Determination of Design of Optimal Actuator Location Based on Control Energy

by

Afrouz Farshad Mehr

Thesis submitted for the degree of
Doctor of Philosophy
in
Control Engineering

January 2018

Systems and Control Research Centre
School of Engineering and Mathematical Sciences
City University of London, Northampton Square, London, EC1V 0HB

I declare that this thesis was composed by myself and that this work has not been submitted for any other degree or professional qualification.

Abstract

The thesis deals with the selection of the sets of inputs and outputs using the energy properties of the controllability and observability of a system and aims to define input and output structures which require minimization of the energy for control and state reconstruction. Such a study explores the energy dimension of the properties of controllability and observability, develops computations for the controllability and observability Gramians for stable and unstable systems and examines measures of the degree of controllability and observability properties using SVD (Singular Value Decomposition) of Gramians to compute the maximal and minimal energy requirements. These characterize the relative degree of controllability and observability under conditions where the available energy is constrained. The notion of energy surfaces in the state space is introduced and this enables the characterization of restricted notions of controllability and observability when the available energy is bounded. The maximal and minimal energy requirements for different input vectors is demonstrated and this provides the basis for the development of strategies and methodologies for selection of systems of inputs and outputs to minimize the energy required for control, respectively state reconstruction. These results enable the development of input, output structure selection methodology using a novel optimization method. This thesis contributes in the further development of the area of systems, or global instrumentation, developed so far based on the assignment of structural characteristics by incorporating the role of energy requirements. The research provides energy based tools for the selection of input and outputs schemes with a main criterion the minimization of the energy required for control and observation and thus provide an alternative approach based on quantitative system properties in characterizing control and state observation as functions of given sets of inputs and output sets. The methodologies developed may be used as design tools where apart from energy requirements other design criteria may be also incorporated for the selection of inputs and outputs. The methodology that is used is based on linear systems theory and tools from numerical linear algebra. The solution to the problems considered here is an integral part of the effort to develop an integrated approach to control and global process instrumentation.

Acknowledgements

I would like to express my sincere gratitude to my advisor Professor Nicos Karcanias for his guidance, advice, help, support, and patience throughout this work. Prof Karcanias has left a deep impression on me, because of his professionalism, pursuit of excellence and dedication to all his work over the years. What I learned from him will always remain with me to help me throughout my professional life. Without his support, this thesis would never have emerged. I am very thankful to him for making this journey possible.

I would like to gratefully acknowledge the enthusiastic help and support of professor George Halikias during this work.

Sincere thanks go to the members of my committee for their patience and support in overcoming numerous obstacles I have been facing through my research.

I am also grateful to the following university staffs: Dr. Youyou Yan, Mrs. Nathalie Chatelain and Dr. Weiping Wu for their helps.

My deepest gratitude and much love goes to my Maman and Baba, Afsaneh and Ebrahim, my brothers, Omid and Arash and my beloved sister Afshan. You have always given me more support and love than I could have asked for, while I was thousands of miles away when you were going through difficult times. I am indebted, forever grateful, and so lucky to have you all!

Finally, there are my friends. We were not only able to support each other by deliberating over our problems and findings, but also happily by talking about things other than just our researches.

To my beloved parents Afsaneh Kiani and Ebrahim Farshadmehr

خدای را بسی شاکرم که از روی کرم پدر و مادری فداکار نصیبم ساخته تا در سایه درخت پر بار وجودشان بیاسایم و از ریشه آنها شاخ و برگ گیرم و از سایه وجودشان در راه کسب علم و دانش تلاش نمایم.

والدینی که بودنشان تاج افتخاری است بر سرم و نامشان دلیلی است بر بودنم چرا که این دو وجود پس از پروردگار مایه هستی ام بوده اند دستم را گرفتند و راه رفتن را در این وادی زندگی پر از فراز و نشیب آموختند.

تقدیم به پدر و مادر عزیز و مهربانم
که در سختی‌ها و دشواری های زندگی همواره یاوری فداکار
و پشتیبانی محکم و مطمئن برایم بوده‌اند.

Table of Contents

List of Tables..... ix

List of Figures..... x

Chapter 1: Introduction.....1

Chapter 2: Literature Review

2.1 Introduction.....6
2.2 Input/Output Structure Selection.....6
2.3 Problem Statement:.....9
2.4 Different Cost Functions:.....10
2.5 Summary.....24

Chapter 3: Overview of Controllability Gramian and its Computation

3.1 Introduction.....25
3.2 Reachability and Reachability Gramian.....26
3.3 Controllability Gramian and its Computation.....28
3.3.1 A Novel Approach for Calculation of Controllability Gramian.....30
3.3.2 Controllability Gramian Assignment with Respect to Control Variables.....35
3.4 Summary.....40

Chapter 4: Minimum Input Energy

4.1 Introduction.....41
4.2 Minimum Input Energy.....42
4.3 Influence of Different Terms on the Value of Minimum Input Energy.....43
4.3.1 Terminal Time t_f.....44
4.3.1.1 Stable System..... 45
4.3.1.2 Unstable System..... 49

4.3.2 Terminal State x_f 51
4.3.2.1 Energy Levels in the Best Case 53
4.3.2.2 Energy Levels in the Worst Case 56
4.4 Summary 60

Chapter 5: Controllability, Energy for Unstable Systems and Measures of Degree of Controllability

5.1 Introduction 61
5.2 Minimum Energy for Unstable Systems 62
5.2.1 Decomposition of Controllability Gramian in LTI Systems with Stable and Anti-Stable Modes 63
5.2.2 Minimum Input Energy for LTI System with Anti-Stable Part 68
5.3 Energy Levels for LTI Systems with Anti-Stable Modes 72
5.4 The Energy and Relative Degree of Controllability 80
5.5 Degree of Controllability 87
5.6 Degree of Disturbance Rejection 91
5.7 Summary 100

Chapter 6: Controllability Gramian and Energy Calculations of the Stable LTI System in Canonical Forms

6.1 Introduction 101
6.2 Controllability Gramian Computation in Controller Canonical Form 102
6.3 Inverse of Controllability Gramian and Energy Calculation in Controller Canonical Form 116
6.4 Controllability Gramian Computation in Diagonal Canonical Form 128
6.5 Summary 131

Chapter 7: Optimal Selection of Input Structure, Minimizing the Average of Input Energy

7.1 Introduction 132
7.2 Optimal Selection of Input to Minimize the Average of Minimum Input Energy ..133
7.2.1 Single Input Matrix Structure Selection Problem 135
7.2.1.a Optimal Input Vector Selection Considering the Finite Terminal Time 136

7.2.1.b Optimal Input Vector Selection for the Case of Stable System 139
7.2.1.c Optimal Input Vector in The Case That A Is a Normal Matrix 141
7.2.1.d Optimal Input Vector in The Case That A Is Stable and Normal 143
7.2.2 Multi-Input Matrix Structure Selection 145
7.2.2.a (Part I) Input Matrix Selection Within a Finite Terminal Time 146
7.2.2.a (Part II) Input Matrix Selection Within a Finite Terminal Time 148
7.2.2.a (Part III) Input Matrix Selection Within a Finite Terminal Time 149
7.2.2.b Optimal Multi-Input Selection in a Stable System 162
7.2.2.c Optimal Input Matrix Selection with Condition On 2Norm of Input Matrix B 163
7.3 Summary:.....166

Chapter 8: Kamineh Algorithm Optimization

8.1 Introduction.....168
8.2 Kamineh Optimization Algorithm (KA).....169
8.2.1 Inspiration.....169
8.2.2 Mathematical Model and Algorithm.....175
8.3 Experimental Results and Discussion.....181
8.3.1 Benchmark Functions 182
8.3.1.1 Unimodal Test Functions..... 184
8.3.1.2 Multimodal Test Functions 190
8.3.2 Analysis of KA Optimizer.....202
8.3.3 Optimization Problems Using the KA Algorithm 203
8.4 Summary:.....230

Chapter 9: Conclusions & Future Research

9.1 Summary and Contribution of Thesis.....232
9.2 Future Directions 234

References..... 236
Appendix.....261

List of Tables:

Table (5.1): Results of Example (5.8) 91
Table (5.2): Distance to uncontrollability for AMB system in Example (5.4) 91
Table (5.3): Distance to uncontrollability for seesaw-inverted pendulum system in Example (5.5) 91
Table (5. 4): DODR results of Example (5.10) 98
Table (5. 5): DODR results of Example (5.11) 99
Table (7. 1): optimum input vector in single input cases, which minimizes the average of $E_{\min}$ 166
Table (7.2): optimum input matrix in multi- input cases, which minimizes the average of $E_{\min}$ 167
Table (8.1): Unimodal test functions 183
Table (8.2): Multimodal test functions with variable dimensions 183
Table (8.3): Fixed-dimension multimodal test functions 184
Table (8.4): Results of unimodal test functions 186
Table (8.5): P-values produced by Wilcoxon's test comparing KA versus GA, PSO, FA, GSA and CAB over all runs for unimodal test functions 186
Table (8.6): Results of high-dimensional multimodal test functions 191
Table (8.7): Results of fixed-dimensional multimodal test functions 192
Table (8.8): P-values produced by Wilcoxon's test comparing KA versus GA, PSO, FA, GSA and CAB over all runs for high-dimensional multimodal test functions 193
Table (8.9): P-values produced by Wilcoxon's test comparing KA versus GA, PSO, FA, GSA and CAB over all runs for fixed-dimensional multimodal test functions 193
Table (8.10): The numerical results of Example (8.1) 204
Table (8.11): The numerical results of Example (8.2) 206
Table (8.12): The numerical results of Example (8.3) 207
Table (8.13): Numerical results for worst-case optimization case of Example (8.4) 211
Table (8.14): Numerical results for average performance optimization case of Example (8.4) ... 212
Table (8.15): Numerical results for Example (8.6) 222
Table (8.16): The eigenvalues of the controllability Gramians in Example (8.8) 228

List of Figures:

Figure (4.1): Controllability and the energy required to transfer states from the origin to terminal states x_f , $\|x_f\| = R$ 42

Figure (4.2): Mass-spring-damper system 46

Figure (4.3): $E_{\min}$ as a function of variable terminal time where $t_f \in [0.5,10]$ 47

Figure (4.4): $E_{\min}$ as a function of terminal time variable, $0 < t_f \leq \infty$ 48

Figure (4.5): The schematic picture of Furuta pendulum 49

Figure (4.6): $E_{\min}$ as a function of final time variable, $0 < t_f < \infty$ 51

Figure (4.7): Energy levels considering $E_{\min}^{\min}$ 55

Figure (4.8): Energy levels for furuta pendulum considering $E_{\min}^{\min}$ 55

Figure (4.9): Energy levels considering $E_{\min}^{\max}$ 57

Figure (4.10): Energy levels for furuta pendulum considering $E_{\min}^{\max}$ 57

Figure (4.11): Minimum input energy as a function of the variables t_f , x_f 58

Figure (4.12): Reachability area changes based on different terminal times for 2 dimensional Example model in (4.16) 59

Figure (5.1): Minimum energy levels with respect to the norm of initial and terminal states (Example (5.1)).....73

Figure (5.2): Minimum energy levels of Example (5.2).....74

Figure (5.3): Max and Min energy levels' directions in Example (5.2).....75

Figure (5.4): Minimum energy levels of Example (5.3).....76

Figure (5.5): Max and Min energy levels' directions in Example (5.3).....77

Figure (5.6): Pure energy values taken by the unstable subsystem when $R_0 = 1$ and $R_f = 0$ in Example (5.2) 78

Figure (5.7): Minimum energy levels when the system steers from the worst initial direction $R_0 (= 100,1000)$ and moves toward the final states located on the sphere with the radius $R_f = 1$ in different directions in Example (5.2) 78

Figure (5.8): Pure energy values taken by the stable subsystem when $R_f = 1$ and $R_0 = 0$ in Example (5.3).....79

Figure (5.9): Minimum energy levels when the system moves from various initial directions $R_0 = 1$ toward the final states located on the sphere with the radius $R_f\,(=100,1000)$ in the worst direction in Example (5.3) 79
Figure (5.10): Schematic of the AMB system 81
Figure (5.11): The schematic of seesaw-pendulum process 85
Figure (5.12): Two-CSTR Process 93
Figure (5.13): Model of bus suspension system (matched disturbance) 98
Figure (5.14): Model of bus suspension system (unmatched disturbance) 99
Figure (7.1): $trace(W_c(0,1))$ with respect to different values of b 139
Figure (7.2): The optimal input vectors and the spectral subspace 150
Figure (7.3): b_2 must be located on the boundary of the cone with centre b_1 154
Figure (7.4): Vector b_3 lies on the intersection of the two cones centered at vectors b_1 and b_2 154
Figure (7.5): $trace\{W_c(0,1)\}$ according to the variable input matrix B in the interval $[-1:0.2:1]^2$ 160
Figure (7.6): $trace\{W_c(0,1)\}$ respect to the variable input matrix B in the interval $[-1:0.1:1]^2$ 160
Figure (7.7): $trace\{W_c(0,1)\}$ based on the angle ε between b_1, b_2 (top), and $trace\{W_c(0,1)\}$ as a function of α_1^2, α_2^2 (bottom) 161
Figure (8.1): Steady states of logistic equation (8.2) 171
Figure (8.2): Graphical resolution for various values of r 173
Figure (8.3): Schematic representation of the bifurcation diagram 174
Figure (8.4): Schematic representation of the KA optimization method 175
Figure (8.5): Change of parameters over the course of iterations. 179
Figure (8.6): The movement of a solution around the destination via change of the parameters over the iterations 180
Figure (8.7): Typical two-dimensional form of unimodal benchmark functions 183
Figure (8.8): Typical two-dimensional form of high-dimensional multimodal benchmark functions 184
Figure (8.9): Typical two-dimensional form of fixed-dimensional multimodal benchmark functions 184
Figure (8.10): Benchmark function F1-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history 187

Figure (8.11): Benchmark function F2-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....187
Figure (8.12): Benchmark function F3-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....188
Figure (8.13): Benchmark function F4-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....188
Figure (8.14): Benchmark function F5-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....189
Figure (8.15): Benchmark function F6-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....189
Figure (8.16): Benchmark function F7-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....190
Figure (8.17): Benchmark function F8-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....194
Figure (8.18): Benchmark function F9-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....194
Figure (8.19): Benchmark function F10-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....195
Figure (8.20): Benchmark function F11-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....195
Figure (8.21): Benchmark function F12-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....196
Figure (8.22): Benchmark function F13-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....196
Figure (8.23): Benchmark function F14-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....197
Figure (8.24): Benchmark function F15-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....197
Figure (8.25): Benchmark function F16-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....198
Figure (8.26): Benchmark function F17-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....198
Figure (8.27): Benchmark function F18-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....199
Figure (8.28): Benchmark function F19-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....199

Figure (8.29): Benchmark function F20-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....200
Figure (8.30): Benchmark function F21-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....200
Figure (8.31): Benchmark function F22-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....201
Figure (8.32): Benchmark function F23-The convergence rate, average fitness of all individuals, trajectory in first dimension and the fitness history.....201
Figure (8.33): Convergence curve of Example (8.1).....204
Figure (8.34): A tension/compression spring.....204
Figure (8.35): A pressure vessel, and its design variables.....205
Figure (8.36): Convergence curve of KA in Example (8.2).....205
Figure (8.37): Convergence curve of KA in Example (8.3).....208
Figure (8.38): Mass-Spring-Damper system.....208
Figure (8.39): Block diagram of the mass-spring-damper system with uncertainty.....209
Figure (8.40): The augmented system G_{aug} 209
Figure (8.41): Three-story building model with supplemental passive viscous dampers.....213
Figure (8.42): The best objective values of Example (8.5) using KA optimization.....215
Figure (8.43): Top-floor displacement response of the 3-storey building.....216
Figure (8.44): Floors' maximum displacement (One damper).....216
Figure (8.45): Inter-story displacement response of the floors for one damper in different locations 217
Figure (8.46): Acceleration of the floors using two dampers 217
Figure (8.47): Floors' maximum displacement (Two dampers).....218
Figure (8.48): Inter-story displacement response of the floors for two dampers in different locations 218
Figure (8.49): Acceleration of the floors using two dampers 219
Figure (8.50): Floors' maximum displacement (Three dampers).....219
Figure (8.51): Inter-story displacement response of the floors for three dampers in different locations.....220
Figure (8.52): Acceleration of the floors using three dampers.....220
Figure (8.53): The input and state of the system with the actuator set $s_{actuator} = \{1, 2, 3, 4\}$ 222
Figure (8.54): The input and state of the system with the actuator set $s_{actuator} = \{1, 2, 3\}$ 223
Figure (8.55): The input and state of the system with the actuator set $s_{actuator} = \{2, 3, 4\}$ 223

Figure (8.56): The input and state of the system with the actuator set $s_{actuator} = \{1, 4\}$ 223
Figure (8.57): The open-loop system of Example (8.7).....224
Figure (8.58): The convergence curve of Example (8.7) 224
Figure (8.59): The graph of cost function of Example (8.7) 225
Figure (8.60): Local minimum of Example (8.7).....225
Figure (8.61): Control force input needed to control the system when using the optimal actuation 226
Figure (8.62): Control force input needed to control the system when using the input vector $b^i, i = 1, 2, 3, 4$ 226
Figure (8.63): The convergence curve of the KA algorithm in Example (8.8) 229

Chapter 1: Introduction

The instrumentation of a process, that is the selection of measurement variables (outputs) and actuation variables (inputs) has a “micro” (local), as well as a “macro” (global) aspect. The “micro” role of instrumentation has been well developed (Finkelstein & Grattan, 1994) and deals with the problem of measurement, or implementation of action upon given physical variables; instrumentation theory and practice deals almost exclusively with the latter problems. The “macro” aspects (Karcanias, 1994) of instrumentation stem from that designing an instrumentation scheme for a given process (classification and selection of input and output variables) expresses the attempt of the “observer” (designer) to build bridges with the “internal mechanism” of the process in order to observe it and/or act upon it. What is considered as the final system, on which Control Systems Design is to be performed, is the object obtained by the interaction of the “internal mechanism” and the specification of the overall instrumentation scheme. Difficulties in control of the final system may be assessed in terms of certain structural characteristics of the final system model (MacFarlane & Karcanias, 1976), (Kouvaritakis & MacFarlane, 1976) and the shaping of the degree of the presence of certain system properties, such as controllability and observability. These structural characteristics are formed through various stages, where the design goes through; however, the process of formation of such structural characteristics, as well as the link between their types, values and nature to control problems is not yet well understood. From the systems viewpoint, global instrumentation is seen as a model structure shaping, design stage as far as the characteristic of the final model is concerned. Given

that the structure of the model determines, in a sense, what can be achieved under compensation, global instrumentation is intimately linked to control design. The problems of control design and overall selection of input, output schemes for a process, referred to here as Global Process Instrumentation (GPI) (Karcanias, 1994), are strongly interrelated and this has been specially recognized in the Process Control area, where issues of selection of input, output schemes have been considered within the area of control structure selection (Morari, et al., 1980), (Morari & Stephanopoulos, 1980), (Morari & Stephanopoulos, 1980), (Govind & G.J. Powers, 1982), (Georgiou & Floundas, 1989). So far, however, there has been no systematic attempt to develop a unifying framework for selection of systems of measurement and actuation variables for processes, where the model structure shaping role of Global Instrumentation is the central feature. The overall research activity here reflects the view that instrumentation and control cannot be seen as independent activities, but as interrelated tasks with an integrated methodology.

The problem of selection of input, output schemes for a process, is part of the overall design of the process, which is of cascade nature and has as main stages, the Process Synthesis, Global Process instrumentation (GPI) and Control System design. It has been argued (Karcanias, 1994) that there is a correspondence between the successive design stage decisions of the cascade design process and the evaluation of structural characteristics of process models. Global Instrumentation plays a crucial role in the shaping of structural characteristics, as well as the boundary values of related system properties. This has the advantage that very frequently many degrees of freedom are available, which may be used for design purposes. The central characteristic of this approach is that we view GPI as a process of shaping further the inherited structure from the process synthesis stage. This expresses a fundamental property in the structure evolution during the design (Karcanias, 2008). Note that the term structure is viewed here as a linear graph and/or as system invariants of the underlined model.

The problem of selection of sets of inputs and outputs has been based so far on the shaping of structural invariants (Karcanias & Vafiadis, 2002), such as the set of invariant zeros (MacFarlane & Karcanias, 1976), (Kouvaritakis & MacFarlane, 1976), (Rosenbrock & Rower, 1970), (Karcanias & Giannakopoulos, 1989), (Karcanias, 1994), (Karcanias, 1996), (Karcanias & Vafiadis, 2002), (Karcanias & Vafiadis, 2002), (Leventides & Karcanias, 2008), (Georgiou & Floudas, 1990). Shaping properties such as controllability and observability, goes beyond the shaping of structural invariants. Structural invariants (Karcanias & Vafiadis, 2002), (Karcanias, 2002) determine the shape of system properties, but do not provide the measure of presence of such properties in a system. In this thesis, we will use energy considerations to evaluate the role of selection of systems of inputs and respectively outputs on the energy requirements for state control and state reconstruction. The study of required energy for control and state reconstruction has been a standard theme in the study of linear systems (Kailath, 1980), (Skelton, 1988). However, this has not being used for evaluation and selection of systems of inputs (location of actuators) and systems of outputs (location of sensors) so far. Such a study is intimately related to the study of degree of controllability and observability of a linear system (Moore, 1981), (Arbel, 1981), or alternatively measuring the distance of a system from the set of uncontrollable, respectively unobservable systems.

The objectives of this thesis are:

Objectives:

- **(i)** Study the energy dependency of the properties of controllability and observability
- **(ii)** Define the relative degree and presence of controllability and observability under conditions where the available energy is constrained.
- **(iii)** Develop strategies and methodologies for selection of systems of inputs and outputs to minimize the energy required for control, respectively state reconstruction.
- **(iv)** Evaluate the effect of systems of inputs and outputs on various other system properties.

This thesis contributes to the further development of the area by clarifying the role of energy requirements in achieving control and state observation as functions of given sets of inputs and outputs and thus provide energy based tools for the selection of input and outputs schemes with a main criterion the minimization of the energy required for control and observation. The methodologies developed may be used as design tools where apart from energy requirements other design criteria may be also incorporated for the selection of inputs and outputs. The methodology that is used is based on Linear Systems theory (Kailath, 1980), (Karcanias, 2002), (Karcanias & Vafiadis, 2002) and tools from Numerical Linear Algebra. The solution to the problems considered here is an integral part of the effort to develop an integrated approach to Control and Global Process Instrumentation (Karcanias, 1994).

The main achievements of the thesis are in the following areas:

- Provide a literature review for link of energy to controllability and observability and a review of methodologies for selection of inputs and outputs.
- Perform an overview of controllability/observability Gramians, and methodologies for their computation.
- Introduce the minimum energy and define the constant energy surfaces linked to controllability and observability.
- Develop computations of Energy for unstable systems and measures of degree of controllability/observability.
- Develop a selection methodology of system inputs, and thus of the input matrix B, to minimize the average required energy for controllability, respectively observability.
- Present a metaheuristic optimization method for the optimal selection of input/output system structure.

The thesis is structured as follows:[1]

[1] Note: (Kalman, 1959) has pointed out that observability and controllability are duals of each other. We can observe some similarities between optimal actuator model and optimal observation model because of this duality. In order to avoid repetition in this study, what is said about actuator selection can be applied to the concept of sensor selection with the proper interpretation.

Chapter 1

This is an introduction chapter, which provides a brief overview of the problem and summarizes the contents of each of the following chapters.

Chapter 2

Describes a review of the literature showing the importance of the input/output selection in control system design. It reviews the current methodologies and criteria for selection of inputs/outputs and states the open challenges.

Chapter 3

The concepts of reachability and reachability gramian are introduced for linear continuous time-invariant systems. The system properties of controllability and controllability Gramian and a novel approach of calculation of controllability gramian are studied in this chapters. The problem of controllability Gramian assignment is investigated.

Chapter 4

The chapter develops the fundamentals for selection of input (and by duality output) structure based on the energy type criteria, and provides answer to the question of the required energy for the transfer of the origin to a given state in the state space within some given final time when the input structure is given. It also introduces first the problem of minimum input energy and then discusses the important related factors which can influence its value.

Further, this chapter presents the definition of the energy levels in the state space characterizing the maximum of minimum input energy required for transfer to a given distance. This reveals the presence of the energy stratification of the state space.

Chapter 5

The chapter discusses the energy calculation for unstable systems and the impact of actuator selection on various issues such as the degree of controllability, and disturbance rejection. Chapter 5 also presents the generalized energy stratification of the state space for possibly unstable systems.

A method of actuator selection is also presented considering the degree of controllability and the degree of disturbance rejection.

The discussed methods which are proposed in chapter 5 are general methods in the sense that they support both stable and unstable systems.

Chapter 6

The chapter considers the case where the system matrix A is Hurwitz and has certain canonical form structures. This chapter presents some Theorems based on which the controllability Gramian of the continuous-time linear time invariant system can be constructed directly based on the coefficients of the characteristic polynomial, along with the value of the trace and some upper bounds for the maximum eigenvalue of the controllability. Furthermore, the chapter investigates how the value of the energy could be derived based on the characteristic polynomial of the system.

Chapter 7

This chapter proposes a strategy for the selection of the proper input matrix B, based on the average of the minimum input energy. This is formulated and solved as an optimization problem for different cases.

First the problem of finding the best single input structure to minimize the energy requirements is discussed and a solution for a general system with a finite terminal time t_f is proposed. Then, the same optimization problem is discussed for the case of stable system with infinite terminal time. This is followed by the input structure selection for the normal system where $AA^* = A^*A$. In the last part of this chapter, the case of multi-input systems subject to different possible conditions are investigated.

The work here provides a new approach to those in existing literature, which considers the problem of input structure selection over a binary set, i.e. where the input matrix B can only take the values {0,1} or can be chosen among the given sets. The current approach developed in this chapter does not rely on such restrictive assumptions.

Chapter 8

The chapter develops a new metaheuristic optimization method based on the logistic equations. Superior performance in terms of exploration, exploitation, and convergence is demonstrated relative to other state-of-the-art methods. These properties make the proposed algorithm powerful and capable of solving complex high-dimensional multi-parameter problems, such as optimal actuator and sensor placement.

Chapter 9

This chapter reviews the contribution of this thesis and draws some conclusions. Directions of further work are discussed.

Chapter 2: Literature Review

2.1. Introduction

In this chapter, we review previous works directly relevant to the problem of sensor and actuator selection. In section 2.3, the problem of optimal actuation in continuous linear time-invariant systems is described. The dual problem can also be formulated to find the optimal location of the sensors. With increasing focus on the particular problem of energy optimization, in section 2.4, we consider some of the available methods on input/output selection (IO) and we review the indexes and the criteria specialized for optimal actuator/sensor placement in continuous linear time-invariant systems. Finally, in section 2.5, we summarize this chapter.

2.2. Input/Output Structure Selection

44 years ago (Foss, 1973) challenged the process control research community as he made the observation that in many areas, application was ahead of theory. He introduced control structure design and the gap between theory and applications in this important area:

“The central issue to be resolved by the new theories are the determination of the control system structure. Which variables should be measured, which inputs should be manipulated and which links

should be made between the two sets. . . . The gap is present indeed, but contrary to the views of many, it is the theoretician who must close it."

A similar observation on the need of control structure design was made by (Findeisen, et al., 1980) (p. 10). Control structure design determines which variables to control, which variables to measure, which inputs to manipulate and which links that should be made between them. (Morari, et al., 1980), (Morari & Stephanopoulos, 1980), (Morari & Stephanopoulos, 1980) introduced new and exciting ideas and theories in the area of control structure design. In the late 1980s (Nett, 1989), (Minto & Nett, 1989) presented a number of lectures about the selection and partitioning of measurements and manipulations for the control of complex systems based on his experience on aero-engine control at General Electric.
A good review of the literature on control structure design can be found in (Larsson & Skogestad, 2000).
While the area of control structure design has received some interest in the literature, e.g. (Georgiou & Floudas, 1989), (Govind & Powers, 1982), (Karcanias, 1996), (Morari & Stephanopoulos, 1980), (Rijnsdorp, 1991), (Skogestad & Postlethwaite, 1996), (Skogested, 2000) and references therein, the gap still remained, and still does to some extent today. Control structure design cannot be compared to the enormous amount of work on controller design, although it is probably the most important in practice. It involves a number of key sub-problems (Karcanias, 1996) which are:

I. The classification of process variables into potential inputs, outputs and referred to as Model Orientation Problem (MOP)
II. Specification of effective sets of inputs, outputs on an oriented model and referred to as Model Projection Problem (MPP)
III. The selection of a control configuration (i.e. CC), which decides on the way we couple effective inputs and outputs for control design purposes (also called the measurement/manipulation partitioning or input/output pairing).

The above sub-problems have received attention in the literature, see e.g. (Wal & Jager, 2001), (Padula & Kincaid, 1999) and references therein. (Chen, 2002) and (Assali, 2008) summarized different methods for the above sub-problems. Furthermore, (Reinschke, 1988), (Lin, 1974), (Siljak, 1991), (Siljak, 1978), (Murota, 2009) and references therein discussed different solution criteria and applications of the IO and CC selection problems. However, most of the attention so far has been focused on the last sub-problem, i.e. control configuration, when heuristics and diagnostic indicators have been used. For the first two sub-problems, which are the focus of this thesis less attention has been given, especially from the Control Theory viewpoint, with the exception of the work in (Karcanias, 1996), (Georgiou & Floudas, 1989), (Karcanias & Giannakopoulos, 1989), (Morari & Stephanopoulos, 1980) on some specific problems.

In control structure design, the determination of the adequate number, place and sensor-actuator type is the input/output selection (i.e. IO selection) process. The system outputs are the set of the measured variables and the controlled variables. The system inputs are the set of manipulated inputs and the exogenous inputs, such as disturbances, sensor noise and reference inputs or the set-points.

In the control structure selection framework and in this thesis outputs are the measured variables and inputs are the manipulated variables.
The Importance of the IO structure selection lies in the fact that it may affect many properties of the system such as zeros locations, cost, maintenance, reliability, controllability, observability and complexity.
A formal definition for the IO selection is given by (Wal & Jager, 2001):

"Select suitable variables u to be manipulated by the controller and suitable variables y to be supplied to the controller"

Qualitative rules for IO selection problem could be found in (Seborg, et al., 1989):

" i) Control the outputs that are not self-regulating, ii) Control the outputs that have favourable dynamic and static characteristics, iii) Select inputs that have large effects on the outputs, iv) Select inputs that rapidly affect the controlled variables. "

Over the last few decades different performance criteria and measurements have been proposed for the selection of actuation variables (inputs) and sensor locations (outputs) (Georges, 1995), (Marx, et al., 2002), (Marx, 2003), (Marx, et al., 2004), (Singh & Hahn, 2005), (Singh & Hahn, 2006). A survey on IO selection is given by (Wal & Jager, 2001). They list several criteria for evaluation of control structure design methods: generality, applicable to nonlinear control systems, controller-independent, direct, quantitative, efficient, effective, simple and theoretically well developed. Ideally, an IO selection method should satisfy all these criteria. After reviewing they conclude that such a method does not exist.
Some discussions on IO selection in the process industry are given in (Morari, 1982), (Shinskey, 1988), (Stephanopoulos, 1984) and (Balchen & Mummé, 1988).

A literature search for optimal actuator or optimal sensor placement methods yields a wide range of publications from different engineering disciplines. Some of these references describe small optimization problems and employ manual optimization techniques or intuitive placement recipes rather than systematic optimization methods. Other references discuss challenging numerical optimization problems and most often use genetic algorithms as the optimization method.

In this thesis, we study the mathematical strategies to determine the optimal actuator locations. A new metaheuristic optimization algorithm is also developed to solve the challenging actuation problems in the case of large scale systems. By duality, the solution methodology may be readily extended to the problem of optimal sensor selection. To verify the results, wherever applicable, numerical examples are designed and the results are compared with those available in the literature.

2.3. Problem Statement:

As it is discussed in the previous sub-section, the location of an actuator/sensor has a tremendous impact on the performance of the controlled system. Misplaced actuators can lead to lack of controllability (Kumar & Narayanan, 2007). The actuators/sensors should therefore be located at positions that optimize certain performance objectives.

In aerospace engineering, the actuators at optimal locations reduce the vibrations (Mehrabian & Yousefi-Koma, 2007). The optimal placement of actuators/sensors is essential for effective control of structural vibration and acoustic noise (Pulthasthan & Pota, 2008). In acoustic problems, an arbitrarily placed actuator can actually increase the sound field locally (Fahroo & M.A. Demetriou, 2000). Civil structures such as high-rise buildings and suspension bridges are designed to protect against earthquake excitation through the placement of actuators at appropriate heights (Abdullah, et al., 2001). The energy consumption is a concern in various control applications, e.g. vibration control of smart structures, chemical process control. (Arbel, 1981), (Peng, et al., 2005) showed that the optimal placement of actuators/sensors improves the performance of the control system significantly and at the same time the energy consumption is minimized. In (Leleu, et al., 2001), (Sung, 2002), (Yue, et al., 2008), (Jha & Inman, 2003), (Hac & Liu, 1993) improved performance of vibration control is observed when actuators/sensors are placed at optimal locations. In (Antoniades & Christo, 2002), an integrated feedback design problem and optimal actuator placement minimizing the control energy was investigated. It is shown in (Morris, 1998) that the optimal actuator placement can reduce noise in a duct. In systems modelled by partial differential equations, optimal location calculations are performed on approximated problems, although the state space for the full model is infinite dimensional. The theory that guarantees optimality of the cost and existence of the optimal actuator location for these models has not been developed in its entirety (Hebrard & Henrot, 2005).

The aim of this thesis is to develop a mathematical framework for calculating optimal actuator locations with regard to the energy-based criteria, and to develop a new evolutionary algorithm to solve the combinatorial problem of optimal actuator/sensor placement in the challenging real world applications.

Consider the following linear time-invariant system in $\mathbb{R}^n$:

$$\begin{cases} \dot{x} = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}, \quad x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{l\times n}, D \in \mathbb{R}^{l\times m} \qquad (2.1)$$

where x denotes the state of the system, A denotes the system dynamics and $u(t)$ is the control applied to the system as a function of time. The effect of control on the state of the system is described by the input matrix B.

In many control systems, the location of actuators/sensors can often be chosen. These locations should be selected in order to optimize the performance criterion of interest.

Consider the situation where there are m actuators with locations that could be varied over some compact set, call it $\Omega \subset \mathbb{R}_a$. Parametrize the actuator locations by a and denote the dependence of the corresponding input matrix with respect to the actuator location by B_a. Note that a is a vector of length m with components in Ω so that a varies over a space denoted by Ω^m. Based on the designer's

interest on the choice of a desired performance measure, a suitable cost function J_a that depends on the actuator location is formulated.

Definition (2.1): The optimal cost J_{opt} over all possible locations is defined as,

$$J_{opt} = \inf_{a \in \Omega^m} J_a \tag{2.2}$$

Also, if it exists, the location $a_{opt} \in \Omega^m$ that satisfies

$$a_{opt} = \arg \inf_{a \in \Omega^m} J_a \tag{2.3}$$

is called the optimal actuator location.

In the rest of this chapter, the background of the formulations of the cost function J_a with regard to the energy-based strategies used in diverse applications will be discussed.

2.4. Different Cost Functions:

Previous studies have used various measures and criteria in order to increase the effectiveness of sensors and actuators and to reduce their numbers and sizes while some specific variables are optimized (Wal & Jager, 2001), (Frecker, 2003), (Padula & Kincaid, 1999). The most commonly used criteria in literature are:

- **Accessibility**

It is a qualitative technique for IO selection based on cause and effect graphs, which can be generated for linear and nonlinear systems. Vertices of the graph represent system variables, disturbances, and control signals. Directed edges depict relations between various variables (Kościelny, et al., 2017). The key idea for IO selection is that a causal path must exist between the manipulated and the controlled variables on the one hand and the measured and the controlled variables on the other hand. A large number of candidate IO sets may be termed viable if they are only assessed for accessibility. So, additional criteria should be invoked (Wal & Jager, 2001).
Accessibility may be chosen as a tool for actuator/sensor placement because of the following features:

I. Faults can be included directly in the model as additional vertices.
II. The graph model is a simple, intuitive, and easily understandable way of describing a process. Mathematical description is not needed, only basic knowledge about the physics of a process. This is an important feature because actuator/sensor placement analysis should be performed at the design stage and a detailed process model can be unavailable.

(Lambert, 1977) used fault-trees to analyse the location of sensors. Even though this work was the first step toward the design of sensor locations based on a diagnostic observability criterion, it had drawbacks such as: i) inability to handle cycles, and ii) the development of a fault tree is in itself an error-prone and time-consuming process. To solve the problem of observability, based on the process graph, (Ali & Narsimhan, 1993), (Ali & Narsimhan, 1995) have presented sensor placement strategies

for linear and bilinear processes (Ali & Narasimhan, 1996). They introduced the concept of reliability and the degree of redundancy (spatial redundancy: measuring more variables than the minimum required to ensure observability, so that there are multiple ways of estimating a variable (Mah, 1990) and hardware redundancy: measuring a variable using more than one sensor) for sensor placement in linear systems. There are also other methods based on linear and bilinear mathematical models for sensor location (Madron & Veverka, 1992). They proposed two objective functions for optimization: overall measurement cost and overall precision of a system and utilized Gauss-Jordan elimination to identify a minimum set of variables that need to be measured in order to observe all important variables. Their work is essentially an extension of the work of (Vaclavek & Loucka, 1976). (Raghuraj, et al., 1999), (Bhushan & Rengaswamy, 2000), (Bhushan & Rengaswamy, 2000b), (Khemliche, et al., 2006), (Bhushan, et al., 2008), (Yang, et al., 2009) also used qualitative (cause-effect) graph theoretic approaches to find the optimal sensor locations in a process.

- **Linear Quadratic Control**

Linear quadratic control is a well-known closed-loop strategy for controller design that minimizes energy of both the control signal and the measured signal. The control is calculated by minimizing a quadratic cost function with penalty on both the state of the system and the control input. This objective is concerned with controlling the initial condition to 0 and disturbances are neglected.

Consider the system (2.1) on $\mathbb{R}^n$. In the case of a single control u, the linear-quadratic (LQ) controller design objective is to find a control $u(t)$ to minimize the quadratic cost functional $J(u, x_0)$ over infinite-time interval, for a given initial state x_0:

$$J(u, x_0) = \int_0^\infty \langle x(t), Qx(t) \rangle + \langle u(t), Ru(t) \rangle \, dt \tag{2.4}$$

where $R > 0$ weights the control cost, Q is a self-adjoint positive semi-definite matrix weighting the state. The linear quadratic control problem is to minimize the cost function (2.4) over all possible controls:

$$\min_{u \in L_2(0,\infty;\mathbb{R}^m)} J(u, x_0) \tag{2.5}$$

The control that achieves this minimum, u_{opt}, is often called the linear quadratic optimal control.

Definition (2.2): The pair (A, B) is stabilizable if there exists $K \in \mathbb{R}^{m \times n}$ so that $A - BK$ is Hurwitz.

□

Definition (2.3): The pair (A, C) is detectable if there exists $F \in \mathbb{R}^{n \times l}$ so that $A - FC$ is Hurwitz.

□

Theorem (2.1): (Morris, 2001) Consider the system given by (2.1). assume the case that there are m actuators with locations that could be varied over the compact set Ω. If $(A, B_a, Q^{1/2})$ is both stabilizable and detectable. Then the infinite-horizon optimization problem (2.5) has a minimum for every given initial condition x_0. Furthermore, there exists a symmetric, semi-positive definite matrix, $P_a \in \mathbb{R}^{n \times n}$, such that:

$$\min_{u \in L_2(0,\infty;\mathbb{R}^m)} J_a(u, x_0) = J_a(u_{opt}, x_0) = x_0^T P_a x_0 \tag{2.6}$$

where P_a is the unique solution of the Algebraic Riccati Equation (ARE):

$$A^T P_a + P_a A - P_a B_a R^{-1} B_a^T P_a + Q = 0 \tag{2.7}$$

The optimal control u_{opt} is as follows:

$$u_{opt} = -R^{-1} B_a^T P_a x(t) \tag{2.8}$$

and the corresponding optimal state feedback, $K = -R^{-1} B_a^T P_a$ is stabilizing, i.e. $A - BK$ is Hurwitz.

□

The proof is given in (Morris, 2011), see Theorem (5.12) and Theorem (5.16). According to the above Theorem, for a particular initial condition x_0, the optimal actuator location problem can be formulated as follows:

$$\inf_{a \in \Omega^m} \; x_0^T P_a x_0 \tag{2.9}$$

The linear quadratic index has been studied in various literature as an objective function to optimise actuator location and feedback gain, (Demetriou, 2000), (Chen, et al., 2012), (Li & Huang, 2013).
The optimal linear-quadratic cost at a given actuator location a, which is given by (2.9) depends on the initial condition x_0. Researchers have used a number of techniques in the past to remove this dependency. The first approach is to consider the worst case initial condition (Devasia, et al., 1993), (Morris, 2011). The cost function is:

$$\max_{\|x_0\|=1} \; \min_{u \in L_2(0,\infty;\mathbb{R}^m)} \; J_a(u, x_0) = \max_{\|x_0\|=1} \; (x_0^T P_a x_0) = \lambda_{\max}(P_a) \tag{2.10}$$

Then the optimal actuator selection problem can be formulated as:

$$\inf_{a \in \Omega^m} \; \lambda_{\max}(P_a) \tag{2.11}$$

Another method is to view the initial condition as a random vector with zero mean and unity variance (Morris, 2011), (Geromel, 1989). The expected optimal linear-quadratic cost is:

$$E\left[\min_{u \in L_2(0,\infty;\mathbb{R}^m)} \; J_a(u, x_0) \right] = E\left[x_0^T P_a x_0 \right] = trace(P_a) \tag{2.12}$$

Thus, the problem of optimal actuator location can be defined as:

$$\inf_{a \in \Omega^m} \; trace(P_a) \tag{2.13}$$

Another approach is to average the cost over a set of linearly independent initial states (Antoniades & Christo, 2002). The effect of a disturbance with fixed frequency content (for instance, a single white noise disturbance) is considered in (Morris & Demetriou, 2010). This leads to a H_2 control problem and if the spatial distribution of the disturbance is unknown, then the cost function (2.10) is used to calculate the optimal actuator location. Linear quadratic control is a popular choice, since the controller is designed simultaneously with the optimal actuator location. A methodology based on the minimisation of the optimal linear quadratic control index as an objective function was proposed by (Devasia, et al., 1992). This objective function was applied to a simply-supported beam in order to optimise the size and location of actuators using a simple search method. They reported that this objective function achieves stability of closed-loop control system and allows a designer to choose various values of weighted matrices of optimal linear quadratic index in order to optimise actuator location and vibration reduction. (Kondoh, et al., 1990) proposed an objective function based on the minimisation of the linear quadratic index to optimise sensor and actuator location and feedback gain

for a cantilever beam. The authors reported that the optimal linear quadratic control has a clear physical meaning and the flexibility to allow the implementation of varying weighted matrices. (Sung, 2002) used the linear quadratic index with the worst case initial condition for placing actuators on a simply supported beam with a moving mass. In (Fahroo & M.A. Demetriou, 2000), the linear quadratic cost function with a random vector uniformly distributed initial condition was minimized for finding the best locations of actuators to reduce the interior noise in an acoustic cavity. To determine the most effective placement and design of an active control system, (Levine & Athans, 1970), (Abdullah, et al., 2001) used the cost function (2.12) to consider both the building response and the control effort. (Kumar & Narayanan, 2007), (Kumar & Narayanan, 2008) considered the optimal placement of collocated piezoelectric actuator/sensor pairs on flexible beams using a model-based linear quadratic regulator (LQR) controller.

- **H_∞ and H_2 Norms**

Noise in the environment and disturbances, which are often unknown, will influence the performance of the control system. Therefore, in many situations, the problem is to reduce the effect of disturbances on the performance. The control system is now described by:

$$\begin{cases} \dot{x} = Ax(t) + Bu(t) + B_d d(t) \\ y(t) = Cx(t) + Du(t) \end{cases}, \quad x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{l\times n}, D \in \mathbb{R}^{l\times m}, B_d \in \mathbb{R}^{n\times k} \quad (2.14)$$

where $d(t)$ is the exogenous input or disturbance. H_∞ and H_2 norm functions are a measure of open and closed loop system frequency response to an external disturbance. These criteria have been used as objective functions to optimise the location of sensors and actuators. The problem is to find a controller $u(t)$ to minimize:

$$\|y\|_{L_2(0,\infty,\mathbb{R}^l)} = \sqrt{\int_0^\infty \|y(t)\|^2 \, dt} \quad (2.15)$$

Since the L_2 norm of y equals the H_2 norm of the Laplace transform of y, this is known as an H_2 controller design problem. The H_2 and H_∞ cost (Geromel, 1989) are closely related to the linear quadratic criteria examined in the previous sub-section. (Leonides, 2012)

The usual orthogonality hypotheses $C^T D = 0$ and $D^T D = I$ are assumed in order to simplify the subsequent equations. The full column rank of matrix D ensures a non-singular map y on the control $u(t)$. It is assumed that all states of the system are available for measurement. This system (2.14) is a special form of the generalised plant configuration, known as the full information problem.

Assume that $d(s)$ and $y(s)$ be the Laplace transforms of the disturbance and the output respectively. Then, minimizing the effect of all disturbances on the output is equivalent to minimizing the size of the transfer function:

$$G_{yd} = \frac{y(s)}{d(s)} \quad (2.16)$$

The notation H_∞ indicates the Hardy space of all functions $G(s)$, which are analytic in the right-half plane Re(s) > 0 and for which:

$$\sup_\omega \lim_{x\to 0} |G(x + j\omega)| < \infty \quad (2.17)$$

Thus, H_∞ norm is defined as:

$$\left\|G(j\omega)\right\|_\infty = \sup_\omega \bar{\sigma}(G(j\omega)) \tag{2.18}$$

where $\bar{\sigma}(G(j\omega))$ denotes the maximum singular value of the transfer function $G(j\omega)$. The details of calculating H_∞ norm can be found in many literature, see e.g. (Morris, 2001), (Zhou & Doyle, 1996).
Let G be the transfer function of the system (2.14):

$$G(s) = C(sI - A)^{-1}\begin{bmatrix} B & B_d \end{bmatrix} + \begin{bmatrix} D & 0 \end{bmatrix} \tag{2.19}$$

With state feedback control:

$$u = -Kx(t) \tag{2.20}$$

The closed loop transfer function from disturbance to output is:

$$G_{yd}(s) = (C - DK)(sI - (A - BK))^{-1}B_d \tag{2.21}$$

Definition (2.4): (Doyle, et al., 1989) The controlled system will have input d (disturbance) and output y (measurement). The fixed attenuation H_∞ control problem for attenuation γ of (2.14) is to construct a stabilizing controller with transfer function $G(s)$ so that the closed loop system G_{yd} with $u(s) = G(s)x(s)$ is L_2 stable and satisfies the bound:

$$\left\|G_{yd}\right\|_\infty < \gamma \tag{2.22}$$

□

Even for stabilizable systems, the fixed attenuation problem cannot be solved for every attenuation since such a controller may not exist. However, if it is solvable, as in the case of linear quadratic control, the control law can be chosen to be constant state feedback. (Özdemir, 2003)

Definition (2.5): The optimal H_∞ control problem for (2.14) with full-information is to find:

$$\hat{\gamma} = \inf \gamma \tag{2.23}$$

over all γ for which the fixed attenuation problem is solvable. The infimum $\hat{\gamma}$ is called the optimal H_∞ attenuation, which can be calculated using a bisection-type algorithm.

□

The H_∞ cost for a particular actuator location a is $\hat{\gamma}_a$ .The value of $\hat{\gamma}_a$ provides the best possible attenuation of the worst-case disturbance for the system at the actuator location a .
Another approach for placing actuators is to use the H_∞ norm of the closed-loop system, which has some advantages:

- The effect of the worst-case disturbance on the output of the system is minimized
- with optimal actuator location, the corresponding optimal state feedback controller is simultaneously designed
- The closed loop system is stabilized

(Gawronski, 1998) discusses the H_∞ criteria in more details.
Consider the performance index:

$$J(u, d; x_0) = \left\|y\right\|^2_{L_2(0,\infty,\mathbb{R}^l)} - \gamma^2 \left\|d\right\|^2_{L_2(0,\infty,\mathbb{R}^k)} \tag{2.24}$$

subject to (2.14) for some $\gamma > 0$. Calculating a controller that achieves the given H_∞ attenuation bound is equivalent to solving the optimization below (Morris, 2001):

$$\max_{d \in \mathbb{R}^k} \min_{u \in \mathbb{R}^m} J(u, d; x_0) \tag{2.25}$$

Consider the first term of the performance index (2.24). The L_2 norm of y equals the H_2 norm of the Laplace transform of y, this is known as an H_2 controller design problem. The H_2 optimal control is the state feedback (Zhou, et al., 1997):

$$u(t) = -B_a^* P_a x(t) \tag{2.26}$$

Subject to stabilizability of (A, B_a). P_a is the unique solution of the Algebraic Riccati Equation (2.7). Therefore, the H_2 optimal actuator location problem is to find the actuator location a that minimizes the corresponding cost (Morris, et al., 2015):

$$\sqrt{trace(B_d^* P_a B_d)} \tag{2.27}$$

Mathematically, the problem is identical to that of minimizing the LQ cost when the initial condition is random with covariance $B_d^* B_d$. If B_a is a continuous function of a, both the optimal actuator location problem with a fixed disturbance location, and the problem where the disturbance is unknown, lead to well-posed optimization problems. Since the optimal cost relies on the norm of the solution to Algebraic Riccati Equation (2.7), well-posedness of this problem follows using techniques and results similar to the linear quadratic case (Morris, et al., 2015), (Morris & Yang, 2015).

In a fluid application, (Chen & Rowley, 2011) used H_2 criteria for optimal actuator and sensor placement with respect to the disturbance and with respect to each other, which is shown to have a significant effect on performance. In (Ambrosio, et al., 2012), an H_2 criterion based on a modal approximation of a structure is applied. They have considered the neglected modes to accomplish spill-over attenuation. In (Guney & Eskinat, 2008), sensor/actuator pair placement on a simply supported beam and H_∞ control with an impulse disturbance was performed. The control signal is shown, but no comparisons were made between locations.

(Liu, et al., 2006), (Hiramoto, et al., 2000), (Arabyan & Chemishkian, 1998), (Chemishkian & Arabyan, 1999) have investigated the optimal locations of sensor and actuator under closed loop control using optimisation algorithms to find the optimal combinations of sensors and actuators. In (Hiramoto, et al., 2000), the product of a frequency weighting term that represents the design specification and the closed-loop transfer function was used for optimal actuator placement.

In (Raja & Narayanan, 2009) H_∞ cost function was used for locating actuators in the vibration control of tensegrity structures. The results showed that the displacement is less with H_∞ criteria than the H_2 performance index. (Silva, et al., 2006) used an H_∞ approach to place actuators on a plate.

(Sweeney, et al., 2005), (Demetriou & Grigoriadist, 2004) have used an analytical expression to compute the upper bound on the H_∞ norm of the controlled system was used to optimize the actuator locations. Using this analytical approach, it was shown in (Demetriou & Grigoriadist, 2007) that the resulting optimal actuator location exhibits spatial robustness. Calculation of optimal actuator locations with spatially varying disturbances was addressed in (Demetriou, 2004).

- **Degree of Controllability/Observability and Gramians**

Controllability refers to the property of being able to steer the state of a dynamical system from any starting point to any terminal point by means of appropriate inputs. Observability is concerned with whether without knowing the initial state, one can determine the state of a system given the input and the output. Controllability and observability of a system, which depend primarily on actuator and sensor location, will have a major influence on the efficiency of the control system and the control effort required to satisfy design requirements.

Definition (2.6): Consider the following linear time-invariant system:

$$\begin{cases} \dot{x} = Ax(t) + B_a u(t) \\ y = C_s x(t) + Du(t) \end{cases}, \qquad A \in \mathbb{R}^{n\times n}, B_a \in \mathbb{R}^{n\times m}, C_s \in \mathbb{R}^{l\times n}, D \in \mathbb{R}^{l\times m} \tag{2.28}$$

B_a denotes a family of input operators that are compact and continuous with respect to actuator location a . This system is said to be controllable at the actuator location a if, for any initial state x_0 and any final state x_f , there exists an input $u(t)$ such that the integral curve $x(t)$ generated by $u(t)$ with $x(0) = x_0$, satisfies $x(t_f) = x_f$ in a finite time $t_f > 0$. Otherwise, the system or the pair (A, B_a) is said to be uncontrollable at the actuator location a .

□

Definition (2.7): Consider the linear time-invariant system in (2.28), C_s denotes a family of output operators that are compact and continuous with respect to sensor location s . This system is called observable at the sensor location s if for any finite time $t_f > 0$, the initial state $x(0) = x_0$ can be determined from the time history of the input $u(t)$ and the output $y(t)$ in the interval $[0, t_f]$. Otherwise, the system or the pair (A, C_s) is said to be unobservable at the sensor location s .

□

Since the performance of the controlled system is highly related to the actuator/sensor locations (Fahroo, 1995), (Morris, 1998), many authors optimized actuator locations based on maximization of observability and controllability (Liu & Hu, 2010), (Peng, et al., 2005), (Qiua, et al., 2007), (Zhang & Erdman, 2006).

A general rule for IO selection method would be to reject candidate IO sets for which (A, B_a) is uncontrollable or (A, C_s) is unobservable. The problem of associating physically meaningful measures of quality with the notions of controllability and observability was discussed a little by (Brown, 1966) and (Monzxngo, 1967). Later, (Johnson, 1969)after a comprehensive review of the problem maximized one special scalar measure of the quality of complete controllability and complete observability for a class of linear dynamical systems. It is shown in (Olshevsky, 2014) that the minimal controllability problem is NP-hard; indeed, this paper utilizes a simple greedy heuristic which sequentially picks variables to maximize the rank increase of the controllability matrix. (Václavek & Loučka, 1976) proposed a sensor selection on steady-state systems using graph theory to ensure observability of

variables. (Caruso, et al., 2003) studied the optimal placement of actuators and sensors for a collocated flexible plate structure. For a given number of available piezoelectric patches of fixed dimensions to be bonded to the plate, the optimal locations were obtained by maximizing the modal controllability and observability of the structure. Various simple tests for this qualitative measure are available (Zhou, et al., 1997). According to Kalman rank condition, the system given in (2.28) is observable and controllable if the controllability and observability matrices (2.29) are full rank.

$$C = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}$$
$$O = \begin{bmatrix} C & AC & A^2C & \dots & A^{n-1}C \end{bmatrix}^T \quad (2.29)$$

A large number of studies has considered classical binary controllability metrics based on Kalman rank (Liu, et al., 2011), (Rajapakse, et al., 2011), (Cowan, et al., 2012), (Nepusz & Vicsek, 2012), (Wang, et al., 2012), (Olshevsky, 2014), (Pequito, et al., 2016). (Morari & Stephanopoulos, 1980) proposed structural state controllability and observability as the selection criteria.

The controllability and observability concepts has been strengthened by the fact that they are robust properties in linear constant systems(Lee & Markus, 1967). That is, the set of all controllable pairs (A, B) is open and dense in the space of all such pairs. This fact was exploited by (Lin, 1974) in formulating his concept of structural controllability, which states that all uncontrollable systems structurally equivalent to a structurally controllable system are a typical. Not only is this concept consistent with physical reality in the sense that values of system parameters are never known precisely, but it is also helpful in testing the properties of controllability using only binary computations. It is a well-known fact that testing controllability is a difficult numerical problem (Nour-Eldin, 1987), and structural controllability offers a computable alternative especially for systems of high dimension. Regarding the duality, the same facts hold for observability.

Definition (2.8): An $n \times n$ matrix $\tilde{M} = \left(\tilde{m}_{ij} \right)$ is said to be a structured matrix if its elements $\tilde{m}_{ij}$ are either fixed zeros or independent free parameters. For example, a 2×3 structured matrix is:

$$\tilde{M} = \begin{bmatrix} \times & \times & 0 \\ \times & 0 & \times \end{bmatrix}$$

□

To relate a numerical $n \times n$ matrix $M = \left(m_{ij} \right)$ to a structured matrix $\tilde{M} = \left(\tilde{m}_{ij} \right)$ we define $n = \{1, 2, \dots, n\}$, $m = \{1, 2, \dots, m\}$ and state the following:

Definition (2.): A numerical matrix M is said to be admissible with respect to a structured matrix $\tilde{M}$, that is, $M \in \tilde{M}$, if and only if $\tilde{m}_{ij} = 0$ implies $m_{ij} = 0$ for all $i \in n$ and $j \in m$.

□

Definition (2.9): Consider the linear system below:

$$S : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \quad (2.30)$$

A structured system $\tilde{S} = (\tilde{A}, \tilde{B}, \tilde{C})$ so that $(A, B, C) \in (\tilde{A}, \tilde{B}, \tilde{C})$. Structural controllability of the system $\tilde{S}$ is defined via the pair $(\tilde{A}, \tilde{B})$ as in (Lin, 1974).

□

Definition (2.10): A pair of matrices $(\tilde{A}, \tilde{B})$ is said to be structurally controllable if there exists a controllable pair (A, B) such that $(A, B) \in (\tilde{A}, \tilde{B})$.

□

Definition (2.11): A pair of matrices $(\tilde{A}, \tilde{C})$ is said to be structurally observable if there exists an observable pair (A, C) such that $(A, C) \in (\tilde{A}, \tilde{C})$.

□

In (Commault & Dion, 2013) and references therein, given the structure of a linear-time invariant system and a possible collection of inputs, the objective is to determine the minimum subset of inputs, which yields the structural controllability. (Commault & Dion, 2013) proposes solution methods to the above constrained minimal input selection problem rely on a two-step optimization procedure, which, in general, leads to suboptimal solutions. Determining feasible solutions to the constrained minimal input selection problem has been a major focus of (Dion, et al., 2003), (Murota, 2009). Similar work is presented in (Boukhobza & Hamelin, 2011), where the analysis of IO selection is investigated via ensuring structural observability for linear systems in descriptor form. The problem of identifying the optimal input/output, which yields the structural controllability/observability was also considered in (Liu, et al., 2011), (Dion, et al., 2002).

(Liu, et al., 2011) states that the minimum number of controlling agents required to achieve structural controllability is related to the number of right unmatched vertices of an associated bipartite matching problem.

However, these criteria have some disadvantages. They are binary qualitative concepts and tell only whether a set of actuators/sensors make the system controllable/observable or not, regardless of other considerations such as the energy required to actually drive the system around the state. Therefore, a measure of the degree of controllability/observability has to be defined to help choose the best IO set. This measure should have the following attributes (Abdel-Mooty & Roorda, 1994):

- It must vanish for uncontrollable/unobservable case
- It must indicate the effectiveness of the selected set of actuators/sensors, which is the ability of the actuators to induce the desired control effect, i.e. enhance the controllability properties, and the ability of the sensors to enhance the detection, i.e. improve the observability properties.
- It must indicate the control/observe cost as a measure of the effort made to achieve the required control/detect level
- It must reflect the control/observe objective

Definition (2.12): A recovery region in the state space was defined as the region that includes all of the initial conditions (or disturbed states) that can be returned to the origin in a finite time using the bounded control forces. The degree of controllability (DOC) was defined as a scalar measure of the size of the recovery region.

□

In other words, the degree of controllability will be determined by the minimum distance from the origin to a state that cannot be brought to the origin in a finite time t_f. More loosely, it is the minimum initial condition disturbance from which the system cannot recover in t_f.

Theorem (2.2): (Morris, 2001) The pair (A, B_a) defined in (2.1) is controllable at a if and only if $W_c^a(0, t_f)$ is positive definite for all $t_f > 0$.

□

where:

$$W_c^a(0, t_f) = \int_0^{t_f} e^{A\tau} B_a B_a^T e^{A^T \tau} d\tau \tag{2.31}$$

The proof is given as Theorem (2.5) in (Morris, 2001). To address the degree of controllability/observability issue various quantitative measures have been proposed in literature. Among the first attempts was the work of (Kalman, et al., 1963), in which a symmetric controllability matrix for time-varying linear systems was defined. That study further defined the determinant and the trace of controllability Gramian as scalar measures of the controllability, see e.g. (Leleu, et al., 2001). (Muller & Weber, 1972), (Arbel, 1981) added the minimum eigenvalue of the controllability Gramian as a third measure. Controllability and observability Gramian have some advantages when compared to controllability/observability matrices. The superior numerical properties of Gramians are due to their symmetry, semi-definiteness and compactness. An interpretation of the Gramians is the geometry of principal components of an ideal sensor/ actuator configuration. The controllability Gramian quantifies the energy required to move the system around the state space.

Relevant results are also found in (Pasqualetti, et al., 2014), where the authors study the controllability of a system with respect to the smallest eigenvalue of the controllability Gramian, and they derive a lower bound on the number of actuators so that this eigenvalue is lower bounded by a fixed value. Nonetheless, they do not provide an algorithm to identify the actuators that achieve this value. They propose a heuristic actuator placement procedure that does not constrain the number of available actuators and does not optimize their control energy objective. A small eigenvalue of the controllability Gramian matrix would lead to at least one state requiring very high control effort. This implies that all the eigenvalues of the controllability Gramian matrix should be as large as possible. (Georges, 1995) extended this idea to nonlinear systems. (Hac & Liu, 1993), (Qiu, et al., 2007) developed an optimization method for finding the optimal locations of piezoelectric actuators based on the degree of controllability using the Gramian matrix as follows:

$$\max_{a \in \Omega^m} \rho = \max_{a \in \Omega^m} \sum_{i=1}^{n} \lambda_i \left(\sqrt[n]{\prod_{i=1}^{n} \lambda_i} \right) \tag{2.32}$$

where m is the number of actuators with locations that could be varied over some compact set $\Omega \subset \mathbb{R}_a$. λ_i, $i = 1, \ldots, n$ denotes the ith eigenvalue of the controllability Gramian. The summation term in (2.32) is the trace of the Gramian matrix. To ensure that all the eigenvalues of the Gramian are high, the geometric mean of all the eigenvalues is included in the objective function. The optimal locations of multiple actuators calculated using (2.32) suppressed the vibrations in a plate (Peng, et al., 2005). An adaptive feedforward controller was designed in (Peng, et al., 2005)after calculating the optimal actuator location.

A term related to the standard deviation of Gramian eigenvalue, $\sigma^{-1}(\lambda_i)$, was often included in the performance index (2.32) as a product term, see e.g., (Jha & Inman, 2003), (Pulthasthan & Pota, 2008), where the objective function is defined as:

$$\rho = \left(\sum_{i=1}^{n} \lambda_i \right) \left(\sqrt[2n]{\prod_{i=1}^{n} \lambda_i} \right) \Big/ \sigma(\lambda_i) \tag{2.33}$$

Here λ_i could be the i[th] eigenvalue of the controllability gramian in the case of actuator locations or the observability gramian when sensor positions are considered. Factor $\left(\sqrt[2n]{\prod_{i=1}^{n} \lambda_i} \right) \Big/ \sigma(\lambda_i)$ in equation (2.5) are used to normalize the factor $\left(\sum_{i=1}^{n} \lambda_i \right)$ which is directly proportional to the total exchanged energy between the plant and actuators or sensors. The term $\left(\sqrt[2n]{\prod_{i=1}^{n} \lambda_i} \right)$ is the geometric mean of the ellipsoid axes length, $\sigma(\lambda_i)$ penalizes locations where there is a poorly controllable or observable state hidden by a highly controllable or observable state or there exist both very high and very small eigenvalues. Optimal sensor and actuator location is obtained at the position where the objective function ρ in equation (2.5) is a maximum.

(Muller & Weber, 1972), (Dochain, et al., 1997), (Waldraff, et al., 1998), (Van den Berg, et al., 2000) presented a sensor placement technique based on a suitable norm of the observability Gramian or the observability matrix. For optimal actuator placement (Choe & Baruh, 1992) minimizes the objective functions based on the entries of the actuator influence matrix, which gives general measures of controllability. (Devasia, et al., 1993), (Jha & Inman, 2003), (Bruant & Proslier, 2005) proposed the maximization of a controllability/observability criterion using the Gramian matrices as well. (Wang & Wang, 2001) suggested the maximization of the control forces transmitted by the actuators to the structure. (Dhuri & Seshu, 2006) proposed a modal controllability index based on the same singular value analysis of the control vector. In an attempt to define a degree of controllability, (Cheng & Pantelides, 1988) suggested a weighted sum of the squares of the modal displacements of seismic buildings at the actuator position, each multiplied by the maximum modal response spectrum value for the design earthquake. However, this criterion does not satisfy the basic requirement that the DOC vanishes when the system becomes uncontrollable.

For actuator placement (Kim & Junkinsf, 1991) introduced a combination of the squares of Hamdan and Nayfeh's modal controllability measures weighted by the respective modes' contributions to a quadratic output cost function. (Bagajewicz & Sanchez, 1999) presented a sensor placement technique with the goal of achieving a certain degree of observability or redundancy for a variable in a system. They introduced the degree of estimability of a variable by merging the concept of degree of redundancy for measurements and degree of observability for unmeasured variables into one single property. (Bruant, et al., 2010) used two modified optimization criteria, ensuring good observability or controllability of the structure, and considering residual modes to limit the spill-over effect. They considered two optimization variables for each piezoelectric device: the location of its centre and its orientation. They implied Genetic algorithms to find the optimal configurations. (Han & Lee, 1999) also used genetic algorithm (GA) to determine the sensor and actuator locations with the consideration of controllability, observability, and spill-over prevention. However, in their work, the dynamic characteristics of both rectangular plate and piezoelectric sensors/actuators were not derived explicitly.

(Lim, 1992) proposed a method based on the combined degree of controllability and observability to select the most suitable set of actuator and sensor locations, which is capable of simultaneously controlling and observing the system to a high degree. The controllability Gramian was computed and the controllable subspace for each actuator location was derived. The observability Gramian was also calculated and the observable subspace for each sensor location was determined. An intersection subspace was formed from the controllable and observable subspaces of each actuator/sensor pair. The proposed optimization cost functional did not incorporate closed-loop control; therefore, any control method could be applied after the actuator positions are determined. Additionally, the approach was computationally efficient because the problem size and design space are proportional to the number of actuators, unlike other methods where the design space is factorially related to the number of actuators. One disadvantage of the method is that a priori knowledge of the system is required.

The use of Gramians as quantitative metrics of controllability in networks is studied in (Rajapakse, et al., 2011), (Yan, et al., 2012), (Sun & Motter, 2013), (Tang, et al., 2012), (Pasqualetti, et al., 2014). Other important studies of controllability measures in networks include (Olshevsky, 2014), (Sorrentino, et al., 2007), (Rahmani, et al., 2009). They have studied the problem of leader selection in networks with consensus dynamics, in which a set of leader states are selected to act as control inputs to the system. Controllability Gramian based approach is an open-loop strategy. Calculation of optimal actuator location does not simultaneously provide a controller. For example, in (Pulthasthan & Pota, 2008), the design of optimal control and the optimal actuator location were treated as two separate problems which is cumbersome.

Definition (2.13): A set function $f : 2^V \to \mathbb{R}$ is called submodular if for all subsets $A \subseteq B \subseteq V$ and all elements $s \notin B$, it holds that:

$$f\left(A \cup \{s\}\right) - f(A) \geq f\left(B \cup \{s\}\right) - f(B) \tag{2.34}$$

or equivalently, if for all subsets:

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{2.35}$$

A set function is called super-modular if the reversed inequalities in (2.34) and (2.35) hold and is called modular if (2.34) and (2.35) hold with equality.

□

The authors in (Summers & Lygeros, 2014) showed that the mapping from possible actuator placements to the trace of the controllability Gramian is a modular set function and therefore a simple optimization can be implemented to compute the metric individually for all possible actuator placement combination, sorts the outcome, and selects the globally optimal subset that minimizes the Gramian energy metric (Summers, 2016). Several classes of Gramian metrics are shown to be submodular in (Summers, et al., 2016), (Summers & Lygeros, 2014), (Cortesi, et al., 2014), (Tzoumas, et al., 2016):

- $-trace(W_c^{-1})$
- $\log|W_c|$
- $rank(W_c)$

In addition, many other problems featuring super-modularity or sub-modularity are discussed in (Pasqualetti, et al., 2014), (Clark, et al., 2014), (Summers, et al., 2015), (Bushnell, et al., 2014), (Yan, et

al., 2015), (Shames & Summers, 2015). Although optimization of submodular functions is difficult, submodularity allows for an approximation guarantee if one uses a simple greedy heuristic for their optimization (Nemhauser, et al., 1978).

The approaches presented in (Summers & Lygeros, 2014) assume that the system is initially stable with strict applications to linear time-invariant systems with the controllability Gramian and its derivatives as the control metric.

(Juang & Rodriguez, 1979) regarded the framework of the optimal control theory for systems subjected to initial disturbances and defined the cost function that is to be minimized as an indication, which relies on the controllability and disturbance sensitivity Gramians, so closed-form solution can be found. Many studies to qualify controllability of a system with external disturbances have been carried out (Stanley, et al., 1985), (Morari, 1983), (Morari, et al., 1985), (Shimizu & Matsubara, 1985), (Skogestad & Morari, 1987), (Hovd & Skogestad, 1992), (Luyben, 1988), (Cao & Rossiter, 1996), (Cao, et al., 1997), (Mirza & Niekerk, 1999). However, these works are mainly defined in frequency domain and hard to be applied to a system with unmatched disturbance on output, though they have an advantage to calculating specific frequency range. Naturally DOC for disturbance rejection in time domain is defined by (Kang & Park, 2009) . (Lee & Park, 2014) extended this idea to unstable systems.

The optimal actuator placement in controlling large structures in space subjected to initial disturbances was considered in (Viswanathan, et al., 1979), (Lindberg & Longman, n.d.), (Longman & Lindberg, 1986), (Viswanathan & Longman, 1983) and (Longman & Horta, 1989). (Viswanathan, et al., 1979) developed some numerical methods for generating the degree of controllability and evaluating the effectiveness of candidate actuator distributions. The method is shown to take on a relatively simple form when spacecraft modal coordinates are used.

(Hughes & Skelton, 1980) used the norms of the rows of the control location matrix as measures of the controllability of the individual modes. (Trajkov & Nestorvic, 2012) analyzed the placement based on controllability and observability criteria. The optimization was conducted based on H2-norm and controllability and observability gramian function which is dependent on vibrational modes. The structure model was designed using finite element and after the reduction of order process, optimization operations were done on the reduced model and the optimal place was suggested for the plate and cantilever beam.

(Vilnay, 1981) and (Ibidapo-Obe, 1985) proposed a method based on the interpretation of the functional relationship (transfer matrix/control influence matrix) between the actuators and modes of the structural system. It is shown that, from the form of the matrix, the controllability and observability of the system with respect to differing locations of the sensors and actuators can be established. In an attempt to consider the control effort in the actuator placement, (Abdei-Rohman, 1984) suggested that the optimal distribution of the sensors and the actuators is that which minimizes the observer gains and the control gains for the same level of control. However, the control gains do not represent completely the control effort, which depends greatly on the type and configuration of the control mechanism. (Sinha, et al., 2013) derived an analytical expression for the finite time controllability and observability Gramian for advection PDE. The selection criteria for the optimal location of actuators and sensors was proposed based on the maximization of Gramians. (Safizadeh, et al., 2010) studied the best position of piezoelectric actuator for active vibration control using controllability Gramian performance index and genetic algorithm. Their proposed method optimizes the controllable performance index in order to minimize the output energy of actuator for limited modes of frequency. In this method, the main responsibility is system's controllability and expressing

an optimum control input so that by applying forces on this optimal place of structure, system, can be damped. (Yang & Chen, 2010), (Yang & Chen, 2010 b) studied the optimal placement of piezoelectric sensor and actuator on a plate by increasing the controlling performance of system. In their work, the performances of two algorithms are used and compared to achieve better optimization result among variables of vibration response deformation precision, control energy and number of actuators. The utilized algorithms are GATSP (Genetic for the TSP) algorithm and HTTSP (Hopfield-Tank for the TSP) algorithm.

2.5. Summary

In this chapter, a brief survey based on some most popular optimization strategies that were used extensively in engineering literature for actuator/sensor placement is done.
An overview of the problem of optimal actuator/sensor selection is presented and some energy-based criteria are described.
The originality of this thesis is to use the reviewed concepts within a systematic framework in order to select the best actuator/sensor location with which the optimum performance is achieved while the required control energy is minimized. A new evolutionary algorithm is developed and merged within the proposed methodologies for the purpose of the control energy optimisation for each actuator/sensor set. The next chapter discusses the controllability Gramian approach in more detail.

Chapter 3: Overview of Controllability Gramian and its Computation

3.1. Introduction

The principal goals of this chapter are to introduce the system properties of controllability and controllability Gramian (and of reachability, reachability Gramian), which play a central role in the study of energy consumption and input/output structure selection, topics that will be studied in the following chapters.

Controllability refers to the ability to manipulate the state by applying appropriate inputs (in particular, by steering the state vector from an initial vector value to a final vector value in finite time). Such is the case, for example, in satellite attitude control, where the satellite must change its orientation by manipulating its thrusters.

In Section 3.2, the concepts of reachability and reachability Gramian are introduced for linear continuous time-invariant systems.

Controllability and controllability Gramians are treated in detail in Section 3.3, where a novel approach of calculation of controllability Gramian is presented.

In the last section, we introduce a new approach for selecting input matrix B, which can assign a desired controllability Gramian to the LTI system.

3.2. Reachability and Reachability Gramian

Consider the LTI state space system:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $u(t) \in \mathbb{R}^m$, Then the state at time t_f is given by:

$$x(t_f) = \Phi(t_f, t_0)x_0 + \int_{t_0}^{t_f} \Phi(t_f, \tau)Bu(\tau)d\tau \tag{3.2}$$

where $\Phi(t,\tau)$ is the state transition matrix of the system, and $x(t_0) = x_0$ is the initial state.

In the time-invariant case considered here the state transition matrix can be defined as:

$$\Phi(t,\tau) = \Phi(t-\tau, 0) = e^{A(t-\tau)} \tag{3.3}$$

Here we are interested in using the input to transfer the state from x_0 to some other value $x(t_f) = x_f$ at some finite time $t_0 \le t_f$. Note that here because of time invariance, only the difference $t_f - t_0 = T$, is relevant (rather than the individual times t_0 and t_f) and we can always take $t_0 = 0$ and $t_f = T$.

Considering (3.3), equation (3.2) can be written as:

$$x_f = e^{AT}x_0 + \int_0^T e^{A(T-\tau)}Bu(\tau)d\tau \tag{3.4}$$

Clearly, there exists $u(t)$, $0 \le t \le T$, that satisfies (3.4) if and only if such transfer of the state is possible. Also, if we define $\hat{x}_f = x_f - e^{AT}x_0$, it follows that the control input $u(t)$, which transfers the states from x_0 at time 0, to x_f at time T will also cause the state to reach $\hat{x}_f$ at T, starting from the origin at time zero.

Definition (3.1): A state x_f is reachable if there exists an input $u(t)$, $0 \le t \le T$, that transfers the states $x(t)$ from the origin at 0 to x_f over some finite time $t_f = T$. A reachable state x_f is sometimes also called controllable from the origin (Antsaklis & Michel, 2007).

□

Definition (3.2): The set of all reachable states $\Re_r$ is the reachable subspace of the LTI system (3.1), or of the pair (A, B). Note that the set of all reachable states x_f contains the origin and constitutes a linear subspace of the state space ($\mathbb{R}^n$, $\mathbb{R}$) (Antsaklis & Michel, 2007).

□

Definition (3.3): The LTI system (3.1), or the pair (A, B) is completely reachable if every state is reachable, i.e., if $\Re_r = \mathbb{R}^n$ (Antsaklis & Michel, 2007).

□

Definition (3.4): For the linear continuous-time invariant system (3.1), the $n \times n$ reachability Gramian is (Lewis, et al., 2012) (Antsaklis & Michel, 2007):

$$W_r(0,T) = \int_0^T e^{(T-\tau)A} B B^* e^{(T-\tau)A^*} d\tau \tag{3.5}$$

□

Lemma (3.1): (Antsaklis & Michel, 2007) W_r is a symmetric, positive semidefinite matrix, for every $T > 0$; i.e., $W_r = W_r^*$ and $W_r \geq 0$.

□

One can also show that the reachable subspace of LTI system (3.1) is exactly the range of the reachability Gramian W_r. This is described as Theorem 3.1, but before that it is worth mentioning that the range of $W_r(0,T)$, is independent of T, it is the same for any finite $T > 0$, and in particular, it is equal to the range of the controllability matrix C, it is stated in Corollary (3.1) below:

Corollary (3.1): (Antsaklis & Michel, 2007) $\Re(W_r(0,T)) = \Re(C)$ for every $T > 0$.

(Note: In discrete case, the reachable (and/or controllable) subspace does not depend on the finite time only when the interval has length larger than or equal to n time steps, when $0 < T < n$, we have $\Re(W_r(0,T)) \subset \Re(C)$ where $C = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$ (Hespanha, 2009).)

□

Theorem (3.1): (Antsaklis & Michel, 2007) Consider the LTI system (3.1), and let the initial condition $x(0) = 0$. There exists an input u that transfers the state $x_0 = 0$ to x_f in finite time T >0 if and only if $x_f \in \Re(C)$, or equivalently, if and only if $x_f \in \Re(W_r(0,T))$. Thus, the reachable subspace $\Re_r = \Re(C) = \Re(W_r(0,T))$. Furthermore, an appropriate u that accomplishes this transfer in time T is given by:

$$u(t) = B^* e^{(T-t)A^*} \eta_1 \tag{3.6}$$

with η_1 such that $W_r(0,T)\eta_1 = x_f$ and $0 \leq t \leq T$.

□

Note that here there is no restriction on time T, other than it has to be finite, so T can be as small as we wish; i.e., the transfer can in theory be accomplished in arbitrarily short time, though we will discuss later that this may be infeasible in practice as it may require control signals of arbitrarily large

magnitude or energy. Thus, it has its own costs, leading to a trade-off due to some physical restrictions, which appears in practical cases.

Next we introduce Corollary (3.2), which is a well-known result that has been used explicitly or implicitly by many authors (Farina & Rinaldi, 2011) (Antsaklis & Michel, 2007) (Guzzella, 2011) (Hespanha, 2009).

Corollary (3.2): The LTI system (3.1), or the pair (A, B) is completely reachable, if and only if:

$$Rank\ \mathrm{C} = n \tag{3.7}$$

or equivalently, if and only if:

$$Rank\, W_r(0,T) = n \tag{3.8}$$

for any finite $T > 0$.

□

Lemma (3.2): (Fabrizio L. Cortesi, Tyler H. Summers, and John Lygero, 2014) Let the LTI system (3.1) or pair (A, B) be completely reachable. Then, there exists an input that will transfer any state x_0 to any other state x_f in some finite time $T > 0$. One such input is given by:

$$u(t) = B^* e^{(T-t)A^*} W_r^{-1}(0,T)[x_f - e^{AT} x_0] \quad , \ 0 \le t \le T \tag{3.9}$$

□

In fact, this input control is not unique; there exist many various control inputs $u(t)$ transferring the state x_0 to the state x_f , in time T , but it can be shown that the control input $u(t)$ given by (3.9), accomplishes this transfer while expending a minimum amount of energy, i.e. it minimizes the cost function $\int_0^T \|u(\tau)\|^2 \, d\tau$, where $\|u(t)\|^2 = u^*(t)u(t)$ denotes the Euclidean norm of $u(t)$.

This result is discussed in full in subsequent chapters, where the effect of the terminal time T and the terminal state x_f on the value of the cost function will also be investigated.

3.3. Controllability Gramian and its Computation

Definition (3.5): A state x_0 is controllable if there exists an input $u(t)$, which transfers the LTI system (3.1) from x_0 at t=0 to the origin in some finite time $T > 0$ (Mellodge & Kachroo, 2008) (Lewis, et al., 2003).

□

Definition (3.6): The set of all controllable states $\Re_c$ is the controllable subspace of the LTI system (3.1) or of the pair (A, B) (Antsaklis & Michel, 2007).

□

Definition (3.7): The continuous linear time invariant system (3.1), or the pair (A, B) is completely state controllable if every state is controllable, i.e., if $\Re_c = \mathbb{R}^n$, i.e. there exists an input $u(t)$ that can transfer the states of the system from an arbitrary initial state x_0 to a final state in a finite interval of time. This definition was given first by Kalman (Kalman, et al., 1963) and has been used by many authors; i.e. (Mellodge & Kachroo, 2008) (Owens, 2015).

□

We now establish the relationship between reachability and controllability for the LTI system (3.1). In view of (3.4), x_0 is controllable when there exists $u(t)$, $0 \le t \le T$, so that:

$$-e^{AT}x_0 = \int_0^T e^{(T-\tau)A}Bu(\tau)d\tau \tag{3.10}$$

Or when $e^{AT}x_0 \in \Re(W_r(0,T))$, or according to Corollary (3.1),

$$e^{AT}x_0 \in \Re(C) \tag{3.11}$$

Recall that x_f is reachable when $x_f \in \Re(C)$. In fact, in the continuous-time case there is no distinction between reachability and controllability, in contrast to discrete-time case, since the state transition matrix is e^{AT} (rather than A^k) and is always invertible.

Lemma (3.3): If $x \in \Re(C)$, then $Ax \in \Re(C)$; i.e., the reachable subspace $\Re_r = \Re(C)$ is A-invariant.

□

The proof is given in (Levine, 1996).

Corollary (3.3): the reachability subspace $\Re_r$ is the smallest A-invariant subspace containing $\mathrm{Im}[B]$.

□

For the proof see (Bittanti & Colaneri, 2009), p. 114.

Theorem (3.2): Consider the LTI system (3.1),

(i) A state x is reachable if and only if it is controllable.
(ii) $\Re_c = \Re_r$.
(iii) The system (3.1), or the pair (A, B), is completely state reachable if and only if it is completely state controllable.

□

For the proof see Theorem (5.20) in (Antsaklis & Michel, 2007). The following Theorem explains the connection of reachability with controllability in the LTI system. Note: Although this Theorem always holds for continuous LTI systems in the discrete time case the Theorem is only valid when the system matrix A is non-singular.

Definition (3.8): (Antsaklis & Michel, 2007) The controllability Gramian in the linear continuous time-invariant system (3.1) is $n \times n$ matrix:

$$W_c(0,T) = \int_0^T e^{A\tau} BB^* e^{A^*\tau} d\tau \tag{3.12}$$

□

It can be verified directly that:

$$W_r(0,T) = -e^{TA} W_c(0,T) e^{TA^*} \tag{3.13}$$

The controllability Gramian is widely used to check the controllability of the linear dynamical systems.

Proposition (3.1): The LTI system (3.1) is controllable:

(i) If and only if $rank(W_c(0,T)) = n$ for some finite $T > 0$

(ii) If and only if $rank(C) = n$, where $C = [B, AB, ..., A^{n-1}B]$

(iii) If and only if $rank([\lambda_i I - A \vdots B]) = n$, where λ_i, $i = 1, 2, .., n$, is an eigenvalue of A.

□

The proof is given in (Klamka, 2009).

Controllability Gramian is related to the energy necessary to transfer the initial state to one final state (Moore, 1981). The Gramians are widely studied in the process of model order reduction (Moore, 1981), (S. Gugercin and A.C. Antoulas, 2003), (Antoulas, 2004).

In the next section, we introduce a novel approach for the computation of controllability Gramian, which presents many advantages not only in decreasing the complexity of the calculation but also in investigating the effect of model uncertainty or actuator noise. It also helps in input energy optimization problem when we are selecting the input matrix B.

3.3.1. A Novel Approach for Calculation of Controllability Gramian

In this part, we introduce a new approach of calculating the controllability Gramian matrix $W_c(0,t_f)$.

The key innovative idea of this method is to split the controllability Gramian $W_c(0,t_f)$ into two distinct independent parts: one part deals with the input matrix B and the other part is a function of the system matrix A. We will show that in the presence of uncertainty, and perturbation, this method helps to examine the effects of the matrices A and B on controllability Gramian (3.12).

By this method, the controllability Gramian can be calculated faster and easier, however it is noteworthy that in the case of large scale systems or multi input systems, this method may result in some large matrices while it still preserves its computation advantages. It is shown later that in the problem of input selection or finding the optimal placement of the actuators we are usually concerned with an eigenvalue optimization problem or a problem of maximization of the trace of controllability Gramian. This means that in most practical cases we do not need to calculate the entire controllability Gramian. Thus, especially in the case of large scale systems our method has significant advantages. This method has been developed by applying the concept of quadratic form matrices and sum of

squares polynomials (SOS polynomials) (Chesi, 2011), (D. Henrion; A. Garull, 2005), (A.A. Ahmadi; P.A. Parrilo, 2009), (Parrilo, 2000), (Parrilo, 2003).

Note: We denote by $p(x) := p[x_1, \ldots, x_n]$ the ring of polynomials in n variables with coefficients in the field $\mathbb{R}$.

Proposition (3.2): (Parrilo, 2003), (Parrilo, 2000) A multivariate polynomial $p(x) \in \mathbb{R}$ in n variables and of degree 2d is a sum of squares (SOS) if and only if there exists a real symmetric positive semidefinite matrix Q such that:

$$p(x) = z^T Q z \tag{3.14}$$

where z is the vector of monomials of degree up to d, we call matrix Q the Gram matrix.

$$Z = [1, x_1, x_2, \ldots, x_n, x_1 x_2, \ldots, x_n^d]^T \tag{3.15}$$

Given a Gram matrix Q, of $rank(Q) = t$, we can construct polynomials $h_1, h_2, \ldots, h_t$, such that:

$$p(x) = \sum_{i=1}^{t} h_i^2 \tag{3.16}$$

□

The construction is given in the Appendix. We also refer reader to (Powers & Wörmann, 1998).

Thus, to find a representation of $p(x)$ as a sum of squares, we just need to find a matrix Q, which satisfies the above relation.

As the controllability Gramian definition (3.8) suggests, the most difficult part of the related computation, which gives rise to complexity, is related to the computation of the integral, in (3.12). thus, the simpler the matrix inside is to integrate the easier and faster are the calculations.

Inspired from quadratic form associated matrices, we introduce Theorem (3.3) below. First, we give two definitions and an intermediate research.

Definition (3.9): (Lam, 2005) Let A be an $n \times n$ symmetric matrix with real entries, and let x be an n vector over the field $\mathbb{R}$. The mapping $Q : x \to x^T A x$ is the quadratic form defined by A.

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

□

According to the definition above for any multivariate polynomial $p(x) \in F$ in n variables and of degree 2d, there exist an associated symmetric matrix $A \in F^{n \times n}$, with the entries:

$$a_{ij} = \frac{coefficient(x_i x_j) + coefficient(x_j x_i)}{2}, 1 \le i, j \le n, i \ne j$$

$$a_{ii} = coefficient(x_i^2), 1 \le i \le n$$

Now regarding Proposition (3.2), we define Lemma (3.4) below whose proof is omitted.

Lemma (3.4): A polynomial $p(x) \in F$ is a sum of squares (SOS) polynomial if and only if its associated matrix $A \in F^{n \times n}$ is a positive semi definite matrix.

□

Definition (3.10): (Loan, 2000) For any matrix $A \in \mathbb{R}^{n \times m}, B \in \mathbb{R}^{p \times q}$ the Kronecker product (i.e., the direct product or tensor product), denoted as $A \otimes B$, is defined by:

$$A \otimes B = [a_{ij}B] = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{bmatrix} \in \mathbb{R}^{np \times mq} \tag{3.17}$$

□

Remark (3.1): (Loan, 2000) The following basic properties easily follow:

i. $(A \otimes B)^T = A^T \otimes B^T$

ii. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$

iii. $(A \otimes B)(C \otimes D) = (AC \otimes BD)$

iv. $trace(A \otimes B) = trace(A)trace(B)$

□

Theorem (3.3): For *n* dimensional LTI system (3.1) with m inputs, the controllability Gramian $W_c(0, t_f) = \int_0^{t_f} e^{A\tau} BB^* e^{A^*\tau} d\tau$, can also be written as:

$$W_c(0, t_f) = (I_n \otimes \Upsilon)^T Q (I_n \otimes \Upsilon) \tag{3.18}$$

Where Υ is the vectorised input matrix and $Q \in \mathbb{R}^{n^2 m \times n^2 m}$ is a real symmetric matrix.

□

Proof:

For simplicity and without loss of generality we consider the case of a 3-dimensional LTI system in (3.1), with two inputs. Let:

$$e^{A\tau} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3.19}$$

According to (3.12), we can define the controllability Gramian as:

$$
\begin{aligned}
W_c(0,t_f) &= \int_0^{t_f} e^{A\tau} BB^* e^{A^*\tau} d\tau \\
&= \int_0^{t_f} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} BB^* \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^* d\tau \\
&= \int_0^{t_f} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}^* \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^* d\tau
\end{aligned}
\tag{3.20}
$$

Clearly, we can simplify (3.20) as:

$$
\left(W_c(0,t_f)\right)_{col,row} = \int_0^{t_f} \sum_{i=1}^{3} a_{col,i} \sum_{j=1}^{3} a_{row,j} B_{ji} d\tau \qquad , B_{ji} = \sum_{l=1}^{2} b_{jl} b_{il} \quad , col, row = 1,2,3 \tag{3.21}
$$

In the same way, for a general n dimensional LTI system with m inputs the controllability Gramian matrix could be defined as:

$$
\left(W_c(0,t_f)\right)_{col,row} = \int_0^{t_f} \sum_{i=1}^{n} a_{col,i} \sum_{j=1}^{n} a_{row,j} B_{ji} d\tau \qquad , B_{ji} = \sum_{l=1}^{m} b_{jl} b_{il} \quad , col, row = 1,\ldots,n \tag{3.22}
$$

Now using this fact that the b 's are independent monomials with respect to τ , one can readily consider $\left(W_c(0,t_f)\right)_{col,row}$, $col, row = 1,\ldots,n$, as a quadratic polynomial, which can be written as:

$$
\left(W_c(0,t_f)\right)_{col,row} = [b_{j1},\ldots,b_{jm}] \int_0^{t_f} \sum_{i=1}^{n} a_{col,i} \sum_{j=1}^{n} a_{row,j} d\tau \begin{bmatrix} b_{i1} \\ \vdots \\ b_{im} \end{bmatrix} \qquad col, row = 1,\ldots,n \tag{3.23}
$$

Then by defining the monomial matrix $\Upsilon = [\underbrace{b_{11}, b_{21}, \ldots, b_{n1}}_{1^{st}\ input}, \underbrace{b_{12}, b_{22}, \ldots b_{n2}}_{2^{nd}\ input}, \ldots, b_{nm}]^*$ the controllability Gramian matrix could be defined as:

$$
\begin{aligned}
W_c(0,t_f) &= \mathrm{T}^* \overbrace{\int_0^{t_f} \sum_{i=1}^{n} a_{corresponding\_col,i} \sum_{j=1}^{n} a_{corresponding\_row,j} d\tau}^{Q_{n^2m\times n^2m}} \mathrm{T} \\
\mathrm{T}_{n^2m\times n} &= \begin{bmatrix} b_{11},\ldots,b_{n1},b_{12},\ldots,b_{n2}\ldots,b_{nm},\bar{0}_{nm},\ldots,\bar{0}_{nm} \\ \bar{0}_{nm},b_{11},\ldots,b_{n1},b_{12},\ldots,b_{n2}\ldots,b_{nm},\ldots,\bar{0}_{nm} \\ \bar{0}_{nm},\bar{0}_{nm},b_{11},\ldots,b_{n1},\ldots,b_{n2}\ldots,b_{nm},\ldots,\bar{0}_{nm} \\ \vdots \\ \bar{0}_{nm},\bar{0}_{nm},\ldots,b_{11},\ldots,b_{n1},b_{12},\ldots,b_{n2}\ldots,b_{nm} \end{bmatrix}^* = (I \otimes \Upsilon)
\end{aligned}
\tag{3.24}
$$

and the Theorem (3.3) is proved.

□

Example (3.1): Assume an LTI system as:

$$A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, B=\begin{bmatrix}0\\ 1\end{bmatrix}$$

A practical example of this system is the linearized pitch motion equation of a rigid spacecraft or a lunch vehicle $I\ddot{\theta}=-TL\sin\delta$ where T is the rocket thrust, L is the distance between the center of mass and thrust, δ is the gimbal angle, θ is the pitch angle and I is the pitch moment of inertia. Consider the transfer of the states from the origin to a sphere of the radius one and center the origin within one second, i.e. $t_f=1$. In this case, according to Theorem (3.3) we have:

$$W_c(0,t_f)=\int_0^{t_f} e^{A\tau}BB^* e^{A^*\tau}d\tau$$

Let us define:

$$e^{A\tau}=e^{\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\tau}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix},\ B=\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}b_{11}\\ b_{21}\end{bmatrix}$$

Then we have:

$$W_c(0,1)=\int_0^1\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}b_{11}\\ b_{21}\end{bmatrix}\begin{bmatrix}b_{11} & b_{21}\end{bmatrix}\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}^* d\tau$$

$$=\int_0^1\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}b_{11}^2 & b_{11}b_{21}\\ b_{11}b_{21} & b_{21}^2\end{bmatrix}\begin{bmatrix}a_{11} & a_{21}\\ a_{12} & a_{22}\end{bmatrix}d\tau$$

$$=\int_0^1\begin{bmatrix}a_{11}b_{11}^2+a_{12}b_{11}b_{21} & a_{11}b_{11}b_{21}+a_{12}b_{21}^2\\ a_{21}b_{11}^2+a_{22}b_{11}b_{21} & a_{21}b_{11}b_{21}+a_{22}b_{21}^2\end{bmatrix}\begin{bmatrix}a_{11} & a_{21}\\ a_{12} & a_{22}\end{bmatrix}d\tau$$

Thus:

$$W_c(0,1)=\int_0^1\begin{bmatrix}{a_{11}}^2b_{11}^2+a_{12}a_{11}b_{11}b_{21}+a_{12}a_{11}b_{11}b_{21}+{a_{12}}^2b_{21}^2 & a_{11}a_{21}b_{11}^2+a_{12}a_{21}b_{11}b_{21}+a_{11}a_{22}b_{11}b_{21}+a_{12}a_{22}b_{21}^2\\ a_{21}a_{11}b_{11}^2+a_{22}a_{11}b_{11}b_{21}+a_{21}a_{12}b_{11}b_{21}+a_{22}a_{12}b_{21}^2 & {a_{21}}^2b_{11}^2+a_{22}a_{21}b_{11}b_{21}+a_{21}a_{22}b_{11}b_{21}+{a_{22}}^2b_{21}^2\end{bmatrix}d\tau$$

$$=\begin{bmatrix}b_{11} & b_{21} & 0 & 0\\ 0 & 0 & b_{11} & b_{21}\end{bmatrix}\int_0^1\begin{bmatrix}{a_{11}}^2 & a_{12}a_{11} & a_{11}a_{21} & a_{12}a_{21}\\ a_{12}a_{11} & {a_{12}}^2 & a_{11}a_{22} & a_{12}a_{22}\\ a_{21}a_{11} & a_{22}a_{11} & {a_{21}}^2 & a_{22}a_{21}\\ a_{21}a_{12} & a_{22}a_{12} & a_{21}a_{22} & {a_{22}}^2\end{bmatrix}d\tau\begin{bmatrix}b_{11} & b_{21} & 0 & 0\\ 0 & 0 & b_{11} & b_{21}\end{bmatrix}^*$$

Therefore considering:

$$\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}=e^{\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\tau}=\begin{bmatrix}1 & \tau\\ 0 & 1\end{bmatrix},\begin{bmatrix}b_{11}\\ b_{21}\end{bmatrix}=\begin{bmatrix}0\\ 1\end{bmatrix}$$

we obtain:

$$W_c(0,1)=\begin{bmatrix}0&1&0&0\\0&0&0&1\end{bmatrix}\int_0^1\begin{bmatrix}1&\tau&0&0\\\tau&\tau^2&1&\tau\\0&1&0&0\\0&\tau&0&1\end{bmatrix}d\tau\begin{bmatrix}0&1&0&0\\0&0&0&1\end{bmatrix}^*$$

Now by calculating the integral we get:

$$W_c(0,1)=\begin{bmatrix}0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0.5&0&0\\0.5&0.333&1&0.5\\0&1&0&0\\0&0.5&0&1\end{bmatrix}\begin{bmatrix}0&1&0&0\\0&0&0&1\end{bmatrix}^*=\begin{bmatrix}0.333&0.5\\0.5&1\end{bmatrix}$$

□

Definition (3.11): A symmetric polynomial matrix $P(x)\in\mathbb{R}[x]^{m\times m}$, $x\in\mathbb{R}^n$ is SOS matrix if there exists a polynomial matrix $M(x)\in\mathbb{R}[x]^{s\times m}$ for some $s\in\mathbb{N}$, such that $P(x)=M^T(x)M(x)$.

□

If $m=1$ we have a standard SOS polynomial. Also, when P is a constant matrix, then the condition simply states that P is positive semidefinite matrix. Thus, SOS matrices generalize both positive semidefinite (constant) matrices and SOS polynomials. (Blekherman, et al., 2013)

Remark (3.2): The controllability Gramian $W_c(0,t_f)$ is a SOS matrix in input matrix B and hence can be written as $W_c(0,t_f)=Q^T(x)Q(x), x=[1;vec(B_{n\times m})]\in\mathbb{R}^{(nm+1)\times 1}$.

□

In the sequel, we will see that remark (3.2) simplifies energy minimization and optimal input selection achieved by maximizing the eigenvalues of the controllability Gramian, which are generally nonlinear functions of its entries.

3.3.2. Controllability Gramian Assignment with Respect to Control Variables

One motivating factor behind this section is the fact that many system properties are related to the controllability Gramian. Moreover, the eigenvalues and their corresponding eigenvectors of the controllability Gramian determine the speed and direction of energy dissipation, a problem, which will be studied in more detail in the following chapters. The basic idea behind this part is to find the optimal control input matrix that assigns a specific controllability Gramian to a LTI system.

In particular, we propose a solution to this assignability problem and specify the best possible set of input matrices B to steer the system to a target state, under energy limitations imposed by a given controllability Gramian.

Theorem (3.4): In the LTI system (3.1) a specified controllability matrix $W_c(0,t_f)$ is assignable if and only if there exist an input matrix B such that:

$$\begin{cases} \Phi_{n^2\times n^2} vec(BB^T) = vec(W_c(0,t_f)), B \in \mathbb{R}^{n\times m} \\ \Phi = \int_0^{t_f} (e^{A\tau} \otimes e^{A\tau}) d\tau = \int_0^{t_f} (e^{(A\oplus A)\tau}) d\tau \end{cases} \quad (3.25)$$

Where m is the number of control inputs.

□

Proof:

Consider the general definition of controllability Gramian defined in this chapter:

$$W_c(0,t_f) = \int_0^{t_f} e^{A\tau} BB^T e^{A^T\tau} d\tau$$

Now based on the Kronecker properties (Loan, 2000) (Jain, 1989) (Petersen & Pedersen, 2012) (Brookes, 2004) we have:

$$vec(A_{n\times m} X_{m\times p} B_{p\times q}) = (B^T \otimes A)_{qn\times pm} vec(X)_{pm\times 1} \quad (3.26)$$

Then we obtain:

$$\begin{cases} \Phi_{n^2\times n^2} vec(BB^T) = vec(W_c(0,t_f)), B \in \mathbb{R}^{n\times m} \\ \Phi = \int_0^{t_f} (e^{A\tau} \otimes e^{A\tau}) d\tau \end{cases} \quad (3.27)$$

and the proof is completed.

□

An important consequence of the above Theorem is that it relates the problem of controllability Gramian assignment to a linear system equations where the matrix of coefficients is itself a Kronecker product. Note that the linear system $Ax = b$ is efficiently solvable if A is a Kronecker product.

For the n dimensional system $e^{A\tau} \in \mathbb{R}^{n\times n}$, and $vec(BB^T)$ can be obtained in $O(n^3)$ flops via the LU factorization of $e^{A\tau}$. Without the exploitation of structure, an $n^2 \times n^2$ system would normally require $O(n^6)$ flops to solve (Loan, 2000).

Now considering the fact that both matrices $W_c(0,t_f)$ and BB^T are symmetric, one may readily simplify (3.27) as:

$$\int_0^{t_f} (e^{A\tau} \otimes e^{A\tau}) d\tau P_{BB^T} vec_{sym}(BB^T) = P_{W_c(0,t_f)} vec_{sym}(W_c(0,t_f)) \quad (3.28)$$

where $P_{BB^T} \in \{0,1\}^{n^2\times\frac{n(n+1)}{2}}$, $P_{W_c(0,t_f)} \in \{0,1\}^{n^2\times\frac{n(n+1)}{2}}$ denote appropriate permutation matrices.

$vec_{sym}(BB^T) \in \mathbb{R}^{\frac{n(n+1)}{2}\times 1}, vec_{sym}(W_c(0,t_f)) \in \mathbb{R}^{\frac{n(n+1)}{2}\times 1}$ are obtained by stacking the sub-diagonal columns of BB^T and $W_c(0,t_f)$ respectively.

This results in a system of linear equations with the special coefficient matrix:

$$\int_0^{t_f} (e^{A\tau} \otimes e^{A\tau}) d\tau P_{BB^T}$$

consisting of $\frac{n(n+1)}{2}$ equations and $\frac{n(n+1)}{2}$ variables, which can be solved almost twice as fast as a system without the symmetry property.

Assume that we are given a designed controllability Gramian according to the energy restrictions on the system. Then the assignability problem is defined as selecting the proper set of input matrices B such that (3.28) is satisfied, which assigns the given controllability Gramian to the system. To do so we just need easily to solve the linear equations based on B.

The difficulty of this approach is that the desired Gramian may not be assignable, i.e. there may be no solution to the linear equations. In this case, one must choose an input matrix B from the set which will closely approximate the desired Gramian. One approach is to minimize the matrix norm of the difference between the desired controllability Gramian and the assignable approximation to this matrix. According to (Wicks & Decarlo, 1990) it is not clear, however, that minimizing some norm of the error matrix implies a good approximation to the eigenvalues and eigenvectors of the desired matrix, nor is it clear that a large error necessarily means a bad approximation. An alternative method could be to approximate the controllability Gramian using least squares by minimizing the Frobenius norm of the error matrix over our variables, which are the elements of $vec_{sym}(BB^T)$.

The following example clarifies the method.

Example (3.2): Consider the single input system below:

$$A = \begin{bmatrix} -2 & -1 \\ 2 & 0 \end{bmatrix}, B = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

using Lyapunov equations or through the method proposed in this chapter we obtain

$$W_c(0,\infty) = \begin{bmatrix} 1.1250 & -0.2500 \\ -0.2500 & 0.7500 \end{bmatrix}.$$

Now assume that we are looking for the input matrix $B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ such that it assigns the controllability Gramian $W_c(0,\infty) = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$ to the system.

According to the above Theorem we have:

$$\int_0^{t_f}(e^{A\tau}\otimes e^{A\tau})d\tau P_{BB^T}vec_{sym}(BB^T)=P_{W_c(0,t_f)}vec_{sym}(W_c(0,t_f))$$

$$\Rightarrow\begin{bmatrix}1/4 & 0 & 0 & 1/8\\ 0 & 1/4 & -1/4 & -1/4\\ 0 & -1/4 & 1/4 & -1/4\\ 1/2 & 1/2 & 1/2 & 3/4\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}b_1^2\\ b_1b_2\\ b_2^2\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}W_1\\ W_2\\ W_3\end{bmatrix}$$

As explained above, due to symmetry the2nd and 3rd rows are identical and one of them can be removed. Then the set of all assignable controllability Gramians in this system is determined by S_W and clearly the given controllability Gramian is not assignable to this system. This we must choose a matrix from the set S_W which will closely approximate the desired controllability Gramian.

$$S_W=\{\begin{bmatrix}W_1 & W_2\\ W_2 & W_3\end{bmatrix}:B=\begin{bmatrix}b_1\\ b_2\end{bmatrix}\in\mathbb{R}^{2\times1}\}$$

Minimizing the Frobenius norm of the difference between the desired matrix and the assignable approximation to this matrix results in:

$$B=\begin{bmatrix}3.1779\\ -1.2195\end{bmatrix},\tilde{W}_c=\begin{bmatrix}2.7106 & -0.3718\\ -0.3718 & 2.2894\end{bmatrix}$$

Note that minimizing the spectral norm in this case also leads to the same result.

The eigenvalues of approximated controllability Gramian are: $\lambda_1=2.9257,\lambda_2=2.0743$ which are close to the eigenvalues of the desired Gramian.

□

Next consider a similar example to the multi input case.

Example (3.3): Consider the 3-dimensional system with 2- inputs:

$$A=\begin{bmatrix}0 & 0 & 2\\ 0 & 0 & 4\\ 0 & 3 & 1\end{bmatrix},B=\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\end{bmatrix}$$

The eigenvalues of this system are: $\lambda_1=-3,\lambda_2=0,\lambda_3=4$. Hence the system is unstable and $W_c(0,t_f)$ becomes unbounded as $t_f\to\infty$ so the Lyapunov equations do not produce proper solution for the controllability Gramian. Hence, we apply the proposed method for $t_f=1$. In this case:

$$W_c(0,1)=\begin{bmatrix}26.4638 & 21.7711 & 21.3257\\ 21.7711 & 69.3133 & 34.2409\\ 21.3257 & 34.2409 & 67.8167\end{bmatrix}$$

with the eigenvalues: $\lambda_1=15.7912,\lambda_2=34.3159,\lambda_2=113.4868$.

Next, we define the problem of finding the input matrix:

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

which assigns the controllability Gramian:

$$W_c(0,1) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Taking a similar approach to the single input case, it easily follows that the given controllability Gramian is not in the set of all assignable Gramians S_W and therefore we need to use an approximation.

By minimizing the spectral norm, we obtain:

$$B = \begin{bmatrix} \text{-}1.5844 & -0.0243 \\ 0.0004 & 2.3652 \\ \text{-}0.0235 & \text{-}1.6908 \end{bmatrix}, \tilde{W}_c = \begin{bmatrix} 3.1329 & \text{-}0.2907 & 0.3252 \\ \text{-}0.2907 & 2.1759 & 0.2827 \\ 0.3252 & 0.2827 & 1.1245 \end{bmatrix}$$

The eigenvalues of the approximating matrix are: $\lambda_1 = 3.2422, \lambda_2 = 2.2082, \lambda_{13} = 0.9829$

If the Frobenius norm of the difference matrix is minimized, we get the optimal input matrix as:

$$B = \begin{bmatrix} 1.7100 & \text{-}0.4690 \\ 0.0018 & \text{-}2.4668 \\ 0 & 1.7750 \end{bmatrix}$$

The corresponding approximate Gramian and its eigenvalues are:

$$\tilde{W}_c = \begin{bmatrix} 3.0019 & 0.0015 & 0.0165 \\ 0.0015 & 2.0198 & \text{-}0.0022 \\ 0.0165 & \text{-}0.0022 & 0.9671 \end{bmatrix}$$

$$\lambda_1 = 3.0020, \lambda_2 = 2.0198\ , \lambda_{13} = 0.9670$$

□

3.4. Summary

In this chapter, we defined the reachability/controllability Gramian for continuous-time LTI system and introduced some of their fundamental properties of them.

Quadratic forms and sum of square polynomials were introduced as well, which led to a new approach for controllability Gramian calculation. In contrast to Lyapunov equations, our method does not imply any restriction on the stability of the system, and this is a noticeable advantage.

Another advantage of our method is that it divides the controllability Gramian matrix into two distinct parts: one parts linked to the system matrix A and the other part is a tensor product of the vectorised input matrix B. Some applications of this method can be demonstrated in perturbation analysis and robust control design. Though the eigenvalues of a matrix are nonlinear functions of its entries it will be attempted to optimize the eigenvalues of the controllability matrix over the elements of the input matrix.

In the last part of this chapter, the controllability Gramian assignment problem was solved by introducing a new approach for selecting the proper input matrix B. This approach can be used for any stable/unstable system with one or more control inputs.

In control design, it is very important to select the proper inputs, to construct an effective control system and to satisfy various constraints and limitations. The controllability assignment is a core problem, which can be considered wherever we have constraints on the input energy, or when the robustness properties of linear system need to be considered.

In the next chapter, the minimum of input energy is introduced as a quadratic form of the inverse of the controllability Gramian. Then we will discuss how the terminal time and the final states may affect the value of the energy. The various energy levels will be depicted in the state space.

Chapter 4: Minimum Input Energy

4.1. Introduction

In this chapter, we develop the fundamentals for selection of input (and by duality output) structure based on the energy type criteria. The fundamental question is the evaluation of the required energy for the transfer of the origin to a given state in the state space within some given final time when the input structure is given. The overall aim is to use these preliminary results to examine in the following parts of the thesis the selection of the input structure under the minimum energy requirement for the worst state. Thus, we introduce first the problem of minimum input energy and then we will discuss the important related factors which can influence the value of this minimum energy. These results lead to the definition of the energy levels in the state space which can be used as an instrument for input structure selection.

The development of the material involves the definition first of the minimal energy and this follows by examining the factors influencing the minimum energy requirements. This involves the influence of the terminal state factor, as well as the role of the required final time that has to be specified for the state transfer. The computations of energy are based on the controllability Gramian and thus we examine separately the case of stable and unstable systems. The essence of our investigation is comparing the minimum of minimum energy E_{min}^{min} transferring the states of a given system from the

origin to a terminal state x_f, such that $\|x_f\| = R$ (= 1) within a fixed terminal time t_f. In fact, we are interested to determine the minimum input energy required to transfer the states of the LTI controllable system from the origin to the states defined by the distance R away from the origin. We compare the maximum of minimum energy E_{min}^{max}, required to transfer the states of a given system from the origin to a terminal state x_f, $\|x_f\| = R$, where R is variable (within a fixed terminal time t_f), that can guaranty to reach a state defined R away from the origin, ignoring the direction of the specific states. This leads to the definition of energy levels in the state space characterizing the maximum of minimum input energy required for transfer to a given distance. This reveals the presence of the energy stratification of the state space.

4.2. Minimum Input Energy

Consider a controllable SISO LTI system, we are interested in transferring the states from the origin to some point on the surface of a sphere with the radius R. The state-space realization of the system is:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases} \qquad A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{l\times n}, D \in \mathbb{R}^{l\times m} \tag{4.1}$$

The energy dissipated by the controller for such transfer is given as the cost function below:

$$E = \|u(t)\|_{[0,t_f]}^2 = \int_0^{t_f} u^T(t)u(t)dt \tag{4.2}$$

We are interested to minimize this cost function, so that the states transfer to the desired terminal, within the defined control time, by expending the least possible input energy.

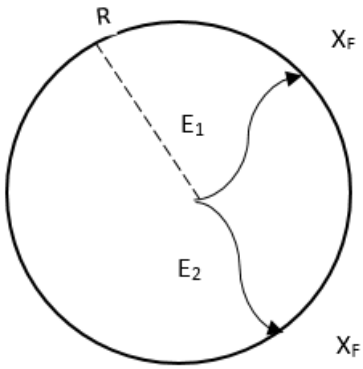


Figure (4.1): Controllability and the energy required to transfer states from the origin to terminal states $x_f, \|x_f\| = R$

As stated in literature, (VanderVelde & Carignan, 1982), (Georges, 1995), (Toan & Georges, 2007) the minimum input energy to transfer the states for the system (4.1) from the origin to the terminal state x_f, such that, $\|x_f\| = R$ (= 1), is given as:

$$E_{\min} = \min \|u(t)\|_{[0,t_f]}^2 = \min \int_0^{t_f} u^T(t)u(t)dt = x_f^T W_c^{-1}(0,t_f)x_f \tag{4.3}$$

Note that the system is assumed controllable and so the indicated inverse of the controllability Gramian is well-defined. The corresponding (unique) optimal control input is:

$$u_{E_{\min}}(t) = u_0(t) = B^T e^{A^T(t_f - t)} W_c^{-1}(0, t_f) x_f \tag{4.4}$$

These results are defined in the Theorem (4.1) below (Fairman, 1998):

Theorem (4.1): In the controllable LTI system (4.1), the minimum energy required to transfer the initial states from the origin to the terminal states located on a sphere around the origin with the radius R within the terminal time t_f is $E_{\min} = x_f^T W_c^{-1}(0, t_f) x_f$ and the required control input for this optimal transfer is:

$$u_{E_{\min}}(t) = u_0(t) = B^T e^{A^T(t_f - t)} W_c^{-1}(0, t_f) x_f .$$

□

The proof can be found in (Fairman, 1998). In different stages of control design, the concept of minimum input energy, may play an important role, e.g. it is important for an airplane or a satellite to complete a defined journey using the minimum fuel, in a smart house minimization of input energy as the cost function usually defines the aim of the design, etc.

As the Figure (4.1) illustrates, to transfer the states of a controllable LTI system from the origin to a terminal point on the surface of a sphere with the radius R, within a terminal time t_f there are many different control inputs in various directions, the Theorem above states that only one of these inputs in a unique direction leads to the minimum energy dissipation. In control design and input selection, it is also important to investigate the factors, which can influence the value of minimum energy corresponding to this optimal direction. This will be discussed in detail in a subsequent section.

4.3. Influence of Different Terms on the Value of Minimum Input Energy

In this part, some significant variables in the value of the minimum input energy are investigated.

Considering the definition of controllability Gramian in (4.5):

$$W_c(0, t_f) = \int_0^{t_f} e^{A\tau} B B^* e^{A^*\tau} d\tau \tag{4.5}$$

As equation (4.3) suggests, in an LTI system, when the matrices A and B are fixed and the system is controllable, two important factors may change the value of minimum input energy:

i. The terminal time: t_f
ii. The terminal state: x_f

It is, of course, clear that if $w_c(0, t_f)$ is nearly singular then a very large energy input is needed to reach certain states aligned with the eigenspace of the controllability Gramian corresponding to its "small" eigenvalues (in this case, the system is hardly controllable.)

4.3.1 Terminal Time t_f

In this part, for the fixed given LTI system (4.1), we consider various transition times to transfer the initial states from the origin to the fixed terminal states x_f such that $\|x_f\| = R$ (= 1), using the minimum energy. In other words, we define the minimum input energy as a function of the variable terminal time.

The results show that the value of the minimum input energy decreases as the terminal time increases, which means that if we give more time to the system it can reach the final states with less input energy requirement.

It is presented in the following Theorem.

Theorem (4.2): In the controllable LTI system (4.1), where the system is fixed, i.e. A and B are invariable matrices, the minimum energy required to transfer the states from the origin to the defined terminal states $x_f, \|x_f\| = R(=1)$, decreases as the terminal time t_f increases, however the rate of this change is not fixed and could be shown as a concave up decreasing graph i.e. the change of the required energy decreases over the terminal time.

Proof:

As stated in (4.3) the minimum input energy for the LTI controllable system (4.1) is defined as follows:

$$E_{\min} = x_f^T W_c^{-1}(0, t_f) x_f$$

By differentiating both sides of the equation above with respect to the terminal time, we get:

$$\frac{\partial E_{\min}}{\partial t_f} = x_f^T \frac{\partial W_c^{-1}(0, t_f)}{\partial t_f} x_f \tag{4.6}$$

Then using identity:

$$\frac{\partial A^{-1}(x)}{\partial x} = -A^{-1} \frac{\partial A(x)}{\partial x} A^{-1} \tag{4.7}$$

Equation (4.6) gives:

$$\begin{aligned} \frac{\partial E_{\min}}{\partial t_f} &= -x_f^T(t) W_c^{-1}(0, t_f) \frac{\partial W_c(0, t_f)}{\partial t_f} W_c^{-1}(0, t_f) x_f(t) \\ &= -x_f^T(t) W_c^{-1}(0, t_f)(BB^T + AW_c(0, t_f) + W_c(0, t_f)A^T) W_c^{-1}(0, t_f) x_f(t) \le 0 \end{aligned} \tag{4.8}$$

Hence the minimum input energy decreases as the final time increases.

□

In other words, the set of states reachable with a cost less than 1 (or any given value) grows as t_f increases.

In this part, we may consider two different cases: In the first case, all eigenvalues of matrix A are located in the open left half plane (OLHP), which means the LTI system is asymptotically stable, while in the second case an unstable system is considered.

We are interested to know how the value of minimum energy is affected by the terminal time t_f, in each of the above cases.

4.3.1.1 The Stable System

For the stable case, we have the following result (Yoshizawa, 2012):

Theorem (4.3): Consider the fixed controllable LTI system (4.1). If the system is asymptotically stable, then we can increase the terminal time to ∞, i.e. $t_f \to \infty$, in this case, the controllability Gramian could be defined as the unique solution of the Lyapunov equation (4.9), and the minimum energy required to move the states from the origin to the defined terminal states $x_f, \|x_f\| = R(=1)$ is also minimized.

$$AW_c(0,\infty) + W_c(0,\infty)A^T = -BB^T \tag{4.9}$$

Proof:

For an asymptotically stable system, whose state matrix A has all its eigenvalues in the OLHP, then, $e^{At} \to 0$ as $t \to \infty$. Thus, there are constants, $M, \sigma > 0$ such that:

$$\left\|e^{At}\right\| = \left\|e^{A^T t}\right\| \le Me^{-\sigma t}, \forall t > 0 \tag{4.10}$$

Then:

$$\begin{aligned} \left\|W_c(0,t_f)\right\| &= \left\|\int_0^{t_f} e^{A\tau} BB^T e^{A^T\tau} d\tau\right\| \le \int_0^{t_f} \left\|e^{A\tau}\right\| \left\|BB^T\right\| \left\|e^{A^T\tau}\right\| d\tau \\ &\le M^2 \left\|BB^T\right\| \int_0^{t_f} e^{-2\sigma\tau} d\tau = M^2 \left\|BB^T\right\| (1-e^{-2\sigma t_f})/2\sigma \\ &\le \frac{1}{2\sigma} M^2 \left\|BB^T\right\| \end{aligned} \tag{4.11}$$

This implies:

$$\lim_{t_f \to \infty} W_c(0,t_f) = W_c(0,\infty) \tag{4.12}$$

and the controllability Gramian matrix $W_c(0,\infty)$ can be obtained as the steady state solution of the equation below:

$$\begin{cases} \dot{W}_c(0,t_f) = BB^T + AW_c(0,t_f) + W_c(0,t_f)A^T \\ W_c(0,0) = 0 \end{cases} \tag{4.13}$$

By setting $\dot{W}_c(0,t_f) = 0$, we then have that the controllability Gramian matrix $W_c(0,\infty)$ satisfies the matrix Lyapunov equation:

$$AW_c(0,\infty) + W_c(0,\infty)A^T = -BB^T \tag{4.14}$$

□

which means that for a stable system the minimum input energy can be calculated in the case $t_f \to \infty$ via the controllability Gramian which in this case is the unique symmetric positive-definite solution of the algebraic Lyapunov equation (4.14). However, in a stable system we have $W_c^{-1}(0,t_f) > 0$, which implies we cannot get anywhere for free.

Example (4.1): An example of a classical linear system is the mass-spring-damper system. Newton's laws of motion for an object moving horizontally on a plane and attached to a wall via a spring and damper gives:

$$m\ddot{y}(t) = u(t) - b\dot{y}(t) - ky(t) \tag{4.15}$$

where $y(t)$ is the position of the center of mass of the object, $\dot{y}(t)$ its velocity, $\ddot{y}(t)$ its acceleration, $u(t)$ is the applied force. Further b is the viscous friction coefficient, k is the spring constant and m is the mass of the moving object. Assume the origin as the initial condition, i.e. $y(t) = \dot{y}(t) = 0$, if we define the target states as $y(t_f) = 1, \dot{y}(t_f) = 0$.We are interested in the minimum energy dissipated by the system considering various terminal times.

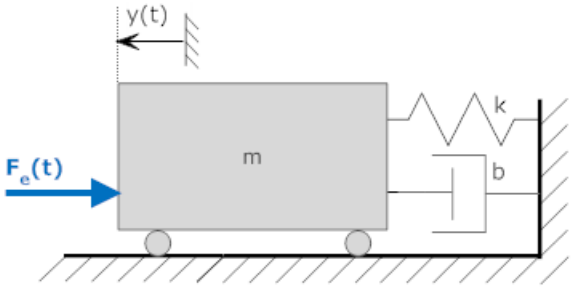


Figure (4.2): Mass-spring-damper system

The state equation is:

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \\ y = [\ 1 \quad 0\ ]x \end{cases} \tag{4.16}$$

where $x_1(t)$ represents the position of the object, and $x_2(t)$ its velocity.

For this example, the controllability test shows that such LTI system is controllable for all b and m:

$$\begin{bmatrix} B & AB \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} & \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{m} \\ \frac{1}{m} & -\frac{b}{m^2} \end{bmatrix} \tag{4.17}$$

Clearly, the system is stable for all values of b, k and m, since all eigenvalues of A are located in OLHP:

$$\begin{aligned} &\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 \\ \frac{k}{m} & \lambda + \frac{b}{m} \end{bmatrix} = \lambda^2 + \lambda \frac{b}{m} - \frac{k}{m} = 0 \\ &\Rightarrow \lambda = real(-\frac{b}{2m} \pm \frac{\sqrt{(\frac{b}{2m})^2 + \frac{k}{m}}}{-1}) < 0 \end{aligned} \tag{4.18}$$

Suppose that $m = 1kg, k = 1N/m$ and $b = 0.2Ns/m$. Considering a range of terminal times $0.5 \le t_f \le 10$, the corresponding value of the minimum input energy was calculated and is illustrated in Figure (4.3). The Figure demonstrates that the minimum energy required to move the mass from the initial condition to the terminal state $y(t_f) = 1, \dot{y}(t_f) = 0$ decreases from 91.6790 to 0.4694, as the terminal time t_f increases from 0.5 to 10.

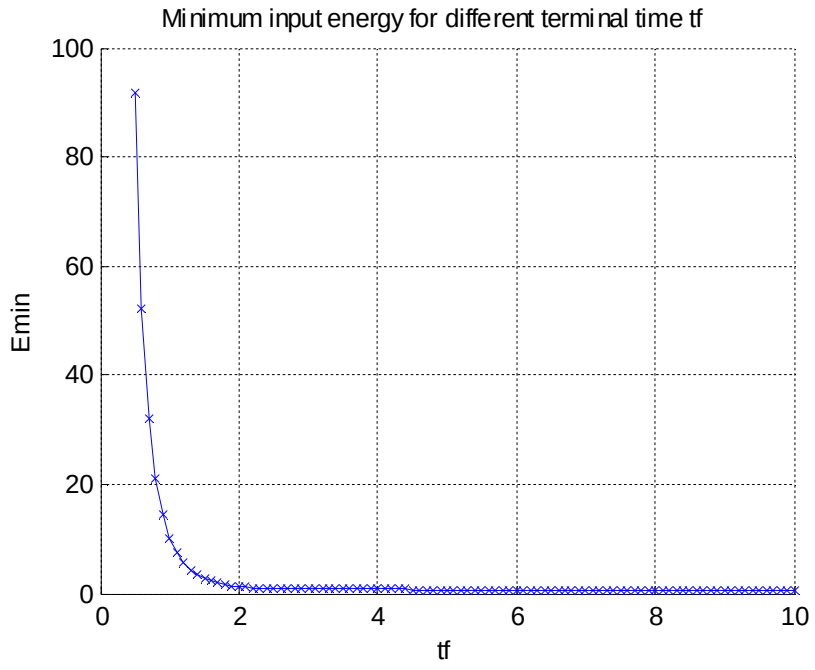


Figure (4.3): $E_{\min}$ as a function of variable terminal time where $t_f \in [0.5, 10]$

Figure (4.4) shows that the minimum energy converges as $t_f \to \infty$. The limiting energy can be calculated via the steady-state controllability Gramian which is the solution of the algebraic Lyapunov equation (4.14).

$$AW_c(0,\infty)+W_c(0,\infty)A^T=-bb^T$$
$$W_c(0,\infty)=\int_0^{\infty} e^{A\tau}bb^T e^{A^T\tau}d\tau=\begin{bmatrix}2.5 & 0\\ 0 & 2.5\end{bmatrix}$$
$$\Rightarrow\begin{bmatrix}0 & 1\\ -1 & -0.2\end{bmatrix}\begin{bmatrix}2.5 & 0\\ 0 & 2.5\end{bmatrix}+\begin{bmatrix}2.5 & 0\\ 0 & 2.5\end{bmatrix}\begin{bmatrix}0 & -1\\ 1 & -0.2\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & -1\end{bmatrix} \qquad (4.19)$$

$$E_{\min}=x_f^T W_c^{-1}(0,\infty)x_f=\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0.4 & 0\\ 0 & 0.4\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=0.4$$

In fact, 0.4 is the minimum possible energy, by which this system can move from the origin to $\begin{bmatrix}1\\ 0\end{bmatrix}$ and is achieved when $t_f \to \infty$.

All the results of this example, supports Theorem (4.3).

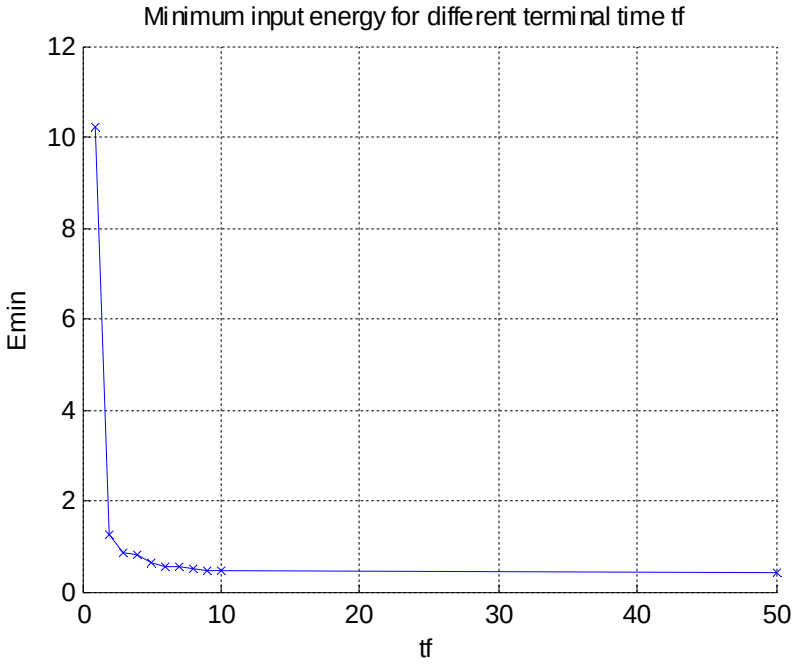


Figure (4.4): $E_{\min}$ as a function of terminal time variable, $0 < t_f \leq \infty$

□

4.3.1.2 Unstable System

If the system is unstable $e^{At}B$ becomes unbounded as $t = t_f \to \infty$. Thus, the controllability Gramian $W_c(0,\infty)$ also becomes unbounded which means that for some x_f, $\lim_{t_f \to \infty} E_{min} \to 0$. This means that some directions do not require any cost of control. In other words, for an unstable system, inverse of controllability Gramian $W_c^{-1}(0,t_f)$ can have a nonzero null space $W_c^{-1}(0,t_f)x_f = 0$ for some $x_f \neq 0$, which means that the system can get to x_f using arbitrarily small input energy. The input just gives a little kick to the initial state and then giving the system enough time, instability carries it out to the desired final state x_f.

Example (4.2): As an example, assume the linear model of a Furuta pendulum (Guzman, et al., 2016), described below:

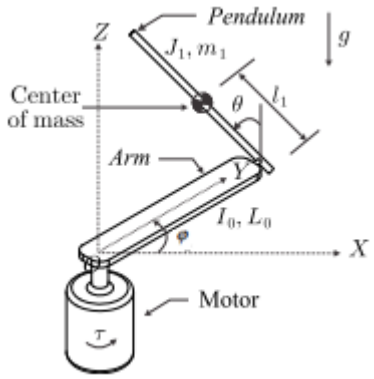


Figure (4.5): The schematic picture of Furuta pendulum

The model equations are:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-gm_1^2 l_1^2 L_0}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{(I_0 + m_1 L_0^2) m_1 l_1 g}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2} & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{(J_1 + m_1 l_1^2)}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2} \\ 0 \\ \frac{-m_1 l_1 L_0}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2} \end{bmatrix} u \qquad (7.20)$$

where, $x_1 = \theta$ is the angle of the pendulum, $x_2 = \dot{\theta}$ is the angular velocity of the pendulum, $x_3 = \phi$ is the angle of the arm carrying the pendulum, and $x_4 = \dot{\phi}$ is the angular velocity of the carrying arm. Now, by defining:

$$\alpha = \frac{-gm_1^2 l_1^2 L_0}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2}, \gamma = \frac{(I_0 + m_1 L_0^2) m_1 l_1 g}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2}$$
$$\beta = \frac{(J_1 + m_1 l_1^2)}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2}, \varepsilon = \frac{-m_1 l_1 L_0}{I_0(J_1 + m_1 l_1^2) + J_1 m_1 L_0^2} \quad (4.21)$$

The linear model of the system reduces to:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \gamma & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ \beta \\ 0 \\ \varepsilon \end{bmatrix} u \quad (4.22)$$

Assume the system is moving from the initial states $\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T$ to $x_f = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T$.

Figure (4.6) illustrates the changes of the minimum input energy for various terminal times $0 < t_f < \infty$ by substituting the following values:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 31.32 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -0.5584 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ -71.23 \\ 0 \\ 191.2 \end{bmatrix} u \quad (4.23)$$

The system above is clearly an unstable, controllable system.

As in can be seen in Figure (4.6), the minimum energy is decreasing as the terminal time increases however because of instability, the value of the minimum energy is changing really close to zero, even if the terminal time is small. It supports the results discussed in this section as for unstable systems.

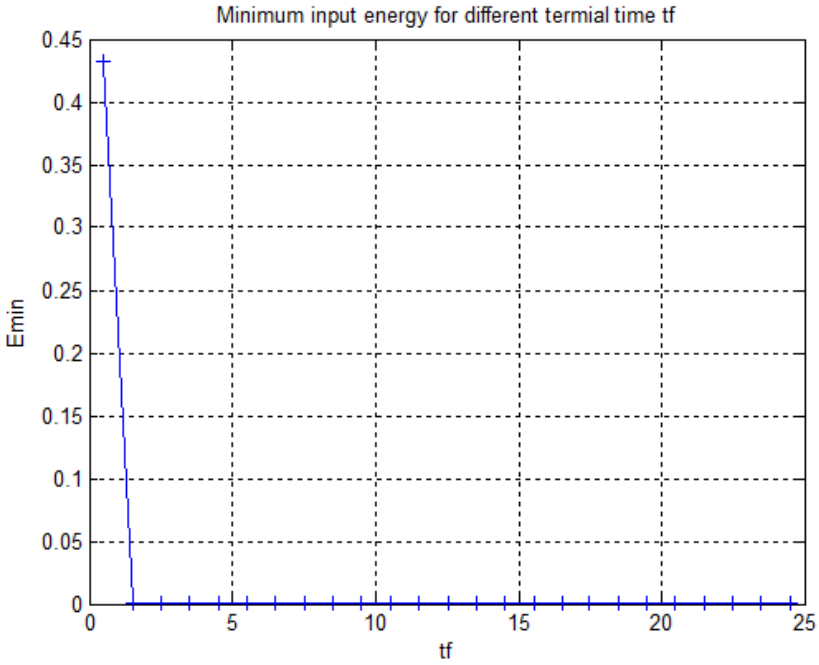


Figure (4.6): $E_{\min}$ as a function of final time variable, $0 < t_f < \infty$

□

4.3.2 Terminal State x_f

In this part, we assume a fixed given LTI system described in (4.1) and we are interested in transferring the initial states from the origin to the variable terminal states x_f such that ($\|x_f\| = R\ (=1)$), within a fixed terminal time t_f=T.

Here the terminal state x_f is assumed a variable point located on the surface of a sphere centerd at the origin with radius R, and we are interested to find the direction for which the corresponding minimum input energy is minimized or maximized.

The results show that the minimum value of the minimum input energy happens in the direction of an eigenvector corresponding to the maximum eiegenvalue of the controllability Gramian, and the maximum of minimum input energy is in the direction of an eigenvector coresponding to the minimum eigenvalue of controllability Gramian. First consider the definitions below:

Definition (4.1): If x is a nonzero vector in $\mathbb{R}^n$ and A is a $n \times n$ dimensional matrix, the Rayleigh quotient of x with respect to A is defined as (Simovici, 2012):

$$RQ_A(x) = \frac{x^T A x}{x^T x} \tag{4.24}$$

□

Definition (4.2): A linear subspace $V \subseteq \mathbb{R}^n$ is called an invariant subspace of A if it satisfies $Ax \in V$ for $\forall x \in V$ (Simovici, 2012).

□

We can now prove the following Lemma (Prasolov, 1994):

Lemma (4.1): If A is a real symmetric matrix and V is an invariant subspace of A, then there is some $x \in V$ such that $RQ_A(x) = \inf\{RQ_A(y), y \in V\}$. Any $x \in V$ that minimizes $RQ_A(x)$ is an eigenvector of A, and the value $RQ_A(x)$ is the corresponding eigenvalue.

□

Proof:

If *x* is a vector and r is a nonzero scalar, then $RQ_A(x) = RQ_A(rx)$, hence every value attained in V by the function $RQ_A(x)$ is attained on the unit sphere $S(V) = \{x \in V, \|x\| = 1\}$.

The function $RQ_A(x)$ is a continuous function on $S(V)$, and $S(V)$ is compact (closed and bounded) so by Weistrass Theorem (Gould, 1957) $RQ_A(x)$ attains its minimum value at some unit vector $x \in S(V)$. Using this quotient rule we can compute the gradient as:

$$\nabla RQ_A(x) = \frac{2Ax - 2(x^T Ax)x}{(x^T x)^2} \tag{4.25}$$

At the vector $x \in S(V)$ where $RQ_A(x)$ attains its minimum value in V, this gradient vector must be orthogonal to V, otherwise, the value of $RQ_A(x)$ would decrease as we move away from x in the direction of any $y \in$ V that has negative dot product with $\nabla RQ_A(x)$.

On the other hand, our assumption that V is an invariant subspace of A implies that the right side of (4.25) belongs to V. The only way that $\nabla RQ_A(x)$ could be orthogonal to V while also belongs to V is if the zero vector, hence:

$$\nabla RQ_A(x) = 0 = 2Ax - 2(x^T Ax)x \Rightarrow Ax = (x^T Ax)x \Rightarrow x^T Ax = \lambda \tag{4.26}$$

so the Lemma is proved.

□

Now considering the Lemma above, equation (4.3) along with the fact that the inverse of controllability Gramian $W_c^{-1}(0, t_f)$ is a symmetric matrix, by defining:

$$S(V) = \{x \in V, \|x\| = 1\} \tag{4.27}$$

it can be concluded that

$$RQ_A(x_f) = x_f^T W_c^{-1}(0, t_f) x_f \tag{4.28}$$

is minimized (maximized) in the direction corresponding to the smallest (largest) eigenvalue of inevrse of controllability Gramian $W_c^{-1}(0,t_f)$.

The results stated above are summarized as shown below:

Remark (4.1): The following relations hold true:

$$\min_{x_f^T x_f=1}\{\min\|u(t)\|^2_{[0,t_f]}\}=\min_{\|x_f\|=1}\{x_f^T W_c^{-1}(0,t_f)x_f\}=\lambda_{\min}(W_c^{-1}(0,t_f))=\frac{1}{\lambda_{\max}\{W_c(0,t_f)\}}$$

$$\Rightarrow E_{\min}^{\min}=\underset{\|x_f\|=1}{\lambda_{\min}}(W_c^{-1}(0,t_f))=\frac{1}{\lambda_{\max}\{W_c(0,t_f)\}} \quad (4.29)$$

and

$$\max_{x_f^T x_f=1}\{\min\|u(t)\|^2_{[0,t_f]}\}=\max_{\|x_f\|=1}\{x_f^T W_c^{-1}(0,t_f)x_f\}=\lambda_{\max}(W_c^{-1}(0,t_f))=\frac{1}{\lambda_{\min}\{W_c(0,t_f)\}}$$

$$\Rightarrow E_{\min}^{\max}=\underset{\|x_f\|=1}{\lambda_{\max}}(W_c^{-1}(0,t_f))=\frac{1}{\lambda_{\min}\{W_c(0,t_f)\}} \quad (4.30)$$

□

Now, based on the results above, in the next section two sub-problems could be defined that is:

i. Energy Levels in the best case, considering "E_{min}^{min}"
ii. Energy Levels in the worst case, considering "E_{min}^{max}"

4.3.2.1 Energy Levels in the Best Case

Here, we are comparing the minimum of minimum energy E_{min}^{min} transferring the states of a given system from the origin to the terminal states x_f (within a fixed terminal time t_f) such that $\|x_f\| = R$ and R is variable.

In other words, in this part we are interested to know how much could be the minimum input energy which transfers the states of the LTI controllable system from the origine to the states defined R away from the origin, according to the Theorem below, this transfer should be done in the direction of the eigenvector corresposding to the minimum eigenvalue of $W_c^{-1}(0,t_f)$ (i.e. the maximum eigenvalue of $W_c(0,t_f)$).

Theorem (4.4): (Prasolov, 1994) A quadratic form can always be reduced to the form:

$$q(x_1,\dots,x_n) = \lambda_1 {x_1}^2 + \cdots + \lambda_n {x_n}^2 \quad (4.31)$$

□

According to the Theorem above (so called Lagrange Theorem), we can define the minimum input energy (4.3) as:

$$E_{\min} = x_f^T W_c^{-1}(0,t_f)x_f = \sum_{i=1}^{n} \lambda_i (W_c^{-1}(0,t_f))x_i^2 \tag{4.32}$$

That is:

$$E_{\min} = x_f^T W_c^{-1}(0,t_f)x_f = \sum_{i=1}^{n} \frac{1}{\lambda_i (W_c(0,t_f))} x_i^2 = \frac{1}{\lambda_1 (W_c(0,t_f))} x_1^2 + \cdots + \frac{1}{\lambda_n (W_c(0,t_f))} x_n^2 \tag{4.33}$$

Then it can be easily inferred from (4.33) that for a given fixed system we have:

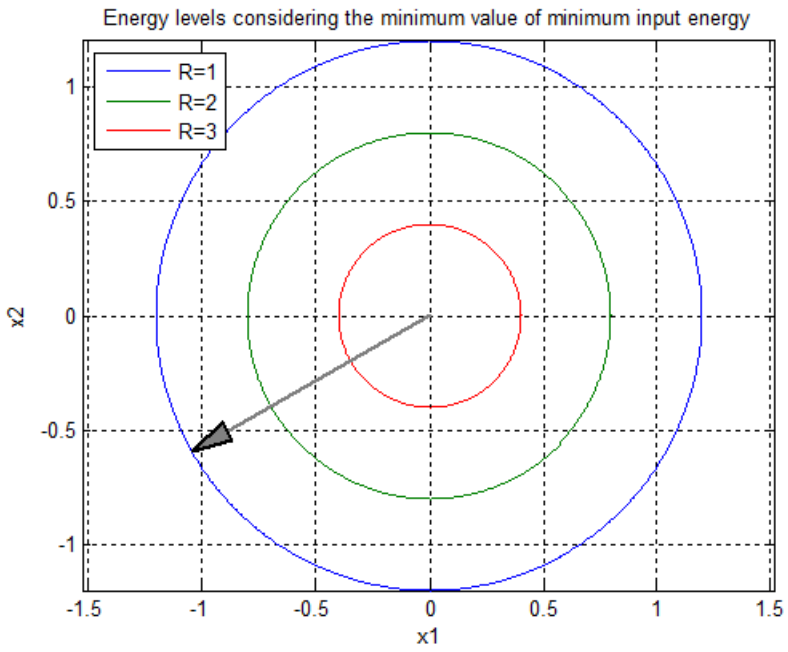
Proposition (4.1): The optimal minimum input energy, which describes the minmimum energy level for $\|x_f\| = R$ is defined along the eigenvector corresponding to the maximum eigenvalue of controllability Gramian $W_c(0,t_f)$,$V_{max} = V\left(\lambda_{max}\left(W_c(0,t_f)\right)\right)$ and its value is:

$$E_{\min}^{\min} = \min_{\|x_f\|=R} \{x_f^T W_c^{-1}(0,t_f)x_f\} = \frac{1}{\lambda_{\max}(W_c(0,t_f))} R^2 \tag{4.34}$$

□

Figure (4.7) shows the levels of minimum input energy for the LTI moving object described in example (4.1), for different values of R , where $\|x_f\|^2 = R$.

In each case the corresponding direction, in which the minimum of minimum energy happens is ploted as well, which is the same as the direction of the eigenvector corresponding to the maximum eigenvalue of the controllability Gramian $W_c(0,t_f)$.



Figure(4.7): Energy levels considering $E_{\min}^{\min}$

In Figure (4.8), the levels for the minium of the minimum input energy is shown for furuta pendulum 3-dimensional system defined in example (4.2).

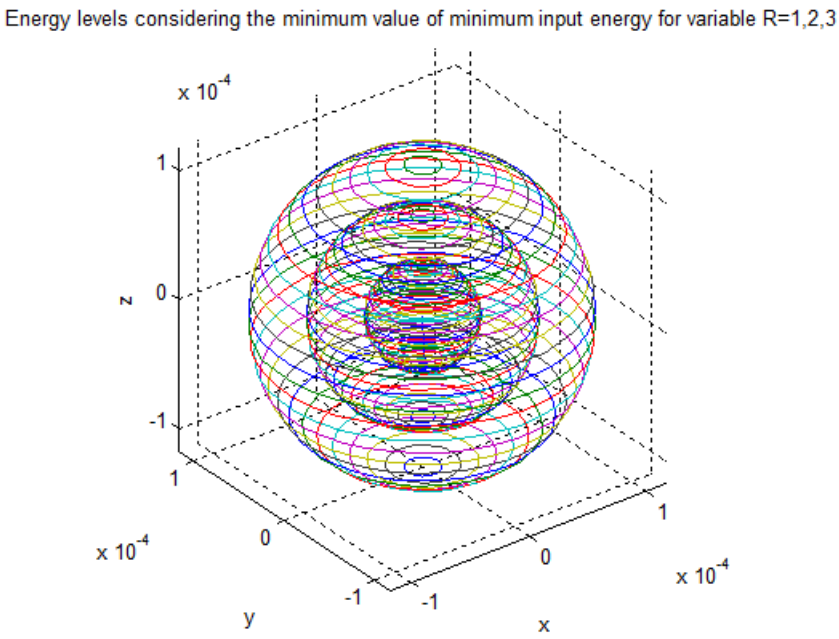


Figure (4.8): Energy levels for furuta pendulum considering $E_{\min}^{\min}$

4.3.2.2 Energy Levels in the Worst Case

In this part, we compare the maximum of minimum energy E_{min}^{max},to transfer the states of a given controllable LTI system from the origin to the terminal states x_f, $\|x_f\| = R$, where R is variable (within a fixed terminal time t_f), it can guaranty if you want to reach a state defined R away from the origin, ignoring the direction of transferring the states, the energy , which you need would be equal or less than a specific value, which is corresponding to the energy level defined for the maxmum of minimum input energy.

This would be exactly the same as the solution proposed in previous part.

Remark (4.2): The maximum of minimum energy for transferring the initial states of controlable LTI system (4.1) from the origin to x_f, $\|x_f\| = R$ could be defined in the direction of eigenvector corresponding to the smallest eigenvalue of controllability Gramian $W_c(0,t_f)$,$V_{min} = V\left(\lambda_{min}\left(W_c(0,t_f)\right)\right)$:

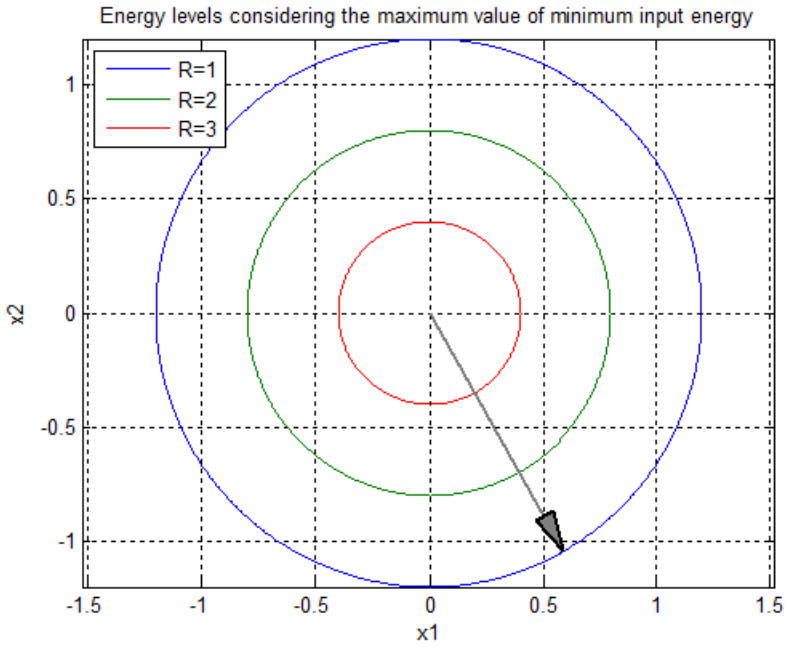
$$E_{\min}^{\max} = \max_{\|x_f\|=R}\{x_f^T W_c^{-1}(0,t_f)x_f\} = \frac{1}{\lambda_{\min}(W_c(0,t_f))}R^2 \tag{4.35}$$

□

The above represents the energy stratification of the state space, which is important in input selection and optimal actuator placement. By optimizing the energy levels of controllable LTI system (4.1), the optimum positions of actuators could be defined.

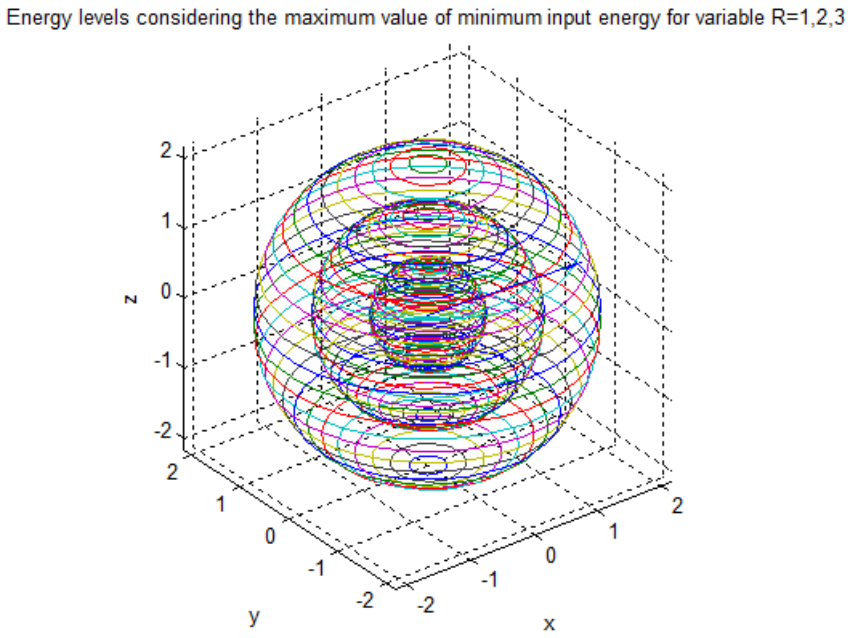
Figure (4.9) illustrates the energy levels for the maximum of minimum input energy for 2-dimensional mass-spring system defined in example (4.1)

The direction, in which these energy levels are calculated, is shown as well, that is the same as the direction of the eigenvector corresponding to the minimum eigenvalue of the controlability Gramian $W_c(0,t_f)$.



Figure(4.9): Energy levels considering $E_{\min}^{\max}$

Figure (4.10) demonstartes the eneregy levels for 3 dimensional pendulum system (see Example (4.2)) consedering the maximum value of minimum energy input for the varius values of $R = \|x_f\|^2 = 1,2,3$



Figure(4.10): Energy levels for furuta pendulum considering $E_{\min}^{\max}$

Now, regarding the resuls achived for the variable trminal times and the variable terminal states in the sequel, we define the minimum energy E_{min} as a function of two variables t_f , x_f , to investigate how the minimum input energy E_{min} varies, if we transfer the states of the LTI controlable system (4.1) , from the origin to the arbitrary terminal states x_f , in varius directions, satisfying the condition $\|x_f\| = R$ wihin different terminal times t_f .

Figure (4.11) shows the changes of minimum input energy in the LTI mass-spring system (example (4.1)), assuming $R = 1$.

The Figure shows that as the terminal time increases the effect of the direction decreeases, since the cost function looks more like a circle, it means that by gowing the terminal time t_f the difference between the minimum and maximum eigenvalues of the ccontrolability Gramian $W_c(0, t_f)$ decreases, i.e. the reachability sets looks more like a circle. This can be seen in Figure (4.12).

In other words, for well-conditioned controllability Gramian matrix, the maximum and minimum energy levels are almost same, which indicates regardless of the trajectory and the direction of movement, there is a minimum value for the energy dissipated to move the states from the origin to the target point x_f, $\|x_f\| = R$ (within a terminal time t_f).

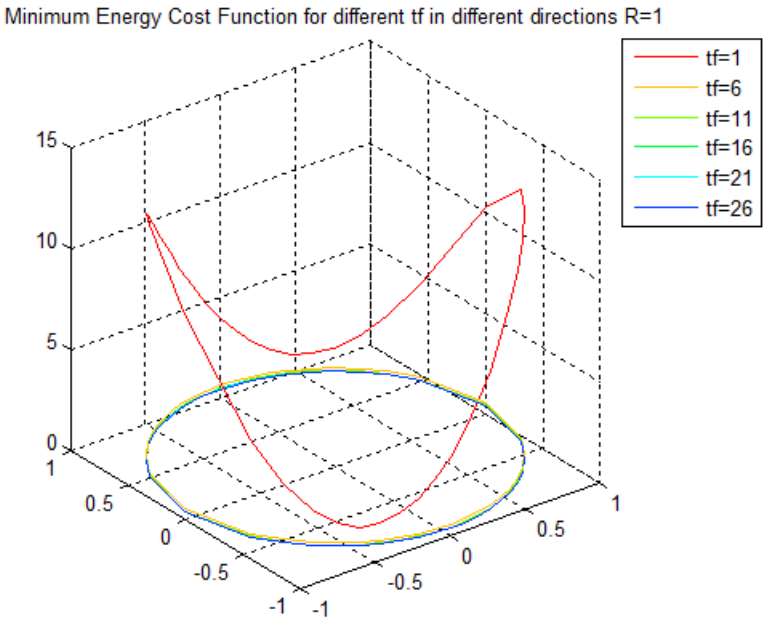


Figure (4.11): Minimum input energy as a function of the variables t_f , x_f

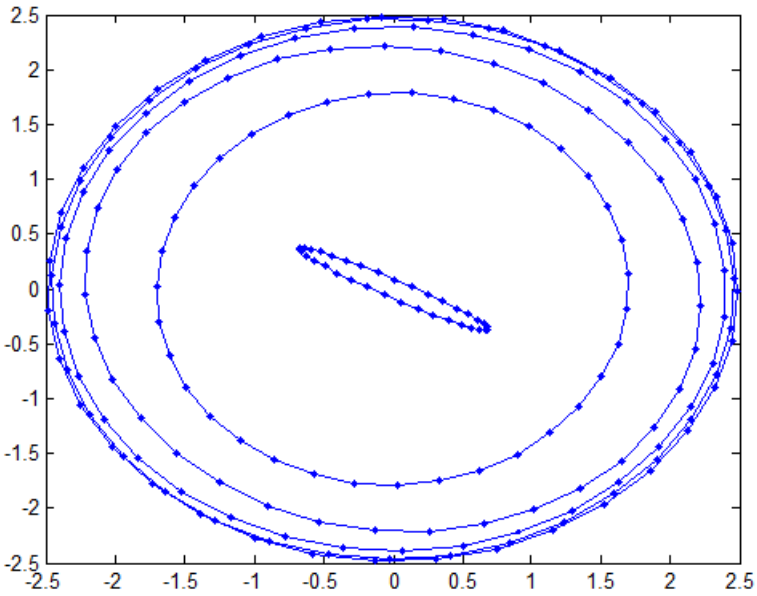


Figure (4.12): Reachability area changes based on different terminal times for 2 dimensional example model in (4.16)

4.4 Summary

In this chapter, we discussed the fundamentals of minimum energy required for transferring the states of a controllable LTI system form the origin to some target states $x_f, \|x_f\| = R$ in the state space within the terminal time t_f . Terminal time and terminal states as two important elements which may affect the value of this minimal energy have also been discussed.

In particular, if controllability Gramian $W_c(0, t_f)$ is nearly singular, then large energy inputs are required to reach the target states belonging to the eigenspace of its least eigenvalue $\lambda_{\min}(W_c(0, t_f))$. This motivated the use of the minimum/maximum eigenvalue of the controllability Gramian in the energy stratification of the state space, in this chapter we adopted the measure of the worst-case control effort when investigating maximum of optimum energy levels.

In the next chapter, we consider the energy calculation for unstable systems and the impact of actuator selection on various issues such as the degree of controllability and disturbance rejection.

Chapter 5: Controllability, Energy for Unstable Systems and Measures of Degree of Controllability

5.1. Introduction

In this chapter, our focus is on the energy calculation for unstable systems and the impact of actuator selection on various issues such as the degree of controllability, and disturbance rejection.

In the previous chapter, we have defined the minimum input energy and introduced the energy levels for stable systems, in this chapter we examine the case where the system is unstable, and we present the generalized energy stratification of the state space for the possibly unstable systems.

Section 5.2 and 5.3 cover the above issues, then in section 5.4 we will investigate the problem of the actuator selection with regard to minimum energy.

One of the most reliable methods for determining actuator locations is the improvement of state controllability of the process (Wal & Jager, 2001), then the problem for the actuator placement could be viewed as the problem of minimizing the input energy required to reach a given state. A deficiency

of this method appears where sufficient information for the initial and final states of the system is not available.

This problem is discussed in section 5.4 in details then optimization of the worst-case method is suggested as a solution.

In section 5.5 we presented a method of actuator selection considering the degree of controllability of linear time-invariant systems. Starting from a set of actuator locations which produce an uncontrollable system, but for which the number of actuators is sufficient to produce controllability, it will usually be the case that moving one of the actuators a distance $\varepsilon > 0$ can produce a controllable system, no matter how small the value ε (and vice versa). It is then important to select the actuators which maximize the degree of controllability.

The similar problem will be considered in section 5.6, where the system is affected by an external disturbance as well. In this case, we define the degree of disturbance rejection by considering both, the controllability Gramian of the system and the disturbance sensitivity Gramian, and we place the actuators such that the presented measure is maximized.

The methods proposed in this chapter are general methods in the sense that they support both stable and unstable systems. The techniques described are successfully verified by various examples.

5.2. Minimum Energy for Unstable Systems

In chapter 4 we have discussed in detail, the minimum energy dissipation needed to transfer the states of an LTI system from the origin to an arbitrary destination x_f within a finite time t_f. This can be calculated as $x_f^T W_c^{-1}(0,t_f)x_f$ where $W_c(0,t_f)$ is the controllability Gramian. According to (Hać & Liu, 1993) in more general case where we are steering the states from any initial states x_0 to the final states x_f within the terminal time t_f ,the minimum energy required for this transferring could be represented as:

$$E_{\min} = (e^{At_f}x_0 - x_f)^T W_c^{-1}(0,t_f)(e^{At_f}x_0 - x_f) \tag{5.1}$$

If the system is asymptotically stable, $e^{At_f} \to 0$ as $t_f \to \infty$, then we can attain the minimum energy as a function of final states only regardless of their initial location. In this case if we choose the origin as the final state $x_f = 0$ then $E_{\min} = 0$, which means bringing the system to the state of equilibrium can be achieved by providing zero energy, as the asymptotically stable system always converges to the origin by itself, if the final time is sufficiently large. This implies that, if the system is asymptotically stable escaping from origin (i.e., the reachability problem) requires more energy than transferring to origin, (i.e., the controllability to 0 problem) because the stable system naturally tends to converge to zero.

Furthermore equation (5.1) states that in undamped systems (i.e. eigenvalues are located on the imaginary axes) where e^{At_f} is an orthonormal matrix, i.e. $e^{At_f}e^{A^T t_f} = I_n$, the minimum input energy required to bringing the system to the origin is $E_{\min} = x_0{}^T W_c^{-1}(0,t_f)x_0$ (Arbel, 1981).

In this section, we use the statement (5.1) to determine the minimum energy for an unstable system, where the system is decomposable to a stable and anti-stable subsystem, i.e. when system matrix A has no eigenvalues on the imaginary axis.

5.2.1. Decomposition of Controllability Gramian in LTI Systems with Stable and Anti-Stable Modes

Consider the unstable LTI system below:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{l\times n}, D \in \mathbb{R}^{l\times m}, x \in \mathbb{R}^{n} \tag{5.2}$$

where A has one or more eigenvalues in the right half plane (RHP) of the complex plane and it has no eigenvalues on the imaginary axis (as for the minimum input energy computation, no energy is required to change the states of undamped modes in sufficiently large time period $[0, t_f]$, thus this assumption imposes no extra restriction).

In general, there are two well-known transformation (Sivasundaram, 2004), (Demenkov & Goman, 2009), (Chen, 1995) suitable for the stable/unstable decomposition of a linear continuous-time system. The one is based on Schur transformation (Castelan, et al., 1996), (Golub & Loan, 1986) and the second one is based on the block Jordan transformation (Chen, et al., 2004). In this thesis, we will use the latter.

Remark (5.1): There exists a similarity transformation $A \to V^{-1}AV = \tilde{A}, V \in \mathbb{R}^{n\times n}$, for which $\tilde{A}$ is block diagonal, i.e. $V^{-1}AV = \Lambda = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix}$, where A_s, A_u are stable and anti-stable system matrices, respectively.

Then in the new coordinate frame one can rewrite the LTI system (5.2) as:

$$\begin{cases} \begin{bmatrix} \dot{x}_s \\ \dot{x}_u \end{bmatrix} = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix} \begin{bmatrix} x_s \\ x_u \end{bmatrix} + \begin{bmatrix} B_s \\ B_u \end{bmatrix} u \\ y = [C_s \quad C_u] \begin{bmatrix} x_s \\ x_u \end{bmatrix} + Du \end{cases} \tag{5.3}$$

with the appropriate dimensions for all matrices.

□

Remark (5.2): For the finite-horizon case $W_c(0, t_f)$ is always well defined as a symmetric positive semidefinite matrix (which is not the case for the infinite-horizon problem), then the controllability Gramian for the finite terminal time $W_c(0, t_f)$ could be derived from the state space realization above as:

$$W_c(0,t_f)=\begin{bmatrix} W_s(0,t_f) & W_{su}(0,t_f) \\ W_{us}(0,t_f) & W_u(0,t_f) \end{bmatrix} \tag{5.4}$$

□

and by the general formula for the inverse of the partitioned matrix we obtain:

Lemma (5.1): Consider the block partitioned controllability Gramian matrix in (5.4) is positive definite, then the inverse of controllability Gramian could be achieved as:

$$W_c^{-1}(0,t_f)=\begin{bmatrix} (W_s-W_{su}W_u^{-1}W_{us})^{-1} & -W_s^{-1}W_{su}(W_u-W_{us}W_s^{-1}W_{su})^{-1} \\ -W_u^{-1}W_{us}(W_s-W_{su}W_u^{-1}W_{us})^{-1} & (W_u-W_{us}W_s^{-1}W_{su})^{-1} \end{bmatrix} > 0 \tag{5.5}$$

$$-W_s^{-1}W_{su}(W_u-W_{us}W_s^{-1}W_{su})^{-1}=\left(-W_u^{-1}W_{us}(W_s-W_{su}W_u^{-1}W_{us})^{-1}\right)^T$$

□

We refer the interested readers to (Horn & Johnson, 2012) for the proof. Next, we define Schur complement Theorem for a positive definite symmetric matrix.

Theorem (5.1): (Fuzhen, 2005) Given any symmetric matrix $M_{n\times n}=\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ the following conditions are equivalent:

i. $M > 0$ (M is positive definite)
ii. $A>0, C-B^TA^{-1}B>0$
iii. $C>0, A-BC^{-1}B^T>0$

□

The complete proof of this Theorem is given in (Fuzhen, 2005). Regarding the time dependency in (5.1) we can use the continuous-time differential Lyapunov equations (CDLE) (Davis, et al., 2010) , then we get the Corollary below:

Corollary (5.1): (Middleton & Goodwin, 1990), (Amato, et al., 2015), (Singh, 2009) Consider the linear time-invariant continuous system (5.2) with no eigenvalues on the imaginary axis. Then the controllability Gramian satisfies the continuous-time differential Lyapunov equation

$$\dot{W}_c(0,t_f)=AW_c(0,t_f)+W_c(0,t_f)A^T+BB^T \tag{5.6}$$

with zero initial condition or equivalently by partitioning $W_c(0,t_f)$:

$$W_c(0,t_f)=\begin{bmatrix} W_s(0,t_f) & W_{su}(0,t_f) \\ W_{us}(0,t_f) & W_u(0,t_f) \end{bmatrix}$$

$$\begin{aligned}
\dot{W}_s &= A_s W_s + W_s A_s^T + B_s B_s^T \\
\dot{W}_{su} &= A_s W_{su} + W_{su} A_u^T + B_s B_u^T \\
\dot{W}_{us} &= A_u W_{us} + W_{us} A_s^T + B_u B_s^T \\
\dot{W}_u &= A_u W_u + W_u A_u^T + B_u B_u^T
\end{aligned} \tag{5.7}$$

□

For the proof see (Middleton & Goodwin, 1990), (Amato, et al., 2015), (Singh, 2009).

Theorem (5.2): Let (5.2) be a controllable LTI system with both stable and anti-stable parts. Then the inverse of controllability Gramian i.e. $W_c^{-1}(0,t_f) > 0$, as is stated in Lemma (5.1), is well-defined in the limit $t_f \to \infty$ and:

$$W_c^{-1}(0,\infty) = \begin{bmatrix} W_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} \tag{5.8}$$

□

Proof:

If the system is controllable, the controllability Gramian (5.4) must be positive definite $W_c(0,t_f) > 0$ (E.Hendricks, et al., 2008) and according to Schur complement Theorem (5.1) we have:

$$(W_u - W_{us}W_s^{-1}W_{su}) > 0 \text{ and } (W_u - W_{us}W_s^{-1}W_{su})^{-1} > 0 .$$

Now set $Q = (W_u - W_{us}W_s^{-1}W_{su})$. Using the chain rule of differentiation, we get:

$$\dot{Q} = (\dot{W}_u - \dot{W}_{us}W_s^{-1}W_{su} - W_{us}(W_s^{-1})\dot{W}_{su} - W_{us}W_s^{-1}\dot{W}_{su}) \tag{5.9}$$

Then considering (CDLE) equations (5.7) we obtain:

$$\begin{aligned}
&\dot{Q} = (\dot{W}_u - \dot{W}_{us}W_s^{-1}W_{su} - W_{us}(W_s^{-1})\dot{W}_{su} - W_{us}W_s^{-1}\dot{W}_{su}) \\
&= A_u W_u + W_u A_u^T + B_u B_u^T - (A_u W_{us} + W_{us} A_s^T + B_u B_s^T) W_s^{-1} W_{su} + W_{us} W_s^{-1} \dot{W}_s W_s^{-1} W_{su} - W_{us} W_s^{-1} (A_s W_{su} + W_{su} A_u^T + B_s B_u^T) \\
&\Downarrow \\
&\dot{Q} = A_u (W_u - W_{us}W_s^{-1}W_{su}) + (W_u - W_{us}W_s^{-1}W_{su}) A_u^T + B_u B_u^T - B_u B_s^T W_s^{-1} W_{su} - W_{us} W_s^{-1} B_s B_u^T + W_{us} W_s^{-1} B_s B_s^T W_s^{-1} W_{su} \\
&= A_u (W_u - W_{us}W_s^{-1}W_{su}) + (W_u - W_{us}W_s^{-1}W_{su}) A_u^T + (W_{us}W_s^{-1}B_s - B_u)(W_{us}W_s^{-1}B_s - B_u)^T \\
&\Downarrow \\
&\dot{Q} = A_u Q + Q A_u^T + (W_{us}W_s^{-1}B_s - B_u)(W_{us}W_s^{-1}B_s - B_u)^T
\end{aligned} \tag{5.10}$$

Therefore, according to the differential Lyapunov equations (5.7) and with regard to the results of chapter 4, we obtain $\dot{Q} \geq 0$.

Furthermore, using the differentiation property of non-singular matrices we have:

$$(Q^{-1})\dot{} = -Q^{-1}\dot{Q}Q^{-1} \tag{5.11}$$

Then the positive definite matrices $\dot{Q}$ and $Q^{-1} = (W_u - W_{us}W_s^{-1}W_{su})^{-1}$ yield:

$$(Q^{-1})^{\cdot} = -(W_u - W_{us}W_s^{-1}W_{su})^{-1}\dot{Q}(W_u - W_{us}W_s^{-1}W_{su})^{-1} \tag{5.12}$$

which clearly defines a negative definite matrix.

Then via the Lyapunov stability Theorem one can readily recognize that $(W_u - W_{us}W_s^{-1}W_{su})^{-1}$ is stable and tends to zero as $t_f \to \infty$.

To verify the stability of off- diagonal blocks in (5.5):

$$-W_s^{-1}W_{su}(W_u - W_{us}W_s^{-1}W_{su})^{-1}, -W_u^{-1}W_{us}(W_s - W_{su}W_u^{-1}W_{us})^{-1}$$

let us define matrices below:

$$\begin{cases} P = (W_u - W_{us}W_s^{-1}W_{su})^{-1} \\ Q = W_s^{-1}W_{su}(W_u - W_{us}W_s^{-1}W_{su})^{-1} \end{cases} \tag{5.13}$$

Considering Theorem (5.1) and the positive definite matrices $W_c(0, t_f)$, $W_c^{-1}(0, t_f)$:

$$i)\begin{cases} W_s > 0 \\ W_s - W_{su}W_u^{-1}W_{us} > 0 \\ W_u > 0 \\ W_u - W_{us}W_s^{-1}W_{su} > 0 \end{cases}$$

$$ii)\begin{cases} P > 0 \\ P - Q^T(W_s - W_{su}W_u^{-1}W_{us})Q > 0 \end{cases} \tag{5.14}$$

According to part (ii) we have:

$$P - Q^T(W_s - W_{su}W_u^{-1}W_{us})Q > 0 \Rightarrow P > Q^T(W_s - W_{su}W_u^{-1}W_{us})Q \tag{5.15}$$

Considering part (i) of equation (5.14) we know that $Q^T(W_s - W_{su}W_u^{-1}W_{us})Q \geq 0$, then due to continuity that $t_f \to \infty$ the statement (5.15) implies that:

$$\lim_{t_f \to \infty} P \geq \lim_{t_f \to \infty} Q^T(W_s - W_{su}W_u^{-1}W_{us})Q \geq 0 \tag{5.16}$$

It was also proved in the previous part that $\lim_{t_f \to \infty} P = 0$ then:

$$0 \geq \lim_{t_f \to \infty} Q^T(W_s - W_{su}W_u^{-1}W_{us})Q \geq 0 \tag{5.17}$$

Hence according to the Squeeze Theorem one can readily conclude:

$$\lim_{t_f \to \infty} Q^T (W_s - W_{su} W_u^{-1} W_{us}) Q = 0 \tag{5.18}$$

Regarding part (i) we have $(W_s - W_{su} W_u^{-1} W_{us}) > 0$ then equation (5.18) implies that the off-diagonal block Q tends to zero as $t_f \to \infty$.

Therefore, with regard to the symmetry, it can readily be verified that both off-diagonal blocks in (5.5) are stable (i.e. they go to zero as $t_f \to \infty$), and they do not appear in the steady state form of the inverse of the controllability Gramian $W_c^{-1}(0, \infty)$.

Now consider the first diagonal block matrix $(W_s - W_{su} W_u^{-1} W_{us})^{-1}$ in (5.5). Using the Sherman–Morrison–Woodbury formula (Press, et al., 2007) we obtain:

$$(W_s - W_{su} W_u^{-1} W_{us})^{-1} = W_s^{-1} + W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} W_{us} W_s^{-1} \tag{5.19}$$

For simplicity, we define matrices P, Q as:

$$\begin{cases} P = (W_u - W_{us} W_s^{-1} W_{su})^{-1} \\ Q = W_s^{-1} W_{su} P \end{cases} \tag{5.20}$$

The statement (5.19) can be written as:

$$(W_s - W_{su} W_u^{-1} W_{us})^{-1} = \underbrace{W_s^{-1}}_{i} + \underbrace{Q P^{-1} Q^T}_{ii} \tag{5.21}$$

Then using the results of the previous parts the second term in (5.21) implies that:

$$\begin{cases} P^{-1}(0, t_f) > 0 \quad , 0 < t_f < \infty \\ \lim_{t_f \to \infty} Q P^{-1} Q^T = 0 \end{cases} \tag{5.22}$$

We know that the constant solution of continuous-time algebraic Lyapunov equations (CALE) is in fact the steady state solution (i.e. $t_f \to \infty$) of (CDLE) in (5.7) (Davis, et al., 2010). Then clearly the $\lim_{t_f \to \infty} W_s$ is the solution of (CALE) $A_s W_s + W_s A_s^T = -B_s B_s^T$ and part (i) in equation (5.21) yields a finite positive definite matrix, i.e. $0 < \lim_{t_f \to \infty} W_s^{-1}$.

Hence, we have shown that in the steady state condition as $t_f \to \infty$ all partitions of:

$$W_c^{-1}(0, t_f) = \begin{bmatrix} (W_s - W_{su} W_u^{-1} W_{us})^{-1} & -W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} \\ -W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} & (W_u - W_{us} W_s^{-1} W_{su})^{-1} \end{bmatrix}$$

tend to zero except the first diagonal block, which would be equal to the positive definite matrix $W_s^{-1}(\infty)$.

□

The result above shows that in the steady state conditions, i.e. $t_f \to \infty$, only the stable modes of LTI system can affect the inverse of the controllability Gramian. However, it is shown in the next section that the steady state of the minimum input energy is dependent to both stable and unstable modes of the system.

5.2.2. Minimum Input Energy for LTI System with Anti-Stable Part

According to (5.1), the minimum input energy is given as:

$$E_{\min} = (e^{At_f}x_0 - x_f)^T W_c^{-1}(0,t_f)(e^{At_f}x_0 - x_f)$$

Then, using Remark (5.2), we get:

$$E_{\min} = (e^{At_f}x_0 - x_f)^T \begin{bmatrix} W_s(0,t_f) & W_{su}(0,t_f) \\ W_{us}(0,t_f) & W_u(0,t_f) \end{bmatrix}^{-1} (e^{At_f}x_0 - x_f) \tag{5.23}$$

Next, we consider Theorem (5.3) in time domain:

Theorem (5.3): In the controllable LTI system (5.3) containing both stable and anti-stable parts, the minimum input energy under steady state conditions, i.e. $t_f \to \infty$ is:

$$E_{\min}(\infty) = \begin{bmatrix} (x_f)_s & (x_0)_u \end{bmatrix} \begin{bmatrix} W_s^{-1} & 0 \\ 0 & W_{-u}^{-1} \end{bmatrix} \begin{bmatrix} (x_f)_s \\ (x_0)_u \end{bmatrix} \tag{5.24}$$

Where:

$$\begin{cases} A_s W_s + W_s A_s^T = -B_s B_s^T \\ (-A_u)W_{-u} + W_{-u}(-A_u^T) = -B_u B_u^T \end{cases} \tag{5.25}$$

□

Proof:

Considering the minimum input energy equation (5.1) we have:

$$E_{\min}(t_f) = \{x_f^T W_c^{-1} x_f - x_f^T W_c^{-1} e^{At_f}x_0 - (e^{At_f}x_0)^T W_c^{-1} x_f + (e^{At_f}x_0)^T W_c^{-1} e^{At_f}x_0\} \tag{5.26}$$

Define:

$$\Rightarrow \begin{cases} v_1 = x_f^T W_c^{-1}(0,t_f)x_f \\ v_2 = -x_f^T W_c^{-1}(0,t_f)e^{At_f}x_0 \\ v_3 = -(e^{At_f}x_0)^T W_c^{-1}(0,t_f)x_f \\ v_4 = (e^{At_f}x_0)^T W_c^{-1}(0,t_f)e^{At_f}x_0 \end{cases} \tag{5.27}$$

Then:

$$E_{\min}(\infty) = \lim_{t_f \to \infty}(v_1 + v_2 + v_3 + v_4) \tag{5.28}$$

Thus, in the steady state:

$$\begin{cases} v_1^{\infty} = \lim\limits_{t_f \to \infty} v_1 = \lim\limits_{t_f \to \infty} x_f^T W_c^{-1}(0, t_f) x_f \\ v_2^{\infty} = \lim\limits_{t_f \to \infty} v_2 = -\lim\limits_{t_f \to \infty} x_f^T W_c^{-1}(0, t_f) e^{At_f} x_0 = 0 \\ v_3^{\infty} = \lim\limits_{t_f \to \infty} v_3 = -\lim\limits_{t_f \to \infty} (e^{At_f} x_0)^T W_c^{-1}(0, t_f) x_f = 0 \\ v_4^{\infty} = \lim\limits_{t_f \to \infty} v_4 = \lim\limits_{t_f \to \infty} (e^{At_f} x_0)^T W_c^{-1}(0, t_f) e^{At_f} x_0 \end{cases} \tag{5.29}$$

According to Theorem (5.2), part v_1^{∞} represents the minimum input energy of the stable modes of the system when $t_f \to \infty$ and is equal to $x_{fs}^T W_s^{-1}(0, \infty) x_{fs}$.

Now consider part the second term:

$$\begin{aligned} &v_2^{\infty} = \lim_{t_f \to \infty} v_2 = \lim_{t_f \to \infty} x_f^T W_c^{-1}(0, t_f)(e^{At_f} x_0) = x_f^T (\lim_{t_f \to \infty} W_c^{-1}(0, t_f) e^{At_f}) x_0 \Rightarrow \\ &v_2^{\infty} = x_f^T (\lim_{t_f \to \infty} \begin{bmatrix} (W_s - W_{su} W_u^{-1} W_{us})^{-1} & -W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} \\ -W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} & (W_u - W_{us} W_s^{-1} W_{su})^{-1} \end{bmatrix} e^{At_f}) x_0 \end{aligned} \tag{5.30}$$

The above equation can be stated as:

$$\begin{aligned} v_2^{\infty} &= x_f^T (\lim_{t_f \to \infty} \begin{bmatrix} (W_s - W_{su} W_u^{-1} W_{us})^{-1} & -W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} \\ -W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} & (W_u - W_{us} W_s^{-1} W_{su})^{-1} \end{bmatrix} \begin{bmatrix} e^{A_s t_f} & \overline{0} \\ \overline{0} & e^{A_u t_f} \end{bmatrix}) x_0 \\ &= x_f^T (\lim_{t_f \to \infty} \begin{bmatrix} (W_s - W_{su} W_u^{-1} W_{us})^{-1} e^{A_s t_f} & -W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} \\ -W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} e^{A_s t_f} & (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} \end{bmatrix}) x_0 \\ \Rightarrow &\begin{cases} i) \lim\limits_{t_f \to \infty} (W_s - W_{su} W_u^{-1} W_{us})^{-1} e^{A_s t_f} = 0 \\ ii) \lim\limits_{t_f \to \infty} W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} e^{A_s t_f} = 0 \\ iii) \lim\limits_{t_f \to \infty} W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} \\ iv) \lim\limits_{t_f \to \infty} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} \end{cases} \end{aligned} \tag{5.31}$$

Then it is easy to show $\lim\limits_{t_f \to \infty} v_2 = 0$.

Consider part (iv) ,as it is described in the proof of Theorem (5.2), the rate of the divergence of exponential matrix $e^{A_u t_f}$ when $t_f \to \infty$ is less than the rate of convergence of $(W_u - W_{us} W_s^{-1} W_{su})^{-1}$ then $\lim_{t_f \to \infty} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} = 0$. For the same reason, part (iii) converges to zero as the final time t_f increases to infinity, i.e. $\lim_{t_f \to \infty} W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} = 0$.

Then part v_2 is stable. By the same method, one can readily prove that part v_3 also tends to zero as $t_f \to \infty$.

Now let us consider the last part v_4, it could itself be broken into four different statements:

$$
\begin{aligned}
& v_4^{\infty} = \lim_{t_f \to \infty} (e^{At_f} x_0)^T W_c^{-1}(0, t_f) e^{At_f} x_0 \\
& \Rightarrow \begin{cases}
i) \lim_{t_f \to \infty} (e^{A_s t_f} x_{0s})^T (W_s - W_{su} W_u^{-1} W_{us})^{-1} e^{A_s t_f} x_{0s} \\
ii) \lim_{t_f \to \infty} (e^{A_s t_f} x_{0s})^T W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} x_{0u} \\
iii) \lim_{t_f \to \infty} (e^{A_u t_f} x_{0u})^T W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_s t_f} x_{0s} \\
iv) \lim_{t_f \to \infty} (e^{A_u t_f} x_{0u})^T (W_u - W_{us} W_s^{-1} W_{su})^{-1} e^{A_u t_f} x_{0u}
\end{cases}
\end{aligned}
\tag{5.32}
$$

We have shown in Theorem (5.2) that:

$$
\lim_{t_f \to \infty} (W_s - W_{su} W_u^{-1} W_{us})^{-1} = W_s^{-1} \tag{5.33}
$$

Now regarding the stability of A_s part (i) yields zero, i.e. $\lim_{t_f \to \infty} (e^{A_s t_f} x_{0s})^T W_s^{-1}(0, t_f) e^{A_s t_f} x_{0s} = 0$.

Part (ii) becomes zero as well in the steady state (i.e. $t_f \to \infty$) since $W_s^{-1} W_{su} (W_u - W_{us} W_s^{-1} W_{su})^{-1}$ is stable and the decay rate of its dynamic in the steady state condition is faster than the increment rate of the exponential term $e^{A_u t_f}$. In a similar way, it can be shown that part (iii) goes to zero as $t_f \to \infty$.

Lastly the block diagonal $(W_u - W_{us} W_s^{-1} W_{su})^{-1}$ could be expanded based on the Sherman–Morrison–Woodbury formula (Press, et al., 2007) then we obtain:

$$
(W_u - W_{us} W_s^{-1} W_{su})^{-1} = W_u^{-1} + W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} W_{su} W_u^{-1} \tag{5.34}
$$

Hence with regard to the previous parts we obtain:

$$
\lim_{t_f \to \infty} (e^{A_u t_f} x_{0u})^T (W_u^{-1} W_{us} (W_s - W_{su} W_u^{-1} W_{us})^{-1} W_{su} W_u^{-1}) e^{A_u t_f} x_{0u} = 0 \Rightarrow
$$

$$
\begin{aligned}
& iv = \lim_{t_f \to \infty} (e^{A_u t_f} x_{0u})^T W_u^{-1}(0, t_f) e^{A_u t_f} x_{0u} = x_{0u}^T (\lim_{t_f \to \infty} \alpha) x_{0u} \\
& \Rightarrow \lim_{t_f \to \infty} \alpha = \lim_{t_f \to \infty} (\int_0^{t_f} e^{-A_u^T t_f} e^{A_u t} B_u B_u^T e^{A_u^T t} e^{-A_u t_f} dt)^{-1} \\
& \Rightarrow \lim_{t_f \to \infty} \alpha = \lim_{t_f \to \infty} (\int_0^{t_f} e^{-A_u (t_f - t)} B_u B_u^T e^{-A_u^T (t_f - t)} dt)^{-1} \\
& \overset{t_f - t = \tau}{\Rightarrow} \lim_{t_f \to \infty} \alpha = \lim_{t_f \to \infty} (\int_0^{t_f} e^{-A_u \tau} B_u B_u^T e^{-A_u^T \tau} d\tau)^{-1} = W_{-u}^{-1}(0, \infty)
\end{aligned}
\tag{5.35}
$$

Then as $t_f \to \infty$ part v_4^{∞} converges to $x_{0u}^T W_{-u}^{-1}(0,\infty) x_{0u}$, which represents the energy needed to regulate the initial condition of the anti-stable subsystem.

□

In fact, the Theorem above supports the results we discussed in the previous chapter stating that in the steady state condition the minimum energy in stable system is a function of terminal time and the terminal states, therefore, in the steady state, i.e. $t_f \to \infty$ the minimum control energy is required, while in anti-stable systems we need no steering energy but an initial excitation to move the states from the origin to the terminal point in the controllable region when $t_f \to \infty$.

As (Zhou, et al., 1999) suggests one can also define the steady state form of the controllability Gramian in the frequency domain through Parseval's Theorem. Similar results can be derived for unstable systems also.

Proposition (5.1): (Zhou, et al., 1999) In the LTI system (5.2) if A is stable, we define the controllability Gramian in the frequency domain as[2]:

$$
W_c = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I - A)^{-1} B B^T (-j\omega I - A^T)^{-1} d\omega \tag{5.36}
$$

□

Corollary (5.2): In LTI system (5.3) W_c defined in Proposition (5.1) can be decomposed as:

$$
W_c = \begin{bmatrix} W_s & 0 \\ 0 & W_{-u} \end{bmatrix} \tag{5.37}
$$

where W_s, W_{-u} are the solutions of Lyapunov equations (5.25).

□

Proof is given in (Zhou, et al., 1999). The above Corollary uses the frequency domain to show that the controllability Gramian of LTI controllable systems with no eigenvalues on the imaginary axes can be decomposed into stable and anti-stable parts. This confirms that minimum input energy, which is

[2] For the simplicity, we use the same symbol for the controllability Gramian in the frequency domain and the time domain.

defined in equation (5.1) can be derived as in (5.24). Theorem (5.3) expresses the same results in the time domain.

5.3. Energy Levels for LTI Systems with Anti-Stable Modes

In this section, we consider the minimum and maximum values of required energy E_{min}^{min} , E_{min}^{max} transferring the states of the LTI system (5.2) with anti-stable modes from the initial point $x_0, \|x_0\| = R_0$ to the terminal states x_f , $\|x_f\| = R_f$ (within a very large terminal time $t_f \to \infty$) such that R_0, R_f are variables.

$$E_{\min}^{\min} = \min E_{min} \ , \quad E_{\min}^{\max} = \max E_{min}$$

In other words, we are interested to determine how much input energy we require to move the states of the LTI controllable system (5.2) from an arbitrary initial state a distance R_0 away from the origin to the states defined at a distance R_f away from the origin.

Assume $x_{0u} \neq 0$, $x_{fs} \neq 0$. According to Theorem (5.3) and Theorem (4.4), in the best case, where we need the minimum value of input energy, i.e. E_{min}^{min}, this transfer is done such that the stable modes of the system are transferred in the direction of the eigenvector corresposding to the maximum eigenvalue of W_s , and the unstbale modes start their movement in the direction of the eigenvector associated with the maximum eigenvalue of W_{-u}. Similarly, in the worst case where we require the maximum energy, this transfer would be done in the direction of the eigenvectors corresponding to the minimum eigenvalues of W_s for stable part and the initial direction of the movement of anti-stable part would be along the eigenvector associated with the minimum eigenvalue of W_{-u}.

Example (5.1): Assume the LTI system as: $A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ where the initial and final states can take values in the interval $[-1,1]$, i.e. $x_0 = \begin{bmatrix} -1 \le x_{01} \le 1 \\ -1 \le x_{02} \le 1 \end{bmatrix}, x_f = \begin{bmatrix} -1 \le x_{f1} \le 1 \\ -1 \le x_{f2} \le 1 \end{bmatrix}$. According to Theorem (5.3) the energy levels could be found as functions of both R_0 and R_f . Figure (5.1) illustrates the energy levels for this unstable system. It is concluded that as the norm of x_{0u} or/and x_{fs} increases the value of energy input increases as well, however this increase depends on the direction of the movement. Wherever $x_{0u} \neq 0$, $x_{fs} \neq 0$ if the movement starts in the direction of eigenvector corresponding to the minimum eigenvalue of W_{-u} for unstable modes and continues to reach the final states of stable modes in the direction of the eigenvector corresponding to the minimum eigenvalue of W_s then we will have the minimum increament, which means the maximum energy for transferring is used. If the unstable modes of system begin their movement along the direction of the eigenvector associated to the maximum eigenvalue of W_{-u} and the stable modes of the system move to the final states in the direction of the eigenvector corresponding to the maximum eigenvalue of W_s , the minimum value of energy woud be required.

In addition, for $R_0 \neq 0, R_f \neq 0$ it is still possible that the system starts its movement in the direction such that $x_{0u} = 0 \Rightarrow \|x_{0s}\| = R_0$ and moves toward the final state along the direction of $x_{fs} = 0 \Rightarrow \|x_{fu}\| = R_f$, in this case the value of the required energy is zero.

Furthermore, as Figure (5.1) illustartes in this perticular example, the value of minimum energy is more sensitive to the norm of initial states, R_0, rather than the distanc of final states from the origin, R_f, i.e. as R_f varies, for a fixed R_0, the value of minimum energy changes faster than the case that R_f is fixed and R_0 can change.

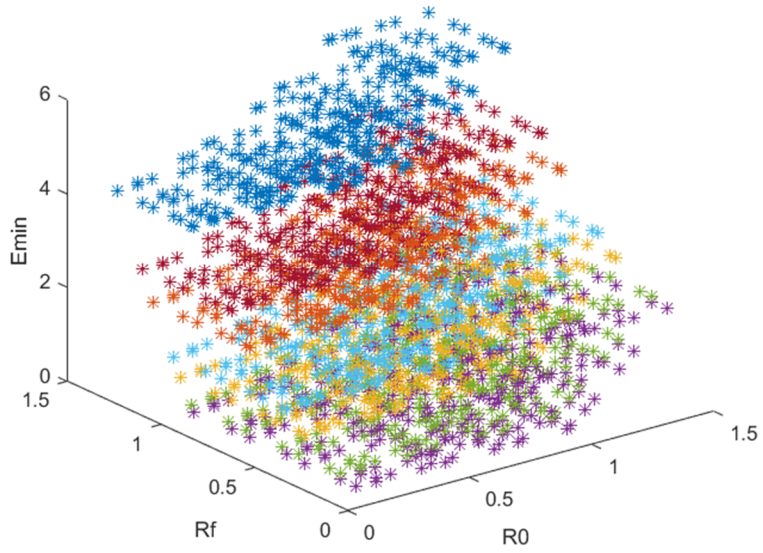


Figure (5.1): Minimum energy levels with respect to the norm of initial and terminal states (Example (5.1))

The importance of the shown energy levels is where we have a constraint on the value of control energy. This determines how far the system may move from the initial states defined by the distance R_0 away from the origin, i.e. $\|x_0\| = R_0$.

□

Additionally, it can be readily concluded from (5.24) that the unstable part contribution vanishes when the unstable modes of the LTI system start their movement from zero. In this special case $V^{-1}x_0 = [x_{0s}, 0] \Rightarrow \|x_{0s}\| = R_0$, $V \in \mathbb{R}^n$ is the transformation defined in Remark (5.1) and x_{0s} represents the stable modes' initial states. Therefore, the input energy would be just a function of the stable part, and the energy levels in new coordination ($x \to V^{-1}x$) could be determined as it is explained in chapter 4. Consider the 3-dimensional unstable system in the next example.

Example (5.2): Consider the LTI system:

$$A=\begin{bmatrix}10 & 0 & 5\\ 0 & -12 & 0\\ 1 & 0 & -30\end{bmatrix}, B=\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}$$

The eigenvalues are $\lambda_1 = -30.1246, \lambda_2 = -12, \lambda_3 = 10.1246$. Assume we wish to move the states from the origin to a final state located on the sphere with radius $R_f\ (=1,2,3)$. The steady state energy levels are shown in Figures below. Figure (5.2) demonstrates the value of minimum energy in the state space. In Figure (5.3) eigenvectors of W_s are shown, which show the directions in which the minimum and maximum of energy levels occur.

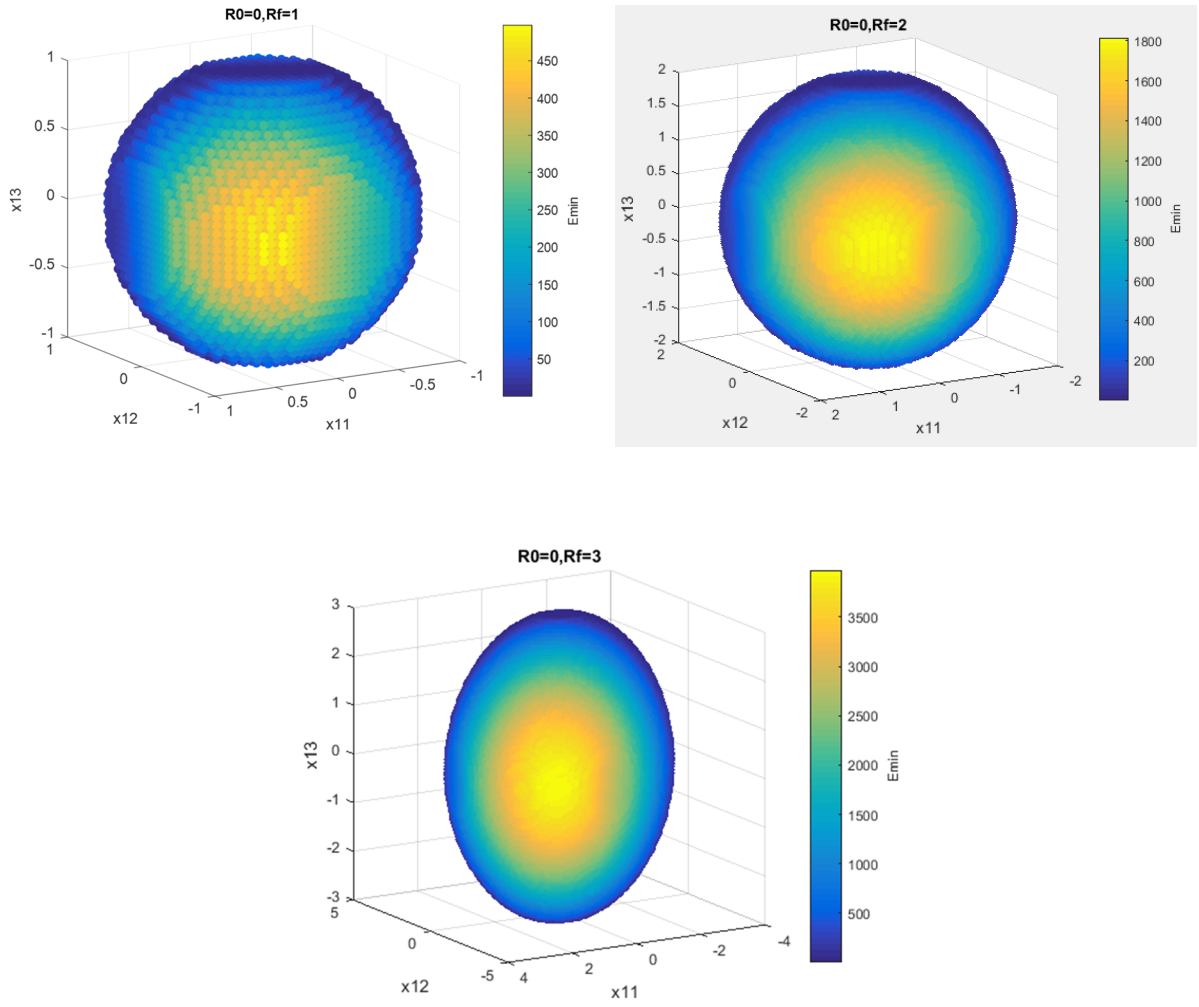


Figure (5.2): Minimum energy levels of Example (5.2)

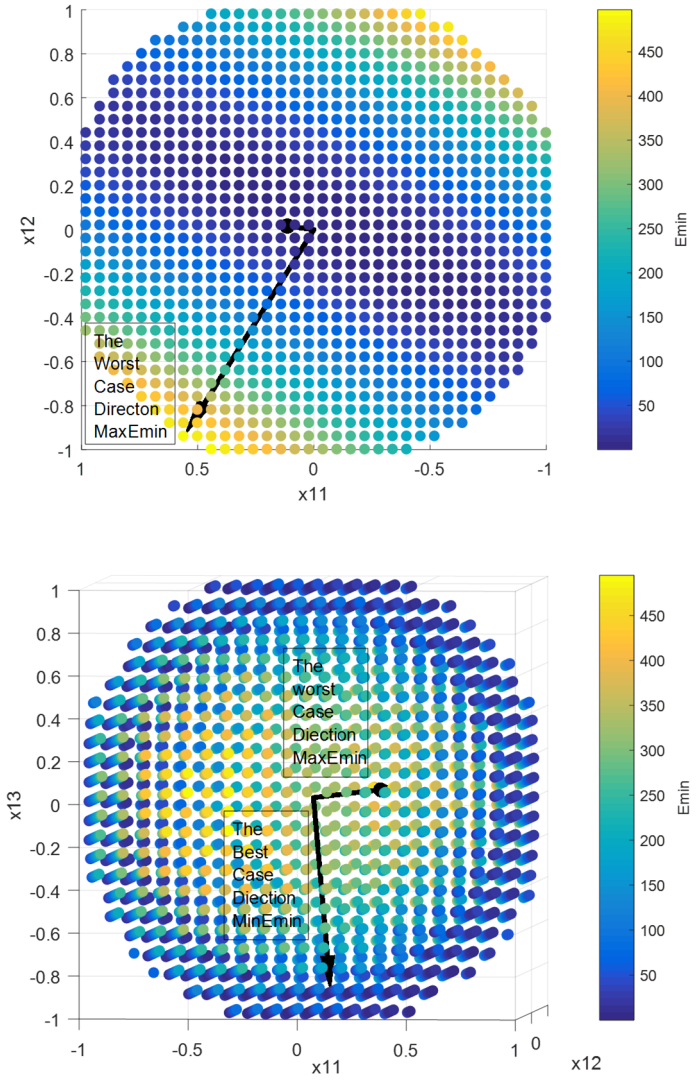


Figure (5.3): Max and Min energy levels' directions in Example (5.2)

□

Conversely, if $x_{0u} \neq 0$, no energy is required to steer the stable modes of the LTI system to the origin, then if $V^{-1}x_f = [0, x_{fu}]$, $\|x_{fu}\| = R_f$, where $V \in \mathbb{R}^n$ is the transformation defined in Remark (5.1) and x_{fu} represents the final states of the unstable subsystem, the input energy could be determined just as a function of the unstable part. Thus, the energy levels can be defined in the new coordinate frame ($x \to V^{-1}x$)as explained in chapter 4. For stable systems, however, the final states would be substitute by the initial states of unstable modes x_{0u}, and the inverse of controllability Gramian can modified as the inverse of controllability Gramian of the anti-stable part as it is described in Theorem (5.3), W_{-u}^{-1}. Figures (5.4), (5.5) show the energy levels for example (5.3) in this particular case, where R_0 varies as the integer in the interval $[1,3]$.

Example (5.3): Assume that we wish to transfer the state of the LTI system:

$$A=\begin{bmatrix}-3 & 0 & 5\\ 1 & 40 & 0\\ 5 & 0 & 10\end{bmatrix}, B=\begin{bmatrix}1\\ 1\\ 0\end{bmatrix}$$

from an initial state on a sphere centered at the origin with the radius $R_0 (=1,2,3)$ to the origin within a large terminal time $t_f \to \infty$. Figure (5.4) describes the minimum energy levels in the state space. These energy levels determine that in the worst case how much energy is needed to transfer the system from the initial states on a sphere with the radius R_0 to the origin. In Figure (5.5) the direction of eigenvectors of W_{-u} are illustrated. As it can be seen the energy required to derive the system from an initial point decreases as $\|x_{0u}\|$ decreases. Furthermore, if $x_{0u} \neq 0$ the required energy decreases in the direction of the eigenvector corresponding to the maximum eigenvalue of W_{-u} and increases along the direction of the eigenvector associated with the minimum eigenvalue of W_{-u}.

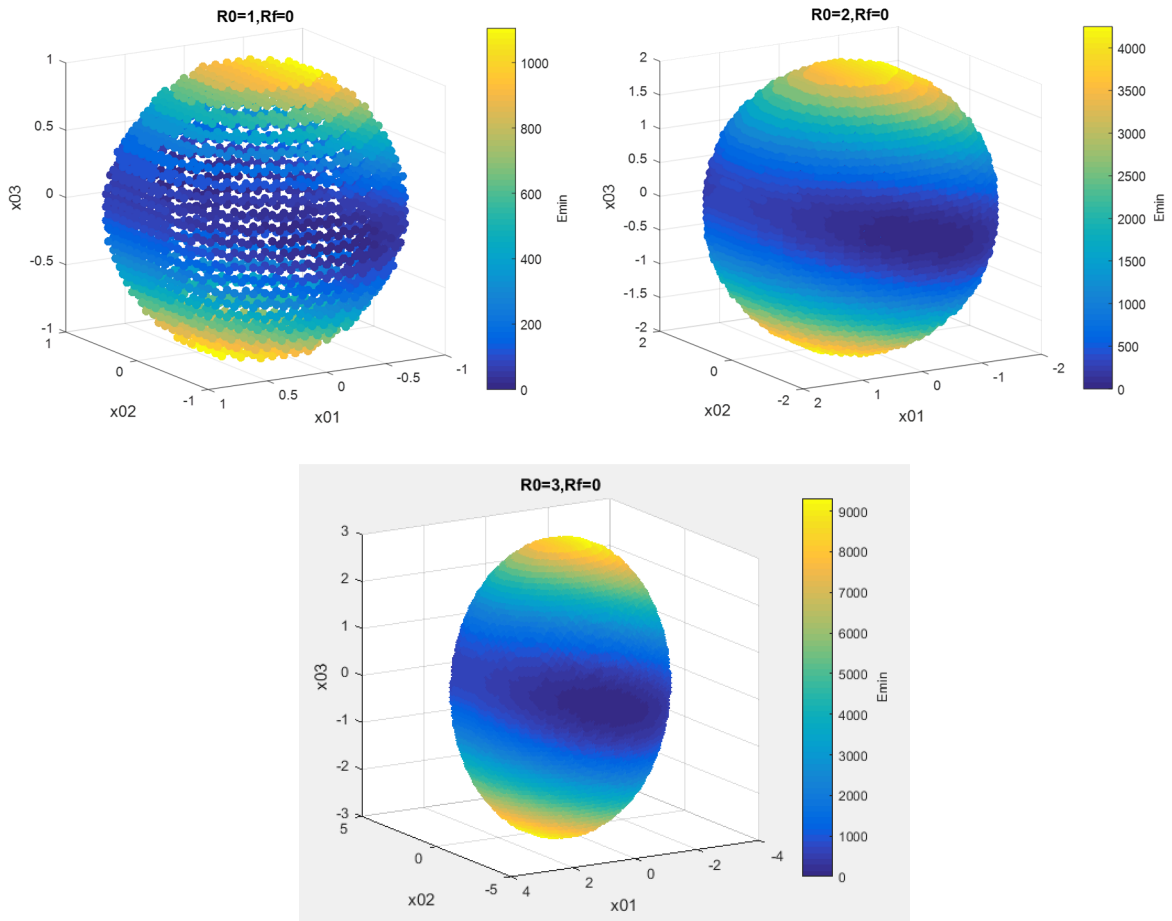


Figure (5.4): Minimum energy levels of Example (5.3)

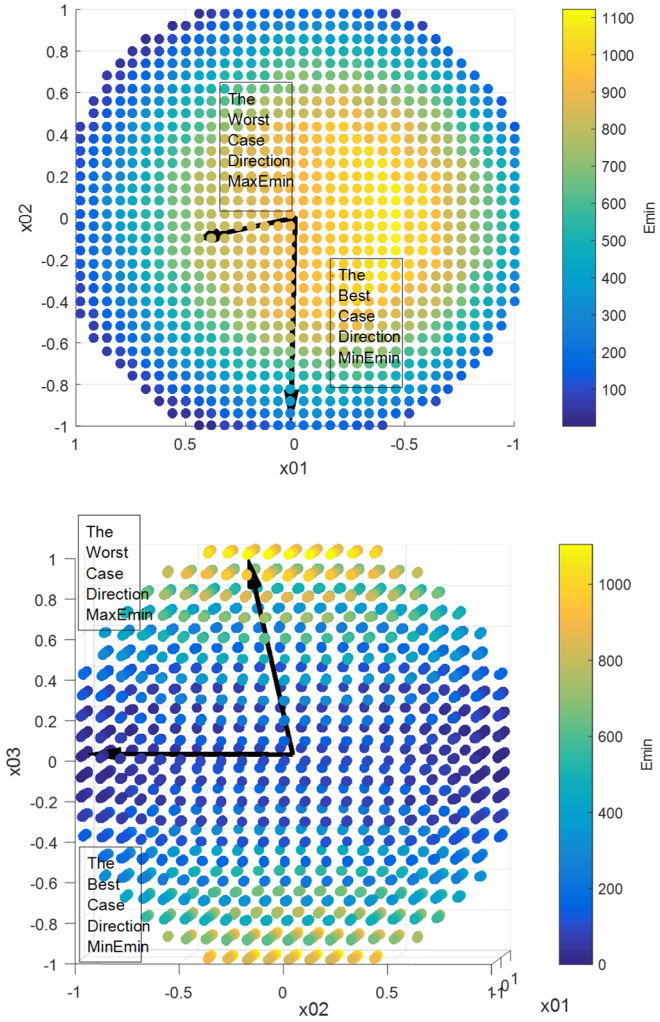


Figure (5.5): Max and Min energy levels' directions in Example (5.3)

□

Furthermore, in the case that the initial state $\|x_0\| = R_0 \neq 0$ is fixed, the minimum input energy levels can be defined based on the scaled extremum eigenvalues of W_s^{-1} via the norm of terminal state vector for the stable part, i.e. $\|x_{fs}\| \leq R_f$, which are added by a constant value $\lambda_{\min}^{W_{-u}} R_0^{\,2}$ related to the unstable part. $\lambda_{\min}^{W_{-u}}$ represents the smallest eigenvalue of W_{-u}. Figure (5.6) and (5.7) describe the energy levels for example (5.2), in the case that R_0 is fixed at the values of 100 and 1000.

The importance of these energy levels is that for an upper bounded input energy they determine how far the system is guaranteed to be transferred from an initial state located on a sphere centered at the origin with radius $R_0 = cte$.

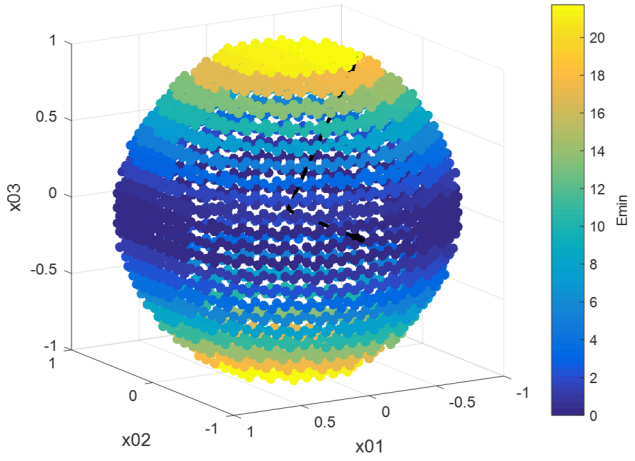


Figure (5.6): Pure energy values taken by the unstable subsystem when $R_0 = 1$ and $R_f = 0$ in Example (5.2)

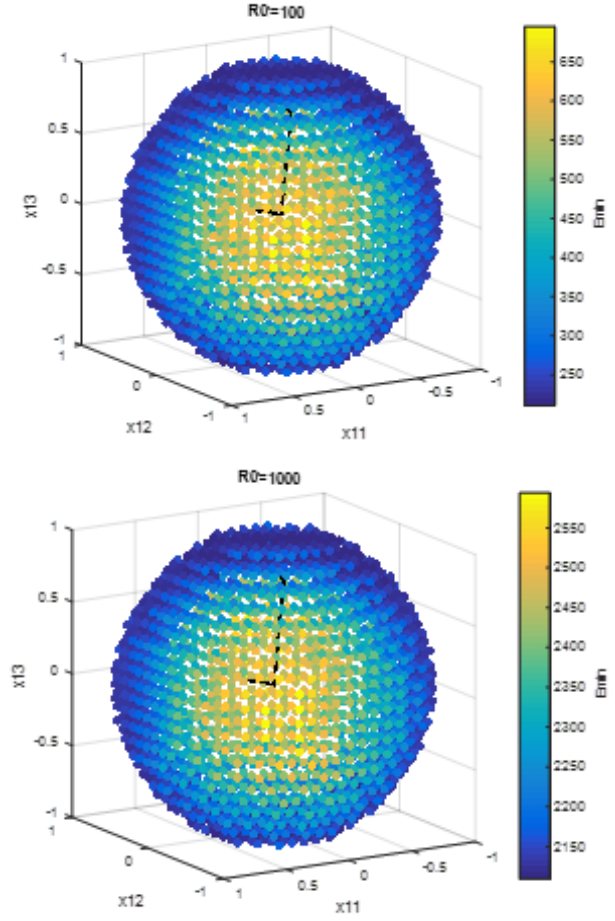


Figure (5.7): Minimum energy levels when the system steers from the worst initial direction $R_0 (= 100, 1000)$ and moves toward the final states located on the sphere with the radius $R_f = 1$ in different directions in Example (5.2)

In a similar manner, if $R_f \neq 0$ is a fixed given value, the minimum input energy levels are defined as the extremum eigenvalues of W_{-u}^{-1} by the scaled norm of unstable modes' initial states $\|x_{0u}\| \leq R_0$, which are increased by a constant value $\lambda_{\min}^{W_s} R_f{}^2$ related to the stable part.

$\lambda_{\min}^{W_s}$ defines the minimum eigenvalue of W_s. In this case, the energy levels describe the maximum distance R_f the system can move away from the origin using a limited amount of input energy. Figure (5.8) shows this minimum energy value when $x_{0u} = 0$ and $R_f = 1$ in example (5.3). Figure (5.9) depicts the energy levels for two fixed values $R_f (=100,1000)$.

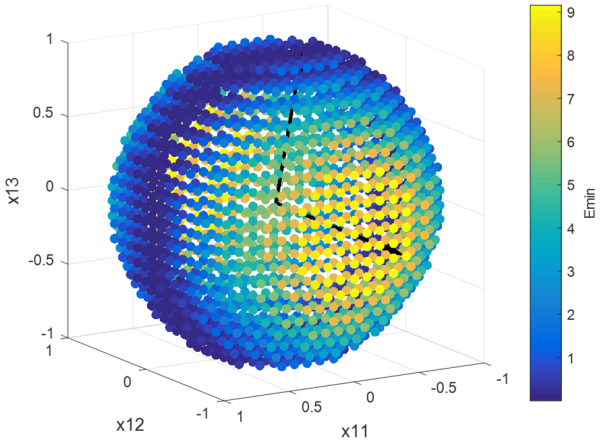


Figure (5.8): Pure energy values taken by the stable subsystem when $R_f = 1$ and $R_0 = 0$ in Example (5.3)

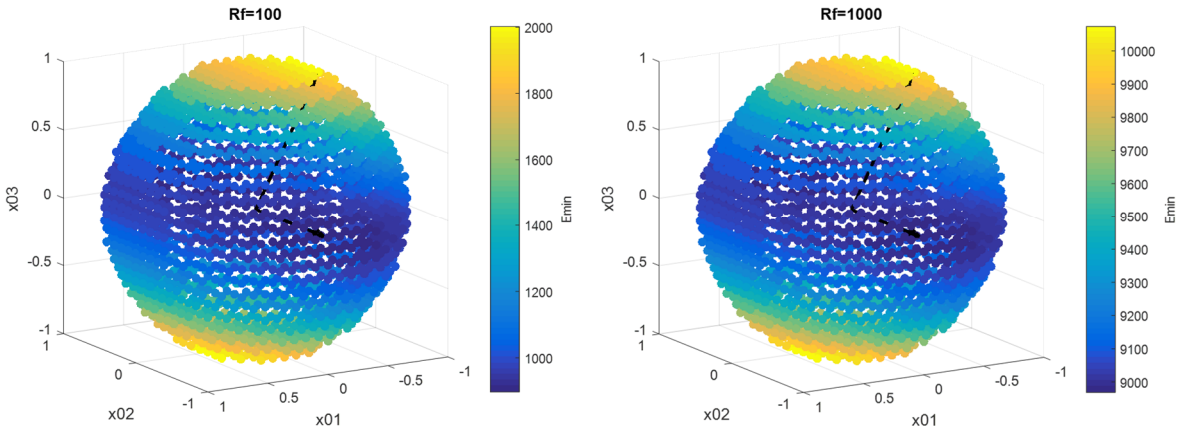


Figure (5.9): Minimum energy levels when the system moves from various initial directions $R_0 = 1$ toward the final states located on the sphere with the radius $R_f (=100,1000)$ in the worst direction in Example (5.3)

The energy levels discussed in this section imply that if the LTI system is anti-stable then the problem of transferring to zero requires more energy than the reachability problem, which deals with the problem of transferring the system from the origin $\|x_0\| = R_0 = 0$ to a terminal state $x_f, \|x_f\| = R_f > 0$, since the anti-stable systems tend to amplify the magnitude of the state.

In the case that the LTI system has a combination of stable and anti-stable parts, the selection of the initial and final states may cause simplifying or complicating the reachability problem and the controllability to 0 problem.

5.4. The Energy and Relative Degree of Controllability

In many practical cases, controllability should not be treated as a binary concept; that either a system is controllable or not. In many classes of control problems, it is also important to know how controllable a system is. Consider an uncontrollable system, if we start from a set including the sufficient number of actuators it will usually be the case that moving one of the actuators a small distance $\varepsilon > 0$ a controllable system will result. For a very small ε, even though technically the system is controllable, in some sense it is almost uncontrollable (Viswanathan, et al., 1984).

According to (Burgmeier, 1992) consider the LTI system:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{l\times n}, D \in \mathbb{R}^{l\times m}, x \in \mathbb{R}^n \tag{5.38}$$

Then $x(t) = x(t_f, t_0, x_0, u), t_f \geq t_0$ denotes the unique solution of this system for a given initial condition $x(t_0) = x_0 \in \mathbb{R}^n$ and $u \in U$ where U is the set of all admissible controls for the system (5.38).

For a given non-empty set of final states $X_f \in \mathbb{R}^n$ we define the X-controllability set of the system at time $t_f \geq t_0$ by $S(t_0, t_f, U, X_f) = \{x_0 \in \mathbb{R}^n, u \in U, x(t_f, t_0, x_0, u) \in X_f\}$.

As simple measures of the size of the X-controllability set at time $t_f \geq t_0$ with respect to a given point $y \in \mathbb{R}^n$ we define two degrees of controllability and uncontrollability as:

Definition (5.1): (Burgmeier, 1992) in LTI system (5.38) the degree of controllability with respect to a given point $y \in \mathbb{R}^n$ is defined as:

$$\rho(t_0, t_f, U, X_f) = \inf\{\|x - y\| : x \in \mathbb{R}^n, x(t_f, t_0, x_0, u)\} \tag{5.39}$$

And the degree of uncontrollability could be stated as:

$$\rho(t_0, t_f, U, X_f) = \sup\{\|x - y\| : x \in \mathbb{R}^n, x(t_f, t_0, x_0, u)\} \tag{5.40}$$

Where $\|.\|$ represents the Euclidian norm.

□

Definition (5.2): (Burgmeier, 1992) The recovery region for time T for normalized system (5.38) is the set $\Re = \{x_f \in X : \exists u \in U, \|u\|_{[0,t_f]} \leq 1, x(t_f, t_0, x_0, u)\}$.

□

Then, the recovery region denotes the reachability set of the system if we consider the problem of transferring the initial states x_0 to the origin, within the finite time t_f using the control input $u, \|u\| \leq 1$.

Therefore, as explained in (Viswanathan, et al., 1984) the recovery region identifies all the final states to which at least one initial state can be transferred in an interval using bounded energy control. Also, the degree of controllability is a scalar measure of the size of the region chosen as the shortest distance from the final point to a given point $y \in \mathbb{R}^n$, which is the desired terminal state.

Then, the control input energy can be employed as the degree of controllability to demonstrate the importance or lack of importance of each control input/actuator.

Example (5.4): Consider the linearized MIMO model of the active magnetic bearing (AMB) system (Noshadi, et al., 2016), (Schweitzer & Maslen, 2009), (Mazenc, et al., 2005):

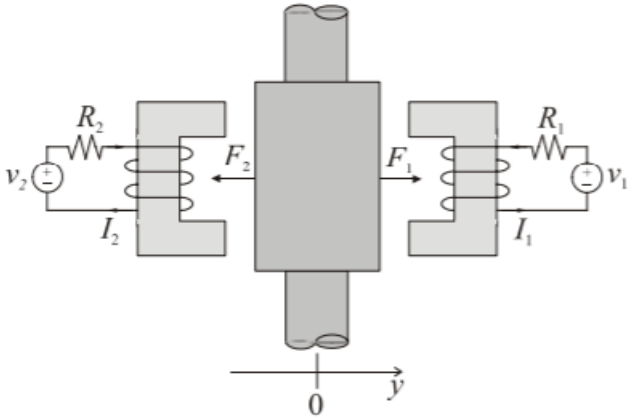


Figure (5.10): Schematic of the AMB system

AMBs can provide contactless suspension of the rotor by attractive forces produced by electromagnets. They are employed in a variety of rotating machines in place of conventional mechanical bearings.

In this example, the 18th order AMB has 4 coil currents as inputs and the outputs are defined as the position and the orientation of rotors.

$$A_{18\times18} = diag\left\{\begin{bmatrix}-0.0359 & 12900\\ -12900 & -0.0359\end{bmatrix},\begin{bmatrix}-0.55 & 4826\\ -4826 & -0.55\end{bmatrix},\begin{bmatrix}-0.69 & 4835\\ -4835 & -0.69\end{bmatrix},\begin{bmatrix}-152 & 335.8\\ -335.8 & -152.9\end{bmatrix}\right.$$
$$\left.\begin{bmatrix}-120.9 & 341.8\\ -341.8 & -120.9\end{bmatrix},\begin{bmatrix}-111.8 & 228.2\\ -228.2 & -111.8\end{bmatrix},\begin{bmatrix}-70.51 & 212.4\\ -212.4 & -70.51\end{bmatrix},191.4,\ 243.4,\ 272.4,\ 284.5\right\}$$

$$B_{18\times4} = \begin{bmatrix}
-0.0243 & 0.0176 & -0.0235 & -0.0193\\
0.0453 & -0.0672 & 0.0486 & 0.0839\\
0.5450 & 0.0348 & -0.5991 & 0.0686\\
-1.4730 & 0.2672 & 1.7560 & 0.3069\\
-0.1713 & -1.4250 & 0.2059 & -1.7710\\
0.0820 & -0.7334 & -0.0522 & -0.8909\\
-5.2730 & 4.2650 & -4.8140 & -3.4230\\
-1.7190 & -3.8070 & 3.0520 & 9.5530\\
-5.1950 & 2.4130 & -8.4050 & -3.1360\\
-1.4100 & 1.6070 & 1.4170 & -3.6310\\
-2.0850 & -1.2980 & -0.3873 & 4.2370\\
-2.0320 & 5.3700 & -1.7270 & 5.5080\\
7.4800 & -0.7283 & -5.0160 & -1.0070\\
1.1000 & 0.9193 & 1.7110 & 1.5220\\
8.7810 & -21.0600 & 0.5281 & 33.6300\\
-44.7700 & -16.0000 & -53.4300 & 1.0640\\
-50.9400 & -8.0200 & 65.9800 & 13.8900\\
34.1700 & -55.9100 & -50.3000 & -44.6800
\end{bmatrix}
C^T_{4\times18} = \begin{bmatrix}
7.2300 & -2.1590 & -1.5950 & 6.9340\\
-5.5360 & 16.7600 & -18.1900 & -20.8700\\
-10.4100 & 0.5644 & 16.4900 & 1.6760\\
-5.5310 & 1.7130 & 3.7640 & 1.2470\\
-1.2190 & -6.6440 & 1.9040 & -1.9720\\
0.7982 & 11.4200 & -1.7800 & 14.4700\\
-1.3220 & -15.4500 & 0.3159 & 9.0980\\
6.2330 & 9.8390 & 4.5670 & -3.8710\\
12.8700 & 16.3700 & 9.2420 & -3.2940\\
8.9640 & -0.9097 & 3.0000 & 1.0810\\
3.3940 & 1.7800 & 1.0590 & -15.4300\\
-3.4690 & -17.5500 & 3.2230 & -10.5200\\
-8.5000 & -1.9170 & 10.1500 & -1.3500\\
4.0930 & -2.0230 & -8.2030 & -2.2680\\
0.6677 & -2.5790 & 0.3406 & 3.1050\\
-2.1520 & -0.2261 & -1.9370 & 0.4813\\
-1.1260 & -1.9160 & 1.4620 & -0.5211\\
0.6849 & -1.9850 & 0.2637 & -1.0290
\end{bmatrix}$$

Active magnetic bearing (AMB) systems are intrinsically unstable as they have some poles in the right half plane (RHP). One can easily decouple the whole model into stable and anti-stable parts using a linear transformation. In this case the eigenvalues are:

$$\lambda_k = \begin{cases}
\text{-0.0001 - 1.2900i , -0.0001 + 1.2900i , -0.0001, - 0.4826i , -0.0001 + 0.4826i}\\
\text{-0.0001 - 0.4835i , -0.0001 + 0.4835i , -0.0152 - 0.0336i , -0.0152 + 0.0336i}\\
\text{-0.0221 , -0.0112 - 0.0228i , -0.0112 + 0.0228i}\\
\text{-0.0071 - 0.0212i , -0.0071 + 0.0212i}\\
0.0191^{>0}\text{ , }0.0243^{>0}\text{ , }0.0272^{>0}\text{ , }0.0285^{>0}\text{, }0.0463^{>0}
\end{cases}$$

$k = 1, 2, \ldots 18$

$A_s \in \mathbb{R}^{13\times13}, A_u \in \mathbb{R}^{5\times5}, B_s \in \mathbb{R}^{13\times4}, B_u \in \mathbb{R}^{5\times4}$ can be found in the Appendix.

Now, let us define the initial condition and the terminal state as:

$$x_0 = \begin{bmatrix}1 & 0 & \cdots & 0 & 0.2 & 0 & 0 & 0\end{bmatrix}^T, x_f = \begin{bmatrix}2 & 0 & \cdots & 0 & 0.3 & 0 & 0 & 0\end{bmatrix}^T$$

We can achieve the values for the degree of controllability proposed in this section by using the Lyapunov equations for the above decoupled matrices, and we get:

$$\begin{cases} E_{\min_1} = E_{s\{B_1\}} + E_{u\{B_1\}} = 217.3633 + 6.9140e{+}04 = 6.9357e{+}04 \\ E_{\min_2} = E_{s\{B_2\}} + E_{u\{B_2\}} = 119.0317 + 1.2020e{+}04 = 1.2139e{+}04 \\ E_{\min_3} = E_{s\{B_3\}} + E_{u\{B_3\}} = 511.1147 + 1.9115e{+}07 = 1.9116e{+}07 \\ E_{\min_4} = E_{s\{B_4\}} + E_{u\{B_4\}} = 95.7876 + 4.7137e{+}03 = 4.8095e{+}03 \end{cases}$$

Here $E_{\min_i}, i = 1,2,3,4$ denote the minimum input energy for i[th] input, and $E_{s\{B_i\}} E_{u\{B_i\}}, i = 1,2,3,4$ are the steady state energy values stored in the stable and anti-stable subsystems related to i[th] input respectively.

The results show that for this initial condition and final state, it is efficient to apply the last input since it corresponds to the minimum energy.

Now let us find DOC by a different set of initial and final state:

$$x_0 = \begin{bmatrix} 1 & 0 & \cdots & 0 & 1 & 0 & 0 & 0 & 2 \end{bmatrix}^T, x_f = \begin{bmatrix} 2 & 0 & \cdots & 0 & 2 & 0 & 0 & 0 & 0.5 \end{bmatrix}^T$$

In this case, we get:

$$\begin{cases} E_{\min_1} = E_{s\{B_1\}} + E_{u\{B_1\}} = 1.5376e{+}04 + 3.1321e{+}08 = 3.1322e{+}08 \\ E_{\min_2} = E_{s\{B_2\}} + E_{u\{B_2\}} = 7.5852e{+}04 + 1.2252e{+}08 = 1.2260e{+}08 \\ E_{\min_3} = E_{s\{B_3\}} + E_{u\{B_3\}} = 7.5184e{+}04 + 1.4773e{+}08 = 1.4780e{+}08 \\ E_{\min_4} = E_{s\{B_4\}} + E_{u\{B_4\}} = 9.8826e{+}04 + 1.4857e{+}08 = 1.4867e{+}08 \end{cases}$$

The achieved results are different from the first case. Here the 2[nd] input leads to less steady state input energy and thus it is more controllable relative to the other inputs. The results also show that this approach is sensitive to the initial and final conditions, i.e. the initial and terminal states may vary the result. In this case the proper initial and terminal conditions that coincide with the control objective should always be selected (Lee & Park, 2014).

□

To remove this sensitivity and increase the reliability of the input selection, we can define the degree of controllability (DOC) as the worst-case metric where we evaluate the maximum of minimum energy, which specifies the largest possible amount of energy required to transfer the states from any initial point to any final states located in the sphere with the center in the origin and the radius R, i.e. $\|x_{sf}\| \le R, \|x_{u0}\| \le R$ (without loss of generality one can simply normalize x_0, x_f). In other words, wherever we have not enough information of the initial and final conditions of the system the presented energy DOC measure for testing the controllability of a system is not satisfactory in the sense it may lead to wrong conclusions.

Therefore, we introduce another controllability measure, based on the maximum value of minimum energy as a metric which seems to be useful, especially in the case that we are designing the system for various initial and final conditions.

We can readily define the above energy measure as a linear combination of eigenvalues of the inverse of controllability Gramians W_s^{-1}, W_u^{-1} for the stable and anti-stable subsystems respectively, where the coefficients are determined through the initial and final states.

In the worst case, according to the explanations given in section 5.3, the normalized initial condition must represent the eigenvector corresponding to the maximum eigenvalue of W_u^{-1} and the terminal states would be in the same direction of the eigenvector corresponding to the maximum eigenvalue of W_s^{-1}. Then DOC measure could be defined as the summation of maximum eigenvalues of W_u^{-1} and W_s^{-1}.

If we use this DOC measure in our example, we get:

$$\begin{cases} E_{\min_1})_{\max} = E_{s\{B_1\}} + E_{u\{B_1\}})_{\max} = 3.6377e{+}04 + 1.5978e{+}08 = 1.5982e{+}08 \\ E_{\min_2})_{\max} = E_{s\{B_2\}} + E_{u\{B_2\}})_{\max} = 1.6471e{+}05 + 2.9201e{+}09 = 2.9202e{+}09 \\ E_{\min_3})_{\max} = E_{s\{B_3\}} + E_{u\{B_3\}})_{\max} = 9.4060e{+}04 + 5.5939e{+}08 = 5.5948e{+}08 \\ E_{\min_4})_{\max} = E_{s\{B_4\}} + E_{u\{B_4\}})_{\max} = 6.8675e{+}04 + 1.7343e{+}10 = 1.7343e{+}10 \end{cases}$$

The results show that the first input could be a reasonable choice if we are interested in the robustness of degree of controllability over the set of all possible initial and final states, though the choosing other inputs might result in a better degree of controllability for some specific initial and final conditions, they do not provide guaranteed degree of controllability, i.e. there might exist some initial or final states which lead to nearly uncontrollable system.

Remark (5.3): Consider the LTI system (5.30) with stable and anti-stable modes, with defined initial and final states the energy optimal degree of controllability (DOC) could be defined as the minimum input energy in the steady state defined by Theorem (5.3).

$$E_{\min}(\infty) = \begin{bmatrix} (x_f)_s & (x_0)_u \end{bmatrix} \begin{bmatrix} W_s^{-1} & 0 \\ 0 & W_u^{-1} \end{bmatrix} \begin{bmatrix} (x_f)_s \\ (x_0)_u \end{bmatrix} \tag{5.41}$$

□

Remark (5.4): Consider the LTI system (5.38) with stable and anti-stable modes. The worst-case energy optimal degree of controllability (WDOC) can be defined as:

$$E_{\min}(\infty))_{\max} = \begin{bmatrix} (V_{\max}^{W_s^{-1}})^T{}_s & (V_{\max}^{W_u^{-1}})^T{}_u \end{bmatrix} \begin{bmatrix} W_s^{-1} & 0 \\ 0 & W_u^{-1} \end{bmatrix} \begin{bmatrix} (V_{\max}^{W_s^{-1}})_s \\ (V_{\max}^{W_u^{-1}})_u \end{bmatrix} \tag{5.42}$$

Where $V_{\max}^{W_s^{-1}}$ shows the eigenvector corresponding to the maximum eigenvalue of W_s^{-1} and $V_{\max}^{W_u^{-1}}$ determines the eigenvector associated with the maximum eigenvalue of W_u^{-1}.

□

Let us investigate further the difference of two proposed DOC measures using the example of linearized seesaw-pendulum.

We refer the interested readers for details on the linearized model of seesaw-pendulum to (Subbotin, 2004), (Jirstrand & Gunnarsson, 2001).

Example (5.5): consider the seesaw-pendulum process in Figure (5.11).

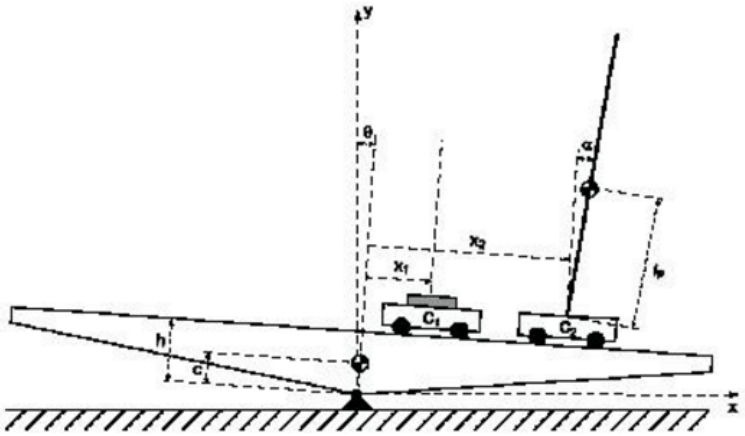


Figure (5.11): The schematic of seesaw-pendulum process

In this system, we have an inverted pendulum coupled with a second cart which can move in parallel on top of an unstable structure - a seesaw. This system has two inputs: a force applied to the cart with the pendulum installed on top of it and a force applied to the load cart which tries to balance the seesaw structure.

Let us define the state vector as:

$$x = \begin{bmatrix} x_1 & \theta & x_2 & \alpha & \dot{x}_1 & \dot{\theta} & \dot{x}_2 & \dot{\alpha} \end{bmatrix}^T$$

where x_1 is the position of the load cart, θ is the angle of the seesaw structure with respect to the vertical axis, x_2 represents the position of the pendulum cart and α denotes the angle of the pendulum with respect to the seesaw symmetry axis, then we can write the linearized model of the system as:

$$\dot{x} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} x + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} [u_1, u_2]$$

Have:

$$A_{11} = 0_{4\times4}$$
$$A_{12} = I_4$$
$$A_{21} = \begin{bmatrix} -ghm_1/J & g - cghm_2/J & -gh(m_2+m_p)/J & 0 \\ gm_1/J & cgm_2/J & g(m_2+m_p)/J & 0 \\ -ghm_1/J & g - cghm_2/J & -gh(m_2+m_p)/J & -gm_p/m_2 \\ -gm_1/J & -cgm_2/J & -g(m_2+m_p)/J & g(m_2+m_p)/l_p m_2 \end{bmatrix}$$
$$A_{22} = \begin{bmatrix} -\alpha(J+m_1h^2)/m_2J & 0 & -\alpha h^2/J & 0 \\ \alpha h/J & 0 & \alpha h/J & 0 \\ -\alpha h^2/J & 0 & -\alpha(J+m_2h^2)/J & 0 \\ -\alpha h/J & 0 & -\alpha(l_p hm_2 - J)/m_2 l_p & 0 \end{bmatrix}$$
$$b_1 = 0_{4\times2}, b_2 = \begin{bmatrix} (J+m_1h^2)/m_1J & h^2/J \\ -h/J & -h/J \\ h^2/J & (J+m_2h^2)/J \\ h/J & (l_p hm_2 - J)/m_2 l_p J \end{bmatrix}$$

Assuming all parameters are equal to 1 then we get the LTI system below:

$$A_{8\times8} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & -2 & 0 & -2 & 0 & -1 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 & 1 & 0 \\ -1 & 0 & -2 & -1 & -1 & 0 & 2 & 0 \\ -1 & -1 & -2 & 2 & -1 & 0 & 0 & 0 \end{bmatrix}, B_{8\times2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 2 & 1 \\ -1 & -1 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}$$

This system has three anti-stable modes. Using a similarity transformation, we can split it into a 5-dimensional stable part and a 3-dimensional anti-stable subsystem.

Let us consider the first case, where the initial and final conditions are given as:

$$x_0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T, x_f = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

Then the results of the first DOC are achieved as:

$$\begin{cases} E_{\min_1} = E_{s\{B_1\}} + E_{u\{B_1\}} = 1.6358\text{e+}05 + 5.8621 = 1.6358\text{e+}05 \\ E_{\min_2} = E_{s\{B_2\}} + E_{u\{B_2\}} = 3.4472\text{e+}33 + 29.9975 = 3.4472\text{e+}33 \end{cases}$$

Recall that $\{B_i\}, i = 1, 2$ means the part of the B matrix related to i[th] input.

As shown in the results, it is efficient to apply force to the cart with pendulum to change the angle of the seesaw structure.

Now consider the second case, where the angle of pendulum changes from 1 to 2:

$$x_0 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T, x_f = \begin{bmatrix} 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

Then, the values of the proposed measure are:

$$\begin{cases} E_{\min_1} = E_{s\{B_1\}} + E_{u\{B_1\}} = 1.4566e{+}06 + 24.1888 = 1.4566e{+}06 \\ E_{\min_2} = E_{s\{B_2\}} + E_{u\{B_2\}} = 4.4568e{+}04 + 1.8663e{+}02 = 4.4755e{+}04 \end{cases}$$

In this case, the results show that it is efficient to apply force to the load cart to change the pendulum's angle.

Note also that DOC result of the system varies with the initial and final conditions.

While the second DOC measure, which illustrates the degree of controllability in the worst case, leads to the same values for all normalized initial and final states:

$$\begin{cases} E_{\min_1})_{\max} = E_{s\{B_1\}} + E_{u\{B_1\}})_{\max} = 1.0588e{+}03 + 32.5079 = 1.0913e{+}03 \\ E_{\min_2})_{\max} = E_{s\{B_2\}} + E_{u\{B_2\}})_{\max} = 1.7695e{+}31 + 246.9892 \simeq \infty \end{cases}$$

The results show that if we use the force to the load cart there exist at least one direction in which the system would be uncontrollable or at least hardly controllable, while if we select the first input (force to the cart with pendulum), even in the worst case the value of control energy does not exceed $1.0913e{+}03$, and the system remains controllable.

□

From this simple example, it becomes apparent that the proposed worst case method (WDOC) is a more meaningful controllability metric rather than the steady state energy measure (EDOC), which is dependent to the initial and final conditions and may produce opposite results according to these conditions.

In summary, if we know the initial and final condition of a system we can easily use the first DOC measure to precisely determine the degree of controllability of the system, however in the case that we are considering the system with various initial and final conditions, the second DOC measure seems more reliable.

5.5. Degree of Controllability

In this section, a novel measure of controllability is proposed, which amounts to generalization of Kalman's rank test to quantify the degree of controllability.

In fact, the proposed (WDOC) approach, in the previous section, determines the distance to uncontrollability, which was firstly defined by (Paige, 1981):

Definition (5.3): The distance to uncontrollability is the spectral norm distance of the pair (A, B) from the set of all uncontrollable pairs:

$$d(A,B) = \min_{s\in\mathbb{C}} \left\{ \left\| [E,F] \right\| : (A+E, B+F) Uncontrollable \right\} \tag{5.43}$$

where $\|.\|$ denotes the spectral norm.

□

This definition has its root in (Hautus, 1970), (Hautus, 1969) characterization of controllability (5.44) and confirms the definition (5.1) involving the degree of controllability, i.e.

$$rank\left(\left[A - sI, B\right]\right) = n, \forall s \in \mathbb{C}, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m} \tag{5.44}$$

According to (Eising, 1984), (Eising, 1984) $d(A,B)$ can be equivalently expressed as:

$$d(A,B) = \min_{s\in\mathbb{C}} \sigma_{\mathrm{n}}\left(\left[A - sI, B\right]\right), A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m} \tag{5.45}$$

where $\sigma_{\min}$ denotes n[th] singular value of augmented matrix $\left[A - sI, B\right]$.

There is a considerable number of research papers which quantify the distance to uncontrollability based on the above definition (Wicks & DeCarlo, 1991) , (Boley & Lu, 1986), (Eising, 1984) (Gahinet & Laub, 1992) (Tarokh, 1992).Special purpose optimization methods have been developed for calculating this distance, however the current methods usually face difficulties due to the presence of numerous local minimum points, so that the algorithm method search for all of them in order to guarantee the optimal solution (Byers, 1989). Furthermore, a good starting point is typically needed to identify the global optimum (Boley, 1987), (Byers, 1989). In addition, some of the existing methods may fail to detect nearly uncontrollable systems (Paige, 1981) (Hu & Davison, 2001).
Next, we propose a new approach based on Kalman's controllability criteria (Ogata, 1997) to solve distance to uncontrollability problem. This new method removes many drawbacks of the current heuristic/numerical methods, and produces more accurate results. Then we will use this novel method to check the validity and accuracy of the worst-case DOC measure (WDOC), which was introduced above.

Lemma (5.2): (Kalman, 1963) A necessary and sufficient condition for (A,B) to be controllable is:

$$rank(C) = rank\left(\left[\begin{array}{cccc} B & AB & A^2B & \cdots & A^{n-1}B \end{array}\right]\right) = n, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m} \tag{5.46}$$

C is called Kalman's controllability matrix or Kalman block matrix (of size $n \times nm$).

□

For the proof see (Kalman, et al., 1963) , (Ogata, 1997). The above Lemma is a useful and simple test but less effort has been spent to generalize it from classical binary rank test to a test which indicates the distance to uncontrollability. In fact, the characterization $d(A,B)$ based on the PBH test has received more attention in literature (Eising, 1984), (Kenney & Laub, 1988). In Theorem (5.4), we propose a new way of calculating distance to uncontrollability based on Kalman's rank condition.

Theorem (5.4): Consider the LTI system described by (5.38). The distance to controllability is:

$$d(A,B)=\min_{s\in\mathbb{C}}\sigma_n(V[I_n \quad \Lambda-sI_n \quad (\Lambda-sI_n)^2 \quad \cdots \quad (\Lambda-sI_n)^{n-1}](I_n\otimes V^{-1}B)) \quad (5.47)$$

where $A=V\Lambda V^{-1}$ is the Eigen-decomposition of system matrix A, and σ_n denotes n[th] singular value of modified Kalman block matrix $\begin{bmatrix} B & (A-sI_n)B & (A-sI_n)^2B & \cdots & (A-sI_n)^{n-1}B \end{bmatrix}$.

□

Proof:

Consider the controllable pair (A,B). According to the definition (5.3) we are looking for the minimum-norm perturbation such that $(A+\Delta A, B+\Delta B)$ is uncontrollable, so according to Kalman's rank condition we have:

$$rank\left(\begin{bmatrix} B+\Delta B & (A+\Delta A)(B+\Delta B) & (A+\Delta A)^2(B+\Delta B) & \cdots & (A+\Delta A)^{n-1}(B+\Delta B) \end{bmatrix}\right)<n \quad (5.48)$$

Additionally, based on (Kienitz, 2012), (Eising, 1984) we know that for uncontrollable pair $(A+\Delta A, B+\Delta B)$ there exists a left-eigenvector/eigenvalue pair (V_i,λ_i) of $(A+\Delta A)$ such that:

$$V_i(A+\Delta A)=\lambda_i V_i \;,\; V_iB=0 \quad (5.49)$$

Furthermore, it is easy to see that the eigenvalues of $A-sI_n$ for any scalar s are those of A subtracted by s, while the eigenvectors remain those of A. Further controllability of pairs (A,B) and $(A-sI_n,B)$ is equivalent, since:

$$V_iA=\lambda_iV_i, V_iB\neq 0 \Rightarrow V_i(A+sI_n)=(\lambda_i-s)V_i, V_iB\neq 0 \quad (5.50)$$

Kalman's controllability matrix for the second pair gives:

$$\begin{bmatrix} B & (A-sI_n)B & (A-sI_n)^2B & \cdots & (A-sI_n)^{n-1}B \end{bmatrix} \quad (5.51)$$

Now considering the Eigen-decomposition of A, i.e. $A=V\Lambda V^{-1}$ (in the case of repeated eigenvalues the singular decomposition is used), we get:

$$\begin{aligned}
&[B \quad V(\Lambda-sI_n)V^{-1}B \quad V(\Lambda-sI_n)^2V^{-1}B \quad \cdots \quad V(\Lambda-sI_n)^{n-1}V^{-1}B]\\
&\Rightarrow [VV^{-1} \quad V(\Lambda-sI_n)V^{-1} \quad V(\Lambda-sI_n)^2V^{-1} \quad \cdots \quad V(\Lambda-sI_n)^{n-1}V^{-1}](I_n\otimes B)\\
&\Rightarrow V[V^{-1} \quad (\Lambda-sI_n)V^{-1} \quad (\Lambda-sI_n)^2V^{-1} \quad \cdots \quad (\Lambda-sI_n)^{n-1}V^{-1}](I_n\otimes B) \quad (5.52)\\
&\Rightarrow (V[I_n \quad \Lambda-sI_n \quad (\Lambda-sI_n)^2 \quad \cdots \quad (\Lambda-sI_n)^{n-1}](I_n\otimes V^{-1})(I_n\otimes B))\\
&\Rightarrow (V[I_n \quad \Lambda-sI_n \quad (\Lambda-sI_n)^2 \quad \cdots \quad (\Lambda-sI_n)^{n-1}](I_n\otimes V^{-1}B))
\end{aligned}$$

Then according to (5.50) and using definition (5.3), calculating the distance to uncontrollability reduces to finding the minimum of n[th] singular values of (5.52) over $s\in\mathbb{C}$ and the proof is completed.

□

Now we illustrate the method introduced above with some numerical examples, each representing a time-invariant linear system of the form (5.38). The examples show the comparable accuracy and validity of our method.
The first example is taken from (Boley & Lu, 1986):

Example (5.6): The LTI system is defined by:

$$A=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, B=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

It is clear by Kalman's criterion that this system is controllable, and by applying the Theorem above we obtain $d(A,B)=0.6616$.

□

In the second example, we start with an uncontrollable system which is studied in various references (Paige, 1981), (Boley, 1987), etc. This example shows the power of our method in detecting the systems very close to uncontrollability where many existing methods may not indicate that fact (in the most cases, our approach is comparable to Eising's method, and in many cases, it produces better results).
For this particular example with repeated eigenvalues, the best current distance to uncontrollability by staircase algorithm is 0.00087689843786 (Boley, 1987). The found solution by Eising's method is 1.1308e-22.

Example (5.7): Consider the pair:

$$A=\begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, B=\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The distance to uncontrollability achieved by the above Theorem is exactly zero, i.e. DOC=0, which is much more accurate rather than the solutions found by other approaches.

□

Now consider the following example, which is borrowed from (Khare, et al., 2012):

Example (5.8): Consider the LTI system:

$$\dot{x}=\begin{bmatrix} -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} x+\begin{bmatrix} 0 \\ t \\ 1 \end{bmatrix} u, t \in \mathbb{R}$$

In this example, the distance to uncontrollability changes as the parameter t varies. Table (5.1) demonstrates the real controllability radius obtained by Theorem (5.4) and compares our method with the results achieved by Eising's method for five different values of t, clearly as t decreases the real controllability radius decreases too (as would be expected since the system is uncontrollable with $t = 0$). The comparison of the results in Table (5.1) confirms the accuracy of our approach.

	t=10	t=2	t=1	t=0.1	t=0
New Theorem (5.4)	0.2192	0.7181	0.3779	0.0388	0
Eising's Method	0.2165	0.7180	0.6063	0.645	3.1659e-05

Table (5.1): Results of Example (5.8)

□

Now let us investigate if the Theorem (5.4) supports the actuator selection which was suggested in Examples (5.4), (5.5) using our WDOC criterion.

Table (5.2) lists the real controllability radius of each actuator selection of AMB system in example (5.4). The results show that the first input produces more controllability. It supports the results of WDOC measure.

	Input 1	Input 2	Input 3	Input 4
$d(A,B)$ **New Method**	0.0175	0.0041	0.0107	0.0037

Table (5.2): Distance to uncontrollability for AMB system in Example (5.4)

The real controllability radius of seesaw-inverted pendulum in example (5.5) is shown in Table (5.3). The results also confirm the input selection by WDOC.

	Input 1	Input 2
$d(A,B)$ **New Method**	0.0147	4.6907e-17

Table (5.3): Distance to uncontrollability for seesaw-inverted pendulum system in Example (5.5)

5.6. Degree of Disturbance Rejection

Consider the linear time-invariant system given in equation (5.36) with an input disturbance d :

$$\begin{cases} \dot{x} = Ax + Bu + B_d d \\ y = Cx + Du \end{cases}, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, B_d \in \mathbb{R}^{n\times p}, C \in \mathbb{R}^{q\times n}, D \in \mathbb{R}^{q\times m} \quad (5.53)$$

Degree of disturbance rejection is defined as minimum energy required to steer initial state x_0 to the final state x_f under the presence of external disturbance d (Lee & Park, 2014).

The state transition equation for above system is given by (Chen, 1984):

$$x_f = e^{At_f} x_0 + \int_0^{t_f} e^{A\tau} B_d d(\tau) d\tau \tag{5.54}$$

(Friedland, 1986) has defined the disturbance sensitivity Gramian as:

$$W_d(0, t_f) = \int_0^{t_f} e^{A\tau} B_d B_d^T e^{A^T \tau} d\tau \tag{5.55}$$

Lemma (5.2): The disturbance energy of the LTI system in equation (5.36) is given by:

$$E^d_{[0,t_f]} = (x_f - e^{At} x_0)^T W_d^{-1}(0, t_f)(x_f - e^{At} x_0) \tag{5.56}$$

where x_0, x_f are initial and terminal states respectively, and $W_d^{-1}(0, t_f)$ is the inverse of disturbance sensitivity Gramian.

□

Proof:

Suppose that $W_d(0, t_f)$ is non-singular for some finite $0 \le t_f$, then the disturbance, which takes the system from some initial state x_0 to the final state x_f is given by:

$$d(t) = B_d^T(t) e^{A(t_f - t)} W_d^{-1}(0, t_f)(x_f - e^{At_f} x_0), 0 \le t \le t_f \tag{5.57}$$

Then the disturbance input energy is defined as:

$$\begin{aligned} E^d_{[0,t_f]} &= \|d(t)\|_2^2 = \int_0^{t_f} d^T(t) d(t) dt \\ &= \int_0^{t_f} d^T(t) B_d^T(t) e^{A^T t} W_d^{-1}(0, t_f)(x_f - e^{At} x_0) dt \\ &= (x_f - e^{At} x_0)^T W_d^{-1}(0, t_f)(x_f - e^{At} x_0) \end{aligned} \tag{5.58}$$

□

In the case that system is transferred from the origin $x_0 = 0$ equation (5.56) could be simplified as:

$$E^d_{[0,t_f]} = x_f{}^T W_d^{-1}(0, t_f) x_f \tag{5.59}$$

From the above Lemma, it can be seen that the inverse of disturbance sensitivity Gramian $W_d^{-1}(0, t_f)$ should be maximized to make the system maximally insensitive to the disturbance. Equivalently, $W_d(0, t_f)$ needs to be minimized to have a minimum effect of disturbance on the system. This could be used to define an appropriate criterion as an indication of the optimal position of the actuators to reject an external disturbance acting on the system.

The degree of disturbance rejection (DODR) is the capability of the system to prevent the disturbance energy to be transferred into the state of the system. This is defined as the trace of the inverse of

controllability Gramian multiplied by the sensitivity Gramian in the steady state (Lee & Park, 2015), (Lee & Park, 2014):

$$\rho = \lim_{t\to\infty} trace(W_c^{-1}(t)W_d(t)) \tag{5.60}$$

Clearly the above measure has a physically meaningful value of input energy and has the potential to be modified to unstable systems by following the procedure proposed in the previous section. W_c, W_d can be obtained from the (CALE) equations:

$$\tilde{A}W_c + W\tilde{A}_c^T = -BB^T, \tilde{A}W_d + W_d\tilde{A}^T = -B_d B_d{}^T \tag{5.61}$$

where $\tilde{A}$ denotes the proper related system matrix, which would be A in the stable case, and $-A$ in anti-stable system. Although this is computationally straightforward we will show that it does not necessarily lead to the correct answer where the growth or decrease in one eigenvalue of $W_c^{-1}W_d$ is large, relative to the increment or decrement of the other eigenvalues. We offer the following illustrative counterexample:

Example (5.9): Consider the linearized model of a two-CSTR process, which is schematically shown in Figure (5.12). A full description of the system and an eight-state model can be found in (Cao & Biss, 1996).

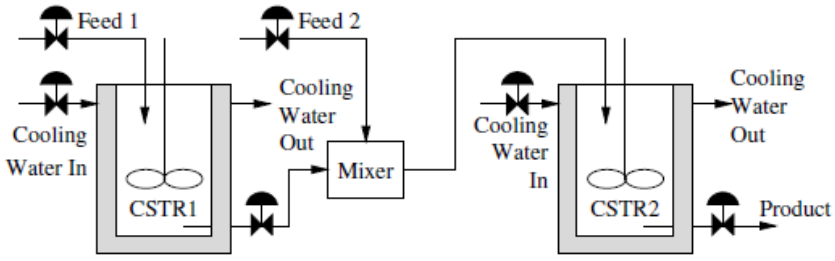


Figure (5.12): Two-CSTR Process

Let us define cooling water temperature fluctuations as the disturbance and assume that we have two options for control inputs:

i. two feed flowrates
ii. two cooling-water flowrates

$$\begin{cases} \dot{x} = Ax + B_i u + B_d d \\ y = Cx + Du \end{cases}$$

$$A=\begin{pmatrix} -17.9751 & -295.8655 & 0 & 0 & 0 & 0 \\ 0.0207 & 0.1889 & 0.0704 & 0 & 0 & 0 \\ 0 & 0.3879 & -0.8000 & 0 & 0 & 0 \\ 0.0977 & 0 & 0 & -18.0088 & -295.8655 & 0 \\ 0 & 0.0617 & 0 & 0.0131 & 0.0433 & 0.0589 \\ 0 & 0 & 0 & 0 & 0.378 & -0.6220 \end{pmatrix}$$

$$B_1=\begin{bmatrix} 17.8996 & -13.7811 \\ -0.0131 & 0.0101 \\ 0 & 0 \\ 17.8636 & 17.8636 \\ -0.0082 & -0.0082 \\ 0 & 0 \end{bmatrix} B_2=\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -0.0294 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -0.0235 \end{bmatrix}, B_d=\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.0137 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.0081 \end{bmatrix}$$

$$C=\begin{bmatrix} 0 & 362.9950 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 362.9950 & 0 \end{bmatrix}$$

If we choose the first input B_1 the degree of disturbance rejection would be obtained as:

$$\rho_1=\lim_{t_f\to\infty} trace(W_c^{-1}W_d)=17.0353$$

While by selecting the second input B_2 we will have:

$$\rho_2=\lim_{t_f\to\infty} trace(W_c^{-1}W_d)=1.0154$$

The two results yield that the control input energy of first input B_1 is larger than that of second input B_2 , which suggest input B_2 is the best selection, while we know that B_d is same in both cases then the degree of controllability could be a good indicator of the disturbance rejection capability of the system and it suggests that input B_1 is a better choice, because the first input causes more controllability, $DOC_1=0.0024, DOC_2=9.6533\text{e-}05$.

□

In following, we introduce a new DODR, based on worst case of steady state input energy. We will show that this new method can be used as a reliable criterion to measure the degree of disturbance rejection of the system, and it can easily be modified for unstable systems.

Definition (5.4): Consider the LTI system in equation (5.53), the degree of disturbance rejection could be defined as:

$$\rho=\frac{trace(W_c(\infty))}{1+trace(W_c(\infty))}+\frac{1}{trace(W_d(\infty))} \tag{5.62}$$

$W_c(\infty)$ denotes the controllability Gramian in the steady state situation for the system without the disturbance, i.e. $\dot{x} = Ax + Bu, d = 0, t_f \to \infty$, and $W_d(\infty)$ demonstrates the sensitivity Gramian of the system without control input when $t_f \to \infty$ i.e. $\dot{x} = Ax + B_d d$, $u = 0$.

□

Consider the continuous time invariant system in equation (5.53), we can write it as:

$$\dot{x} = Ax + \begin{bmatrix} B & B_d \end{bmatrix} \begin{bmatrix} u \\ d \end{bmatrix} \tag{5.63}$$

Then with regard to the definition of controllability Gramian we obtain:

$$W_{sys}(0, t_f) = \int_0^{t_f} e^{A\tau} \begin{bmatrix} B & B_d \end{bmatrix} \begin{bmatrix} B \\ B_d \end{bmatrix} e^{A^T\tau} d\tau \tag{5.64}$$

Thus, we have:

$$\begin{aligned} & trace(W_{sys}(0, t_f)) = trace \int_0^{t_f} e^{A\tau} \begin{bmatrix} B & B_d \end{bmatrix} \begin{bmatrix} B \\ B_d \end{bmatrix} e^{A^T\tau} d\tau \\ & \Rightarrow trace(W_{sys}(0, t_f)) = trace \int_0^{t_f} e^{A\tau} BB^T e^{A^T\tau} d\tau + trace \int_0^{t_f} e^{A\tau} B_d B_d{}^T e^{A^T\tau} d\tau \end{aligned} \tag{5.65}$$

Now assume we define:

$$k = trace(W_{sys}(0, t_f)), a = trace \int_0^{t_f} e^{A\tau} BB^T e^{A^T\tau} d\tau, b = trace \int_0^{t_f} e^{A\tau} B_d B_d{}^T e^{A^T\tau} d\tau$$

Then (5.65) can be written as:

$$k - b = a$$

According to $a \geq \dfrac{a}{1+a}$ we get:

$$k - b \geq \frac{a}{1+a} \tag{5.66}$$

One can add $0 \leq \dfrac{1}{b}$ to both sides of statement (5.66), then we obtain:

$$
\begin{aligned}
&\Rightarrow k-b+\frac{1}{b}\geq\frac{a}{1+a}+\frac{1}{b}\\
&\Rightarrow \frac{kb-b^2+1}{b}\geq\frac{a}{1+a}+\frac{1}{b}\\
&\Rightarrow \frac{ab+b^2-b^2+1}{b}\geq\frac{a}{1+a}+\frac{1}{b}\\
&\Rightarrow \frac{ab+1}{b}\geq\frac{a}{1+a}+\frac{1}{b}\geq 0
\end{aligned}
\tag{5.67}
$$

Hence, with regard to minimum input energy optimization problem we know that the left-hand side of the statement (5.67) should be maximized that is a non-convex optimization, then using a relaxation we can optimize input energy by maximizing the lower bound $\rho=\frac{a}{1+a}+\frac{1}{b}$:

$$
\max(a+\frac{1}{b})\Rightarrow\max(\frac{ab+1}{b})\Rightarrow\max(\frac{a}{1+a}+\frac{1}{b}) \tag{5.68}
$$

$$
\Rightarrow \rho=\frac{trace(W_c(\infty))}{1+trace(W_c(\infty))}+\frac{1}{trace(W_d(\infty))} \tag{5.69}
$$

Thus statement (5.62) is verified.

□

Remark (5.5): In the case that A is stable we can easily use the Lyapunov equations to get DODR defined in (5.4):

$$
\begin{cases}
AW_{sys}(\infty)+W_{sys}(\infty)A^T=-\begin{bmatrix}B & B_d\end{bmatrix}\begin{bmatrix}B\\B_d\end{bmatrix}\\
AW_c(\infty)+W_c(\infty)A^T=-BB^T\\
AW_d(\infty)+W_d(\infty)A^T=-B_dB_d{}^T
\end{cases}
\tag{5.70}
$$

Furthermore, as discussed in previous sections, in the case that A is unstable, i.e. A has some eigenvalues on the right half plane (RHP), we may decompose the system into stable and anti-stable parts, and then we can easily use the Lyapunov equations to find the controllability and sensitivity Gramians of each subsystem separately:

$$
\begin{cases}
A_sW_{sys.s}(\infty)+W_{sys.s}(\infty)A_s^T=-\begin{bmatrix}B_s & B_{ds}\end{bmatrix}\begin{bmatrix}B_s\\B_{ds}\end{bmatrix},\ A_uW_{sys.u}(\infty)+W_{sys.u}(\infty)A_u^T=\begin{bmatrix}B_u & B_{du}\end{bmatrix}\begin{bmatrix}B_u\\B_{du}\end{bmatrix}\\
A_sW_{c.s}(\infty)+W_{c.s}(\infty)A_s^T=-B_sB_s^T \quad , \quad A_uW_{c.u}(\infty)+W_{c.u}(\infty)A_u^T=B_uB_u^T\\
A_sW_{d.s}(\infty)+W_{d.s}(\infty)A_s^T=-B_{ds}B_{ds}^T \quad , \quad A_uW_{d.u}(\infty)+W_{d.u}(\infty)A_u^T=B_{du}B_{du}^T
\end{cases}
\tag{5.71}
$$

□

Therefore, our method is not computationally intensive either. It can be readily solved using the Lyapunov equations.

Now let us test the validity of the proposed DODR measure in example (5.9):

The eigenvalues of A are: $\lambda_A = \begin{cases} -17.6314 \\ -17.7915 \\ -0.1142 \\ -0.1283 \\ -0.8406 \\ -0.6678 \end{cases}$ By choosing the first input B_1 we obtain $\rho = 0.9699 + 444.8508 = 445.8207$ while the second input B_2 yields $\rho = 0.0122 + 444.8508 = 444.8630$ thus the results suggest the CSTR plant with B_1 would be less affected by the disturbance d.

Now consider another example cited from (Johnston & Barton, 1985):

Example (5.10): Consider the linear system with three candidate inputs, u described by the following state space model:

$$\begin{cases} \dot{x} = Ax + Bu + B_d d \\ y = Cx \end{cases}$$

$$A = \begin{bmatrix} -1 & -2 & 3 & 0 & 0 & 1 \\ 0 & -2 & 10 & 0 & 2 & 0 \\ 0 & 0 & -20 & 10 & 0 & -20 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ 4 & 0 & 0 & 0 & -10 & 2 \\ 1 & 1 & 0 & 0 & 3 & -15 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} -4 \\ 1 \\ 5 \\ 1 \\ -4 \\ -2 \end{bmatrix}, B_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, d = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The aim is to select two inputs from the three candidates by considering the disturbance rejection properties of each input. Table (5.4) demonstrates the results of our method in measuring the system's capability in disturbance rejection. In the original work Johnston and Barton claimed that

$\{B_1, B_3\}$ is the best input set. However, later (Cao, et al., 1997) through Input-Disturbance Gain Deviation (IDGD) method in the frequency domain revealed that this selection appears to be wrong and showed that the best performance is expected from selecting the input set $\{B_1, B_2\}$. The results of our method also support this statement, because our proposed DODR is maximized by the selection of input set $\{B_1, B_2\}$, which means that this approach maintains the physical meanings of DODR as the maximum energy requirement for the disturbance.

	$\{B_1, B_2\}$	$\{B_1, B_3\}$	$\{B_2, B_3\}$
DODR New Method	1.3240	1.2624	1.3210

Table (5.4): DODR results of Example (5.10)

□

To evaluate the performance and usefulness of the proposed measure, let us consider another numerical example describing a continues-time linear model of bus suspension system. In this example, we will consider the cases when the disturbance is matched and unmatched. Next, we will identify the variation of the measure according to the variation of a model parameter.

Example (5.11): Consider a 1/4 bus model (one of the four wheels) to simplify the problem to a one-dimensional spring-damper system:

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k}{m_1} & -\frac{b}{m_1} & \frac{k}{m_1} & \frac{b}{m_1} \\ 0 & 0 & 0 & 1 \\ \frac{k}{m_2} & \frac{b}{m_2} & -\frac{2k}{m_2} & -\frac{b}{m_2} \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} u + B_d d \\ y = \begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix} x \end{cases}$$

Assume body mass $m1 = 2500\, kg$, suspension mass $m2 = 320\, kg$, spring rate $k = 10000\ N/m$ damping constant $b = 140000\, Ns/m$ and the deformation of the tire as the disturbance. The system is controllable and asymptotically stable.

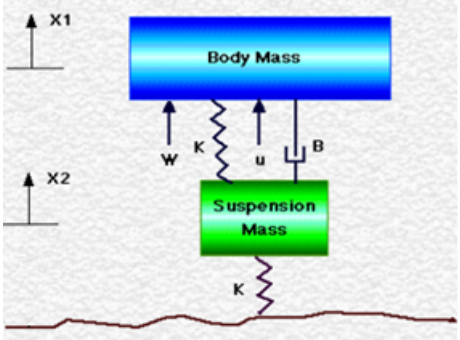


Figure (5.13): Model of bus suspension system (matched disturbance)

We first study the case when the matching condition is satisfied; that is, the disturbance is in the range of input matrix B (Figure (5.13)). For the matched disturbance, $B_d = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^T$ then we get $\rho = 0.9474 + 0.0555 = 1.0029$.

Next, we will calculate the measure for the case of the unmatched disturbance depicted in Figure (5.14).

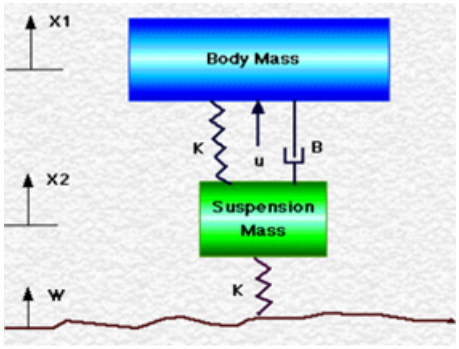


Figure (5.14): Model of bus suspension system (unmatched disturbance)

In this case, the disturbance matrix is given by $B_d = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$, then the proposed DODR is $\rho = 0.9474 + 3.3857 = 4.3331$.

Thus, it is found that the capability of disturbance rejection of the system would be degraded when a matched disturbance exists, i.e. $B_d = B$, since the matched disturbance can directly affect the system, then the control input energy is equal to the disturbance energy, while in unmatched case the disturbance may require larger energy to affect the system.

Now assume that the spring constant k changes from zero to infinity while the matched disturbance $B_d = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^T$ is applied to the system. The results from changing the spring constant k are shown in Table (5.5). The results reveal that as the spring constant becomes larger our proposed DODR measure increases. That means that the degree of disturbance rejection improves. Those results are expected from the physical understanding of the system. The small spring constant means that the masses are almost not linked to each other. Thus, the disturbance may highly affect the system. As the spring constant k increases the two masses m_1 and m_2 are connected tightly to each other. Accordingly, the DODR measure increases.

	K=0	K=0.01	K=0.1	K=1	K=10	
DODR New Method	1	1+2.5329e-13	1+2.5328e-11	1+2.532e-9	32.5240e-7	
	K=100	**K=10^3**	**K=10^4**	**K=10^5**	**K=10^6**	**K=inf**
DODR New Method	1.00002446	1.0019	1.0029	1.2699	5.4894	21.3538

Table (5.5): DODR results of Example (5.11)

5.7. Summary

In this chapter, we defined the minimum energy of unstable systems in the steady state form. Two different energy-based measures were introduced to check the degree of controllability for both stable and unstable systems. According to Kalman's controllability rank condition a new quantitative measure of DOC is also presented, which verified the validity of the energy-based controllability metrics.

In the chapter, it was explained how these measures can be employed to determine optimal actuator locations for good degree of controllability. A measure for the degree of disturbance rejection was also proposed. The approach depends on computation of the controllability Gramian and the disturbance sensitivity Gramian.

In the next chapter, we consider the stable LTI systems and propose an approach to find the value of the minimum input energy using the coefficients of the characteristic polynomial of the system. This enables to determine the trace and some upper-bounds for the maximum eigenvalue of the controllability Gramian.

Chapter 6: Controllability Gramian and Energy Calculations of the Stable LTI System in Canonical Forms

6.1. Introduction

In this chapter, we consider the case where the LTI system is stable and it has certain canonical form structure. We derive a simple structure for the controllability Gramian based on the coefficients of the characteristic polynomial of the system.

The aim of this chapter is to present the interesting links between the value of minimum input energy and the place of the eigenvalues of a stable system. Moreover, an expression for the trace of the controllability Gramian is derived as a simple function of the coefficients of the characteristic polynomial using the fact that the controllability Gramian of a stable LTI system is the solution of a Lyapunov equation.

In section 6.2 we discuss the case where the system is defined in the controller companion form. An upper bound for the maximum eigenvalue of the controllability Gramian is proposed based on the Routh Hurwitz table. The inverse controllability Gramian and the minimum input energy using the coefficients of the characteristic polynomial of a stable system in the controller companion form is discussed in section 6.3. Finally, in section 6.4 an approach is proposed to find the controllability Gramian of a diagonal LTI system.

6.2. Controllability Gramian Computation in Controller Canonical Form

The controllability Gramian of stable linear continuous time-invariant systems:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times 1}, C \in \mathbb{R}^{1\times n}, D \in \mathbb{R}^{1\times 1} \tag{6.1}$$

can be described by the solution of the following Lyapunov equation:

$$AW_c + W_c A^T = -BB^T \tag{6.2}$$

A number of methods for solving the Lyapunov equation have been presented in the literature (Barnett & Storey, 1970), (Barnett & Storey, 1968), (Wu, et al., 2006), (Zhou & Duan, 2005), (Hauksdóttir, et al., 2008), (Hauksdóttir & Sigurðsson, 2009). In this section through the Kronecker product we present a simple method for finding the controllability Gramian directly based on the coefficients of the characteristic polynomial.

Lemma (6.1): (Barnett & Storey, 1970) The Lyapunov equation (6.2) can be written as:

$$(A \oplus A)vec(W_c) = -vec(Q) \tag{6.3}$$

where $Q = BB^T$ and $vec(.)$ is the vector obtained by stacking columns of matrix over each other. $(A \oplus A)$ is the Kronecker sum of matrix A. (A is assumed to be in controller companion form)

(Note: The Kronecker sum of two matrices is defined as: $(A_{n\times n} \oplus B_{m\times m}) = (I_m \otimes A) + (B \otimes I_n)$ (Laub, 2005))

□

The proof can be found in (Barnett & Storey, 1968), (Laub, 2005).

Lemma (6.2): Consider the LTI system (6.1). There is a non-singular similarity matrix T such that:

$$\begin{cases} T^{-1}AT = A_c \\ T^{-1}b = b_c \end{cases} \Rightarrow \begin{cases} AT = TA_c \\ b = Tb_c \end{cases} \tag{6.4}$$

where:

$$A_c = \begin{bmatrix} 0_{(n-1)\times 1} & & I_{(n-1)\times(n-1)} \\ -\beta_0 & -\beta_1 \quad \cdots \quad & -\beta_{n-2} \quad -\beta_{n-1} \end{bmatrix}, b_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Matrix T can be described as:

$$T = \begin{bmatrix} b & Ab & A^2b & \cdots & A^{n-1}b \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_{n-1} & 1 \\ \beta_2 & \beta_3 & \cdots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{6.5}$$

where $\beta_i, i = 1, \ldots, n-1$ are the coefficients of the characteristic polynomial of the system, i.e.

$$p(s) = s^n + \beta_{n-1}s^{n-1} + \cdots + \beta_0 \tag{6.6}$$

□

Proof: (Kailath, 1980)

Let $T = \begin{bmatrix} t_1 & t_2 & t_3 & \ldots & t_n \end{bmatrix}$ and define the similarity transformation:

$$A \to T^{-1}AT = A_c$$

Consider the characteristic polynomial of the system:

$$\left|sI - A\right| = s^n + \beta_{n-1}s^{n-1} + \cdots + \beta_0 \tag{6.7}$$

According to (6.5) we obtain:

$$AT = TA_c \Rightarrow \begin{bmatrix} At_1 & At_2 & At_3 & \ldots & At_n \end{bmatrix} = \begin{bmatrix} -\beta_0 t_n & t_1 - \beta_1 t_n & t_2 - \beta_2 t_n & \ldots & t_{n-1} - \beta_{n-1} t_n \end{bmatrix} \tag{6.8}$$

and:

$$b = Tb_c \Rightarrow b = T\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = t_n \tag{6.9}$$

Then (6.5) shows that T is the product of the controllability matrix of the system and the Hankel matrix of the characteristic polynomial's coefficients. In vector form:

$$\begin{cases} b = t_n \\ Ab + \beta_{n-1}b = t_{n-1} \\ At_{n-1} + \beta_{n-2}b = t_{n-2} \Rightarrow T = \begin{bmatrix} b & Ab & A^2b & \cdots & A^{n-1}b \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_{n-1} & 1 \\ \beta_2 & \beta_3 & \cdots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \\ \vdots \\ At_2 + \beta_1 b = t_1 \end{cases} \tag{6.10}$$

which concludes the proof.

□

Using the similarity transformation T, Lemma (6.1) yields:

Corollary (6.1): Consider the stable LTI system (6.1). Then the Lyapunov equation (6.2) can be written as:

$$(A_c \oplus A_c)(T \otimes T)^{-1} vec(W_c) = -vec(Q_c) \tag{6.11}$$

where A_c and Q_c are defined in (6.4).

Considering the equation above, the controllability Gramian can be written as:

$$W_c = -vec^{-1}\left((T \otimes T)\underbrace{(A_c \oplus A_c)^{-1} vec(Q_c)}_{vec(W_c^c)} \right) \tag{6.12}$$

or equivalently:

$$W_c = TW_c^c T^T$$

Here W_c^c defines the controllability Gramian in the controller companion form, and vec^{-1} is the inverse vec operator reshaping the vector into a n-by-n matrix.

□

Note: In the Corollary above system is stable, i.e. A_c has no eigenvalues in the open right-half plain (ORHP) and then $\lambda_i + \lambda_j \neq 0, \forall i, j = 1, \ldots, n$ where λ_i defines the i[th] eigenvalue of A_c. Hence the inverse of the Kronecker sum $(A_c \oplus A_c)$ is well-defined.

Proposition (6.1): Let $A_c \in \mathbb{R}^{n \times n}$ be in controller companion form and assume that $\lambda_i + \lambda_j \neq 0, \forall i, j = 1, \ldots, n$. Then:

$$(A_c \oplus A_c)^{-1}{}_{n^2 \times n^2} = \begin{bmatrix} \bar{O}_{11} & \bar{O}_{12} & \cdots & \bar{O}_{1n} \\ \bar{O}_{21} & \bar{O}_{22} & \cdots & \bar{O}_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{O}_{n1} & \bar{O}_{n2} & \cdots & \bar{O}_{nn} \end{bmatrix}$$

where the blocks $\bar{O}_{ij}$ are defined recursively as:

$$\begin{cases} \bar{O}_{11} = (\sum_{i=0}^{n} (-1)^{i+1} \beta_i A_c^{\,i})^{-1} \sum_{i=1}^{n} (-1)^{i-1} \beta_i A_c^{\,i-1}, \begin{cases} \bar{O}_{21} = I - A_c \bar{O}_{11} \rightarrow \textit{first} - \textit{underdiagblock} \\ \bar{O}_{i1} = (-1)^i A_c^{\,i-2} \bar{O}_{21}, i = 3, \ldots, n \end{cases} \\ \bar{O}_{12} = (\sum_{i=0}^{n} (-1)^{i+1} \beta_i A_c^{\,i})^{-1} \sum_{i=2}^{n} (-1)^{i-2} \beta_i A_c^{\,i-2}, \begin{cases} \bar{O}_{32} = I - A_c \bar{O}_{22} \rightarrow \textit{first} - \textit{underdiagblock} \\ \bar{O}_{i2} = (-A_c)^{i-1} \bar{O}_{12}, i = 2 \rightarrow \textit{diagblock} \\ \bar{O}_{i2} = (-A_c)^{i-3} \bar{O}_{32}, i = 4, \ldots, n \end{cases} \\ \bar{O}_{13} = (\sum_{i=0}^{n} (-1)^{i+1} \beta_i A_c^{\,i})^{-1} \sum_{i=3}^{n} (-1)^{i-3} \beta_i A_c^{\,i-3}, \begin{cases} \bar{O}_{43} = I - A_c \bar{O}_{33} \rightarrow \textit{first} - \textit{underdiagblock} \\ \bar{O}_{i3} = (-A_c)^{i-1} \bar{O}_{13}, i = 2,3 \rightarrow \textit{upperdiag} - \textit{and} - \textit{diagblocks} \\ \bar{O}_{i3} = (-A_c)^{i-4} \bar{O}_{43}, i = 5, \ldots, n \end{cases} \\ \vdots \\ \bar{O}_{1n} = (\sum_{i=0}^{n} (-1)^{i+1} \beta_i A_c^{\,i})^{-1}, \{ \bar{O}_{in} = (-A_c)^{i-1} \bar{O}_{1n}, i = 2, \ldots, n \end{cases}$$

(6.13)

□

Proof:

From the Kronecker sum definition (Laub, 2005) we have:

$$(A_c \oplus A_c) = \big((A_c \otimes I) + (I \otimes A_c)\big) = \begin{bmatrix} A_c & I & 0_{n\times n} & \cdots & \cdots & 0_{n\times n} \\ 0_{n\times n} & A_c & I & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & & \ddots & & & \vdots \\ 0_{n\times n} & 0_{n\times n} & \cdots & A_c & & I \\ -\beta_0 I & -\beta_1 I & \cdots & -\beta_{n-2} I & & -\beta_{n-1} I + A_c \end{bmatrix} \quad (6.14)$$

Then, we define the inverse of this matrix as:

$$(A_c \oplus A_c)^{-1}{}_{n^2\times n^2} = \begin{bmatrix} \bar{O}_{11} & \bar{O}_{12} & \cdots & \bar{O}_{1n} \\ \bar{O}_{21} & \bar{O}_{22} & \cdots & \bar{O}_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{O}_{n1} & \bar{O}_{n2} & \cdots & \bar{O}_{nn} \end{bmatrix} \quad (6.15)$$

Thus:

$$\underbrace{\begin{bmatrix} A_c & I & 0_{n\times n} & \cdots & \cdots & 0_{n\times n} \\ 0_{n\times n} & A_c & I & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & & \ddots & & & \vdots \\ 0_{n\times n} & 0_{n\times n} & \cdots & A_c & & I \\ -\beta_0 I & -\beta_1 I & \cdots & -\beta_{n-2} I & & -\beta_{n-1} I + A_c \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} \bar{O}_{11} & \bar{O}_{12} & \cdots & \bar{O}_{1n} \\ \bar{O}_{21} & \bar{O}_{22} & \cdots & \bar{O}_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{O}_{n1} & \bar{O}_{n2} & \cdots & \bar{O}_{nn} \end{bmatrix}}_{N} = I_{n^2\times n^2} \quad (6.16)$$

Appling Gaussian elimination with partial pivoting to equation (6.16), we can readily obtain $(A_c \oplus A_c)^{-1}$ as it is described in equation (6.13).

By multiplying the first row of M by the columns of N we obtain:

$$\begin{cases} A_c\bar{O}_{11} + \bar{O}_{21} = I_n \Rightarrow \bar{O}_{21} = I - A_c\bar{O}_{11} \\ A_c\bar{O}_{12} + \bar{O}_{22} = 0_n \Rightarrow \bar{O}_{22} = -A_c\bar{O}_{12} \\ \vdots \\ A_c\bar{O}_{1n} + \bar{O}_{2n} = 0 \Rightarrow \bar{O}_{2n} = -A_c\bar{O}_{1n} \end{cases} \quad (6.17)$$

Take the second row of M , and multiply by the columns of N :

$$\begin{cases} A_c\bar{O}_{21} + \bar{O}_{31} = 0_n \Rightarrow \bar{O}_{31} = -A_c\bar{O}_{21} \\ A_c\bar{O}_{22} + \bar{O}_{32} = I_n \Rightarrow \bar{O}_{32} = I - A_c\bar{O}_{22} \\ \vdots \\ A_c\bar{O}_{2n} + \bar{O}_{3n} = 0 \Rightarrow \bar{O}_{2n} = -A_c\bar{O}_{3n} \end{cases}$$

By multiplying, in the third row we get:

$$\begin{cases} A_c\bar{O}_{31}+\bar{O}_{41}=0_n \Rightarrow \bar{O}_{41}=-A_c\bar{O}_{31}=A_c{}^2\bar{O}_{21} \\ A_c\bar{O}_{32}+\bar{O}_{42}=0 \Rightarrow \bar{O}_{42}=-A_c\bar{O}_{32} \\ A_c\bar{O}_{33}+\bar{O}_{43}=I_n \Rightarrow \bar{O}_{43}=I-A_c\bar{O}_{33} \\ \vdots \\ A_c\bar{O}_{3n}+\bar{O}_{4n}=0 \Rightarrow \bar{O}_{4n}=-A_c\bar{O}_{3n}=A_c{}^2\bar{O}_{2n} \end{cases}$$

Similarly, multiplication of the rest of the rows gives:

$$\vdots$$

$$\begin{cases} A_c\bar{O}_{(n-1)1}+\bar{O}_{n1}=0_n \Rightarrow \bar{O}_{n1}=-A_c\bar{O}_{(n-1)1} \\ A_c\bar{O}_{(n-1)2}+\bar{O}_{n2}=0 \Rightarrow \bar{O}_{n2}=-A_c\bar{O}_{(n-1)2} \\ \vdots \\ A_c\bar{O}_{(n-1)n}+\bar{O}_{nn}=I_n \Rightarrow \bar{O}_{nn}=I-A_c\bar{O}_{(n-1)n} \end{cases}$$

$$\begin{cases} -\beta_0\bar{O}_{11}-\beta_1\bar{O}_{21}-\cdots-\beta_{n-1}\bar{O}_{n1}+A_c\bar{O}_{n1}=0 \Rightarrow (\sum_{i=0}^{n}(-1)^{i+1}\beta_i A_c^i)\bar{O}_{11}=\sum_{i=1}^{n}(-1)^{i-1}\beta_i A_c^{i-1}) \\ -\beta_0\bar{O}_{12}-\beta_1\bar{O}_{22}-\cdots-\beta_{n-1}\bar{O}_{n2}+A_c\bar{O}_{n2}=0 \Rightarrow (\sum_{i=0}^{n}(-1)^{i+1}\beta_i A_c^i)\bar{O}_{12}=\sum_{i=2}^{n}(-1)^{i-2}\beta_i A_c^{i-2}) \\ \vdots \\ -\beta_0\bar{O}_{1n}-\beta_1\bar{O}_{2n}-\cdots-\beta_{n-1}\bar{O}_{nn}+A_c\bar{O}_{nn}=I \Rightarrow (\sum_{i=0}^{n}(-1)^{i+1}\beta_i A_c^i)\bar{O}_{1n}=I \end{cases}$$

and the proof is complete.

□

Remark (6.1): The Proposition above, allows to calculate the inverse of n^2-dimensional matrix $(A_c \oplus A_c)^{-1} \in \mathbb{R}^{n^2\times n^2}$ by inverting the n-dimensional matrix $(\sum_{i=0}^{n}(-1)^{i+1}\beta_i A_c{}^i)$ where $|sI-A_c|=\beta_n s^n+\beta_{n-1}s^{n-1}+\cdots+\beta_0, \beta_n=1.$

□

Corollary (6.2): Consider the stable LTI system (6.4) in controller companion form. The controllability Gramian $W_c^c(0,\infty)$ can be readily derived as the matrix produced by the inverse vectorization of the last column of $(A_c \oplus A_c)^{-1}$, i.e.

$$W_c^c(0,\infty)=-vec^{-1}\left(\left[\bar{O}_{1n}(:,n) \quad \bar{O}_{2n}(:,n) \quad \cdots \quad \bar{O}_{nn}(:,n)\right]^T\right) \tag{6.18}$$

where:

$$(A_c \oplus A_c)^{-1}{}_{n^2\times n^2} = \begin{bmatrix} \bar{O}_{11} & \bar{O}_{12} & \cdots & \bar{O}_{1n} \\ \bar{O}_{21} & \bar{O}_{22} & \cdots & \bar{O}_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \bar{O}_{n1} & \bar{O}_{n2} & \cdots & \bar{O}_{nn} \end{bmatrix} \tag{6.19}$$

and $\bar{O}_{in}(:,n), i = 1,\ldots,n$ describes the last column of the block matrix $\bar{O}_{in}$.

□

Remark (6.2): Corollary (6.2) shows that the full structure of $(A_c \oplus A_c)^{-1}$ is not needed for the calculation of the controllability Gramian $W_c^c(0,\infty)$.

□

Remark (6.3): Consider the controllability Gramian of the stable LTI system (6.1) in controller companion form. Then the trace of the controllability Gramian is equal to:

$$trace(W_c^c(0,\infty)) = -\sum_{i=1}^{n} o_{in}(\bar{O}_{in}) \tag{6.20}$$

where $(A_c \oplus A_c)^{-1} > 0$ is defined in (6.19) and $o_{in}(\bar{O}_{in})$ denotes the element in ith row and last column of the block matrix $\bar{O}_{in}$, for $i = 1,\ldots,n$.

□

Remark (6.4): An upper bound on the maximum eigenvalue of the controllability Gramian of the stable LTI system (6.4) in controller form is equal to:

$$\begin{aligned} &\lambda_{\max}(W_c^c(0,\infty)) \le \sqrt{trace(W_c^c(0,\infty))^2} \\ &= \sqrt{\bar{e}_n{}^T \bar{O}_{1n}^T (I + A_c^T A_c + (A_c^2)^T A_c^2 + (A_c^3)^T A_c^3 + \cdots + (A_c^{(n-1)})^T A_c^{n-1}) \bar{O}_{1n} \bar{e}_n} \\ &= \sqrt{\bar{e}_n{}^T \sum_{i=1}^{n} \bar{O}_{in}^T \bar{O}_{in} \bar{e}_n} \\ &\Downarrow \\ &\lambda_{\max}(W_c^c(0,\infty)) \le \sqrt{\sum_{i=1}^{n} (\bar{O}_{in}(:,n))^T \bar{O}_{in}(:,n)} \end{aligned} \tag{6.21}$$

where $\bar{e}_n = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}_{1\times n}^T$. $\bar{O}_{in}(:,n)$ represents the last column of the block matrix $\bar{O}_{in}$ and $\bar{O}_{in}$ denotes the block matrix in the ith row and nth column of the matrix $(A_c \oplus A_c)^{-1}, i = 1,\ldots,n$ in (6.19)

□

Example (6.1): Consider the stable LTI system below:

$$A=\begin{bmatrix}-1.6673 & -0.0868 & -0.0345\\ 0.0924 & -1.2321 & -1.6134\\ -0.0139 & 1.6137 & -1.2334\end{bmatrix}, B=\begin{bmatrix}1.0306\\ 0.3275\\ 0.6521\end{bmatrix}$$

Using the similarity matrix:

$$T=\begin{bmatrix}4.2601 & 2.4902 & 1.0306\\ -0.9425 & -0.0069 & 0.3275\\ 2.3626 & 2.4050 & 0.6521\end{bmatrix}$$

The system can be put in controller companion form:

$$A_c=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -6.8913 & -8.2417 & -4.1328\end{bmatrix}, B_c=\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}$$

Then according to Corollary (6.1) we obtain the controllability Gramian $W_c(0,\infty)$ as:

$$W_c(0,\infty)=vec^{-1}(-(T\otimes T)(A_c\oplus A_c)^{-1}vec(Q_c))$$

Then, using Proposition (6.1) we have:

$$(A_c\oplus A_c)^{-1}=\begin{bmatrix}\bar{O}_{11} & \bar{O}_{12} & \bar{O}_{13}\\ \bar{O}_{21} & \bar{O}_{22} & \bar{O}_{23}\\ \bar{O}_{31} & \bar{O}_{32} & \bar{O}_{33}\end{bmatrix}, vec(Q_c)_{9\times1}=\begin{bmatrix}0 & 0 & \cdots & 1\end{bmatrix}^T$$

where:

$$\bar{O}_{11}=\left(A_c^2-\beta_2A_c+\beta_1I\right)\left(A_c^3-\beta_2A_c^2+\beta_1A_c-\beta_0I\right)^{-1}\begin{cases}\bar{O}_{21}=I-A_c\bar{O}_{11}\\ \bar{O}_{31}=-A_c\bar{O}_{21}\end{cases}$$

in which:

$$\begin{cases}\bar{O}_{11}=\begin{bmatrix}-0.9130 & -0.3765 & -0.0727\\ 0.5000 & -0.3146 & -0.0761\\ 0.5236 & 1.1267 & 0\end{bmatrix}\\ \bar{O}_{21}=\begin{bmatrix}0.5000 & 0.3146 & 0.0761\\ -0.5236 & -0.1267 & 0.0000\\ 0.0000 & -0.5236 & -0.1267\end{bmatrix}\\ \bar{O}_{31}=\begin{bmatrix}0.5236 & 0.1267 & 0.0000\\ -0.0000 & 0.5236 & 0.1267\\ -0.8716 & -1.0432 & -0.0000\end{bmatrix}\end{cases}$$

Also:

$$\bar{O}_{12} = \left(-A_c + \beta_2 I\right)\left(A_c^3 - \beta_2 A_c^2 + \beta_1 A_c - \beta_0 I\right)^{-1} \begin{cases} \bar{O}_{22} = -A_c \bar{O}_{12} \\ \bar{O}_{32} = I - A_c \bar{O}_{22} \end{cases}$$

$$\begin{cases} \bar{O}_{12} = \begin{bmatrix} -0.3765 & -0.2074 & -0.0457 \\ 0.3146 & 0.0000 & -0.0184 \\ 0.1267 & 0.4662 & 0.0761 \end{bmatrix} \\ \bar{O}_{22} = \begin{bmatrix} -0.3146 & -0.0000 & 0.0184 \\ -0.1267 & -0.4662 & -0.0761 \\ 0.5236 & 0.5000 & -0.1516 \end{bmatrix} \\ \bar{O}_{32} = \begin{bmatrix} 1.1267 & 0.4662 & 0.0761 \\ -0.5236 & 0.5000 & 0.1516 \\ -1.0432 & -1.7721 & -0.1267 \end{bmatrix} \end{cases}$$

and finally:

$$\bar{O}_{13} = \left(A_c^3 - \beta_2 A_c^2 + \beta_1 A_c - \beta_0 I\right)^{-1} \begin{cases} \bar{O}_{23} = -A_c \bar{O}_{13} \\ \bar{O}_{33} = A_c^2 \bar{O}_{13} \end{cases}$$

$$\begin{cases} \bar{O}_{13} = \begin{bmatrix} -0.0727 & -0.0457 & -0.0111 \\ 0.0761 & 0.0184 & -0.0000 \\ 0.0000 & 0.0761 & 0.0184 \end{bmatrix} \\ \bar{O}_{23} = \begin{bmatrix} -0.0761 & -0.0184 & 0.0000 \\ -0.0000 & -0.0761 & -0.0184 \\ 0.1267 & 0.1516 & -0.0000 \end{bmatrix} \\ \bar{O}_{33} = \begin{bmatrix} 0.0000 & 0.0761 & 0.0184 \\ -0.1267 & -0.1516 & 0.0000 \\ -0.0000 & -0.1267 & -0.1516 \end{bmatrix} \end{cases}$$

Hence the controllability Gramian $W_c(0,\infty)$ is readily obtained by Corollary (6.1) as:

$$W_c(0,\infty) = vec^{-1}(-(T \otimes T)(A_c \oplus A_c)^{-1} vec(Q_c))$$

In this example:

$$W_c(0,\infty) = \begin{bmatrix} 0.3138 & -0.0004 & 0.2279 \\ -0.0004 & 0.0375 & 0.0046 \\ 0.2279 & 0.0046 & 0.1758 \end{bmatrix}$$

□

Example (6.2): Consider the stable LTI system below:

$$A=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \text{-}727.8173 & \text{-}185.0919 & \text{-}5.8585 \end{bmatrix}, B=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

According to Remark (6.3) the trace of the controllability Gramian is obtained as:

$$trace(W_c^c(0,\infty)) = -(o_{13}(\bar{O}_{13}) + o_{23}(\bar{O}_{23}) + o_{33}(\bar{O}_{33})) = -(0\text{-}0.0014\text{-}0.2596) = 0.2610$$

where:

$$\bar{O}_{13} = \left(A_c^3 - \beta_2 A_c^2 + \beta_1 A_c - \beta_0 I\right)^{-1}, \begin{cases} \bar{O}_{23} = -A_c \bar{O}_{13} \\ \bar{O}_{33} = A_c^2 \bar{O}_{13} \end{cases} \Rightarrow$$

$$\bar{O}_{13} = \left(A_c^3 - 5.8585 A_c^2 + 185.0919 A_c - 727.8173 I\right)^{-1}, \begin{cases} \bar{O}_{23} = -A_c \bar{O}_{13} \\ \bar{O}_{33} = A_c^2 \bar{O}_{13} \end{cases} \Rightarrow$$

$$\begin{cases} \bar{O}_{13} = \begin{bmatrix} \text{-}0.0007 & \text{-}0.0001 & \text{-}0.0000 \\ 0.0082 & 0.0014 & \text{-}0.0000 \\ 0.0000 & 0.0082 & 0.0014 \end{bmatrix} \\ \bar{O}_{23} = \begin{bmatrix} \text{-}0.0082 & \text{-}0.0014 & 0.0000 \\ \text{-}0.0000 & \text{-}0.0082 & \text{-}0.0014 \\ 1.0207 & 0.2596 & \text{-}0.0000 \end{bmatrix} \Rightarrow \\ \bar{O}_{33} = \begin{bmatrix} 0.0000 & 0.0082 & 0.0014 \\ \text{-}1.0207 & \text{-}0.2596 & 0.0000 \\ \text{-}0.0000 & \text{-}1.0207 & \text{-}0.2596 \end{bmatrix} \end{cases}$$

□

Example (6.3): Consider the stable LTI system below:

$$A_c=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \text{-}6.8913 & \text{-}8.2417 & \text{-}4.1328 \end{bmatrix}, B_c=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

For this stable controller form system, the corresponding matrices below are calculated in example (6.1):

$$\bar{O}_{13} = \left(A_c^3 - 4.1328 A_c^2 + 8.2417 A_c - 6.8913 I \right)^{-1} \begin{cases} \bar{O}_{23} = -A_c \bar{O}_{13} \\ \bar{O}_{33} = A_c^2 \bar{O}_{13} \end{cases} \Rightarrow$$

$$\begin{cases} \bar{O}_{13} = \begin{bmatrix} -0.0727 & -0.0457 & -0.0111 \\ 0.0761 & 0.0184 & -0.0000 \\ 0.0000 & 0.0761 & 0.0184 \end{bmatrix} \\ \bar{O}_{23} = \begin{bmatrix} -0.0761 & -0.0184 & 0.0000 \\ -0.0000 & -0.0761 & -0.0184 \\ 0.1267 & 0.1516 & -0.0000 \end{bmatrix} \\ \bar{O}_{33} = \begin{bmatrix} 0.0000 & 0.0761 & 0.0184 \\ -0.1267 & -0.1516 & 0.0000 \\ -0.0000 & -0.1267 & -0.1516 \end{bmatrix} \end{cases}$$

Then according to Remark (6.4) an upper-bound of the maximum eigenvalue of the controllability Gramian is:

$$\sqrt{trace(W_c^c(0,\infty))^2} = \sqrt{\sum_{i=1}^{3} (\bar{O}_{i3}(:,3))^T \bar{O}_{i3}(:,3)}$$

$$= \sqrt{\begin{bmatrix} -0.0111 & 0 & 0.0184 \end{bmatrix} \begin{bmatrix} -0.0111 \\ 0 \\ 0.0184 \end{bmatrix} + \begin{bmatrix} 0 & -0.0184 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -0.0184 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.0184 & 0 & -0.1516 \end{bmatrix} \begin{bmatrix} 0.0184 \\ 0 \\ -0.1516 \end{bmatrix}}$$

Hence:

$$\sqrt{trace(W_c^c(0,\infty))^2} = \sqrt{0.0241} = 0.1554$$

Further, according to Corollary (6.2) we know that the controllability Gramian of the system in controller companion form can be obtained by the last column of $(A_c \oplus A_c)^{-1}$ as:

$$W_c^c(0,\infty) = \begin{bmatrix} 0.0110 & -0.0000 & -0.0184 \\ -0.0000 & 0.0184 & 0.0000 \\ -0.0184 & 0.0000 & 0.1516 \end{bmatrix}$$

The eigenvalues of the controllability Gramian are $\lambda_1 = 0.1540, \lambda_2 = 0.0184, \lambda_3 = 0.0087$ and they verify the upper bound for the maximum eigenvalue of $W_c^c(0,\infty)$, i.e.

$$\lambda_1 = 0.1540 \le \sqrt{trace(W_c^c(0,\infty))^2} = 0.1554$$

□

In the sequel, a simple method for calculating the controllability Gramian of the stable LTI system in companion form is proposed. This method involves the solution of a linear system of equations based on the system's characteristic polynomial.

Definition (6.1): (Xiao, et al., 1992) A square $n \times n$ matrix $P = [p_{ij}]$ is said to be a Xiao matrix, if it has only n independent elements, and has the property:

$$p_{ij} = \begin{cases} 0, i + j \neq 2k, k = 1, \ldots, n \\ (-1)^{\frac{(i-j)}{2}} p_{kk}, i + j = 2k, k = 1, \ldots, n \end{cases} \tag{6.22}$$

□

Proposition (6.2): (Xiao, et al., 1992) Let the stable LTI system (6.1) be in controller companion form. The controllability Gramian $W_c^c(0, \infty)$ is a Xiao matrix and can be obtained as the solution of the quasi Hankel form linear system of equations below:

If the dimension of the system is even, i.e. $A \in \mathbb{R}^{n \times n}$, $n = 2k$, $k = 1, 2, 3, \ldots$

$$\begin{bmatrix} -\beta_0 & \beta_2 & -\beta_4 & \beta_6 & \cdots & \cdots & \cdots & 0 \\ 0 & -\beta_1 & \beta_3 & -\beta_5 & \cdots & \cdots & \cdots & 0 \\ 0 & \beta_0 & -\beta_2 & \beta_4 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \beta_1 & -\beta_3 & \cdots & \cdots & \cdots & 0 \\ & & \vdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & & -2\beta_1 & \cdots & -2\beta_{n-3} & -2\beta_{n-1} \end{bmatrix}_{n \times n} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix} \tag{6.23}$$

If the dimension of the system is odd, i.e. $A \in \mathbb{R}^{n \times n}$, $n = 2k - 1$, $k = 1, 2, 3, \ldots$

$$\begin{bmatrix} -\beta_0 & \beta_2 & -\beta_4 & \beta_6 & \cdots & \cdots & \cdots & 0 \\ 0 & -\beta_1 & \beta_3 & -\beta_5 & \cdots & \cdots & \cdots & 0 \\ 0 & \beta_0 & -\beta_2 & \beta_4 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \beta_1 & -\beta_3 & \cdots & \cdots & \cdots & 0 \\ & & \vdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & & -2\beta_0 & \cdots & -2\beta_{n-3} & -2\beta_{n-1} \end{bmatrix}_{n \times n} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix} \tag{6.24}$$

Here $\beta_i, i = 0, \ldots, n$ are the coefficients of the characteristic polynomial, and the vector x includes the elements of the Xiao matrix $W_c^c(0, \infty)$:

In an even dimensional system, i.e. $A \in \mathbb{R}^{n \times n}$, $n = 2k$, $k = 1, 2, 3, \ldots$

$$W_c^c(0,\infty)=\begin{bmatrix} x_1 & 0 & -x_2 & 0 & x_3 & \cdots & 0 \\ 0 & x_2 & 0 & -x_3 & \ddots & \ddots & \vdots \\ -x_2 & 0 & x_3 & \ddots & \ddots & & x_{n-2} \\ 0 & -x_3 & \ddots & \ddots & & -x_{n-2} & 0 \\ x_3 & \ddots & \ddots & & x_{n-2} & 0 & -x_{n-1} \\ \vdots & \ddots & & -x_{n-2} & 0 & x_{n-1} & 0 \\ 0 & \cdots & x_{n-2} & 0 & -x_{n-1} & 0 & x_n \end{bmatrix} \tag{6.25}$$

In an odd dimensional system, i.e. $A \in \mathbb{R}^{n\times n}, n=2k-1$, $k=1,2,3,\ldots$

$$W_c^c(0,\infty)=\begin{bmatrix} x_1 & 0 & -x_2 & 0 & x_3 & \cdots & x_{\frac{(n+1)}{2}} \\ 0 & x_2 & 0 & -x_3 & \ddots & \ddots & \vdots \\ -x_2 & 0 & x_3 & \ddots & \ddots & & x_{n-2} \\ 0 & -x_3 & \ddots & \ddots & & -x_{n-2} & 0 \\ x_3 & \ddots & \ddots & & x_{n-2} & 0 & -x_{n-1} \\ \vdots & \ddots & & -x_{n-2} & 0 & x_{n-1} & 0 \\ x_{\frac{(n+1)}{2}} & \cdots & x_{n-2} & 0 & -x_{n-1} & 0 & x_n \end{bmatrix} \tag{6.26}$$

□

The proof is given in (Xiao, et al., 1992).

In the above Proposition matrices (6.23), (6.24) are called quasi Hankel since by permuting rows and columns so that all the odd numbered rows and columns precede the even numbered ones, the matrix transforms into a 2-block diagonal matrix, each block being almost a Hankel matrix.

Regarding Proposition (6.2) a simple method is proposed for the calculating the controllability Gramian in asymptotically stable systems, which involves evaluating only the elements of a Routh table and does not involve the solution of a linear system of equations or matrix inversion (Sreeram & Agathoklis, 1991).

Proposition (6.3): In n-dimensional stable LTI system in controller companion form, the zero-plaid structured controllability Gramian $W_c^c(0,\infty)$ described in (6.25), (6.26) can be obtained as:

$$\begin{cases} x_n = \dfrac{0.5}{R_{n1}} \\ x_{n-k} = \dfrac{-\sum_{i=1}^{m-1}(-1)^i R_{n-k,i+1}x_{n-k+i}}{R_{n-k,1}}, k=1,\ldots,n-1 \end{cases} \tag{6.27}$$

where R_{ij} denotes the entry in the Routh table of the reciprocal system corresponding to the ith row and jth column and m denotes the number of elements in the (n-k)th row of the Routh table.

□

The proof is given in (Sreeram & Agathoklis, 1991).

Note: Consider the characteristic polynomial of an LTI system $p(s) = s^n + \beta_{n-1}s^{n-1} + \cdots + \beta_1 s + \beta_0$. Then the Routh table for the reciprocal characteristic polynomial is defined as: $\tilde{p}(s) = s^n p(s^{-1})$.

Example (6.4): Consider the stable LTI system in the previous example. The reciprocal polynomial of the system is:

$$\tilde{p}(s) = \beta_0 s^3 + \beta_1 s^2 + \beta_2 s + 1 = 6.8913s^3 + 8.2417s^2 + 4.1328s + 1$$

Then the Routh table of the reciprocal system is:

$$\begin{cases} s^3 & 6.8913 & 4.1328 \\ s^2 & 8.2417 & 1 \\ s^1 & 3.2966 & 0 \\ s^0 & 1 \end{cases}$$

Then according to Proposition (6.3) we have:

$$W_c^c(0,\infty) = \begin{bmatrix} x_1 & 0 & -x_2 \\ 0 & x_2 & 0 \\ -x_2 & 0 & x_3 \end{bmatrix}$$

where:

$$x_3 = \frac{0.5}{3.2966} = 0.1517$$
$$x_2 = \frac{1 \times 0.1517}{8.2417} = 0.0184$$
$$x_1 = \frac{4.1328 \times 0.0184}{6.8913} = 0.0110$$

□

Remark (6.5): The trace of the controllability Gramian in the controller companion form of the stable LTI system (6.1), is equal to:

$$trace(W_c^c(0,\infty)) = \sum_{k=1}^{n} x_k \tag{6.28}$$

where:

$$\begin{cases} x_n = \dfrac{0.5}{R_{n1}} \\ x_i = \dfrac{-\sum_{j=1}^{m-1}(-1)^j R_{i,j+1} x_{i+j}}{R_{i,1}}, i = 1,\ldots,n-1 \end{cases} \tag{6.29}$$

Here R_{ij} denotes the entry in the Routh table corresponding to the i[th] row and j[th] column and m denotes the number of elements in the j[th] row of the Routh table.

□

Remark (6.6): An upper bound on the maximum eigenvalue of the controllability Gramian of the stable LTI system (6.1) in controller form is equal to:

$$\lambda_{\max}(W_c^c(0,\infty)) \le \sqrt{\sum_{i=1}^{n} q_i x_i^2} \tag{6.30}$$

where:

$$\begin{cases} x_n = \dfrac{0.5}{R_{n1}} \\ x_i = \dfrac{-\sum_{j=1}^{m-1}(-1)^j R_{i,j+1} x_{i+j}}{R_{i,1}}, i = 1,\ldots,n-1 \end{cases} \tag{6.31}$$

and:

$$\begin{cases} q_i = 2i-1, i \le \text{int}\{\dfrac{n}{2}\} \\ q_i = 2n-(2i-1), i > \text{int}\{\dfrac{n}{2}\} \end{cases} \tag{6.32}$$

Here R_{ij} denotes the entry in the Routh table corresponding to the i[th] row and j[th] column and m denotes the number of elements in the j[th] row of the Routh table. Further, $\text{int}\{\alpha\}$ is defined $\dfrac{\alpha}{2}$ if $\alpha = 2k, k = 1,2,\ldots$ and $\dfrac{\alpha+1}{2}$ if $\alpha = 2k-1, k = 1,2,\ldots.$

□

The above Remark states that the eigenvalues of the controllability Gramian of the stable LTI system in controller form is bounded from above by a function of x_i s defined in Proposition (6.3).

Example (6.5): Consider the LTI system in example (6.2). The Routh table of the reciprocal system is:

$$\begin{cases} s^3 & 727.8173 & 5.8585 \\ s^2 & 185.0919 & 1 \\ s^1 & 1.9263 & 0 \\ s^0 & 1 \end{cases}$$

According to Remark (6.5) we have:

$$trace(W_c^c(0,\infty)) = \frac{0.5}{\underbrace{1.9263}_{0.2596}} + \frac{0.2596}{\underbrace{185.0919}_{0.0014}} + \frac{5.8585 \times 0.0014}{727.8173} = 0.2596 + 0.0014 + 1.1290\text{e-}05 = 0.26101129$$

Remark (6.6) determines also an upper bound of the maximum eigenvalues of the controllability Gramian as:

$$\lambda_{\max}(W_c^c(0,\infty)) = 0.2596 \le \sqrt{x_1^2 + 3x_2^2 + x_3^2} = \sqrt{0.0674} = 0.2596$$

□

6.3. Inverse of Controllability Gramian and Energy Calculation in Controller Canonical Form

In this section, we introduce a simple algorithm for finding the inverse controllability Gramian of a SISO stable system based on the Routh table. This helps in feedback design when a constraint on the input control energy is imposed. By relating the minimum input energy to the location of the eigenvalues of the system.

Note that in the previous section we introduced some methods for finding the controllability Gramian $W_c^c(0,\infty)$ in controller companion form, which has close resemblance to our method for finding the inverse of controllability Gramian of a SISO stable LTI system in phase variable form (Controller companion form) $L = (W_c^c(0,\infty))^{-1}$. Our method is important since it does not involve solution of a linear system of equations or matrix inversion and hence is a computationally efficient way for calculating the control input energy in various problems of feedback control design or pole placement with input energy constraints, compared to other techniques. This result is illustrated by a number of numerical examples.

Theorem (6.1): Consider the SISO stable LTI system (6.1) in controller companion form. The inverse of the controllability Gramian $L = (W_c^c(0,\infty))^{-1}$ can be derived as the zero-plaid structured matrix in terms of the characteristic polynomial's coefficients:

$$|sI - A| = s^n + \beta_{n-1}s^{n-1} + \cdots + \beta_0 \tag{6.33}$$

If the dimension of the system is odd, i.e. $A \in \mathbb{R}^{n\times n}, n = 2k+1, k = 0,1,2,\ldots$:

$$\begin{cases} s^n & 1 & \beta_{n-2} & \dots & \beta_1 \\ s^{n-1} & \beta_{n-1} & \beta_{n-3} & \dots & \beta_0 \\ s^{n-2} & & & & \\ \vdots & & & & \\ s^1 & & & & \end{cases} \tag{6.34}$$

and:

$$L = L^T = \begin{bmatrix} L_1 & 0 & L_2 & 0 & \cdots & 0 & L_{k+1} \\ & L_{k+2} & 0 & L_{k+3} & \cdots & L_{(k+1-1)+k+1} & 0 \\ & & L_{2k+2} & 0 & \cdots & 0 & L_{((2k+1)-(k+2)+1)+(2k+1)} \\ & & & L_{3k+2} & 0 & \cdots & 0 \\ & & & & L_{4k+1} & \cdots & L_{(4k-(3k+2)+1)+4k} \\ & & & & & \ddots & \vdots \\ & & & & & & L_{(\frac{2k(k+1)-n}{2}+n)} \end{bmatrix} \tag{6.35}$$

or:

$$L = L^T = \begin{bmatrix} 2\beta_0\beta_1 & 0 & 2\beta_0\beta_3 & 0 & \cdots & 0 & 2\beta_0 \\ 0 & 2\beta_2\beta_1 - 2\beta_0\beta_3 & 0 & 2\beta_4\beta_1 - 2\beta_0\beta_5 & \cdots & 2\beta_{n-1}\beta_1 - 2\beta_0 & 0 \\ & 0 & 2\beta_2\beta_3 - (2\beta_4\beta_1 - 2\beta_0\beta_5) & 0 & \cdots & 0 & 2\beta_2 \\ & & 0 & \ddots & \cdots & 2\beta_{n-1}\beta_3 - 2\beta_2 & 0 \\ & & & & \ddots & \vdots & 2\beta_4 \\ & & & & & \vdots & \vdots \\ & & & & & 0 & 2\beta_{n-1} \end{bmatrix} \tag{6.36}$$

If the dimension is even, i.e. $A \in \mathbb{R}^{n \times n}, n = 2k$, $k = 0,1,2,\dots$:

$$\begin{cases} s^n & 1 & \beta_{n-2} & \dots & \beta_0 \\ s^{n-1} & \beta_{n-1} & \beta_{n-3} & \dots & 0 \\ s^{n-2} & & & & \\ \vdots & & & & \\ s^1 & & & & \end{cases} \tag{6.37}$$

Then:

$$L = L^T = \begin{bmatrix} L_1 & 0 & L_2 & 0 & \cdots & L_k & 0 \\ & L_{k+1} & 0 & L_{k+2} & \cdots & 0 & L_{((k-1)+1)+k} \\ & & L_{2k+1} & 0 & \cdots & L_{(2k-(k+1))+2k} & 0 \\ & & & L_{3k} & 0 & \cdots & L_{((3k-1)-(2k+1)+1)+(3k-1)} \\ & & & & L_{4k-1} & \cdots & 0 \\ & & & & & \ddots & \vdots \\ & & & & & & L_{(\frac{n(n+2)}{4})} \end{bmatrix} \tag{6.38}$$

or:

$$L = L^T = \begin{bmatrix} 2\beta_0\beta_1 & 0 & 2\beta_0\beta_3 & 0 & \cdots & 2\beta_0\beta_{n-1} & 0 \\ 0 & 2\beta_2\beta_1 - 2\beta_0\beta_3 & 0 & 2\beta_4\beta_1 - 2\beta_0\beta_5 & \cdots & 0 & 2\beta_1 \\ & 0 & 2\beta_2\beta_3 - (2\beta_4\beta_1 - 2\beta_0\beta_5) & 0 & \cdots & 2\beta_{n-1}\beta_2 - 2\beta_1 & 0 \\ & & 0 & \ddots & \cdots & 0 & 2\beta_3 \\ & & & & \ddots & \vdots & 0 \\ & & & & & \vdots & \vdots \\ & & & & & 0 & 2\beta_{n-1} \end{bmatrix} \tag{6.39}$$

□

Proof:

Assuming that A is Hurwitz and using the system's controllability Lyapunov equation we get:

$$AW_c^c + W_c^c A^T = -bb^T \Rightarrow LA + A^T L = -Lbb^T L \tag{6.40}$$

The system is in controller form then if $A \in \mathbb{R}^{n\times n}, n = 2k+1$, $k = 0,1,2,\ldots$ according to above Lyapunov equation we have:

$$\underbrace{\begin{bmatrix} L_1 & L_2 & L_3 & \ldots & L_n \\ L_2 & L_{n+1} & L_{n+2} & \ldots & L_{2n-1} \\ L_3 & L_{n+2} & L_{2n} & \ldots & L_{3n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_n & L_{2n-1} & L_{3n-2} & \ldots & L_{\frac{n(n+1)}{2}} \end{bmatrix}_{n\times n} \begin{bmatrix} 0_{(n-1)\times 1} & & I_{(n-1)\times(n-1)} \\ -\beta_0 & -\beta_1 \quad \cdots \quad -\beta_{n-2} & -\beta_{n-1} \end{bmatrix}_{n\times n}}_{H} + H_{n\times n}^T = -L \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}_{n\times n} L \tag{6.41}$$

Furthermore, note that since the controllability Gramian in controller companion form is symmetric and zero plaid (Sreeram & Agathoklis, 1991), (Xiao, et al., 1992), as it is stated in the previous section, the same must hold true for L. Thus (6.35) reduces to:

$$\begin{cases} \underbrace{\begin{bmatrix} L_1 & 0 & L_2 & 0 & \cdots & 0 & L_{k+1} \\ 0 & L_{k+2} & 0 & L_{k+3} & \cdots & L_{(k+1-1)+k+1} & 0 \\ L_2 & 0 & L_{2k+2} & 0 & \cdots & 0 & L_{((2k+1)-(k+2)+1)+(2k+1)} \\ 0 & L_{k+3} & 0 & L_{3k+2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & 0 & L_{4k+1} & \cdots & L_{(4k-(3k+2)+1)+4k} \\ 0 & L_{2k+1} & 0 & \vdots & \vdots & \ddots & \vdots \\ L_{k+1} & 0 & L_{3k+1} & 0 & L_{5k-1} & \cdots & L_{(\frac{2k(k+1)-n}{2}+n)} \end{bmatrix}}_{H} \begin{bmatrix} 0_{(n-1)\times 1} & I_{(n-1)\times(n-1)} \\ -\beta_0 & -\beta_1 \quad \cdots \quad -\beta_{n-2} \quad -\beta_{n-1} \end{bmatrix}_{n\times n} \Rightarrow \\ H + H^T_{n\times n} = -L \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}_{n\times n} L \end{cases}$$

(6.42)

By calculating the right hand-side we obtain:

$$-\begin{bmatrix} L_{k+1}^2 & 0 & L_{k+1}L_{3k+1} & 0 & \cdots & 0 & L_{k+1}L_{(\frac{2k(k+1)-n}{2}+n)} \\ 0 & 0 & \cdots & & & 0 & 0 \\ L_{3k+1}L_{k+1} & 0 & L_{3k+1}^2 & 0 & L_{3k+1}L_{5k-1} & \cdots & L_{3k+1}L_{(\frac{2k(k+1)-n}{2}+n)} \\ 0 & 0 & \cdots & & & 0 & 0 \\ L_{5k-1}L_{k+1} & 0 & L_{5k-1}L_{3k+1} & 0 & L_{5k-1}^2 & \cdots & L_{5k-1}L_{(\frac{2k(k+1)-n}{2}+n)} \\ \vdots & & & & & \ddots & \vdots \\ L_{(\frac{2k(k+1)-n}{2}+n)}L_{k+1} & 0 & L_{(\frac{2k(k+1)-n}{2}+n)}L_{3k+1} & 0 & L_{(\frac{2k(k+1)-n}{2}+n)}L_{5k-1} & \cdots & L_{(\frac{2k(k+1)-n}{2}+n)}^2 \end{bmatrix}$$

(6.43)

Hence equation (6.42) defines the elements of symmetric matrix L as:

$$\begin{cases} -2L_{k+1}\beta_0 = -L_{k+1}{}^2 \\ L_1 - L_{k+1}\beta_1 = 0 \\ -L_{k+1}\beta_2 - L_{3k+1}\beta_0 = -L_{k+1}L_{3k+1} \\ L_2 - L_{k+1}\beta_3 = 0 \\ -L_{k+1}\beta_4 - L_{5k-1}\beta_0 = -L_{k+1}L_{5k-1} \\ \vdots \\ -L_{k+1}\beta_{n-1} - L_{\frac{2k(k+1)-n}{2}+n}\beta_0 = -L_{k+1}L_{\frac{2k(k+1)-n}{2}+n} \\ L_{k+2} + L_2 - L_{3k+1}\beta_1 = 0 \\ L_{k+3} + L_3 - L_{5k-1}\beta_1 = 0 \\ \vdots \\ -2L_{\frac{2k(k+1)-n}{2}+n}\beta_{n-1} = -L^2_{\frac{2k(k+1)-n}{2}+n} \end{cases} \Rightarrow \begin{cases} L_{k+1} = 2\beta_0 \\ L_1 = 2\beta_1\beta_0 \\ L_{3k+1} = 2\beta_2 \\ L_2 = 2\beta_3\beta_0 \\ L_{5k-1} = 2\beta_4 \\ \vdots \\ L_{\frac{2k(k+1)-n}{2}+n} = 2\beta_{n-1} \\ L_{k+2} = 2\beta_2\beta_1 - 2\beta_3\beta_0 \\ L_{k+3} = 2\beta_4\beta_1 - 2\beta_5\beta_0 \\ \vdots \\ L_{\frac{2k(k+1)-n}{2}+n} = 2\beta_{n-1} \end{cases} \tag{6.44}$$

Thus (6.36) is verified.

Now consider the case that $A \in \mathbb{R}^{n\times n}, n = 2k$, $k = 0,1,2,\ldots$. Regarding the zero-plaid structure of matrix L we can write the Lyapunov equation (6.40) as:

$$\underbrace{\begin{bmatrix} L_1 & 0 & L_2 & 0 & \cdots & L_k & 0 \\ 0 & L_{k+1} & 0 & L_{k+2} & \cdots & 0 & L_{((k-1)+1)+k} \\ L_2 & 0 & L_{2k+1} & 0 & \cdots & L_{(2k-(k+1))+2k} & 0 \\ 0 & L_{k+2} & 0 & L_{3k} & 0 & \cdots & L_{((3k-1)-(2k+1)+1)+(3k-1)} \\ \vdots & \vdots & \vdots & 0 & L_{4k-1} & \cdots & 0 \\ & & & \vdots & \vdots & \ddots & \vdots \\ 0 & L_{2k} & 0 & L_{4k-2} & 0 & \cdots & L_{\frac{n(n+2)}{4}} \end{bmatrix}}_{H_{n\times n}} \begin{bmatrix} 0_{(n-1)\times 1} & & I_{(n-1)\times(n-1)} \\ -\beta_0 & -\beta_1 & \cdots & -\beta_{n-2} & -\beta_{n-1} \end{bmatrix} \tag{6.45}$$

which implies that:

$$H+H_{n\times n}^{T}=-\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & L_{2k}^{2} & 0 & L_{2k}L_{4k-2} & 0 & L_{2k}L_{6k-6} & 0 & \cdots & L_{2k}L_{(\frac{n(n+2)}{4})} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & L_{2k}L_{4k-2} & 0 & L_{4k-2}^{2} & 0 & L_{4k-2}L_{6k-6} & 0 & \cdots & L_{4k-2}L_{(\frac{n(n+2)}{4})} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & L_{2k}L_{6k-6} & 0 & L_{4k-2}L_{6k-6} & 0 & L_{6k-6}^{2} & 0 & \cdots & L_{6k-6}L_{(\frac{n(n+2)}{4})} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & L_{2k}L_{(\frac{n(n+2)}{4})} & 0 & L_{4k-2}L_{(\frac{n(n+2)}{4})} & 0 & L_{6k-6}L_{(\frac{n(n+2)}{4})} & 0 & \cdots & L_{(\frac{n(n+2)}{4})}^{2}
\end{bmatrix}_{n\times n} \tag{6.46}$$

Therefore, the nonzero elements of the symmetric matrix L are readily obtained by Gaussian elimination with partial pivoting to equation (6.46) as:

$$\begin{cases}
-2\beta_0 L_{k+1}=-L_{k+1}^{2} \\
L_1-\beta_1 L_{k+1}=0 \\
-\beta_2 L_{k+1}-\beta_0 L_{3k+1}=-L_{k+1}L_{3k+1} \\
L_2-\beta_3 L_{k+1}=0 \\
-\beta_4 L_{k+1}-\beta_0 L_{5k-1}=-L_{k+1}L_{5k-1} \\
\vdots \\
-\beta_{n-1} L_{k+1}-\beta_0 L_{\frac{n(n+2)}{4}}=-L_{k+1}L_{\frac{n(n+2)}{4}} \\
L_{k+2}+L_2-\beta_1 L_{3k+1}=0 \Rightarrow L_{k+2}=\beta_1 L_{3k+1}-L_2 \\
\vdots \\
L_{k+3}+L_3-\beta_1 L_{5k-1}=0 \Rightarrow L_{k+3}=\beta_1 L_{5k-1}-L_3 \\
\vdots \\
L_{2k+2}+L_{k+3}-\beta_3 L_{3k+1}=0 \Rightarrow L_{2k+2}=\beta_3 L_{3k+1}-L_{k+3} \\
\vdots
\end{cases} \Rightarrow \begin{cases}
L_{k+1}=2\beta_0 \\
L_1=2\beta_1\beta_0 \\
L_{3k+1}=2\beta_2 \\
L_2=2\beta_3\beta_0 \\
L_{5k-1}=2\beta_4 \\
\vdots \\
L_{\frac{n(n+2)}{4}}=2\beta_{n-1} \\
L_{k+2}=2\beta_1\beta_2-2\beta_3\beta_0 \\
\vdots \\
L_{k+3}=2\beta_1\beta_4-2\beta_5\beta_0 \\
\vdots \\
L_{2k+2}=2\beta_3\beta_2-2\beta_1\beta_4+2\beta_5\beta_0 \\
\vdots
\end{cases} \tag{6.47}$$

Thus (6.39) is proved.

□

Considering the above Theorem, it is of interest to mention that the minimum energy and the inverse of the controllability Gramian of SISO stable LTI system can be directly obtained from the Routh–Hurwitz table through a simple algorithm.

1) At the beginning, we define the system in controller companion form. Then:

2) Define the first two rows of Routh–Hurwitz for the SISO stable LTI system

$A_c \in \mathbb{R}^{n\times n}, n=2k+1, k=0,1,2,\ldots$	$A_c \in \mathbb{R}^{n\times n}, n=2k, k=0,1,2,\ldots$
$\begin{matrix} s^n & 1 & \beta_{n-2} & \cdots & \beta_1 \\ s^{n-1} & \beta_{n-1} & \beta_{n-3} & \cdots & \beta_0 \\ s^{n-2} & & & & \\ \vdots & & & & \\ s^0 & & & & \end{matrix}$	$\begin{matrix} s^n & 1 & \beta_{n-2} & \cdots & \beta_0 \\ s^{n-1} & \beta_{n-1} & \beta_{n-3} & \cdots & 0 \\ s^{n-2} & & & & \\ \vdots & & & & \\ s^0 & & & & \end{matrix}$

3) The last column of L is obtained by multiplying the second row of coefficients in the Routh–Hurwitz table.

$A_c \in \mathbb{R}^{n\times n}, n=2k+1, k=0,1,2,\ldots$	$A_c \in \mathbb{R}^{n\times n}, n=2k, k=0,1,2,\ldots$
$\begin{bmatrix} \times & 0 & \cdots & \times & 0 & 2\beta_0 \\ 0 & \times & \cdots & 0 & \times & 0 \\ \times & 0 & \times & \ddots & 0 & 2\beta_2 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \times & \ddots & \ddots & \ddots & 0 \\ 2\beta_0 & 0 & \cdots & \times & 0 & 2\beta_{n-1} \end{bmatrix}$	$\begin{bmatrix} \times & 0 & \cdots & 0 & \times & 0 \\ 0 & \times & \ddots & \times & 0 & 2\beta_1 \\ \vdots & \ddots & \times & \ddots & \times & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \times & 0 & \ddots & \ddots & \times & 0 \\ 0 & \cdots & 0 & \times & 0 & 2\beta_{n-1} \end{bmatrix}$

4) The first row of L can be readily determined based on the last column elements and the first row of coefficients in the Routh–Hurwitz table as:

$A_c \in \mathbb{R}^{n\times n}, n=2k+1, k=0,1,2,\ldots$
$2\beta_0\beta_1 \quad 0 \quad 2\beta_0\beta_3 \quad \cdots \quad 2\beta_0\beta_{n-4} \quad 0 \quad 2\beta_0\beta_{n-2} \quad 0 \quad 2\beta_0$
$A_c \in \mathbb{R}^{n\times n}, n=2k, k=0,1,2,\ldots$
$2\beta_0\beta_1 \quad 0 \quad 2\beta_0\beta_3 \quad \cdots \quad 0 \quad 2\beta_0\beta_{n-3} \quad 0 \quad 2\beta_0\beta_{n-1} \quad 0$

5) The rest of the elements are determined by the opposite sign of their upper right hand side element plus a proper coefficient of the last column's arrays. These coefficients are selected from the first row of the corresponding Routh–Hurwitz table:

$$A_c \in \mathbb{R}^{n\times n}, n = 2k+1, k = 0,1,2,\ldots$$

$$\begin{array}{cccccc}
\beta_1 \circ & \cdots & \beta_{n-4} \circ & 0 & \beta_{n-2} \circ & 0 \quad \circ \\
0 & \cdots & 0 & \beta_1 * - \beta_{n-2} \circ & 0 & \beta_1 \times - \circ \quad 0 \\
 & \ddots & \vdots & \vdots & \vdots & \vdots \quad \vdots \\
 & & \beta_{n-6} \bullet & 0 & \beta_{n-2}\Delta & 0 \quad \Delta \\
 & & & \beta_{n-4} \bullet + \beta_{n-6} * - \beta_{n-2}\Delta & 0 & \beta_{n-6} \times - \Delta \quad 0 \\
 & & & 0 & \beta_{n-2} \bullet - (+\beta_{n-6} \times - \Delta) & 0 \quad \bullet \\
 & & & \beta_{n-4} * - (+\beta_{n-2} \bullet - (+\beta_{n-6} \times - \Delta)) & 0 & \beta_{n-4} \times - \bullet \quad 0 \\
 & & & & \beta_{n-2} * - (+\beta_{n-4} \times - \bullet) & 0 \quad * \\
 & & & & & \beta_{n-2} \times - * \quad 0 \\
 & & & & & 0 \quad \times
\end{array}$$

$$A_c \in \mathbb{R}^{n\times n}, n = 2k, k = 0,1,2,\ldots$$

$$\begin{array}{cccccc}
\beta_0 \circ & \cdots & 0 & \beta_0 * & 0 & \beta_0 \times \quad 0 \\
 & \beta_2 \circ & \cdots & 0 & \beta_{n-2} \circ - \beta_0 \times & 0 \quad \circ \\
 & & 0 & \beta_1 * - (+\beta_{n-2} \circ - \beta_0 \times) & 0 & \beta_2 \times - \circ \quad 0 \\
 & & \vdots & \vdots & \vdots & \vdots \quad \vdots \\
 & & \beta_{n-6} \bullet & 0 & \beta_{n-2}\Delta & 0 \quad \Delta \\
 & & & \beta_{n-4} \bullet + \beta_{n-6} * - \beta_{n-2}\Delta & 0 & \beta_{n-6} \times - \Delta \quad 0 \\
 & & & 0 & \beta_{n-2} \bullet - (+\beta_{n-6} \times - \Delta) & 0 \quad \bullet \\
 & & & \beta_{n-4} * - (+\beta_{n-2} \bullet - (+\beta_{n-6} \times - \Delta)) & 0 & \beta_{n-4} \times - \bullet \quad 0 \\
 & & & & \beta_{n-2} * - (+\beta_{n-4} \times - \bullet) & 0 \quad * \\
 & & & & & \beta_{n-2} \times - * \quad 0 \\
 & & & & & 0 \quad \times
\end{array}$$

6) In the last step of the algorithm, we just need to use the non-singular similarity matrix:

$$T = \begin{bmatrix} b & Ab & A^2 b & \cdots & A^{n-1} b \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_{n-1} & 1 \\ \beta_2 & \beta_3 & \cdots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

to find the inverse of the controllability Gramian $W_c^{-1} = T^{-T} L T^{-1}$, then the minimum input energy is readily obtained as $E_{\min} = x_f^T W_c^{-1} x_f$.

□

The importance of this method lies in the fact that given a stable linear time-invariant continuous-time system in the frequency domain (i.e. transfer function $G(s)=\frac{b(s)}{a(s)}$) or the time domain (i.e. state space), we are able to calculate the minimum input energy directly in terms of the coefficients of the transfer function (i.e. $a(s), b(s)$) or equivalently based on the Routh-Hurwitz table.

Let us to review the details of the proposed method by some examples:

Example (6.6): As in the first example, consider the stable 4-dimensional LTI system below:

$$A=\begin{bmatrix} -1.1577 & 0.0161 & -0.5121 & -0.1164 \\ 0.0161 & -0.9540 & -0.1244 & 0.1954 \\ -0.5121 & -0.1244 & -1.7388 & 0.0544 \\ -0.1164 & 0.1954 & 0.0544 & -2.0123 \end{bmatrix}, B=\begin{bmatrix} -1.0690 \\ -0.8095 \\ -2.9440 \\ 0 \end{bmatrix}$$

We are interested to calculate the minimum energy to transfer the states of the system from the origin to the final state $x_f=\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}^T$.

To be able to use our algorithm we consider the controller model of the system as:

$$A_c=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -3.2289 & -10.5455 & -12.1920 & -5.8628 \end{bmatrix}, B_c=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Then regarding the dimension of the system ($n=4 \Rightarrow n=2k, k=2$), based on Theorem (6.1) we can define the symmetric matrix L as:

$$L=\begin{bmatrix} L_1 & 0 & L_2 & 0 \\ \times & L_3 & 0 & L_4 \\ \times & \times & L_5 & 0 \\ \times & \times & \times & L_6 \end{bmatrix}$$

Then in the first step of the algorithm, we write the Routh–Hurwitz table for the system:

$$\begin{cases} s^4 & 1 & 12.1920 & 3.2289 \\ s^3 & 5.8628 & 10.5455 & 0 \\ s^2 & & & \\ s^1 & & & \\ s^0 & & & \end{cases}$$

In the second step, the last column of matrix L is determined as the second row of the Routh–Hurwitz coefficients multiplied by two:

$$L_6 = 2 \times 5.8628 = 11.7256$$
$$L_4 = 2 \times 10.5455 = 21.0910$$

According to the next step of the algorithm we determine the nonzero elements in the first row of matrix L:

$$L_2 = 3.2289 \times L_6 = 3.2289 \times 11.7256 = 37.8608$$

$$L_1 = 3.2289 \times L_4 = 3.2289 \times 21.0910 = 68.1007$$

In the fourth step, we calculate the rest nonzero elements of the symmetric matrix L as:

$$L = \begin{bmatrix} 68.1007 & 0 & 37.8608 & 0 \\ \times & 12.1920 \times 21.0910 - 37.8608 & 0 & 21.0910 \\ \times & \times & 12.1920 \times 11.7256\text{-}21.0910 & 0 \\ \times & \times & \times & 11.7256 \end{bmatrix}$$

Hence, we get the inverse of the controllability Gramian in the controller form as:

$$L = \begin{bmatrix} 68.1007 & 0 & 37.8608 & 0 \\ 0 & 219.2807 & 0 & 21.0910 \\ 37.8608 & 0 & 121.8675 & 0 \\ 0 & 21.0910 & 0 & 11.7256 \end{bmatrix}$$

Now in the last step, we can easily use the similarity transformation T in Lemma (6.2) to obtain the inverse of the controllability Gramian W_c^{-1} and to calculate the minimum input energy of the system:

$$T = \begin{bmatrix} -1.0690 & 2.7322 & -6.0760 & 12.8639 \\ -0.8095 & 1.1215 & -1.7816 & 3.1611 \\ -2.9440 & 5.7671 & -11.5769 & 23.4956 \\ 0 & -0.1940 & 0.6054 & -1.4895 \end{bmatrix} \begin{bmatrix} 10.5455 & 12.1920 & 5.8628 & 1.0000 \\ 12.1920 & 5.8628 & 1.0000 & 0 \\ 5.8628 & 1.0000 & 0 & 0 \\ 1.0000 & 0 & 0 & 0 \end{bmatrix}$$

or:

$$T = \begin{bmatrix} -0.7199 & -3.0907 & -3.5351 & -1.0690 \\ -2.1476 & -5.0760 & -3.6244 & -0.8095 \\ -5.1108 & -13.6589 & -11.4930 & -2.9440 \\ -0.3051 & -0.5319 & -0.1940 & 0 \end{bmatrix}$$

Hence:

$$W_c^{-1} = T^{-T} L T^{-1} = 1.0\text{e+}06 \times \begin{bmatrix} 0.0168 & -0.0327 & 0.0028 & 0.1423 \\ -0.0327 & 0.5428 & -0.1369 & -1.4594 \\ 0.0028 & -0.1369 & 0.0366 & 0.3482 \\ 0.1423 & -1.4594 & 0.3482 & 4.1250 \end{bmatrix}$$

and finally:

$$E_{\min} = x_f^T W_c^{-1} x_f = 1.6759\text{e+}04$$

□

Example (6.7): In the second example, we calculate the minimum input energy for the 5-dimensional stable system below:

$$A = \begin{bmatrix} -1.0860 & 1.2673 & 0.2865 & -2.1291 & -0.7724 \\ -0.9140 & -0.4367 & -0.4485 & 0.5639 & 0.6074 \\ -0.5004 & 0.0405 & -0.5401 & -0.4105 & 0.0616 \\ 2.0285 & 0.0340 & 0.3304 & -0.9072 & -0.9572 \\ 1.2718 & -0.3384 & 0.1786 & 0.3489 & -0.3977 \end{bmatrix}, B = \begin{bmatrix} -0.2725 \\ 1.0984 \\ -0.2779 \\ 0 \\ 0 \end{bmatrix}$$

Assume the terminal state is defined as: $x_f = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T$.

In the beginning, we need to define the controller form of the system:

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -0.1409 & -2.0769 & -9.2391 & -11.6151 & -3.3676 \end{bmatrix}, B_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Noting that the dimension of the system is odd ($n = 5 \Rightarrow n = 2k+1, k = 2$), we define the structure of the symmetric matrix L using Theorem (6.1) as:

$$L = \begin{bmatrix} L_1 & 0 & L_2 & 0 & L_3 \\ \times & L_4 & 0 & L_5 & 0 \\ \times & \times & L_6 & 0 & L_7 \\ \times & \times & \times & L_8 & 0 \\ \times & \times & \times & \times & L_9 \end{bmatrix}$$

In the first step, we write the first two rows of Routh–Hurwitz table for the system:

$$\begin{cases} s^5 & 1 & 11.6151 & 2.0769 \\ s^4 & 3.3676 & 9.2391 & 0.1409 \\ s^3 & & & \\ s^2 & & & \\ s^1 & & & \\ s^0 & & & \end{cases}$$

Then the nonzero elements in the last column of L are determined as the second row's elements of the Routh–Hurwitz multiplied by two:

$$L_9 = 2 \times 3.3676 = 6.7352$$
$$L_7 = 2 \times 9.2391 = 18.4782$$
$$L_3 = 2 \times 0.1409 = 0.2818$$

In the next step of the algorithm, the first row of matrix L is described as:

$$L_2 = 11.6151 \times 0.2817 = 3.2720$$
$$L_1 = 2.0769 \times 0.2817 = 0.5851$$

Then in the third step of the algorithm, we readily obtain the rest of nonzero elements of the symmetric matrix L as:

$$L = \begin{bmatrix} L_1 & 0 & L_2 & 0 & L_3 \\ \times & 2.0769L_7\text{-}L_2 & 0 & 2.0769L_9 - L_3 & 0 \\ \times & \times & 11.6151L_7\text{-}(2.0769L_9 - L_3) & 0 & L_7 \\ \times & \times & \times & 11.6151L_9 - L_7 & 0 \\ \times & \times & \times & \times & L_9 \end{bmatrix}$$

Thus:

$$L = \begin{bmatrix} 0.5851 & 0 & 3.2720 & 0 & 0.2818 \\ 0 & 35.1051 & 0 & 13.7066 & 0 \\ 3.2720 & 0 & 200.9196 & 0 & 18.4782 \\ 0 & 13.7066 & 0 & 59.7518 & 0 \\ 0.2818 & 0 & 18.4782 & 0 & 6.7352 \end{bmatrix}$$

Now in the last stage, we can readily calculate the minimum input energy as:

$$T = \begin{bmatrix} -0.2725 & 1.6084 & 0.0997 & -14.9854 & 35.1977 \\ 1.0984 & -0.1060 & -2.3810 & 5.2822 & 9.3239 \\ -0.2779 & 0.3309 & -0.7859 & -1.4947 & 11.2815 \\ 0 & -0.6072 & 4.6541 & -6.4990 & -26.2780 \\ 0 & -0.7679 & 2.2339 & 1.5276 & -23.9877 \end{bmatrix} \begin{bmatrix} 2.0769 & 9.2391 & 11.6151 & 3.3676 & 1.0000 \\ 9.2391 & 11.6151 & 3.3676 & 1.0000 & 0 \\ 11.6151 & 3.3676 & 1.0000 & 0 & 0 \\ 3.3676 & 1.0000 & 0 & 0 & 0 \\ 1.0000 & 0 & 0 & 0 & 0 \end{bmatrix}$$

or:

$$T = \begin{bmatrix} 0.1836 & 1.5140 & 2.3513 & 0.6908 & -0.2725 \\ 0.7597 & 6.1815 & 10.0204 & 3.5931 & 1.0984 \\ -0.3997 & -2.8647 & -2.8989 & -0.6048 & -0.2779 \\ 0.2832 & 2.1218 & 2.6093 & -0.6072 & 0 \\ 0.0098 & 0.1321 & -0.3519 & -0.7679 & 0 \end{bmatrix}$$

Hence:

$$W_c^{-1} = T^{-T} L T^{-1} = 1.0\text{e+}03 \times \begin{bmatrix} 0.4138 & 0.0699 & -0.0493 & -0.5970 & 1.1412 \\ 0.0699 & 0.0246 & 0.0176 & -0.1033 & 0.2358 \\ -0.0493 & 0.0176 & 0.0825 & 0.0978 & -0.0763 \\ -0.5970 & -0.1033 & 0.0978 & 0.9071 & -1.6991 \\ 1.1412 & 0.2358 & -0.0763 & -1.6991 & 3.4126 \end{bmatrix}$$

and then:

$$E_{\min} = x_f^T W_c^{-1} x_f = 413.7565$$

□

6.4. Controllability Gramian Computation in Diagonal Canonical Form

In this section, we investigate the controllability Gramian of the stable linear continuous time-invariant system (6.1) as the solution of the Lyapunov equation (6.2), in the case that matrix A is diagonalizable.

Let T be a non-singular transformation matrix such that:

$$T^{-1} A T = \sum$$

where $\sum$ is diagonal matrix.

Note that T is the matrix of the eigenvectors and $\sum$ is the diagonal matrix of the eigenvalues of A.

Define a new state vector $\tilde{x} = T^{-1} x$, then the linear system (6.1) transforms to:

$$\begin{cases} \dot{\tilde{x}} = \sum \tilde{x} + T^{-1} B u \\ y = CT\tilde{x} + Du \end{cases} \tag{6.48}$$

which represents n decoupled first order equations.

Denoting the solution of the Lyapunov equation associated with above system as W_c^{diag}, the controllability Gramian of the original system (6.1) will be equal to:

$$W_c = T W_c^{diag} T^T \tag{6.49}$$

The Proposition below presents the explicit solution of the Lyapunov equation in diagonal canonical form.

Proposition (6.4): Consider the stable LTI system (6.48). The controllability Gramian $W_c^{diag}(0,t_f)$ is obtained as:

$$W_c^{diag}(0,t_f))_{i,j} = \left(T^{-1}B\right)_i \left(T^{-1}B\right)_j^T \frac{\left(e^{(\lambda_i+\lambda_j)t_f}-1\right)}{\lambda_i+\lambda_j} \tag{6.50}$$

where $\left(T^{-1}B\right)_i, \left(T^{-1}B\right)_j^T$ represent i[th] and j[th] rows of $T^{-1}B$ and λ_i denotes i[th] eigenvalue of A and T is the eigenvectors matrix of A.

□

Proof:

According to the definition of the controllability Gramian, we have:

$$W_c^{diag}(0,t_f) = \int_0^{t_f} e^{\Sigma\tau} T^{-1} B B^T T^{-T} e^{\Sigma\tau} d\tau \tag{6.51}$$

Since Σ is a diagonal matrix, exponentiation can be performed simply by exponentiating each of the diagonal elements:

$$W_c^{diag}(0,t_f) = \int_0^{t_f} \begin{bmatrix} e^{\lambda_1\tau} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2\tau} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n\tau} \end{bmatrix} \underbrace{T^{-1}BB^TT^{-T}}_{H} \begin{bmatrix} e^{\lambda_1\tau} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2\tau} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n\tau} \end{bmatrix} d\tau \tag{6.52}$$

Using the Kronecker product we obtain:

$$vec(W_c^{diag}(0,t_f)) = (\int_0^{t_f} \left(\begin{bmatrix} e^{\lambda_1\tau} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2\tau} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n\tau} \end{bmatrix} \otimes \begin{bmatrix} e^{\lambda_1\tau} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2\tau} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n\tau} \end{bmatrix} \right) d\tau) vec(H) \tag{6.53}$$

where:

$$vec(W_c^{diag}(0,t_f)) = (\int_0^{t_f} e^{(\Sigma\oplus\Sigma)\tau} d\tau) vec(H) \tag{6.54}$$

or:

$$vec(W_c^{diag}(0,t_f)) = (\int_0^{t_f} \begin{bmatrix} e^{2\lambda_1\tau} & 0 & \cdots & 0 \\ 0 & e^{\lambda_1+\lambda_2\tau} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{2\lambda_n\tau} \end{bmatrix} d\tau) vec(H) \tag{6.55}$$

Equation (6.55) yields:

$$\begin{cases} W_c^{diag}(0,t_f))_{1,1} = (\int_0^{t_f} e^{2\lambda_1\tau} d\tau)\left(T^{-1}B\right)_1 \left(T^{-1}B\right)_1^T \\ W_c^{diag}(0,t_f))_{2,1} = (\int_0^{t_f} e^{(\lambda_1+\lambda_2)\tau} d\tau)\left(T^{-1}B\right)_2 \left(T^{-1}B\right)_1^T \\ \vdots \\ W_c^{diag}(0,t_f))_{i,j} = (\int_0^{t_f} e^{(\lambda_i+\lambda_j)\tau} d\tau)\left(T^{-1}B\right)_i \left(T^{-1}B\right)_j^T \\ \vdots \\ W_c^{diag}(0,t_f))_{n,n} = (\int_0^{t_f} e^{(\lambda_n+\lambda_n)\tau} d\tau)\left(T^{-1}B\right)_n \left(T^{-1}B\right)_n^T \end{cases} \quad (6.56)$$

and the Proposition is proved.

□

Corollary (6.3): The solution of the Lyapunov equation of the stable LTI system (6.41) is:

$$W_c^{diag}(0,\infty))_{i,j} = -\left(T^{-1}B\right)_i \left(T^{-1}B\right)_j^T (\lambda_i + \lambda_j)^{-1} \quad (6.57)$$

where $\left(T^{-1}B\right)_i, \left(T^{-1}B\right)_j^T$ represent i[th] and j[th] rows of $T^{-1}B$, λ_i denotes i[th] eigenvalue of A and T is the eigenvectors matrix of A.

□

Corollary (6.4): Consider the stable LTI system (6.48) in diagonal form. Then the trace of the controllability Gramian when $t_f \to \infty$ is obtained as:

$$trace(W_c^{diag}(0,\infty)) = -0.5\sum_{i=1}^{n} \frac{\left(T^{-1}BB^TT^{-T}\right)_{ii}}{\lambda_i} \quad (6.58)$$

Again $\left(T^{-1}B\right)_i, \left(T^{-1}B\right)_j^T$ denote the i[th] and j[th] rows of $T^{-1}B$, λ_i is i[th] eigenvalue of A and T is the eigenvector matrix of A.

□

6.5. Summary

Two different canonical form structures of the system have been discussed in this chapter:

- Phase-variable canonical form
- Diagonal canonical form

We considered the stable LTI system in the above canonical form structures and introduced the interesting links that the entries of the controllability Gramian of the system have to the entries of the Routh Hurwitz table. Moreover, an expression for the minimum input energy is developed as a simple function of the coefficients of the characteristic polynomial from the fact that it is connected to the inverse of the controllability Gramian.

Furthermore, we devised an approach to determine the value of the trace and an upper-bound for the maximum eigenvalue of the controllability Gramian based directly on the coefficients of the characteristic polynomial of the system.

In the next chapter, the problem that will be considered is the variability of the input structure. This will lead to the investigation of the selection of the optimal input structure under the minimum average energy requirement. A variety of such problem will be examined next.

Chapter 7: Optimal Selection of Input Structure, Minimizing the Average of Input Energy

7.1. Introduction

In this Chapter, we propose a novel approach for selection of input (and by duality output) structure based on the energy type criteria. The main aim of this chapter is to propose a strategy for the selection of the proper input matrix B, based on the average of the minimum input energy. The overall aim is to use these results to define the proper number and location of the sensors and actuators in a control system.

In the following parts of the chapter we discuss the input structure selection problem in different cases. First, we consider the case that the input matrix B is a single input vector. We look for the best input structure to minimise the energy requirements and we propose a solution for a general system and a finite terminal time t_f. Then, in the next part we assume the system to be stable and that the terminal time tends to infinity, using appropriate Lyapunov equations. This is followed by the input structure selection for a normal system where $AA^* = A^*A$. Finally, in the case of multi-input systems, we define the proper set of

input vectors that minimize the average of the minimum input energy, and we investigate the problem subject to different possible conditions.

The work here provides a new approach to those in existing literature, which consider the problem of input structure selection over a binary set, where the input matrix B can only take the values {0,1} (V. Tzoumas, 2015), (Kumar & Narayanan, 2008), (Shaker & Tahavori, 2013) or can be chosen among the given sets (Summers & Lygeros, 2014), (Cortesi, et al., 2015). In contrast, the current approach developed in this chapter doesn't have such limitations.

7.2. Optimal Selection of Input to Minimize the Average of Minimum Input Energy

In previous chapters, the minimum input energy was defined as a function of the inverse of the controllability Gramian and the effect of input structure on the value of minimum energy requirements was examined. In this chapter, we are looking for the best selection of the input matrix B, such that the average of the minimum energy is minimized. The mathematical formulation of this optimization problem is presented in (7.1).

$$\min_{\substack{B\in\beta \\ \|B\|=1}} \{ \underset{\|x_f\|=1}{ave}\, E_{\min[0,t_f]} \} = \min_{\substack{B\in\beta \\ \|B\|=1}} \{ \underset{\|x_f\|=1}{ave}\, x_f^T W_c^{-1}(0,t_f) x_f \} \tag{7.1}$$

where β is the set of all possible inputs and $E_{\min[0,t_f]}$ is the minimum input energy evaluated over the interval $\left[0,t_f\right]$. x_f defines the final state and $W_c^{-1}(0,t_f)$ is the inverse of the controllability Gramian. $\|.\|$ denotes the standard Euclidean 2-norm.

Theorem (7.1): The average value of the minimum input energy E_{min} over the unit hypersphere is equal to $n^{-1}trace(W_c^{-1}(0,t_f))$ where n denotes the dimension of the system.

□

Proof:

Based on the results of chapter 4, E_{min}, the minimum energy required for the state trajectory to reach the final state x_f from the origin in a terminal time t_f, was found as:

$$E_{\min[0,t_f]} = x_f^T W_c^{-1} x_f \tag{7.2}$$

The average value of $E_{\min[0,t_f]}$ over the unit hypersphere is obtained by integrating (7.2), $\left\|x_f\right\| = 1$, (Summers & Lygeros, 2014):

$$aveE_{\min[0,t_f]} = \frac{\int_{\|x_f\|=1} x_f^T W_c^{-1}(0,t_f) x_f dx}{\int_{\|x_f\|=1} dx} = \frac{trace(W_c^{-1}(0,t_f))}{n} \quad (7.3)$$

and hence the Theorem is proved.

□

Note that the trace of $W_c(0,t_f)$is inversely related to the trace of $W_c^{-1}(0,t_f)$, and thus maximizing $trace\left(W_c(0,t_f)\right)$ effectively minimizes the average of E_{min} required to move around the state space in all directions (Le, et al., 2015), (Cortesi, et al., 2015).

For an asymptotically stable strictly-proper system its H_2norm is an important design parameter and is defined as (Antoulas, 2005):

$$\|H\|_2 = (trace \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) H^*(j\omega) d\omega)^{\frac{1}{2}} \quad (7.4)$$

This can be interpreted as the RMS[3] response of the system when it is driven by a white noise input.

Proposition (7.1): (Summers & Lygeros, 2014) Let $H(s)$ be an asymptotically stable strictly proper matrix function with realization (A,B,C). Then:

$$W_c(0,\infty) = \int_0^{\infty} e^{A\tau} BB^* e^{A^*\tau} d\tau$$

is the controllability Gramian of the system, from which the system H_2norm is obtained as:

$$trace\{CW_c(0,\infty)C^*\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} trace\{H(jw)H^*(jw)\} dw = \|H(s)\|_2^2 \quad (7.5)^4$$

□

Proof:

The controllability Gramian corresponding to a stable LTI system can be expressed in the frequency domain using Parseval's Theorem as:

$$W_c(0,\infty) = \int_0^{\infty} e^{A\tau} BB^* e^{A^*\tau} d\tau = \frac{1}{2\pi} \int_{-\infty}^{\infty} (jwI - A)^{-1} BB^* (-jwI - A^*)^{-1} dw \quad (7.6)$$

Now if we multiply both sides by output matrix C , we get:

[3] RMS norm is not a norm, but only a semi-norm, since we can have nonzero signals with zero RMS value

[4] For simplicity of notation, quantities in the time and frequency domains will be denoted by the same symbol.

$$CW_c(0,\infty)C^* = \frac{1}{2\pi}\int_{-\infty}^{\infty} C(jwI-A)^{-1}BB^*(-jwI-A^*)^{-1}C^* dw = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(jw)G^*(jw)dw$$

Then it can be easily inferred from the above equation:

$$trace\{CW_c(0,\infty)C^*\} = \frac{1}{2\pi} trace\int_{-\infty}^{\infty} G(jw)G^*(jw)dw \qquad (7.7)$$

which proves the result.

□

The Proposition above shows that the H_2 norm of a stable strictly proper system is equal to the L_2 norm of its (causal) impulse response, i.e. the energy of the system response to a unit impulse input, and can be represented as a weighted trace of the controllability Gramian $W_c(0,\infty)$.
Thus, maximizing the H_2 norm, minimizes a weighted average energy required to move around the state space, with certain directions weighted differently, which can be encoded into the output matrix.

In the following, we discuss the proper selection of input matrix B, such that the average energy of the trajectory between the origin and a terminal state uniformly distributed over the unit hypersphere, which is formulated in (7.1) is minimized.
Thus, the main aim in this chapter is the optimization of the average of the minimum energy, over the set of all possible input matrices B.
According to what was stated above, the optimization problem can be equivalently expressed as the maximization of the controllability Gramian's trace over the set of all inputs defined over the interval $\left[0, t_f\right]$ and satisfying the condition $\|B\| = 1$:

$$\max_{\substack{B\in\beta \\ \|B\|=1}} \{trace(W_c(0,t_f))\} \qquad (7.8)$$

Note again that β is the set of all possible inputs.
In this chapter, considering the dimension of input matrix B, the solution to the presented optimization problem (7.8) is investigated in two main parts:

- **Single input:**
  Where matrix B is an input vector as $b_{n\times 1}$
- **Multiple inputs:**
  Where input matrix B is a matrix of dimension $n\times m$

7.2.1. Single Input Matrix Structure Selection Problem

Consider the case of a single input LTI system. The aim is to select the proper input vector $b_{n\times 1}$, such that the trace of the controllability Gramian $W_c(0,t_f)$ is maximized. This guarantees the least value of the average of the minimum input energy E_{min}.

According to the equation (7. 8), we describe the optimization problem in this section as:

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{aveE_{\min[0,t_f]}\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,t_f))\} \tag{7. 9}$$

where β is the set of all possible inputs and $E_{\min[0,t_f]}$ is the minimum input energy evaluated over the interval $\left[0,t_f\right]$.

In the following, this problem is investigated by considering various separate cases, such as the finite and infinite terminal times, and different properties for the system matrix A.

7.2.1.a. Optimal Input Vector Selection Considering the Finite Terminal Time

Here, we investigate the solution to the optimization problem (7.9) for the general single input LTI system, while the terminal time t_f is finite. No special assumptions on system matrix A are imposed. Note that the system may be stable, or unstable.

Problem:

Consider a LTI SISO system, find

$\max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,t_f))\}$, where β is the set of all possible input vector

Theorem (7. 2): In a LTI single input system, the eigenvector corresponding to the maximum eigenvalue of matrix Q defined below describes the optimal input vector b, which minimizes the average of the minimum input energy, i.e.

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{aveE_{\min[0,t_f]}\} = \lambda_{\max}(Q)$$

$$Q_{n\times n} = \int_0^{t_f} e^{A^* t} e^{At} dt \tag{7.10}$$

$$\arg\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{aveE_{\min[0,t_f]}\} = b_{n\times 1}^{opt} = V_{\max}(Q)$$

β is the set of all possible inputs. $\lambda_{\max}(Q)$, $V_{\max}(Q)$ denote the largest eigenvalue of Q and the corresponding eigenvector respectively.

□

Proof:

Considering the definition of the controllability Gramian:

$$W_c(0,t_f) = \int_0^{t_f} e^{At} b b^* e^{A^* t} dt \tag{7. 11}$$

we get:

$$trace(W_c(0,t_f)) = trace(\int_0^{t_f} e^{At} b b^* e^{A^* t} dt) \tag{7.12}$$

Using the linearity property of trace, the equation above can be stated as:

$$trace(W_c(0,t_f)) = \int_0^{t_f} trace(e^{At} b b^* e^{A^* t}) dt \tag{7.13}$$

Using the commutative property of trace, this can be rewritten as:

$$trace(W_c(0,t_f)) = \int_0^{t_f} trace(b^* e^{A^* t} e^{At} b) dt \tag{7.14}$$

Since $b^* e^{A^* t} e^{At} b$ is scalar, the trace function can be ignored, and we get:

$$trace(W_c(0,t_f)) = \int_0^{t_f} b^* e^{A^* t} e^{At} b dt \tag{7.15}$$

and since b and b^* are constant, we obtain:

$$trace(W_c(0,t_f)) = b^* (\int_0^{t_f} e^{A^* t} e^{At} dt) b \tag{7.16}$$

By defining the n-dimensional matrix Q:

$$Q_{n\times n} = \int_0^{t_f} e^{A^* t} e^{At} dt \tag{7.17}$$

we get:

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,t_f]}\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{trace(W_c(0,t_f))\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{b^* Q b\} = \lambda_{\max}(Q) \tag{7.18}$$

and:

$$\arg\max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{trace(W_c(0,t_f))\} = b_{n\times 1}^{opt} = V_{\max}(Q) \tag{7.19}$$

Here $\lambda_{max}(Q)$ is the largest eigenvalue of Q and $V_{max}(Q)$ is the eigenvector corresponding to it.

The equation above shows that in a LTI single input system, the eigenvector corresponding to the maximum eigenvalue of $Q_{n\times n} = \int_0^{t_f} e^{A^* t} e^{At} dt$, describes the input matrix b, which optimizes the average of the minimum input energy.

□

Example (7.1): (Rubin, 2016) Consider the LTI system below, which may have a real-world example such as a rocket in vertical motion:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + b_{2\times 1} u$$

Now the problem can be defined as the selection of input vector b , which maximizes the trace of controllability Gramian, i.e. the minimum average of input energy.

Based on the Theorem above, the eigenvector corresponding to the maximum eigenvalue of matrix Q in (7.17) corresponds to the optimal input vector b , which leads to the maximum trace of the controllability Gramian. In addition, as one can easily see, the system is unstable since the eigenvalues of the system are located in the origin, $\lambda_1 = \lambda_2 = 0$. Let $t_f = 1$ be the (finite) terminal time. Then based on (7.17) we get:

$$Q = \int_0^1 e^{A^* t} e^{At} dt = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1.3333 \end{bmatrix}$$

The characteristic polynomial of Q is:

$$\left|\lambda I - A\right| = \left\| \begin{bmatrix} \lambda - 1 & -0.5 \\ -0.5 & \lambda - 1.3333 \end{bmatrix} \right| \Rightarrow \lambda^2 - 2.3333\lambda + 1.0833 = 0$$

and hence the two eigenvalues and corresponding eigenvectors are:

$$\lambda_1 = 1.6937,\ \lambda_2 = \ 0.6396$$

and

$$V_1 = \begin{bmatrix} 0.5847 \\ 0.8112 \end{bmatrix}, V_2 = \begin{bmatrix} \text{-}0.8112 \\ 0.5847 \end{bmatrix}$$

respectively.

Then based on Theorem (7.2) choosing $b = V_1 = \begin{bmatrix} 0.5847 \\ 0.8112 \end{bmatrix}$, we achieve the maximum trace of the controllability Gramian, that is equal to 1.6937. This is illustrated in Figure (7.1) for vector b described uniformly in the interval $[-1,1]$ to verify the validity of this result.

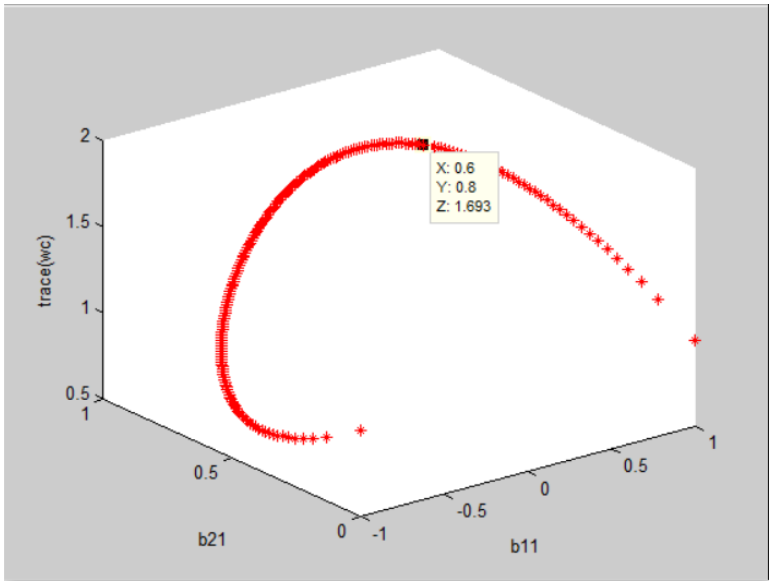


Figure (7.1): $trace(W_c(0,1))$ with respect to different values of b

□

The importance of Theorem (7.2) is that it can be effectively used to find the optimal input vector for any LTI-SISO system, stable or unstable, with the minimum possible complexity in computation. After matrix Q is calculated, the eigenvector corresponding to its largest eigenvalue is the solution of the optimization problem.

7.2.1.b. Optimal Input Vector Selection for the Case of Stable System

Problem:

Given a stable LTI SISO system, find

$\max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{trace(W_c(0,\infty))\}$, where β is the set of all possible input vector

For a single input LTI system, we now assume that the system is stable (i.e., all eigenvalues of A are in the open left half plane). We also consider the limit of the terminal time t_f tends to infinity, i.e. $t_f \to \infty$.

Proposition (7.2): For single input LTI system, optimal selection of the input vector b , which maximizes the trace of the controllability Gramian $W_c(0,\infty)$, i.e. minimizes the average of the minimum input energy, is given by the normalized eigenvector corresponding to the maximum eigenvalue of the unique solution to the Lyapunov equation below:

$$A^*Q + QA + I = 0 \tag{7.20}$$

Further, the value of the maximum trace is equal to the maximum eigenvalue of Q i.e.

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,\infty]}\} = \lambda_{\max}(Q) \tag{7.21}$$

$$\arg\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,\infty]}\} = b_{n\times 1}^{opt} = V_{\max}(Q) \tag{7.22}$$

$\lambda_{\max}(Q)$, $V_{\max}(Q)$ denote the largest eigenvalue of Q and the corresponding eigenvector respectively. β denotes the set of all possible input vector.

□

Proof:

For simplicity, the largest eigenvalue of Q is assumed to be distinct. The result is still true if this condition is violated but in this case the solution is not uniquely defined. The controllability Gramian has been defined as the integral (7.23):

$$W_c(0,\infty) = \int_0^\infty e^{At} b b^* e^{A^* t} dt \tag{7.23}$$

If the LTI system is stable, i.e. A is Hurwitz, the controllability Gramian $W_c(0,\infty)$ is the unique solution for the Lyapunov equation (Raczyński & Stanisławski, 2012):

$$A^* W_c(0,\infty) + W_c(0,\infty) A + bb^* = 0 \tag{7.24}$$

Similar to section (7.2.1.a), we consider the trace of the controllability Gramian as:

$$trace(W_c(0,\infty)) = b^*(\int_0^\infty e^{A^* t} e^{At} dt) b \tag{7.25}$$

and then by introducing the constant matrix Q as in (7.26) below, we derive the optimum input vector which minimizes the average of the minimum input energy:

$$\begin{aligned} trace(W_c(0,\infty)) &= b^* Q b \\ Q &= \int_0^\infty e^{A^* t} e^{At} dt \end{aligned} \tag{7.26}$$

Clearly Q is the unique solution of the Lyapunov equation:

$$A^* Q + QA + I = 0 \tag{7.27}$$

and the solution of trace optimization problem (7.9) is obtained as:

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,\infty]}\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{trace(W_c(0,\infty))\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{b^* Q b\} = \lambda_{\max}(Q) \tag{7.28}$$

and

$$\arg\max_{\substack{b_{n\times1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,\infty))\} = b_{n\times1}^{opt} = V_{\max}(Q) \tag{7.29}$$

Again β is the set of all possible inputs.

□

7.2.1.c. Optimal Input Vector in The Case That "A" Is a Normal Matrix

Problem:

Consider the standard LTI SISO system and assume that A is normal. Find

$\max_{\substack{b_{n\times1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,t_f))\}$, where β is the set of all possible input vectors

Definition (7.1): In LTI system, if A is normal, then it satisfies the property in (7.30):

$$AA^* = A^*A \tag{7.30}$$

Note that $A + A^*$ is always a normal matrix and it is diagonalizable by unitary matrices.

□

Lemma (7. 1): Consider the LTI system, in which A is normal. Then the optimal input vector b, which minimizes the average input energy within finite time $t_f > 0$, is given by the eigenvector, i.e. $V_{\max}(A+A^*)$, corresponding to the maximum eigenvalue of $A+A^*$, i.e. $\lambda_{\max}(A+A^*)$.

Furthermore, maximum value of the trace of the controllability Gramian is equal to:

$$\frac{e^{\lambda_{\max}t_f}-1}{\lambda_{\max}} \tag{7.31}$$

and

$$\arg\min_{\substack{b_{n\times1}\in\beta \\ \|b\|=1}}\{aveE_{\min[0,t_f]}\} = b_{n\times1}^{opt} = V_{\max}(A+A^*) \tag{7.32}$$

where β represents the set of all possible inputs.

□

Proof:

Similar to part (7.2.1.a) by using the linearity property of the trace we define the optimization problem as the maximization of (7.33):

$$trace(W_c(0,t_f)) = b^*(\int_0^{t_f} e^{A^*t}e^{At}dt)b \tag{7.33}$$

Considering the normal assumption for the finite terminal time t_f, we have:

$$Q = \int_0^{t_f} e^{A^*t}e^{At}dt = \int_0^{t_f} e^{(A^*+A)t}dt \tag{7.34}$$

and hence:

$$Q = \int_0^{t_f} P\begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix} P^{-1}dt \tag{7.35}$$

Where$\lambda_i$defines the i[th] eigenvalue of $A + A^*$, and P, P^{-1} are orthogonal eigenvector matrices. Now by calculating the integral we get:

$$Q = P\begin{pmatrix} \frac{e^{\lambda_1 t}-1}{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{e^{\lambda_n t}-1}{\lambda_n} \end{pmatrix} P^{-1} \tag{7.36}$$

Then using matrix Q above, the solution of the optimization problem is obtained as:

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{aveE_{\min[0,t_f]}\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,t_f))\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{b^*Qb\} = \lambda_{\max}(Q) = \frac{e^{\lambda_1 t}-1}{\lambda_1} \tag{7.37}$$

and

$$\arg\max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,t_f))\} = b_{n\times 1}^{opt} = V_{\max}(Q) = V_{\max}(A+A^*) \tag{7.38}$$

where β denotes the set of all possible inputs, λ_1 is the maximum eigenvalue of matrix $A+A^*$, and $V_{\max}$ is its corresponding eigenvector.

□

7.2.1.d. Optimal Input Vector in The Case That "A" Is Stable and Normal

Problem:

Consider a stable LTI SISO system and assume A is stable and normal. Find

$\max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}}\{trace(W_c(0,\infty))\}$, where β is the set of all possible input vectors

Lemma (7.2): (Karcanias, et al., 2010) if $A\in\mathbb{C}^{n\times n}$ is asymptotically stable and normal, then it is also strongly stable, which is defined by the condition:

$$\lambda_i(A+A^*)<0, i=1,\ldots,n.$$

□

Proof:

Using the real Schur form of matrix A (Bai & Demmelt, 1992):

$$U^*A^*U = block-diag\{\lambda_1,\ldots,\lambda_k,\ldots,\begin{bmatrix}\sigma_i & -\omega_i\\ \omega_i & \sigma_i\end{bmatrix},\ldots\} \tag{7.39}$$

where λ_i are the real eigenvalues and $\sigma_i\pm j\omega_i$ the complex eigenvalues of A.

Thus:

$$U^*(A+A^*)U = block-diag\{2\lambda_1,\ldots,2\lambda_k,\ldots,\begin{bmatrix}2\sigma_i & 0\\ 0 & 2\sigma_i\end{bmatrix},\ldots\}$$

Then, by introducing $\hat{A}$:

$$\frac{A+A^*}{2}=\hat{A}$$

we have:

$$U^*\hat{A}U = block-diag\{\lambda_1,\ldots,\lambda_k,\ldots,\sigma_i,\sigma_i,\ldots\} \tag{7.40}$$

and from asymptotic stability $\lambda_i<0$, $\sigma_i<0$ and this establishes the negative-definiteness of $\hat{A}$ and the Lemma is proved.

□

The above result establishes a sufficient condition for strong stability in terms of the property of normality.

According to (Ricardo, 2009), the spectral Theorem for normal matrices states:

Theorem (7.3): For $A \in \mathbb{C}^{n\times n}$, there exists a unitary $n \times n$ matrix $P \in \mathbb{R}^{n\times n}$ such that $P^* AP = D$ (D is a diagonal matrix) if and only if A is normal.

□

Proof:

By the Schur decomposition, we can write any matrix as $A = P^* TP$, where P is unitary and T is upper-triangular. If A is normal clearly T is normal as well, since:

$$AA^* = A^* A \Rightarrow P^* TT^* P = P^* T^* TP \Rightarrow TT^* = T^* T \tag{7.41}$$

which implies that T must be diagonal. The converse is obvious.

□

Now, by combining the results of the last two sections, we can examine the problem of optimum input selection in the case that the LTI system is both normal and stable. We have the following results:

Proposition (7.3): Consider an LTI single input system with A is a normal and Hurwitz. Then, the optimal input vector b, which maximizes the trace of the controllability Gramian, i.e. minimizes the average of the input energy, is the eigenvector corresponding to the largest eigenvalue of $A + A^*$ and the maximum value of the trace is $-\lambda_1^{-1}$, where $\lambda_1 (< 0)$ is the maximum eigenvalue of $A + A^*$. Equivalently:

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,\infty]}\} = \frac{-1}{\lambda_1} \tag{7.42}$$

and

$$\arg\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,\infty]}\} = b_{n\times 1}^{opt} = V_{\max}(A + A^*) \tag{7.43}$$

where β denotes the set of all possible input vectors and $V_{\max}$ (i.e. optimal input b) is the normalised eigenvector associated with λ_1.

□

Proof:

Due to the assumption of stability for the system, we can increase the terminal time to infinity $t_f \to \infty$, and similar to the last part we define matrix Q as:

$$Q = \int_0^\infty e^{A^* t} e^{At} dt = \int_0^\infty e^{(A^*+A)t} dt \tag{7.44}$$

Note: the identity $e^{A^* t} e^{At} = e^{(A^*+A)t}$ follows directly from the assumed normal property of A.

Then, regarding Theorem (7.3), Q in (7.44) can be written as (7.45):

$$Q = P \begin{pmatrix} \frac{-1}{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{-1}{\lambda_n} \end{pmatrix} P^* \tag{7.45}$$

Thus, by combination the properties of stability and normality for the LTI system, the optimal input vector minimizing the average of input energy is achieved by:

$$\min_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{aveE_{\min[0,\infty]}\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{trace(W_c(0,\infty))\} = \max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{b^* Q b\} = \lambda_{\max}(Q) = \frac{-1}{\lambda_1} \tag{7.46}$$

and

$$\arg\max_{\substack{b_{n\times 1}\in\beta \\ \|b\|=1}} \{trace(W_c(0,\infty))\} = b_{n\times 1}^{opt} = V_{\max}(Q) = V_{\max}(A + A^*) \tag{7.47}$$

where β is the set of all possible inputs, $\lambda_1 (< 0)$ is the largest eigenvalue of $A + A^*$ and $V_{\max}$ is its corresponding eigenvector.

□

7.2.2 Multi-Input Matrix Structure Selection

In this part, we assume that input $B \in \mathbb{R}^{n\times m}$, is a $n \times m$ matrix:

$$\begin{cases} \dot{x} = A_{n\times n} x + B_{n\times m} u \\ y = C_{l\times n} x + D_{l\times m} u \end{cases} \tag{7.48}$$

The aim is to select the proper input matrix $B = [\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_m], \underline{b}_i \in \mathbb{R}^{n\times 1}$ such that the trace of controllability Gramian matrix $W_c(0, t_f)$ is maximized. Again, this is equivalent to the minimization of the average of minimum input energy E_{min}, i.e.

$$\min_{B_{n\times m}\in\beta} \{aveE_{\min[0,t_f]}\} = \max_{B_{n\times m}\in\beta} \{trace(W_c(0,t_f))\} \tag{7.49}$$

Now β is the set of all possible input matrices of appropriate dimensions.

In this part, we also investigate the problem subject to various assumptions defined on input matrix B. In each case, we suggest the selection of the proper set, over which the average energy minimization (i.e. maximization of the trace of the controllability Gramian) is carried out.

7.2.2.a. Input Matrix Selection Within a Finite Terminal Time Part I

Problem:

For the LTI MIMO system (7.48), find

$\max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F\le 1}}\{trace(W_c(0,t_f))\}$, where β is the set of all possible input matrices

In this part, we find the optimal input structure for the general case, where $B = [\underline{b}_1, \underline{b}_2, \dots, \underline{b}_m], \underline{b}_i \in \mathbb{R}^{n\times 1}$ is the input matrix such that $\|B\|_F \le 1$, and matrix $Q = \int_0^{t_f} e^{A^*t} e^{At} dt$ has n eigenvalues ordered as $\lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n > 0$. It is assumed for simplicity that λ_1 is distinct.

Here we are looking for the input matrix B that maximizes the trace of the controllability Gramian, $W_c(0,t_f)$, i.e. minimizes the average input energy.

Theorem (7.4): Consider the system in (7.48) and assume that $Q = \int_0^{t_f} e^{A^*t} e^{At} dt$ has a maximum eigenvalue of multiplicity one. Then, all optimal input $B = [\underline{b}_1, \underline{b}_2, \dots, \underline{b}_m], \underline{b}_i \in \mathbb{R}^{n\times 1}$, such that $\|B\|_F \le 1$, which minimize the average input energy, i.e. maximize the trace of the controllability Gramian $W_c(0,t_f)$, can be defined as:

$$\begin{aligned} b_1 &= \alpha_1 V_1 \\ b_2 &= \alpha_2 V_1 \\ &\vdots \\ b_m &= \alpha_m V_1 \end{aligned} \qquad (7.50)$$

where V_1 is the normalized eigenvector corresponding to the largest eigenvalue of Q and the α_i's $i = 1, \dots, m$ are scalars satisfying the condition $\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2 = 1$.

□

Proof:

We know that:

$$trace\{W_c(0,t_f)\} = trace\{\int_0^{t_f} e^{At} BB^* e^{A^*t} dt\} = trace\{B^* \int_0^{t_f} e^{A^*t} e^{At} dt B\} \qquad (7.51)$$

Defined Q as:

$$Q = \int_0^{t_f} e^{A^*t} e^{At} dt \qquad (7.52)$$

Then by substituting (7.52) into (7.51) we obtain:

$$trace\{W_c(0,t_f)\} = trace\{B^*QB\} \tag{7.53}$$

Equivalently:

$$trace\{W_c(0,t_f)\} = trace\{\begin{pmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_m^* \end{pmatrix} Q \begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix}\} \tag{7.54}$$

or:

$$trace\{W_c(0,t_f)\} = trace\{\begin{pmatrix} b_{1_{1\times n}}^* Q_{n\times n} b_{1_{n\times 1}} & \cdots & \times \\ \vdots & \ddots & \vdots \\ \times & \cdots & b_{m_{1\times n}}^* Q_{n\times n} b_{m_{n\times 1}} \end{pmatrix}_{m\times m}\} \tag{7.55}$$

and hence:

$$trace\{W_c(0,t_f)\} = \sum_{i=1}^{m} b_i^* Q b_i \tag{7.56}$$

Then according to the Rayleigh quotient eigenvalue problem (Golub & Loan, 1983), (Watkins, 1982), (Uper, July 2002), the solution for the maximization of the trace of controllability Gramian is achieved by selecting input vectors b_i, $i = 1,\ldots,m$, along the eigenvector corresponding to the largest eigenvalue of Q:

$$\begin{aligned} &\max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F \le 1 \\ i=1,2,..}} \{trace(W_c(0,t_f))\} = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F = \le 1 \\ i=1,2,..}} \{trace(B^*QB)\} = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F = 1 \\ i=1,2,..}} \sum_{i=1}^{m} b_i^* Q b_i \\ &= \max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F = 1}} \{b_1^* b_1 \times \frac{b_1^* Q b_1}{b_1^* b_1} + b_2^* b_2 \times \frac{b_2^* Q b_2}{b_2^* b_2} + \cdots + b_m^* b_m \times \frac{b_m^* Q b_m}{b_m^* b_m}\} \end{aligned} \tag{7.57}$$

so that:

$$\begin{aligned} &b_1 = \alpha_1 V_1, b_2 = \alpha_2 V_1, \ldots, b_m = \alpha_m V_1 \\ &\|B\|_F = 1 \Rightarrow \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_m^2 = 1 \end{aligned} \tag{7.58}$$

where β is the set of all possible input matrices, V_1 is the eigenvector corresponding to λ_1, the maximum eigenvalue of Q.

Note that optimality over the set $\|B\|_F \le 1$ occurs on the boundary, i.e. $\|B\|_F = 1$.

□

Then all columns of the optimal solution set are aligned along the direction of V_1 while we can distribute their lengths arbitrarily subject to the condition $\sum_{i=1}^{m}\|b_i\|^2=1$. This generalizes the solution of the single-input problem. In the next subsection, we investigate the optimal input structure $B\in\mathbb{R}^{n\times m}$ where the largest eigenvalue of Q has multiplicity $r\le n$.

7.2.2.a. Input Matrix Selection Within a Finite Terminal Time Part II

Problem:

Consider the LTI MIMO system (7.48) and suppose that $Q=\int_0^{t_f}e^{A^*t}e^{At}dt$ has maximum eigenvalue λ_1 of multiplicity $r\le n$. Find

$\max_{\substack{B_{n\times m}\in\beta\\ \|B\|_F\le 1}}\{trace(W_c(0,t_f))\}$, where β is the set of all possible input matrices

In this section, we find the optimal input matrix $B=[\underline{b}_1,\underline{b}_2,\dots,\underline{b}_m],\underline{b}_i\in\mathbb{R}^{n\times 1}$, such that $\|B\|_F\le 1$, and matrix $Q=\int_0^{t_f}e^{A^*t}e^{At}dt$ has n eigenvalues ordered as $\lambda_1=\lambda_2=\cdots\lambda_r>\lambda_{r+1}\ge\cdots\ge\lambda_n>0$, where λ_1 is the maximum eigenvalue of multiplicity $r\le n$.

Lemma (7.3): For the LTI MIMO system (7.48), the columns of the optimal input matrix $B=[\underline{b}_1,\underline{b}_2,\dots,\underline{b}_m],\underline{b}_i\in\mathbb{R}^{n\times 1}$, $\|B\|_F\le 1$, which maximizes the trace of controllability Gramian $W_c(0,t_f)$, are equal to:

$$b_i=\sum_{k=1}^{r}\alpha_{ik}V_k,\ i=1,\dots,m \tag{7.59}$$

where $V_k,k=1,\dots,r$ are independent (orthonormal) eigenvectors corresponding to the maximum eigenvalue of $Q=\int_0^{t_f}e^{A^*t}e^{At}dt$ and the α_i's $i=1,\dots,m$ are the constant coefficients that satisfy the condition $\sum_{i=1}^{m}\sum_{k=1}^{r}\alpha_{ik}^2=1$. In this case $rank(B)=min(r,m)$, and the maximum value for the trace of $W_c(0,t_f)$ is equal to the largest eigenvalue of Q.

The dimension of optimal input matrix depends on the maximum eigenvalue of Q, assumed of multiplicity r. If $m\le r$, then B is full rank and the optimal solution would be m

dimensional input matrix. Otherwise, $m-r$ of input vectors $b_i, i=1,\ldots,m$, would be dependent.

□

Proof:

Since the proof of this Lemma is quite similar to the proof of Theorem (7.4), then it is neglected.

□

Remark (7.1): Suppose that $m \le r \le n$. Thus, among all optimal solutions with $rank(B)=m$ the best (in the sense that the condition number of B is equal to one) is:

$$B=[b_1,\ldots,b_m]=\frac{1}{\sqrt{m}}[V_1,V_2,\ldots,V_m] \tag{7.60}$$

where $V_1,V_2,\ldots,V_m$ are m orthonormal eigenvectors corresponding to the maximum eigenvalue of Q.

□

7.2.2.a. Input Matrix Selection Within a Finite Terminal Time Part III

Problem:

Consider the LTI MIMO system (7.48) and suppose that $Q=\int_0^{t_f} e^{A^*t}e^{At}dt$ has maximum eigenvalue λ_1 of multiplicity $r<m\le n$. Find

$\max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F\le 1 \\ rank(B)=m}} \{trace(W_c(0,t_f))\}$, where β is the set of all possible input matrices

Here we complete the solution of finding the optimal matrix B over a finite terminal time t_f, by imposing the constraint on the rank of input matrix B to guarantee a full rank multiple input solution.

In other words, in this section we are looking for a full rank input matrix $B = [\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_m], \underline{b}_i \in \mathbb{R}^{n\times 1}$, $\|B\|_F \le 1$, that maximizes the trace of the controllability Gramian $W_c(0,t_f)$, i.e. minimizes the average of input energy, where r is the order of the maximum eigenvalue of $Q=\int_0^{t_f} e^{A^*t}e^{At}dt$ is less than $rank(B)=m$, i.e. $r<m\le n$.

Definition (7.2): Inner product of two vectors a, b is defined as:

$$\langle a, b \rangle = \|a\| \|b\| \cos\theta \tag{7.61}$$

From the Cauchy–Schwarz inequality it follows immediately that:

$$0 \le \left| \langle a, b \rangle \right| \le \|a\| \|b\| \tag{7.62}$$

□

Next, we seek to find an optimal input matrix $B_{n\times m}$ =[$\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_m$] of full rank, which maximizes the trace of the controllability Gramian $W_c(0, t_f)$ subject to constraint $\|B\|_F \le 1$. For simplicity, we assume that $r < rank(b) = m$, where r is the multiplicity of the largest eigenvalue of $Q = \int_0^{t_f} e^{A^* t} e^{At} dt$. To define a meaningful problem, we need to redefine the constraint region as described below.

In effect, this replaces the full rank constraint on B by a numerical-rank constraint of a certain tolerance. Specifically, the optimal solution requires that the first r input vectors of matrix B are aligned in the directions of the r independent eigenvectors corresponding to the maximum eigenvalue of Q. Let E be the spectral subspace generate by theses eigenvectors. A numerical rank constraint on B can be imposed by constraining the input vectors b_i, $i = r+1, \ldots, m$ to lie in directions whose angles with any vector in the spectral subspace E exceeds same minimum value 0< ε .

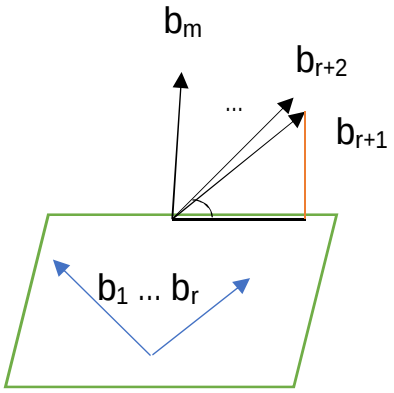


Figure (7.2): The optimal input vectors and the spectral subspace

Thus, the problem becomes:

$$\max_{\substack{\alpha_i^j, i=1,\ldots,1+(m-r) \\ j=1,\ldots,m}} \{ \sum_{i=1}^{1+(m-r)} \beta_i \lambda_i \} \tag{7.63}$$

where:

$$\beta_i = \sum_{j=1}^{m} \alpha_i^j, i = 1,\ldots,1+(m-r)$$

$$\sum_{j=1}^{r} (\alpha_1^j)^2 + \sum_{\substack{j=r+1 \\ i=1,\ldots,j}}^{m} (\alpha_i^j)^2 = 1 \tag{7.64}$$

$$0 < \varepsilon \le \cos\theta_i = \frac{\left|\langle b_k, b_l \rangle\right|}{\|b_k\|\|b_l\|}, i = 1,\ldots,(m-r)((m-r)+1)/2, k, l = r, r+1,\ldots,m$$

Then, the optimal solution is obtained by selecting:

$$\begin{aligned}
b_1 &= \alpha_1^1 V_1 \\
&\vdots \\
b_r &= \alpha_r^r V_r \\
b_{r+1} &= \sum_{i=1}^{r} \alpha_i^{r+1} V_i + \alpha_{r+1}^{r+1} V_{r+1} \\
b_{r+2} &= \sum_{i=1}^{r} \alpha_i^{r+2} V_i + \alpha_{r+1}^{r+2} V_{r+1} + \alpha_{r+2}^{r+2} V_{r+2} \\
&\vdots \\
b_m &= \sum_{i=1}^{r} \alpha_i^m V_i + \alpha_{r+1}^m V_{r+1} + \alpha_{r+2}^m V_{r+2} + \cdots + \alpha_{(m-r)+1}^m V_{(m-r)+1}
\end{aligned} \tag{7.65}$$

Here $V_i, i = 1,\ldots,1+(m-r)$, is the normalized eigenvector corresponding to i[th] largest eigenvalue of Q, λ_i denotes the i[th] eigenvalue of Q, $i = 1,\ldots,1+(m-r)$, i.e. $\lambda_1 = \cdots = \lambda_r > \lambda_{r+1} \ge \cdots \ge \lambda_m > 0$.

Note that $\alpha_i^j, i = 1,\ldots,1+(m-r), j = 1,\ldots,m$ are constant coefficients, which satisfy condition $\sum_{j=1}^{r} (\alpha_1^j)^2 + \sum_{\substack{j=r+1 \\ i=1,\ldots,j}}^{m} (\alpha_i^j)^2 = 1$. For notational simplicity coefficients

$\beta_i = \sum_{j=1}^{m} \alpha_i^j, i = 1,\ldots,1+(m-r)$ have also been defined.

Note further that in the special case where $r = 1$, (i.e. Q has a non-repeated maximum eigenvalue), the optimal input matrix is defined as:

$$B = \{\alpha_1^1 V_1, \alpha_1^2 V_1 + \alpha_2^2 V_2, \alpha_1^3 V_1 + \alpha_2^3 V_2 + \alpha_3^3 V_3, \ldots, \alpha_1^m V_1 + \alpha_2^m V_2 + \alpha_3^m V_3 + \cdots + \alpha_m^m V_m\} \tag{7.66}$$

where $V_1, V_2, \ldots V_m$ are the normalized eigenvectors corresponding to the m largest eigenvalues of $Q = \int_0^{t_f} e^{A^* t} e^{At} dt$ respectively. We can now state and prove the following Theorem.

Theorem (7.5): Under the constraint of matrix $B \in \beta$ given in equation (7.65) above, the trace of the controllability Gramian, $W_c(0,t_f)$, is maximized as:

$$\max_{\substack{B_{n\times m}\in\beta \\ \|b_i\|=\frac{1}{\sqrt{m}} \\ i=1,\ldots,m}} \{trace(W_c(0,t_f))\} = \sum_{i=1}^{1+(m-r)} \beta_i \lambda_i \tag{7.67}$$

□

Proof:

It is assumed without loss of generality that the eigenvalues of Q are distinct. Otherwise, the $\underline{b_i}$'s could be selected as orthonormal bases of the corresponding spectral subspace spanned by the clusters of the repeated eigenvalues.

Similar to single input section, first we define the trace of controllability Gramian as:

$$trace\{W_c(0,t_f)\} = trace\{\int_0^{t_f} e^{At}BB^*e^{A^*t}dt\} = trace\left\{B^*\int_0^{t_f} e^{A^*t}e^{At}dtB\right\} \tag{7.68}$$

Again define:

$$Q = \int_0^{t_f} e^{A^*t}e^{At}dt \tag{7.69}$$

Then:

$$trace\{W_c(0,t_f)\} = trace\{B^*QB\} = trace\{\begin{pmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_m^* \end{pmatrix} Q \begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix}\} \tag{7.70}$$

or equivalently:

$$trace\{W_c(0,t_f)\} = trace\{\begin{pmatrix} b_{1_{1\times n}}^* Q_{n\times n} b_{1_{n\times 1}} & \cdots & \times \\ \vdots & \ddots & \vdots \\ \times & \cdots & b_{m_{1\times n}}^* Q_{n\times n} b_{m_{n\times 1}} \end{pmatrix}_{m\times m}\} \tag{7.71}$$

which implies that:

$$trace\{W_c(0,t_f)\} = \sum_{i=1}^{m} b_i^* Q b_i \tag{7.72}$$

Then considering the Rayleigh quotient eigenvalue problem (Golub & Loan, 1983), (Watkins, 1982), (Uper, July 2002), the trace of the controllability Gramian is maximized as follows:

$$\max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F\le 1}} \{trace(W_c(0,t_f))\}$$

$$= \max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F=1}} \left\{ b_1^* b_1 \times \left(\frac{b_1^* Q b_1}{b_1^* b_1} \right) + b_2^* b_2 \times \left(\frac{b_2^* Q b_2}{b_2^* b_2} \right) + \cdots + b_m^* b_m \times \left(\frac{b_m^* Q b_m}{b_m^* b_m} \right) \right\} = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|_F\le 1}} \sum_{i=1}^{m} b_i^* Q b_i \tag{7.73}$$

It follows using Theorem (7.4) that one optimal solution can be obtained by choosing b_1 in the direction of the normalized eigenvector corresponding to the maximum eigenvalue of Q, i.e. V_1, $b_2 = \alpha_1^2 b_1 + \alpha_2^2 V_2$, where V_2 is the normalized eigenvector corresponding to the second largest eigenvalue of Q.

Using a similar procedure, we can define $b_3 = \alpha_1^3 b_1 + \alpha_2^3 b_2 + \alpha_3^3 V_3$, where V_3 is the normalized eigenvector corresponding to the third largest eigenvalue of Q. In general:

$$\begin{aligned} b_1 &= \alpha_1^1 V_1 \\ b_2 &= \alpha_1^2 V_1 + \alpha_2^2 V_2 \\ &\vdots \\ b_m &= \alpha_1^m V_1 + \alpha_2^m V_2 + \alpha_3^m V_3 + \cdots + \alpha_m^m V_m \end{aligned} \tag{7.74}$$

Nothing that;

$$\|B\|_F = 1 \Rightarrow \sum_{\substack{i=1,..,j \\ j=1,..,m}} (\alpha_i^j)^2 = 1 \tag{7.75}$$

and the Cauchy–Schwarz inequality:

$$0 \le \left| \langle b_i, b_j \rangle \right| \le \|b_i\| \|b_j\| \tag{7.76}$$

Show the optimality of (7.65).

□

A geometric interpretation of Theorem (7.5) is given below.

Let b_1 be defined in the direction of the eigenvector V_1. To satisfy the numerical rank constraint on input matrix B and minimize the angle between b_1 and b_2, b_2 must be located on the boundary of the cone shown in Figure (7.3) below. Constraining b_2 to lie outside this cone ensures that the angle between b_1 and b_2 is always greater than the required tolerance ε.

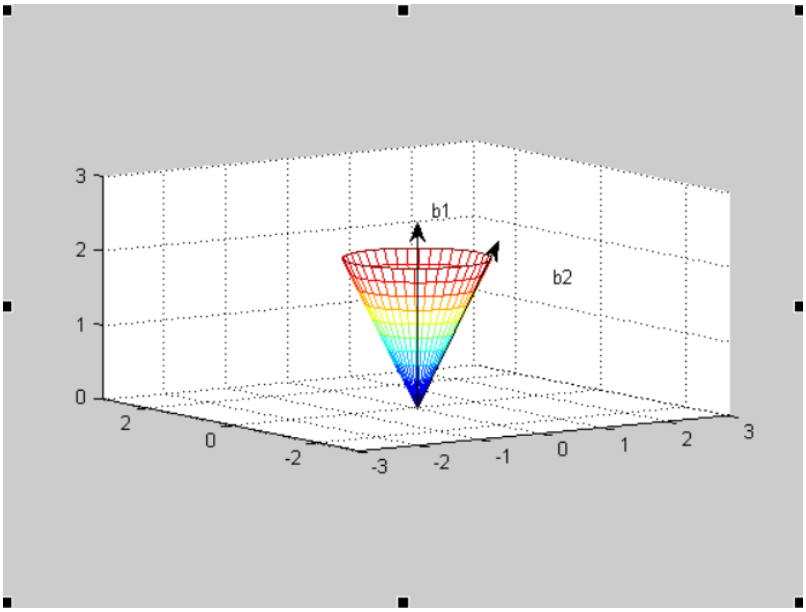


Figure (7.3): b_2 must be located on the boundary of the cone with center b_1

Similarly, with b_1 and b_2 fixed, b_3 is selected to lie on the intersection of the two cones centered at vectors b_1 , b_2 respectively as shown in Figure (7.4).

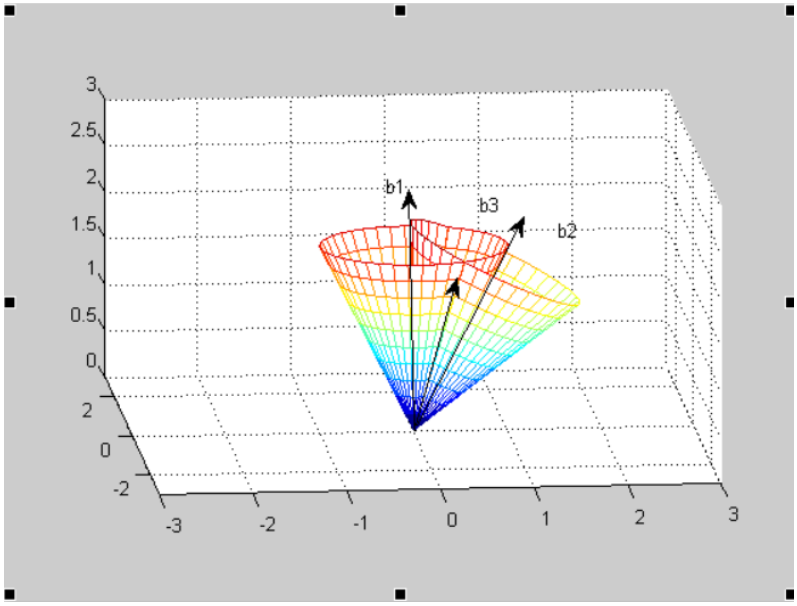


Figure (7.4): Vector b_3 lies on the intersection of the two cones centered at vectors b_1 and b_2

This procedure could be continued to construct all $b_i, i = 1, \ldots, m$.

An alternative approach for defining the sub-optimal solution is based on the following Lemma:

Lemma (7.4): For any matrix $A_{n\times m}$, $m \le n$, $rank(A) = m$, if and only if the minimum singular value of A is greater than zero, i.e. $\sigma_m > 0$.

□

The proof is given in the Appendix. Using the above Lemma, we have:

$$rank(B) = m \Rightarrow \sigma_m(B) \ge \varepsilon > 0 \tag{7.77}$$

Then the requirement that the numerical rank of B is m can be expressed as $\sigma_m(B) \ge \varepsilon$ for a sufficiently small number $\varepsilon > 0$ then the corresponding optimization problem takes the form:

$$\max_{\substack{\alpha_i^j, i=1,\ldots,1+(m-r) \\ j=1,\ldots,m}} \{ \sum_{i=1}^{1+(m-r)} \beta_i \lambda_i \} \tag{7.78}$$

Subject to:

$$0 < \varepsilon \le \sigma_m(B)$$

Here $\alpha_i^j, i = 1,\ldots,1+(m-r),\ j = 1,\ldots,m$ are constant coefficients, which satisfy condition $\sum_{j=1}^{r}(\alpha_1^j)^2 + \sum_{\substack{j=r+1 \\ i=1,\ldots,j}}^{m}(\alpha_i^j)^2 = 1$ and coefficients β_i are defined as:

$$\beta_i = \sum_{j=1}^{m} \alpha_i^j, i = 1,\ldots,1+(m-r)$$

Then, the optimal solution is obtained by selecting:

$$\begin{aligned}
b_1 &= \alpha_1^1 V_1 \\
&\vdots \\
b_r &= \alpha_r^r V_r \\
b_{r+1} &= \sum_{i=1}^{r} \alpha_i^{r+1} V_i + \alpha_{r+1}^{r+1} V_{r+1} \\
b_{r+2} &= \sum_{i=1}^{r} \alpha_i^{r+2} V_i + \alpha_{r+1}^{r+2} V_{r+1} + \alpha_{r+2}^{r+2} V_{r+2} \\
&\vdots \\
b_m &= \sum_{i=1}^{r} \alpha_i^m V_i + \alpha_{r+1}^m V_{r+1} + \alpha_{r+2}^m V_{r+2} + \cdots + \alpha_{(m-r)+1}^m V_{(m-r)+1}
\end{aligned} \tag{7.79}$$

where $V_i, i = 1,\ldots,1+(m-r)$, is the normalized eigenvector corresponding to λ_i, the i[th] largest eigenvalue of Q, i.e. $\lambda_1 = \cdots = \lambda_r > \lambda_{r+1} \ge \cdots \ge \lambda_m > 0$.

An alternative formulation of the optimization problem relies on the determinant of the Gram matrix of B.

Definition (7.3): If $A \in \mathbb{C}^{m\times n}$, then the square matrix $G = A^* A \in \mathbb{C}^{n\times n}$ is known as the Gram matrix of A and $|G|$ as the Gram determinant of A (Mirsky, 2012).

□

Theorem (7.6): (Mirsky, 2012) If G is the Gram matrix of A, then $rank(G) = rank(A)$.

□

The proof is given in the Appendix.

Proposition (7.4): (Peter Lancaster, Miron Tismenetsky, 1985) Let $B = [\underline{b}_1, \underline{b}_2, \dots, \underline{b}_m]$, $\underline{b}_i \in \mathbb{R}^{n\times 1}$, $m \le n$. Then, $rank(B) = m$ if and only if the Gram matrix of B is non-singular, i.e. $|G| \neq 0$.

□

The proof of the above Proposition, which is given by (Peter Lancaster, Miron Tismenetsky, 1985) can be found in the Appendix, however, an easier proof might be using Theorem (7.6), it states that $rank(G) = rank(B)$ then if $rank(B) = m$, it implies that $rank(G) = m$, and G is a symmetric positive definite matrix, which means Gram determinant of B is nonzero, i.e. $\prod_{i=1}^{m} \lambda_i \neq 0$ where λ_i's are the eigenvalues of G. Conversely, if G is non-singular, with regard to Definition (7.3) G is a symmetric positive definite matrix, then $rank(G) = m$ thus using the Theorem above $rank(B) = m$, and the proof of Proposition is completed.

□

Note that if $B \in \mathbb{R}^{n\times m}$, then $G_{ij} = \underline{b}_i^* \underline{b}_j$. Thus, if θ_{ij} is the angle between vectors $\underline{b}_i$ and $\underline{b}_j$ and $\alpha, \beta \in \mathbb{R}$ such that $\alpha \le \theta_{ij} \le \beta$, then it follows from the monotonicity of function $\cos^{-1}(.)$ that:

$$b_i b_j \cos\beta \le G_{ij} \le b_i b_j \cos\alpha \tag{7.80}$$

Which is a pair of linear inequalities on the elements of G. Thus, we can always maximize or minimize angle θ_{ij} by maximizing or minimizing G_{ij} (Stephen P. Boyd, Lieven Vandenberghe, 2004).

Using $|G|$ as a measure f the numerical rank of B, the optimization problem can be formulated as:

$$\max_{\substack{\alpha_i^j, i=1,\dots,1+(m-r) \\ j=1,\dots,m}} \{ \sum_{i=1}^{1+(m-r)} \beta_i \lambda_i \} \tag{7.81}$$

Such that:

$$\beta_i = \sum_{j=1}^{m} \alpha_i^j, i = 1,\ldots,1+(m-r)$$

$$\sum_{j=1}^{r} (\alpha_1^j)^2 + \sum_{\substack{j=r+1 \\ i=1,\ldots,j}}^{m} (\alpha_i^j)^2 = 1 \tag{7.82}$$

$$0 < \varepsilon \le |G|$$

where:

$$\begin{aligned}
b_1 &= \alpha_1^1 V_1 \\
&\vdots \\
b_r &= \alpha_r^r V_r \\
b_{r+1} &= \sum_{i=1}^{r} \alpha_i^{r+1} V_i + \alpha_{r+1}^{r+1} V_{r+1} \\
b_{r+2} &= \sum_{i=1}^{r} \alpha_i^{r+2} V_i + \alpha_{r+1}^{r+2} V_{r+1} + \alpha_{r+2}^{r+2} V_{r+2} \\
&\vdots \\
b_m &= \sum_{i=1}^{r} \alpha_i^m V_i + \alpha_{r+1}^m V_{r+1} + \alpha_{r+2}^m V_{r+2} + \cdots + \alpha_{(m-r)+1}^m V_{(m-r)+1}
\end{aligned} \tag{7.83}$$

To summarize the results of this sub-section, the optimization of the trace of the controllability Gramian $W_c(0,t_f)$ under the constraint that $\|B\|_F \le 1$ and the input matrix $B = [\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_m], \underline{b}_i \in \mathbb{R}^{n\times 1}$ has full (numerical) rank, i.e. $rank(B) = m$, is the solution of one of the suboptimal optimization problems proposed in (7.65), (7.79) or (7.83).

Example (7.2): Consider the problem of multiple input structure selection of a heating furnace (Yadykin, 2011), which can be represented by the following equations:

$$\dot{x} = \begin{bmatrix} -0.5 & 0 \\ 0 & -1 \end{bmatrix} x + Bu$$

$$y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x$$

subject to the constraints: $rank(B) = 2$ and $B = [b_1, b_2]$ $\underline{b}_i \in \mathbb{R}^{2\times 1}$, i = 1,2, $\|B\|_F^2 = \|b_1\|^2 + \|b_2\|^2 = 1$. Assume the terminal time $t_f = 1$.

According to the results stated above the optimal solution can be obtained by minimizing the positive angle between the two input vectors, while choosing the first input vector along the eigenvector corresponding to the maximum eigenvalue of $Q = \int_0^{t_f} e^{A^* t} e^{At} dt$.

In this case:

$$Q = \int_0^1 e^{A^* t} e^{At} dt = \begin{bmatrix} 0.6321 & 0 \\ 0 & 0.4323 \end{bmatrix}$$

and

$$\lambda_1 = 0.6321 \Rightarrow V_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\lambda_2 = 0.4323 \Rightarrow V_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Thus, the optimization problem can be stated as:

$$\max_{\alpha_1^1, \alpha_1^2, \alpha_2^2} \{0.6321((\alpha_1^1)^2 + (\alpha_1^2)^2) + 0.4323(\alpha_2^2)^2\}$$

Such that:

$$\begin{aligned} & b_1 = \alpha_1^1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \alpha_1^1 \neq 0 \\ & b_2 = \alpha_1^2 b_1 + \alpha_2^2 b_2, \alpha_1^2, \alpha_2^2 \neq 0 \\ & 0 < \varepsilon \leq \cos^{-1} \frac{\left|\alpha_1^2\right|}{\sqrt{\alpha_1^2 + \alpha_2^2}} \\ & (\alpha_1^1)^2 + (\alpha_1^2)^2 + (\alpha_2^2)^2 = 1 \end{aligned} \tag{7.84}$$

Taking the minimum angle ε between b_1 and b_2 as $\varepsilon = 0.0367\pi$, then by solving optimization problem in (7.84), the suboptimal solution for B is obtained as:

$$\varepsilon = \cos^{-1}(0.9993) = 0.0367\pi \Rightarrow \alpha_1^1 = 0.3134, \alpha_1^2 = -0.9490, \alpha_2^2 = 0.0349$$

and hence:

$$B = \begin{bmatrix} 0.3134 & -0.9490 \\ 0 & 0.0349 \end{bmatrix}$$

In this case the maximum trace of controllability Gramian $W_c(0,1)$ is:

$$\begin{aligned} & \max\{trace(W_c(0,1)\} = \{0.6321((\alpha_1^1)^2 + (\alpha_1^2)^2) + 0.4323(\alpha_2^2)^2\} \\ & = 0.6321(0.3134^2 + 0.9490^2) + 0.4323(0.0349^2) = 0.6319 \end{aligned}$$

Similarly, if we solve the optimization problem (7.84) for a minimum angle $\varepsilon = 0.0736\pi$, the input matrix will be achieved as:

$$B = \begin{bmatrix} 0.1432 & 0.9870 \\ 0 & -0.0728 \end{bmatrix}$$

for which the value of trace of $W_c(0,1)$ is 0.6311.

In Figures (7.5) the validity of the optimization results is investigated using a brute force method. The values of b_1, b_2 are continuously changed between -1 and 1 (with a step of 0.2), in this case, the maximum value of trace of controllability Gramian $W_c(0,1)$ is obtained as 0.6229, where the suboptimal input matrix is $B = \begin{bmatrix} 0.8 & 0.56 \\ 0.08 & 0.2 \end{bmatrix}$ and the angle between b_1, b_2 is $\varepsilon = 0.2435\pi$. The corresponding results obtained by reducing the step to 0.1 are shown in Figure (7.6). In this case, the maximum value of trace of controllability Gramian $W_c(0,1)$ is obtained as 0.6229 and the suboptimal solution is $B = \begin{bmatrix} -0.64 & -0.76 \\ 0.08 & 0.08 \end{bmatrix}$, here the angle between b_1, b_2 is $\varepsilon = 0.0195\pi$. This confirms that as the accuracy increases b_1 tends to align itself with the eigenvector corresponding to the maximum eigenvalue of Q, $V_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and the angle between b_1, b_2, tends to zero.

Figure (7.7) illustrates changes in the value of trace $\{W_c(0,1)\}$, when b_1 is selected along the eigenvector corresponding to the maximum eigenvalue of Q and b_2 is formed as a linear combination of first and second eigenvectors of Q, i.e. $b_1 = \alpha_1^1 V_1, b_2 = \alpha_1^2 V_1 + \alpha_2^2 V_2$. In the first subplot, the values of trace are demonstrated as a function of $\cos(\varepsilon)$, while the second subplot displays the values of trace for different values of α_1^2, α_2^2. This figure also confirms that we can reach the suboptimal solution by decreasing the angle between input vectors b_1, b_2, i.e. increasing $\cos(\varepsilon)$. The maximum value of trace in Figure (7.7) is equal to 0.6320 and it is reached where input matrix is $B = \begin{bmatrix} 0.7680 & -0.64 \\ 0 & 0.024 \end{bmatrix}$ and $\varepsilon = 0.0374\pi$, comparing this results with the results in plot (7.6) it can be easily inferred that though the optimal angle between input vectors is greater in Figure (7.7), however since in Figure (7.7) b_1 is selected align the eigenvector corresponding to the maximum eigenvalue of Q the optimal value of trace (0.6320) is greater than optimal result (0.6296), shown in Figure (7.6).

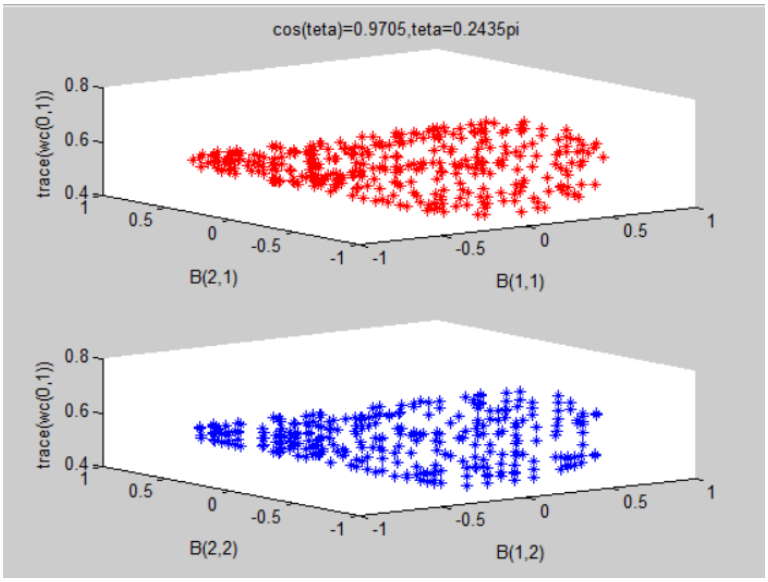


Figure (7.5): $trace\{W_c(0,1)\}$ according to the variable input matrix B in the interval $[-1:0.2:1]^2$

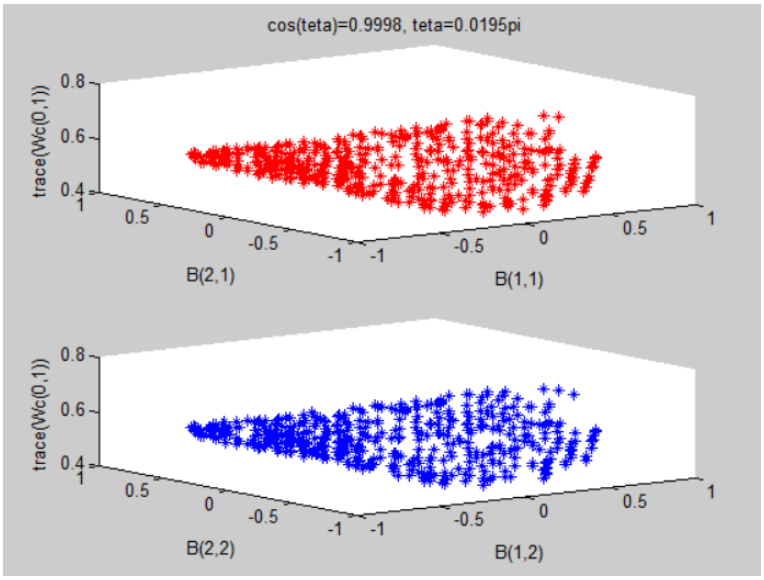


Figure (7.6): $trace\{W_c(0,1)\}$ respect to the variable input matrix B in the interval $[-1:0.1:1]^2$

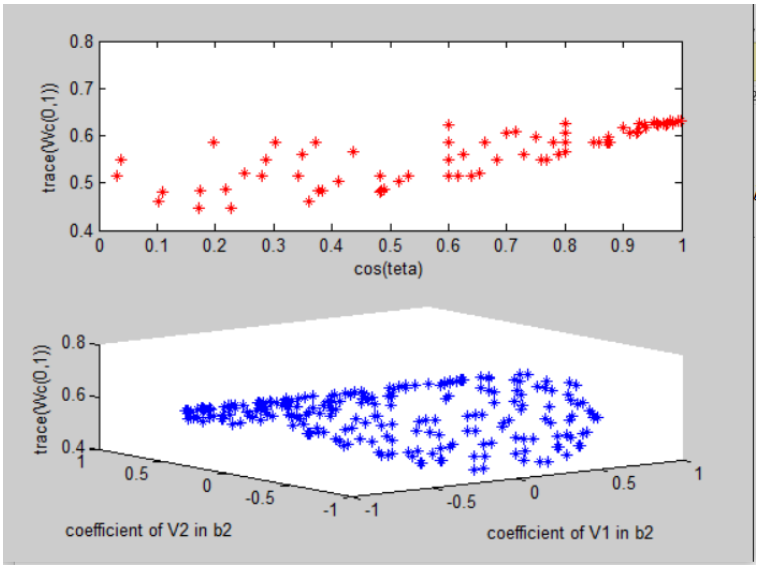


Figure (7.7): $trace\{W_c(0,1)\}$ based on the angle ε between b_1, b_2 (top), and $trace\{W_c(0,1)\}$ as a function of α_1^2, α_2^2 (bottom)

Next, consider optimization method (7.79), in which numerical rank is imposed via a lower bound on the minimum singular value of input matrix B:

$$\max_{\alpha_1^1, \alpha_1^2, \alpha_2^2} \{0.6321((\alpha_1^1)^2 + (\alpha_1^2)^2) + 0.4323(\alpha_2^2)^2\}$$

where:

$$b_1 = \alpha_1^1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \alpha_1^1 \neq 0$$
$$b_2 = \alpha_1^2 b_1 + \alpha_2^2 b_2, \alpha_1^2, \alpha_2^2 \neq 0$$
$$0 < \varepsilon \leq \sigma_m(B), B = [b_1, b_2]$$
$$(\alpha_1^1)^2 + (\alpha_1^2)^2 + (\alpha_2^2)^2 = 1$$

The optimal input matrix for $\varepsilon = 0.017$ is given as $B = \begin{bmatrix} 0.1776 & 0.9794 \\ 0 & 0.0958 \end{bmatrix}$, and the corresponding value of $trace\{W_c(0,1)\}$ is 0.6303. In this case, the positive angle between b_1, b_2 is equal to $\cos^{-1}(0.9952) = 0.098\pi$. By decreasing ε to 0.0011, i.e. $\varepsilon = 0.0011$, changes the suboptimal solution to $B = \begin{bmatrix} 0.0434 & -0.9988 \\ 0 & 0.0243 \end{bmatrix}$, while the optimal value of $trace\{W_c(0,1)\}$ becomes 0.6320. Further the positive angle between input vectors b_1, b_2 reduces to $\cos^{-1}(0.9997) = 0.0245\pi$.

Repeating the procedure with the Gram determinant constraint (optimization (7.83)) the problem becomes:

$$\max_{\alpha_1^1,\alpha_1^2,\alpha_2^2}\{0.6321((\alpha_1^1)^2+(\alpha_1^2)^2)+0.4323(\alpha_2^2)^2\}$$

Such that:

$$b_1=\alpha_1^1\begin{bmatrix}1\\0\end{bmatrix},\alpha_1^1\neq 0$$
$$b_2=\alpha_1^2 b_1+\alpha_2^2 b_2,\alpha_1^2,\alpha_2^2\neq 0$$
$$0<\varepsilon\leq\left|G(B)\right|,B=[b_1,b_2]$$
$$(\alpha_1^1)^2+(\alpha_1^2)^2+(\alpha_2^2)^2=1$$

The solution of this problem for $\varepsilon = 0.0022$ is:

$$B=\begin{bmatrix}-0.8753 & 0.4805\\ 0 & 0.0542\end{bmatrix}$$

$$trace\{W_c(0,1)\}=0.6315$$

The positive angle between input vectors b_1,b_2 is $\cos^{-1}(0.9937)=0.1123\pi$ and the minimum singular value of B is 0.0475.

Setting $\varepsilon = 1.2771\text{e-}04$ gives:

$$B=\begin{bmatrix}-0.3279 & -0.9441\\ 0 & -0.0345\end{bmatrix}$$

$$trace\{W_c(0,1)\}=0.6319$$

The positive angle between input vectors b_1,b_2 reduces to $\cos^{-1}(0.9993)=0.0374\pi$ and the minimum singular value of B decreases to 0.0113.

In conclusion, the example confirms the validity of the results stated in this section, and shows that the optimization problems (7.65), (7.79), (7.83) produce equivalent suboptimal solutions.

7.2.2.b. Optimal Multi-Input Selection in a Stable System

Problem:

For stable LTI MIMO system (7.48) find

$\max_{\substack{B_{n\times m}\in\beta\\ \|B\|_F\leq 1}}\{trace(W_c(0,\infty))\}$, where β is the set of all possible input matrices

In this sub-section, we assume the multi input LTI system (7.48) is stable and $B = [b_1, \ldots, b_m]$, $b_i \in \mathbb{R}^{n\times 1}$ with $\|B\|_F \leq 1$ and $rank(B) = m$ for. In this case, Q can be calculated as the unique solution to the Lyapunov equation:

$$A^*Q + QA + I = 0 \tag{7.85}$$

Using the results described in the previous sub-section the average of minimum input energy can be obtained by solving one of sub-optimization problems stated in (7.65), (7.79), (7.83).

Lemma (7.5): If A is stable, then $trace\{W_c(0,\infty)\} = \left\|(sI - A)^{-1}B\right\|_2$, where $\|.\|_2$ denotes the H_2 norm of the system.

□

Proof:

Let $h(t) = e^{At}B$ be the impulse response of the system:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = x \end{cases} \tag{7.86}$$

Then:

$$\begin{aligned} \|h\|_2^2 &= trace\{\int_0^\infty h(t)h^*(t)dt\},\ h(t) = e^{At}B \\ &= trace\{\int_0^\infty e^{At}BB^*e^{A^*t}dt\} \\ &= trace\{B^*\int_0^\infty e^{A^*t}e^{At}dtB\} \end{aligned} \tag{7.87}$$

Hence:

$$\|H(s)\|_2^2 = \left\|(sI - A)^{-1}B\right\|_2^2 \tag{7.88}$$

So, if we assume all states are observed in output, i.e. $C = I$ then we can write the optimization problem of minimizing the average of input energy as the maximization problem of H_2 norm of the stable LTI system.

□

7.2.2.c. Optimal Input Matrix Selection Subject to the Condition On 2-Norm of the Input Matrix B

Problem:

For the LTI MIMO system (7.48), find

$\delta_0 = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|\leq 1 \\ rank(B)=m}} \{trace(W_c(0,t_f))\}$, where β is the set of all possible input matrices

In this part, we look for an optimal solution to the problem of multi input structure selection corresponding to the constraint $\|B_{n\times m}\| \le 1$, $\|.\|$ denotes the spectral norm (i.e. maximum singular value) and $rank(B) = m$ note that the condition $\|B_{n\times m}\| \le 1$ can be expressed equivalently as $B^*B \le I$.

Theorem (7.7): The optimal matrix B, which maximizes the trace of controllability Gramian subject to $rank(B) = m$ and $B^*B \le I$ is $B = [\underset{\downarrow b_1}{V_1(Q)}, \underset{\downarrow b_2}{V_2(Q)}, ..., \underset{\downarrow b_m}{V_m(Q)}]$ where V_i, $i = 1,\ldots,m$ is the eigenvector corresponding to the i^{th} largest eigenvalue of $Q = \int_0^{t_f} e^{A^*t} e^{At} dt$. The optimal value of δ_0, in this case, is $\sum_{i=1,\ldots,m} \lambda_i$, where λ_i denotes the i^{th} largest eigenvalue of Q.

□

Proof:

The constraint $\|B\| \le 1$ can be written as:

$$B^*B \le I \Rightarrow \begin{bmatrix} b_1^*b_1 & \cdots & b_1^*b_m \\ \vdots & \ddots & \vdots \\ b_m^*b_1 & \cdots & b_m^*b_m \end{bmatrix} \le I_{m\times m} \tag{7.89}$$

Then by considering the trace of the controllability Gramian, and defining $Q = \int_0^{\infty} e^{A^*t} e^{At} dt$ we have:

$$trace(W_c(0,t_f)) = trace(B^*QB) = trace\left(\begin{pmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_m^* \end{pmatrix} Q \begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix} \right) \tag{7.90}$$

$$= \begin{pmatrix} b_{1_{1\times n}}^* Q_{n\times n} b_{1_{n\times 1}} & \cdots & \times \\ \vdots & \ddots & \vdots \\ \times & \cdots & b_{m_{1\times n}}^* Q_{n\times n} b_{m_{n\times 1}} \end{pmatrix}_{m\times m} = \sum_{i=1}^{m} b_i^* Q b_i$$

Thus, the optimization problem is equivalent to:

$$\min_{\substack{B_{n\times m}\in\beta \\ \|B\|\le 1 \\ rank(B)=m}} \{aveE_{\min[0,t_f]}\} = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|\le 1 \\ rank(B)=m}} \{trace(W_c(0,t_f))\} = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|\le 1 \\ rank(B)=m}} \{trace(B^*QB)\} = \max_{\substack{B_{n\times m}\in\beta \\ \|B\|\le 1 \\ rank(B)=m}} \sum_{i=1}^{m} b_i^* Q b_i \tag{7.91}$$

This is equivalent to the Rayleigh quotient form of $\sum_{i=1}^{m} \lambda_i(Q)$.

The results follow by expanding $B_{n\times m} = X_{n\times n} Y_{n\times m}$ where X is the matrix whose columns are the eigenvectors of Q[5], and $Y \in \mathbb{R}^{n\times m}$. Then:

$$trace(B^* QB) = trace(Y^* \Lambda Y) = trace(\Lambda_{n\times n} Y_{n\times m} Y^*_{m\times n}) \tag{7.92}$$

where Λ is the diagonal matrix of the eigenvalues. The condition $B^*B = I$, implies $Y^*Y = I_{m\times m}$ and so $trace\, Y^*Y = trace\, YY^* = trace\, I = m$.

Hence $trace(\Lambda\, YY^*)$ is maximized when $Y = \begin{bmatrix} I \\ 0 \end{bmatrix}$, which concludes the proof.

□

[5] Q is symmetric matrix then it is diagonalizable

7.3. Summary:

In this chapter, we proposed a novel solution to the problem of optimum input structure selection, considering the average of minimum input energy, and we solved the problem for two main cases, single input where input matrix B is a vector, and the case of multi inputs, where input matrix B has the dimension of $n \times m$.

We also considered various interesting cases, in which matrix A has special properties or the input matrix B satisfies different constraints.

The results of this chapter are summarized in tables (7. 1) and (7. 2).

These results can be used as the solution to the problem of actuator localization, where we are looking for the optimum location of actuators, considering the minimum achievable input energy averaged over all possible directions.

$b \in \mathbb{R}^{n\times 1}$	$\max trace\{W_c(0,t_f)\}$	$\arg\max trace\{W_c(0,t_f)\}$
1) $b \in \mathbb{R}^{n\times 1}$ t_f finite	$\lambda_{max}(Q)$	$V_{max}(Q)$
2) $b \in \mathbb{R}^{n\times 1}$ $t_f \to \infty$ A stable	$\lambda_{max}(Q)$ $A^*Q + QA + I = 0$	$V_{max}(Q)$
3) $b \in \mathbb{R}^{n\times 1}$ t_f finite A normal	$\lambda_{max}(Q) = \frac{e^{\lambda_1 t_f} - 1}{\lambda_1}$ $\lambda_1 = \lambda_{max}(A + A^*)$	$V_{max}(A + A^*)$
4) $b \in \mathbb{R}^{n\times 1}$ $t_f \to \infty$ A stable A normal	$\lambda_{\max}(Q) = -\lambda_1^{-1}$ $\lambda_1 = \lambda_{max}(A + A^*) < 0$	$V_{max}(A + A^*)$

Table (7.1): Optimum input vector in single input cases, which minimizes the average of $E_{\min}$

$B \in \mathbb{R}^{n\times m}$	max $trace\{W_c(0,t_f)\}$	arg max $trace\{W_c(0,t_f)\}$
1) $B \in \mathbb{R}^{n\times m}$ $\|B\|_F \leq 1$ $rank(B) = m$ t_f finite	$\sum_{i=1}^{m} \beta_i \lambda_i$ A weighted function of m first largest eigenvalues of Q	b_i is the linear function of $\{V_{max1}(Q), \dots, V_{maxi}(Q)\}$ $V_{maxi}, i = 1, \dots, m$ Eigenvectors corresponding to λ_i^*, ith largest eigenvalue of Q
2) $B \in \mathbb{R}^{n\times m}$ $\|B\|_F \leq 1$ $rank(B) = m$ A stable $t_f \to \infty$	If: $C = I$ $trace\{W_c(0,\infty)\} = \|h(t)\|_2^2$ $= \|((sI - A)^{-1})B\|_2^2$ $A^*Q + QA + I = 0$	b_i is the linear function of $\{V_{max1}(Q), \dots, V_{maxi}(Q)\}$ $V_{maxi}, i = 1, \dots, m$ Eigenvectors corresponding to λ_i^*, ith largest eigenvalue of Q
3) $B \in \mathbb{R}^{n\times m}$ $\|B_{n\times m}\| \leq 1$ $rank(B) = m$ t_f finite	$\sum_{i=1}^{m} \lambda_i^*(Q)$ $\lambda_i^*$are the m largest eigenvalues of Q	$B = [V_{max1}(Q), \dots, V_{maxm}(Q)]$ $V_{maxi}, i = 1, \dots, m$ Eigenvectors corresponding to λ_i^*, ith largest eigenvalue of Q

Table (7.2): Optimum input matrix in multi- input cases, which minimizes the average of $E_{\min}$

Chapter 8: Kamineh Algorithm Optimization

8.1. Introduction

In this chapter, a novel metaheuristics optimization algorithm namely Kamineh Algorithm (KA) is proposed. The KA starts from a random initial population and requires them to move towards the best solution using Logistic equation like functions.

A metaheuristic is formally defined as an iterative generation process which guides a subordinate heuristic by combining intelligently different concepts for exploring and exploiting the search space, learning strategies are used to structure information in order to find efficiently near-optimal solutions (Osman & Laporte, 1996).

Metaheuristics have admirable ability to avoid trapping in local optima, this important property is due to the stochastic nature of metaheuristic algorithms, which make them suitable for optimizing challenging real problems where the search space is usually unknown and very complex with a massive number of local optima.

Different metaheuristic algorithms exist in literature, some more popular ones are Genetic algorithms (GAs), Simulated annealing (SA), Differential evolution (DE), Ant colony optimization (ACO), Bat algorithms (BA), Particle swarm optimization (PSO), Tabu search (TS), Gravitational search algorithm (GSA), and Firefly algorithms (FA). Some algorithms have also

been proposed integrating or improving the current techniques, to make them more efficient and applicable.

Two main components of any metaheuristic algorithm are intensification and diversification, or exploitation and exploration. Exploration is the ability of the algorithm to search for new individuals far from the current solution in the search space and to explore the search space on a global scale. Exploitation is to search the surrounding search area nearby the current solution, something like local search. Exploitation ensures that the solutions will converge to the optimality, whereas the exploration via randomization avoids the solutions being trapped at local optima and, at the same time, increases the diversity of the solutions.
Finding an algorithm that can handle both (exploitation and exploration) is still challenging because they are two different contrast objectives.
The No Free Lunch (NFL) Theorem of (Wolpert & Macready, 1997) states that "There is no strategy or algorithm that generally behaves better than another for the entire set of possible problems". (Ho & Pepyne, 2002) also expressed that: "Universal optimizers are impossible". This motivated us to develop a new effective metaheuristic algorithm called KA (Kamineh Algorithm), using the mathematical logistic equation like functions, which shows superior performance in terms of exploration, exploitation, and convergence comparable with other state-of-the-art methods. It makes KA a powerful optimization algorithm in solving complex multi-parameter optimization problems, such as the actuators (sensors) placement. In this thesis, the optimization criteria for the selection of actuators are based on the eigenvalues of controllability Gramian matrix.

The structure of this chapter is as follow:

Section 8.2 outlines the proposed KA optimization method. The results and discussion of benchmark functions and some case studies are presented in Sections 8.3. Finally, Section 8.4 concludes the work.

8.2. Kamineh Optimization Algorithm (KA)

In this section, the inspiration of the proposed method is first discussed. Then, the mathematical model is provided.

8.2.1. Inspiration

The logistic equation (also called Verhulst model or logistic growth curve) is a model of population growth first published by (Verhulst, 1845), (Verhulst, 1847). The model is continuous and when is translated into mathematics, results to the differential equation below:

$$\frac{dP(t)}{dt} = rP(t)\left[1-\frac{P(t)}{K}\right], P(0) = P_0 \tag{8.1}$$

where t denotes time, P_0 is the initial population and r, K are constants associated with the growth rate and the carrying capacity of the population.

However, in many applications, when modelling a realistic problem, one may decide to describe the continuous-time logistic equations in terms of discrete-time model. Thus, the differential equation (8.1) can be formulated instead, as an initial value problem of a difference equation below:

$$P_{n+1} = rP_n\left(1 - P_n\right) \tag{8.2}$$

where r is the Malthusian parameter (rate of maximum population growth). The above statement is non-linear, since it involves a term P_n^2. Because of this non-linearity, the equation has remarkable non-trivial properties, but cannot be solved analytically. The discrete version of the logistic equation has the fascinating properties and can be used in many real-world examples such as:

- The concentration of oxygen in the lungs after the ith breath
- The concentration in the blood of a drug after the ith dose
- The size of a population of mosquitoes in year n
- The number of cells in a bacterial culture on day i

The steady state P_{ss} of difference equation (8.2) is defined by:

$$P_{n+1} = P_n = P_{ss} \tag{8.3}$$

which returns:

$$P_{ss} = rP_{ss}\left(1 - P_{ss}\right) \Rightarrow \begin{cases} P_{ss1} = 0 \\ P_{ss2} = 1 - \dfrac{1}{r} \end{cases} \tag{8.4}$$

Now consider a small perturbation ε that moves the system out of its steady state:

$$\begin{aligned} &P_n = P_{ss} + \varepsilon_n \Rightarrow \\ &P_{n+1} = P_{ss} + \varepsilon_{n+1} \Rightarrow \\ &\varepsilon_{n+1} = f\left(P_{ss} + \varepsilon_n\right) - P_{ss} \end{aligned} \tag{8.5}$$

Since $\varepsilon_n \ll P_{ss}$ we can use the Taylor expansion around P_{ss} :

$$f\left(P_{ss} + \varepsilon_n\right) = f\left(P_{ss}\right) + \left(\frac{df}{dP}\right)_{P=P_{ss}} P_n + O\left(P_n^2\right) \tag{8.6}$$

Then equation (8.5) can be approximated as:

$$\varepsilon_{n+1} \simeq \left(\frac{df}{dP}\right)_{P=P_{ss}} P_n \tag{8.7}$$

Clearly, if $\left|\left(\frac{df}{dP}\right)_{P=P_{ss}}\right| < 1$, the steady state is stable and the perturbation tends to zero as n increases, otherwise the steady state is unstable. In the case of the logistic equation (8.2), for the first steady state, i.e. $P_{ss1} = 0$ we have:

$$\left(\frac{df}{dP}\right)_{P=0} = (r - 2rP)_{P=0} = r \tag{8.8}$$

Thus, the steady state $P_{ss1} = 0$ is stable when $0 < r < 1$.

Similarly, for the second steady state $P_{ss2} = 1 - \frac{1}{r}$, we have:

$$\left(\frac{df}{dP}\right)_{P=1-\frac{1}{r}} = (r - 2rP)_{P=1-\frac{1}{r}} = 2 - r \tag{8.9}$$

Thus, the steady state $P_{ss2} = 1 - \frac{1}{r}$ is stable when $1 < r < 3$.

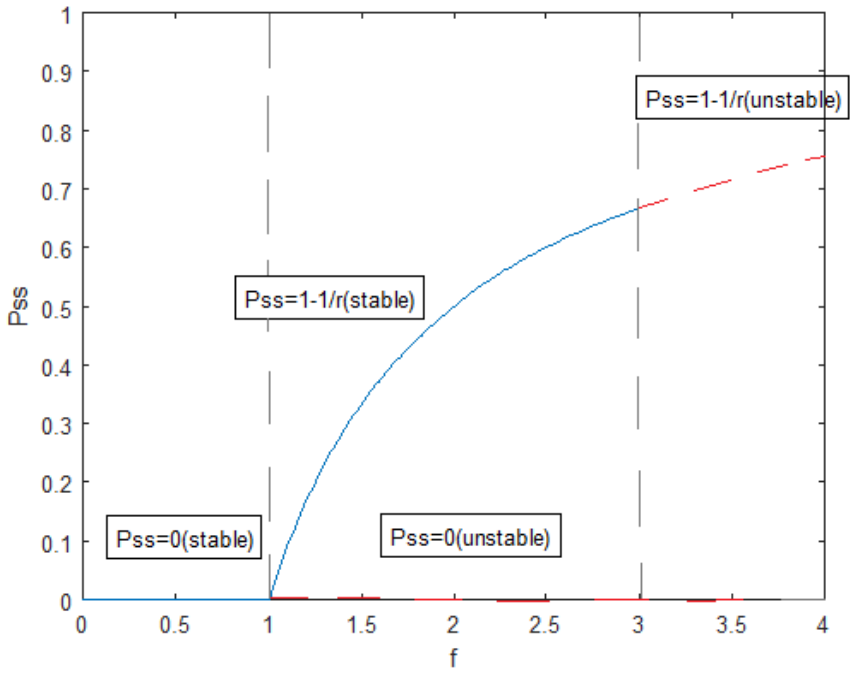
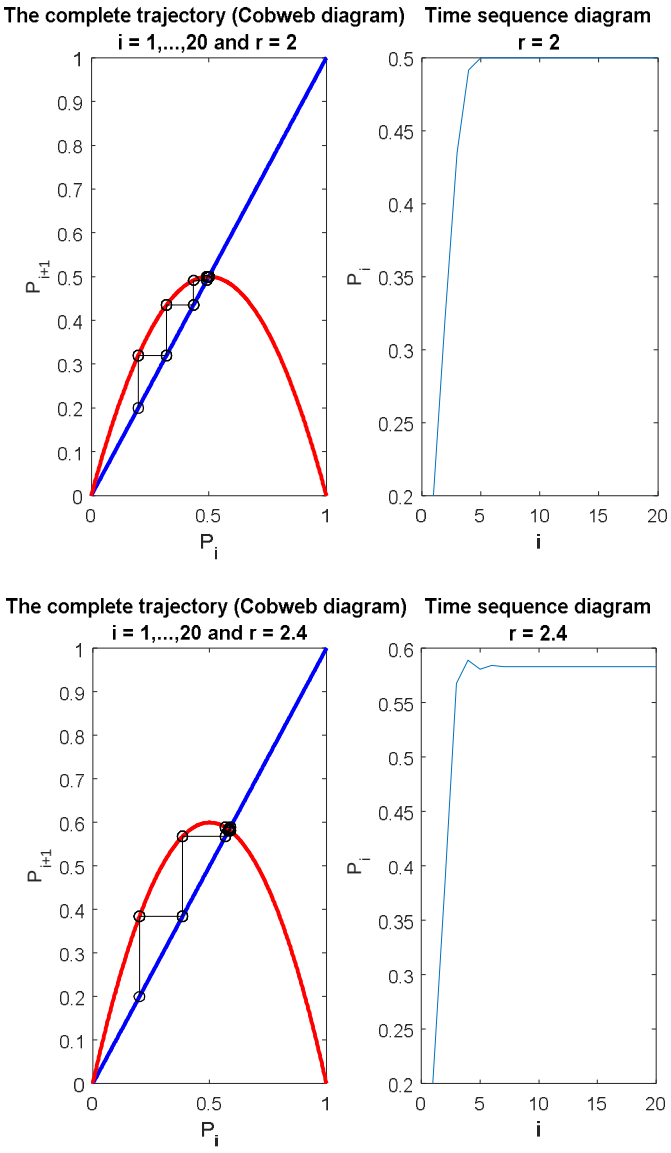


Figure (8.1): Steady states of logistic equation (8.2)

In Figure (8.2), several time sequences, corresponding to different values of $0 < r$ are shown. When *r* increases, the steepness of the parabola increases as well, which makes the slope of this tangent steeper, so that eventually the stability condition is violated. The steady state then becomes unstable and the oscillations appear. (May, 1976) proves that as *r* increases just beyond $r = 3$, stable oscillations of period 2 (sometimes called two-point cycles) appear. A stable oscillation means a periodic behavior that is maintained despite small perturbations

and period 2 implies that successive generations alternate between two fixed values of P. When r further increases, the periodic solution becomes unstable and then higher period oscillations appear. When all the cycles become unstable, chaos is observed.

The schematic representation shown in Figure (8.3) illustrates that as r increases, the system (8.2) undergoes successively cycles of period $2, 4, 8, \ldots$ and ultimately leads to a chaotic behavior.



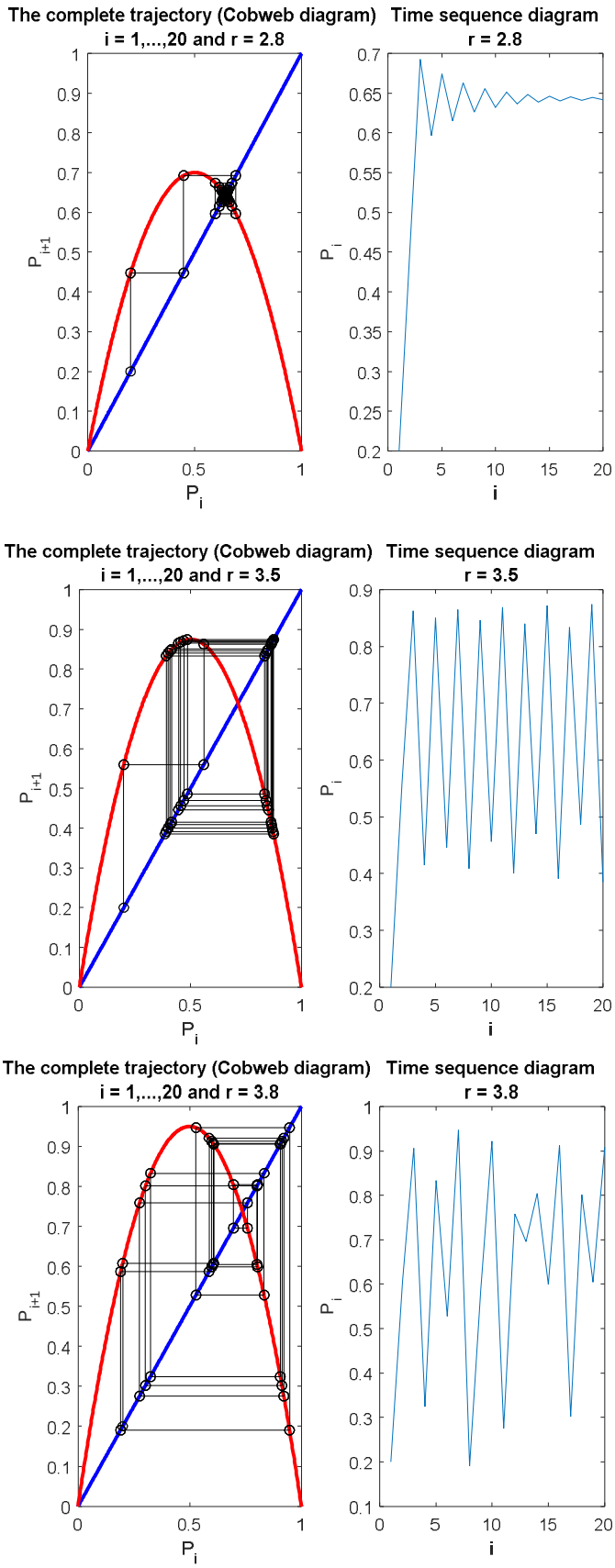


Figure (8.2): Graphical resolution for various values of r

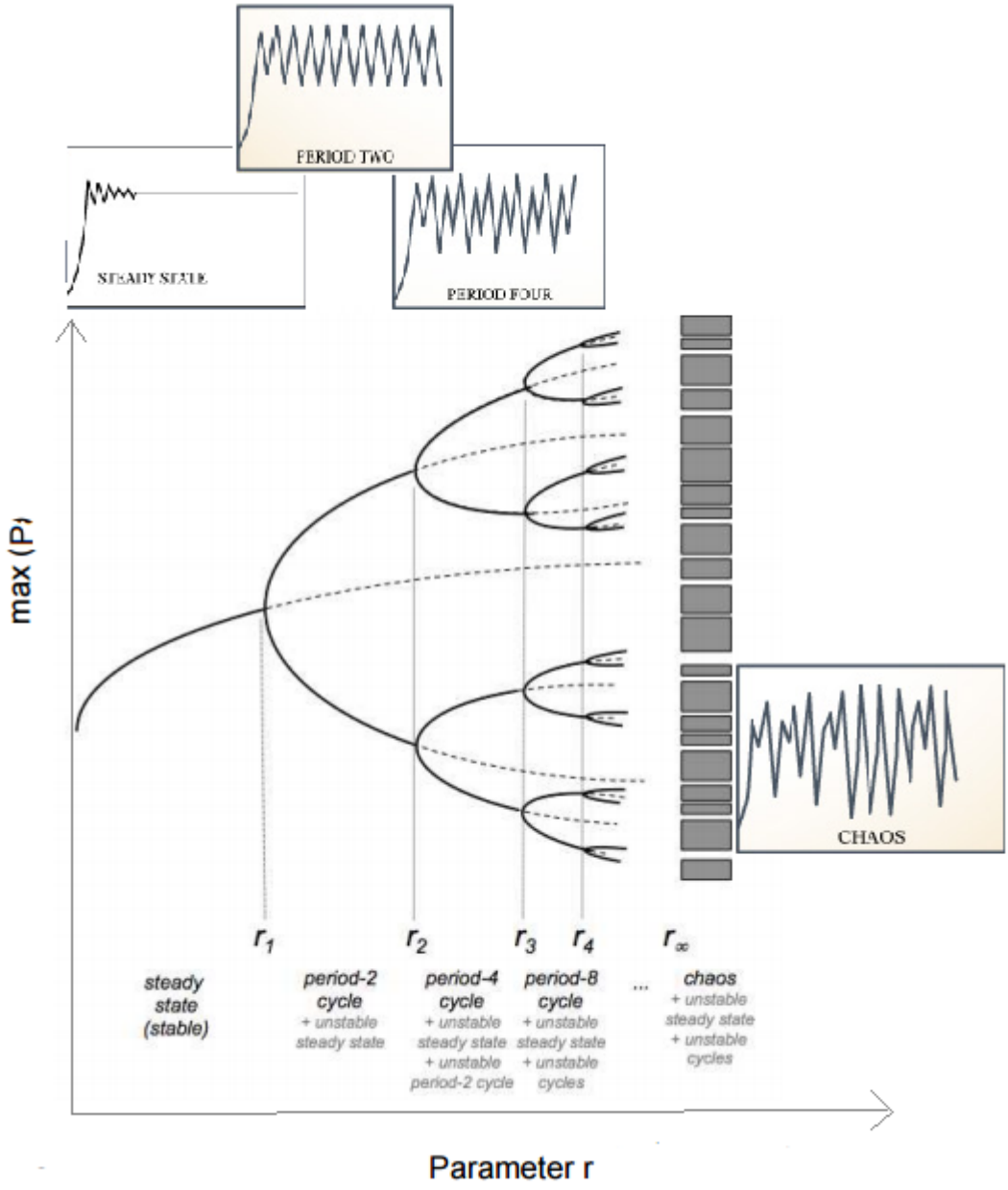


Figure (8.3): Schematic representation of the bifurcation diagram

In KA optimization method, an agent represents a diversity of species with different populations. Over the iterations, the species reach new populations different from the previous ones. The change of each species' population is correlated with the fitness of the agent and can affect all its other species' populations as well. The species on the same agent share their individual information with others to reach their destination population, which leads to the optimum fitness value.

A function of the distance between the current agent's fitness and the best existing objective value updates the rate of maximum population growth of the species in each iteration. Then the closer the objective value is to the current optimum fitness, the smaller the Malthusian parameter r is and the species are closer to the stable steady state region. Moreover, as r increases the chaotic growth of the species guarantees that KA algorithm intrinsically benefits

from high exploration and local optima avoidance. However, in the exploitation phase when r is small there are gradual changes due to the stable behavior and the cyclic oscillations of the logistic equations and random variations are considerably less than those in the exploration phase. This enhances the convergence capability of KA method towards the optimal solution.

In the next subsection, KA optimization method is proposed based on the mathematical model.

8.2.2. Mathematical Model and Algorithm

As shown in Figure (8.4), KA algorithm starts with randomly generated populations of species for each candidate/agent called individual and the fitness function is evaluated for each candidate.

We represent the set of current populations in a matrix as:

$$X^T = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} = \begin{bmatrix} \bar{X}_1 & \bar{X}_2 & \dots & \bar{X}_n \end{bmatrix} \tag{8.10}$$

where n is the number of agents and m is the number of species (i.e. dimension). For all agents $\bar{X}_i, i = 1, \dots, n$, we also assume a vector corresponding to the fitness values as:

$$f_X = \begin{bmatrix} f_{\bar{X}_1} & f_{\bar{X}_2} & \dots & f_{\bar{X}_n} \end{bmatrix} \tag{8.11}$$

Each agent passes through the fitness function and the obtained output updates the corresponding fitness value of the associated agent in the matrix f_X.

We also assume that there is an array for storing the optimum populations for the best agent $X_{best} = \bar{X}_i, i = 1, \dots, n$ and the corresponding fitness value $f_{\bar{X}_i}$.

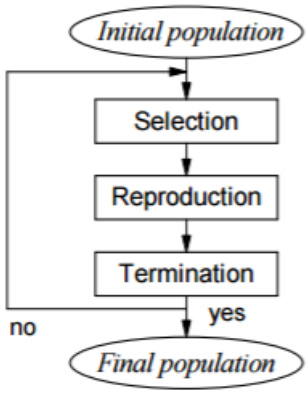


Figure (8.4): Schematic representation of the KA optimization method

The KA algorithm approximates the global solution of the optimization problem based on the triplet:

$$\left(\text{Initialze}, \text{Update}, \text{Terminate}\right) = \left(I, U, T\right)$$

I is a function that generates some random populations of the species for individuals and corresponding fitness values:

$$I : \varnothing \rightarrow \left\{X,\ f_X\right\} \tag{8.12}$$

Any random distribution can be used in I function. What we used in KA algorithm is as follows:

$$\begin{aligned} &X(:, j) = rand(n,1) \times (ub_j - lb_j) + lb_j, \quad j = 1, \ldots, m \\ &f_X = Obj\ (X) \end{aligned} \tag{8.13}$$

where $ub_j, lb_j,\ j = 1, \ldots, m$ denote the upper-bound and lower-bound of the j[th] species' population (i.e. the j[th] decision variable) respectively. Obj defines the objective function, which is used for the optimization problem. $X(:, j)$ is a n-dimensional vector, which consists the j[th] species' population of all individuals and $rand(n,1)$ is a n-dimensional random vector.

The U function, which is the main function, moves the agents around the search space. This function receives the populations x_{ij} where $i = 1, \ldots, n,\ j = 1, \ldots, m$ and returns their updated ones eventually:

$$U : x_{ij}(t) \rightarrow x_{ij}(t+1) \tag{8.14}$$

The T function returns $Stop$ if the termination criterion is satisfied and $Continue$ if the termination criterion is not satisfied. The KA algorithm terminates the optimization process when the iteration counter goes higher than the maximum number of iterations by default. However, any other termination condition can also be considered easily wherever it is needed, e.g. maximum number of function evaluation or the accuracy of the global optimum obtained.

$$T : X \rightarrow \left\{Stop, Continue\right\} \tag{8.15}$$

The U function iteratively run until the T function returns $Stop$ .With defined three-tuple $\left(I, U, T\right)$ the general framework of the KA algorithm is defined as:

$X = I(\ )$
While $T = Continue$
for $i = 1, \ldots, n$
Update $\bar{X}_{best}$, $f_{\bar{X}_{best}}$
Update species' populations of i-th individual $x_{ij} = U(x_{ij}, f_{\bar{X}_i}), \quad j = 1, \ldots, m$
end
end

(8.16)

As mentioned in the previous subsection the inspiration of this algorithm is the fascinating properties of the discrete version of the logistic equation. In order to mathematically model this behavior, we update each species with respect to its current value of population via the equation below:

$$x_{ij}(t+1) = \begin{cases} r x_{ij}(t)(1 - x_{ij}(t)) & O(t) = 1 \\ rand \times x_{best,j}(t) & O(t) = 0 \end{cases} \tag{8.17}$$

where x_{ij} indicates the j[th] species' population of the individual i and $x_{best,j}$ denotes the j[th] species' population of the current optimal solution. $rand$ denotes a random number generated with uniform distribution over the interval of [0,1]. $O(t)$ is a stochastic function defined as:

$$O(t) = \begin{cases} 1 & \text{if } rand \geq 0.5 \\ 0 & \text{if } rand < 0.5 \end{cases} \tag{8.18}$$

In equation (8.17), *r* is the rate of maximum population growth and is defined as:

$$r = \begin{cases} 1 < \beta_1 \leq 2 & S(f_{\bar{X}_i}, f_{\bar{X}_{best}}) \leq \alpha_1 \\ 2 < \beta_2 \leq 3.4 & \alpha_1 < S(f_{\bar{X}_i}, f_{\bar{X}_{best}}) \leq \alpha_2 \\ 3.4 < \beta_3 \leq 4 & \alpha_2 < S(f_{\bar{X}_i}, f_{\bar{X}_{best}}) \end{cases} \tag{8.19}$$

where $f_{\bar{X}_i}$ is the objective value of the associated agent and $f_{\bar{X}_{best}}$ denotes the fitness value of the current optimal solution. $S(f_{\bar{X}_i}, f_{\bar{X}_{best}})$ dictates a logarithmic function of the difference between $f_{\bar{X}_i}$ and $f_{\bar{X}_{best}}$:

$$S(f_{\bar{X}_i}, f_{\bar{X}_{best}}) = \log\left(1 + \left| f_{\bar{X}_i} - f_{\bar{X}_{best}} \right|\right) \times rand \tag{8.20}$$

The *rand* parameter in (8.20) brings a random weight for the logarithmic function, i.e. $\log\left(1 + \left| f_{\bar{X}_i} - f_{\bar{X}_{best}} \right|\right)$ in order to stochastically determine the effect of the difference between the current solution and the best one in defining the behavior of the logistic equations, which are used to update the current populations via (8.17).

In KA method, we employ two user-controlled parameters to guide the searching behavior:

- $\beta_i, i = 1, 2, 3$: This parameter defines the rate of maximum population growth and determines the population's behavior.
- $0 < \alpha_1 < \alpha_2 < 1$: defines random values in (0,1). This parameter controls the probability of the species' population changing rate and assists KA method to show a more random behavior throughout optimization through a balanced exploration and exploitation.

Choosing proper parameters of KA for numerical examples and real-world optimization problems can be done through different schemes:

- The trial-and-error scheme: May reveal the best possible performance over the parameter space at the expense of high computational cost. In real-world optimization problems, evaluating the fitness function may take a long time, much longer than evaluating our benchmark functions, then it might be impractical to use this scheme.
- Fixed parameter schemes: As the optimization benchmark set is generally divided into three categories (i.e., unimodal, high-dimensional multimodal, and low-dimensional multimodal), we try to determine a parameter value combination generally suitable for each category. This combination remains constant throughout the whole search (Yao, et al., 1999), (Lam, et al., 2012), (Price, et al., 2005).
- Deterministic parameter schemes: Change the parameter values throughout the search using some pre-defined rules (Lam, et al., 2012), (Chen, et al., 2012).
- Adaptive parameter schemes: Change the parameter values by adaptively learning the impact of changing parameters on the searching performance throughout the search (Qi, et al., 2009). Some schemes encode the parameters into the solution and evolve the parameters together with the population (Vrugt, et al., 2009).

Here, we use the deterministic scheme to choose the parameters, which can consistently lead to satisfactory results on a wide range of different kinds of optimization problems. According to (Bergh & Engelbrecht, 2006), there should be abrupt changes in the movement of search agents over the initial steps of optimization. This assists a meta-heuristic to explore the search space extensively. Therefore, in the KA algorithm, the first step is to combine the random solutions in the set of solutions abruptly with a high rate of randomness to find the promising regions of the search space. Hence, at the beginning of the optimization, the appropriate values of the parameters are selected, which lead to a higher rate of maximum population growth, i.e. r . However, according to (Bergh & Engelbrecht, 2006) the changes should be reduced to emphasize exploitation at the end of optimization. Therefore, after finding the promising regions of the search space, the KA algorithm converges to the global optimum by gradual changes in the random solutions. The random variations are considerably less than those in the exploration phase. Then, the parameters change adaptively such that r ,i.e. the rate of maximum population growth, decreases over the course of iterations.
Figure (8.5) shows that how the parameters of the KA algorithm change adaptively through the iterations. The following formulas are utilized in this regard:

$$\beta_1=2\left(1-\frac{t}{T}\right),\ \beta_2=3.4\left(1-\frac{t}{T}\right),\ \beta_3=4\left(1-\frac{t}{T}\right)$$
$$\alpha_1=\min\left(\left(1-\frac{t}{T}\right),\alpha_2-0.3\right),\quad \alpha_2=\min\left(\left(1-\frac{t}{T}\right)+0.3,0.8\right) \tag{8.20}$$

where t is the current iteration and T is the maximum number of iterations.

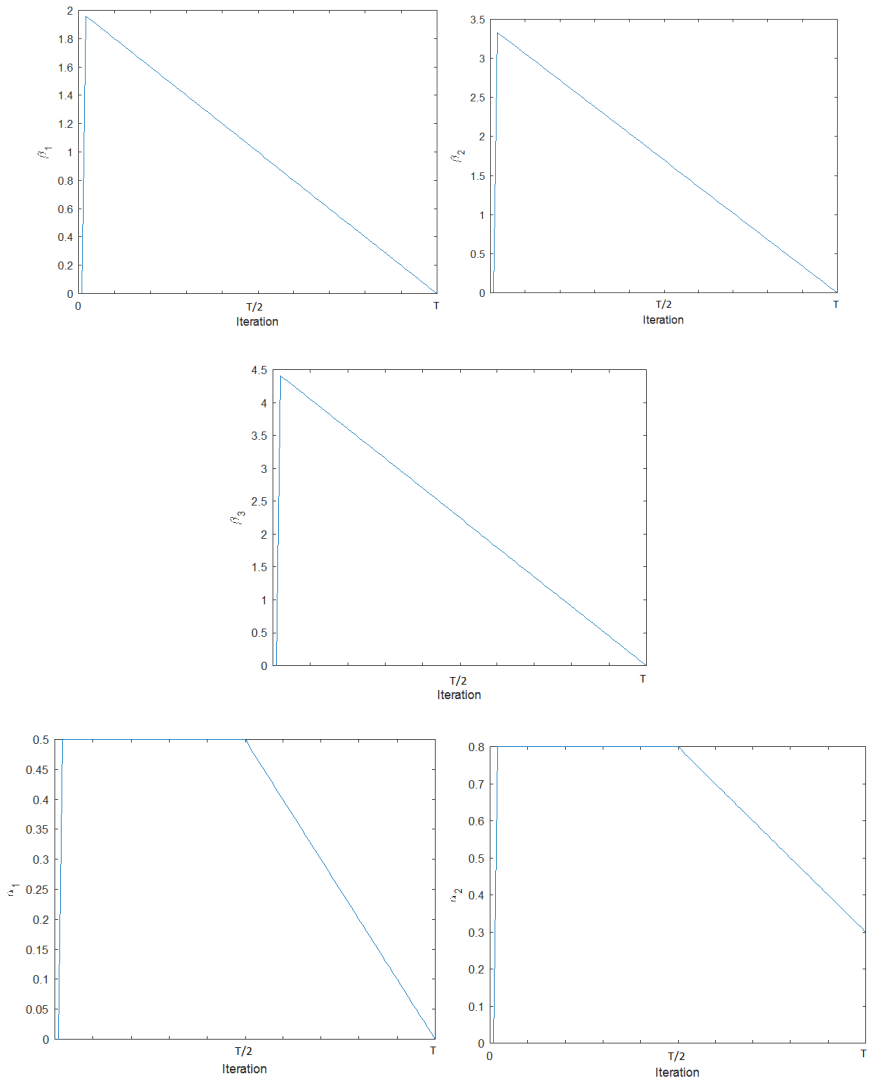


Figure (8.5): Change of parameters over the course of iterations.

Figure (8.6) illustrates how the parameters in equation (8.20) change the behavior of the logistic equation and decrease the range of the movement of the solutions over the iterations. It may be inferred from figure (8.6) that the KA algorithm explores different regions of the search space with a higher degree of randomness via increasing the range of the chaotic behavior of the logistic equation in (8.17), when the optimization starts, so this can assist resolving local optima stagnations. Then the algorithm gradually increases the probability of

the steady state behavior of the logistic equation through the iterations to emphasize more exploitation around the most promising regions of the search space. Then, using the adaptive parameters in the proposed method leads to a smooth transition from exploration to exploitation during optimization process, which guarantees the improved accuracy of local search over the iterations and the convergence of the proposed algorithm. In Figure (8.6), the red point shows a search agent, the blue star is the best solution found so far and the three colored circles around the red point describe the search spaces associated with 3 different logistic equation regions:

- Steady state
- Cyclic
- Chaos

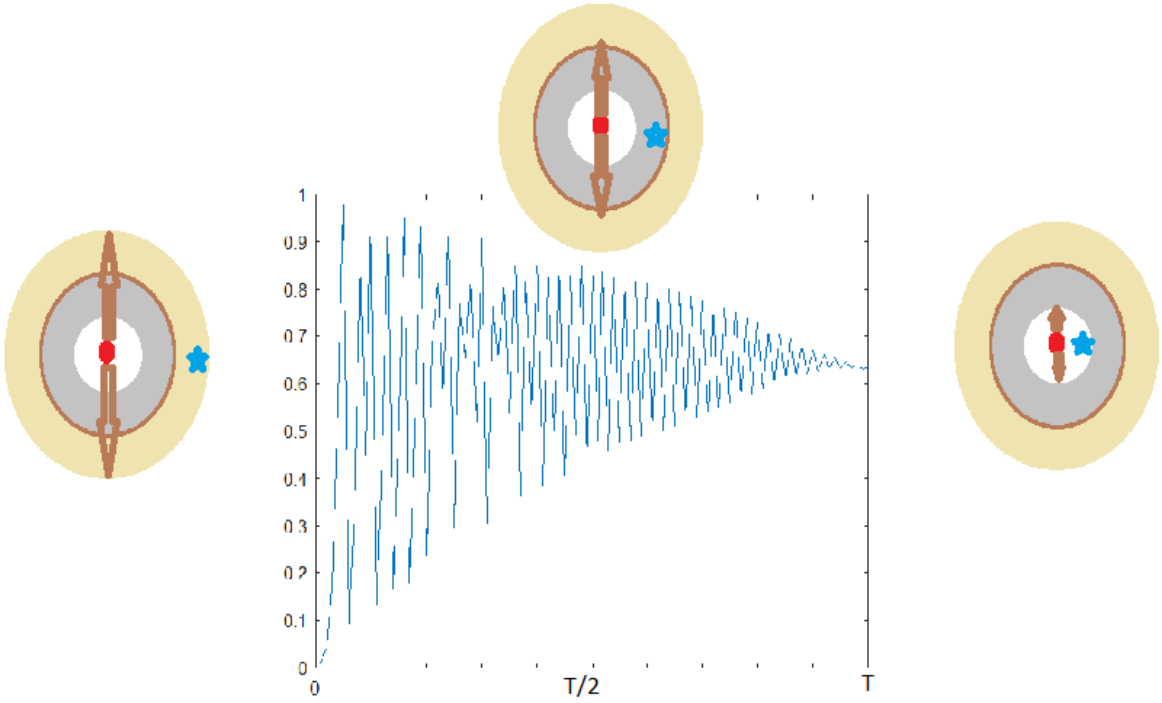


Figure (8.6): The movement of a solution around the destination via change of the parameters over the iterations

After all, the pseudo code of the KA algorithm in (8.16) can be presented as follows:

Step 1: Initialization

Step 2: Iteration

Do

for each individual repeat

Update $\overline{X}_{best}$, $f_{\overline{X}_{best}}$

Update the parameters β_k, α_l, $k=1,2,3,\ l=1,2$ (8.21)

Update the rate of maximum population growth, i.e. r

Update the population of the species $x_{ij}=U(x_{ij}, f_{\overline{X}_i});\ j=1,\ldots m$

end

While stopping criteria

Step 3: Return the best solution obtained

The general steps defined in (8.21) shows that the KA algorithm starts the optimization process with creating a set of random solutions. The algorithm then saves the best solutions obtained so far, assigns it as the destination point, and updates other solutions with respect to it. Mean-while, the parameters and the rate of maximum population growth, i.e. r , are updated to emphasize more exploration of the search space at the beginning and higher exploitation of the search space as the iteration increases. The KA algorithm terminates the optimization process when the termination criterion is satisfied and the best solution obtained so far is returned as the global optimum.

The next section employs a wide range of test problems to investigate, analyze and confirm the effectiveness of the KA algorithm.

8.3. Experimental Results and Discussion

For any new optimization algorithm, it is essential to validate its performance and compare it with other existing algorithms over a good set of test functions.

In this section, the proposed KA algorithm is tested against different sets of benchmark functions and compared with five current popular metaheuristic methods introduced over a long range of time starting from 1975 onwards: Genetic Algorithms (GA) (Holland, 1975), Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995), Firefly Algorithm (FA) (Yang, 2008), Gravitational Search Algorithm (GSA) (Rashedi & Nezamabadi-pour, 2009), and Collective Animal Behavior Algorithm (CAB) (Cuevas, et al., 2012).

In this work, the maximum number of iterations and the size of population for the KA algorithm are set to 100 and 80 respectively. The results are averaged over 30 independent runs and the mean value and standard deviation of the best solutions are obtained.

The performance of optimization algorithms were compared considering three criteria: 1) the final results (i.e. average value of cost functions and the value of standard deviations), 2) the convergence rate and 3) the results produced by non-parametric statistical Wilcoxon rank sum

test (Wilcoxon, 1945), (Garc´ıa, et al., 2009). We can also use the following efficiency measure to evaluate the performance of the algorithm (de-los-Cobos-Silva, et al., 2015):

$$\text{efficiency} = \begin{cases} 1 - \left| \dfrac{f_{X_{best}} - f_{\min}}{f_{\min}} \right| & \text{if } f_{\min} \neq 0 \\ 1 - \left| f_{X_{best}} \right| & \text{if } f_{\min} = 0 \end{cases} \tag{8.22}$$

where $f_{X_{best}}$ is the best value of fitness found by the algorithm and $f_{\min}$ is the optimum value of the cost function. (Note: considering the accuracy of Matlab, when the efficiency is greater than 0.999999, it is rounded to 1.)

8.3.1. Benchmark Functions

Test functions are important to evaluate reliability, efficiency and validation of optimization algorithms. There have been many test or benchmark functions reported in the literature; however, there is no standard list or set of benchmark functions. The number of test functions in most papers varied from a few to about two-dozen, and ideally, the test functions used should be diverse and unbiased to cover a wide variety of problems, such as unimodal, multimodal, separable, non-separable and multi-dimensional problems to confidently make sure that the achieved results of our algorithm are not by chance. (Jamil & Yang, 2013)

A complete detailed description of optimization test functions can be found in (Jamil & Yang, 2013), (Ali, et al., 2005), (Adorio & Dilman, 2005), (Botev, et al., 2004), (Yang, 2010).

In this work, we are using a comprehensive set of 23 functions, which are used in most literature (Cuevas, et al., 2012), (Yao & Liu, 1996), (Gaviano, et al., 2003), (Ali, et al., 2005) to test the performance of the proposed approach. Such functions are classified into three different categories: unimodal test functions (F1-F7), which have only one local optimum and allow to evaluate the exploitation capability of the investigated optimization algorithm, multimodal test functions with variable dimensions (F8-F13) and multimodal fixed-low-dimensional test functions (F14-F23), which both have more than one local optimum. These functions are used to test the capability of the KA algorithm to escape from local minima, and to investigate the exploration ability of the algorithm. The test functions are given in Tables (8.1), (8.2) and (8.3), where *V-no* is the dimension of the function, *Range* is the boundary of the function's search space and f_{min} shows the optimum value of the function. Figures (8.7), (8.8) and (8.9) represent two-dimensional form of some benchmark functions out of each category.

Function	V_no	Range	f_{min}
$F_1(x) = \sum_{i=1}^{n} x_i^2$	30	[−100,100]	0
$F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30	[−10,10]	0
$F_3(x) = \sum_{i=1}^{n} (\sum_{j-1}^{i} x_j)^2$	30	[−100,100]	0
$F_4(x) = \max_i\{\lvert x_i \rvert,\ 1 \le i \le n\}$	30	[−100,100]	0
$F_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[−30,30]	0
$F_6(x) = \sum_{i=1}^{n} ([x_i + 0.5])^2$	30	[−100,100]	0
$F_7(x) = \sum_{i=1}^{n} ix_i^4 + random[0, 1)$	30	[−1.28,1.28]	0

Table (8.1): Unimodal test functions

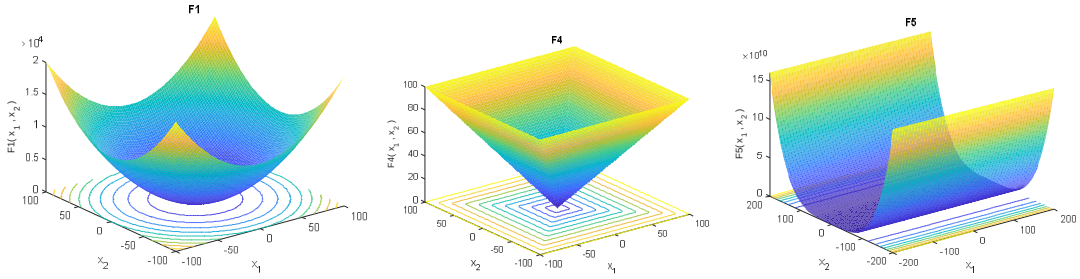


Figure (8.7): Typical two-dimensional form of unimodal benchmark functions

Function	V_no	Range	f_{min}
$F_8(x) = \sum_{i=1}^{n} -x_i \sin(\sqrt{\lvert x_i \rvert})$	30	[-500,500]^n	-12569.5
$F_9(x) = \sum_{i=1}^{n} [x_i^2 - 10\cos(2\pi x_i) + 10]$	30	[-5.12,5.12]^n	0
$F_{10}(x) = -20\exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n} \cos(2\pi x_i)) + 20 + e$	30	[-32,32]^n	0
$F_{11}(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos(\frac{x_i}{\sqrt{i}}) + 1$	30	[-600,600]^n	0
$F_{12}(x) = \frac{\pi}{n}\{10\sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2[1 + 10\sin^2(\pi y_{i+1})] + (y_n - 1)^2\} + \sum_{i=1}^{n} u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{x_i+1}{4}$ $u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$	30	[-50,50]^n	0
$F_{13}(x) = 0.1\{\sin^2(3\pi x_1) + \sum_{i=1}^{n} (x_i - 1)^2[1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2[1 + \sin^2(2\pi x_n)]\} + \sum_{i=1}^{n} u(x_i, 5, 100, 4)$	30	[-50,50]^n	0

Table (8.2): Multimodal test functions with variable dimensions

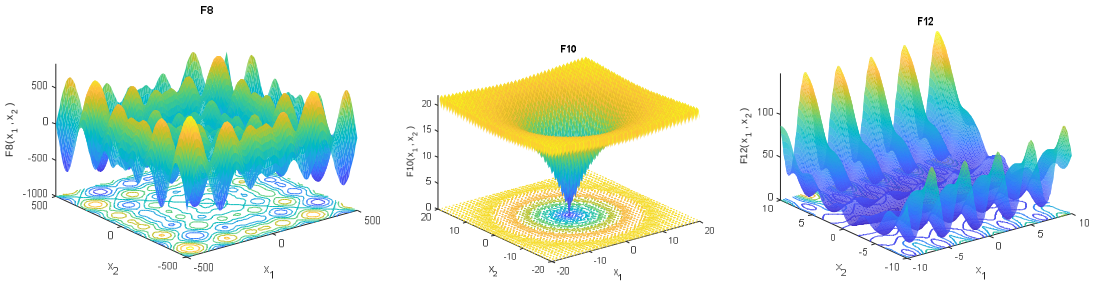


Figure (8.8): Typical two-dimensional form of high-dimensional multimodal benchmark functions

Function	V_no	Range	f_{min}
$F_{14}(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j+\sum_{i=1}^{2}(x_i - a_{ij})^6}\right)^{-1}$	2	$[-65,65]$	1
$F_{15}(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1(b_i^2+b_i x_2)}{b_i^2+b_i x_3+x_4}\right]^2$	4	$[-5,5]$	0.00030
$F_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	2	$[-5,5]$	-1.0316
$F_{17}(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$	2	$[-5,5]$	0.398
$F_{18}(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$ $\times[30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	2	$[-2,2]$	3
$F_{19}(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{3} a_{ij}(x_j - p_{ij})^2\right)$	3	$[1,3]$	-3.86
$F_{20}(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{6} a_{ij}(x_j - p_{ij})^2\right)$	6	$[0,1]$	-3.32
$F_{21}(x) = -\sum_{i=1}^{5} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.1532
$F_{22}(x) = -\sum_{i=1}^{7} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.4028
$F_{23}(x) = -\sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.5363

Table (8.3): Fixed-dimension multimodal test functions

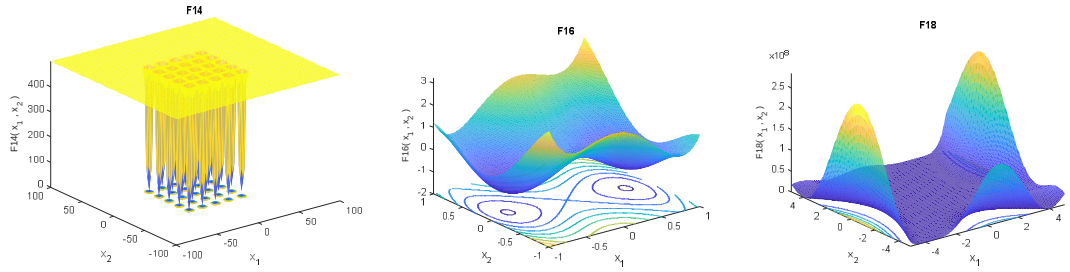


Figure (8.9): Typical two-dimensional form of fixed-dimensional multimodal benchmark functions

8.3.1.1. Unimodal Test Functions

Functions F1-F7 in Table (8.1) are unimodal functions. Unimodal functions are those functions which have only single local minima and these functions are suitable for testing the convergence rate and exploitation of algorithms.

Minimization results of unimodal benchmark functions are given in Table (8.4) (due to accuracy of Matlab software, the values below 10^{-16} are assumed to be 0).

As Table (8.4) shows, the KA algorithm outperforms the other algorithms on 6 out of 7 unimodal test functions. The results are followed by the CAB, GSA, PSO, FA and GA algorithms. Only in the case of function F6 the CAB algorithm provides the better mean value over 30 runs, however, in this case, the standard deviation (i.e. SD) of the CAB algorithm is much higher than the SD value in the KA method. F6 poses a difficulty for the KA algorithm, since the flatness of the function does not give the algorithm any information to direct the search process towards the minima. Furthermore, the KA algorithm does provide very competitive results compared to the CAB on F5 and F7. The benchmark F7 is a noisy function, which contains a random term. The best results are those of KA. As Table (8.4) shows all algorithms are able to optimize F1 with a more or less acceptable error. Unlike the other algorithms, KA and GSA are able to reach the global optimum without any error. Although none of the six tested algorithms could optimize F5 which is a non-separable function the KA was able to provide a better result. Note that in general, inseparable functions are relatively difficult to solve, when compared with their separable counterpart, where each variable is independent of the other variables. If all the variables are independent, then a sequence of n independent optimization processes can be performed and then each design variable or parameter can be optimized independently (Salomon, 1996). F3 is also a non-separable function, for which the CAB algorithm comes close to the global optimum (which is equal to zero) with an error of about 10^{-9}, while the best result obtained by the KA algorithms without any error. Some algorithms such as GSA and GA are not able to find good solutions for this benchmark.

Moreover, to significantly compare the algorithms, we conducted a pairwise comparison to detect significant performance between all the algorithms with the non-parametric Wilcoxon rank test (Wilcoxon, 1945), (Garc´ıa, et al., 2009). The analysis is performed considering a 5% significance level over the mean value data. Table (8.5) reports the p-values produced by Wilcoxon's test for the pairwise comparison of the five groups. Such groups are formed by KA versus GA, KA versus PSO, KA versus FA, KA versus GSA, and CAB versus CAB. As a null hypothesis, it is assumed that there is no significant difference between mean values of the two algorithms. The alternative hypothesis considers a significant difference between the mean values of both approaches. The most of p-values reported in Table (8.5) are less than 0.05 (5% significance level), which is a strong evidence against the null hypothesis, indicating that the KA results are statistically significant and that it has not occurred by coincidence i.e., due to the normal noise contained in the process.

In total, the results of Table (8.4) show that the KA algorithm is able to provide very competitive results on the unimodal benchmarks. The p-values in Table (8.5) also prove that the superiority is significant in the majority of the cases. This testifies that the proposed algorithm has a high exploitation ability, which is due to the integrated adaptive growth rate parameter that assists the KA optimization method to provide very good exploitation. High exploitation assists the KA algorithm to rapidly converge towards the optimum and exploit it accurately as can also be inferred from Figures (8.10)- (8.16).

	KA		GA		PSO	
	Mean	SD	Mean	SD	Mean	SD
F1	0	0	0.8078	0.4393	0.000136	0.000202
F2	0	0	0.269483	0.23788	0.042144	0.045421
F3	0	0	0.13902	0.121161	0	0.317039
F4	0	0	1.59375	1.21348	1.086481	0.862029
F5	6.2295	0.0813	369.7545	342.8893	96.71832	60.11559
F6	3.6042e-05	9.8070e-06	6.984222	7.010388	0.167918	0.868638
F7	9.4939e-05	2.9781e-05	0.047174	0.043587	0.010073	0.003263

	FA		GSA		CAB	
	Mean	SD	Mean	SD	Mean	SD
F1	0.039615	0.01449	0	0	8.2e-14	5.9e-14
F2	0.050346	0.012348	0.055655	0.194074	1.43e-9	9.9e-10
F3	0.049273	0.019409	50.9289	20.6008	3.51e-12	7.4e-11
F4	0.145513	0.031171	1.8027e-09	2.5326e-10	4.96e-10	2.93e-06
F5	21.75892	14.47251	67.54309	62.22534	25.16	7.1944
F6	0.05873	0.014477	0.089441	0.04339	1.03e-12	1
F7	0.000853	0.000504	0.018	0.0055	0.089441	0.0012

Table (8.4): Results of unimodal test functions

KA versus	GA	PSO	FA	GSA	CAB
F1	8.0065e-09	5.4955e-05	8.0065e-09	N/A	3.9926e-06
F2	2.1025e-07	5.4987e-04	8.0065e-09	5.4987e-04	3.9926e-06
F3	3.9926e-06	0.4643	2.1025e-07	8.0065e-09	0.2534
F4	8.0065e-09	8.0065e-09	8.0065e-09	8.0065e-09	0.2534
F5	1.6098e-04	6.7956e-08	6.7956e-08	0.0012	1.2009e-06
F6	3.5557e-08	2.4844e-04	6.7956e-08	6.7956e-08	0.2853
F7	1.5997e-05	6.7956e-08	8.4848e-09	6.7956e-08	6.7956e-08

Table (8.5): P-values produced by Wilcoxon's test comparing KA versus GA, PSO, FA, GSA and CAB over all runs for unimodal test functions

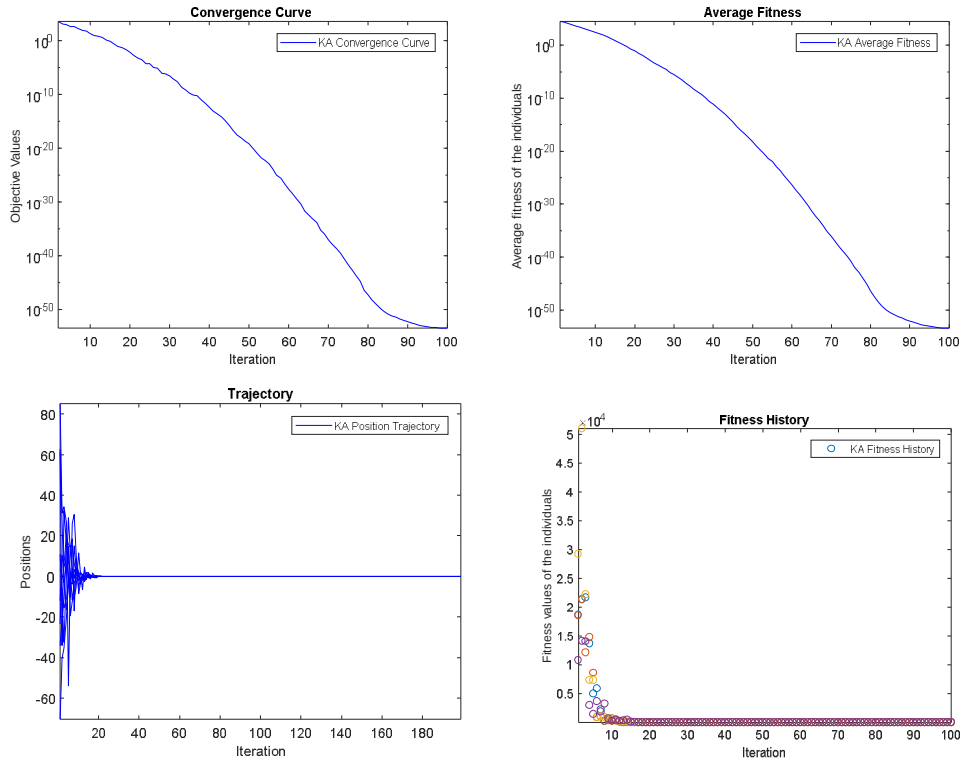


Figure (8.10): Benchmark function F1-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

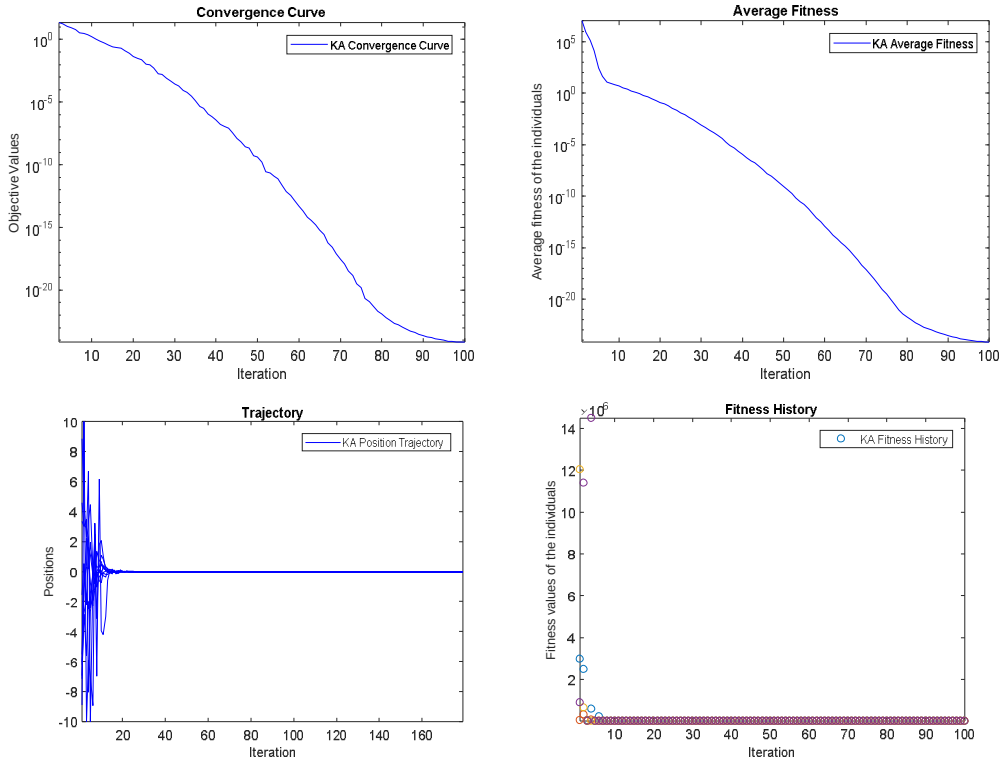


Figure (8.11): Benchmark function F2-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

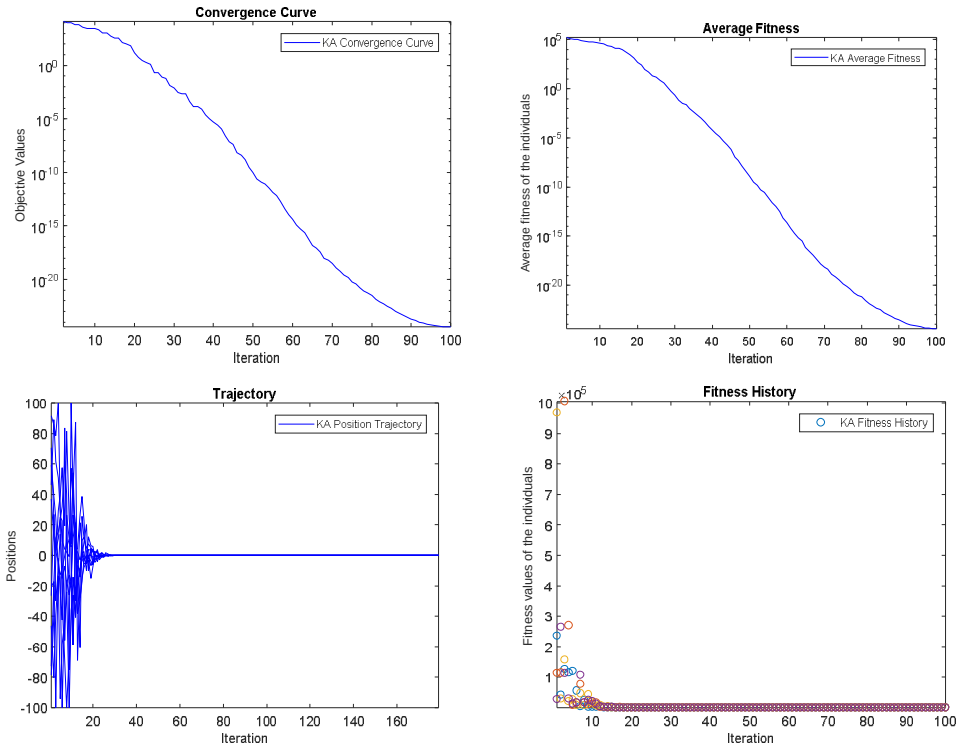


Figure (8.12): Benchmark function F3-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

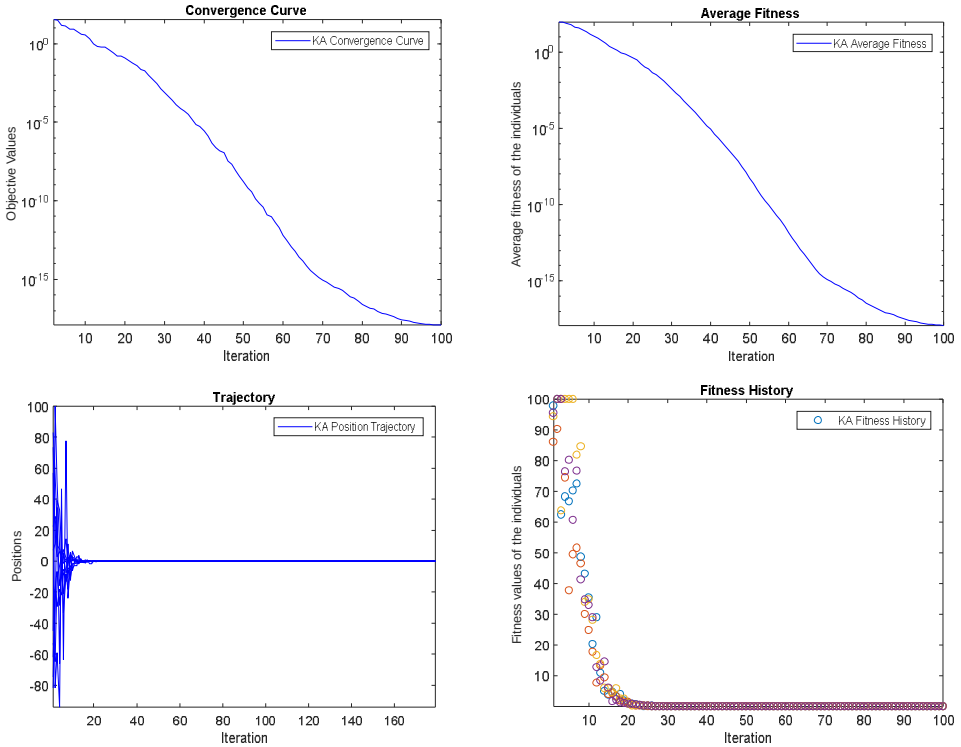


Figure (8.13): Benchmark function F4-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

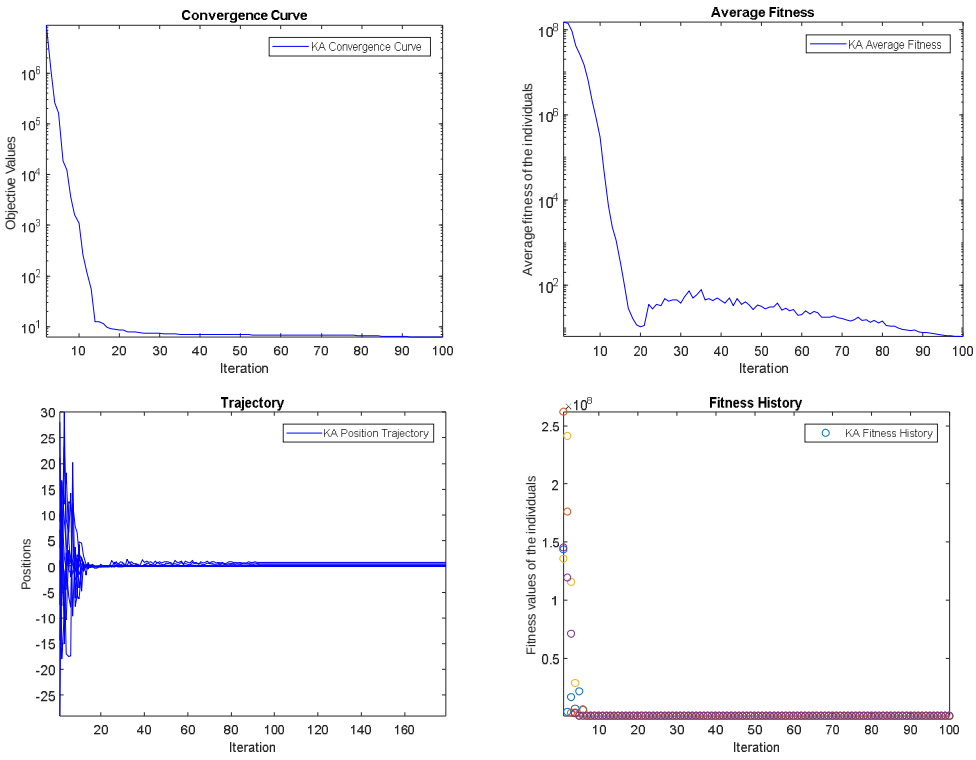


Figure (8.14): Benchmark function F5-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

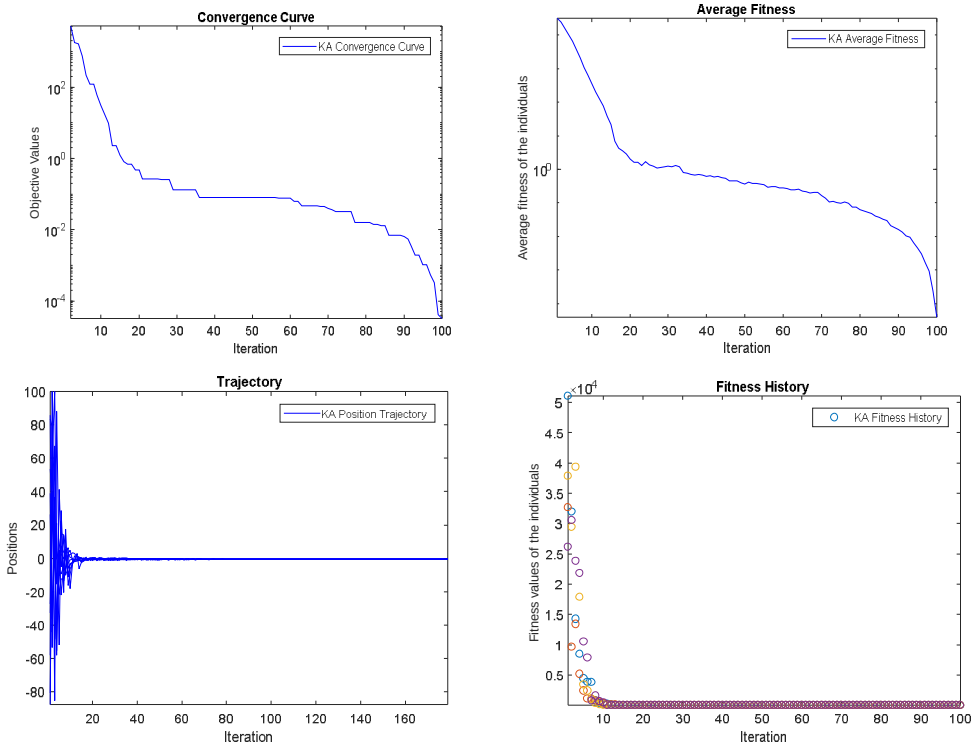


Figure (8.15): Benchmark function F6-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

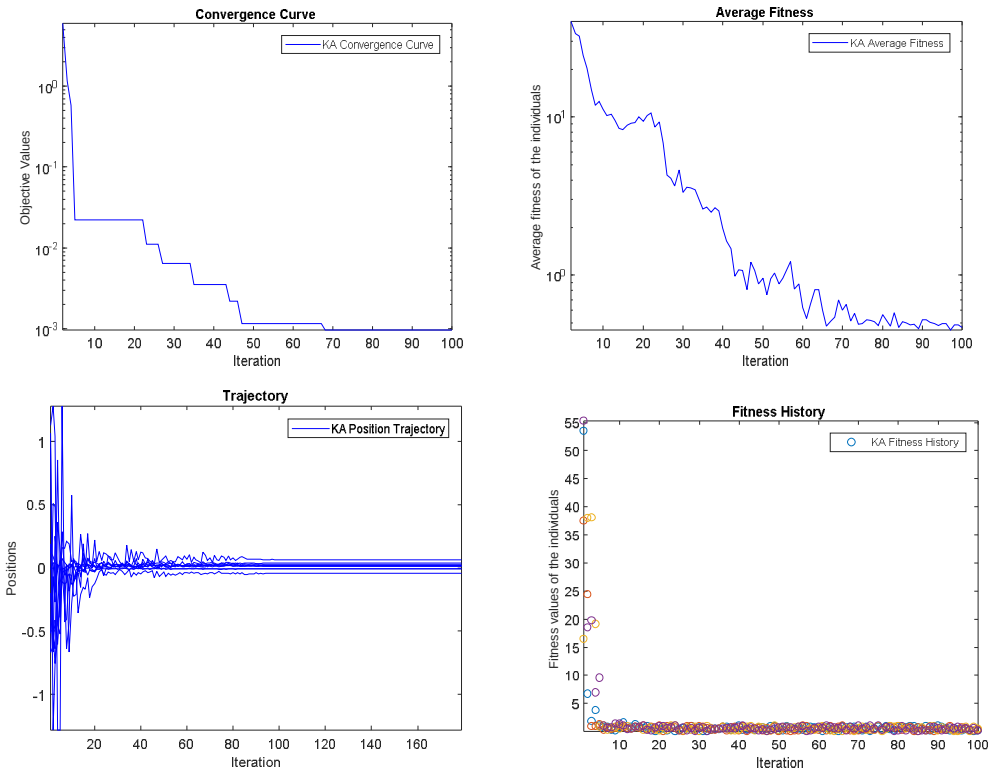


Figure (8.16): Benchmark function F7-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

8.3.1.2. Multimodal Test Functions

Functions F8-F23 given in Table (8.2), (8.3) are multimodal functions. Multimodal functions are those functions which have more than one local optimum. These functions are used to test the ability of an algorithm to escape from any local minimum. If the exploration process of an algorithm is poorly designed, then it cannot search the function landscape effectively. This, in turn, leads to an algorithm getting stuck at a local minimum. Multimodal functions with many local minima are among the most difficult class of problems for optimization algorithms. The results reported in Table (8.6), (8.7) show that KA algorithm has a very good exploration capability, since it is the most efficient or the second-best algorithm in the majority of test problems. This is due to integrated mechanisms of exploration in the proposed algorithm that leads this algorithm towards the global optimum.

F9 is one of the multimodal functions that is difficult to optimize. Thus, only the KA algorithm was able to reach the global optimum. The only function that the KA algorithm cannot optimize is F8, which is a complex multimodal function with many local minima and very difficult to optimize. However, as the results show, the KA optimizer outperformed the other algorithms in the optimization of this benchmark function. The results of 30 runs with KA for the multimodal functions are often better than the results found by concurrent algorithms, with the exception of PSO and CAB which show a performance almost equal to that of the F16-F20 KA results. In benchmark functions F10 and F14 the area that contains the global minima are very small, when compared to the whole search space. Optimizing this kind of functions is really challenging. In addition, KA is the only one who can optimize F9-F11 and the

results found by the KA algorithm for F12, F13, F15, F20, F23 are the best. Only KA and CAB with theirs two variants reached the solutions very close to the global minimum of the function F21-F23. For the benchmark F15 the global minimum is located very close to the local minima. This function is notoriously difficult. Furthermore, in optimization, a problem that algorithms may suffer is the scaling problem with many orders of magnitude differences between the domain and the function hyper-surface (Junior, et al., 2004), it can be seen in benchmark F18. Only KA could optimize it without any error. The results obtained with KA in Tables (8.6), (8.7) evidence that even the worst outcomes of KA outperformed those of the other algorithms in the majority of the test. The p-values reported in Tables (8.8) and (8.9) also show that the KA algorithm has significantly better results. Therefore, it can be concluded that the KA algorithm is efficient algorithm with a high level of exploration, which assists it to explore the promising regions of the search space and to avoid all of the local optima and approach the global optimum.

	KA		**GA**		**PSO**	
	Mean	SD	Mean	SD	Mean	SD
F8	-3004.4716	400.9120	-2091.64	2.47235	-1367.01	146.4089
F9	0	0	0.659271	0.815751	0.278588	0.218991
F10	0	0	0.956111	0.807701	1.11e-09	2.39e-11
F11	0	0	0.487809	0.217782	0.009215	0.007724
F12	1.2623e-06	2.993e-06	0.110769	0.002152	0.006917	0.026301
F13	5.7287e-06	9.5067e-04	0.129	0.068851	0.006675	0.008907

	FA		**GSA**		**CAB**	
	Mean	SD	Mean	SD	Mean	SD
F8	-1245.59	353.2667	-1108.01	574.7	-1200	50.42824
F9	0.263458	0.182824	25.96841	7.470068	0.001	8.45e-06
F10	0.168306	0.050796	0.062087	0.23628	0.0031	0.003096
F11	0.099815	0.024466	27.70154	5.040343	0.122678	0.049673
F12	0.126076	0.263201	1.799617	0.95114	5.60e-06	1.58e-8
F13	0.00213	0.001238	8.899084	7.126241	9.5067e-06	6.09e-07

Table (8.6): Results of high-dimensional multimodal test functions

	KA		GA		PSO	
	Mean	SD	Mean	SD	Mean	SD
F14	0.998	3.3e-16	2.111973	2.498594	3.627168	2.560828
F15	0.00031007	6.2200e-06	0.0005	0.00032	0.000577	0.000222
F16	-1.0316	0	-1.03	4.9e-07	-1.03163	6.25e-16
F17	0.39789	0	0.398	1.5e-07	0.397887	0
F18	3	0	3.02	0.11	3	1.33e-15
F19	-3.8623	0	-3.86	1.4e-5	-3.86278	2.58e-15
F20	-3.3219	0.026	-2.98105	0.376653	-3.26634	0.060516
F21	-10.1489	1.0000e-04	-5.52	1.59	-6.8651	3.019644
F22	-10.3908	0.0117	-5.53	2.12	-8.45653	3.087094
F23	-10.5326	0.0094	-6.57	3.14	-9.95291	1.782786

	FA		GSA		CAB	
	Mean	SD	Mean	SD	Mean	SD
F14	1.22	0.56	5.859838	3.831299	0.999	0.0003
F15	0.00058	0.000324	0.003673	0.001647	2.2e-3	8.8000e-04
F16	-1.03163	4.2e-07	−1.03163	4.88e-16	-1.0316	3.1e-13
F17	0.397914	2.7e-05	0.397887	0	0.397887	9.9e-09
F18	3	4.22e-15	3	4.17e-15	3	2e-15
F19	-3.85616	0.002706	-3.86278	2.29e-15	-3.86278	0.0127
F20	-3.27	0.059	-3.31778	0.023081	-3.2369	0.0471
F21	-6.04918	3.629551	-5.95512	3.737079	-10.1532	2.5e-06
F22	-8.18178	3.829202	-9.68447	2.014088	-10.4028	3.9e-07
F23	-9.34238	2.414737	-10.5364	2.6e-15	-10.5323	7e-4

Table (8.7): Results of fixed-dimensional multimodal test functions

KA versus	GA	PSO	FA	GSA	CAB
F8	6.7956e-08	6.7956e-08	6.7956e-08	9.1728e-08	6.7956e-08
F9	5.4987e-04	5.4955e-05	5.4955e-05	8.0065e-09	8.0065e-09
F10	5.4987e-04	8.0065e-09	8.0065e-09	0.0215	5.4955e-05
F11	8.0065e-09	5.4955e-05	8.0065e-09	8.0065e-09	8.0065e-09
F12	6.7956e-08	0.5979	0.0315	6.7956e-08	1.2009e-06
F13	6.7956e-08	1.0373e-04	1.4309e-07	1.5997e-05	0.1075

Table (8.8): P-values produced by Wilcoxon's test comparing KA versus GA, PSO, FA, GSA and CAB over all runs for high-dimensional multimodal test functions

KA versus	GA	PSO	FA	GSA	CAB
F14	0.0388	6.6725e-05	0.0058	2.1882e-15	6.4918e-18
F15	5.8913e-04	3.7961e-13	3.2874e-08	5.5727e-10	0.3790
F16	1.2118e-12	1.1275e-12	1.2118e-12	1.1092e-12	0.6411
F17	1.2118e-12	1.6853e-14	2.2081e-06	1.6853e-14	1.2118e-12
F18	0.6411	0.6293	0.0960	0.6409	0.3393
F19	1.2118e-12	1.1661e-12	1.2118e-12	1.1882e-12	0.2706
F20	2.6816e-07	0.1568	0.9246	0.3794	0.0771
F21	3.0199e-11	1.1077e-06	1.1077e-06	9.5139e-06	3.0199e-11
F22	3.0199e-11	9.5139e-06	1.1077e-06	0.0315	1.5997e-05
F23	1.1077e-06	9.5139e-06	0.0271	0.0263	0.1537

Table (8.9): P-values produced by Wilcoxon's test comparing KA versus GA, PSO, FA, GSA and CAB over all runs for fixed-dimensional multimodal test functions

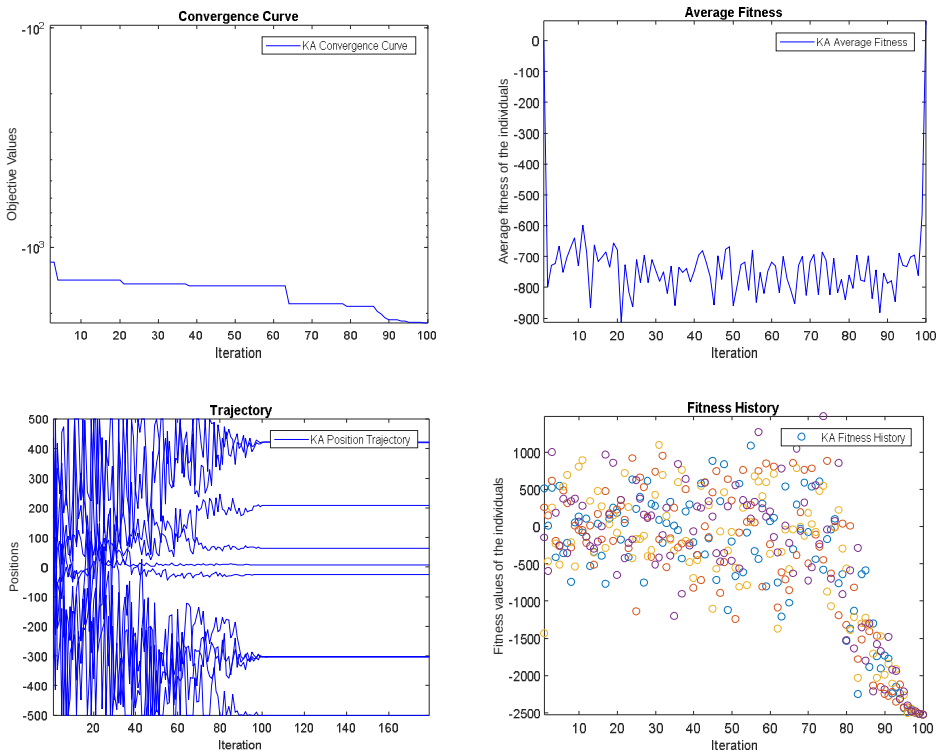


Figure (8.17): Benchmark function F8-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

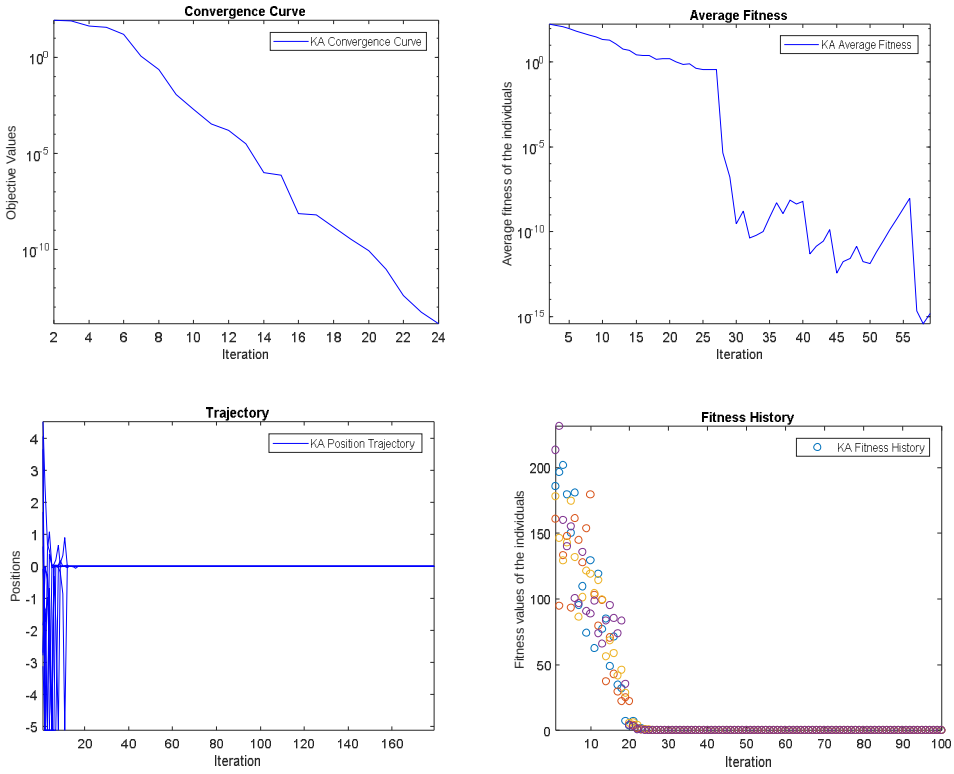


Figure (8.18): Benchmark function F9-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

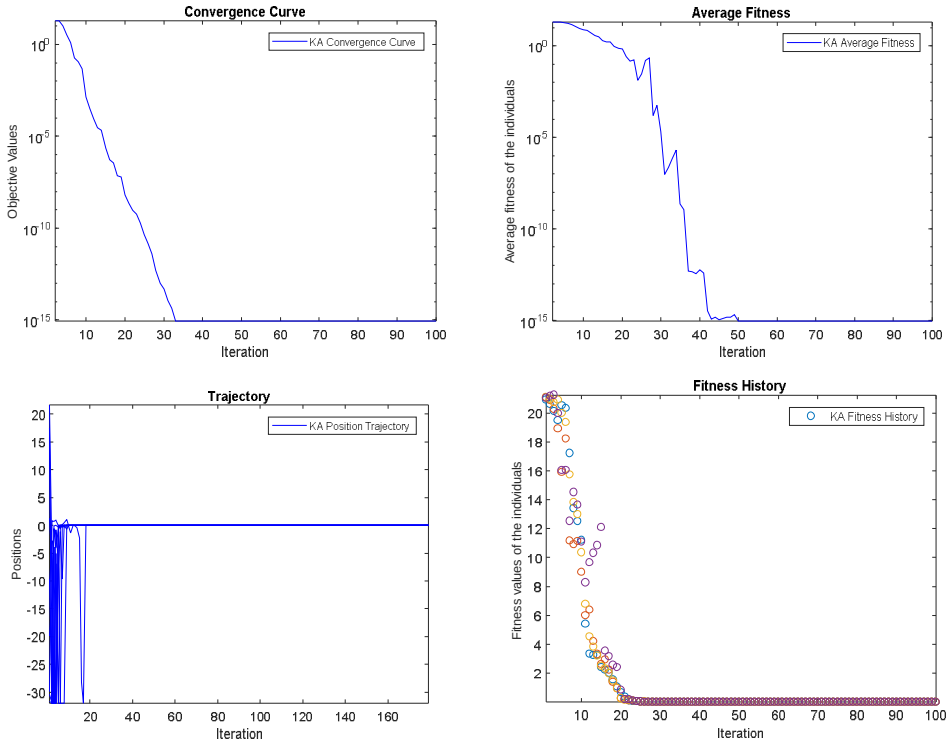


Figure (8.19): Benchmark function F10-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

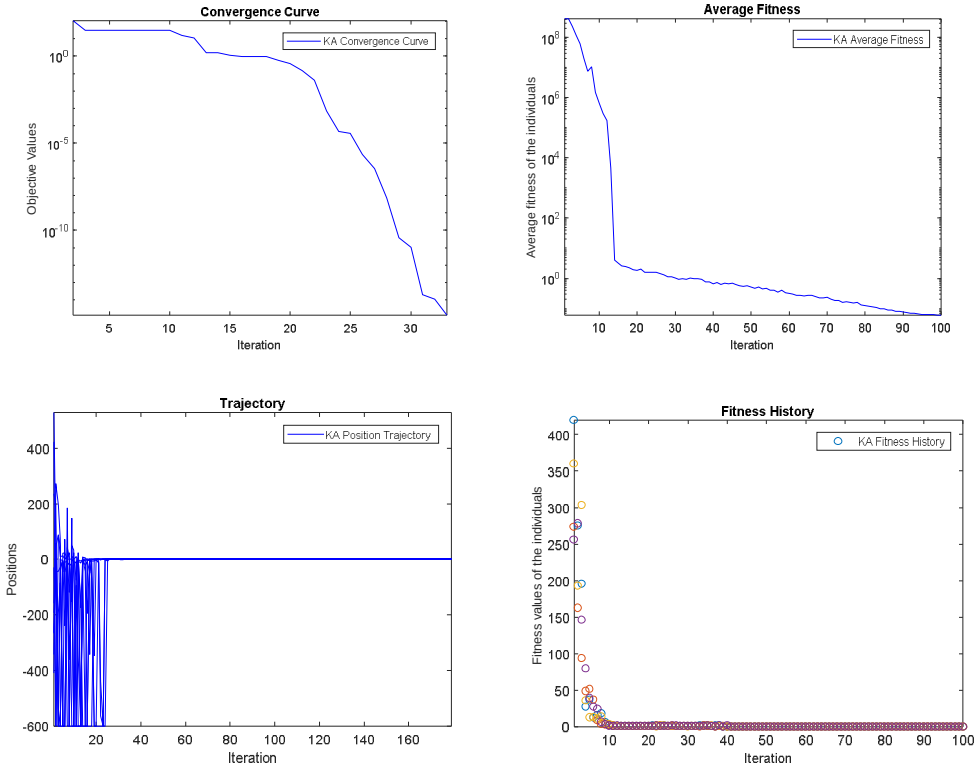


Figure (8.20): Benchmark function F11-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

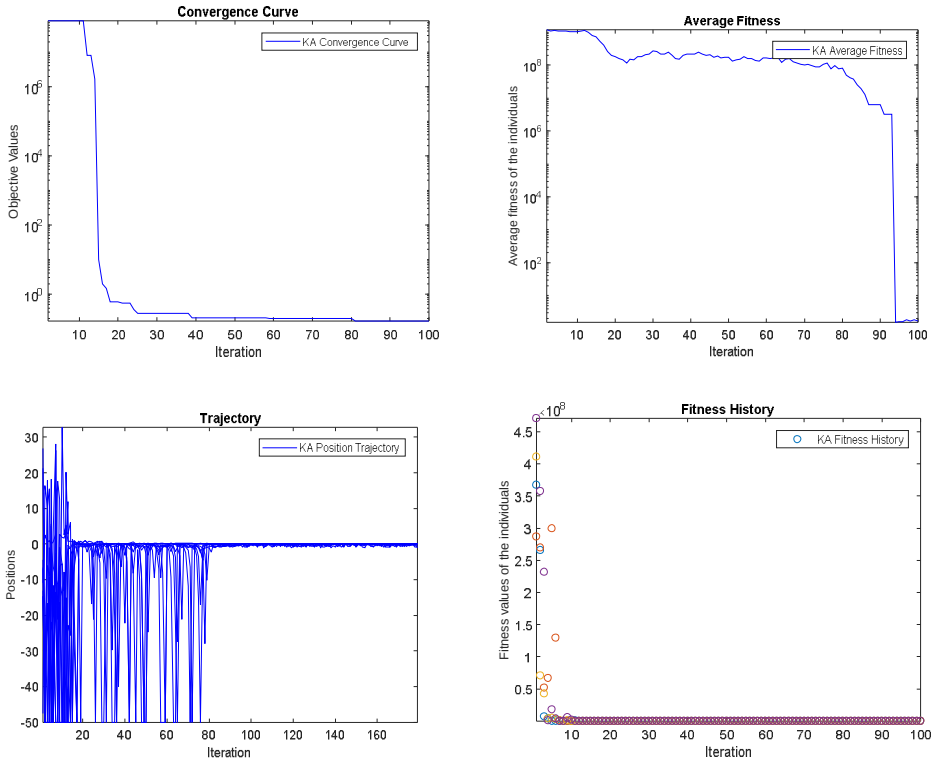


Figure (8.21): Benchmark function F12-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

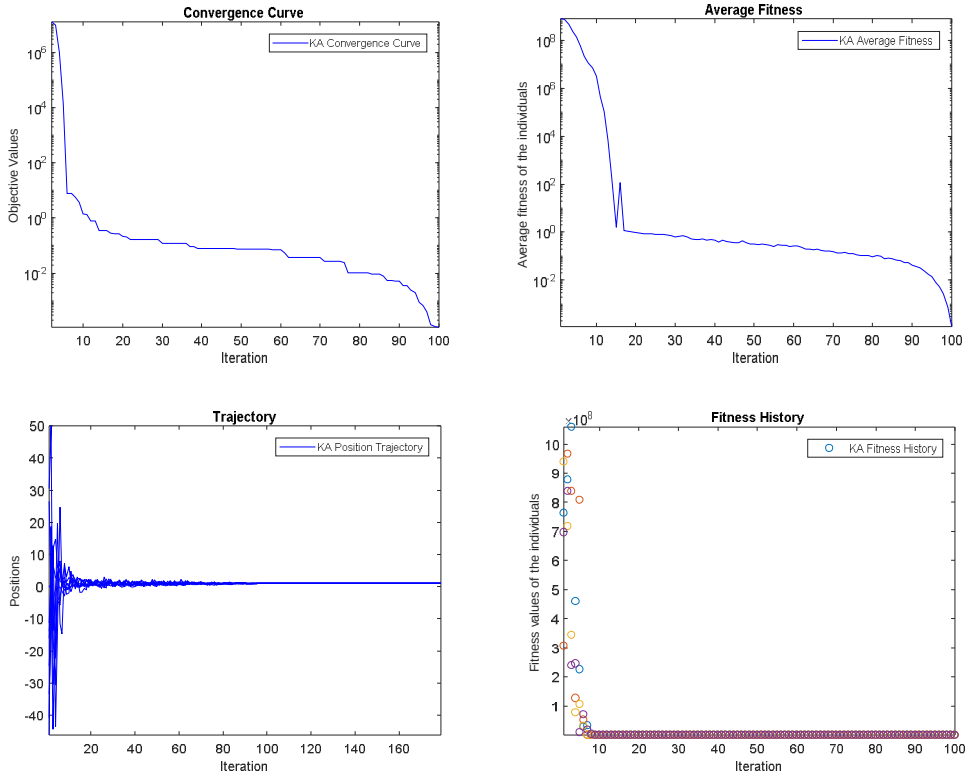


Figure (8.22): Benchmark function F13-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

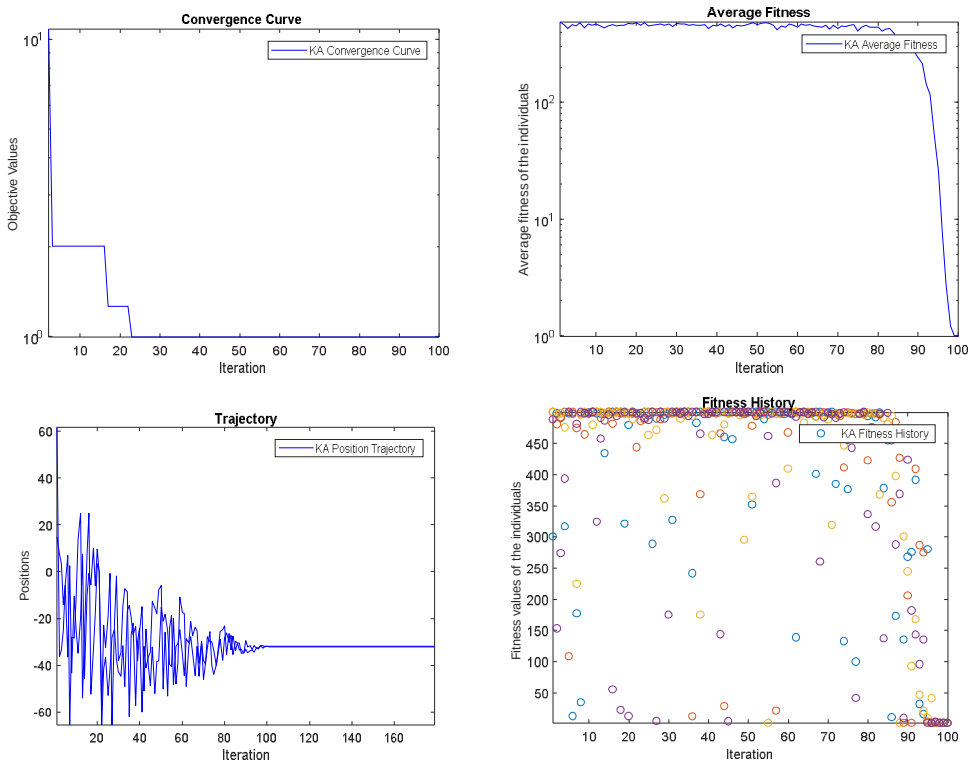


Figure (8.23): Benchmark function F14-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

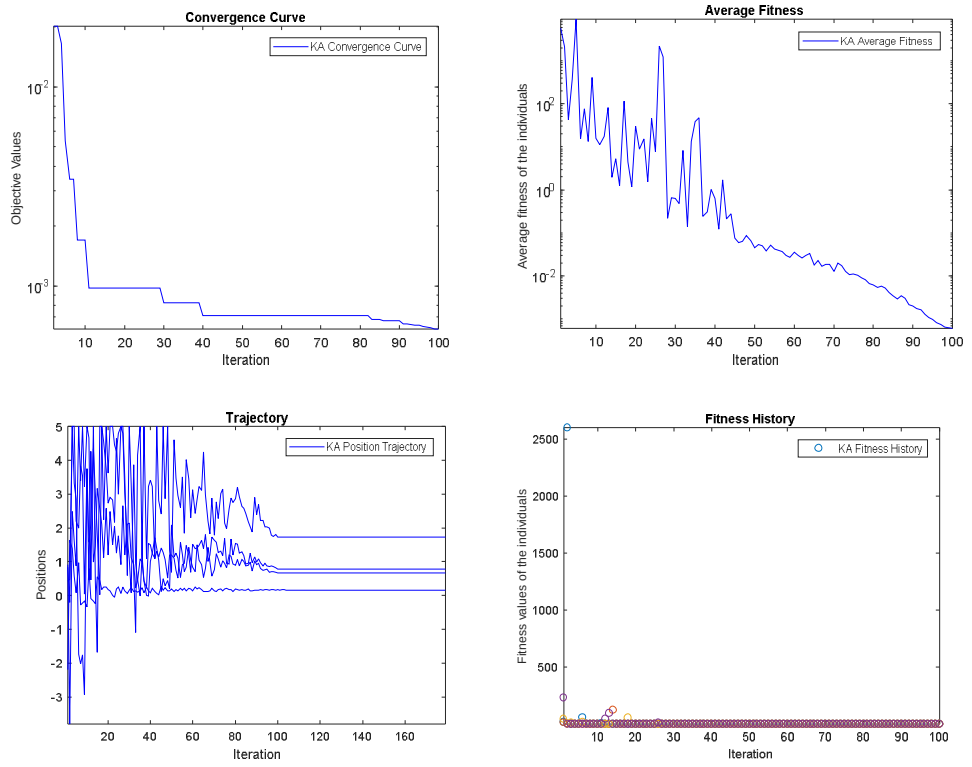


Figure (8.24): Benchmark function F15-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

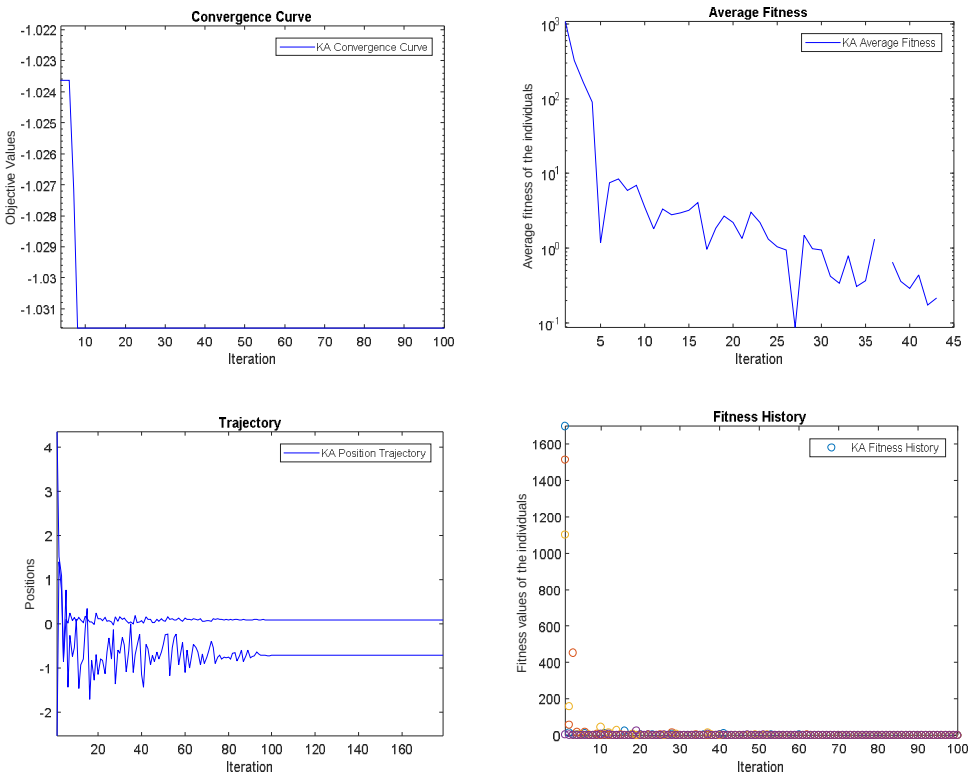


Figure (8.25): Benchmark function F16-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

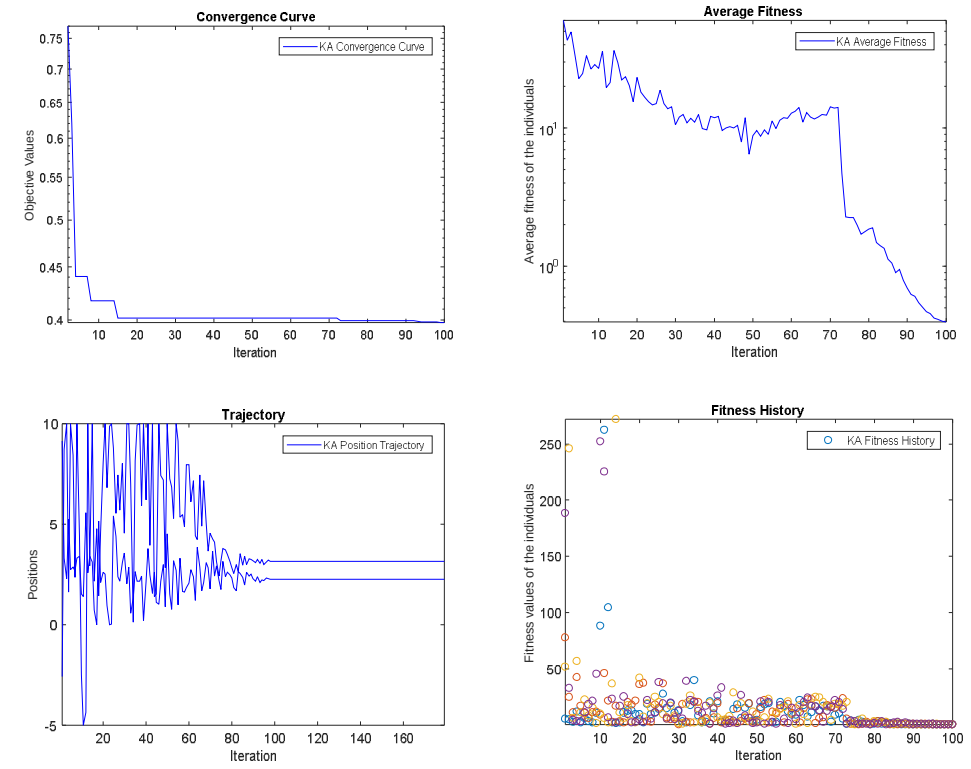


Figure (8.26): Benchmark function F17-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

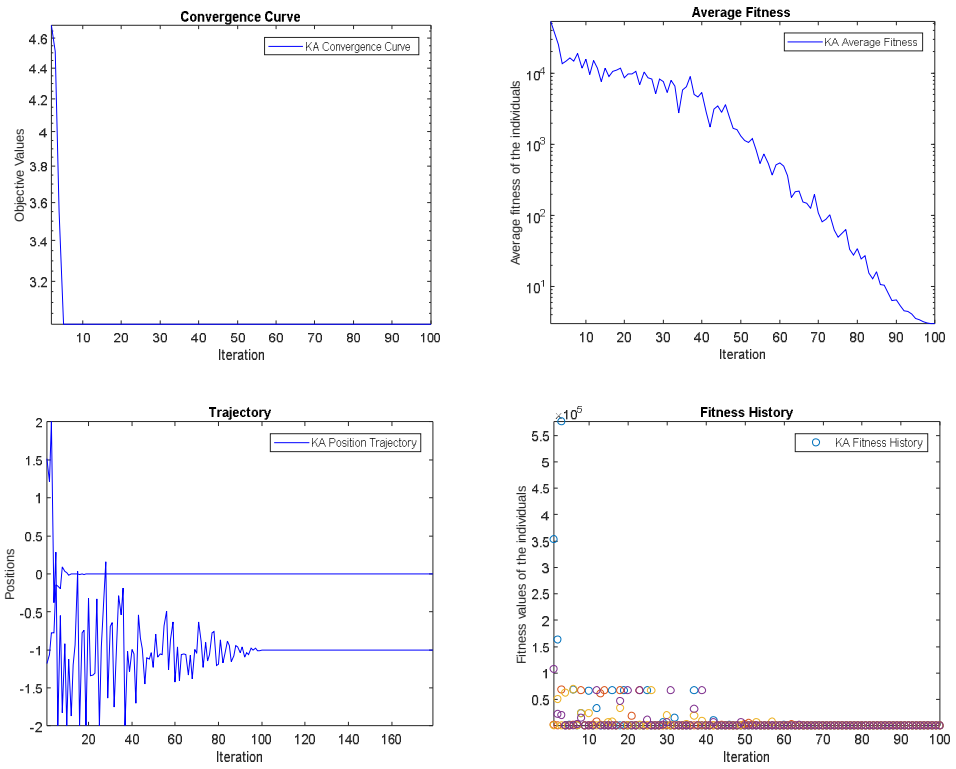


Figure (8.27): Benchmark function F18-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

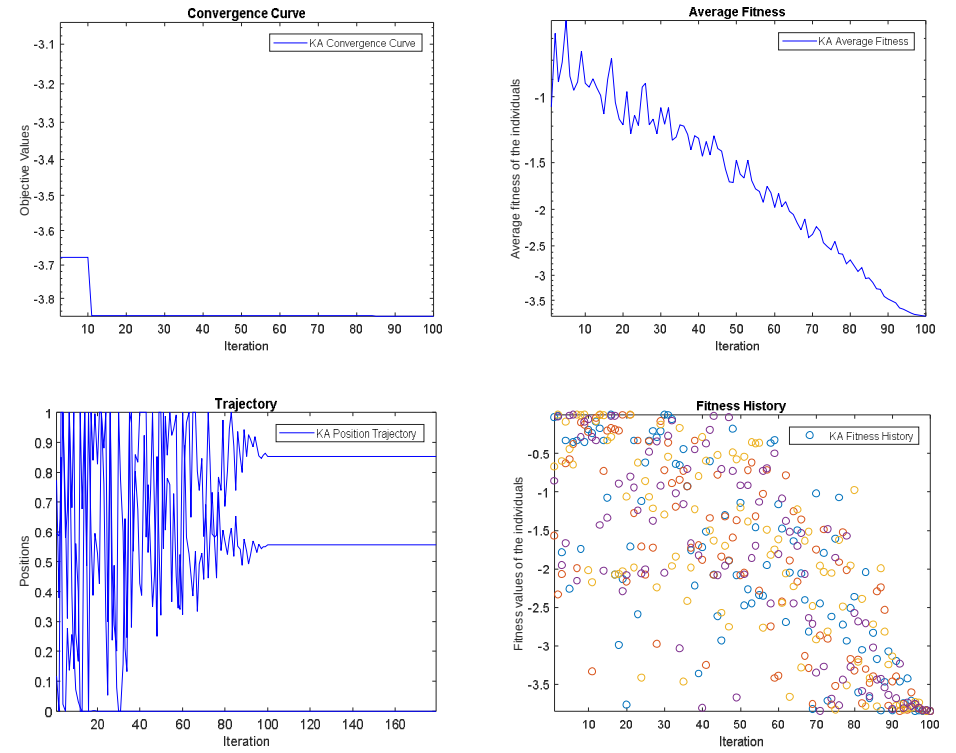


Figure (8.28): Benchmark function F19-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

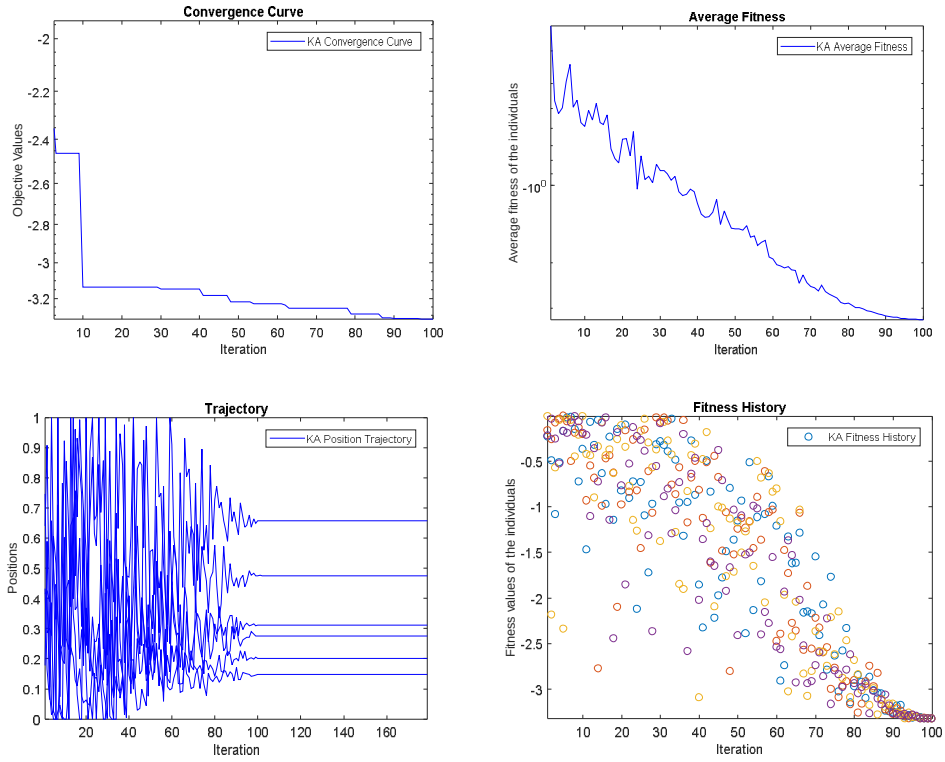


Figure (8.29): Benchmark function F20-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

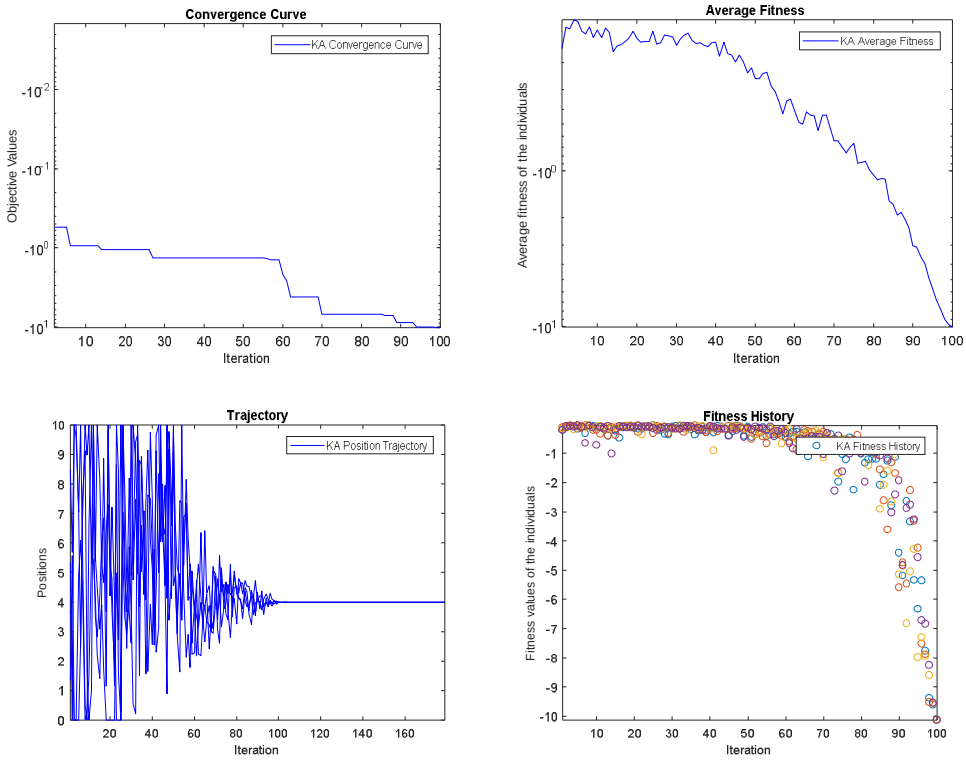


Figure (8.30): Benchmark function F21-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

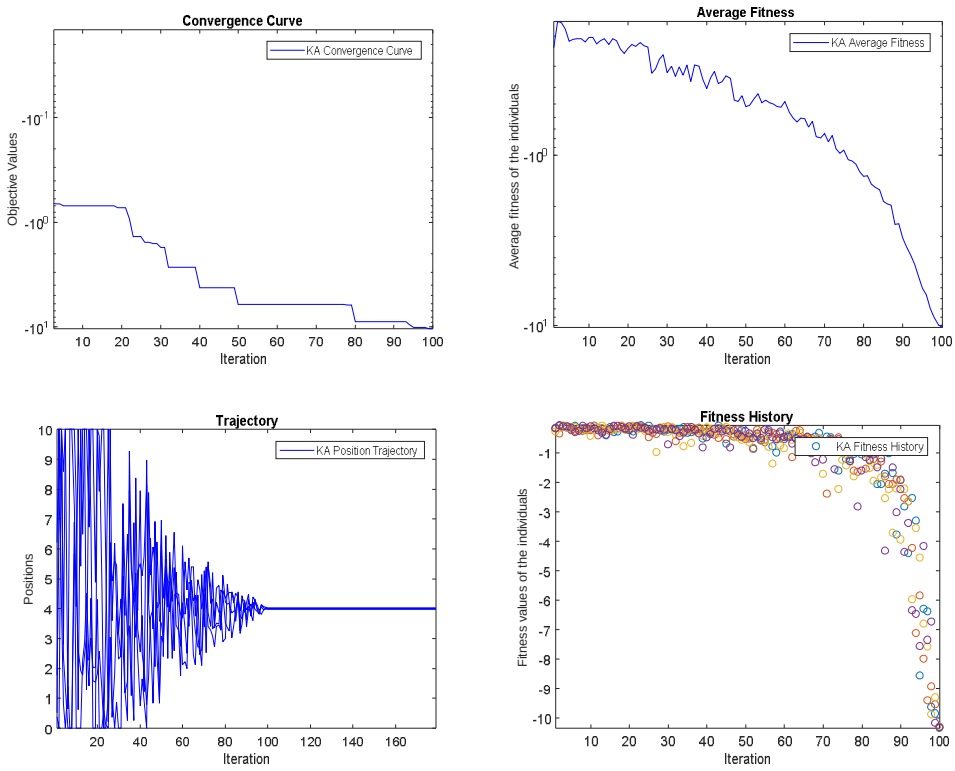


Figure (8.31): Benchmark function F22-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

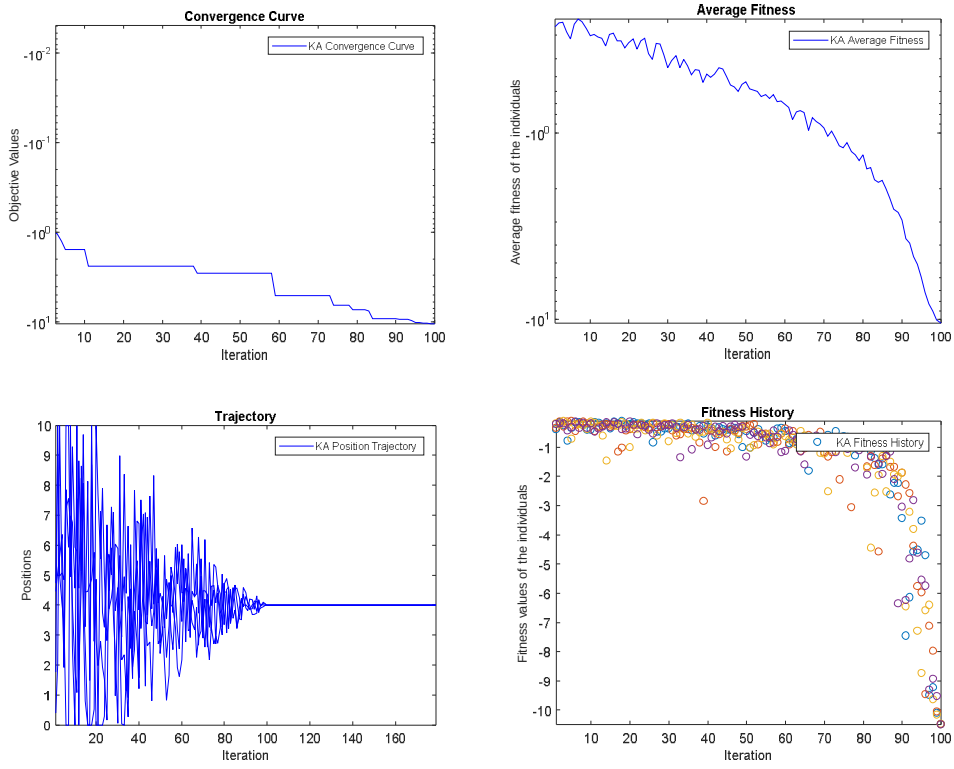


Figure (8.32): Benchmark function F23-The convergence rate, average fitness of all individuals, trajectory and the fitness history.

8.3.2. Analysis of KA Optimizer

This subsection aims for identifying and confirming the convergence and potential behavior of the KA algorithm when solving real problems. Although the results discussed in the previous section prove the high performance of the KA algorithm, there are several quantitative metrics, which can be employed to confidently confirm the performance of this algorithm in solving real problems such as actuator selection. In other words, the change of populations of the species and the obtained fitness values during optimization should be monitored to observe:

- how the average fitness of search agents changes from the first to the last iteration
- how the individuals converge towards the global optima in the search space
- how the algorithm improves the initial random solutions, and their fitness values over the iterations
- if there are abrupt changes in the populations of species in the initial stages of optimization to explore the search space and gradual changes in the final steps of iteration to exploit the search space

The second plots of the first column in Figures (8.10)-(8.32) show the fluctuations of the variables in the search agents. It can be seen that the populations of the species face abrupt fluctuations in the early steps of optimization. However, the sudden changes are decreased gradually over the iterations. This confirms that the search candidates first explore the search space to find the promising regions and then converge around the best solution obtained in the exploration phase. There is a question here as how to make sure that all of the search agents are improved during optimization despite the rapid and steady changes in the aforementioned plots. In order to confirm the improvement of all solutions, the average fitness of all search agents during optimization is illustrated in the first plots of the second column in Figures (8.10)-(8.32). It can be seen that the average fitness of all individuals tends to be decreased over the course of iterations. The interesting pattern that can be observed in this plots is the high fluctuation of the average fitness in the exploration phase and low changes in the average fitness in the last iterations, when the exploitation phase begins. Deterioration of the fitness of some of the search agents is inevitable in the exploration. However, the general pattern in the average plots illustrates that the fitness of candidates has a descending behavior over the iterations. This proves that the proposed KA algorithm is able to eventually improve the fitness of initial random solutions for a given optimization problem. In the previous section, it was claimed that the search agents of the KA algorithm tend explore the promising regions of the search space and exploit the best one eventually. However, the convergence behavior of the algorithm was not observed and verified. Although this can be inferred indirectly from the trajectory and average fitness, the convergence curve of KA is depicted in the first plots of Figures (8.10)-(8.32). These plots illustrate the best solutions obtained so far by KA during optimization. The descending trend is quite evident in the convergence curve of KA on all of the test functions investigated. This strongly evidences the ability of the proposed algorithm in obtaining a better approximation of the global optimum over the course of iterations. All the results and discussions of this section prove that the KA algorithm is suitable for solving real-world optimization problems. This is due to the integrated adaptive parameter selection scheme in the algorithm and employed mechanism of updating the best feasible solutions obtained during the change of populations of individuals, in which

the species are guided towards promising feasible regions of the search space. It leads to boosting the exploration of the feasible areas of a search space
and therefore, KA can solve the real-world challenging optimization problems efficiently with low computational cost.

8.3.3. Optimization Problems Using the KA Algorithm

As the proposed algorithm considers optimization problem as black-boxes, it is readily incorporable to problems in different fields subject to proper problem formulation. In this section, first, some well-studied engineering design problems that have been solved by various optimization methods in the literature are used to examine the efficiency of KA algorithm. Then we used the KA algorithm to solve the problem of optimal actuator placement in some real-world examples.

Example (8.1): (Kaveh & Talatahari, 2010) Consider the well-known optimization problem of tension/compression spring design as depicted in Figure (8.34). The objective of this optimization problem is to minimize the weight of tension/compression spring. The optimization problem can be formulated as:

$$\begin{cases} \min \quad J(x) = (x_3 + 2)x_1^2 x_2 \\ h_1(x) = 1 - \dfrac{x_2^3 x_3}{71785 x_1^4} \le 0 \\ h_2(x) = \dfrac{4x_2^2 - x_1 x_2}{12566(x_2 x_1^3 - x_1^4)} - \dfrac{1}{5108 x_1^2} \le 0 \\ h_3(x) = 1 - \dfrac{140.45 x_1}{x_2^2 x_3} \le 0 \\ h_4(x) = \dfrac{x_1 + x_2}{1.5} - 1 \le 0 \\ 0.05 \le x_1 \le 2, \quad 0.25 \le x_2 \le 1.3, \quad 2 \le x_3 \le 15 \end{cases}$$

Here, the decision variables are the mean coil diameter $D\,(= x_1)$, the wire diameter $d\,(= x_2)$, and the number active coils $N\,(= x_3)$. The results of KA optimization method are illustrated in Table (8.10) and compared to some existing results of the other optimization methods.

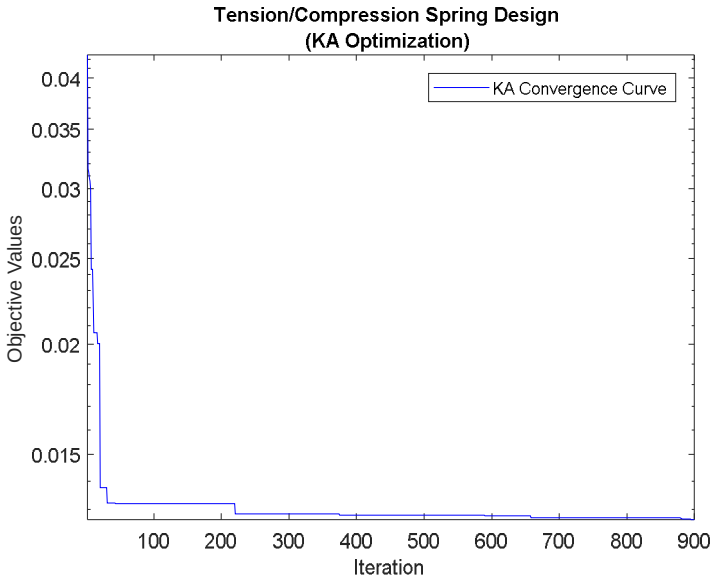


Figure (8.33): Convergence curve of Example (8.1)

Figure (8.34): A tension/compression spring

	x_1	x_2	x_3	$J_{opt}(x)$
KA Optimization	0.0515778	0.354025	11.45	0.012667
Belegundu (Belegundu, 1982)	0.050000	0.315900	14.250000	0.0128334
Arora (Arora, 2016)	0.053396	0.399180	9.1854000	0.0127303
GA (Coello, 2000)	0.051480	0.351661	11.632201	0.0127048
GA (Coello & Montes, 2002)	0.051989	0.363965	10.890522	0.0126810
PSO (He & Wang, 2007)	0.051728	0.357644	11.244543	0.0126747
ES (Montes & Coello, 2008)	0.051643	0.355360	11.397926	0.012698

Table (8.10): The numerical results of Example (8.1)

□

Example (8.2): Consider a pressure vessel design problem (Sandgren, 1988), (Kaveh, 2016), where the objective is to minimize the cost of fabricating a pressure vessel which is clapped at both ends by hemispherical heads as depicted in Figure (8.35). The design variables are the

thickness of the shell $T_s(=x_1)$, the thickness of the head $T_h(=x_2)$, the inner radius $R(=x_3)$ and the length of cylindrical section of the vessel $L(=x_4)$. The optimization problem can be defined as:

$$\begin{cases} \min \quad J(x) = 0.622x_1x_3x_4 + 1.7781x_2x_3^2 + 3.1661x_1^2 + 19.84x_1^2x_3 \\ h_1(x) = -x_1 + 0.0193x_3 \le 0 \\ h_2(x) = -x_2 + 0.0095x_3 \le 0 \\ h_3(x) = -\pi x_3^2x_4 - \frac{4}{3}\pi x_3^3 + 1296000 \le 0 \\ h_4(x) = x_4 - 240 \le 0 \\ 0 \le x_1 \le 99, \quad 0 \le x_2 \le 99, \quad 10 \le x_3 \le 200, \quad 10 \le x_4 \le 200 \end{cases}$$

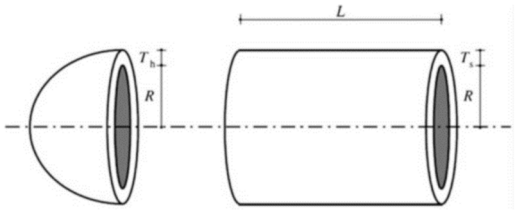


Figure (8.35): A pressure vessel, and its design variables

The results of KA optimization method are illustrated in Table (8.11) and compared to some existing results of the other optimization methods.

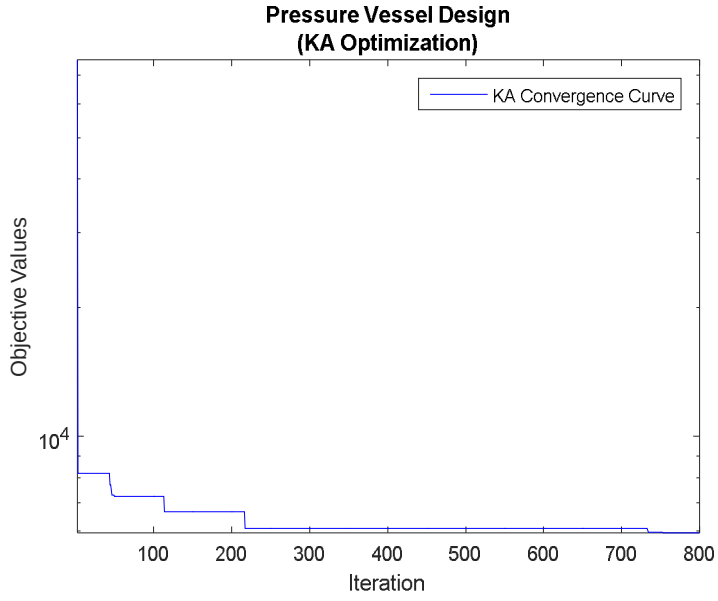


Figure (8.36): Convergence curve of KA in Example (8.2)

	T_s	T_h	R	L	J_{opt}
KA Optimization	0.804799	0.4026198	41.68858	181.8004	5949.0906
PSO (He & Wang, 2007)	0.8125	0.4375	42.091266	176.746500	6061.0777
GA (Coello, 2000)	0.8125	0.4345	40.323900	200.000000	6288.7445
GA (Coello & Montes, 2002)	0.8125	0.4375	42.097398	176.654050	6059.9463
GA (Deb, 1997)	0.9375	0.5000	48.329000	112.679000	6410.3811
ES (Montes & Coello, 2008)	0.8125	0.4375	42.098087	176.640518	6059.7456
DE (Li, et al., 2007)	0.8125	0.4375	42.098411	176.637690	6059.7340
ACO (Kaveh & Talatahari, 2010)	0.8125	0.4375	42.103624	176.572656	6059.0888
Lagrangian Multiplier (Kannan & Kramer, 1994)	1.1250	0.6250	58.291000	43.6900000	7198.0428
Branch-bound (Sandgren, 1988)	1.1250	0.6250	47.700000	117.701000	8129.1036

Table (8.11): The numerical results of Example (8.2)

□

In the next example, we apply KA optimization method to solve the problem of designing a robust PID controller. The results obtained are compared to those of ALPSO (Bellman, 2003). The problem is mathematically formulated as follows:

$$\begin{aligned} \min \quad & J(x) = \arg\max_{\lambda_i(x)} \left\{ \mathrm{Re}(\lambda_i(x)), \forall i \right\} \\ s.t & \begin{cases} h_1(x) = \sup\limits_{\omega \geq 0} \left| \left[W_S(s) \right]_{s=j\omega} \left[S(s,x) \right]_{s=j\omega} \right| - 1 \leq 0 \\ h_2(x) = \sup\limits_{\omega \geq 0} \left| \left[W_T(s) \right]_{s=j\omega} \left[T(s,x) \right]_{s=j\omega} \right| - 1 \leq 0 \\ lb_i \leq x_i \leq ub_i, \quad i = 1, \ldots, 4 \end{cases} \end{aligned} \tag{8.23}$$

where $\lambda_i(x)$ denotes the i[th] pole of the closed-loop system. $W_S(s)$ and $W_T(s)$ are weighting matrices.

$x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T$ is the vector of variables, lb_i , ub_i are the lower and upper bounds respectively. s is the Laplace variable, ω is the frequency (rad/s). $S(s,x)$ denotes the sensitivity function defined as:

$$S\left(s,\ x\right)\ =\ 1/\left(1\ +\ L\left(s,\ x\right)\right).$$

$T(s,x)$ is the closed-loop system, i.e.:

$$T\left(s,x\right) = L\left(s,\ x\right)/\left(1\ +\ L\left(s,\ x\right)\right)$$

and $L\left(s,\ x\right)$ is the open-loop transfer function:

$$L\left(s,\ x\right) = P\left(s\right)K\left(s,x\right)$$

where $P\left(s\right)$ is the transfer function of the system and $K\left(s,x\right)$ the transfer function of the PID controller:

$$K\left(s,x\right) = 10^{x_1}\left(1+10^{-x_2}s^{-1}+\frac{10^{x_3}s}{1+0^{(x_3-x_4)}s}\right) \tag{8.24}$$

This optimization problem has been solved for a magnetic levitation system in the next example. For more details see (Bellman, 2003) and references therein.

Example (8.3): The transfer function of a magnetic levitation system is defined as:

$$P(s) = \frac{7.147}{(s-22.55)(s+20.9)(s+13.99)}$$

The objective is to design the robust PID controller:

$$K\left(s,x\right) = 10^{x_1}\left(1+10^{-x_2}s^{-1}+\frac{10^{x_3}s}{1+0^{(x_3-x_4)}s}\right)$$

Let us define the weighting matrices in (8.23) as:

$$W_S(s) = \frac{5}{s+0.1}\ ,\ W_T(s) = \frac{43.867(s+0.066)(s+88)(s+31.4)}{(s+10^4)^2}$$

Consider the search space as follows:

$$2 \le x_1 \le 4\ ,\ -1 \le x_2 \le 1\ \ ,\ \ -1 \le x_3 \le 1\ \ ,\ \ 1 \le x_4 \le 3$$

The results are shown in Table (8.12), from which we can observe that the better value of the objective function obtained with ALPSO is due to the violation of the constraint $h_1(x)$; this is not the case at all in our solution for which all constraints are satisfied. Therefore, the best solution found by KA method is significantly better than the solution found by ALPSO. Figure (8.37) shows the convergence curve of the KA optimization for this example.

	x_1	x_2	x_3	x_4	$h_1(x)$	$h_2(x)$	J_{opt}
KA optimization	3.2614	-0.8085	-0.7433	2.3026	-9.1778e-04	-0.0045	-1.7124
ALPSO	3.2548	-0.8424	-0.7501	2.3137	6.1 e-03	-4 e-04	-1.7197

Table (8.12): The numerical results of Example (8.3)

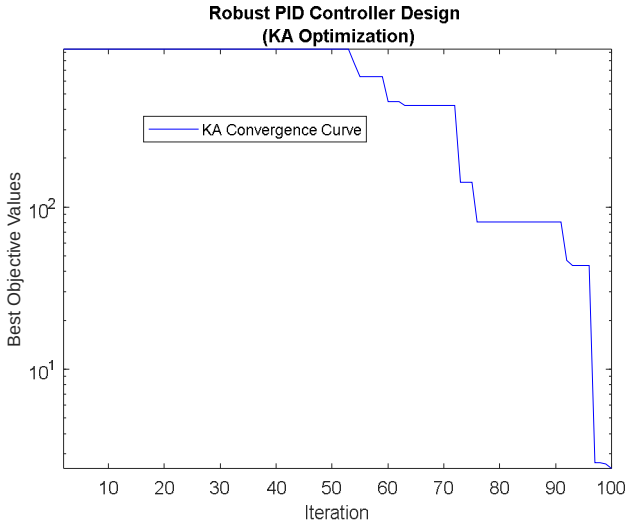


Figure (8.37): Convergence curve of KA in Example (8.3)

□

Example (8.4): Consider the control of the mass-damper-spring system shown in the figure below:

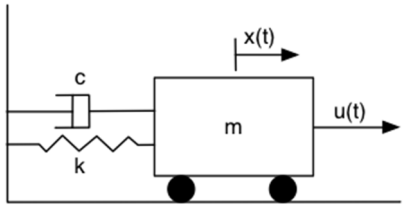


Figure (8.38): Mass-Spring-Damper system

By applying the fundamental principle of dynamics, we get the following equation of motion $m\ddot{x}+c\dot{x}+kx=u$ where x is the position of the mass from its equilibrium position, and u is the control
variable which represents the force applied to the mass. The state space representation of the system is then given by:

$$\begin{cases}\dot{x}=\begin{bmatrix}0 & 1\\ -k/m & -c/m\end{bmatrix}x+\begin{bmatrix}0\\ 1/m\end{bmatrix}u\\ y=\begin{bmatrix}1 & 0\end{bmatrix}x\end{cases}$$

where x_1 is the position and x_2 the velocity of the mass. The parameters m, c and k are not known exactly, but it is assumed that their values are within the following known intervals: $m\in[m_{\min},m_{\max}]$, $c\in[c_{\min},c_{\max}]$ and $k\in[k_{\min},k_{\max}]$. These parameters can be also written as:

$$\begin{cases} m = \underbrace{\dfrac{m_{\max}+m_{\min}}{2}}_{m_o} + \dfrac{m_{\max}-m_{\min}}{2}\delta_m, & \|\delta_m\| \le 1 \\ c = \underbrace{\dfrac{c_{\max}+c_{\min}}{2}}_{c_o} + \dfrac{c_{\max}-c_{\min}}{2}\delta_c, & \|\delta_c\| \le 1 \\ k = \underbrace{\dfrac{k_{\max}+k_{\min}}{2}}_{k_o} + \dfrac{k_{\max}-k_{\min}}{2}\delta_k, & \|\delta_k\| \le 1 \end{cases}$$

The block diagram of the system is shown below:

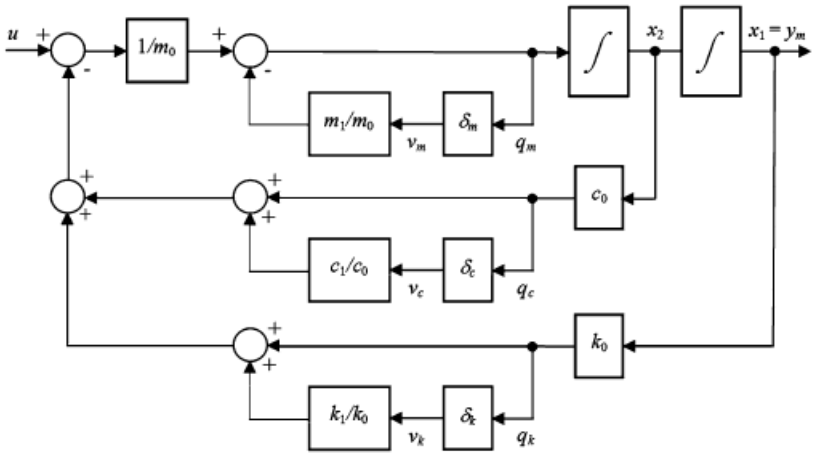


Figure (8.39): Block diagram of the mass-spring-damper system with uncertainty

Then, the corresponding uncertain system can be described as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y \end{bmatrix} = \tilde{G} \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix}$$

$$\tilde{G} \in \Gamma = \left\{ G(\Delta) = F_u(sys, \Delta) : \Delta = \begin{bmatrix} \delta_m & 0 & 0 \\ 0 & \delta_c & 0 \\ 0 & 0 & \delta_k \end{bmatrix}, \|\Delta\|_\infty \le 1 \right\}$$

where the matrices *sys* and Δ are:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ q_m \\ q_c \\ q_k \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -k_o/m_o & -c_o/m_o & -m_d/m_o & -c_d/c_o m_o & -k_d/k_o m_o & 1/m_o \\ -k_o/m_o & -c_o/m_o & -m_d/m_o & -c_d/c_o m_o & -k_d/k_o m_o & 1/m_o \\ 0 & c_0 & 0 & 0 & 0 & 0 \\ k_o & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ v_m \\ v_c \\ v_k \\ u \end{bmatrix}$$

where:

$$\underbrace{\begin{bmatrix} q_m \\ q_c \\ q_k \end{bmatrix}}_{q} = \Delta \underbrace{\begin{bmatrix} v_m \\ v_c \\ v_k \end{bmatrix}}_{v}$$

Hence, the extended or augmented system G_{aug}, which is a perfectly known LTI system is defined as:

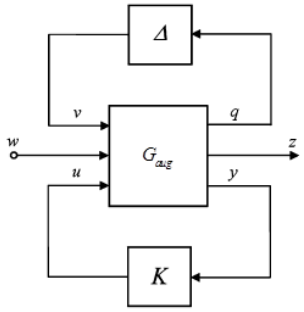


Figure (8.40): The augmented system G_{aug}

$$G_{aug} = \left[\begin{array}{c|ccc} A & B_v & B_w & B_u \\ \hline C_q & D_{qv} & D_{qw} & D_{qu} \\ C_z & D_{zv} & D_{zw} & D_{zu} \\ C_y & D_{yu} & D_{yw} & D_{yu} \end{array}\right]$$

$$A = \begin{bmatrix} 0 & 1 \\ -k_o/m_o & -c_o/m_o \end{bmatrix}$$

$$B_v = \begin{bmatrix} 0 & 0 & 0 \\ -m_d/m_o & -c_d/c_0 m_o & -k_d/k_0 m_o \end{bmatrix},\ B_w = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ B_u = \begin{bmatrix} 0 \\ 1/m_o \end{bmatrix}$$

$$C_q = \begin{bmatrix} -k_o/m_o & -c_o/m_o \\ 0 & c_o \\ k_o & 0 \end{bmatrix},\ C_z = C_y = \begin{bmatrix} -1 & 0 \end{bmatrix}$$

$$D_{qu} = \begin{bmatrix} 1/m_o \\ 0 \\ 0 \end{bmatrix},\ D_{zv} = D_{yv} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix},\ D_{zw} = D_{yw} = 1,\ D_{zu} = D_{yu} = 0$$

Therefore, the multi-objective robust structured control design problem can be stated as follows. For the uncertain system $G \in \Gamma$, defined above, find a robust PID controller K, which minimizes the decay rate while ensuring the robust stability of the closed-loop system. The optimization problem in KA can be formulated based on the minimization of the maximum decay rate as:

$$\begin{cases} \min\limits_{K} \ \alpha \\ \Phi\left(\left[\tilde{G}^{(i)}, K\right]\right) = \max\limits_{1\le j\le n} \mathrm{Re}\left[\lambda_j(A_{cl}(K,\Delta^{(i)}))\right] \\ \max\limits_{1\le i\le \eta}\left\{\Phi\left(\left[\tilde{G}^{(i)}, K\right]\right)\right\} \le \alpha < 0 \\ \tilde{G}^{(i)} \in \left[\tilde{G}^{(1)}, \tilde{G}^{(2)}, \ldots, \tilde{G}^{(\eta)}\right] \\ K \in K_{\tilde{G}} \\ 0 \le K_P \le 150,\ 0 \le K_I \le 150,\ 0 \le K_D \le 150,\ 0.0001 \le \tau \le 0.1 \end{cases}$$

or based on the average decay rate minimization as follows:

$$\begin{cases} \min\limits_{K} \ \alpha \\ \Phi\left(\left[\tilde{G}^{(i)}, K\right]\right) = \max\limits_{1\le j\le n} \mathrm{Re}\left[\lambda_j(A_{cl}(K,\Delta^{(i)}))\right] \\ \eta^{-1}\sum\limits_{i=1}^{\eta}\left\{\Phi\left(\left[\tilde{G}^{(i)}, K\right]\right)\right\} \le \alpha < 0 \\ \tilde{G}^{(i)} \in \left[\tilde{G}^{(1)}, \tilde{G}^{(2)}, \ldots, \tilde{G}^{(\eta)}\right] \\ K \in K_{\tilde{G}} \\ 0 \le K_P \le 150,\ 0 \le K_I \le 150,\ 0 \le K_D \le 150,\ 0.0001 \le \tau \le 0.1 \end{cases}$$

A_{cl} is the state matrix of the nominal closed-loop system and λ_j, $i = 1, \ldots, n$ is the i[th] eigenvalue of A_{cl}. K_G is the set of structured controllers that stabilize the uncertain system $\tilde{G}$. η denotes the number of samples, which can be determined using the inequality below (Toscano, 2013):

$$\eta \ge \ln(1-\rho)/\ln(1-e)$$

where $\rho \in (0,1)$, $e \in (0,1)$ are two given numbers, which state with a confidence ρ, the obtained solution is a probable global minimum of the objective function to level e. $\tilde{G}^{(i)}$ is the sample system with the uncertainty $\Delta^{(i)}$. Table (8.13) shows the results of the worst-case decay rate optimization and the results of the average decay rate optimization are also given in Table (8.14). The obtained results evidence that KA algorithm outperforms the other methods.

	K_P	K_I	K_D	τ	J_{opt}
KA Optimization	150	58.12941	131.5966	0.0007594358	-0.71043
HKA	150	59.14118	132.7368	0.0001095862	-0.62616
GA	148.1950	48.5724	148.6651	0.0064	-0.58759

Table (8.13): Numerical results for worst-case optimization case of Example (8.4)

	K_P	K_I	K_D	τ	J_{opt}
KA Optimization	150	58.6768	132.0986	0.0001009429	-0.7221
HKA	150.0	58.7	130.5	0.01	-0.722
GA	150	58.23111	130.8706	0.009314293	-0.70898

Table (8.14): Numerical results for average performance optimization case of Example (8.4)

□

The next example discusses the problem of optimal distribution of actuators in a three-story building. Let us assume the actuators are the viscous dampers. Viscous dampers greatly reduce the earthquake excitation of a structure and permit it to remain linearly elastic during a seismic event. They work by adding damping to a structure, which significantly lowers its resonant response to an earthquake and reduces stress and deflection because the force from the damping is completely out of phase with stresses. This provides a significant decrease in earthquake excitation (Lee, 2016). By installing as many dampers as necessary, damping can be selectively inserted into the structure at optimal positions. Several dangerous modes are effectively reduced, and resonance effects are eliminated. Therefore, damper locations can be regarded as the problem of maximizing the efficiency taking into account the energy dissipated in dampers as well as minimizing the number of devices, which has great influence on costs in case of retrofitting. (Wu & Ou, 1997), (Takewaki, et al., 1999). Clearly, minimum vibration magnitudes are criteria for the effectiveness (Takewaki, 2000).

Example (8.5): Consider the three-story building presented in (Gluck, et al., 1996), (Takewaki, 1997):

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = -M\,\underline{1}\,\ddot{x}_g + \Gamma u(t)$$

$$M = \begin{bmatrix} 200.4 & 0 & 0 \\ 0 & 200.4 & 0 \\ 0 & 0 & 178 \end{bmatrix} (Kg)\ ,\quad K = 10^3 \times \begin{bmatrix} 238.932 & -119.466 & 0 \\ -119.466 & 238.932 & -119.466 \\ 0 & -119.466 & 119.466 \end{bmatrix} (N/m)$$

$$C = \begin{bmatrix} 264.99 & -78.09 & -16.08 \\ -78.09 & 246.89 & -92.15 \\ -16.08 & -92.15 & 162.02 \end{bmatrix}$$

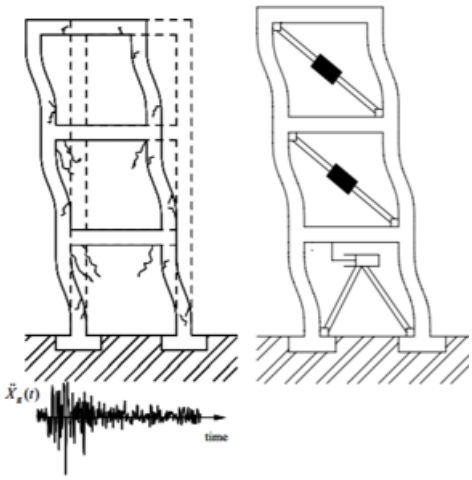


Figure (8.41): Three-story building model with supplemental passive viscous dampers

where x is the n-dimensional displacement vector relative to the ground. M, C, K are the mass, damping and stiffness matrices respectively. Γ is the $n \times m$ distribution matrix for the damper force $u(t)$ and $\underline{1}$ is an n-dimensional distribution vector for the ground acceleration $\ddot{x}_g$, with all elements being unity. m denotes the number of actuators.
An augmented system state equation can be formulated as:

$$\dot{z} = Az + Bu + W$$

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix}, B = \begin{bmatrix} 0 \\ M^{-1}\Gamma \end{bmatrix}$$

Here, $z = \begin{bmatrix} x^T & \dot{x}^T \end{bmatrix}^T$ is the state vector. $W = \begin{bmatrix} 0 & -\underline{1} \end{bmatrix} \ddot{x}_g$ denotes the earthquake induced disturbance.
The response of the building structure, under the effect of random earthquake excitation, largely depends on the system matrix A and the control effort u . In fact, installation of dampers will affect the control gain matrix B . A zero entry of Γ is related to the absence of a damper installed at the corresponding story.
As (Hac & Liu, 1993) states the controllability Gramian $W_c(0, t_f)$ of the above system tends to a diagonal form when the terminal time, t_f , is large enough. Therefore, we have:

$$(W_c)_{ii} = (W_c)_{i+N+N^r, i+N+N^r} = (W_c^u)_{ii}$$

$$= \sum_{j=1}^{m} \frac{b_{ij}^2}{4\zeta_i \omega_i} = \frac{1}{4\zeta_i \omega_i} \sum_{j=1}^{m} b_{ij}^2 \qquad i = 1, \ldots, N$$

and

$$(W_c)_{i+N, i+N} = (W_c)_{i+2N+N^r, i+2N+N^r} = (W_c^r)_{ii}$$

$$= \sum_{j=1}^{m} \frac{(b_{ij}^r)^2}{4\zeta_i^r \omega_i^r} = \frac{1}{4\zeta_i^r \omega_i^r} \sum_{j=1}^{m} (b_{ij}^r)^2 \qquad i = 1, \ldots, N^r$$

where $(W_c^u)_{ii}$, $(W_c^r)_{ii}$ equal to the energy transmitted from the actuators to the structure for the i[th] used and residual eigenmode. N denotes the number of the eigenmodes, which are considered, and N^r is the number of residual eigenmodes.

Hence, according to the previous chapters one approach to determine the actuator placement is to minimize the control energy required by maximizing a measure of the controllability Gramian W_c.

However, one must consider that if one eigenvalue of $(W_c^r)_{ii}$ corresponding to a residual mode is high, the induced spill-over effect can be important. As our objective is to control the modes considered, without exciting residual modes, by transmitting a maximum control force with a minimum energy, and to achieve optimum suppression of quake-induced vibrations, the control problem is cast as a multi-objective optimization process. To manage this in our algorithm, we defined the objective function as a single fitness by the use of importance weights. Thus, the optimization problem can be defined as follows (Arbel, 1981), (Hac & Liu, 1993), (Collet, 2001):

$$\min_b \quad J = \alpha_1 \max_{i=1,\ldots,N^r} \left(W_c^r\right)_{ii} - \alpha_2 \min_{i=1,\ldots,N} \left(W_c^u\right)_{ii} + \alpha_3 \int_0^{t_f} z^T z \, dt$$

Here α_i, $i = 1, 2, 3$ is a weighting constant, z is the state vector of displacements and velocities and $\int_0^{t_f} z^T z \, dt$ reflects upon the desire to minimize the structural response. This optimization problem has m decision variables (the locations of the actuators) and $N + N^r$ optimization functions. This criterion ensures global controllability of the system for the N first eigenmodes and try to minimize the global excitation of the N^r residual modes.

In this example, we considered El Centro earthquake, on soil, with a peak ground acceleration of 0.30g, and the North- East (21º) component recorded during the 1952 Kern County earthquake, Taft, on rock, with a peak ground acceleration of 0.16g.

According to (Bruant & Proslier, 2005), as the components of W_c have not the same range then the above optimization problem becomes, by using the homogeneous components:

$$\min_b \quad J = \alpha_1 \max_{i=1,\ldots,N^r} \frac{\left(W_c^r\right)_{ii}}{\max\limits_{locations} \left(W_c^r\right)_{ii}} - \alpha_2 \min_{i=1,\ldots,N} \frac{\left(W_c^u\right)_{ii}}{\max\limits_{locations} \left(W_c^u\right)_{ii}} + \alpha_3 \int_0^{t_f} z^T z \, dt$$

Hence, the objective function in the KA algorithm can be formulated as:

$$\min_b \quad J = \alpha_1 \max_{i=1,\ldots,N^r} \frac{\sum_{j=1}^{m} (b_{ij}^r)^2}{\max\limits_{locations} \sum_{j=1}^{m} (b_{ij}^r)^2} - \alpha_2 \min_{i=1,\ldots,N} \frac{\sum_{j=1}^{m} b_{ij}^2}{\max\limits_{locations} \sum_{j=1}^{m} b_{ij}^2} + \alpha_3 \int_0^{t_f} z^T z \, dt$$

or (regarding the type of the selected seismic, Elcentro):

$$\min_{b} \; J = \alpha_1 \max_{i=1,\ldots,N^r} \frac{\sum_{j=1}^{m} (b_{ij}^r)^2}{\max\limits_{locations} \sum_{j=1}^{m} (b_{ij}^r)^2} - \alpha_2 \min_{i=1,\ldots,N} \frac{\sum_{j=1}^{m} b_{ij}^2}{\max\limits_{locations} \sum_{j=1}^{m} b_{ij}^2} + \alpha_3 \sum_{k=1}^{number\ of\ floors} \max_{t}(z_k)$$

In this example, for simplicity, we consider that we are given a fixed number of the actuators. We are interested in finding the optimal distribution for the available actuators. The results are shown in Figures (8.43)-(8.52). Figure (8.42) demonstrates the fluctuation in the convergence curve of the optimization due to the multi-objective intrinsic of the problem.
The results of this particular example show that when only one damper is placed this should be located at the first story in order to obtain the best overall performance. When more than one damper is available, the best damper placement is one damper per story; if the number of dampers is less than the number of stories, one damper per story beginning at the lowest story is the best choice.

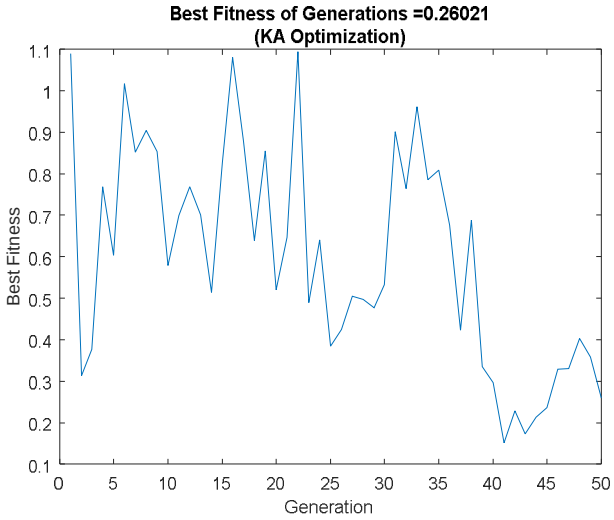


Figure (8.42): The best objective values of Example (8.5) using KA optimization

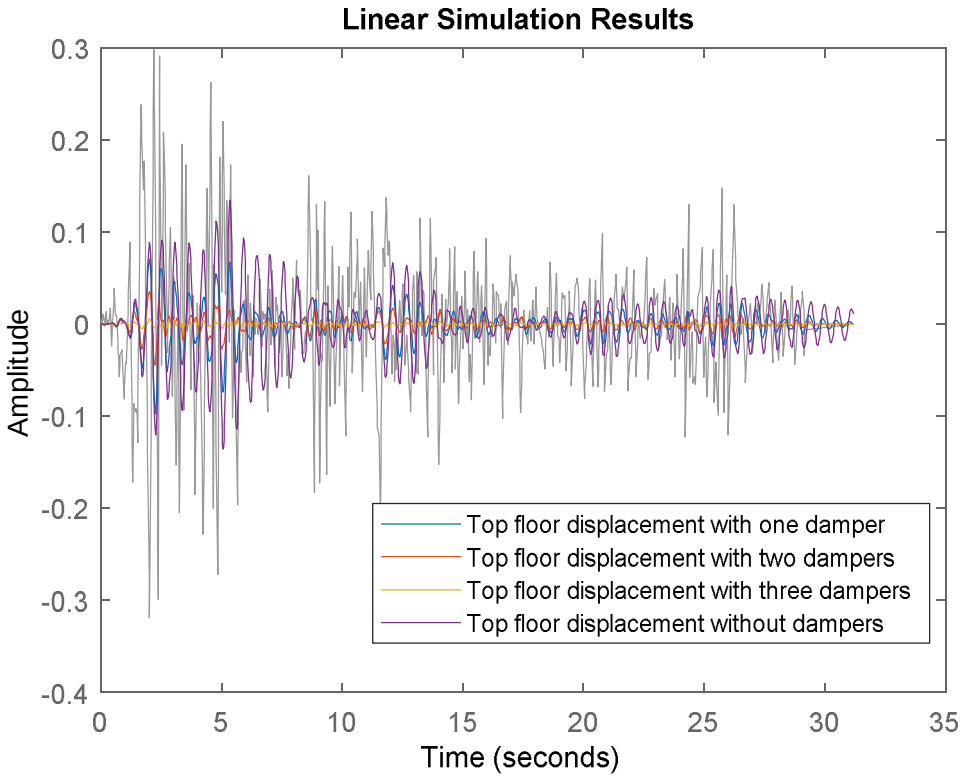


Figure (8.43): Top-floor displacement response of the 3-storey building

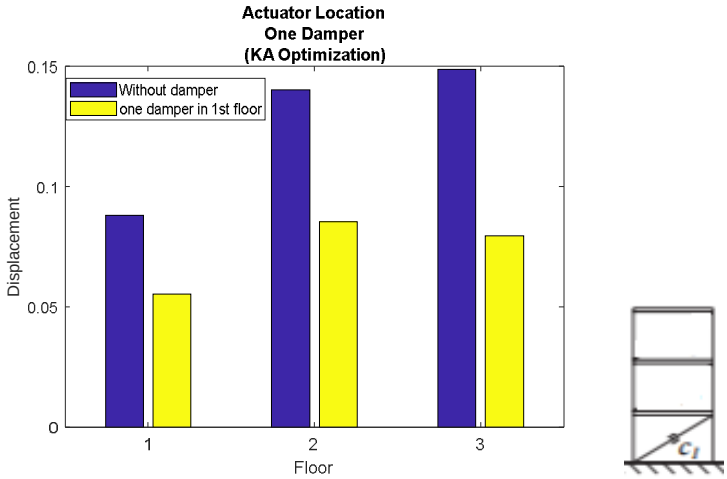


Figure (8.44): Maximum displacement of the floors (One damper)

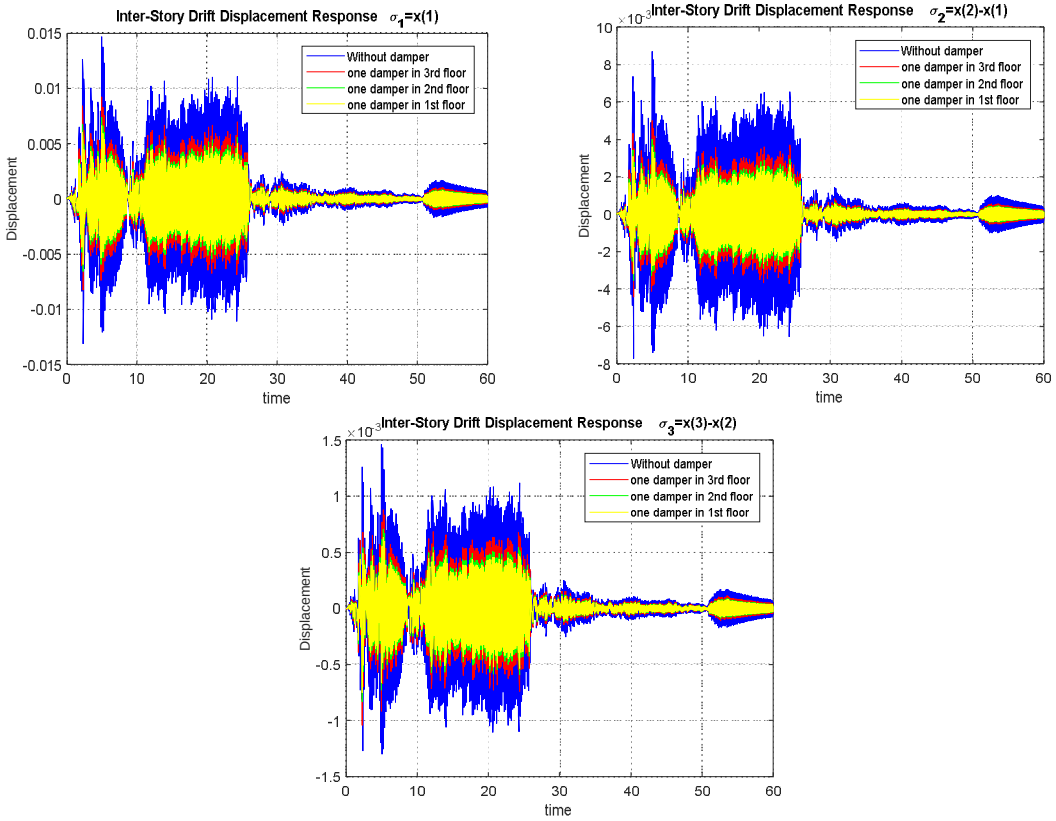


Figure (8.45): Inter-story displacement response of the floors for one damper in different locations

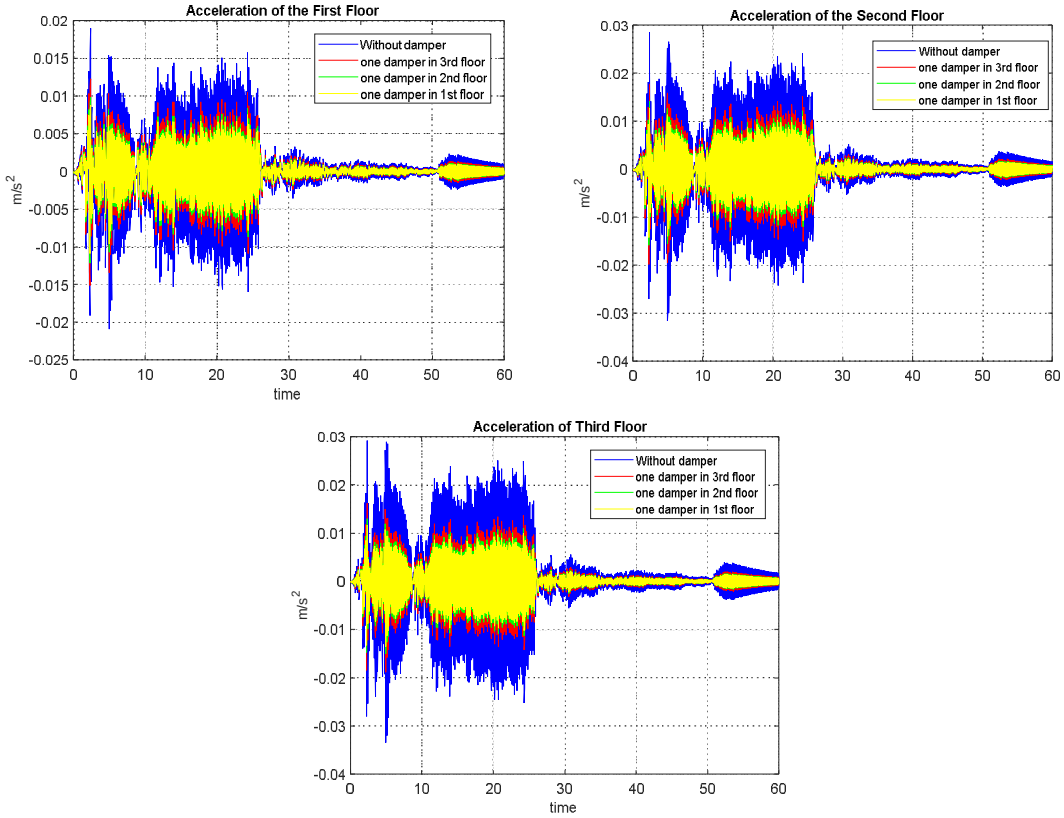


Figure (8.46): Acceleration of the floors using two dampers

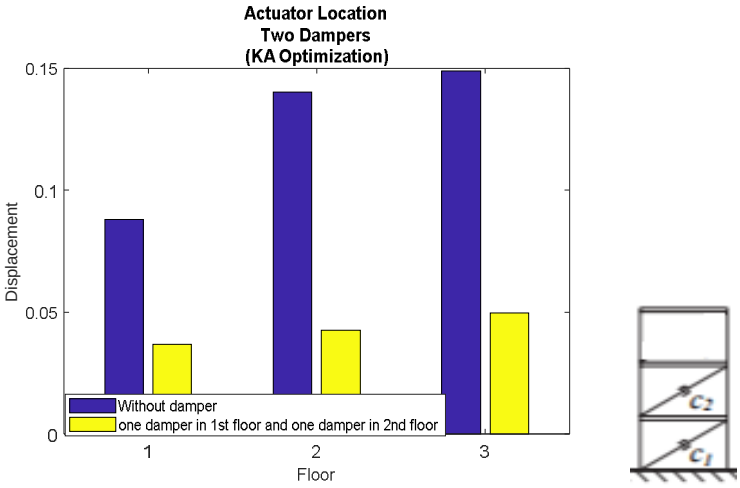


Figure (8.47): Floors' maximum displacement (Two dampers)

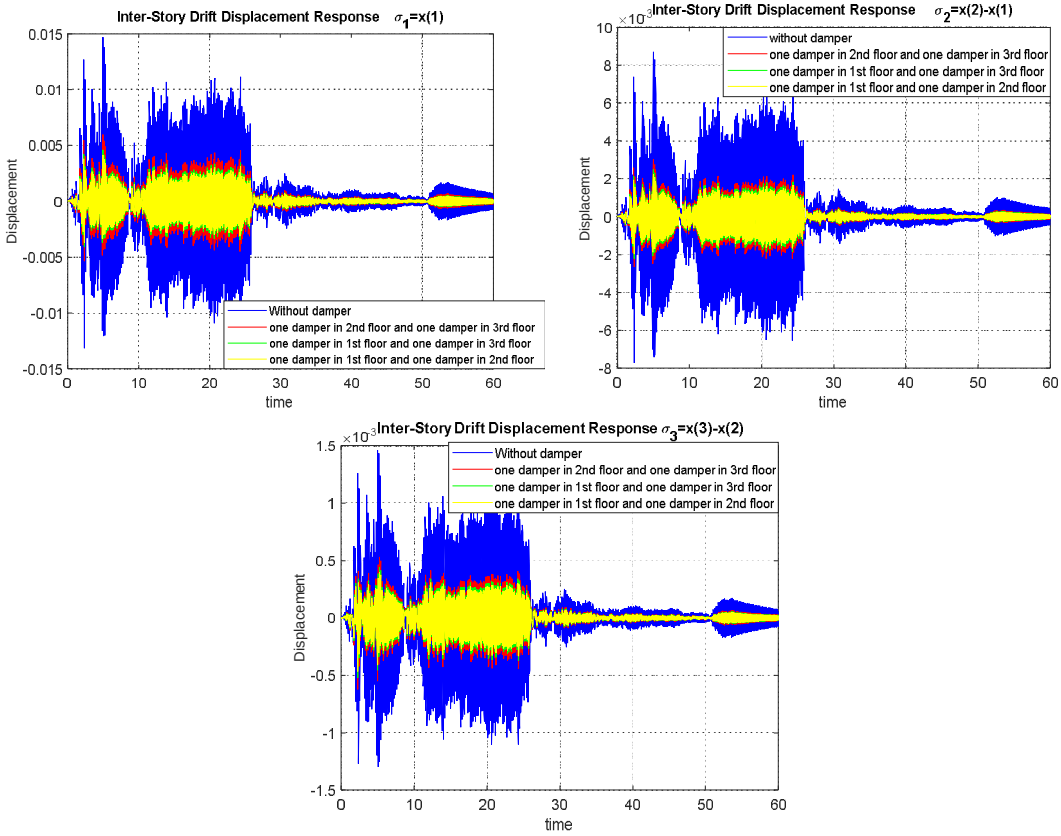


Figure (8.48): Inter-story displacement response of the floors for two dampers in different locations

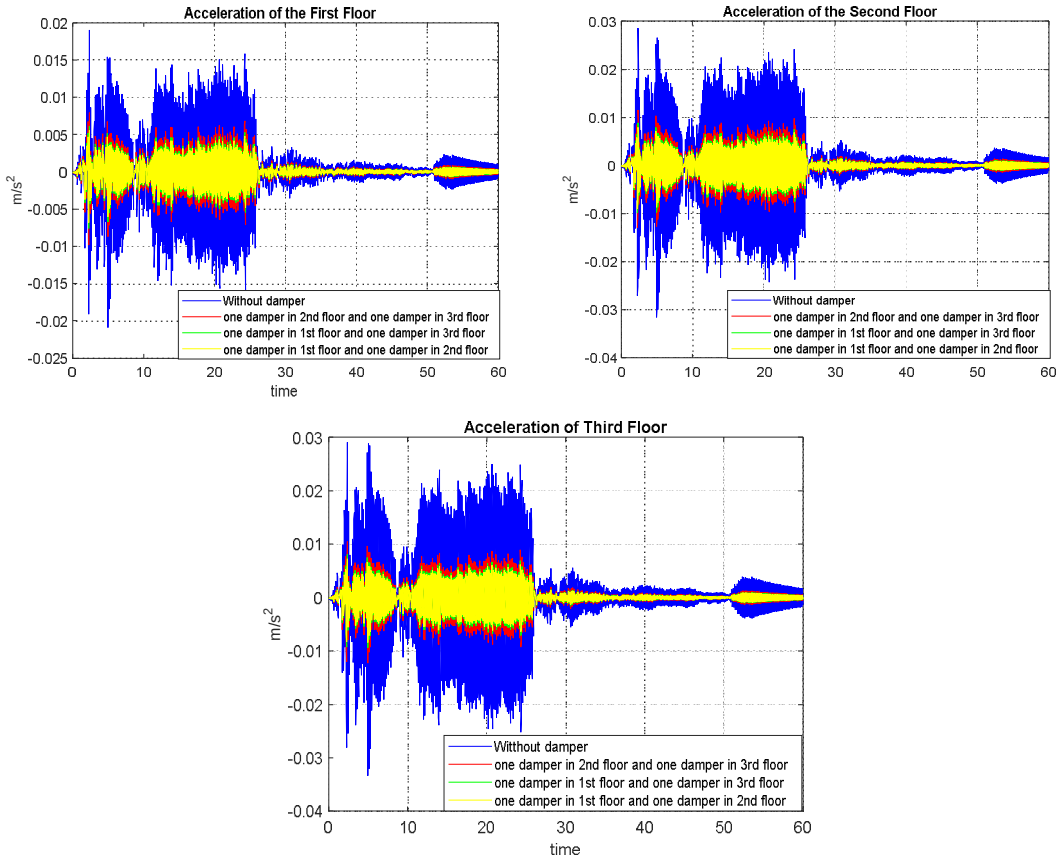


Figure (8.49): Acceleration of the floors using two dampers

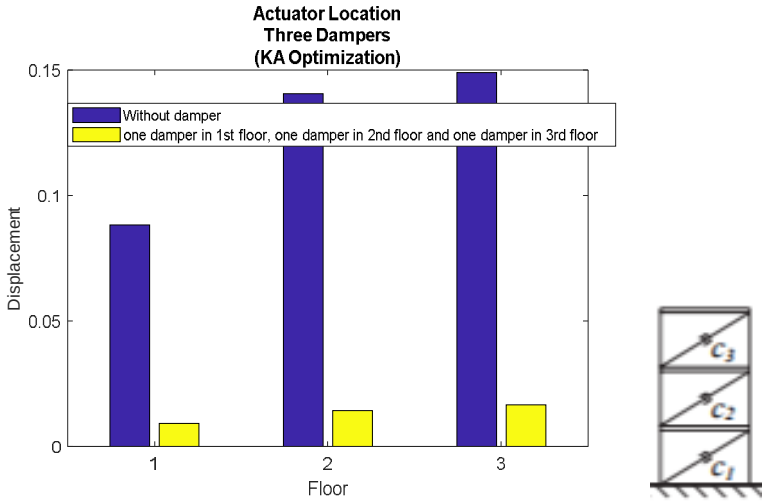


Figure (8.50): Floors' maximum displacement (Three dampers)

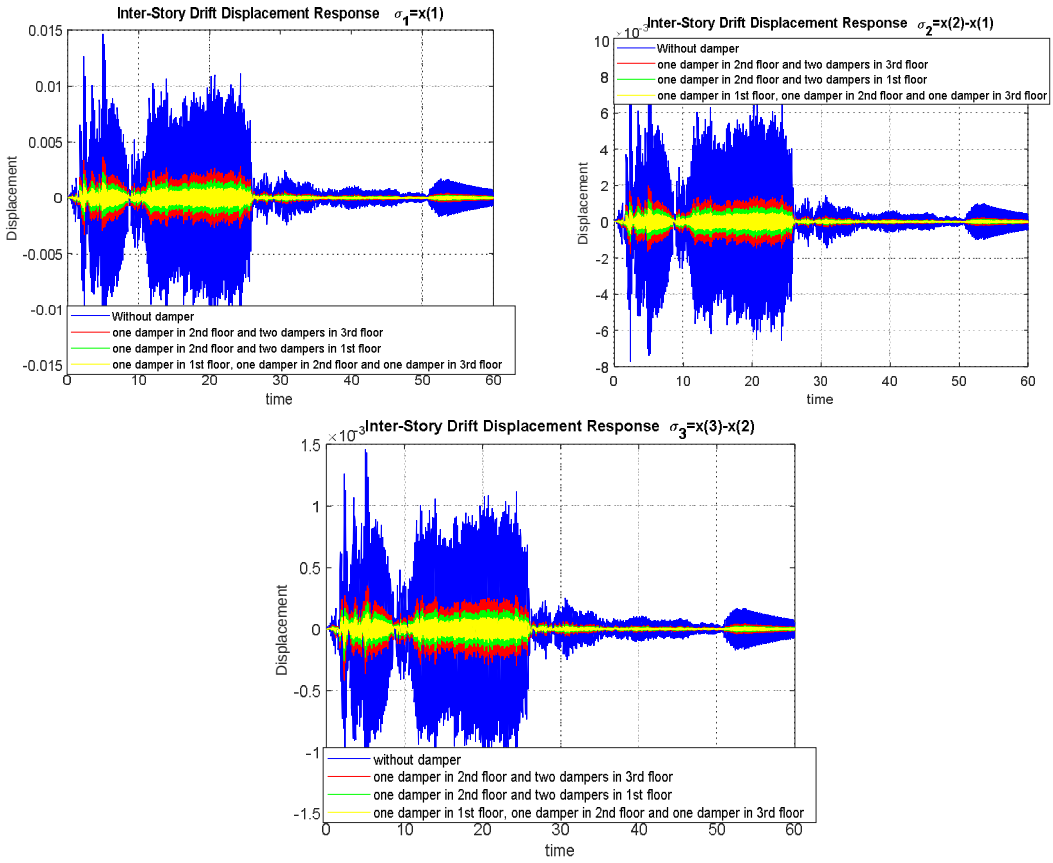


Figure (8.51): Inter-story displacement response of the floors for three dampers in different locations

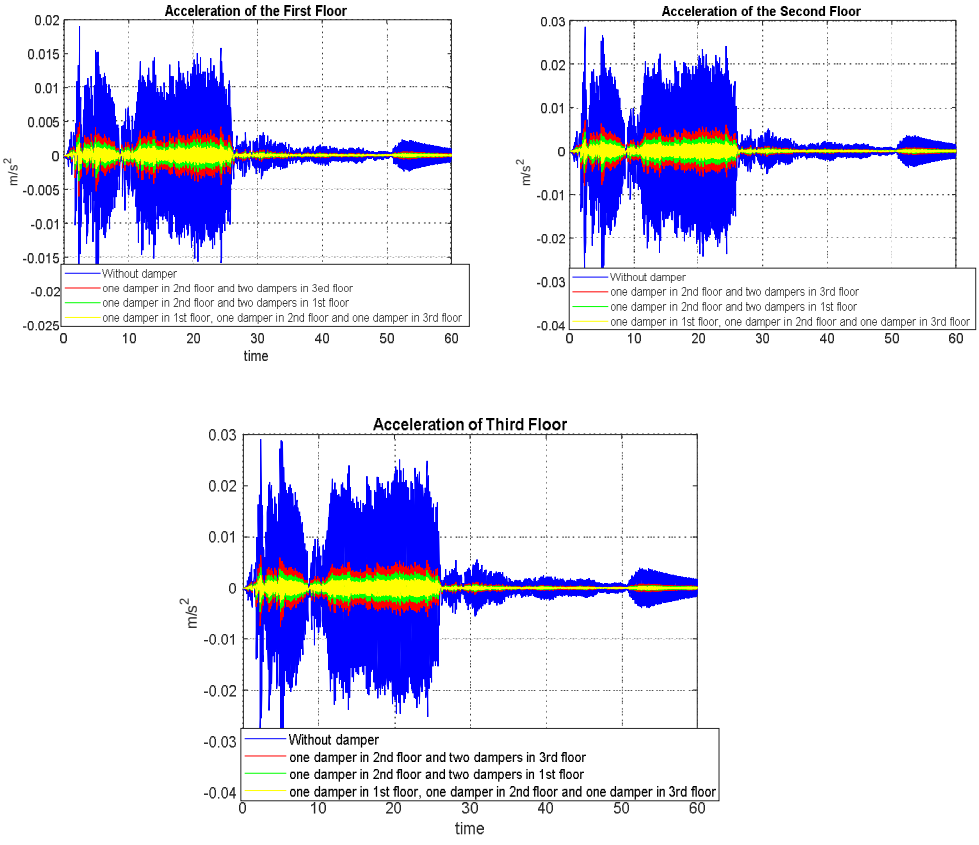


Figure (8.52): Acceleration of the floors using three dampers

□

In the next example, we consider the actuator placement problem for uncertain linear systems, where the aim is to minimize the control input energy to drive the system from an arbitrary initial condition on the unit sphere to the origin in the worst case, i.e. $\min_{B\in\beta} E_{\min}^{\max}$.

$E_{\min}^{\max}$ depicts the maximum value of the minimum input energy and β is the set of all possible input matrices. According to the previous chapters, the above optimization problem can be viewed as the optimization problem of minimizing the inverse of the smallest eigenvalue of the associated controllability Gramian via the placement of an actuator.
In spite of its simple and natural formulation, this problem is difficult to solve.

Example (8.6): (Sojoudi, et al., 2009) , (Sojoudi, et al., 2009 (b)) Consider the uncertain fourth-order LTI system below:

$$A(\alpha_1,\ \alpha_2)=\begin{bmatrix} -0.65 & -0.6 & -1 & 1.5 \\ -0.05 & -0.95 & -1.85 & -0.4 \\ -0.7 & 1.6 & -3.05 & 0.2 \\ -1.5 & -0.1 & -2.85 & -0.35 \end{bmatrix}\alpha_1+\begin{bmatrix} -0.75 & 1.4 & 0 & 2.3 \\ -1.35 & -1.75 & -0.65 & -1.1 \\ 0.8 & 2.2 & -3.25 & -2.5 \\ -0.5 & 1.7 & 0.15 & 0.45 \end{bmatrix}\alpha_2$$

$$B(\alpha_1,\ \alpha_2)=\begin{bmatrix} -1.2 & -0.3 & -0.85 & 1.55 \\ 0.35 & -1.05 & -1.3 & -1.8 \\ 0.25 & 1.2 & -2.5 & -0.8 \\ -0.3 & 1.45 & -1 & -0.3 \end{bmatrix}\alpha_1+\begin{bmatrix} 0.45 & 1.7 & 0.85 & 0.75 \\ -1.7 & -0.7 & 0.65 & 0.7 \\ 0.55 & 1 & -0.75 & -1.7 \\ -0.2 & 0.25 & 1.15 & 0.75 \end{bmatrix}\alpha_2$$

where α_1 and α_2 are the uncertain parameters of the system, which belong to the polytope:

$$S=\{(\alpha_1,\ \alpha_2)\,|\,\alpha_1+\alpha_2=1,\ \alpha_1,\alpha_2\geq 0\}$$

Regard $A(\alpha_1,\alpha_2)$ is robustly Hurwitz over the region S . Assume that four actuators are available, the aim is to find the minimum set of actuators, which play a vital role in controlling the system. To this end, for the given set of actuator $s_{actuator}\subset\{1,\ 2,\ 3,\ 4\}$, we consider four cases as follows:

- {1, 2, 3, 4}
- {1, 2, 3},
- {2, 3, 4},
- {1, 4}

The optimization problem can be formulated using SDP as:

$$\begin{cases} \min\limits_{\alpha_1,\alpha_2}\ x \\ W_c(\alpha_1,\alpha_2)-xI\geq 0 \\ x>0 \\ \alpha_1+\alpha_2=1,\ \alpha_1,\alpha_2\geq 0 \end{cases}$$

Here, $W_c(\alpha_1, \alpha_2)$ denotes the controllability Gramian, which is dependent to the values of the parameters $(\alpha_1, \alpha_2) \in S$.

The optimal results obtained by applying KA algorithm are given in Table (8.15).

	(α_1, α_2)	obj
{1, 2, 3, 4}	0.23665 , 0.25416	0.25311
{1, 2, 3}	0.22288 , 0.15434	0.19853
{2, 3, 4}	0.23629 , 0.4307	0.074542
{1, 4}	0.82059 , 0.172	0.0074269

Table (8.15): Numerical results for Example (8.6)

The results show that the last actuator can be neglected due to its weak contribution to the value of the minimum input energy in the worst case. In contrast, the first actuator is not neglectable as removing it increases the maximum value of the minimum input energy drastically. Moreover, in the case that the second and third actuators are removed concurrently, although the system remains robustly controllable by the remaining inputs, the required control energy would be huge.

The resulting input and state of the system for the interval $t \in [-10, 0]$ are plotted in Figures (8.53)-(8.56), which confirm the results obtained in Table (8.15). One can observe that if the complete set of actuators, i.e. $s_{actuator} = \{1, 2, 3, 4\}$ is selected the worst-case optimal input of the system has the minimum overshoot occurring at $t = 0$ (about 2 in magnitude) while if both actuators 2 and 3 are removed, i.e. $s_{actuator} = \{1, 4\}$, the worst-case optimal input of the system has a very large overshoot (about 10 in magnitude).

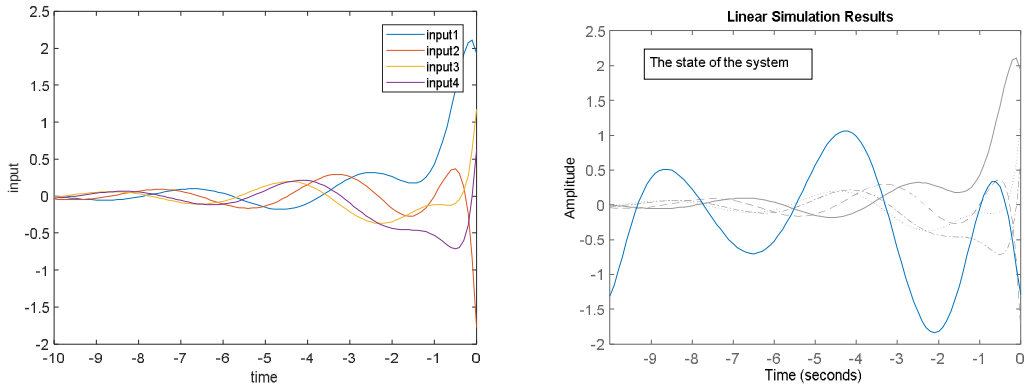


Figure (8.53): The input and state of the system with the actuator set $s_{actuator} = \{1, 2, 3, 4\}$

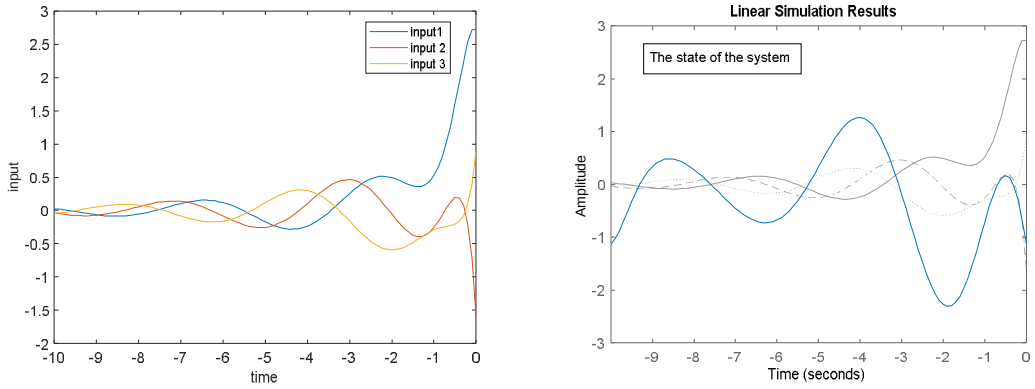


Figure (8.54): The input and state of the system with the actuator set $s_{actuator} = \{1, 2, 3\}$

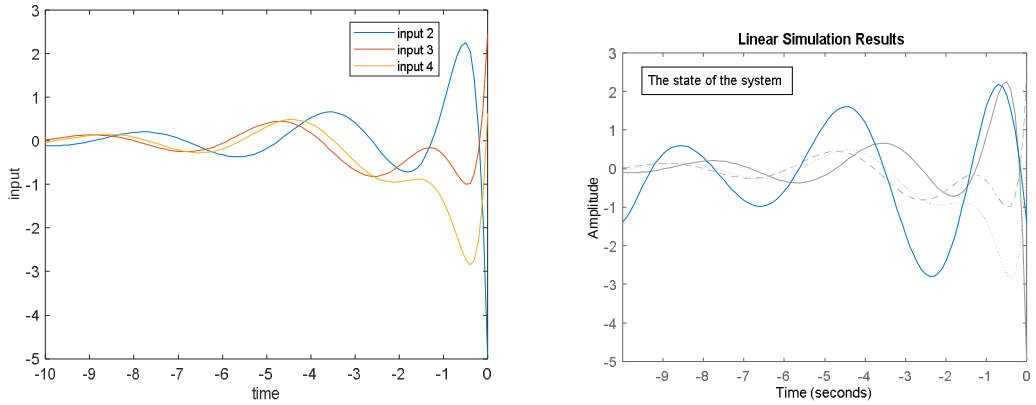


Figure (8.55): The input and state of the system with the actuator set $s_{actuator} = \{2, 3, 4\}$

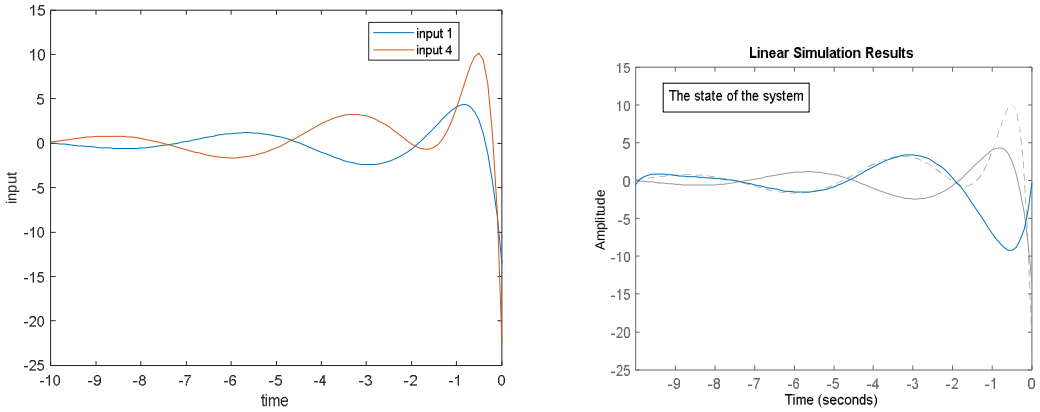


Figure (8.56): The input and state of the system with the actuator set $s_{actuator} = \{1, 4\}$

□

Example (8.7): The purpose of this example is to accomplish the pole placement, finding the optimal location of actuators in a mechanical system to control the vibration response with the minimum control effort.

Consider the open-loop system shown in figure below:

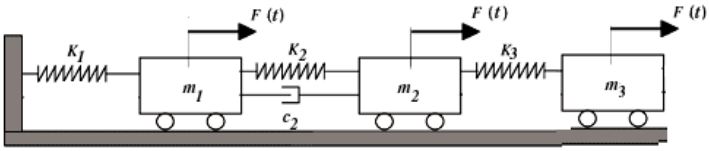


Figure (8.57): The open-loop system of Example (8.7)

Define the mass damping and stiffness matrix for this system as:

$$M=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix},\ C=\begin{bmatrix}0.2 & -0.2 & 0\\ -0.2 & 0.2 & 0\\ 0 & 0 & 0\end{bmatrix},\ K=\begin{bmatrix}15 & -10 & 0\\ -10 & 15 & -5\\ 0 & -5 & 5\end{bmatrix}$$

We wish to assign the eigenvalues of the system to the set:

$$\Lambda=\{\text{-1},\ \text{-3},\ \text{-0.5}\pm 2\text{i},\ \text{-0.75}\pm 5\text{i}\}$$

by using optimal actuation with the minimum control effort. Therefore, the optimization problem can be defined as:

$$\begin{cases}\min\limits_{b} \ \|F\| \\ |A_{cl}-\lambda I|=0, \quad \forall\lambda\in\Lambda \\ \|b\|\le 1\end{cases}$$

where A_{cl} denotes the closed-loop system and F is the associated feedback gain. Note that the force selection vector could theoretically be made very large to allow the position and velocity gain vectors to be very small, with the same control effect. However, physically this would not minimize the actuation needed. To consider this, we defined the last constraint $\|b\|\le 1$. Using the KA optimization algorithm, we obtain:

$$b_{opt}=[0.5828\ \ \text{-0.3996}\ \ 0.7076]^T$$

$$F_{opt}=[\text{-10.8490}\ \ 14.0029\ \ \text{-7.4201}\ \ 2.3196\ \ \text{-2.7915}\ \ \text{-11.4613}]$$

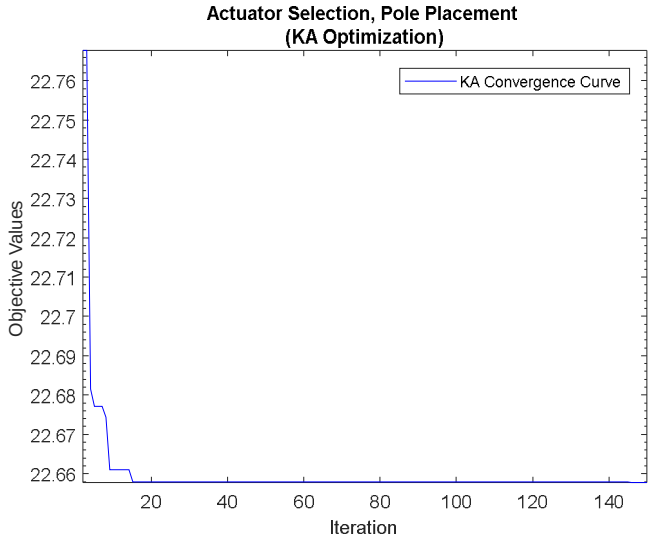


Figure (8.58): The convergence curve of Example (8.7)

Figure (8.59) shows the change of cost function when the entries of the input matrix b change over the interval [-1,1]. The graph of cost function in the neighborhood of the input matrix b associated with the global optimization is also shown in Figure (8.60).

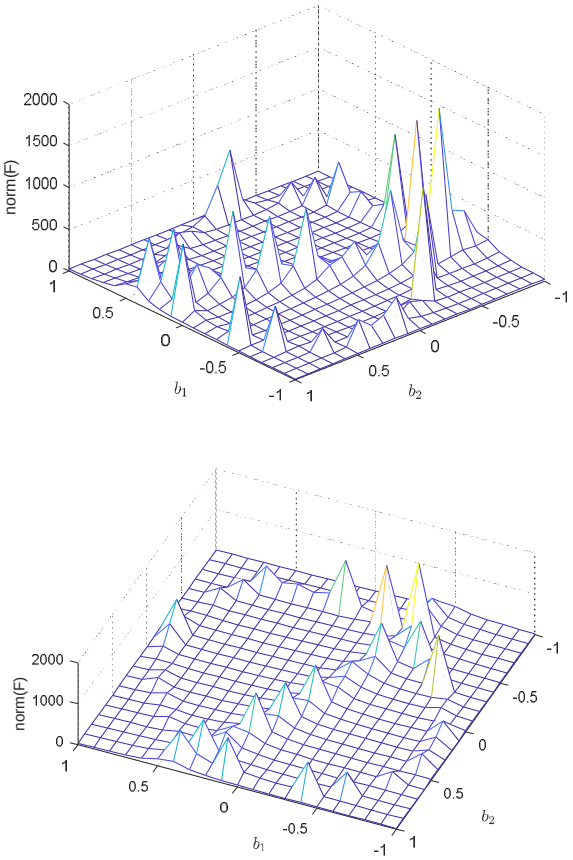


Figure (8.59): The graph of cost function of Example (8.7)

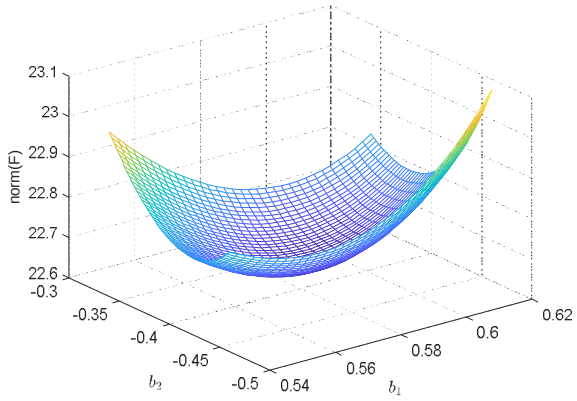


Figure (8.60): Local minimum of Example (8.7)

The validity of the optimization results is investigated using brute force method. The values of b_1, b_2 are continuously changed between -1 and 1 with a step of 0.01 and $b_3 = \sqrt{b_1^2 + b_2^2}$, in

this case, the minimum value of the feedback gain associated with the desired eigenvalues is obtained when the input matrix b is:

$$b = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}^T = \begin{bmatrix} 0.5800 & -0.4000 & 0.7096 \end{bmatrix}^T$$

It yields the feedback gain vector as follows:

$$F = \begin{bmatrix} -10.9032 & 14.0101 & -7.3940 & 2.2987 & -2.8320 & -11.4136 \end{bmatrix}$$

Furthermore, given an initial position and velocity of:

$$x_0 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T, \ \dot{x}_0 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$$

using the optimal input b and the associated gain vector, the control input necessary over time is shown in Figure (8.61). To compare the results, Figure (8.62) demonstrates the control input for some other random selected input matrices:

$$b^1 = \begin{bmatrix} 0.6948 & 0.3171 & 0.9502 \end{bmatrix}^T, \ b^2 = \begin{bmatrix} 0.0344 & 0.4387 & 0.3816 \end{bmatrix}^T$$

$$b^3 = \begin{bmatrix} 0.7655 & 0.7952 & 0.1869 \end{bmatrix}^T, \ b^4 = \begin{bmatrix} 0.4898 & 0.4456 & 0.6463 \end{bmatrix}^T$$

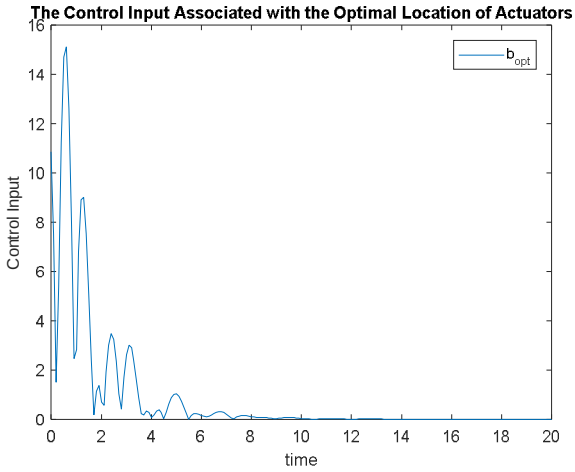


Figure (8.61): Control force input needed to control the system when using the optimal actuation

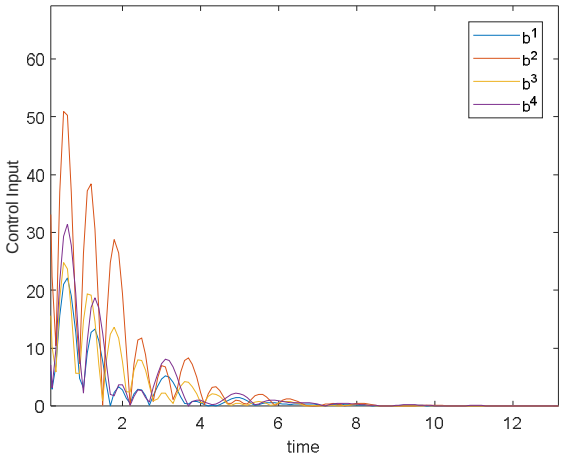


Figure (8.62): Control force input needed to control the system when using the input vector $b^i, i = 1, 2, 3, 4$

□

As it is discussed in chapter 3, many system properties are related to the controllability Gramian. Moreover, the eigenvalues and their corresponding eigenvectors of the controllability Gramian determine the speed and direction of energy dissipation, therefore, it is worth investigating the optimal location of the actuators, which use the minimum feedback gain, assigning a specific controllability Gramian to the LTI system. This assignability problem defines the best possible set of input matrices to steer the LTI system to a target state, under energy limitations imposed by a given controllability Gramian.

Example (8.8): Consider the mass-spring-damper system in the previous example. We apply the KA optimization to find the optimal actuation, which assigns the below controllability Gramian to the system when the control terminal time is very large $t_f \to \infty$:

$$W_c^o = \begin{bmatrix} 0.1372 & 0.1809 & 0.2113 & 0.0000 & -0.0020 & 0.0019 \\ 0.1809 & 0.2567 & 0.3014 & 0.0020 & 0.0000 & 0.0032 \\ 0.2113 & 0.3014 & 0.3926 & -0.0019 & -0.0032 & -0.0000 \\ 0.0000 & 0.0020 & -0.0019 & 0.2489 & 0.1467 & 0.1558 \\ -0.0020 & 0.0000 & -0.0032 & 0.1467 & 0.2737 & 0.2292 \\ 0.0019 & 0.0032 & -0.0000 & 0.1558 & 0.2292 & 0.4616 \end{bmatrix}$$

Let A_{cl}, A_{ol} denote the state matrices of the closed-loop and the open loop system respectively. Using the Lyapunov equation, we obtain:

$$A_{cl}W_c^o + W_c^o A_{cl}^T = -BB^T$$
$$(A_{ol} - BF)W_c^o + W_c^o (A_{ol} - BF)^T = -BB^T$$

Then, we have:

$$A_{ol}W_c^o + W_c^o A_{ol}^T + BB^T = BFW_c^o + W_c^o F^T B^T$$

Using the Kronecker sum, we get:

$$vec(A_{ol}W_c^o + W_c^o A_{ol}^T + BB^T) = \left((W_c^o \otimes B) + (B \otimes W_c^o)\right) vec(F)$$

B is the input matrix, F is the feedback gain matrix, $vec(.)$ is the vectorization operator and $\otimes$ denotes the Kronecker product. Then F can be readily obtained by solving a linear system of equations.

Hence, the assignability problem can be formulated as:

$$\begin{cases} \min\limits_{B} \left\| W_c - W_c^o \right\|_F + \|F\| \\ W_c = -vec^{-1}((A_{cl} \oplus A_{cl})^{-1} vec(BB^T)) \\ \|B\| \le 1 \\ A_{cl} \in H \end{cases}$$

$\|.\|_F$ denotes the frobenious norm, A_{cl} is the state matrix of the closed-loop system and H is the set of Hurwitz matrices with the appropriate dimension. $vec^{-1}(.)$ is de-vectorization operator.

Using the KA optimizer, the optimal actuation is obtained as:

$$B = \begin{bmatrix} 0.5030 & 0.4993 & 0.7055 \end{bmatrix}^T$$

which results in the feedback matrix:

$$F = \begin{bmatrix} 0.0007 & -0.0133 & 0.0050 & -0.4041 & 0.2113 & -0.3628 \end{bmatrix}$$

and produces the controllability Gramian:

$$W_c = \begin{bmatrix} 0.1434 & 0.1881 & 0.2208 & -0.0000 & -0.0026 & 0.0002 \\ 0.1881 & 0.2680 & 0.3148 & 0.0026 & 0.0000 & 0.0023 \\ 0.2208 & 0.3148 & 0.4089 & -0.0002 & -0.0023 & -0.0000 \\ -0.0000 & 0.0026 & -0.0002 & 0.2714 & 0.1419 & 0.1647 \\ -0.0026 & 0.0000 & -0.0023 & 0.1419 & 0.2842 & 0.2362 \\ 0.0002 & 0.0023 & -0.0000 & 0.1647 & 0.2362 & 0.4718 \end{bmatrix}$$

The convergence curve is shown in Figure (8.63). To verify the accuracy of the results, Table (8.16) compares the eigenvalues of the desired controllability Gramian with the eigenvalues of the controllability Gramian produced by the KA optimizer:

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
W_c^o	0.0055	0.0228	0.1002	0.1720	0.7121	0.7582
W_c	0.0066	0.0227	0.1146	0.1795	0.7333	0.7909

Table (8.16): The eigenvalues of the controllability Gramians in Example (8.8)

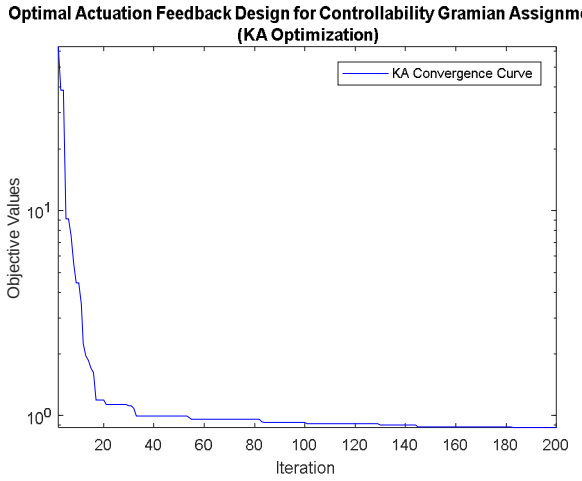


Figure (8.63): The convergence curve of the KA algorithm in Example (8.8)

□

8.4. Summary:

In this chapter, a novel optimization algorithm was proposed as an alternative for solving optimal actuator selection problems among the current techniques in the literature. In the proposed algorithm, the logistic equation like functions fluctuated the solutions outwards or towards the current optimum point via switching amongst the steady state, cyclic and chaotic behavior to guarantee exploration and exploitation of the search space, respectively. A combination of random and adaptive variables also facilitated divergence and convergence of the search agents in the algorithm. To benchmark the performance of the proposed algorithm, a set of well-known test cases including unimodal, fixed- dimensional multi-modal, and varying-dimensional multi-modal functions were employed to test exploration, exploitation, local optima avoidance, and convergence of the KA algorithm for the various convex, non-convex, separable, non-separable cases. To quantitatively and qualitatively verify the performance of the KA algorithm several performance metrics were employed:

- The average and standard deviation of the results, which define the convergence curve
- The average fitness of solutions over the iterations
- The trend of the change of the solutions throughout the iterations
- Non-parametric statistical Wilcoxon rank sum test results

The results of the metrics proved that KA algorithm required the search individuals to change abruptly in the initial stage of optimization and gradually in the final steps. The results showed that this behavior caused exploration of the search space extensively and exploitation of the most promising regions. The convergence curves and the average fitness of solutions also confirmed the improvement of initial random populations and the best current solution obtained during the optimization by the KA algorithm.

Finally, a wide range of challenging case studies is employed to demonstrate the performance of the proposed algorithm in solving real-world problems with constrains. The results demonstrate the merits of KA in solving real problems.

Amongst the case studies, the KA algorithm is used to solve the problem of the optimal actuation in a typical 3-storey building to decrease seismic effects.

The problem of input selection in optimal feedback design is also investigated in two different cases:

- Considering the pole placement
- Assigning the controllability Gramian

The KA algorithm also demonstrates the ability to solve the problem of optimal actuation for the robust DOC in the case of uncertainty.

It can be concluded that the KA algorithm not only is a very suitable alternative compared the current selection methods in the literature for solving optimal actuator problems, but also it can be applied in many different optimization problems. Therefore, the KA algorithm is offered to researchers in different fields.

This chapter opens up several research directions for future studies. The binary version of this algorithm can be proposed to solve problems with binary objectives. The KA algorithm can

also be hybridized with other methods and algorithms in the field of optimization to improve its performance. Finally, investigation of the application of KA in different fields would be a valuable contribution.

Chapter 9: Conclusions & Future Research

9.1. Summary and Contribution of Thesis

This thesis has contributed to the development of methodologies for the selection of the input and output structures, that is selection of the matrix B and the matrix C in a continuous-time LTI system described by the triple $S(A,B,C)$. This has been achieved by considering the implications of such selections as far as the required energy to achieve control and state reconstruction. The main criterion for this selection has been the minimization of the required energy for control and state reconstruction and the notions of relative degree of controllability and observability have been instrumental for the development of the overall methodology. The methodologies developed may be used as design tools where apart from energy requirements other design criteria may be also incorporated for the selection of inputs and outputs. The research undertaken in this thesis is an integral part of the effort to develop an integrated approach to Control and Global Process Instrumentation. The developed methodology has been supported by computational tools for the relevant energy indicators (Gramians) and the development of an optimization based approach that can be used for the design process.

A general research agenda to determine the quality of input and output structure based on fundamental criteria such as performance, stability and so on has been presented, which highlights the significance of the input-output structure selection to meet a certain control objective. The developed research agenda is novel and has introduced new aspects to the design of desirable input and output structures by exploring the state input, state output criteria based on energy and linking them to the

relative degree of controllability and observability. In the case of controllability, the transfer a state from the origin to some point, i.e. to a terminal state location on the boundary of a disc of radius R from the origin has been the key issue, whereas, in the case of observability, it has been the evaluation of the cost of energy to reconstruct the initial state from the distance R from the origin. To address the above challenges properly, it has been required to provide solutions to some specific problems such as:

I. The computation of the energy required for each of the above cases for a given time of transition and a fixed value of R as a function of terminal time and R .
II. Expressing the functions derived in the first problem as a function of the input matrix, output matrix, and then evaluate the quality of the given input-output structure selection, it should be as a function of the input-output matrix.

The main contributions of the thesis have been in the following areas:

- Provided a literature review for link of energy to controllability and observability and a review of methodologies for selection of inputs and outputs. The review on relative measure of controllability and observability properties covers measures such as Kalman: controllability and observability matrices rank tests, gramians: integral definitions (finite time) as well as the algebraic expression (Lyapunov and Sylvester for infinite time). These reviews also include Gramians (controllability, observability and cross gramians) for unstable systems with no poles on the imaginary axis. Additionally, eigenvectors measures, and input and output restriction pencil, which are measures of controllability and observability independent of state feedback.
- Proposed an overview of Controllability/ Observability Gramians, and methodologies for their computation. The proposed approach is based on the integral definitions (finite time) of the controllability Gramian hence it does not imply any restriction on the stability of the system. Another advantage of this method is its simplicity to be used in perturbation analysis and robust control design. The controllability Gramian assignment problem is also discussed.
- Provided a solution of minimum energy, minimum time problems for controllability and observability which has led to the definition of the constant energy surfaces linked to controllability and observability. This has been linked to clarifying how to transfer states from the origin to some point at defined time on the boundary of a disc of radius R along with the transfer energy requirement may be achieved.
- Has produced computational tools for the Energy for unstable systems and measures of degree of controllability/observability. The proposed measures have been linked to the minimum required control input energy to change the state from an arbitrary initial condition to an arbitrary final condition. The improved measure has been introduced by generalizing the Kalman controllability and observability matrices rank tests. A new quantitative measure of the degree of disturbance rejection (DODR) has also been proposed representing how

controllable the system is under the presence of persistent disturbance. The measures are general in the sense that they cover both stable and unstable systems.

- Has developed some new methods for the computation of controllability Gramian in the stable system, which has particular canonical form structures. The value of minimum energy has been evaluated directly via the coefficients of the characteristic polynomial of the system. The trace and some upper bounds for the largest eigenvalue of the controllability Gramian has been derived based on the coefficients of the Routh Hurwitz's table.
- Has introduced a number of partial problems linked to the overall selection of system of inputs, and thus of the input matrix B, to minimize the average required energy for controllability respectively observability. The new proposed approach provided an extension to the existing literature, in which the problem of input structure selection is investigated over the binary set $\{0,1\}$.
- Has developed an optimization method for the optimal selection of input/output system structure for large-scale systems. The proposed optimization algorithm exploring and exploiting the search space in order to find efficiently near-optimal solution of the problem of IO selection. We focused on the problem of cardinality-constrained actuator placement for minimum control effort, where the optimal actuator set is selected so that an input energy metric is minimized. While this is an NP-hard combinatorial problem, our proposed algorithm has proved its efficiency in converging for optimal solution.

9.2. Future Directions

The research agenda developed for the selection of the input and output structures has a number of problems which have not yet been handled in this thesis and remain open. A number of topics that require further research are:

- Energy methodology for the selection of the input, output structure is based on energy requirements. A related issue is the gain requirements for feedback design which leads to an alternative problem formulation. The link between energy and gain requirements is not well understood and this may also form the basis for further research. Of particular interest is exploring the link between energy and gain restrictions.
- Linked to the above investigation is exploring the relations between relative degrees of controllability and observability and the restrictions on the gains for state feedback design, state reconstruction respectively.
- The computational tools based on optimization need further development. The associated optimization problem is non-convex and computing approximate solutions in an efficient way is still challenging.
- The selection of the input and output structures has been considered so far as unconstrained problems. In many design cases, certain variables have to be used for control, measurement respectively and thus constrained versions emerge which are linked to constrained optimization problems.
- Two challenging theoretical problems that require attention are linked to introducing restricted notions of controllability and observability under energy, or gain constraints.

- The basic ideas developed for linear systems can be transferred with minor modification to the case of nonlinear systems. This is achieved by generating an efficient problem relaxation based on the steady state. The results of this thesis can also be extended to some other types of the system, e.g. discrete- time systems, time variant systems.
- when the parameters of the system are subject to perturbation in order to find a dominant subset of inputs and outputs for any given large-scale system, it is needed to determine the effectiveness of each input and output in the overall operation of the control system. Thus, the investigating the robust controllability (observability) degree of the system is worth pursuing.

References

1) Abdei-Rohman, M., 1984. Design of optimal observers for structural control. *IEE Proceedings D - Control Theory and Applications ,* 131(4), pp. 158-163.

2) Abdel-Mooty, M. & Roorda, J., 1994. Optimal configuration of active-control mechanisms. *Journal of engineering mechanics,* 120(3), pp. 535-556.

3) Abdullah, M., Richardson, A. & Hanif, J., 2001. Placement of sensors/actuators on civil structures using genetic algorithms. *Earthquake Engineering and Structural Dynamics,* Volume 30, pp. 1167-1184.

4) Adorio, E. & Dilman, U., 2005. *MVF - Multivariate Test Function Library in C for Unconstrained Global Optimization Methods.* [Online]
Available at: http://www.geocities.ws/eadorio/mvf.pdf

5) Agathoklis, P. & Sreeram, V., 1990. Identification and model reduction from impulse response data. *International Journal of Systems Science ,* 21(8), pp. 1541-1552.

6) Ahmadi, A.A & Parrilo, P.A., 2009. *Sum of Squares and Polynomial Convexity.* s.l., 48th IEEE Conference on Decision and Control.

7) Ali, M., Khompatraporn, C. & Zabinsky, Z., 2005. A numerical evaluation of several stochastic algorithms on selected continuous global optimization test problems. *Journal of Global Optimization,* 31(4), p. 635–672.

8) Ali, Y. & Narasimhan, S., 1996. Sensor Network Design for Maximizing Reliability of Bilinear Processes. *AIChE,* 42(9).

9) Ali, Y. & Narsimhan, S., 1993. Sensor network design for maximizing reliability of linear processes. *AIChE,* Volume 39, p. 821.

10) Ali, Y. & Narsimhan, S., 1995. Redundant Sensor Network Design for Linear Processes. *AIChE,* Volume 41, pp. 2237-2249.

11) Alvarez, J., Romangnoli, J. & Stephanopoulous, G., 1981. Variable measurements structures for the control of a tubular reactor. *Chemical Engineering Science,* 36(10), pp. 1695-1712.

12) Amato, F., Tommasi, G. D. & Pironti, A., 2015. *Necessary and Sufficient Conditions for Input-Output Finite-Time Stability of Impulsive Dynamical Systems.* Chicago, IL, USA, 2015 American Control Conference.

13) Ambrosio, P., Resta, F. & Ripamonti, F., 2012. An H2-norm approach for the actuator and sensor placement in vibration control of a smart structure. *Smart Materials and Structures,* 21(12), p. 125016.

14) Antoniades, C. & Christo, P., 2002. Integrated optimal actuator/sensor placement and robust control of uncertain transport-reaction processes. *Computers and Chemical Engineering,* Volume 26, pp. 187-203.

15) Antoulas, A., 2004. *Approximation of large-scale dynamical systems.* Philadelphia: SIAM Book series "Advances in Design and Control".

16) Antsaklis, P. & Michel, A., 2007. *A Linear Systems Primer.* illustrated, reprint ed. Birkhäuser Boston: Springer Science & Business Media.

17) Arabyan, A. & Chemishkian, S., 1998. *H∞*-optimal mapping of actuators and sensors in flexible structures. s.l., *Proceedings of the IEEE Conference on Decision and Control,* pp. 821-826.

18) Arbel, A., 1981. Controllability measures and actuator placement in oscillatory system. *International Journal of Control ,* 33(3), pp. 565-574.

19) Arora, J., 2016. *Introduction to optimum design.* 4 ed. New York : Elsevier Science Academic Press.

20) Assali, W., 2008. *Optimal Selection of Measurements and Manipulated Variables for Production Control.* University of Maryland, College Park, MD: Ph.D. Thesis.

21) Bagajewicz, M., 1997. Design and retrofit of sensor networks in process plants. *AIChE,* 43(9), pp. 2300-2306.

22) Bagajewicz, M., 2000. *Process Plant Instrumentation: Design and Upgrade.* illustrated ed. Lancaster, Pennsylvania, USA: CRC Press.

23) Bagajewicz, M. & Sanchez, M., 1999. Design and Upgrade of Nonredundant and Redundant Linear Sensor Networks. *AIChE,* 45(9), pp. 1927-1938.

24) Balchen, J. & Mummé, K., 1988. New York: Von Nostrand Reinhold.

25) Barnett, S. & Storey, C., 1968. Some Applications of the Lyapunov Matrix Equation. *J. Inst. Maths Applies ,* Volume 4, pp. 33-42.

26) Barnett, S. & Storey, C., 1970. *Matrix Methods in Stability Theory.* illustrated ed. the University of Michigan: Thomas Nelson & Sons Ltd.

27) Belegundu, A., 1982. *A study of mathematical programming methods for structural optimization.* Iowa, USA: Ph.D. thesis, Department of Civil and Environmental Engineering, University of Iowa.

28) Bellman, R., 2003. *Dynamic programming.* paperback edition ed. New York. Dover: Princeton University Press.

29) Benner, P. & Damm, T., 2011. Lyapunov Equations, Energy Functionals, and Model Order Reduction of Bilinear and Stochastic Systems. *SIAM J. Control Optim,* 49(2), p. 686–711.

30) Benqlilou, C., Graells, M., Muslin, E. & Puigjaner, L., 2004. Desig and Retrofit of Reliable Sensor Networks. *Ind. Eng. Chem. Res.,* 43(25), pp. 8026-8036.

31) Berat Doğan, , Tamer Ölmez , 2015. A new metaheuristic for numerical function optimization: Vortex Search algorithm. *Information Sciences,* Volume 293, p. 125–145.

32) Bergh, F. v. d. & Engelbrecht, A., 2006. A study of particle swarm optimization particle trajectories. *Information Sciences,* 176(8), pp. 937-971.

33) Bhushan, M., Narasimhan, S. & Rengaswamy, R., 2008. Robust Sensor Network Design for Fault Diagnosis. *Computers & Chemical Engineering,* Volume 32, pp. 1067-1084.

34) Bhushan, M. & Rengaswamy, R., 2000b. Design of sensor location based on various fault diagnostic observability and reliability criteria. *Computers and Chemical Engineerin,* Volume 24, pp. 735-741.

35) Bhushan, M. & Rengaswamy, R., 2000. Design of sensor network based on the SDG of the process for efficient fault diagnosis. *Industrial & Engineering Chemistry Research,* 39(4).

36) Bittanti, S. & Colaneri, P., 2009. *Periodic Systems: Filtering and Control.* Illustrated ed. s.l.:Springer Science & Business Media.

37) Blekherman, G., Parrilo, P. A. & Thomas, R. R., 2013. *Semidefinite Optimization and Convex Algebraic Geometry.* illustrated ed. s.l.:SIAM.

38) Blum, Christian; Roli, Andrea, 2003. Metaheuristics in Combinatorial Optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR),* 35(3), pp. 268-308.

39) Boley, D., 1987. Computing rank-deficiency of rectangular matrix pencials. *IEEE Conference Decision and Control,* pp. 207-214.

40) Boley, D. & Lu, W., 1986. Measuring how far a controllable system is from an uncontrollable one. *IEEE Trans. Automat. Contr,* Volume AC-31, pp. 249-251.

41) Botev, Z. et al., 2004. *the GLOBAL library at The Cross-Entropy Toolbox.* [Online] Available at: http://www.maths.uq.edu.au/CEToolBox/

42) Boukhobza, T. & Hamelin, F., 2011. Observability analysis and sensor location study for structured linear systems in descriptor form with unknown inputs. *Automatica,* 47(12), p. 2678–2683.

43) Brookes, M., 2004. *Matrix Reference Manual by Mike Brookes.* [Online] Available at: http://www.psi.toronto.edu/matrix/matrix.html [Accessed 20 May 2004].

44) Brown, R., 1966. Not Just Observable, But How Observable?. *National Electronic Conf,* Volume 22, pp. 709-714.

45) Bruant, I., Gallimard, L. & Nikoukar, S., 2010. Optimal piezoelectric actuator and sensor location for active vibration control, using genetic algorithm. *Journal of Sound and Vibration,* 329(10), pp. 1615-1635.

46) Bruant, I. & Proslier, L., 2005. Optimal Location of Actuators and Sensors in Active Vibration Control. *Journal of Intelligent Material Systems and Structures,* 16(3), pp. 197-206.

47) Burgmeier, P., 1992. Operations Research '91-Extended Abstracts of the 16th Symposium on Operations Research held at the University of Trier at September 9–11,1991. In: *Degrees of Controllability.* Germany: Physica-Verlag HD, pp. 182-185.

48) Bushnell, L., Clark, A. & Poovendran, R., 2014. A supermodular optimization framework for leader selection under link noise in linear multiagent systems. *IEEE Transactions on Automatic Control,* 59(2), p. 283–296.

49) Byers, R., 1989. Detecting nearly uncontrollable pairs. *Proceedings of the International Symposium MTNS-89,* pp. 447-457.

50) Cao, Y. & Biss, D., 1996. An extension of singular value analysis for assessing manipulated variable constraints. *Journal of Process Control,* 6(1), pp. 37-48.

51) Cao, Y. & Rossiter, D., 1996. Input Screening Method for Disturbance Rejection. *Proceedings of Control '96, CP427, United Kingdom, Automatic Control Council, Exeter, England, u.K.,* pp. 497-502.

52) Cao, Y., Rossiter, D. & Owens, D., 1997. Input Selection for Disturbance Rejection Under Manipulated Variable Constraints. *Computers & Chemical Engineering,* 21(Supplement), pp. S403-S408.

53) Caruso, A., Galeani, S. & Menini, L., 2003. *On sensor/actuator placement for collacated flexible plate.* s.l., IEEE Mediterranean Conference on Control and Automation .

54) Castelan, E., Silva, J. G. d. & Cury, J., 1996. A reduced-order framework applied to linear systems with constrained controls. *IEEE Transactions on Automatic Control,* 41(2), pp. 249-255.

55) Chemishkian, S. & Arabyan, A., 1999. Intelligent algorithms for *H*∞-optimal placement of actuators and sensors in structural control. s.l., *Proceedings of the American Control Conference, IEEE*, pp. 1812-1816.

56) Chen, B., 1995. A simple algorithm for the stable/unstable decomposition of a linear discrete-time system. *International Journal of Control ,* 61(1), pp. 255-260.

57) Chen, B., Lin, Z. & Shamash, Y., 2004. *Linear Systems Theory: A Structural Decomposition Approach.* illustrated ed. Boston: Springer Science & Business Media.

58) Chen, C., 1984. *Linear System Theory and Design.* New York: Holt, Rinehart, and Winston.

59) Chen, D., Zheng, S. & Wang, H., 2012. Genetic algorithm based LQR vibration wireless control of laminated plate using photostrictive actuators. *Earthquake Engineering and Engineering Vibration,* 11(1), pp. 83-90.

60) Cheng, F. & Pantelides, C., 1988. *Optimal placement of actuators for structural control,* Buffalo, N.Y.: National Center for Earthquake Engineering Research, State Univ. of New York at Buffalo.

61) Chen, K. & Rowley, C., 2011. H2-optimal actuator and sensor placement in the linearised complex Ginzburg–Landau system. *Journal of Fluid Mechanics,* Volume 681, pp. 241-260.

62) Chen, R., 2002. *An Optimal Control Based Plantwide Control Design Methodology and its Applications.* University of Maryland, College Park, MD: Ph.D. Thesis, .

63) Chen, W. et al., 2012. Particle swarm optimization with an aging leader and challengers. *IEEE Transactions on Evolutionary Computation,* 17(2), pp. 241-258.

64) Chesi, G., 2011. SOS Polynomials. In: *Domain of Attraction.* London: Springer London, pp. 3-44.

65) Chmielewski, D., Palmer, T. & Manousiouthakis, V., 2002. On the theory of optimal sensor placement. *AIChE,* 48(5), pp. 1001-1012.

66) Choe, K. & Baruh, H., 1992. Actuator Placement in Structural Control. *Journal of Guidance, Control, and Dynamics,* 15(1), pp. 40-48.

67) Clark, A., Alomair, B., Bushnell, L. & Poovendran, R., 2014. Minimizing convergence error in multi-agent systems via leader selection: A supermodular optimization approach. *IEEE Transactions on Automatic Control,* 59(6), p. 1480–1494.

68) Coello, C., 2000. Use of a self-adaptive penalty approach for engineering optimization. *Computers in Industry,* 41(2), pp. 113-127.

69) Coello, C. & Montes, E., 2002. Constraint-handling in genetic algorithms through the use of dominance-based tournament selection. *Advanced Engineering Informatics,* 16(3), pp. 193-203.

70) Colantuoni, G. & Padmanabhan, L., 1977. Optimal sensor location for tubular-flow reactor systems. *Chemical Engineering Science,* 32(9), pp. 1035-1049.

71) Collet, M., 2001. Shape optimization of piezoelectric sensors dealing with spill-over instability. *IEEE Transactions on Control Systems Technology,* 9(4), pp. 654-662.

72) Commault, C. & Dion, J., 2013. Input addition and leader selection for the controllability of graph-based systems. *Automatica,* Volume 49, pp. 3322-3328.

73) Cortesi, F., Summers, T. & Lygeros, J., 2014. Submodularity of energy related controllability metrics. *Los Angeles, CA, IEEE Conference on Decision and Control,* p. 2883–2888.

74) Cowan, N. et al., 2012. Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PLOS ON,* 7(6), pp. 1-6.

75) Cuevas, E. et al., 2012. An Algorithm for Global Optimization Inspired by Collective Animal Behavior. *Discrete Dynamics in Nature and Society,* Volume 2012, pp. 1-24.

76) D. Henrion; A. Garull, 2005. *Positive Polynomials in Control.* Berlin: Springer Science & Business Media.

77) Davis, J., Gravagne, I., Marks, R. & Ramos, A., 2010. Algebraic and dynamic Lyapunov equations on time scales. *IEEE System Theory (SSST).*

78) Deb, K., 1997. GeneAS: A robust optimal design technique for mechanical component design. In: D.Dasgupta & Z. Michalewicz, eds. *Evolutionary Algorithms in Engineering Applications.* Berlin: Springer, pp. 497-514.

79) de-los-Cobos-Silva, S. et al., 2015. An Efficient Algorithm for Unconstrained Optimization. *Mathematical Problems in Engineering,* Volume 2015, pp. 1-17.

80) Demenkov, M. & Goman, M., 2009. *Suppressing Aeroelastic Vibrations via stability region maximization and numerical continuation techniques.* University of Manchester, UKACC International Conference on Control, p. https://ukacc.group.shef.ac.uk/proceedings/control2008/Proceedings.html#We07_04.

81) Demetriou, M., 2000. *Numerical algorithm for the optimal placement of actuators and sensors for flexible structures.* Chicago, IL, USA, USA, American Control Conference.

82) Demetriou, M., 2004. Integrated actuator/sensor placement and hybrid controller design of flexible structures under worst case spatiotemporal disturbance variations. *Journal of Intelligent Material Systems and Structures,* 15(12), pp. 901-921.

83) Demetriou, M. & Grigoriadist, K., 2004. Collocated actuator placement in structural systems using an analytical bound approach. s.l., *Proceeding of the 2004 American Control Conference,* pp. 1604-1609.

84) Demetriou, M. & Grigoriadist, K., 2007. Utilizing spatial robustness measures for the optimization of a pzt-actuated flexible beam. In: D. K. Lindner, ed. *Modeling, Signal Processing, and Control for Smart Structures 2007.* s.l.:Proc. of SPIE Vol. 6523, 65230N.

85) Devasia, S., Meressi, T., Paden, B. & Bayo, E., 1992. *Piezoelectric actuator design for vibration suppression: placement and sizing.* Tucson, AZ, USA, USA, Proceedings of the 31st IEEE Conference on Decision and Control.

86) Devasia, S., Meressi, T., Paden, B. & Bayo, E., 1993. Piezoelectric actuator design for vibration suppression: placement and sizing. *Journal of Guidance Control and Dynamics,* 16(5), p. 859–864.

87) Dhuri, K. & Seshu, P., 2006. Piezo actuator placement and sizing for good control effectiveness and minimal change in original system dynamics. *Smart Materials and Structures,* Volume 15, p. 1661–1672.

88) Dion, J., Commault, C. & Woude, J. d., 2002. Characterization of generic properties of linear structured systems for efficient computations. *Kybernetika,* 38(5), p. 503–520.

89) Dion, J., Commault, C. & Woude, J. v. d., 2003. Generic properties and control of linear structured systems: a survey. *Automatica,* 39(7), pp. 1125-1144.

90) Dochain, D., Tali-Mammar, N. & Babary, J., 1997. On modeling,monitoring and control of fixed bed bioreactors. *Computers & Chemical Engineering,* 21(11), pp. 1255-1266.

91) Doyle, J., Glover, K., Khargonekar, P. & Francis, B., 1989. State-space solutions to standard H2 and Hinf control problems.. *IEEE Transactions on Automatic Control,* 34(8), pp. 831-847.

92) Eising, R., 1984. Between controllable and uncontrollable. *Systems & Control Letters,* 4(5), pp. 263-264.

93) Eising, R., 1984. The distance between a system and the set of uncontrollable systems. In: *Mathematical Theory of Networks and Systems-Proceedings of the MTNS-83 International Symposium Beer Sheva, Israel, June 20–24, 1983.* s.l.:Springer Berlin Heidelberg, pp. 303-314.

94) Esmat Rashedi, Hossein Nezamabadi-pour, Saeid Saryazdi, 2009. GSA: A Gravitational Search Algorithm. *Information Sciences,* 179(13), p. 2232–2248.

95) Fabrizio L. Cortesi, Tyler H. Summers, and John Lygero, 2014. *Submodularity of Energy Related Controllability Metrics.* s.l., s.n.

96) Fahroo, F., 1995. Optimal location of controls for an acoustic problem. *New Orleans, LA, USA, USA, Proc. of the 34th IEEE Conference on Decision and Control*, p. 3765 –3766.

97) Fahroo, F. & M.A. Demetriou, 2000. Optimal actuator/sensor location for active noise regulator and tracking control problems. *Journal of Computational and Applied Mathematics,* 114(1), pp. 137-158.

98) Fairman, F., 1998. *Linear Control Theory: The State Space Approach.* illustrated, reprint ed. Kingston, Ontario, Canada: John Wiley & Sons.

99) Farina, L. & Rinaldi, S., 2011. *Positive Linear Systems: Theory and Applications.* annotated ed. s.l.:John Wiley & Sons.

100) Findeisen, W. et al., 1980. *Control and Coordination in Hierarchical Systems.* s.l.:John Wiley & Sons.

101) Finkelstein, L. & Grattan, K.T.V., 1994. *Concise Encyclopedia of Measurement and Instrumentation (Advances in Systems Control and Information Engineering).* s.l.:Pergamon.

102) Foss, C., 1973. Critique of chemical process control theory. *AIChE Journal,* 19(2), pp. 209-214.

103) Frecker, M., 2003. Recent advances in optimization of smart structures and actuators. *Journal of Intelligent Material Systems and Structures,* Volume 14, p. 207–216.

104) Friedland, B., 1986. *Control system design: an introduction to state-space methods.* New York: McGraw-Hill.

105) Fuzhen, Z., 2005. *The Schur Complement and Its Applications.* US: Springer.

106) Gahinet, P. & Laub, A., 1992. Algebraic Riccati equations and distance to the nearest uncontrollable pair. *SIAM J. Contr. Optim,* 30(4), pp. 765-786.

107) Gamal Abd El-Nasser A. Said, Abeer M. Mahmoud, El-Sayed M. El-Horbaty, 2014. A Comparative Study of Meta-heuristic Algorithms for Solving Quadratic Assignment Problem. *(IJACSA) International Journal of Advanced Computer Science and Applications,* 5(1).

108) Garc´ıa, S., Molina, D., Lozano, M. & Herrera, F., 2009. A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the CEC'2005 special session on real parameter optimization. *Journal of Heuristics,* 15(6), p. 617–644.

109) Gaviano, M., Kvasov, D., Lera, D. & Sergeyev, Y., 2003. Software for generation of classes of test functions with known local and global minima for global optimization. *Association for Computing Machinery. Transactions on Mathematical Software,* 29(4), p. 469–480.

110) Gawronski, W., 1998. *Dynamics and Control of Structures: A Modal Approach.* New York: Springer-Verlag.

111) Georges, D., 1995. *The Use of Observability and Controllability Gramians or Functions.* New Orleans, LA, IEEE, pp. 3319 - 3324.

112) Georges, D., 1995. *The use of observability and controllability gramians or functions for optimal sensor and actuator location in finite-dimensional systems.* New Orleans, LA, USA, USA, Proceedings of the 34th IEEE Conference on Decision and Control.

113) Georgiou, A. & Floudas, C., 1990. Structural properties of large scale systems. *Int. J. Control,* Volume 51, pp. 169-187.

114) Georgiou, A. & Floudas, C. A., 1989. Structural analysis and synthesis of feasible control systems: Theory and applications. *Chemical Engineering Research and Design,* Volume 67, pp. 600-618.

115) Georgiou, A. & Floundas, C., 1989. Structural Analysis and Synthesis of feasible Control Systems: Theory and Applications. *Chem. Eng. Res. Des.,* Volume 67, pp. 600-618.

116) Geromel, J., 1989. Convex analysis and global optimization of joint actuator location and control problems. *IEEE Transactions on Automatic Control,* 34(7), pp. 711-720.

117) Gluck, N., Reinhorn, A., Gluck, J. & Levy, R., 1996. Design of Supplemental Dampers for Control of Structures. *Journal of Structural Engineering, ASCE,* 122(12), pp. 1394-1399.

118) Golub, G. H. & Loan, C. V., 1986. *Matrix Computations.* London: North Oxford Academic publ.

119) Gould, S. H., 1957. *Variational Methods for Eigenvalue Problems: An Introduction to the Methods of Rayleigh, Ritz, Weinstein, and Aronszajn.* Mathematical expositions ed. s.l.:University of Toronto Press.

120) Govind, R. & G.J. Powers, 1982. Control Systems Synthesis Strategies. *AICHE,* Volume 28, pp. 60-73.

121) Gugercin, S. & Antoulas, A.C., 2003. On balancing related model reduction methods and the corresponding error. *Int. Journal of Control.*

122) Guney, M. & Eskinat, E., 2008. Optimal actuator and sensor placement in flexible structures using closed-loop criteria. *Journal of Sound and Vibration,* Volume 312, pp. 210-233.

123) Guzman, V., Cruz, M. & A.R.S.Ortigoza, 2016. Linear State Feedback Regulation of a Furuta Pendulum: Design Based on Differential Flatness and Root Locus. *IEEE Access,* Volume 4, pp. 8721 - 8736.

124) Guzzella, L., 2011. *Analysis and Synthesis of Single-Input/Single-Output Control Systems.* revised ed. s.l.:vdf Hochschulverlag AG.

125) Hac, A. & Liu, L., 1993. Sensor and actuator location in motion control of flexible structures. *Journal of Sound and Vibration,* 167(2), p. 239–261.

126) Han, J. & Lee, I., 1997. Active Damping Enhancement of Composite Plates with Electrode Designed Piezoelectric Materials. *Journal of Intelligent Material Systems and Structures,* 8(3), pp. 249-259.

127) Han, J. & Lee, I., 1999. Optimal placement of piezoelectric sensors and actuators for vibration control of a composite plate using genetic algorithms. *Smart Mat. Struct.,* 8(2), p. 257–267.

128) Harris, T., Macgregor, J. & Wright, J., 1980. Optimal sensor location with an application to a packed bed tubular reactor. *AIChE,* 26(6), pp. 910-916.

129) Hauksdóttir, A. & Sigurðsson, S., 2009. *The continuous closed form controllability Gramian and its inverse.* Hyatt Regency Riverfront, St. Louis, MO, USA, 2009 American Control Conference .

130) Hauksdóttir, A. et al., 2008. Closed Form Solutions of the Sylvester and the Lyapunov Equations - Closed Form Gramians. *Seattle, The 2008 American Control Conference*, pp. 2585-2590.

131) Hautus, M., 1969. Controllability and observability conditions of linear autonomous systems. *Nederl. Akad. Wetensch. Proc., Ser., A72,* pp. 443-448.

132) Hautus, M., 1970. Stabilization controllability and observability of linear autonomous systems. *Indagationes Mathematicae (Proceedings),* Volume 73, pp. 448-455.

133) Hebrard, P. & Henrot, A., 2005. A spillover phenomenon in the optimal location of actuators. *SIAM Journal on Control and Optimization,* 44(1), pp. 349-366.

134) He, Q. & Wang, L., 2007. An effective co-evolutionary particle swarm optimization for constrained engineering design problems. *Engineering Applications of Artificial Intelligence,* 20(1), pp. 89-99.

135) Hendricks, E. & Jannerup, O. & Sørensen, P., 2008. *Linear Systems Control: Deterministic and Stochastic Methods.* illustrated ed. Berlin: Springer Science & Business Media.

136) Hespanha, J., 2009. *Linear Systems Theory.* annotated ed. s.l.:Princeton University Press.

137) Hiramoto, K., Doki, H. & Obinata, G., 2000. Optimal sensor/actuator placement for active vibration control using explicit solution of algebraic Riccati equation. *Journal of Sound and Vibration,* 229(5), pp. 1057-1075.

138) Holland, J., 1975. *Adaptation in Natural and Artificial Systems.* First edition ed. s.l.:Cambridge, MA: MIT Press.

139) Horn, R. & Johnson, C., 2012. *Matrix Analysis.* illustrated, revised ed. s.l.:Cambridge University Press.

140) Hovd, M. & Skogestad, S., 1992. Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design. *Automatica,* 28(5), pp. 989-996.

141) Ho, Y. & Pepyne, D., 2002. Simple explanation of the no-free-lunch theorem and its implications. *J Optim Theory Appl,* 115(3), p. 549–570.

142) Hu, G. & Davison, E., 2001. *A Real Radius Measure for Controllability.* Arlington, VA, s.n., pp. 3144-3148.

143) Hughes, P. & Skelton, R., 1980. Controllability and observability for flexible spacecraft. *Journal of Guidance, Control, and Dynamics,* 3(5), pp. 452-459.

144) Jain, A., 1989. Block Matrices and Kronecker Products. In: *Fundamentals of digital image processing.* Englewood Cliffs: Prentice Hall.

145) Jamil, M. & Yang, X., 2013. A literature survey of benchmark functions for global optimization problems. *Journal of Mathematical Modelling and Numerical Optimisation,* 4(2), p. 150–194.

146) Jha, A. & Inman, D., 2003. Optimal sizes and placements of piezoelectric actuators and sensors for an inflated torus. *Journal of Intelligent Material Systems and Structures,* 14(9), p. 563–576.

147) Jirstrand, M. & Gunnarsson, J., 2001. *The Mathematica Journal-Code generation for simulation and control applications.* [Online]
Available at: http://www.mathematica-journal.com/issue/v8i2/features/codegeneration/contents/html/Links/index_lnk_2.html
[Accessed 2001].

148) Johnson, C., 1969. Optimization of a certain quality of complete controllability and observability for linear dynamical systems. *ASME Trans J. Basic Eng,* 91(2), pp. 228-238.

149) Johnston, R. & Barton, G. W., 1985. Control system development without dynamic simulation. *IChemE Symposium Series No. 92,* pp. 443-456.

150) Jorgensen, S., Goldschmidt, L. & Clement, K., 1984. A sensor location procedure for chemical processes. *Computers & Chemical Engineering,* 3(3-4), pp. 195-204.

151) Juang, J. & Rodriguez, G., 1979. *Formulations and applications of large structure actuator and sensor placements.* Blacksburg, VA, VPI&SUIAIAA Symp., on Dynamics and control of large flexible spacecraft.

152) Junior,]., Silva, R., Mundim, K. & Dardenne, L., 2004. Performance and Parameterization of the Algorithm Simplified Generalized Simulated Annealing. *Genet. Mol. Biol.,* 27(4), pp. 616-622.

153) K Ramesh Kumar and S Narayanan, 2008. Active vibration control of beams with optimal placement of piezoelectric sensor/actuator pairs. *IOPScience, Smart Materials and Structures,* 17(5).

154) Kailath, T., 1980. *Linear Systems.* illustrated ed. The University of Michigan: Prentice-Hall Information and System Sciences Series.

155) Kalman, R., 1959. On the general theory of control systems. *IRE Transactions on Automatic Control,* 4(3), pp. 110-110.

156) Kalman, R., 1963. Mathematical description of linear Dynamical systems. *SIAM J. Control Optim. ,* Volume 1, pp. 152-192.

157) Kalman, R. E., Ho, T. C. & Narendra, K. S., 1963. Controllability of linear dynamical systems. *Contributions to Differential Equations,* 1(2), p. 189–213.

158) Kang, O. & Park, Y., 2009. New Measure Representing Degree of Controllability for Disturbance Rejection. *Journal of Guidance, Control, and Dynamics,* 32(5), pp. 1658-166l.

159) Kannan, B. & Kramer, S., 1994. An augmented Lagrange multiplier based method for mixed integer discrete continuous optimization and its applications to mechanical design. *Journal of Mechanical Design,* 116(2), pp. 405-411.

160) Karcanias, N., 1994. Global Process Instrumentation: Issues and Problems of a System and Control Theory Framework. *Measurement,* 14(1), pp. 103-113.

161) Karcanias, N., 1994. The selection of input and output schemes for a systems and the Model Projection Problems. *KYBERNETICA,* 30(6), pp. 585-596.

162) Karcanias, N., 1996. Control Problems in Global Process Instrumentation : A structural Approach. *6th European Symposium on Computer Aided Process Engineering (ESCAPE-6) Rhodes and Computers & Chemical Engineering,* Volume 2, pp. 1101-S1106.

163) Karcanias, N., 1996. Control problems in global process instrumentation: A structural approach. *Computers Chem. Engng,* Volume 20, pp. 1101-1106.

164) Karcanias, N., 2002. Multivariable Poles and Zeros. *Control Systems, Robotics and Automation from Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http://www.eolss.net], [Retrieved October 26,2005],* Volume VII.

165) Karcanias, N., 2008. Structure Evolving Systems and Control in Integrated Design. *IFAC Annual Reviews in Control,* 32(2), pp. 161-182.

166) Karcanias, N. & Giannakopoulos, C., 1989. Necessary and sufficient conditions for zero assignment by constant squaring down. *Linear Algebra and its Applications,* Volume 122-124, pp. 415-446.

167) Karcanias, N. & Giannakopoulos, C., 1989. Necessary and Sufficient Conditions for Zero Assignment by Constant Squaring Down. *Linear Algebra and Its Applications,* Volume 122-124, pp. 415- 446.

168) Karcanias, N. & Vafiadis, D., 2002. Canonical Forms for State Space Descriptions. *Control Systems, Robotics and Automation, from Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http://www.eolss.net], [Retrieved October 26,2005],* Volume VII.

169) Karcanias, N. & Vafiadis, K., 2002. Derivation of Effective Transfer Function Models By Input-Output Variables Selection. *KYBERNETICA,* Volume 38, pp. 657-683.

170) Karcanias, N. & Vafiadis, K., 2002. *Model Orientation and Well Conditioning of System Models: System and Control Issues.* Barcelona, 15th IFAC World Congress.

171) Kaveh, A., 2016. *Advances in Metaheuristic Algorithms for Optimal Design of Structures.* 2, illustrated ed. Gewerbestrasse, Switzerland: Springer.

172) Kaveh, A. & Talatahari, S., 2010. An improved ant colony optimization for constrained engineering design problems. *Engineering Computations. Int J Comput Aided Eng.,* 27(1), p. 155–182.

173) Kaveh, A. & Talatahari, S., 2010. Optimal design of skeletal structures via the charged system search algorithm. *Structural and Multidisciplinary Optimization,* 41(6), p. 893–911.

174) Kennedy, J. & Eberhart, R., 1995. Particle Swarm Optimization. *IEEE International Conference on Neural Networks,* Volume IV, p. 1942–1948.

175) Kenney, C. & Laub, A., 1988. Controllability and stability radii for companion form systems. *Mathematics of Control, Signals and Systems,* 1(3), p. 239–256.

176) Khare, S., Pillai, H. & Belur, M., 2012. Computing the radius of controllability for state space systems. *Elsevier Systems & Control Letters,* pp. 327-333.

177) Khemliche, M., Bouamama, B. & Haffaf, H., 2006. Sensor placement for component diagnosability using bond-graph. *Sensors and Actuators A: Physical,* 132(2), pp. 547-556.

178) Kienitz, K., 2012. *Controllability and Observability of Linear Dynamic Systems Revisited.* Barcelona, IEEE.

179) Kim, Y. & Junkinsf, J., 1991. Measure of Controllability for Actuator Placement. *American Institute of Aeronautics and Astronautics,* 14(5), pp. 895-902.

180) Klamka, J., 2009. System Characteristics: Stability, Controllability, Observability. In: *Control System, Robotics and Automation-Volume VII: Advanced Control Systems-I.* s.l.:EOLSS Publications, pp. 232-247.

181) Kondoh, S., Yatomi, C. & Inoue, K., 1990. The Positioning of Sensors and Actuators in the Vibration Control of Flexible Systems. *JSME international journal. Ser. 3, Vibration, control engineering, engineering for industry,* 33(2), pp. 145-152.

182) Kościelny, J., Sztyber, A. & Syfert, M., 2017. *Graph description of the process and its applications.* s.l.: In: Mitkowski W., Kacprzyk J., Oprzędkiewicz K., Skruch P. (eds) Trends in Advanced Intelligent Control, Optimization and Automation. KKA 2017. Advances in Intelligent Systems and Computing, vol 577. Springer, Cham.

183) Koumboulis, F. & Mertzios, B., 1999. On Kalman's controllability and observability criteria for singular systems. *Circuits, Systems and Signal Processing,* 18(3), pp. 269-290.

184) Kouvaritakis, B. & MacFarlane, A., 1976. Geometric approach to analysis and synthesis of system zeros; Part II: non-square systems. *Int. J. Control,* Volume 23, pp. 167-181.

185) Kumar, K. & Narayanan, S., 2007. The optimal location of piezoelectric actuators and sensors for vibration control of plates. *Smart Materials and Structures,* 16(6), pp. 2680-2691.

186) Kumar, K. R. & Narayanan, S., 2008. Active vibration control of beams with optimal placement of piezoelectric sensor/actuator pairs. *Smart Materials and Structures,* 17(5), pp. 1-15.

187) Kumar, S. & Seinfeld, S., 1978. Optimal location of measurements in tubular reactors. *Chemical Engineering Science,* 33(11), pp. 1507-1516.

188) Lam, A. S., Li, V. K. & Yu, J. Q., 2012. Real-Coded Chemical Reaction Optimization. *IEEE Transactions on Evolutionary Computation* 16(3), pp. 339-353.

189) Lambert, H., 1977. Fault Trees for Locating Sensors in Process Systems. *Chemical Engineering Progress,* Volume 81.

190) Lam, T. Y., 2005. *Introduction to Quadratic Forms Over Fields.* illustrated ed. s.l.:American Mathematical Soc.

191) Larsson, T. & Skogestad, S., 2000. Plantwide control—A review and a new design procedure. *Modeling, Identification and Control,* 21(4), pp. 209-240.

192) Laub, A., 2005. *Matrix Analysis for Scientists and Engineers.* s.l.:SIAM: Society for Industrial and Applied Mathematics.

193) Ibidapo-Obe, O., 1985. Optimal actuators placements for the active control of lexible structures. *Journal of Mathematical Analysis and Applications,* 105(1), pp. 12-25.

194) Lee, D., 2016. *Taylor Devices, Inc.*. [Online]
Available at: http://www.taylordevices.com/fluidviciousdamping.html

195) Lee, E. & Markus, L., 1967. *Foundations of Optimal Control Theory.* New York: John Wiley and Sons.

196) Lee, H. & Park, Y., 2014. Degree of controllability for linear unstable systems. *Journal of Vibration and Control,* pp. 1-7.

197) Lee, H. & Park, Y., 2014. *Degree of Disturbance Rejection Capability for Linear Anti-Stable Systems.* Seoul, South Korea, IEEE, 14th International Conference on Control, Automation and Systems (ICCAS), pp. 154-156.

198) Lee, H. & Park, Y., 2015. *Degree of Disturbance Rejection Capability for Linear Marginally Stable Systems.* Busan, South Korea, IEEE, 15th International Conference on Control, Automation and Systems (ICCAS 2015), pp. 307-310.

199) Lee, I. & Han, J., 1996. *Optimal placement of piezoelectric actuators in intelligent structures using genetic algorithms.* Lyons, 3rd Int. Conf. on Intelligent Materials, pp. 872-875.

200) Leleu, S., Abou-Kandil, H. & Bonnassieux, Y., 2001. Piezoelectric actuators and sensors location for active control of flexible structures. *IEEE Transactions on Instrumentation and Measurement,* 50(6), pp. 1577-1582.

201) Leonides, C., 2012. *Control and Dynamic Systems V57: Multidisciplinary Engineering Systems: Design and Optimization Techniques and Their Application: Advances in Theory and Application.* Control and Dynamic Systems: Advances in Theory and Applicat, ISSN 0090-5267 ed. s.l.:Academic Press.

202) Leventides, J. & Karcanias, N., 2008. Structured Squaring Down and Zero Assignment. *International Journal Control,* Volume 81, pp. 294-306.

203) Levine, W. & Athans, M., 1970. On the determination of the optimal constant output feedback gains for the linear multivariablesystems. *IEEE Transactions on Automatic Control,* 15(1), pp. 44-48.

204) Levine, W. S., 1996. *The Control Handbook.* illustrated ed. s.l.:CRC Press.

205) Lewis, F., Dawson, D. & Abdallah, C., 2003. *Robot Manipulator Control: Theory and Practice.* 2, revised ed. s.l.:CRC Press.

206) Lewis, F., Vrabie, D. & Syrmos, V., 2012. *Optimal Control.* 3 ed. s.l.:John Wiley & Sons.

207) Li, L., Huang, Z., Liu, F. & Wu, Q., 2007. A heuristic particle swarm optimizer for optimization of pin connected structures. *Computers & Structures,* 85(7-8), pp. 340-349.

208) Lim, K., 1992. Method for Optimal Actuator and Sensor Placement for Large Flexible Structures. *Journal of Guidance, Control, and Dynamics,* 15(1), pp. 49-57.

209) Lin, C., 1974. Structural controllability. *IEEE Transactions on Automatic Control,* 19(3), pp. 201-208.

210) Lindberg, R. & Longman, R., 1981. Aspects of the degree of controllability-applications to simple systems. *Adv.Astronautical Sci,* Volume 46, pp. 871-891.

211) Liu, K. & Yao, Y., 2016. *Robust Control: Theory and Applications.* Singapore: John Wiley & Sons.

212) Liu, W., Hou, Z. & Demetriou, M., 2006. A computational scheme for the optimal sensor/actuator placement of flexible structures using spatial H2 measures. *Mechanical Systems and Signal Processing,* 20(4), pp. 881-895.

213) Liu, X. & Hu, J., 2010. On the placement of actuators and sensors for flexible structures with closely spaced modes. *Sci. China Technol. Sci,* 53(7), p. 1973–1982.

214) Liu, Y., Slotine, J. & Barabasi, A., 2011. Controllability of complex networks. *Nature,* 473(7346), pp. 167-173.

215) Li, W. & Huang, H., 2013. Integrated optimization of actuator placement and vibration control for piezoelectric adaptive trusses. *Journal of Sound and Vibration,* 332(1), pp. 17-32.

216) Loan, C. V., 2000. The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics,* 123(1-2), pp. 85-100.

217) Longman, R. & Horta, L., 1989. Actuator placement by degree of controllability including the effect of actuator mass. *Dynamics and control of large structures; Proceedings of the Seventh VPI&SU Symposium, Blacksburg, VA, May 8-10, 1989 (A91-23726 08-18). Blacksburg, VA, Virginia Polytechnic Institute and State University,* pp. 245-260.

218) Longman, R. & Lindberg, R., 1986. *The search for appropriate actuator distribution criterion in large space structures control.* Boston, MA: Narendra K.S. (eds) Adaptive and Learning Systems, Springer.

219) Luyben, W. L., 1988. The Concept of 'Eigenstructure in Process Control. *Industrial and Engineering Chemistry Research,* Volume 27, pp. 206-208.

220) MacFarlane, A. & Karcanias, N., 1976. Poles and zeros of linear multivariable systems: A survey of the Algebraic Geometric and Complex Variable Theory. *Int. J. Control,* Volume 24, pp. 33-74.

221) Madron, F. & Veverka, V., 1992. Optimal selection of measuring points in complex plants by linear models. *AIChE ,* 38(2), p. 227–236.

222) Mah, R., 1990. *Chemical Process Structures and Information Flows.* Butterworths series in chemical engineering ed. Stoneham, England: Elsevier.

223) Marx, B., 2003. *"Contribution a´ la commande et au diagnostic de syste´mes alge´bro-diffe´rentiels line´aires" Contribution to the analysis and structured ordering of large systems.* France: PhD Thesis, Institut National Polytechnique de Grenoble (INPG).

224) Marx, B., Koenig, D. & Georges, D., 2002. *Optimal sensor/actuator location for descriptor systems using Lyapunov-like equations.* Las Vegas, NV, USA, USA, Proceedings of the 41st IEEE Conference on Decision and Control.

225) Marx, B., Koenig, D. & Georges, D., 2004. *Optimal Sensor and Actuator Location for Descriptor Systems using Generalized Gramians and Balanced Realizations.* Boston. Massachusett, Proceedings of the 2004 American Control Conferenc.

226) May, R., 1976. Simple mathematical models with very complicated dynamics. *Nature,* Volume 261, pp. 459-467.

227) Mazenc, F., Queiroz, M. d. & Malisoff, M., 2005. On Active Magnetic Bearing Control with Input Saturation. *44th IEEE Conference on Decision and Control-the European Control Conference,* pp. 1090-1095.

228) Mehrabian, A. & Yousefi-Koma, A., 2007. Optimal positioning of piezoelectric actuators on a smart fin using bio-inspired algorithms. *Aerospace Science and Technology,* Volume 11, pp. 174-182.

229) Mellodge, P. & Kachroo, P., 2008. *Model Abstraction in Dynamical Systems: Application to Mobile Robot Control.* illustrated ed. s.l.:Springer.

230) Miao, J., 1991. General expressions for the Moore-Penrose inverse of a 2 × 2 block matrix. *Linear Algebra and its Applications,* Volume 151, pp. 1-15.

231) Middleton, R. & Goodwin, G., 1990. *Digital Control and Estimation: A Unified Approach.* Englewood Cliffs, New Jersey: Prentice Hall .

232) Minto, D. & Nett, C., 1989. *A quantitative approach to the selection and partitioning of measurements and manipulations for the control of complex systems.* Pittsburgh, PA, American Control Conference.

233) Mirjalili, S., 2015. The Ant Lion Optimizer. *Advances in Engineering Software,* Volume 83, pp. 80-98.

234) Mirjalili, S. & Mirjalili, S.M. and Yang, X.S., 2014. Binary bat algorithm. *Neural Computing and Applications,* 25(3), p. 663–681.

235) Mirza, M. & Niekerk, J., 1999. Optimal Actuator Placement for Active Vibration Control with Known Disturbances. *Journal of Vibration and Control,* 5(5), pp. 709-724..

236) Montes, E. & Coello, C., 2008. An empirical study about the usefulness of evolution strategies to solve constrained optimization problems. *International Journal of General Systems,* 37(4), pp. 443-473.

237) Monzxngo, R., 1967. A note on sensitivity of system observability. *IEEE Trans. Aut. Control,* Volume 12, pp. 314-315.

238) Moore, B., 1981. Principal component analysis in linear systems. *IEEE Transactions on Automatic Control,* 26(1), pp. 17-32.

239) Moore, B., 1981. Principle component analysis in linear systems: Controllability, observability and model reduction. *IEEE Tran. sAutomat. Control,* Volume 26, pp. 17-32.

240) Morari, M., 1982. Integrated plant control: A solution at hand or a research topic for the next decade. *New York, Chemical Process Control II : AIChE*, pp. 467-495.

241) Morari, M., 1983. Design of Resilient Process Plants 3: A General Framework for the Assessment of Dynamic Resilience. *Chemical Engineering Science,* 38(11), pp. 1881-189l.

242) Morari, M., Arkun, Y. & Stephanopoulos, G., 1980. Studies in the synthesis of control structures for chemical processes: Part I: Formulation of the problem. Process decomposition and the classification of the control tasks. Analysis of the optimizing control structures. *AIChE Journal,* 26(2), pp. 220-232.

243) Morari, M., Grimm, W., Oglesby, M. & Prosser, I., 1985. Design of Resilient Processing Plants 7: Design of Energy Management System for Unstable Reactors-NewInsights. *Chemical Engineering Science,* 40(2), pp. 187-198.

244) Morari, M. & O'Dowd, M., 1980. Optimal sensor location in the presence of nonstationary noise. *Automatica,* 16(5), pp. 463-480.

245) Morari, M. & Stephanopoulos, G., 1980. Studies in the synthesis of control structures for chemical processes. Part II: Structural aspects and the synthesis of alternative feasible control schemes. *AIChE Journal,* 26(2), pp. 232-246.

246) Morari, M. & Stephanopoulos, G., 1980. Studies in the synthesis of control structures for chemical processes: Part III: Optimal selection of secondary measurements within the framework of state estimation in the presence of persistent unknown disturbances. *AIChE Journal,* 26(2), pp. 247-259.

247) Morris, K., 1998. Noise reduction achievable by point control. *ASME Journal on Dynamic Systems, Measurement and Control,* 120(2), p. 216–223.

248) Morris, K., 2001. *Introduction to Feedback Control.* illustrated ed. the University of Michigan: Harcourt/Academic Press.

249) Morris, K., 2011. Linear quadratic optimal actuator location. *IEEE Transactions on Automatic Control,* Volume 56, p. 113–124.

250) Morris, K. & Demetriou, M., 2010. Using H2 control metrics for the optimal actuator location of infinitedimensional systems. s.l., *Proceeding of the 2010 American Control Conference*, pp. 4899-4904.

251) Morris, K., Demetriou, M. & Yang, S., 2015. Using H2-control performance metrics for optimal actuator location in distributed parameter systems. *IEEE Transactions on Automatic Control,* 60(2), p. 450–462.

252) Morris, K. & Yang, S., 2015. Comparison of actuator placement criteria for control of structures. *Journal of Sound and Vibration,* Volume 353, pp. 1-18.

253) Muller, P. & Weber, H., 1972. Analysis and optimization of certain quantities of controllability and observability for linear dynamic system. *Automatica,* 8(3), pp. 237-246.

254) Murota, K., 2009. *Matrices and Matroids for Systems Analysis.* illustrated ed. Berlin Heidelberg: Springer Science & Business Media.

255) Nemhauser, G., Wolsey, L. & Fisher, M., 1978. An analysis of approximations for maximizing submodular set functions. *Mathematical Programming,* 14(1), p. 265–294.

256) Nepusz, T. & Vicsek, T., 2012. Controlling edge dynamics in complex networks. *Nature Physics,* 8(7), p. 568–573.

257) Nett, C. N., 1989. *A quantitative approach to the selection andpartitioning of measurements and manipulations for the control ofcomplex systems.* Pasadena, USA: Presentation at the Caltech Control Workshop.

258) Noshadi, A. et al., 2016. System Identification and Robust Control of Multi-Input Multi-Output Active Magnetic Bearing Systems. *IEEE Transactions on Control Systems Technology,* 24(4), pp. 1227-1239.

259) Nour-Eldin, H., 1987. *Linear Multivariable Systems Controllability and Observability: Numerical Aspects.* M.G. Singh (ed.) ed. Oxford, UK: Systems and Control Encyclopedia, Pergamon Press.

260) Ogata, K., 1997. *Modern Control Engineering.* 4th edition (2002) ed. s.l.:PrenticeHall.

261) Olshevsky, A., 2014. The minimal controllability problem. *IEEE Transactions on Control of Network Systems,* 1(3), p. 249–258.

262) Omatu, S., Koide, S. & Soeda, T., 1978. Optimal sensor location for a linear distributed parameter system. *IEEE Transactions on Automatic Control,* 23(4), pp. 665-673.

263) Osman, I. & Laporte, G., 1996. Metaheuristics: a bibliography. *Annals of Operations Research,* 63(5), pp. 513-623.

264) Owens, D., 2015. *Iterative Learning Control: An Optimization Paradigm.* illustrated ed. s.l.:Springer.

265) Özdemir, N., 2003. State-space solutions to standard H∞ control problem. *Academic Journal,* 52(2), pp. 35-53.

266) Padula, S. & Kincaid, R., 1999. *Optimization strategies for sensor and actuator placement,* Virginia 23681: National Aeronautics and Space Administration Langley Research Center, Tech. Rep..

267) Paige, C., 1981. Properties of numerical algorithms related to computing controllability. *IEEE Transactions on Automatic Control,* pp. 130-138.

268) Parrilo, P., 2003. Semidefinite programming relaxations for semialgebraic. *Springer, Mathematical Programming,* 96(2), pp. 293-320.

269) Parrilo, P. A., 2000. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* California: PhD Thesis,California Institute of Technology.

270) Pasqualetti, F., Zampieri, S. & Bullo, F., 2014. Controllability metrics limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems,,* 1(1), pp. 40-52.

271) Peng, F., Ng, A. & Hu, Y., 2005. Actuator Placement Optimization and Adaptive Vibration Control of Plate Smart Structures. *Journal of Intelligent Material Systems and Structures,* 16(3), pp. 263-271.

272) Pequito, S., Kar, S. & Aguiar, A., 2016. A Framework for Structural Input/Output and Control Configuration Selection in Large-Scale Systems. *IEEE Transactions on Automatic Control,* 61(2), pp. 303-318.

273) Petersen, K. & Pedersen, M., 2012. *The Matrix Cookbook.* [Online]
Available at:
http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf
[Accessed 15 Nov 2012].

274) Powers, V. & Wörmann, T., 1998. An Algorithm for Sums of Squares of Real Polynomials. *Pure and Applied Algebra,* 127(1), pp. 99-104.

275) Prasolov, V., 1994. *Problems and Theorems in Linear Algebra.* s.l.:American Mathematical Soc.

276) Press, W., Teukolsky, S., Vetterling, W. & Flannery, B., 2007. *Numerical Recipes: The Art of Scientific Computing.* 3rd Edition ed. New York: Cambridge University Press.

277) Price, K., Storn, R. & Lampinen, J.A., 2005. *Differential Evolution: A Practical Approach to Global Optimization.* New York, NY, USA. 1 ed. :Springer-Verlag Berlin Heidelberg.

278) Pulthasthan, S. & Pota, H., 2008. The optimal placement of actuator and sensor for active noise control of sound-structure interaction systems. *Smart Materials and Structures,* 17(3), pp. 1-11.

279) Qi, A., Huan, V. & Suganthan, P., 2009. Differential evolution algorithm with strategy adaptation for global numerical optimization. *IEEE Transactions on Evolutionary Computation,* 13(2), pp. 398-417.

280) Qiua, Z., Zhang, X., Wub, H. & Zhang, H., 2007. Optimal placement and active vibration control for piezoelectric smart flexible cantilever plate. *Journal of Sound and Vibration,* 301(3-5), pp. 521-543.

281) Qiu, Z., Zhang, X., Wu, H. & Zhang, H., 2007. Optimal placement and active vibration control for piezoelectric smart flexible cantilever plate. *Journal of Sound and Vibration,* 301(3-5), pp. 521-543.

282) Raghuraj, R., Bbushan, M. & Rengaswamy, R., 1999. Location of sensors in complex chemical plants based on fault diagnostic observability criteria. *American Institute of Chemistry Engineering Journa,* 45(2).

283) Rahmani, A., Ji, M., Mesbahi, M. & Egerstedt., M., 2009. Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journalon Control and Optimization,* 48(1), p. 162–186.

284) Raja, M. & Narayanan, S., 2009. Simultaneous optimization of structure and control of smart tensegrity structures. *Journal of Intelligent Material Systems and Structures,* 20(1), pp. 109-117.

285) Rajapakse, I., Groudine, M. & Mesbahi, M., 2011. Dynamics and control of state-dependent networks for probing genomic organization. *Proceedings of the National Academy of Sciences,* 108(42), p. 17257–17262.

286) Rao, S., Pan, T. & Venkayya, V., 1991. Optimal placement of actuators in actively controlled structures using genetic algorithms. *AIAA journal,* 29(6), pp. 942-943.

287) Rashedi, E. & Nezamabadi-pour, H., 2009. GSA: A Gravitational Search Algorithm. *Information Sciences: an International Journal,* 179(13), pp. 2232-2248.

288) Reinschke, K. J., 1988. *Multivariable Control A Graph-Theoretic Approach.* Lecture Notes in Control and Information Sciences ed. s.l.:Springer-Verlag.

289) Rijnsdorp, J., 1991. *Integrated Process Control and Automation.* Amsterdam: Elsevier Science Ltd.

290) Romagnoli, J., Alvarez, J. & Stephanopolus, G., 1981. Variable measurement structures for process control. *International Journal of Control,* 33(2), pp. 269-289.

291) Rosenbrock, H. & Rower, B., 1970. Allocation of poles and zeros. *Proceedings of the Institution of Electrical Engineers,* 117(9), pp. 1879 - 1886.

292) Safizadeh, M. R., Darus, I. M. & Mailah, M., 2010. *Optimal placement of piezoelectric actuator for active vibration control of flexible plate.* Kuala Lumpur, Malaysia, Malaysia, 2010 International Conference on Intelligent and Advanced Systems (ICIAS).

293) Salomon, R., 1996. Re-evaluating genetic algorithm performance under coordinate rotation of benchmark functions. A survey of some theoretical and practical aspects of genetic algorithms. *Biosystems,* 39(3), pp. 263-278.

294) Sandgren, E., 1988. Nonlinear integer and discrete programming in mechanical design. *Journal of Mechanical Design,* 112(2), pp. 223-229.

295) Schweitzer, G. & Maslen, E., 2009. *Magnetic Bearings: Theory, Design, and Application to Rotating Machinery.* Verlag Berlin Heidelberg: Springer.

296) Seborg, D., Edgar, T., Mellichamp, D. & Doyle, F., 1989. *Process Dynamics and Control.* Third Edition ed. USA: John Wiley & Sons, Inc..

297) Seif, Z. & Ahmadi, M.B., 2015. An opposition-based algorithm for function optimization. *Engineering Applications of Artificial Intelligence,* Volume 37, p. 293–306.

298) Shaker H.R. & Tahavori, M., 2013. Optimal sensor and actuator location for unstable systems. *Journal of Vibration and Control,* Volume 19, pp. 15-20.

299) Shames, I. & Summers, T., 2015. Rigid network design via submodular set function optimization. *IEEE Transactions on Network Science and Engineering,* 2(3), pp. 84-96.

300) Shimizu, K. & Matsubara, M., 1985. Directions of Disturbances and Modeling Errors on the Control Quality in Distillation Systems. *Chemical Engineering Communications,* 37(1), pp. 67- 9l.

301) Shinskey, F., 1988. *Process Control System Application Design and Tuning.* 4, illustrated ed. New York: McGraw Hill.

302) Siljak, D., 1978. *Large-scale dynamic systems: stability and structure.* illustrated ed. The University of Michigan: North-Holland.

303) Siljak, D., 1991. *Decentralized control of complex systems.* Boston: Academic Press.

304) Silva, S., Lopes, V. & Brennan, M., 2006. Design of a control system using linear matrix inequalities for the active vibration control of a plate. *Journal of Intelligent Material Systems and Structures,* 17(1), p. 81{93.

305) Simovici, D. A., 2012. *Linear Algebra Tools for Data Mining.* s.l.:World Scientific.

306) Singh, A. & Hahn, J., 2005. Determining Optimal Sensor Locations for State and Parameter Estimation for Stable Nonlinear Systems. *Ind. Eng. Chem. Res,* 44(15), p. 5645–5659.

307) Singh, A. & Hahn, J., 2006. Sensor Location for Stable Nonlinear Dynamic Systems: Multiple Sensor Case. *Ind. Eng. Chem. Res,* 45(10), p. 3615–3623.

308) Singh, T., 2009. *Optimal Reference Shaping for Dynamical Systems: Theory and Applications.* illustrated ed. s.l.:CRC Press.

309) Sinha, S., Vaidya, U. & Rajaram, R., 2013. *Optimal placement of actuators and sensors for control of nonequilibrium dynamics.* Zürich, Switzerland, 2013 European Control Conference (ECC).

310) Sivasundaram, S., 2004. *Advances in Dynamics and Control.* Nonlinear Systems in Aviation, Aerospace, Aeronautics and Astronautics Series ed. s.l.:CRC Press.

311) Skelton, R., 1988. *Dynamic Systems Control: Linear Systems Analysis and Synthesis.* New York: John Wiley & Sons.

312) Skogestad, S. & Morari, M., 1987. The Effect of Disturbance Directions on Closed-Loop Performance. *Industrial and Engineering Chemistry Research,* 26(10), pp. 2029-2035.

313) Skogestad, S. & Postlethwaite, I., 1996. *Multivariable Feedback Control.* Chichester : Wiley.

314) Skogestad, S. & Postlethwaite, I., 2005. Multivariable Feedback Control:Analysis and Design 2nd Edition. In: *Chapter 4.* s.l.: Chichester: John Wiley & Sons.

315) Skogested, S., 2000. Plantwide control: The search for the self-optimizing control structure. *Journal of Process Control,* 10(5), pp. 487-507.

316) Sojoudi, S., Lavaei, J. & Aghdam, A., 2009 (b). Robust Controllability and Observability Degrees of Polynomially Uncertain Systems. *Automatica,* 45(11), pp. 2640-2645.

317) Sojoudi, S., Lavaei, J. & Aghdam, A., 2009. *Robust controllability and observability degrees of polynomially uncertain systems.* Shanghai, China, 48th IEEE Conference on Decision and Control.

318) Sontag, E. D., 1991. Mathematical System Theory-Kalman's Controllability Rank Condition: From Linear to Nonlinear. In: *Chapter 7, The Influence of R. E. Kalman.* s.l.:Springer Berlin Heidelberg, pp. 453-462.

319) Sorrentino, F., Bernardo, M. d., Garofalo, F. & Chen, G., 2007. Controllability of complex networks via pinning. *Physical Review-Series E,* 75(4), p. 046103.

320) Sreeram, V. & Agathoklis, P., 1991. Solution of Lyapunov equation with system matrix in companion form. *IEE Proceedings D - Control Theory and Applications,* 138(6), pp. 529 - 534.

321) Stanley, G., Marino-Galarraga, M. & McAvoy, T., 1985. Shortcut Operability Analysis I: The Relative Disturbance Gain. *Industrial and Engineering Chemistry Process Design and Development,* 24(4), pp. 1181-1188.

322) Stephanopoulos, G., 1984. *Chemical Process Control: An Introduction to Theory and Practice.* illustrated, reprint ed. The University of Michigan: Prentice-Hall.

323) Subbotin, M., 2004. *Balancing an Inverted Pendulum on a Seesaw,* Santa Barbara: University of California.

324) Summers, T., 2016. *Actuator placement in networks using optimal control performance metrics.* s.l., in to appear, IEEE Conference on Decision and Control.

325) Summers, T., Cortesi, F. & Lygeros, J., 2016. On submodularity and controllability in complex dynamical networks. *IEEE Transactions on Control of Network Systems,* 3(1), pp. 1-11.

326) Summers, T. & Lygeros, J., 2014. Optimal sensor and actuator placement in complex dynamical networks. *IFAC Proceedings Volumes,* 47(3), p. 3784–3789.

327) Summers, T. & Lygeros, J., 2014. *Optimal Sensor and Actuator Placement in Complex Dynamical Networks.* Cape Town, South Africa, Preprints of the 19th World Congress The International Federation of Automatic Control.

328) Summers, T., Shames, I., Lygeros, J. & Dorfler, F., 2015. Topology design for optimal network coherence. s.l., *European Control Conference (ECC). IEEE,* p. 575–580.

329) Sung, Y., 2002. Modelling and control with piezo actuators for a simply supported beam under a moving mass. *Journal of Sound and Vibration,* 250(4), pp. 617-626.

330) Sun, J. & Motter, A., 2013. Controllability transition and nonlocality in network control. *Physical Review Letters,* 110(20), p. 208701.

331) Sweeney, R., Demetriou, M. & Grigoriadist, K., 2005. Hinf control of a piezo-actuated flexible beam using an analytical bound approach. s.l., *Proceeding of the 2005 American Control Conference,* pp. 2505-2509.

332) T.H.Summers & J.Lygeros, 2014. *Optimal Sensor & Actuator Placement in Complex Dynamical Networks.* Cape Town, South Africa, s.n.

333) Takewaki, I., 1997. Optimal Damper Placement for Minimum Transfer Functions. *Earthquake Engineering and Structural Dynamics, John Wiley & Sons, Ltd,* 26(11), p. 1113–1124.

334) Takewaki, I., 2000. Optimal damper placement for critical excitation. *Probabilistic Engineering Mechanics,* 15(4), pp. 317-325.

335) Takewaki, I., Yoshitomi, S., Uetani, K. & Tsuji, M., 1999. Non-monotonic optimal damper placement via steepest direction search. *Earthquake Engineering & Structural Dynamics, John Wiley & Sons, Ltd.,* 28(6), p. 655–670.

336) Tang, Y., Gao, H., Zou, W. & Kurths, J., 2012. Identifying controlling nodes in neuronal networks in different scales. *PLOS ONE,* 7(7), p. e41375.

337) Tarokh, M., 1992. Measures for controllability, observability, and fixed modes. *IEEE Trans. Auto. Contr,* Volume 37, pp. 1268-1273.

338) Toan, N. & Georges, D., 2007. *An energy approach to optimal selection of controllers/sensors in power system.* Singapore, IEEE, pp. 106-111.

339) Toscano, R., 2013. *Structured Controllers for Uncertain Systems: A Stochastic Optimization Approach.* illustrated ed. Saint-Etienne, France: Springer Science & Business Media.

340) Trajkov, M. & Nestorvic, T., 2012. *Optimal Placement of Piezoelectric Actuators and Sensors for Smart Structures..* Porto/Portuga, 15th International Conference on Experimental Mechanics.

341) Tzoumas, V., Rahimian, M.A., Pappas, G.J. & Jadbabaie, A., 2016. Minimal actuator placement with bounds on control effort. *IEEE Transactions on Control of Network Systems,* 3(1), pp. 67-78.

342) Tzoumas, V., Rahimian, M.A., Pappas, G.J. & Jadbabaie, A., 2015. Minimal Actuator Placement with Optimal Control Constraints. *Chicago, s.n.*

343) Václavek, V. & Loučka, M., 1976. Selection of Measurements Necessary to Achieve Multicomponent Mass Balances in Chemical Plant. *Chemical Engineering Science,* 31(12), pp. 1199-1205.

344) Van den Berg, F., Hoefsloot, H., Boelens, H. & Smilde, A., 2000. Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria. *Chemical Engineering Science,* 55(4), pp. 827-837.

345) VanderVelde, W. E. & Carignan, C. R., 1982. *Placement of Control System Actuators.* Arlington, VA, USA, IEEE, pp. 7-15.

346) Verhulst, P., 1845. Mathematical Researches into the Law of Population Growth Increase. *Nouveaux Mémoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles,* Volume 18, pp. 1-42.

347) Verhulst, P., 1847. Deuxième mémoire sur la loi d'accroissement de la population. *Mémoires de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique,* Volume 20, pp. 1-32.

348) Vilnay, O., 1981. Design of modal control of structures. *Journal of Engineering Mechanics,* 109(1), pp. 907-915.

349) Viswanathan, C. & Longman, R., 1983. The determination of the degree of controllability for dynamic systems with repeated eigenvalues. *The American Institute of Aeronautics and Astronautics (AIAA),* 50(part II), pp. 1091-1111.

350) Viswanathan, C., Longman, R. & Likins, P., 1979. A definition of the degree of controllability- a criterion for actuator placement. *Dynamics and control of large flexible spacecraft; Proceedings of the Second Symposium, Blacksburg, Va., June 21-23, 1979. (A80-31656 12-18) Blacksburg, Va., Virginia Polytechnic Institute and State University,* pp. 369-384.

351) Viswanathan, C. N., Longman, R. W. & Likins, P. W., 1984. A degree of controllability definition : Fundamental concepts and application to modal systems. *Journal of Guidance, Control, and Dynamics,* pp. 222-230.

352) Vrugt, J., Robinson, B. & Hyman, J., 2009. Self-adaptive multimethod search for global optimization in real-parameter spaces. *IEEE Transactions on Evolutionary Computation,* 13(2), p. 243–259.

353) Waldraff, W., Dochain, D., Bourrel, S. & Magnus, A., 1998. On the use of observability measures for sensor location in tubular reactor. *Journal of Process Control,* 8(5-6), pp. 497-505.

354) Wal, M. v. d. & Jager, B. d., 2001. A review of methods for input/output selection. *Automatica,* 37(4), pp. 487-510.

355) Wang, Q. & Wang, C., 2001. A controllability index for optimal design of piezoelectric actuators in vibration control of beam structures. *Journal of Sound and Vibration,* 242(3), p. 507–518.

356) Wang, W., Ni, X., Lai, Y. & Grebogi, C., 2012. Optimizing controllability of complex networks by minimum structural perturbations. *Physical Review E 85, 026115,* 85(2), pp. 026115-1-5.

357) Wicks, M. & Decarlo, R., 1990. Gramian Assignment Based on the Lyapunov Equation. *IEEE Transactions on Automatic control,* 35(4), pp. 465-468.

358) Wicks, M. & DeCarlo, R., 1991. Computing the distance to uncontrollable system. *IEEE Transactions on Automatic Control,* 36(1), pp. 39-49.

359) Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics,* Volume 1, p. 80–83.

360) Wolpert, D. & Macready, W., 1997. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation,* 1(1), pp. 67-82.

361) Wu, A., Duan, G. & Yu, H., 2006. On Solutions of the Matrix Equations $XF - AX = C$ and $XF - A\bar{X} = C$. *Applied Mathematics and Computation,* Volume 183, pp. 932-941.

362) Wu, B. & Ou, .., 1997. Optimal placement of energy dissipation devices for three-dimensional structures. *Engineering Structures,* 19(2), pp. 113-125.

363) Xiao, C., Feng, Z. & Shan, X., 1992. On the solution of the continuous-time Lyapunov matrix equation in two canonical forms. *IEE Proceedings D - Control Theory and Applications,* 139(3), pp. 286-290.

364) Xiao, C., Feng, Z. & Shan, X., 1992. On the solution of the continuous-time Lyapunov matrix equation in two canonical forms. *IEE Proceedings D - Control Theory and Applications,* 139(3), pp. 286 - 290.

365) Yang, F., Xiao, D. & Shah, S., 2009. Optimal Sensor Location Design for Reliable Fault Detection in Presence of False Alarms. *Sensors,* 9(11), p. 8579–8592.

366) Yang, J. & Chen, G., 2010 b. *Actuator Placement and Configuration Direct Optimization in Plate Structure Vibration Control System.* s.l., International Conference on Measuring Technology and Mechatronics Automation.

367) Yang, J. & Chen, G., 2010. *Optimal Placement and Configuration Direction of Actuators in Plate Structure Vibration Control System.* Wuhan, China, 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR).

368) Yan, G. et al., 2012. Controlling complex networks: How much energy is needed?. *Physical Review Letters,* 108(218703), pp. 1-9.

369) Yan, G. et al., 2015. Spectrum of controlling and observing complex networks. *Nature Physics,* Volume 11, p. 779–786.

370) Yang, X., 2008. *Nature-Inspired Metaheuristic Algorithms.* illustrated ed. Cambridge, UK: Luniver Press.

371) Yang, X. S., 2010. *Test problems in optimization, in: Engineering Optimization: An Introduction with Metaheuristic Applications.* Cambridge: John Wiley & Sons.

372) Yao, X. & Liu, Y., 1996. *Fast Evolutionary Programming.* MIT Press, s.n.

373) Yao, X., Liu, Y. & Lin, G., 1999. Evolutionary programming made. *IEEE Transactions on Evolutionary Computation,* 3(2), pp. 82 - 102.

374) Yoshizawa, T., 2012. *Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions.* illustrated ed. New York: Springer Science & Business Media.

375) Yue, H., Deng, Z. & Tzou., H., 2008. Optimal actuator locations and precision micro control actions on free paraboloidal membrane shells. *Communications in Nonlinear Science and Numerical Simulation,* Volume 13, pp. 2298-2307.

376) Zhang, H. & Ding, F., 2013. On the Kronecker Products and Their Applications. *Journal of Applied Mathematics,* 2013(Article ID 296185), p. 8 pages.

377) Zhang, X. & Erdman, A., 2006 . Optimal placement of piezoelectric sensors and actuators for controlled flexible linkage mechanisms. *Trans. ASME., J. Vib. Acoust.,* 128(2), pp. 256-260.

378) Zhou, B. & Duan, G., 2005. An Explicit Solution to the Matrix Equation AX − XF = BY. *Linear Algebra and its Applications,* Volume 402, pp. 345-366.

379) Zhou, K. & Doyle, J., 1996. *Essentials of Robust Control.* New Jersey: Prentice Hall.

380) Zhou, K., J.C. Doyle & Glover, K., 1997. *Robust and optimal control.* illustrated ed. Upper Saddle River, NJ: Prentice Hall.

381) Zhou, K., Salomon, G. & Wu, E., 1999. Balanced realization and model reduction for unstable systems. *International Journal of Robust and Nonlinear Control,* pp. 183-198.

Appendix chapter 3:

Proposition : (Parrilo, 2003), (Parrilo, 2000) A multivariate polynomial $p(x) \in \mathbb{R}$ in n variables and of degree 2d is a sum of squares if and only if there exists a real symmetric positive semidefinite matrix Q such that:

$$p(x) = z^T Q z$$

where z is the vector of monomials of degree up to d, we call matrix Q the Gram matrix.

$$Z = [1, x_1, x_2, \ldots, x_n, x_1 x_2, \ldots, x_n^d]^T$$

Given Gram matrix Q, of $rank(Q) = t$, we can construct polynomial $h_1, h_2, \ldots, h_t$, such that:

$$p(x) = \sum_{i=1}^{t} (h_i)^2$$

Proof:

If $p(x) = \sum_{i=1}^{t} (h_i)^2$ is true and is SOS, then as above we take $Q = AA^T$, where A is the matrix whose columns are the coefficients of the h_i s.

Now suppose there exist a real symmetric psd matrix Q, such that $p(x) = z^T Q z$ and rank(Q)=t, since Q is real symmetric of rank t, there exists a real matrix V and a diagonal $D = diag(d_1, \ldots, d_t, 0, \ldots 0)$, $d_i \neq 0$ such that $Q = V^T D V$, since Q is psd we have $d_i > 0$, then:

$$p(x) = z^T V^T D V z$$

Now considering $V = v_{i,j}, i = 1, \ldots t$, we have:

$$h_i = \sqrt{d_i} \sum_{i=1}^{k} v_{i,j} z^{\beta_i} \in \mathbb{R}$$

It follows that $p(x) = \sum_{i=1}^{t} (h_i)^2$.

□

Appendix chapter 5:

Proposition: In an LTI system with no undamped modes, the controllability Gramian in the frequency domain is defined as[6]:

$$W_c = \frac{1}{2\pi}\int_{-\infty}^{\infty}(j\omega I - A)^{-1}BB^T(-j\omega I - A^T)^{-1}d\omega$$

And it can be expressed as the block diagonal form:

$$W_c = V\begin{bmatrix} W_s & 0 \\ 0 & W_u \end{bmatrix}V^T$$

Proof:

Consider the steady state definition of controllability Gramian in the time domain:

$$W_c(0,\infty) = \int_0^{\infty} e^{At}BB^T e^{A^T t}dt$$

Then based on the Parseval's Theorem we can convert the definition above to the statement below:

$$W_c = \frac{1}{2\pi}\int_{-\infty}^{\infty}(j\omega I - A)^{-1}BB^T(-j\omega I - A^T)^{-1}d\omega$$

The statement above is clear for the stable systems, but let's investigate what happens if the system has some eigenvalues on the right half plane.

Through non-singular similarity transformation V (the eigenvectors' matrix) the LTI system is decomposed into stable and anti-stable diagonal blocks as it is explained before:

$$\Lambda = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix}, \Lambda = V^{-1}AV$$

By imposing the similarity transformation V we will obtain:

$$W_{cnew}(0,\infty) = \int_0^{\infty} e^{V\Lambda V^{-1}t}VB_{new}B_{new}^T V^T e^{V^{-1}\Lambda V^T t}dt$$

$$\Rightarrow W_{cnew}(0,\infty) = V\int_0^{\infty} e^{\Lambda t}B_{new}B_{new}^T e^{\Lambda t}dtV^T \Rightarrow VW_{cnew}V^T = W_c$$

Using the Parseval's Theorem in the frequency domain we have:

$$W_c = V\frac{1}{2\pi}\int_{-\infty}^{\infty}(j\omega I - \Lambda)^{-1}B_{new}B_{new}^T(-j\omega I - \Lambda^T)^{-1}d\omega V^T$$

[6] For the simplicity we use the same symbol for the controllability Gramian in the frequency domain and the time domain.

Considering the clockwise contour Γ that encircles the right-half plane, we will have:

$$W_c = V\frac{1}{2\pi}\oint_{\Gamma}\begin{bmatrix}(sI-A_s)^{-1}B_sB_s^T(-sI-A_s^T)^{-1} & (sI-A_s)^{-1}B_sB_u^T(-sI-A_u^T)^{-1}\\(sI-A_u)^{-1}B_uB_s^T(-sI-A_s^T)^{-1} & (sI-A_u)^{-1}B_uB_u^T(-sI-A_u^T)^{-1}\end{bmatrix}dsV^T$$

Clearly off-diagonal blocks both are analytic in the right half plane (RHP) then their integral would be zero.

And the diagonal blocks both are stable. The stability of the first block is clear regarding this fact that it represents the stable modes of the system. For the second diagonal block that is related to the anti-stable subsystem, one can readily prove that:

$$\oint_{\Gamma}(sI-A_u)^{-1}B_uB_u^T(-j\omega-A_u^T)^{-1}d\omega = \oint_{\Gamma}(-(-sI+A_u))^{-1}B_uB_u^T(-(sI+A_u^T))^{-1}ds \overset{s\to -s}{=}$$

$$\ldots -\oint_{\Gamma}(sI-(-A_u))^{-1}B_uB_u^T(-sI-(-A_u^T))^{-1}ds = \oint_{\Gamma}(sI-(-A_u))^{-1}B_uB_u^T(-sI-(-A_u^T))^{-1}ds$$

□

Example (5.4):

$$A_{s13\times 13} = (a_{s1}+a_{s2}+a_{s3})$$

$$a_{s1} = \begin{bmatrix}
\text{-0.0000 + 1.2900i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{-0.0000 - 1.2900i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{-0.0001 + 0.4826i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{-0.0001 - 0.4826i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{-0.0001 + 0.4835i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}\\
\text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i} & \text{0.0000 + 0.0000i}
\end{bmatrix}$$

$$a_{s2}=\begin{bmatrix}0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ -0.0001 - 0.4835i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & -0.0152 + 0.0336i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & -0.0152 - 0.0336i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & -0.0221 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & -0.0112 + 0.0228i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\end{bmatrix}$$

$$a_{s3}=\begin{bmatrix}0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ -0.0112 - 0.0228i & 0.0000 + 0.0000i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & -0.0071 + 0.0212i & 0.0000 + 0.0000i\\ 0.0000 + 0.0000i & 0.0000 + 0.0000i & -0.0071 - 0.0212i\end{bmatrix}$$

$$A_{u5\times 5}=\begin{bmatrix}462.7000 & 0 & 0 & 0 & 0\\ 0 & 191.4000 & 0 & 0 & 0\\ 0 & 0 & 243.4000 & 0 & 0\\ 0 & 0 & 0 & 272.4000 & 0\\ 0 & 0 & 0 & 0 & 284.5000\end{bmatrix}$$

$$B_{s13\times1} = \begin{bmatrix} -0.0172 & 0.0124 & -0.0166 & -0.0136 \\ -0.0172 & 0.0124 & -0.0166 & -0.0136 \\ -1.0416 & 0.1889 & 1.2417 & 0.2170 \\ -1.0416 & 0.1889 & 1.2417 & 0.2170 \\ -0.1211 & -1.0076 & 0.1456 & -1.2523 \\ -0.1211 & -1.0076 & 0.1456 & -1.2523 \\ -3.7286 & 3.0158 & -3.4040 & -2.4204 \\ -3.7286 & 3.0158 & -3.4040 & -2.4204 \\ -1.4743 & -0.9178 & -0.2739 & 2.9960 \\ -1.4743 & -0.9178 & -0.2739 & 2.9960 \\ 0.7778 & 0.6500 & 1.2099 & 1.0762 \\ 0.7778 & 0.6500 & 1.2099 & 1.0762 \\ -4.6704 & 2.8426 & -4.9413 & -4.7850 \end{bmatrix}$$

$$B_{u5\times5} = \begin{bmatrix} 2.6764 & -0.5699 & 6.9452 & -0.3500 \\ 8.7810 & -21.0600 & 0.5281 & 33.6300 \\ -44.7700 & -16.00 & -53.4300 & 1.0640 \\ -50.9400 & -8.020 & 65.9800 & 13.8900 \\ 34.1700 & -55.910 & -50.3000 & -44.6800 \end{bmatrix}$$

□

Appendix chapter 7:

Lemma: For any matrix $A_{n\times m}$, $m \le n$, $rank(A) = m$, if and only if the minimum singular value of A is greater than zero, i.e. $\sigma_m > 0$.

Proof:

Based on singular value decomposition (Gene H. Golub, Charles F. Van Loan, 1996), (Horn, Roger A and Johnson, Charles R, 2012), (Marcus, M. and Minc, H, 1988), for any matrix $A_{n\times m}$ has a factorization of the form $U\Sigma V^*$, where U is a $n\times n$ unitary matrix, Σ is a $n\times m$ rectangular diagonal matrix with non-negative real numbers on the diagonal, and V is a $m\times m$ unitary matrix. The diagonal entries σ_i's of Σ are the singular values of A.

Now assume that $rank\left(A\right) = m$ then clearly $rank(\Sigma) = m$, which means that all singular values of A must be greater than zero, (i.e. $\sigma_m > 0$). Conversely, if minimum singular value of A is greater than zero, then the number of nonzero columns of Σ, which shows the rank of A would be equal to m, i.e. $rank(\Sigma) = rank(A) = m$, and the Lemma is proved.

□

Theorem: If G is the Gram matrix of A, then $rank(G) = rank(A)$.

Proof: (Mirsky, 2012)

Let x be a vector such that $A^*Ax = 0$. Then $x^*A^*Ax = 0$ and so we have $(A^*x)^*Ax = 0$, i.e. $\left|Ax\right|^2 = 0$. Therefore, we have $Ax = 0$. Conversely, if x satisfies the relation $Ax = 0$, then clearly $A^*Ax = 0$. It follows that the homogeneous system $Ax = 0$ and $Gx = 0$ are equivalent. Hence, $n - rank(A) = n - rank(G)$ where n is the number of columns of A. The proof of the Theorem is completed.

□

Proposition: Let $B = [\underline{b}_1, \underline{b}_2, \dots, \underline{b}_m]$, $\underline{b}_i \in \mathbb{R}^{n\times 1}$, $m \le n$. Then, $rank(B) = m$ if and only if the Gram matrix of B is non-singular, i.e. $\left|G\right| \neq 0$.

Proof: (Peter Lancaster, Miron Tismenetsky, 1985)

Assume $rank(B) < m$, i.e. $[\underline{b}_1, \underline{b}_2, \dots, \underline{b}_m]$ are linearly dependent, let $a = [a_1, a_2, \dots, a_m]$ be a nonzero vector such that $\sum_{i=1}^{m} a_i b_i = 0$ then for all j, $\sum_{i=1}^{m} a_i b_j . b_i = 0$ so a^T is in the kernel of the Gram matrix. Conversely, if the Gram matrix is singular, then there exists a nonzero vector $a = [a_1, a_2, \dots, a_m]$, such that $\sum_{i=1}^{m} a_i b_j . b_i = 0$, let $w = \sum_{i=1}^{m} a_i b_i = 0$ Then w is orthogonal to every

b_j, and therefore orthogonal to itself. That is, $w.w = 0$ then $w = 0$, thus $[\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_m]$ are linearly dependent, and $rank(B) < m$.

□