control of Spatially Invariant Systems”, *in Proceedings of the ACC* Arlington, VA June 25-27, 2001
- [66] G. A. de Castro and F. Paganini, “An IQC characterization of decentralized Lyapunov functions in spatially invariant arrays”, *in Proceedings of 41st CDC.*, 2002 pp. Vegas, NV
- [67] Dimitry Gorinevsky and Gunter Stein. “Structured Uncertainty Analysis of Robust Stability for Multidimensional Array System”, *IEEE Transaction on Automatic Control*. Vol. 48, pp. 1557-1568. 2003.
- [68] M.Khamash and J.B. Pearson. “Performance Robustness of Discrete-Time Systems with Structured Uncertainty”, *IEEE Transactions on Automatic Control*. (Vol. 36, No. 4, April 1991.)
- [69] J. S. Shamma. “Robust Stability with Time-Varying Structured Uncertainty”, *IEEE Transactions on Automatic Control*. (Vol. 39, No. 4, April 1994.)
- [70] M.A. Dahleh, I.J. Diaz-Bobillo. “Control of Uncertain Systems: A Linear Programming Approach”, *Prentice-Hall, Inc.* (1995.)
- [71] R. Chandra and R. DAndrea. “A Scaled Small Gain Theorem with Applications to Spatially Interconnected Systems”, *IEEE Transactions on Automatic Control*. (Volume 51, Issue 3, March 2006 Pages:465 - 469)
- [72] Martha A. Gallivan, David G. Goodwin, Richard M. Murray . “Modeling and Control of Thin Film Morphology Using Unsteady Processing Parameters: Problem Formulation and Initial Results”, *in Proceedings of CDC.*, 2001 Orlando, Florida
- [73] Ni, D. and P. D. Christofides, ”Multivariable Predictive Control of Thin Film Deposition Using a Stochastic PDE Model”, *Ind. & Eng. Chem. Res.*, Vol. 44, pp: 2416-2427, 2005.

- [74] M. Dahleh and M. A. Dahleh, "On slowly time-varying systems", *Automatica* 27 (1991) 201-205.
- [75] P. G. Voulgaris and M. A. Dahleh, "On l^∞ to l^∞ performance of slowly time-varying systems", *Systems and Control Letters* 24 (1995) 243-249.
- [76] G. Zames and L.Y. Wang, "Local-global double algebras for slow H-infinity adaptation, Part I: Inversion and stability", *IEEE Trans. Automatic Control* Vol. 36, No. 2, pp.130-142, 1991
- [77] N.K.Bose, "Applied Multidimensional Systems Thoery", Van Nostrand Reinhold Company, 1982
- [78] Hoffman, K., "Banach Spaces of Analytic functions", Prentice-Hall, Englewood Cliffs, NJ, 1962
- [79] S. Jiang , P. G. Voulgaris, L.E. Holloway, L.A. Thompson "Distributed control of large segmented telescopes", In Proc. ACC, 2006
- [80] C. G. Lopes, A. H. Syed "Incremental Adaptive Strategies Over Distributed Networks", IEEE Transactions on Signal Processing, Vol 55, No. 8, Aug 2007
- [81] C. G. Lopes, A. H. Syed "Diffusion Least-Mean Squares Over Adaptive Networks", In Proc. IEEE ICASSP, Hawaii, 2007,
- [82] L. Xiao, S. Boyd, S.Lall "A Space-Time Diffusion Scheme for Peer-to-Peer Least Squares Estimation", In Proc. The Fifth International Conference on Information Processing in Sensor Networks, 2006
- [83] S. S. Ram, A Nedić, V. V. Veeravalli "Distributed Subgradient Projection Algorithm for Convex Optimization", In Proc. ICASSP, 2009
- [84] P. Massioni, M. Verhaegen "Subspace Identification of Circulant Systems", *Automatica* 2008, Vol 44, 2825-2833
- [85] B. Golub, M. O. Jackson "A Naïve Learning in Social Networks and the Wisdom of Crowds", To Appear *American Economic Journal Microeconomics*
- [86] M. Dahleh, M.A. Dahleh "A Class of Adaptive Controllers with Application to Robust Adaptive Control", *LIDS, MIT Report: LIDS-*; 1759, 1988
- [87] G.C. Goodwin, P.J. Ramadge, and P.E. Caines, "Discrete-time multi-variable adaptive control", *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 3, pp. 449-456.

- [88] Keum Shik Hong, IEEE, and Joseph Bentsman “Direct Adaptive Control of Parabolic Systems: Algorithm Synthesis and Convergence and Stability Analysis”, *IEEE Trans. Automat. Contr.* Vol. 39. No. 10, Oct 1994
- [89] Mustafa H. Khammash, “Necessary and Sufficient Conditions for the Robustness of Time-Varying Systems with Applications to Sampled-Data Systems”, *IEEE Trans. Automat. Contr.*, Vol. 38, No. 1, Jan 1993
- [90] Voulgaris, Petros G. and Dahleh, Munther A. and Valavani, Lena S., “Robust adaptive control: a slowly varying systems approach”, *Automatica*, Vol 30, No. 9, 1994, 1455-1461
- [91] L. Y. Wang, “How conservative is frozen-time interpolation?”, *Systems & Control Letters*, Vol. 18, pp. 277-283, 1992.
- [92] Azeem Sarwar and Petros G. Voulgaris “The two dimension l_1 model matching problem: optimal closed loop is not FIR”, *Unpublished Note*, 2009.
- [93] M. Vidyasagar, “Control System Synthesis: A Factorization approach”, *The MIT Press*, 1985.
- [94] D. E. Dudgeon, “Two Dimensional Recursive Filtering”, *D.Sc. Dissertation*, M.I.T Cambridge, Mass. May 1974
- [95] N.K.Bose, “Applied Multidimensional Systems Theory”, *Van Nostrand Reinhold Company*, 1982
- [96] Geir E. Dullerud and Raffaello D’Andrea, “Distributed Control of Heterogeneous Systems”, *IEEE Transaction on Automatic Control*, Vol. 49, No. 12, pp. 2113-2128, 2004.

APPENDIX A

DETAILS FOR NUMERICAL EXAMPLE GIVEN IN CHAPTER 5

For clarity of exposition, we have only considered parametric uncertainty on the lumped parameters of the plant model given in [6]. A time step of 0.01 msec was used to obtain a discrete time approximation of the plant model. The closed loop $M = \{A^M, B^M, C^M, D^M\}$ is a stable LSTI system. To ease the calculations for the example at hand, we considered only an immediate neighbor interaction model for M . Although limited for the purpose of the example, this is in fact a very good approximation of M as rest of the interactions are very small. The corresponding $\{A^M, B^M, C^M, D^M\}$ matrices are given below.

$$\begin{aligned}
 A_0^M &= \begin{pmatrix} 0.99 & 1 \times 10^{-5} & 0 & 0 & 0 & 0 & 0 & 0 \\ -1.14 & 0.99 & 3.04 \times 10^{-5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.99 & -6.40 & -0.01 & -1.41 \times 10^{-5} & 3235.89 & 0 \\ 0 & 0 & 0 & 0.99 & 1.00 \times 10^{-5} & 0 & 0 & 0 \\ 0 & 0 & 0.82 & -1.14 & 0.99 & -3.20 \times 10^{-4} & 0 & 0 \\ 0 & 0 & 0.02 & -15005.07 & -17.02 & 0.96 & 7.58 \times 10^6 & 0 \\ 0 & 0 & 0 & -1 \times 10^{-5} & 0 & 0 & 0.99 & 0 \\ 0 & 1.20 \times 10^{-7} & 0 & 0 & -5.13 \times 10^{-11} & 0 & 1.59 \times 10^{-14} & 0.99 \end{pmatrix} \\
 A_1^M &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.1 & -1.4 \times 10^{-15} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.1 & -2.9 \times 10^{-5} & -2.7 \times 10^{-8} & 0.02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2.3 \times 10^{-15} & 0.1 & -1.4 \times 10^{-15} & 0 & 0 & 0 \\ 0 & 0 & 0 & -133.7 & -0.1 & -6.2 \times 10^{-5} & 49.0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1.4 \times 10^{-15} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -9.0 \times 10^{-15} & 0 \end{pmatrix} \\
 B_0^M &= \begin{pmatrix} 0 & 0 \\ -0.01 & -4.10 \times 10^{-8} \\ 0_{8 \times 1} & 0_{8 \times 1} \end{pmatrix} & B_1^M &= \begin{pmatrix} 0 & 0 \\ 0.00124 & 4.30 \times 10^{-23} \\ 0_{8 \times 1} & 0_{8 \times 1} \end{pmatrix} \\
 C_0^M &= \begin{pmatrix} 1.00 & -3.33 \times 10^{-18} & 0_{1 \times 8} \\ 0 & 1.00 & 0_{1 \times 8} \end{pmatrix} & C_1^M &= \begin{pmatrix} -7.22 \times 10^{-16} & 1.70 \times 10^{-18} & 0_{1 \times 8} \\ 0 & 1.00 & 0_{1 \times 8} \end{pmatrix}
 \end{aligned}$$

Since there is spatial symmetry in the system, we have $A_1^M = A_{-1}^M$. Also

$B_1^M = B_{-1}^M$, and $C_1^M = C_{-1}^M$. The D term is simply zero.

AUTHOR'S BIOGRAPHY

Azeem Sarwar was born on September 10, 1980 in Pakistan. In November 2000, Azeem graduated from National University of Sciences and Technology (NUST), Pakistan with his bachelor's degree in Mechanical Engineering with highest distinction. He then went on to join the energy sector in Pakistan working for Hagler Bailly Pakistan (associated with P.A. Consulting UK) for nearly two and a half years. He then joined Orient Petroleum Incorporated, an oil and gas exploration and production company, as a *senior pipeline engineer* where he worked for about a year. In the Fall of 2003, he enrolled in the Department of Mechanical Science and Engineering at the University of Illinois at Urbana-Champaign. In the summer of 2004, he started working on modeling, control design, and robustness analysis of microcantilever arrays under the guidance of Professor Petros G. Voulgaris, and received a Master of Science degree in October 2006.

He continued his Ph.D study in Mechanical Engineering with Professor Petros G. Voulgaris. He worked on spatiotemporal systems with a focus on gradual variations, identification, adaptation, and robustness. While working on his Ph.D he received a Masters degree in Mathematics from University of Illinois at Urbana-Champaign in August 2008. During his Ph.D study he spent the summer of 2007 at the Innovation Center of EATON Corporation in Minneapolis as a *summer intern* where he worked on flow force analysis, based on FEA, of one of EATON's industrial proportional valves. In the summer of 2008, he again joined the Innovation Center of EATON Corporation in Minneapolis as *EATON's Engineering and Technology Excellence Intern* where he worked on developing and testing control algorithms on a prototype of a high speed proportional valve. His Masters and Ph.D research works have resulted in several journal publications. His current research interests include exploring the interplay of control theory, and micro/nano/biological systems.