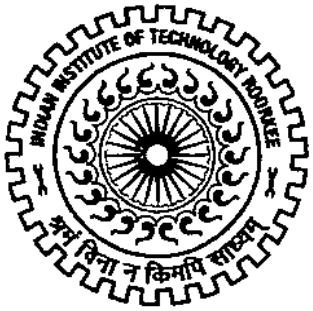
INTELLIGENT ADAPTIVE OPTIMAL CONTROL OF DYNAMICAL SYSTEMS

Ph.D. THESIS

by

LAL BAHADUR PRASAD



DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247 667 (INDIA)
JULY, 2014

INTELLIGENT ADAPTIVE OPTIMAL CONTROL OF DYNAMICAL SYSTEMS

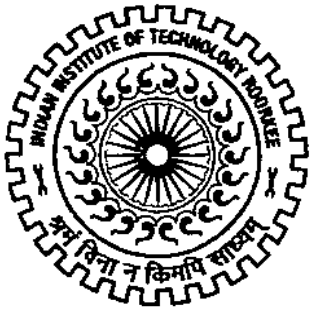
A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree of

DOCTOR OF PHILOSOPHY
in
ELECTRICAL ENGINEERING

by

LAL BAHADUR PRASAD



DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247 667 (INDIA)
JULY, 2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **"INTELLIGENT ADAPTIVE OPTIMAL CONTROL OF DYNAMICAL SYSTEMS"** in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Electrical Engineering of Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from July, 2009 to July, 2014 under the supervision of Dr. Hari Om Gupta, Director, JIIT, Sector-128, Noida and Dr. Barjeev Tyagi, Associate Professor, Department of Electrical Engineering, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

(LAL BAHADUR PRASAD)

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(Barjeev Tyagi)
Supervisor

(Hari Om Gupta)
Supervisor

Date: July, 2014

The Ph.D. Viva-Voce Examination of **Mr. LAL BAHADUR PRASAD**, Research Scholar, has been held on …….

Signature of Supervisor's Chairman, SRC Signature of External Examiner

Head of the Deptt./Chairman, ODC

ABSTRACT

Most of the dynamical systems such as power systems, missile systems, robotic systems, inverted pendulum, industrial processes, chaotic circuits etc. are highly nonlinear in nature. The control of such systems is a challenging task. Intelligent adaptive optimal control is a viable recent approach. Intelligent adaptive optimal control has been emerged from the integration of adaptive control and optimal control methodologies with intelligent computational techniques. In this research work the performance investigation of intelligent adaptive optimal control of dynamical systems is presented. The applications of control schemes for dynamical systems control are implemented considering certain examples of linear and nonlinear dynamical systems to attempt this research investigation.

The performance of controlled systems is desired to be optimal which should be valid also when applied in the real situation. Adaptive control which is able to deal with uncertainties is generally not optimal. Optimal control is off-line, and needs the knowledge of system dynamics for its design. Thus, to have both features of control design, it is desired to design online adaptive optimal control.

Policy Iteration (PI) is a computational intelligence technique that belongs to a class of reinforcement learning (RL) algorithms; solves Hamilton-Jacobi-Bellman (HJB) equation by direct approach. Based on actor-critic structure, PI algorithm consists of two-step iteration: policy evaluation and policy improvement. These two steps of policy evaluation and policy improvement are repeated until the policy improvement step no longer changes the actual policy and thus converging to the optimal control. PI algorithm starts by evaluating the cost of a given initial admissible (stabilizing) control policy to converge towards state feedback optimal control. The infinite horizon optimal solution using HJB and algebraic Riccati equation (ARE) which gives linear quadratic regulator (LQR) require the complete knowledge of the system dynamics. Also these techniques give offline solution. The online PI algorithm solves online the continuous-time optimal control problem without using the knowledge of system internal dynamics, the information which is extracted from real-time dynamics by online measurement of sampled states along state trajectory. The knowledge of internal state dynamics is not needed for evaluation of cost or the update of control policy; and only the knowledge of input-to-state dynamics is required for updating the control policy. Thus, it is a partially model-free approach. The adaptive critic design (ACD) using online PI technique gives an online infinite horizon adaptive optimal control solution for continuous-time linear time invariant (LTI) systems and continuous-time affine nonlinear systems. By using neural networks to parameterize actor and critic for online implementation, this control scheme becomes a high-level intelligent control scheme.

In the PI algorithm the critic is trained to approximate the solution of Lyapunov equation at the policy evaluation step, and the actor is trained to approximate the control policy at the

policy improvement step, and the critic and actor are sequentially updated taking other one constant. In the generalized PI algorithm either one or both of the policy evaluation and policy improvement steps are not required to complete before the next step is started. The online synchronous policy iteration algorithm uses simultaneous continuous-time tuning for the actor and critic neural networks. The online synchronous PI algorithm which needs the knowledge of the system dynamics, solve the optimal control problem online using real-time measurements of closed-loop signals. Using neural networks approximations for critic and actor both it gives an online intelligent adaptive optimal control solution for the continuous-time dynamical systems.

This research work contributes by presenting the comprehensive performance investigation of the different control schemes for continuous-time linear time-invariant (LTI) systems and affine nonlinear systems. The following objectives have been considered in this research work.

1. Optimal control of nonlinear inverted pendulum dynamical system using PID controller & LQR.
2. Intelligent control of nonlinear inverted pendulum dynamical system using Mamdani and TSK fuzzy inference systems.
3. Optimal control using LQR for automatic generation control of two-area interconnected power system.
4. Intelligent control using fuzzy-PI controller for automatic generation control of two-area interconnected nonlinear power system.
5. Intelligent control of process system using radial basis function.
6. Adaptive optimal control using policy iteration technique for LTI systems.
7. Adaptive optimal control using policy iteration technique for affine nonlinear systems.
8. Intelligent adaptive optimal control using synchronous policy iteration technique for LTI systems.
9. Intelligent adaptive optimal control using synchronous policy iteration technique for affine nonlinear systems.

These research objectives are briefly described as below.

Linear quadratic regulator (LQR), an optimal control technique and PID control method which are generally used for control of the linear dynamical systems have been used in this research work to control the nonlinear dynamical system. The modeling and control design of nonlinear inverted pendulum-cart dynamic system using PID controller & LQR have been presented for both cases of without and with disturbance input. The simulation results and performance analysis have been presented which justify the comparative advantages of optimal control using PID+LQR method. The pendulum stabilizes in upright position and cart

reaches the desired position quickly & smoothly even under the continuous disturbance input justify that the control schemes are simple, effective & robust.

Due to the capabilities of generalization, function approximation, learning and adaptation etc. the neural networks are applied for various control, identification, and estimation applications. In this research work the indirect adaptive control of a nonlinear process system using radial basis function neural networks (RBFNN) is presented. The liquid level control problem of a surge tank is considered as a process system. Two RBFNNs are used to model this affine nonlinear system to approximate the internal state dynamic function and input-to-state dynamic function respectively. The RBFNN controller provides a satisfactory response.

Fuzzy control has an impact in the control community because of the simple approach it provides to use heuristic control knowledge for nonlinear control problems. Fuzzy control is an intelligent control technique which uses the human expert knowledge to make the control decisions. In this research work, the modeling, control design and performance analysis of fuzzy control for nonlinear inverted pendulum-cart dynamic system without & with disturbance input are presented. The fuzzy control methods using Mamdani and Takagi-Sugeno-Kang (TSK) fuzzy inference systems have been implemented to control the cart position and stabilize the inverted pendulum in vertically upright position. The comparative performance analysis of these fuzzy control methods have also been done with PID control method. The simulation results justify the comparative advantages of fuzzy control methods. The pendulum stabilizes in vertically upright position and cart approaches the desired position even under the continuous disturbance input justify that the control schemes are effective & robust. The analysis of the responses of the control schemes gives that the performance of PD-FLC using Mamdani type FIS is better than PID controller, and the performance of TSK FLC is better than both. The response of direct fuzzy control using TSK FIS is more smooth & fast than both PID control & Mamdani PD-fuzzy control.

Electrical power systems are complex nonlinear dynamic systems. As the system parameters can’t be completely known under the presence of nonlinearities and uncertainties, the controller designed based on a fixed-parameter linearized model may not work properly for the actual plants. Thus, it is required to take into account the system nonlinearities and parametric uncertainties in the control design. In view of this aspect of investigation, this research work presents the modeling, simulation and performance analysis of automatic generation control (AGC) of two-area interconnected nonlinear power system using fuzzy-PI controller. The conventional integral control is also presented for comparing results. The simulation results and analysis justify the comparative advantages of fuzzy control method.

In this research work the application of policy iteration technique based adaptive critic scheme for adaptive optimal control of continuous-time LTI dynamical systems is presented. The control scheme is implemented considering various practical examples of LTI systems- general SISO LTI system, higher order LTI system- a mechanical system, load frequency control of power system, automatic voltage regulator of power system, and DC motor speed control system. The systems modeling, analysis, and simulation results are presented for load frequency control of power system for both of system models without and with integral control, automatic voltage regulator of power system for both models of without and with sensor, DC motor speed control system for both of system models without and with integral compensator. Analyzing the simulation results obtained for these applications, it is observed that critic parameter matrix P and actor parameter K obtained from adaptive critic scheme using PI technique are converging adaptively to optimal values which are mostly same to that obtained from LQR approach. Also in case of change in system parameter in real situation the controller adapts it and converges to same optimal values. Thus the actor K and critic P parameters remain unchanged. The structural change introduced in system dynamics by including integral control/compensator is augmenting the system behavior such as of its credit that removing the steady state error in closed loop responses. The structural change in system will not be adapted by the proposed controller but it will adapt the change in system parameters in real situation at any moment of time which is demonstrated by simulating also with change in system parameters at certain instant of time. The comparative performance investigation of adaptive critic control scheme and linear quadratic regulator is also presented. Thus, adaptive optimal control scheme is partially model-free, effective & robust.

In this research work the application of PI technique based adaptive critic scheme for adaptive optimal control of continuous-time affine nonlinear dynamical systems is presented. The cost function approximation using neural network is used for online implementation of PI algorithm. The application of control scheme is implemented considering the state regulation problem for certain general affine nonlinear systems and certain practical examples of affine nonlinear systems- single-link manipulator, inverted pendulum, Vander Pol oscillator. The simulation results and performance analysis are presented from which it is observed that the system states converge towards the equilibrium point at origin, and the control signal remains bounded converging towards zero. The cost function approximation neural networks weights are adjusted to the optimal values which give the critic parameters converging adaptively to optimal values and thus the control policy is adaptive optimal. The online PI algorithm requires an initial stabilizing controller for converging to the optimal solution. The simulation results and performance analysis demonstrate the effectiveness of online policy iteration technique based adaptive critic control scheme.

In this research work the applications of online synchronous PI technique using neural networks for adaptive optimal control of continuous-time LTI systems and affine nonlinear systems are presented. The application of online synchronous PI based control scheme is implemented for two practical examples of LTI systems- load frequency control of power system, and automatic voltage regulator of power system. The application of online synchronous PI based control scheme is also implemented for affine nonlinear systems considering the state regulation problem for certain general affine nonlinear systems and two practical examples of affine nonlinear systems- single-link manipulator, and Vander Pol oscillator. The simulation results and performance analysis are presented which demonstrate the effectiveness of online synchronous PI based adaptive critic control scheme. The online synchronous PI based adaptive critic design using neural networks provides an intelligent adaptive optimal control of continuous-time dynamical systems.

ACKNOWLEDGEMENTS

With my deep conscience and spirit I bow my head and thank to almighty God who gave me the opportunity to study and gain knowledge & wisdom, and this life of intellectual. In that process I got the opportunity for doctoral study in this premere institute for which I am highly thankful to almighty God. By his blessings, I feel a sense of self-satisfaction and pleasure of having accomplished my Ph.D research work entitled “Intelligent Adaptive Optimal Control of Dynamical Systems”.

I am pleased to express my profound sense of gratitude and sincere thanks to my Ph.D supervisor Dr. Hari Om Gupta, Professor, Department of Electrical Engineering, for his invaluable active expert guidance, inspiration and constant encouragement throughout the course of this research work. His exemplary patience, concern and support have resulted in the completion of this work to my fullest satisfaction.

I am pleased to express my profound sense of gratitude and sincere thanks to my Ph.D supervisor Dr. Barjeev Tyagi, Associate Professor, Department of Electrical Engineering, for his active expert guidance, constant support, encouragement, and invaluable suggestions throughout the research work.

With my deep hearted gratitude, I express my sincere thanks to Prof. M. K. Vasantha, Professor Emeritus, Department of Electrical Engineering, for his inspiration, motivation, and constant moral support that developed the spirit & zeal to work for my Ph.D and the learning which I received from him through the classes I associated him. I am pleased to express my sincere thanks to him for his constant moral support in my disturbing and troubled days of adverse situations.

I express my deep gratitude to unanimous eminent academicians & researchers whose literatures and publications in the subject area provided the knowledge base and inspired to learn and work for this research assignment. Especially, I express my gratitude and thanks to Dr. Kyriakos G. Vamvoudakis, Faculty Project Research Scientist, CCDC, Department of Electrical and Computer Engineering, University of California, USA, for his generous help in clearing my doubts and improving my conceptual understanding on the topic under this research work.

I express my deep sense of gratitude and sincere thanks to my SRC internal expert member, Dr. G. N. Pillai, Professor, Department of Electrical Engineering, for his invaluable suggestions, motivation, help and support at every step I do need. I express my deep sense of gratitude and sincere thanks to my SRC external expert member, Dr. R. Mitra, Professor Emeritus, Department of Electronics & Communication Engineering, for his invaluable suggestions, and support at every step I do need. I wish to express my sincere thanks to Dr. Vishal Kumar, Assistant Professor & Chairman SRC, Department of Electrical Engineering, for his invaluable suggestions and support at every step I do need.

I sincerely express my thanks to Dr. S. P. Srivastava, Professor & Head, Department of Electrical Engineering, for his invaluable suggestions, support and official help at every step I do need. I sincerely express my thanks to Dr. Rama Bhargwa, Professor & Coordinator, QIP centre, for her invaluable suggestions, support and official help at every step I do need.

I sincerely express my thanks to Dr. Pramod Agarwal, Professor & previous Head of Department of Electrical Engineering, Dr. Vinod Kumar, Professor & previous Coordinator, QIP centre and previous Head of Department of Electrical Engineering, for their invaluable suggestions, support and official help at every step I do need. I express my sincere thanks to Dr. R. S. Anand, Dr. Rajendra Prasad, and Dr. N. P. Padhi, Professors in Department of Electrical Engineering, for their invaluable suggestions, motivation and moral support.

I feel pleasure to express my sincere thanks to all the faculty members of Electrical Engineering Department for their invaluable encouragement, support and timely suggestions. I sincerely thank to all the staff of Electrical Engineering Department, for their invaluable support and official help, and providing me all the necessary information when I required. I sincerely thank to all the staff of QIP Centre, IIT Roorkee for their prompt, sincere, invaluable support and official help, and providing me all the necessary information when I needed.

I am highly thankful to my friends at IIT Roorkee Dr. Vinod Kumar Yadav, Dr. N. S. Beniwal, Dr. Manoj Panda, Dr. Rakesh Maurya, Dr. Emjee Puthooran, Dr. Jitendra Virmani, Mr. Derminder Singh, Mr. Puskar Tripathi, Mr. Kuldeep Dhiman, Mr. Rahul Agrawal, Mr. Yatindra Kumar, Mr. Aurbindo Panda, Mr. Sanjiv Kumar, and many others who helped and cooperated me in resolving certain problems either in my research work or at personal levels.

I am very much grateful to my family friends for their motivation, moral support and help in all respect. I express my deep gratitude and thanks to my family friend Dr. Pramod Kumar Sharma & his family for their invaluable support and help during stay at IIT Roorkee.

I am deep hearted grateful to my family for motivation, moral support and making me free for this work. I express my deep hearted gratitude to my father Shri Neur Prasad, my mother Smt. Naina Devi, my younger brothers Mr. Hari Lal Bhaskar and Mr. Mukesh Kumar Bhaskar, my younger sisters Ms. Neeta and Ms. Reeta, my father-in-law Shri Suraj Deo Prasad, my mother-in-law Smt. Kalapati Devi, my wife Smt. Karuna, my little three & half years old lovely daughter Ms. Mugdha, and my relatives for their invaluable constant motivation, moral support and cooperation by which this dream could be realized.

I dedicate this thesis to Lord Buddha's teaching "Appa Deepo Bhav" and my beloved parents Shri Neur Prasad and Smt. Naina Devi.

LAL BAHADUR PRASAD

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AC	Actor-Critic, Adaptive Critic
ACD	Adaptive Critic Design
ACE	Area Control Error
ADP	Approximate/Adaptive Dynamic Programing
AGC	Automatic Generation Control
AI	Artificial Intelligence
ANN	Artificial Neural Network
ANFIS	Adaptive Neuro-Fuzzy Inference System
ARE	Algebraic Riccati Equation
AVR	Automatic Voltage Regulator
COG	Centre of Gravity
DAC	Direct Adaptive Control
DC	Direct Current
DP	Dynamic Programing
ES	Expert System
FIS	Fuzzy Inference System
FL	Fuzzy Logic
FLC	Fuzzy Logic Controller
GA	Genetic Algorithm
GPI	Generalized Policy Iteration
HJB	Hamilton-Jacobi-Bellman
IAC	Indirect Adaptive Control
IP	Inverted Pendulum
LFC	Load Frequency Control
LQR	Linear Quadratic Regulator
LTI	Linear Time Invariant
MIMO	Multiple-Input–Multiple-Output
MRAC	Model Reference Adaptive Control
NN	Neural Network
PD	Proportional and Derivative
PDE	Partial Differential Equation
PE	Persistence of Excitation
PI	Proportional and Integral, Policy Iteration
PID	Proportional Integral Derivative
PSO	Particle Swarm Optimization
RBFNN	Radial Basis Function Neural Network

RL	Reinforcement Learning
SISO	Single-Input–Single-Output
SMC	Sliding Mode Control
SNAC	Single Network Adaptive Critic
STR	Self-Tuning Regulator
SVM	Support Vector Machine
TSK	Takagi-Sugeno-Kang
VFA	Value Function Approximation
VI	Value Iteration
UUB	Uniformly Ultimately Bounded
UCO	Uniform Complete Observability
ZOH	Zero-Order Hold

LIST OF SYMBOLS

$x(t)$	State vector, cart position
$y(t)$	Output vector
$u(t)$	Control vector, force on cart
f, g, h	Smooth nonlinear vector valued functions
$x_d(t)$	Desired state
$e(t)$	State error
$x(0), x_0$	Initial conditions of system states
A, B, C, D	State space matrices
U_c	Controlability region
$\mathcal{R}$	Set of real numbers
J	Performance index
Q	Positive semi-definite symmetric constant matrix of cost function
R	Positive definite symmetric constant matrix of cost function, Speed regulation due to governor action
P	Positive definite symmetric constant matrix from solution of ARE
K_p	Proportional gain
K_i	Integral gain
K_d	Derivative gain
γ	Adaptation rate
$\theta(t)$	Inverted pendulum tilt angle
F_x, F_y	Force components in x and y directions acting on inverted pendulum
Δf	Incremental frequency deviation in power system
ΔP_g	Incremental change in generator output
ΔX_g	Incremental change in governor valve position
ΔP_c	Incremental change in speed changer position
ΔP_d	Load disturbance
ΔP_{tie}	Incremental change in tie-line power flow
$ACE(t)$	Area control error
$\eta(\Delta X_g(t))$	Nonlinear function of governor valve position limits
ΔV_t	Terminal voltage of AVR system
ΔV_{ref}	Reference voltage of AVR system
ω	Angular velocity of DC motor

i_a	Armature current of DC motor
T_L	Load disturbance torque on DC motor shaft
$\phi(\bullet)$	Radial basis function, activation function
c	Radial centre
σ	Spread parameter of Gaussian function
W, w	Neural network weights
$RicP$	Positive definite symmetric constant matrix from solution of ARE
P	Positive definite symmetric constant matrix from solution of adaptive critic scheme
$RicK$	LQR gain from solution of ARE
K	Actor gain from solution of PI based adaptive critic scheme, LQR gain
V, V^{μ}	Cost function
$K_i, \ \mu, u$	Control policy
λ	Costate vector which denotes gradient of cost function with respect to x
L	Utility function of performance index
H	Hamiltonian functional
φ	Vector of activation functions
W_1, W_{c}	Critic NN weights
W_2, W_{a}	Actor NN weights

CHAPTER 1

INTRODUCTION

This chapter introduces the research work. The recent trends state-of-the-art literature review on the research work is presented. The research motivation, scope of work and problem formulation, author's contribution and thesis organization are also presented in this chapter.

1.1 MOTIVATION

The control systems are decision-making systems that are designed to provide autonomy to dynamic systems with desired system response and performance. In general, the control problem consists of obtaining dynamic models of systems, and using these models to determine control laws or strategies to achieve the desired system response and performance. It is a challenging task to design control algorithm that be simple and guarantee the stability and robustness in closed-loop system in real situations. With technological advancement, it is the need of time to design control systems that be capable to maintain acceptable performance levels under significant unanticipated uncertainties.

Most of the dynamical systems such as power systems, missile systems, robotic systems, inverted pendulum, industrial processes, chaotic circuits etc. are inherently highly nonlinear in nature. The control of such systems is a challenging task. The conventional control strategies are not able to provide proper control solution under the presence of nonlinearities and uncertainties in the controlled systems. Traditionally, the nonlinear control system design has been dominated by linear control theory, which relies on the key assumption of small range of operation of dynamic system for the linear model to be valid. The linear control theory has even provided many reliable and effective control systems, the demand for nonlinear control theory has recently been increasing due to several reasons. The high-performance real-world applications demands control systems with much more stringent design specifications with capabilities of handling nonlinearities and rejecting disturbances & uncertainties in a large range of operation of dynamic system. The optimal control theory synthesizes a control policy which results in the best possible behaviour with respect to the prescribed performance criterion satisfying some physical constraints. Adaptive control design which deals with uncertain systems or time-varying systems mainly systems with known dynamic structure but unknown constant or slowly-varying parameters; provides adaptive controllers that are inherently nonlinear whether developed for linear systems or for nonlinear systems. The nonlinear control design has become a relatively simple task by utilizing the advances in computational techniques. The emergence of intelligent control from the intelligent computational techniques such as neural networks and fuzzy logic systems has provided novel solutions for identification and control applications with a greater degree of autonomy.

For the high-performance control systems with stringent design specifications the intelligent adaptive optimal control is a viable recent approach. Intelligent adaptive optimal control has been emerged from the integration of adaptive control and optimal control methodologies with intelligent computational techniques. The performance investigation of intelligent adaptive optimal control of dynamical systems is motivating to consider in this research work. Thus, the applications of control schemes for dynamical systems control are implemented considering certain examples of linear and nonlinear dynamical systems to attempt this research investigation.

1.2 OVERVIEW

The control literature has influx of various control design concepts over the time. The advances in computational techniques and digital technology have made the control system automatic, fast, and more efficient & reliable, and offering a wide spectrum of choices for control schemes for various practical applications. The Proportional-Integral-Derivative (PID) control gives the simplest and yet the most efficient solution to various real-world control problems. Even more than 90% of industrial controllers are still implemented based on this, particularly at the lowest levels, as no other controllers match with the simplicity, clear functionality, applicability, and ease of use offered by the PID controller. Even though the demand of better level of performance with better functionality, under uncertainties motivate for the development of alternate algorithms of control design, especially at the higher levels in the controlled system.

The performance of the controlled systems is desired to be optimal which should be valid also when applied in the real situation. Optimal control designed with approximate system model will not give optimal performance when applied in the real situation, as it will not be sensitive to changes in system dynamics. Thus, for optimal performance of system in real situation, the adaptation of control parameters is desired. Adaptive control has objective of maintaining consistent performance of systems which have known structure but unknown constant or slowly time-varying parameter values. Adaptive control is an online design approach, and which is able to deal with uncertainties is generally not optimal in the sense of minimizing a formal performance function as specified for optimal control. Optimal control is offline, and needs the knowledge of system dynamics for its design. Thus, to have both features of control design, it is desired to design online adaptive optimal control. Adaptive optimal control is designed either by adding optimality features to adaptive control (e.g. the adaptation of control parameters is done by seeing the desired performance improvement reflected by an optimality criterion functional) or by adding adaptive features to optimal control (e.g. the optimal control policy is improved relative to the adaptation of the parameters of system model).

Emulating certain characteristics as generalization, heuristics, learning, adaptation, and evolution etc. of intelligent biological systems have lead to the recent developments of intelligent computational techniques. The integration of automatic control design with intelligent computational techniques has led to the intelligent control methodologies. The intelligent control has emerged as a viable recent approach giving novel solutions to the various control system problems. The intelligent controllers can drive uncertain complex systems with greater degree of autonomy than the available classical control schemes. The intelligent control is mostly task-oriented and rule-based as the dependencies are generally too complex to admit an analytical representation whereas the classical control is rooted in the differential or difference equations mathematical theory. The intelligent adaptive control and intelligent optimal control systems have emerged by the synthesis of intelligent computational techniques with adaptive control and optimal control design concepts. Neural networks ability to approximate any continuous function to a desired degree of accuracy through learning from data makes it suitable for various identification and control applications. Fuzzy systems ability to represent heuristic knowledge by fuzzy linguistic inputs and rules makes it suitable for various decision and control applications.

The reinforcement learning (RL) is a third approach to adaptive optimal control, strongly related with direct and indirect adaptive optimal control methods from a theoretical point of view. RL is a computational intelligence and machine learning approach based on the idea that successful control decisions may be remembered, by means of a reinforcement signal, such that they become more likely to be used a second time. RL algorithms provide a natural approach to solve the optimal control problem; can be implemented by means of function approximation structures, such as neural networks, which can be trained to learn the solution of Hamilton-Jacobi-Bellman (HJB) equation. There are basically two ways of solving the associated optimal control problem; one is Pontryagin's minimum principle and the other is Bellman's dynamic programming (DP). DP provides a computational technique to apply the principle of optimality to a sequence of decisions which define an optimal control policy. However, the offline backward-in-time solution of associated HJB equation has a computational complexity. Adaptive/approximate dynamic programming (ADP) overcomes these issues. The combining of concepts of ADP, RL, and backpropagation lead to the concept of adaptive critic design (ACD) as a way of solving dynamic programming problems in forward-in-time. ACD utilizes two parametric structures known as the actor and the critic. The actor parameterizes the control policy. The critic approximates a value-related cost function and captures the effect that the control law will have on the future cost which describes the performance of control system. At any given time, the critic provides guidance to improve the control policy, the actor to update the critic. Actor-critic online learning algorithms solve the optimal control problems.

Policy Iteration (PI), a computational intelligence technique refers to a class of RL algorithms consisting of two-step iteration: policy evaluation and policy improvement. These two steps of policy evaluation and policy improvement are repeated until the policy improvement step no longer changes the actual policy and thus converging to the optimal control. Instead of solving HJB equation by direct approach, the PI algorithm starts by evaluating the cost of a given initial admissible control policy, which is often accomplished by solving a nonlinear Lyapunov equation. This updated cost is then used to obtain an updated improved control policy which will have a lower associated cost. This is often accomplished by minimizing a Hamiltonian function with respect to the updated cost. Value Iteration (VI), does not require an initial stabilizing control policy. Generalized Policy Iteration is a family of optimal learning techniques which has PI at one extreme. In generalized PI, at each step one does not completely evaluate the cost of a given control, but only updates the current cost estimate towards that value. VI in fact belongs to the family of generalized PI techniques. The PI technique based adaptive critic scheme performs adaptive optimal control without using complete knowledge of the system dynamics. The online PI algorithm solves the optimal control problem, along a single state trajectory, does not require knowledge of the system internal dynamics, and thus giving a direct adaptive optimal control scheme. The adaptive optimal control using PI method relies on identification of the cost function associated with a given control policy followed by policy improvement in the sense of minimizing the identified cost, whereas the regular adaptive controllers rely on online identification of the system dynamics followed by model based controller design. By using neural networks to parameterize actor and critic for online implementation, this control scheme becomes a high-level intelligent control scheme. The synchronous policy iteration algorithm is a form of generalized continuous-time policy iteration algorithm. The online synchronous PI algorithm involves the simultaneous online training of both actor and critic networks. The initiation of the tuning process for the parameters of either actor or critic does not require complete convergence of the adaptation process of critic or actor respectively, which is regularly required in PI algorithms where only one neural network, critic or actor, is tuned until convergence, while the other neural network is held constant. This algorithm requires the complete knowledge of the system model. The adaptive critic design using online synchronous PI technique with neural networks approximations of cost and policy provides an intelligent adaptive optimal control.

In this research work the comprehensive performance investigation of intelligent adaptive optimal control of dynamical systems is considered. Though it is a wide area of research, this thesis attempts to contribute some research works into this subject matter. To attempt this research investigation nine research objectives are considered. In this thesis the performance investigations of various control schemes for continuous-time linear time-

invariant (LTI) systems and affine nonlinear systems considering applications of various general and practical systems examples are presented. In this thesis the comparative performance analyses of control schemes for certain applications are also presented. The state-of-the-art literature review, objectives of thesis, author's research contribution, and organization of thesis respectively are presented in the following sections of this chapter.

1.3 STATE-OF-THE-ART LITERATURE REVIEW: RECENT TRENDS

The recent trends state-of-the-art literature review on the subject is categorically presented in the following subsections.

1.3.1 System Modelling, Identification, and Control Design

The system modelling, identification, and control design are related problems in the real-world applications. The formulation of control system design strategies requires all the information characterizing the process to be controlled. The mathematical models describe the behaviour of dynamical systems. The system modelling and simulation are interrelated steps taken before a new prototype design is tested for any real-world application. The simulation of appropriate model of dynamical system provides numerical insights into its behaviour. The differential or difference equations which describe the system dynamics form the mathematical model. The mathematical model is not unique and depends on the perspective of system analysis and design, the nature of signals such as continuous-time or discrete-time, and the nature of system parameters etc. In the control system literature [1-10], the system dynamics is widely represented by the differential or difference equations, the state variable representation, the transfer function model, and state-space model; and the system analysis is performed using frequency-domain or time-domain approaches. The appropriate mathematical model is considered for control system design and analysis as per its perspective. In case not all state variables of system are measurable, the state observer is designed to obtain the observed model [2-9]. In case system dynamics is unknown or system parameters are all unknown or some are known and some are unknown, the system model is identified using some procedures of system identification and parameter estimation [11-14].

The real-world applications have highly nonlinear dynamical systems exhibiting complex behaviour. The control design of nonlinear dynamical systems is traditionally, based on linearized models and linear control theory. However, the linear controller may exhibit significant performance degradation or even instability due the nonlinearities and uncertainties present is the system. Thus, for nonlinear systems control design the most appropriate models are nonlinear ones. Recently, intelligent computational techniques have been used to develop novel system models and controllers for various identification and control applications. The soft computing (SC) techniques such as neural networks (NN), and

fuzzy logic (FL) etc. have given novel solutions of modelling, identification and control problems for various nonlinear dynamical systems [13-16].

Most of the real-world dynamical systems such as biomedical systems, complex industrial processes, electrical power systems, ecological & environmental systems, and missile guidance systems etc. are the large-scale systems. A system is said to be large-scale system if it can be decoupled or partitioned into a number of interconnected systems or small-scale systems for either computational or practical reasons. Alternatively, a system is large-scale when its dimensions are so high such that the conventional techniques of modelling, analysis, control system design and computation fail to give accurate solutions with reasonable computational efforts. The mathematical models of large-scale systems are of high-order which may pose difficulties in its analysis, synthesis, identification and control design. In order to deal with such higher-order systems, an obvious method is to approximate it by a low-order model which reflects the important characteristics of the original high-order system as time constant, damping ratio, natural frequency, etc. The model order reduction techniques simplify the understanding and analysis of the system, reduce the computation time and complexity, and economize the hardware synthesis. The reduced order models and the model order reduction techniques have been widely used in control engineering environment. The various model order reduction techniques are presented in the literature [17-20]. The development of computational techniques, simulation techniques, and digital control design techniques have simplified the modelling, analysis, and control design with real-time practical implementation [3-5, 7-10, 21, 22].

The various methods of control design for dynamical systems are presented in the literature [7, 9, 10, 15]. Some of these control design methods are:

(i) Classical control: lead-lag compensation, Bode and Nyquist methods, root-locus design, etc.
(ii) Proportional-integral-derivative (PID) control (or its variants P or PI).
(iii) State-space methods: State feedback control, observers, etc.
(iv) Optimal control: Linear quadratic regulator (LQR), Pontryagin's minimum principle or dynamic programming, etc.
(v) Robust control: H2 or H∞ methods, quantitative feedback theory, loop shaping, etc.
(vi) Nonlinear methods: Feedback linearization, Lyapunov redesign, sliding mode control, backstepping, etc.
(vii) Adaptive control: Model reference adaptive control (MRAC), self-tuning regulator (STR), nonlinear adaptive control, etc.
(viii) Discrete event systems: Petri nets, supervisory control, Infinitesimal perturbation analysis, etc.
(ix) Intelligent control: Neural network control, fuzzy control etc.

1.3.2 PID Control

The PID control is the most simple feedback control scheme for linear dynamical systems. The PID or its variants P or PI are widely used in industrial control applications. The design and tuning of PID controllers have been continued to be research problems. The state-of-the-art of PID control system analysis, design, and technology, and its future issues are presented in [23, 24]. The empirical tuning method proposed by Ziegler-Nichols (1942) is traditionally used for linear time-invariant (LTI) systems. The Ziegler-Nichols tuning methods give moderately good tuning in only restricted situations. Since PID controller is linear, it does not give satisfactory performance for nonlinear systems. The development of novel design and automatic tuning methods are important research problems. Using various control and optimization strategies, various PID control design and tuning methods such as auto-tuning, self-tuning, etc. are presented in the literature. A method to compute the entire set of stabilizing PID controller parameters for an arbitrary (including unstable) linear time delay system is presents in [25]. [26] presents a new control structure with a tuning method to design a PID load frequency controller for single area and multi-area power systems in the presence of uncertainties in plant parameters using relay based identification technique to estimate power system dynamics. The fractional order PID (FOPID) control design and tuning algorithms have taken attention of researchers recently, and are presented in literature [27, 28]. The tuning rules of Ziegler-Nichols type for fractional PIDs are presented in [27]. The FOPID to improve stability and response of load frequency control (LFC) and automatic generation control (AGC) system is presented in [28]. Recently various intelligent computational techniques have been applied for PID control design for various applications and are presented in literature [29-39]. To determine the PID controller parameters online, a fuzzy gain scheduling scheme is presented in [29], and a function-based evaluation approach for fuzzy-PID controller is presented in [30]. A parallel fuzzy P + fuzzy I + fuzzy D controller is presented in [31]. A robust intelligent tracking controller (RITC) for a class of unknown nonlinear systems comprising a neural controller using a PID-type learning algorithm in the sense of Lyapunov function, and the robust controller to achieve L_2 tracking performance with desired attenuation level is presented in [32]. For speed and position control of DC motor, an optimal integral state feedback control with PID controller using GA and Kalman filter is presented in [33]. An intelligent PID control design scheme using the least squares-support vector machine (LS-SVM) is presented in [34]. The optimal fuzzy PID controller design using particle swarm optimization (PSO) with Q-learning algorithm for active automobile suspension system presented in [35]. The PID control designs for AVR systems are presented using PSO-PID in [36], PSO-fuzzy-PID in [37, 38], the velocity relaxed and craziness based PSO-PID in [39], and improved evolutionary non-dominated sorting genetic algorithm II (NSGA II) based fractional order (FO) $PI^{\lambda}D^{\mu}$ in [40, 41].

1.3.3 Optimal Control

The performance of the dynamical systems being controlled is desired to be optimal. There are many optimization & optimal control techniques which are presented in the literature for linear & nonlinear dynamical systems [6, 7, 9, 10, 42, 43]. The optimal control provides the best possible behaviour of dynamical system with respect to the specified performance criterion satisfying some physical constraints. The performance criterion (or cost function) may have several forms based on the control application. The optimal control theory which is an extension of calculus of variations is a mathematical optimization approach to derive optimal control policy. The Euler-Lagrange equation provides the necessary condition for extremum of a functional using first variation method, and the sufficient condition using second variation of a functional. The Lagrange multiplier method is a powerful method in optimization. The Hamiltonian formalism is another powerful approach of optimization and to find optimal control policy. The optimal control theory is mainly due to the works of maximum principle of L. S. Pontryagin (1956), which provides a necessary condition for optimality; and dynamic programming of Richard E. Bellman (1953), that resulted in the Hamilton-Jacobi-Bellman (HJB) equation, which is a sufficient condition. R. E. Kalman (1960) provided the linear quadratic regulator (LQR) and linear quadratic Gaussian (LQG) theory to design optimal feedback controllers. He presented optimal filtering and estimation theory leading to Kalman filter. The solution of algebraic Riccatti equation (ARE) determines the LQR and LQG which give infinite horizon optimal control solution for linear time-invariant systems [42, 43]. The H_2 and H_∞ optimal control theories are robust optimal control theories which are developed using minimization of H_2 and H_∞ norms respectively [44]. The problems of optimization of performance criterion and optimal control policy design have continuously been the prime concern for control system designers and researchers. The optimal control designs using certain approaches for certain applications are presented in [45-49]. The optimal control problem with a continuous inequality constraint on the state and the control is presented in [45]. A LQR design for linear stochastic systems with probabilistic uncertainty in the parameters is presented in [46]. The LQG design for optimal control of pneumatic Stewart-Gough platform is presented in [47]. The optimal control design using applications of intelligent computational techniques are also presented in [35, 48, 49]. The optimal control for stochastic linear singular system using neural networks solution of the matrix Riccati differential equation (MRDE) is presented in [48]. The optimal control design using particle swarm optimization (PSO) for power system stabilizers is presented in [49].

1.3.4 Adaptive Control

The control of dynamical systems in the presence of uncertainties, structural perturbations, and environmental variations is a major problem of real-world applications. The control solution for such systems is obtained using adaptive control algorithms. There

are many adaptation & adaptive control techniques which are presented in the literature for linear & nonlinear dynamical systems [50-54]. Adaptive control is an online design approach which maintains consistent performance of systems which have known structure but unknown constant or slowly time-varying parameter values. An adaptive system has an adaptation mechanism that automatically compensate for variations in operating conditions, process dynamics or disturbances, in order to maintain an optimal performance of the system. The high-performance applications motivated the research & development in adaptive control schemes. Some fundamental developments in adaptive control are: Gain-scheduling in 1950s for aircraft autopilots, optimalizing control or peak-holding control by Draper and Li (1951), control system with adaptive characteristics by Benner and Drenick (1955), model reference adaptive control (MRAC) by Witaker et al. (1958), self-oscillating adaptive system by Li and van der Velde (1960), variable structure systems by Petrov et al. (1963), sliding mode control (SMC) by A. F. Filippov (1960s), Dual control by the application of dynamic programming in control design for systems with probabilistic uncertainties by Bellman (1960-61), partitioning approach by Lainiotis (1971), self-tuning regulator (STR) by Astrom and Wittenmark, and augmented error approach for stable MRAC system design by Monopoli (1974) etc. [50]. Other developments in control theory such as the state-space and stability theory, stochastic control theory, proofs for stability of adaptive systems, control design using Lyapunov function, system identification, robust control, and learning control, etc. made further advances to adaptive control design methods. The adaptive control using feedback linearization that transforms nonlinear adaptive control problem into a linear adaptive control problem to apply linear control methods is developed as an approach for nonlinear control design for feedback linearizable nonlinear system with known dynamics. The indirect adaptive control (IAC) and direct adaptive control (DAC) schemes are structurally two basic approaches of adaptive control which were developed based on presence or non-presence of system identification in control design. MRAC and STR both with IAC and DAC both are widely explored adaptive schemes [52]. With significant theoretical and algorithmic advancements the adaptive control systems have wide range of current and potential industrial applications [50, 52]. The adaptive control problem has continuously been the prime concern for control system designers and researchers. [53-55] present the state-of-the-art of theory and applications of adaptive control. A comparative study of model reference adaptive control and fuzzy model reference learning control techniques for inverted pendulum system is presented in [56]. The sliding mode control which is a robust adaptive control approach with various methodologies for various applications is presented in [57-64]. Recently, the intelligent control techniques are widely used in the framework of adaptive control design for various applications [15, 61-64]. The

literature review on sliding mode control using soft computing techniques is presented in [63, 64].

1.3.5 Intelligent Control

The recent development in the area of artificial intelligence (AI), such as artificial neural network (ANN), fuzzy logic (FL), expert systems (ES), evolutionary computational techniques such as genetic algorithm (GA), and particle swarm optimization (PSO) etc., and machine learning such as support vector machine (SVM), etc., all of which belong to soft computing techniques, and are commonly known as intelligent computational techniques have led to the development of intelligent control schemes by the integration of these intelligent computational techniques with automatic control design [7, 9, 10, 15, 16, 65-77].

The intelligent computational techniques have continuously attracted researchers of diverse applications areas seeking intelligent novel solutions of different problems. The application of intelligent computational techniques in control engineering framework has continuously been attracted the control researchers for getting intelligent control solutions for various dynamical systems applications. Recently several researchers have explored the intelligent control methodologies for various applications [7, 9, 10, 13-16, 29-43, 48, 49, 61-79, 81-99].

Neural networks (NN) [7, 10, 15, 43, 65, 66, 69-74] represent an important paradigm for classifying patterns or approximating complex nonlinear process dynamics. These properties clearly indicate that NN exhibit some intelligent behaviour, and are good candidate models for nonlinear processes, for which no perfect mathematical model is available. The NN have given novel solutions to various dynamical systems modelling, identification, and control problems [7, 10, 13-15, 43, 65, 66, 69-74, 76, 78, 79, 82, 87, 98]. Fuzzy logic (FL) systems [7, 9, 10, 15, 43, 66, 67, 69-74] attempt to approximate the human knowledge and the associated reasoning process (i.e. process of knowledge representation). In this sense the FL systems in intelligent control serves to represent and process the control knowledge of a human in a given plant. The FL systems have given promising solutions to various modelling and control problems of various applications [7, 9, 10, 15, 29-31, 43, 61, 66-74, 78, 81, 82, 84-86, 90, 97, 99]. Genetic algorithms (GA) [7, 15, 43, 66, 70, 72, 94] represent an optimization approach which is analogous to biological evolution through natural selection, crossover, and mutation, where a search is made to “evolve” a solution algorithm that will retain the “most-fit” components. These features make GA applicable in intelligent control, particularly when optimization is an objective. GA has been used in various optimization and control problems [7, 15, 40, 41, 43, 66, 70, 72, 82, 94]. Particle swarm optimization (PSO) [95] is an evolutionary computation technique that was inspired by the social behaviour of bird flocking and fish schooling; is a population-based optimization algorithm that is not much affected by the size and nonlinearity of the problem, and can converge to the optimal solution

in many problems where most analytical methods fail to converge. The PSO has given novel solutions to various optimization and control problems of dynamical systems [35-39, 49, 95, 96]. Support vector machine (SVM) [7] is a supervised learning algorithm that automatically improves the performance with experience at a given task. SVM has given novel intelligent control solutions of various applications [7, 34].

In recent years, the hybrid intelligent systems by the integration of intelligent computational techniques (NN, FL, GA, etc.) in a complementary hybrid framework (hybrid neuro-fuzzy, fuzzy-GA, neuro-GA, neuro-fuzzy-GA systems etc.) have been the research trends applications for solving complex problems [15, 16]. Adaptive neuro-fuzzy inference system (ANFIS), a hybrid neuro-fuzzy system has been widely used for various modelling, identification, and control applications [15, 16, 88, 91, 92, 100]. Other hybrid intelligent systems as PSO-fuzzy [35, 37], fuzzy-wavelet neural [14], GA-fuzzy [101], GA-fuzzy-neural [102], and RBFNN-PSO [96] have also been used for various applications and presented in literature.

The intelligent control is a wider area of research and applications that uses the various control theories and computational intelligent systems. The literature reviews on intelligent control specifically with adaptive control, optimal control, and adaptive optimal control are presented in the following subsections.

1.3.6 Intelligent Adaptive Control

The intelligent adaptive control is emerged from the integration of adaptive control schemes with intelligent computational techniques. The different intelligent computational techniques such as NN, FL, GA etc or hybrid systems such as ANFIS etc are applied in adaptation mechanism of adaptive control schemes mainly MRAC or STR and in IAC or DAC control structures, resulting in intelligent adaptive control. The intelligent adaptive control has continuously been an emerging research area. There are various industrial applications of intelligent adaptive control systems. The intelligent adaptive control schemes using various methodologies for various applications have been presented in [13, 15, 16, 32, 35, 67-74, 91-93, 100-115].

1.3.7 Intelligent Optimal Control

The intelligent optimal control is emerged from the integration of optimal control schemes with intelligent computational techniques. The different intelligent computational techniques or hybrid systems are used to optimize the specified performance criterion of optimal control system. The evolutionary computational techniques have played a major role in the optimization and optimal control applications. Recently several researchers have explored the intelligent computational techniques mainly evolutionary computational

techniques using various optimal control methodologies for dynamical systems applications [7, 15, 36-41, 116-121].

1.3.8 Intelligent Adaptive Optimal Control

The intelligent adaptive optimal control is emerged from the synergistic integration of adaptive control algorithms and optimal control theory with intelligent computational techniques having all the features of three technologies. The adaptive optimal control is designed by adding either optimality features to adaptive control or adaptive features to optimal control. The adaptive optimal control design with application of intelligent computational techniques gives intelligent adaptive optimal control. For a given dynamical plant and the corresponding performance index, there are basically two ways of solving the associated optimal control problem; one is Pontryagin's minimum principle and the other is Bellman's dynamic programming [6, 7, 10, 42, 71]. Recently several researchers have tried to explore the intelligent computational techniques with adaptive and optimal control design by applying certain methodologies for certain applications [7, 13, 15, 16, 34, 35, 43, 48, 49, 61-64, 67-74, 79, 87-89, 91-93, 100-113, 116-122]. A time-optimal control law using neural network with dynamic programming and adaptive control law using neural network for conical tank level control is proposed in [122].

The recent development in the area of computational intelligence and machine learning techniques provided a third approach to adaptive optimal control namely the reinforcement learning (RL) [123-125] based on the idea that successful control decisions may be remembered, by means of a reinforcement signal, such that they become more likely to be used a second time; the idea originating from experimental animal learning, where it has been observed that the dopamine neurotransmitter acts as a reinforcement informational signal which favours learning at the level of the neuron. The RL algorithm which was generally to find optimal control policies for Markovian systems with discrete state and action spaces is strongly related with direct and indirect adaptive optimal control methods from a theoretical point of view. The RL algorithm characterizes a learning problem which is in fact the adaptive optimal control problem to find a control policy based on reward information which characterizes the performance of a given controller. The RL algorithms provide a natural approach to solve the optimal control problem; can be implemented by means of function approximation structures, such as neural networks, which can be trained to learn the solution of Hamilton-Jacobi-Bellman (HJB) equation [123-126]. Bellman's dynamic programming (DP) provides a computational technique to apply the principle of optimality to a sequence of decisions which define an optimal control policy or trajectory giving an optimal controller in state feedback form, may provide the best approach to find optimal control strategies for highly constrained nonlinear systems. Even though a feedback form is robust to noise and model uncertainties, however, the solution of the HJB equation associated with

DP has a computational complexity that grows exponentially with the number of state variables, the problem known as “curse of dimensionality” [124, 127-130], and it is an offline process where the problem is solved from the end point and is approached in backward direction. To overcome these issues, in 1977, Werbos proposed Adaptive/ Approximate Dynamic Programming (ADP) [124-129, 131-141]. Combining the concepts of ADP, RL, and backprogation [142, 143], he introduced an approach for ADP called adaptive critic design (ACD) as a way for solving dynamic programming problems forward-in-time. Werbos defined actor-critic online learning algorithms to solve the optimal control problem based on Value Iteration (VI), and defined a family of VI algorithms as ADP algorithms. He used a critic neural network for value function approximation (VFA) and an actor neural network for approximation of the control policy. Based on functional architectures and critic choices, the several versions of ACD are presented recently in the literature [123, 124, 129, 130, 142-178]. Some works present ACDs as supervised learning systems and reinforcement learning systems [124, 143]. The online learning algorithms such as Policy Iteration (PI) [123-126, 155-167], Value Iteration (VI) [124, 125, 168], and Generalized Policy Iteration (GPI) [154, 157], have given solution of the optimal control problems for linear [155-158, 161, 166, 169] and nonlinear systems [123, 154, 158-165, 167, 168]. GPI is a family of optimal learning techniques which has PI at one extreme [125, 158] and VI belongs to it. In GPI, at each step the cost of a given control is not completely evaluated, but only the current cost estimate towards that value is updated. Likewise, one does not fully update the control policy to the greedy policy for the new cost estimate, but only updates the policy towards the greedy policy. ACD utilizes two parametric structures known as the actor and the critic. The actor parameterizes the control policy. The critic approximates a value-related cost function and captures the effect that the control law will have on the future cost which describes the performance of control system. At any given time, the critic provides guidance to improve the control policy, and the actor to update the critic.

Policy Iteration (PI) algorithms consist two-step iteration: policy evaluation and policy improvement. These two steps of policy evaluation and policy improvement are repeated until the policy improvement step no longer changes the actual policy and thus converging to the optimal control. Instead of solving HJB equation by direct approach, the PI algorithm starts by evaluating the cost of a given initial admissible control policy, which is often accomplished by solving a nonlinear Lyapunov equation. This updated cost is then used to obtain an updated improved control policy which will have a lower associated cost [125, 155-157, 159-162]. This is often accomplished by minimizing a Hamiltonian function with respect to the updated cost [124, 125, 155, 156, 158-160, 166]. This is the so-called `greedy policy' with respect to the updated cost [158]. It is noted that the infinite horizon cost can be evaluated only in the case of admissible and stabilizing control policies. Admissibility is in fact

a condition for the control policy which is used to initialize the algorithm. PI algorithm requires an initial stabilizing control policy, but VI does not require an initial stabilizing control policy [124, 125, 158].

Adaptive optimal control using various approaches has been presented recently in literature [122-178]. Adaptive critic designs for various control applications applying certain approaches in both discrete-time and continuous-time frameworks are described recently in certain papers [122-126, 128-130, 132, 135-178]. The ACD using SVM based tree type neural network as critic is presented in [147]. The application of radial basis function neural network for ACD is presented in [148]. The ACD using PSO based actor and neural network based critic is presented in [149]. Adaptive optimal control using PI technique is presented for linear systems in [155-158, 161, 162, 166, 169], and for nonlinear systems using neural networks in actor-critic configuration in [123, 158-165, 167]. Adaptive optimal control using synchronous PI technique is presented for linear systems in [157, 158, 161, 162], and for nonlinear systems in [158, 161, 162]. As from applications point of view PI algorithm is implemented for optimal load frequency control of a power system in [156, 169], for F-16 aircraft in [155, 158, 162], for linear hyperbolic PDE systems in [166], and for general nonlinear systems in [158-162]. The convergence guarantees of the continuous-time PI technique to the optimal control are given in [123, 154-162, 167-169]. The optimal control for discrete-time affine nonlinear systems using VI technique is presented in [168]. A simplified version of adaptive critic architecture called ‘single network adaptive critic (SNAC)’ which uses only one network instead of two required in a standard adaptive critic design for affine systems is proposed in [172-178]. The SNAC design is presented using neural networks in [172-177], and using Takagi-Sugeno fuzzy systems in [178]. The applications of SNAC design are presented for linear systems in [174], and for nonlinear systems in [172-178].

1.3.9 Dynamical Systems Applications

The literature reviews on particularly the control of dynamical systems considered in this thesis for control applications are briefly presented in this subsection. The various dynamical systems problems for control applications considered in this research study are: nonlinear inverted pendulum-cart system, nonlinear surge tank liquid level process system, automatic generation control of two-area power system linear and nonlinear both, load frequency control of single-area power system, automatic voltage regulator system, DC motor speed control system, spring-mass-damper mechanical system, single-link robotic manipulator, Vander Pol’s oscillator, and various general linear and nonlinear systems.

The inverted pendulum system is a classical benchmark for the implementation of control schemes and performance analysis. The inverted pendulum-cart dynamical system has been taken by certain researchers for implementing the various control schemes which are presented in literature [2, 9, 56, 61, 62, 85, 86, 179-181]. The swing-up of inverted

pendulum by energy control is presented in [179]; and the swing-up and stabilization of a cart-pendulum system under restricted cart track length is presented in [180]. The fuzzy control of inverted pendulum is presented in [85, 86]; and fuzzy hierarchical swing-up and sliding position control in [61]. The adaptive fuzzy sliding mode control of inverted pendulum is presented in [62]. The control of inverted pendulum cart system using a Lyapunov function based approach is presented in [181]; and using model reference adaptive control and fuzzy model reference learning control techniques for comparative study in [56]. The nonlinear surge tank liquid level process system is also a benchmark control problem that is discussed in [71].

The automatic generation control (AGC) of power systems is a widely studied problem. The load frequency control (LFC) is the main subsystem of AGC system which is an important problem in power systems [182-187]. The various control schemes for AGC or LFC systems are presented in the literature [26, 28, 97-99, 182-186, 188-198]. The fuzzy logic control of AGC systems is presented in [97, 186]. The adaptive LFC using dynamic neural network is presented in [98]. The PID control of LFC system is presented in [26]; and FOPID control of LFC and AGC systems is presented in [28]. The robust LFC design using fuzzy logic for uncertain power systems is proposed in [99]. [188-198] presents the robust LFC design for power systems using various approaches. The H_∞ control approach is presented in [195-198].

The automatic voltage regulator (AVR) and power system stabilizer (PSS) are important subsystems of power system to maintain voltage limits and stability [182-184]. Recently the control design using various approaches for AVR system and PSS has attracted researchers [36-41, 175, 176, 199-203]. PSO-PID controller design and comparison of performance with GA-PID for AVR system using a new performance criterion for obtaining optimal controller parameters is presented in [36]. The AVR system using PSO-fuzzy-PID controller is presented in [37]. The performance of intelligent fuzzy based coordinated control of the AGC loop and the excitation loop equipped with PID controlled AVR system and PSS controlled AVR system is investigated in [38] using craziness based PSO (CRPSO) as optimizing tool to get the optimal tuning of PSS parameters as well as the gains of PID controllers and for on-line, off-nominal operating conditions to obtain the off-nominal optimal gains of PID controllers and PSS parameters the Takagi Sugeno fuzzy logic (TSFL) has been applied. [39] presents PID controlled AVR system, and PSS controlled AVR system using velocity update relaxation PSO (VURPSO) and position, velocity updating strategy and craziness PSO (CRPSO), and also compares the performance with GA based approach. A fractional order (FO) $PI^{\lambda}D^{\mu}$ controller for an AVR system using an improved evolutionary non-dominated sorting genetic algorithm II (NSGA II) that is augmented with a chaotic map for greater effectiveness for the multi-objective optimization problem, is presented in [40] and a

frequency domain design approach for a fractional order PID (FOPID) controller for an AVR system using NSGA-II augmented with a chaotic henon map that is used for the multi-objective optimization based design procedure, is presented in [41]. The steady state voltage stability assessment of power systems with AVR voltage limits is discussed in [199]. The effect of time delays on the stability of generator excitation control system is investigated in [200]. A parameter tuning method for AVR system using online measured data of the excitation control system with parameter optimization technique is presented in [201]. The design of PSS for a small-signal stability study using sliding mode control (SMC) techniques is presented in [202]. The design of PSS using single network adaptive critic (SNAC) is proposed in [175, 176]. In [203] the design of a sub-optimal nonlinear feedback controller for power systems based on the approximate solution of the HJB equation is presented.

The DC motor speed control system is discussed in [7, 204]. The single-link manipulator [71, 178], mass-spring-damper mechanical system [1], and Vander Pol's oscillator [173, 178] are also taken as benchmark systems for the application of various control schemes.

1.4 OBJECTIVES OF THESIS

Based on the recent trends state-of-the-art of the control system design and applications the objectives of this research work have been proposed. The emerging trends of optimal control, adaptive control, and intelligent control systems are motivating. The dynamical systems such as inverted pendulum, industrial processes, and power systems have wide applications in academics & industry. Thus the emerging trends of control theory and applications provide a lot of scope of research works on performance investigation of control schemes for dynamical systems and exploring novel control algorithms. To carry out the research objectives the following problem formulation and scope of work have been considered in this thesis.

1.4.1 Problem Formulation

Consider the general nth order dynamical systems described in state-space form as following:

$$\dot{x}(t) = f(t, x(t), u(t)) \quad , \quad x(t_0) = x_0 \tag{1.1}$$

$$y(t) = h(t, x(t), u(t)) \tag{1.2}$$

where f and h are smooth nonlinear vector valued functions and are also assumed to be bounded, $x(t) = [x_1, x_2, ..., x_n]^T \in \mathcal{R}^n$ is the state vector, $y(t) = [y_1, y_2, ..., y_p]^T \in \mathcal{R}^p$ is the output vector, and $u(t) = [u_1, u_2, ..., u_m]^T \in \mathcal{R}^m$ is the control input vector.

For the control input affine dynamic systems (1.1) and (1.2) can be written as

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))u_i(t) \quad \text{or}$$

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \quad , \quad x(t_0) = x_0 \tag{1.3}$$

$$y(t) = h(x(t)) \tag{1.4}$$

where $f(x(t)) \in \mathbb{R}^n$, and $g(x(t)) \in \mathbb{R}^{n \times m}$ are smooth nonlinear functions. It is assumed that these functions will be bounded. In order to system (1.3) and (1.4) to be controllable, it is required that $g(x) \neq 0$ for x in a certain controllability region $U_C \subset \mathbb{R}^n$. It is assumed that $0 < g(x) < \infty$ for $x \subset U_C$.

If the functions $f(x(t))$, $g(x(t))$, and $h(x(t))$ are linear functions then (1.3) and (1.4) approache to represent the linear dynamical systems. The state-space model of general continuous-time linear time-invariant (LTI) systems is as follows:

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1.5}$$

$$y(t) = Cx(t) + Du(t) \tag{1.6}$$

with the initial condition $x(t_0) = x_0$, $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^p$, and $u(t) \in \mathbb{R}^m$ are the state, output and input vectors respectively. Similarly, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, and $D \in \mathbb{R}^{p \times m}$ are constant matrices.

Consider the state regulation problem for dynamical systems. Let $x_d(t)$ is the desired state then relative to desired state $x_d(t)$ the instantaneous system state error is given by

$$e(t) = x(t) - x_d(t) \tag{1.7}$$

If the system coordinates are transformed such that the desired state becomes the origin of the state space, then the new state $x(t)$ itself will be error. The control objective is to determine a control law $u^*(t)$ such that the system state is regulated to the desired state with minimizing a performance index

$$J = \frac{1}{2}\int_0^\infty [x^T(t)Qx(t) + u^T(t)Ru(t)]dt \tag{1.8}$$

and the instantaneous system state error vanishes as $\lim_{t \to \infty} \| e(t) \| = 0$ and thus the system error dynamics is asymptotically stable. The matrix Q is nxn real, symmetric, positive definite (or positive semi-definite) constant matrix and matrix R is mxm real, symmetric, positive definite constant matrix.

In case of the output tracking problem the control objective is to force the system output vector y to follow a given reference signal vector y_r under the constraint that all signals involved must be bounded. The state regulation problems and state tracking or output tracking problems are convertible to each other.

In this research work, an infinite horizon intelligent adaptive optimal control solution for dynamical systems is proposed. The performance investigations of optimal control, adaptive control, intelligent control, adaptive optimal control, and intelligent adaptive optimal control schemes for dynamical systems are considered in this research work.

1.4.2 Scope of Work

The scope of work undertaken in this research investigation is as following:

1. The exhaustive literature survey relevant to the research area of intelligent adaptive optimal control of dynamical systems.
2. Mathematical modelling of dynamical systems both LTI systems and nonlinear systems. Consider the continuous-time LTI dynamical systems preferably power systems automatic generation control, load frequency control, automatic voltage regulator, dc motor speed control systems etc. and continuous-time affine nonlinear dynamical systems preferably inverted pendulum, robotic manipulator, Vander Pol's oscillator, etc. for applications of control schemes.
3. The modelling, simulation, control design, and practical implementation of control schemes.
4. The performance investigation of optimal control scheme for dynamical systems.
5. The performance investigation of adaptive control scheme for dynamical systems.
6. The performance investigation of intelligent control techniques for dynamical systems using intelligent computational techniques particularly neural networks, and fuzzy logic systems.
7. The performance investigation of adaptive optimal control schemes for dynamical systems.
8. The performance investigation of intelligent adaptive optimal control of dynamical systems.
9. The comparative performance investigation with conventional approaches.
10. The investigation on stability of the proposed control system with robust tracking and state regulation performance.
11. The investigations of novel control algorithms for dynamical systems.

1.5 CONTRIBUTIONS OF RESEARCH WORK

This research work contributes by presenting the comprehensive performance investigation of the different control schemes for continuous-time linear time-invariant (LTI) systems and affine nonlinear systems. The research objectives which have been considered and contributions made in this research work are highlighted as following:

1. Optimal control of nonlinear inverted pendulum dynamical system using PID controller & LQR.
   - The modeling and control design of nonlinear inverted pendulum-cart dynamic system using PID controller & LQR for both cases of without and with disturbance input.
   - A simple optimal control approach for nonlinear dynamical systems using PID controller and LQR is proposed.
   - The comparative performance investigation of PID control, 2PID+LQR control, and 1PID+LQR control for nonlinear inverted pendulum dynamical system is presented.
2. Intelligent control of nonlinear inverted pendulum dynamical system using Mamdani and TSK fuzzy inference systems.
   - The modeling, control design and performance analysis of fuzzy control for nonlinear inverted pendulum-cart dynamic system for both cases of without & with disturbance input.
   - Fuzzy control using Mamdani and Takagi-Sugeno-Kang (TSK) fuzzy inference systems (FIS) have been implemented.
   - The comparative performance analysis of PD-fuzzy control using Mamdani FIS, and direct fuzzy control using TSK FIS has been presented. The comparative performance analysis with PID control method is also presented.
3. Optimal control design using LQR for automatic generation control (AGC) of two-area interconnected power system.
   - The modelling and optimal control design using LQR for AGC of two-area interconnected power system is presented.
   - The performance analyses of control schemes of conventional integral control, optimal control using LQR, and optimal control using integral controller and LQR for AGC system of two-area interconnected power system are presented.
4. Intelligent control using fuzzy-PI controller for automatic generation control of two-area interconnected nonlinear power system.
   - The modeling, control and performance analysis of AGC of two-area interconnected nonlinear power system using fuzzy-PI controller is presented.
   - Fuzzy-PI controller using Mamdani FIS is implemented.
   - The performances of fuzzy-PI control and PID control methods are compared for AGC system of two-area interconnected power system for both linear and nonlinear system models.
5. Intelligent control of process system using radial basis function neural networks.
   - The indirect adaptive control of a nonlinear process system, a surge tank liquid level control system using radial basis function neural networks (RBFNN) is presented.

- Feedback linearization method is implemented for control design using RBFNNs.
- The performance analysis of intelligent control of affine nonlinear systems using RBFNNs is presented.

6. Adaptive optimal control using policy iteration technique for LTI systems.
   - The adaptive optimal control using policy iteration (PI) technique based adaptive critic scheme for continuous-time LTI systems is presented.
   - The PI technique based adaptive critic control scheme is implemented for applications of general and practical examples of LTI systems- (i) a general LTI SISO system, (ii) a higher order LTI system- a mechanical system, (iii) load frequency control (LFC) of power system, (iv) automatic voltage regulator (AVR) of power system, and (v) DC motor control system.
   - The comprehensive performance analysis of adaptive optimal control of continuous-time LTI systems with adaptive critic scheme using PI technique is presented.
   - The performance analysis of control scheme is presented for both of system models of LFC system without and with integral control.
   - The performance analysis of control scheme is presented for both of system models of AVR system neglecting sensor dynamics and including sensor dynamics.
   - The performance analysis of control scheme is presented for both of system models of DC motor control system without and with integral compensator.
   - The effects of including integral control in LFC system, sensor dynamics in AVR system and integral compensator in DC motor control system are analysed.
   - The comparative performance investigation of adaptive critic control scheme and linear quadratic regulator is also presented.
   - The performance of adaptive critic control scheme under structural change in system dynamics is investigated.
   - The PI based adaptive critic controller adapts the change in system parameters in real situation at any moment of time is demonstrated.
   - The control scheme is partially model-free is demonstrated.
   - The novel adaptive optimal control design for AVR system is proposed.
   - The novel adaptive optimal control design for DC motor control system is proposed.
   - The novel modeling of DC motor control system with integral compensator is proposed.

7. Adaptive optimal control using policy iteration technique for affine nonlinear systems.
   - The adaptive optimal control using PI technique based adaptive critic scheme for continuous-time affine nonlinear dynamical systems is presented.
   - The neural network approximation of cost function is used for online implementation of PI algorithm with the Kronecker product quadratic polynomial basis vector considered to be the activation function vector.
   - The application of control scheme is implemented considering state regulation problem for two general and three practical examples of affine nonlinear systems- (i) single-link manipulator, (ii) inverted pendulum-cart system, and (iii) Vander Pol oscillator system.
   - The performance of online PI based adaptive critic control scheme which gives intelligent adaptive optimal control solution for affine nonlinear systems is analysed.
   - The online PI algorithm provides a partially model-free approach is demonstrated.
8. Intelligent adaptive optimal control using synchronous policy iteration technique for LTI systems.
   - The intelligent adaptive optimal control using synchronous PI technique for LTI systems applications is presented.
   - The cost function and control policy are approximated using neural networks.
   - The application of online synchronous PI based control scheme is implemented for two practical examples of LTI systems- (i) load frequency control of power system, and (ii) automatic voltage regulator of power system.
   - The performance of synchronous PI based control scheme is analysed for both of system models of LFC system without and with integral control.
   - The performance of synchronous PI based control scheme is analysed for both of system models of AVR system neglecting sensor dynamics and including sensor dynamics.
   - The performance of synchronous PI based adaptive critic control scheme under structural change in system dynamics is investigated.
   - For convergence of critic NN and actor NN, the requirement of persistence of excitation (PE) condition is investigated.
9. Intelligent adaptive optimal control using synchronous policy iteration technique for affine nonlinear systems.
   - The intelligent adaptive optimal control using synchronous PI technique for continuous-time affine nonlinear systems applications is presented.
   - The neural networks approximations of critic function and actor function are used.
   - The application of online synchronous PI based control scheme is implemented considering state regulation problem for three general affine nonlinear systems with

stronger nonlinearities, and two practical examples of affine nonlinear systems- (i) single-link manipulator system, and (ii) Vander Pol oscillator system.

- The performance of synchronous PI based adaptive critic control scheme for affine nonlinear systems applications is analysed.
- For convergence of critic NN and actor NN, the requirement of PE condition is investigated also in the control applications of affine nonlinear systems.

1.6 ORGANIZATION OF THESIS

This research thesis is organized in seven chapters. The brief overview of each chapter is as following:

Chapter 1 presents the relevance & general introduction of the research work. The recent trends state-of-the-art literature review on the research work is presented in this chapter. The research motivation, problem formulation and scope of work, author's contribution and thesis organization are also presented in this chapter.

Chapter 2 Chapter 2 deals with optimal and adaptive control of dynamical systems. In this chapter the performance analysis of control application problems of the optimal control of nonlinear inverted pendulum-cart dynamical system using PID controller and LQR, the optimal control of automatic generation control of two-area power system, and adaptive control of dynamical systems are presented.

Chapter 3 deals with intelligent control of dynamical systems. In this chapter the performance analysis of control application problems of the intelligent control of nonlinear inverted pendulum-cart dynamical system using fuzzy logic systems, the fuzzy-PI control for automatic generation control of two-area nonlinear power system, and neural network control of a process system using radial basis function neural networks are presented.

Chapter 4 describes the adaptive optimal control using policy iteration technique for continuous-time LTI systems. The policy iteration technique with proof of convergence of algorithm is discussed. The performance analysis of adaptive critic control scheme using policy iteration technique for LTI systems is presented. The comparative performance analysis of adaptive critic scheme and linear quadratic regulator is also presented. The application of control scheme is implemented for certain examples of general and practical continuous-time LTI systems.

Chapter 5 describes the adaptive optimal control using policy iteration technique for continuous-time affine nonlinear systems. The neural network approximation of cost function for implementation of online policy iteration technique is presented. The application of control scheme is implemented for certain examples of general and practical continuous-time affine nonlinear systems.

Chapter 6 discusses the intelligent adaptive optimal control using synchronous policy iteration technique for continuous-time LTI systems and affine nonlinear systems. The neural

network approximations of cost function and control policy is presented. The application of control scheme is implemented for certain examples of general and practical continuous-time LTI systems and affine nonlinear systems.

Chapter 7 presents the conclusions and future scope of research work.

At the end a list of publications from this research work and a list of references have been given.

Appendices A and B respectively present the continuous-time smooth function and Lipschitz continuity for the completeness of comprehension on the subject matters.

CHAPTER 2

OPTIMAL AND ADAPTIVE CONTROL OF DYNAMICAL SYSTEMS

This chapter describes the optimal and adaptive control schemes of dynamical systems. The theoretical background on the control schemes of PID control, optimal control, adaptive control and their applications are briefly described for completeness of the topic under discussion. The performance analysis of control schemes is presented for linear and nonlinear dynamical systems applications.

2.1 INTRODUCTION

The dynamic systems are described by ordinary differential or difference equations in contrast to static systems, which are described by algebraic equations. The dynamic systems change with time or any other independent variable according to the dynamic relations. The behaviour of dynamic system is completely determined by the system states for any future time with the knowledge of initial conditions and knowledge of input for any future time. Most of the dynamical systems, physical, chemical, biological, and economical, can be modelled by mathematical relations, such as deterministic and/or stochastic differential and/or difference equations. The dynamic systems are classified based on the nature of their dynamic equations.

The dynamical systems may be steered from one state to another state by the application of some external inputs or controls. For the control system design problems, the classical (or conventional) control theory and modern control theory have given various control solutions for dynamical systems [1-10]. As per the desired system performances in the control system design, the theories of robust control, optimal control, nonlinear control, adaptive control, intelligent control etc. have been developed in the control literature. The basic concepts of PID control, optimal control, and adaptive control schemes, and their implementation for dynamical systems control applications are presented in the respective following sections of this chapter.

2.2 PID CONTROL

The proportional-integral-derivative (PID) control is the combination of three basic control actions, proportional (P) control action, integral (I) control action and derivative (D) control action. This combined control action has the advantages of each of the three individual control actions. The combinations two terms, proportional control action with integral control action, and proportional control action with derivative control action give proportional-integral (PI) control action and proportional-derivative control action respectively. The equation of PID control action is given by

$$u(t) = K_p e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt} \tag{2.1}$$

or

$$\begin{aligned} u(t) &= K_p e(t) + \frac{K_p}{T_i} \int_0^t e(t)dt + K_p T_d \frac{de(t)}{dt} \\ &= K_p \left(e(t) + \frac{1}{T_i} \int_0^t e(t)dt + T_d \frac{de(t)}{dt} \right) \end{aligned} \tag{2.2}$$

where K_p is the proportional gain, K_i is the integral gain, K_d is the derivative gain, T_i is the integral time constant, T_d is the derivative time constant, $e(t)$ is the actuating error signal, the input signal to the controller, and $u(t)$ is the control signal, the controller output signal.

The transfer function of PID control action is given by

$$G_{pid}(s) = \frac{U(s)}{E(s)} = K_p + \frac{K_i}{s} + K_d s = K_p \left(1 + \frac{1}{T_i s} + T_d s \right) \tag{2.3}$$

The PID control has the functionally of all the three terms P, I, and D to their credits in both the steady-state and transient-state responses of the controlled systems. It takes the past (I), present (P), and future (D) information of the control error into account so that in many cases it is able to provide satisfactory control performance. The “three-term” functionalities are highlighted as [24]: The proportional control action provides an overall control action proportional to the error signal through the all-pass gain factor, the integral control action reduces the steady-state errors through low-frequency compensation by an integrator, and the derivative term improves the transient response through high-frequency compensation by a differentiator. The process of selecting the controller parameters to meet given performance specification is known as controller tuning. The conventional tuning method for PID controllers (setting values of K_p, T_i, and T_d) is the Ziegler-Nichols tuning method[1-5, 7-10] for LTI systems based on experimental step responses or based on the value of K_p that results in marginal stability when only proportional control is used. This is an empirical tuning method. The tuning methods for PID controllers are based on their nature and usage in the particular applications with given objectives. These methods may be grouped as: trial and error, analytical, heuristic, frequency response methods, optimization, and adaptive tuning methods. The PID control is the most commonly used dynamic systems control technique. The PID controller is very simple and can easily be implemented using pneumatic, hydraulic, mechanical, electronic devices, and of course software. The PID controllers are very robust to plant uncertainties. There are various issues of PID controllers,

such as noise filtering and high frequency roll-off, set-point weighting and two-degree-of-freedom, windup, tuning and computer implementation [23].

2.3 OPTIMAL CONTROL

In our daily life we are used to optimization of time and resources for optimum utilization. The optimization theory gave the development of optimal control theory. The optimal control is concerned with the control design in the best possible way. The control input given to the system for this best situation is called optimal control, and the measure of best way or performance is called performance index or cost function. Thus, when a system is controlled in an optimum way satisfying a given performance index, we have an optimal control system.

The optimization techniques can be of different types based on various factors as the approach (algebraic or geometric), the interest (single or multiple), the nature of signals (deterministic or stochastic), and the stage (single or multiple). The optimization can be further classified as static optimization and dynamic optimization. The static optimization is concerned with system control under steady-state conditions, i.e. the system variables are not changing with respect to time. The system is described by algebraic equations. In this case, the optimization techniques used are: ordinary calculus, Lagrange multipliers, linear and nonlinear programming. The dynamic optimization is concerned with optimal control of system under dynamic conditions, i.e. the system variables are changing with respect to time and thus time is involved in system description. The system is described by differential or difference equations. In this case, the optimization techniques used are: variational calculus, Pontryagin principle, dynamic programming, and search techniques [10, 42].

The main objective of optimal control is to determine control signals that will cause a process (plant) to satisfy some physical constraints and at the same time extremize (maximize or minimize) a chosen performance criterion (performance index or cost function). The optimal control problem is to find a control which causes the dynamical system to reach a target or follow a state variable (or trajectory) and at the same time extremize a performance index which may take several forms. The various forms of performance indices are based on minimum time, minimum fuel, minimum energy, and minimum target error (terminal target) optimal control problems [10, 42, 43].

The formulation of optimal control problem requires

1. a mathematical description (or model) of the system to be controlled (generally in state variable form),
2. a specification of the performance index, and
3. a statement of boundary conditions and the physical constraints on the states and/or controls.

The optimal control problem is formulated in general as

Consider a linear time-invariant dynamic system described by the state equation

$$\dot{x}(t) = Ax(t) + Bu(t) \quad , \quad x(0) = x_0 \tag{2.4}$$

where $x(t)$ is state variable, and $u(t)$ is control variable.

The control objective is to find the optimal control $u^*(t)$ that causes the system (2.4) to give the trajectory $x^*(t)$ that optimizes or extremizes (minimizes or maximizes) a performance index

$$J = x^T(t_f)Fx(t_f) + \int_{t_0}^{t_f} [x^T(t)Qx(t) + u^T(t)Ru(t)]dt \tag{2.5}$$

where t_0 is initial time, t_f is final time, R is a positive definite matrix, and Q and F are positive definite and positive semidefinite matrices respectively. This form of performance index is called quadratic form.

In case of a nonlinear system described by

$$\dot{x}(t) = f(x(t), u(t), t) \quad , \quad x(0) = x_0 \tag{2.6}$$

the state trajectory $x^*(t)$ optimizes the general performance index

$$J = S(x(t_f), t_f) + \int_{t_0}^{t_f} V(x(t), u(t), t)dt \tag{2.7}$$

with some constraints on the control variables $u(t)$ and/or the state variables $x(t)$. The final time t_f may be fixed, or free, and the final (target) state may be fully or partially fixed or free.

There are two ways to obtain the solutions of optimal control problems either by using Pontryagin's minimum principle, which provides a necessary condition for optimality, or by solving the Hamilton-Jacobi-Bellman (HJB) equation, which is a sufficient condition [42]. Although mathematically elegant, both approaches have a major disadvantage of the need of complete knowledge of the system dynamics for getting the solution through an offline process. In the case when only an approximate model of the system is available, the optimal controller derived with respect to the system's model will not perform optimally when applied for the control of the real process. Thus, adaptation of the controller parameters is highly desired such that system operation becomes optimal with respect to the behaviour of the real plant.

For the applications of optimal control for dynamical systems, the optimal control of nonlinear inverted pendulum dynamical system using PID controller and LQR is presented in section 2.5, and the optimal control design for automatic generation control of two-area interconnected power system using integral controller and LQR is presented in section 2.6.

2.4 ADAPTIVE CONTROL

An adaptive control system uses a control scheme that is capable of modifying its behavior in response to changes in the process dynamics and the disturbance character.

Adaptive controllers have been extensively used in several industries including chemical, aerospace, automotive, and pulp and paper [50, 52, 54]. Most current techniques for designing control systems are based on a good understanding of the plant under consideration and its environment. However, in a number of instances, the plant to be controlled is too complex and the basic physical processes in it are not fully understood. Hence, control design methods need to be augmented with an identification technique aimed at obtaining a progressively better understanding of the plant to be controlled. Adaptive control is a technique of applying some system identification techniques to obtain a model of the process and its environment from input/output experiment and using this model to design a controller. An adaptive system is one in which in addition to the basic (feedback) structure, explicit measures are taken to automatically compensate for variations in the operating conditions, for variations in the process dynamics or for variations in the disturbances, in order to maintain an optimal performance of the system. The ordinary feedback control also attempts to reduce the effects of disturbances and plant uncertainty, but a constant-gain feedback system is not an adaptive system.

Adaptive control is a method of designing a controller with some adjustable parameters and an embedded mechanism for adjusting these parameters. Adaptive controllers have been used mainly to improve the controller's performance online. The controller becomes nonlinear because of the parameter adjustment mechanism. It has, however, a very special structure. Since general nonlinear systems are difficult to deal with it makes sense to consider special classes of nonlinear systems. An adaptive control system can be thought of as having two loops- one loop is a normal feedback with the process and the controller, and the other loop is the parameter adjustment loop. The parameter adjustment loop is often slower than the normal feedback loop. A block diagram of an adaptive system is shown in Fig. 2.1 [52].

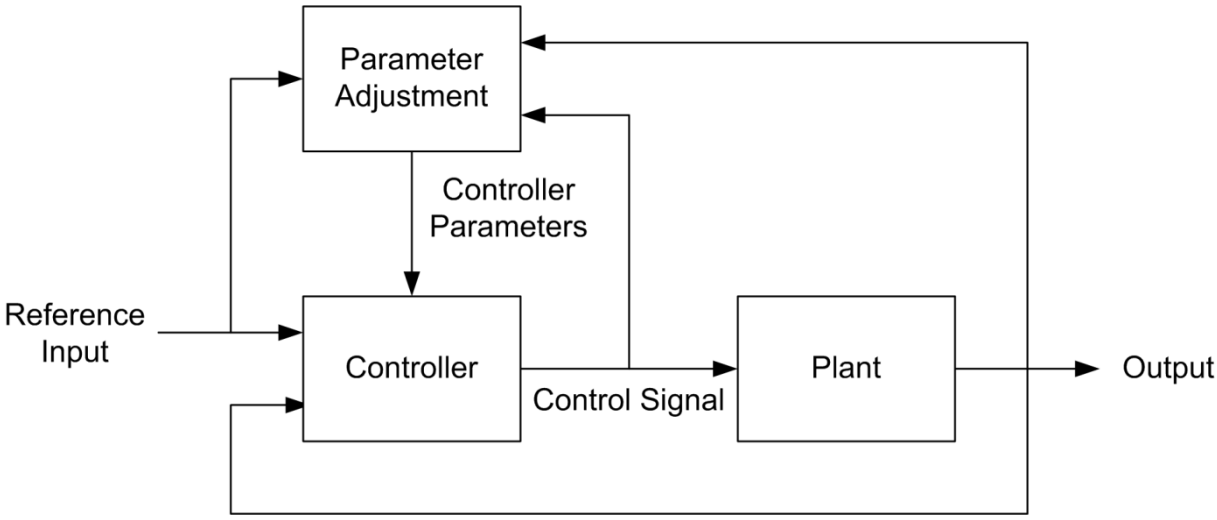


Fig. 2.1 Block diagram of an adaptive control system.

For each control cycle, the adaptive algorithm is normally implemented in three basics steps: (1) Observable data are collected to calculate the controller's performance, (2) The controller's performance is used as a guidance to calculate the adjustment to a set of controller parameters, (3) The controller's parameters are then adjusted to improve the performance of the controller in the next cycle. Normally, an adaptive controller is designed based on one of the available techniques. Each technique is originally designed for a specific class of dynamic system. The controller is then adjusted as data are collected during run time to extend its effectiveness to control a large class of dynamic systems [50, 52].

The rapid growth in the design of integrated and powerful information processors has made the use of adaptive controllers even more versatile. Traditionally, there are four basic approaches for adaptive control [50, 52, 53]: (i) gain scheduling, (ii) model reference adaptive control (MRAC), (iii) self-tuning regulator (STR), and (iv) dual control.

Gain scheduling is a method of adjusting the control signal based on a known look-up table describing changes of a dynamic system. This type of adaptive control system is based on the adjustment of controller parameters in response to the operating conditions of a process. This control scheme is particularly useful when the variations of the process dynamics are predictable. In fact, for a class of dynamic systems, it is possible that an explicit model of the system can be accurately described every time the operating conditions of the system take new values. A block diagram of gain scheduling adaptive scheme is shown in Fig. 2.2 [52]. The two main drawbacks of this method are related to its open loop behaviour and to the discrete assignment of controller gains according to a look-up table. Indeed, for intermediate operating conditions, no explicit control gains are assigned to the system, and the control designer must apply interpolation techniques to avoid instabilities.

The model reference adaptive control (MRAC) is an adaptive control scheme capable of handling processes with unpredictable changes [52]. This control scheme is based on the design of an adaptive scheme whose objective is to drive the error signal between the response of the process and that of the reference model to zero. Thus, the model reference adaptive system (MRAS) is a method of comparing the performance of the actual system against an assumed mathematical model that describes the actual system, and designing control input to drive this comparison error to zero. The block diagram MRAC is shown in Fig. 2.3 [52]. The inner loop, which is faster one, is used for regulation of the process, while the outer loop is designed for adjustment of the parameters of the inner loop controller to drive the error signal to zero. It is observed that there are some instability problems in applying MRAC [52]. The key problem with MRAS is to determine the adjustment mechanism so that a stable system, which brings the error to zero, is obtained. This problem is nontrivial. The following parameter adjustment mechanism, called the MIT rule, was used in the original MRAS:

$$\frac{d\theta}{dt} = -\gamma e \frac{\partial e}{\partial \theta} \tag{2.8}$$

where y is the closed-loop system output, y_m is the output of a reference model which specifies the desired closed-loop response, $e = y - y_m$ denotes the model error and θ is a controller parameter. The quantity $\frac{\partial e}{\partial \theta}$ is the sensitivity derivative of the error with respect to parameter θ. The parameter γ determines the adaptation rate. In practice it is necessary to make approximations to obtain the sensitivity derivative. The MIT rule can be regarded as a gradient scheme to minimize the squared error loss function $J(\theta) = \frac{1}{2}e^2$.

Self-tuning regulator (STR) [52] is another adaptive control scheme characterized by its ability to handle dynamic processes that may be subjected to unpredictable changes in system parameters. STR is a method of updating the parameters of a model that describes the plant based on observed data, and channeling the updated information into the controller that is designed based on these parameters. STR uses the outputs of a recursive identification scheme for plant parameters (outer loop) to adjust, through a suitable adaptation algorithm, the parameters of a controller located in the regulation loop of the system (inner loop). The block diagram of STR is shown in Fig. 2.4 [52]. There are some similarities in terms of inner and outer loop structuring between STR and MRAC. The main difference between the two schemes is, however, that while the STR design is based on an explicit separation between identification and control, the MRAC design uses a direct update of controller parameters to achieve asymptotic decay of the error signal to zero. In view of this fundamental difference in design, MRAC is referred to as a direct adaptive control scheme, while STR is known as an indirect adaptive control scheme.

Dual control [52] is a method of extending adaptive control to stochastic model dealing with uncertainties. The controller can be regarded as being composed of two parts: a nonlinear estimator and a feedback controller. The estimator generates the conditional probability distribution of the state, called the hyperstate of the problem. The feedback controller is a nonlinear function that maps the hyperstate into the space of control variables. This function could be computed offline. The hyperstate must, however, be updated online. The structural simplicity of the solution is obtained at the price of introducing the hyperstate, which is a quantity of very high dimension. Updating of the hyperstate generally requires solution of a complicated nonlinear filtering problem. The block diagram of a dual controller is shown in Fig. 2.5 [52].

The adaptive control problem is to find a method of adjusting the controller when the characteristics of the process and its environment are unknown or changing. The way adaptation law is combined with control law gives the adaptive control methodologies which

include direct adaptive control (DAC) and indirect adaptive control (IAC) algorithms [52]. Direct adaptive systems adjust the controller parameters without explicit identification. Indirect adaptive systems use the results of identification of the process parameters in an optimization procedure to compute the controller settings. In DAC the controller parameters are changed directly without the characteristics of the process and its disturbances first being determined. In IAC methods the process model and possibly the disturbance characteristics are first determined. The controller parameters are designed on the basis of this information.

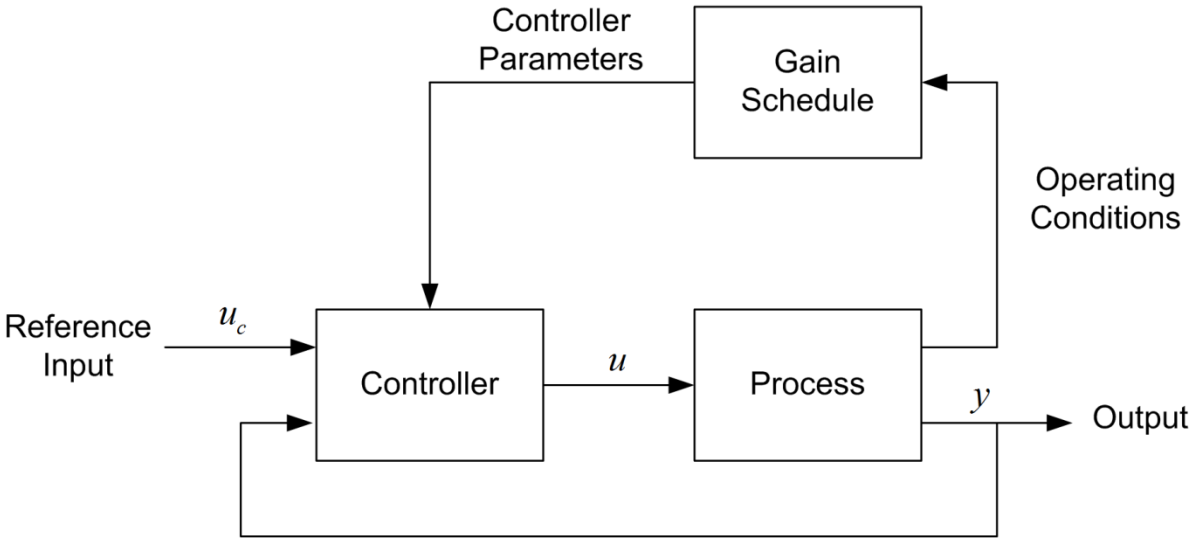


Fig. 2.2 Block diagram of a system with gain scheduling.

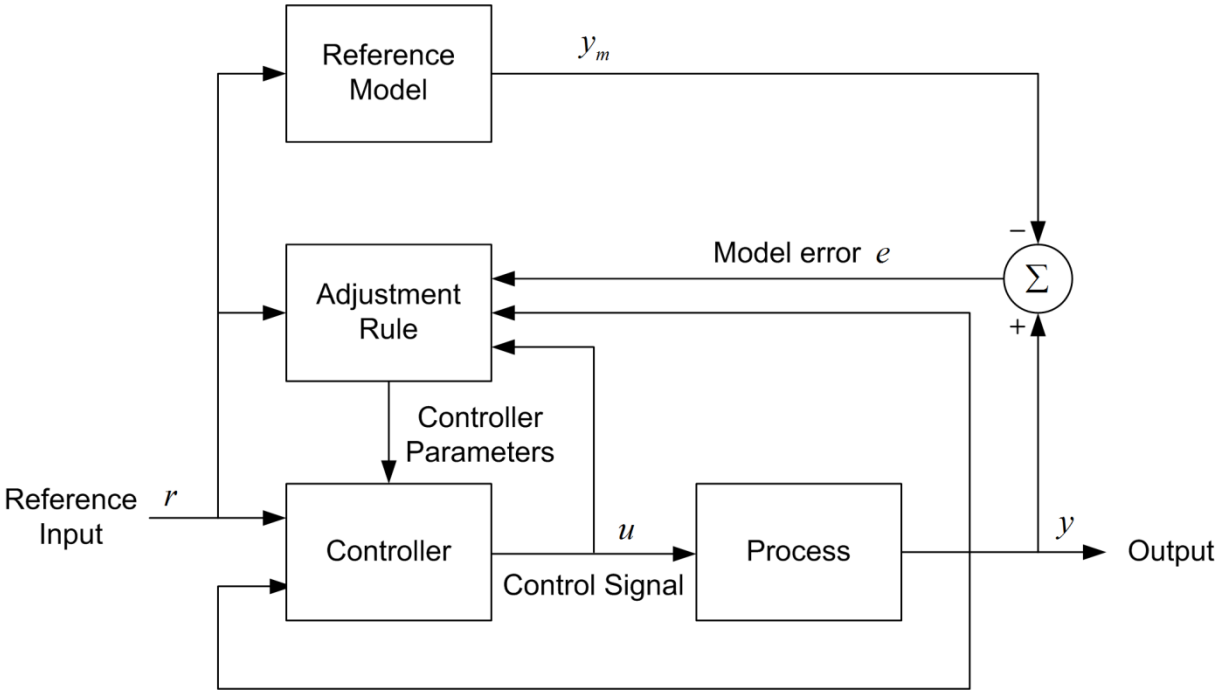


Fig. 2.3 Block diagram of a model reference adaptive control system (MRAC).

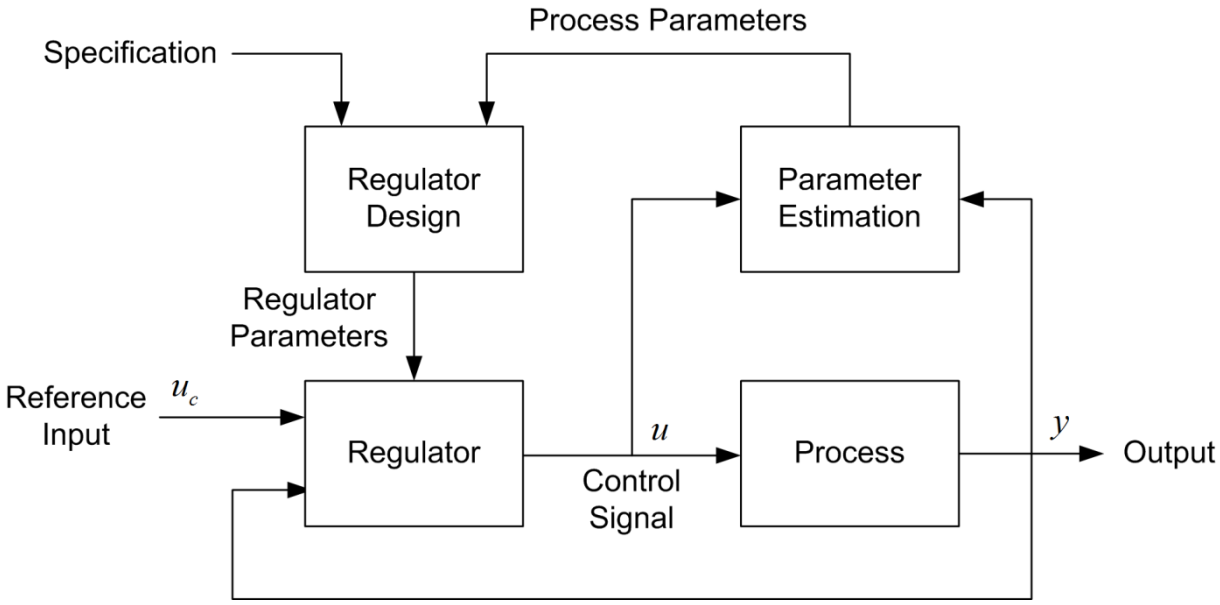


Fig. 2.4 Block diagram of self-tuning regulator (STR).

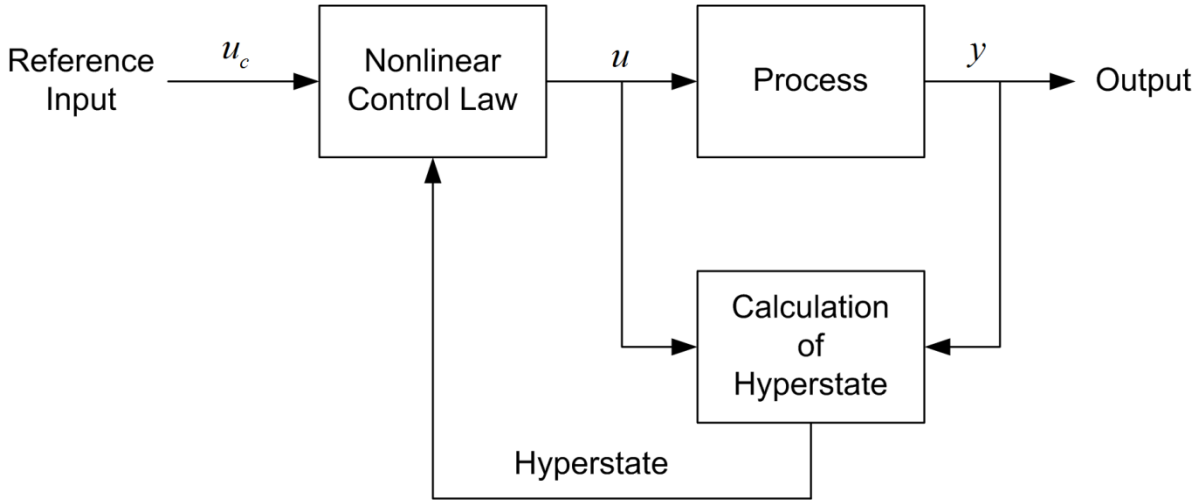


Fig. 2.5 Block diagram of a dual controller.

A typical application of adaptive control was to calibrate a system at startup. In this case, a controller is also designed for a specific class of dynamic systems. However, the parameters that characterize the dynamic behavior of a particular system might not be known in advance. A controller is then designed and arbitrary values are assigned to initialize these parameters. After a few control cycles, parameters are adjusted to converge to the actual parameters of the system. This approach is often used for cases in which a system is designed to handle a variable payload. The payload is different each time, e.g., a crane is used to pick up a sizeable object. The payload will alter the basic dynamic behavior of a dynamic system. Adaptive control is normally used to calibrate these parameters that characterize such dynamic behavior [50, 52, 53].

2.4.1 Indirect Adaptive Control of Dynamical Systems

The indirect adaptive control (IAC) schemes use the explicit plant model identifier to tune controller parameters. The plant model identifier represents the actual plant property; the controller is designed such that the controller stabilizes the plant model identifier although the state feedback is taken from the actual plant. In case of neural networks control approach the plant model identifier is parameterized by neural networks. The conceptual framework for the indirect adaptive control for continuous time system is shown in Fig. 2.6 [71]. The scheme for discrete-time system is also similar except the continuous-time variables are replaced by discrete-time variables.

The indirect adaptive control is implemented in the following two stages:

1. System identification: The plant model is explicitly identified.
2. Control: The controller is designed so that the explicit model identified in step 1 is stable in the closed loop.

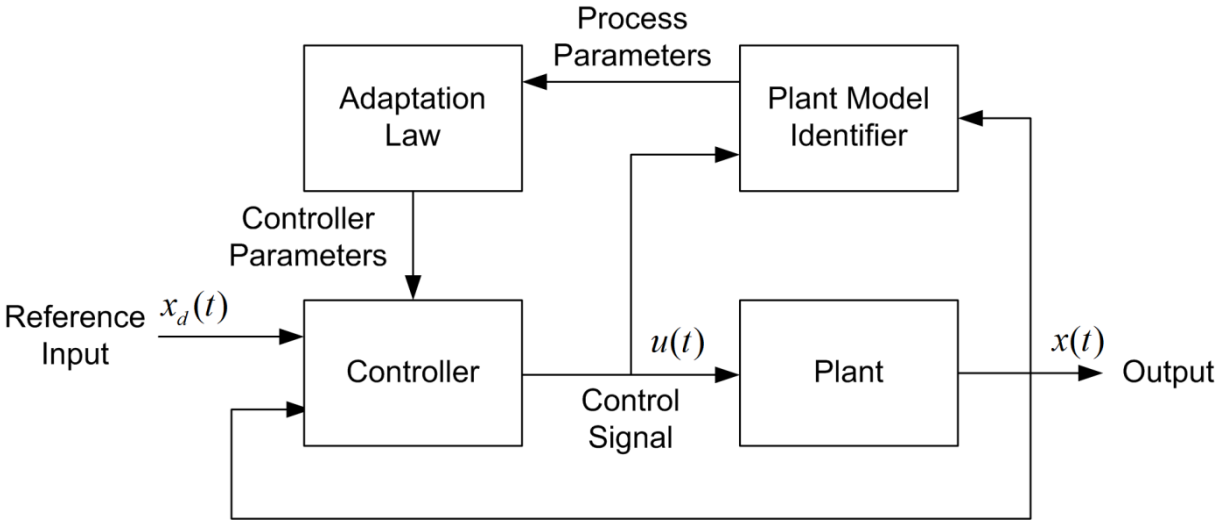


Fig. 2.6 Schematic block diagram of indirect adaptive control.

To elaborate the indirect adaptive control design process, consider an example of indirect adaptive control of a first-order linear system[71]. Let the first-order discrete-time plant is described by

$$x(k+1) = ax(k) + bu(k) \tag{2.9}$$

It is assumed that the plant parameters a and b are unknown. Design an indirect adaptive control scheme for this system. The parameter vector is represented as $\hat{\theta} = [\hat{a} \;\; \hat{b}]^T$ that consists of unknown parameters. The regression vector is represented as $\phi(k) = [x(k) \;\; u(k)]^T$ which consists of the previous plant state and previous control input. Using the recursive least square (RLS) algorithm, the plant parameters can be estimated as

$$\hat{\theta}(k+1) = \hat{\theta}(k) + P(k)\phi(k)[1+\phi^T(k)P(k)\phi(k)]^{-1}[x(k+1) - \phi^T(k)\hat{\theta}(k)] \tag{2.10}$$

$$P(k+1) = P(k) - P(k)\phi(k)[1+\phi^T(k)P(k)\phi(k)]^{-1}\phi^T(k)P(k) \tag{2.11}$$

where $P(k)$ is the covariance matrix. Let us assume that the plant follows a desired trajectory $x_d(k+1)$, the one-step ahead controller is designed as

$$u(k) = \frac{1}{\hat{b}}\left(x_d(k+1) - \hat{a}x(k)\right) \tag{2.12}$$

At every sampling instant k, the parameter vector is updated using (2.10), and is substituted in the control law (2.12). The parameter vector is updated using regression vector $\hat{\theta} = [\hat{a} \;\; \hat{b}]^T$ and the control input $u(k)$ is computed using (2.12). The simulation results for actual parameter vectors $a = 0.6$ and $b = 1$ are shown in Fig. 2.7, Fig. 2.8, and Fig. 2.9 where Fig. 2.7 shows learning of parameters through time instant *k*, Fig. 2.8 shows $x(k)$ tracking desired trajectory $x_d(k)$, and Fig. 2.9 shows the control input $u(k)$.

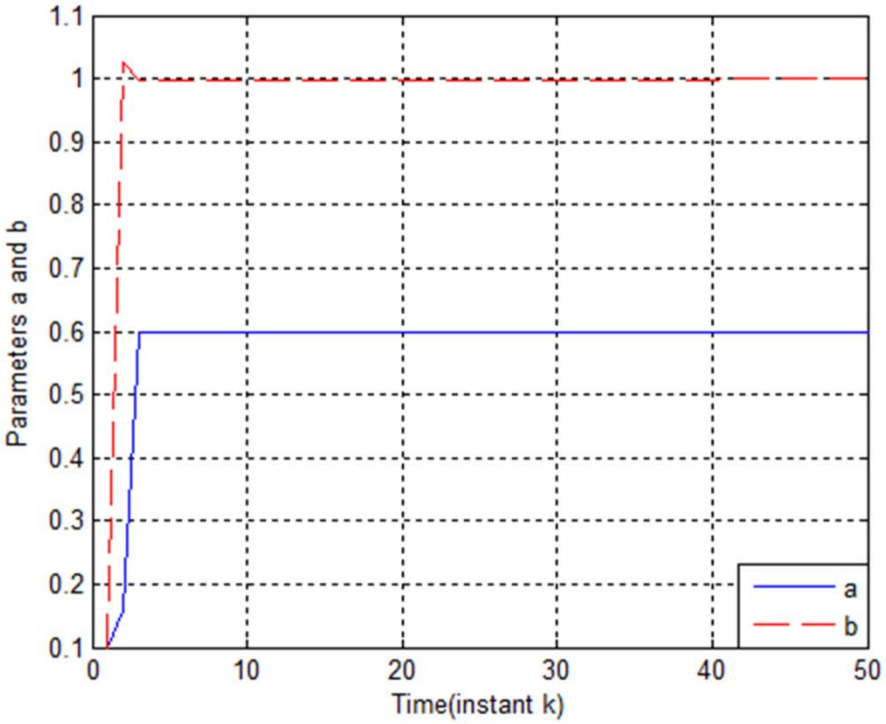


Fig. 2.7 Learning of parameters.

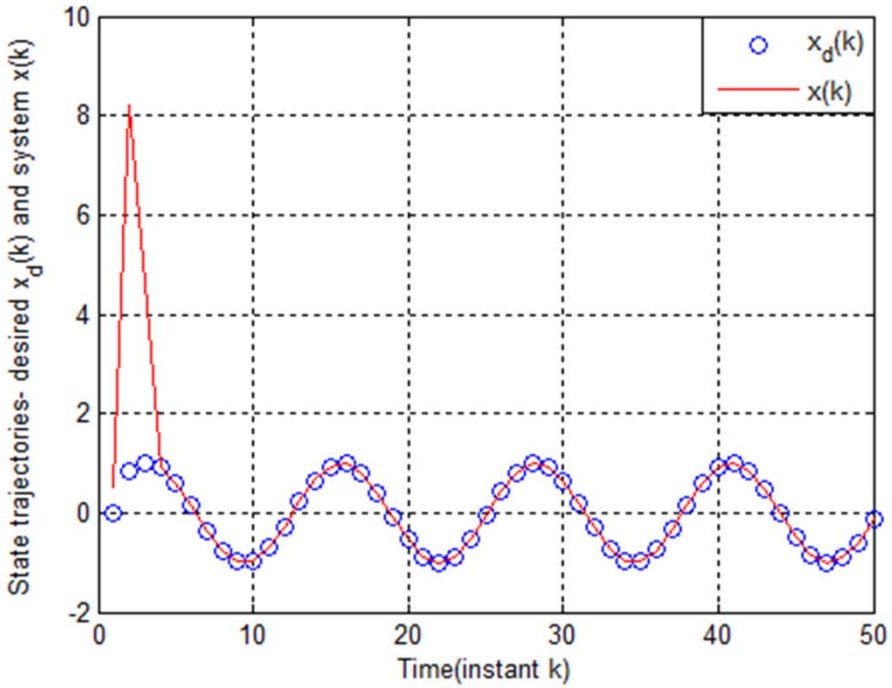


Fig. 2.8 Trajectory tracking.

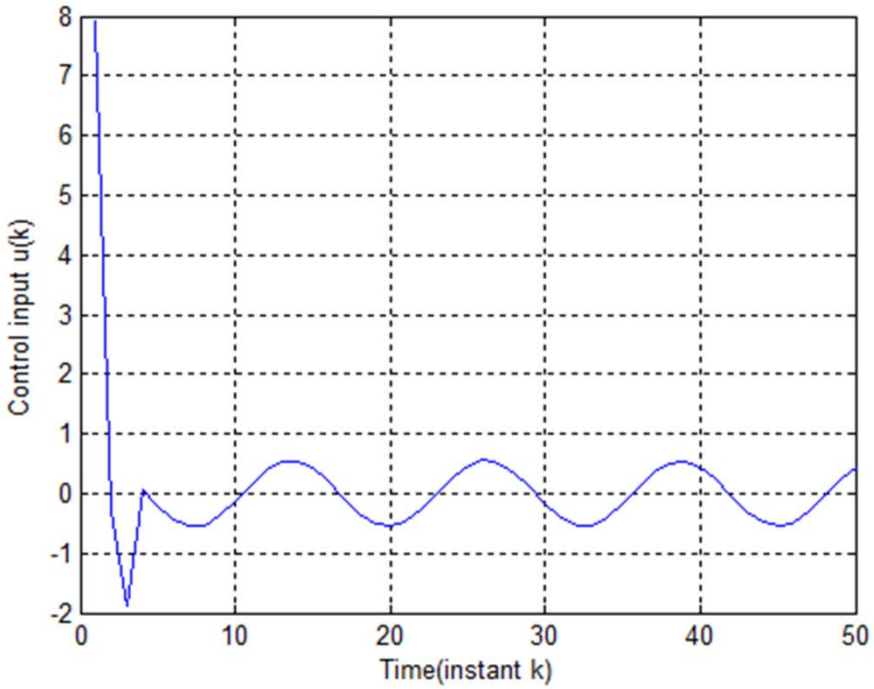


Fig. 2.9 Control input.

2.4.2 Direct Adaptive Control of Dynamical Systems

The direct adaptive control scheme (DAC) directly determines the controller parameters without system identification. In DAC, the plant model is parameterized in terms of the controller parameters that are estimated directly without intermediate calculations involving plant parameter estimates. This control approach is also referred as an implicit adaptive control because the design is based on the estimation of an implicit plant model,

and explicit plant identification is not needed. A schematic block diagram of direct adaptive control is shown in Fig. 2.10 [71].

To elaborate the direct adaptive control design process, consider an example of direct adaptive control of a first-order linear system [71]:

$$\dot{y} = -a_p y + b_p u \tag{2.13}$$

where y is the plant output, u is the plant input, a_p and b_p are the plant parameters. Let us assume that the plant parameters are unknown. The desired performance of the system is specified by the following model:

$$\dot{y}_m = -a_m y_m + b_m r \tag{2.14}$$

where r is the reference signal, and a_m and b_m are known constants.

The objective of the adaptive control design is to formulate a control law and an adaptation law such that the error between the system output and the model output converges to zero. It is assumed that the sign of b_p is known.

Let us choose the following control law:

$$u = \hat{a}_r r + \hat{a}_y y \tag{2.15}$$

where $\hat{a}_r$ and $\hat{a}_y$ are variable controller parameters for which we need to derive an update law. Combining (2.13) and (2.15), the closed loop dynamics becomes

$$\dot{y} = -(a_p - \hat{a}_y b_p) y + \hat{a}_r b_p r \tag{2.16}$$

The choice of control law (2.15) allows the possibility of perfect model matching if the plant parameters are known. In that case if we choose the following values of controller parameters

$$a_r^* = \frac{b_m}{b_p} \quad , \quad a_y^* = \frac{a_p - a_m}{b_p} \tag{2.17}$$

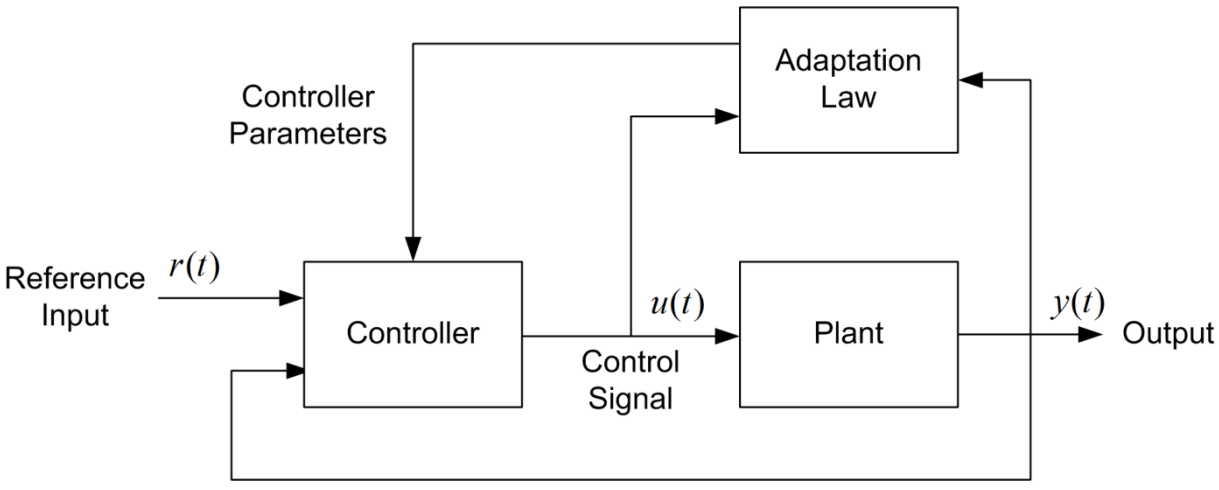


Fig. 2.10 Schematic block diagram of direct adaptive control.

the closed loop dynamics would become identical to reference model (2.14). But since the parameters a_p and b_p are unknown, the control input will achieve this objective adaptively based on the tracking error $e = y - y_m$. Let us define the parameter errors as $\tilde{a}_r = \hat{a}_r - a_r^*$ and $\tilde{a}_y = \hat{a}_y - a_y^*$. Combining (2.16) and (2.14), the error dynamic can be written as

$$\dot{e} = -a_m e + b_p(\tilde{a}_r r + \tilde{a}_y y) \tag{2.18}$$

To analyse the error convergence, consider the following Lyapunov function candidate:

$$V = \frac{1}{2}e^2 + \frac{1}{2\alpha}|b_p|\tilde{a}_r^2 + \frac{1}{2\beta}|b_p|\tilde{a}_y^2 \tag{2.19}$$

Taking the derivative of (2.19), we can write

$$\begin{aligned} \dot{V} &= e\dot{e} + \frac{1}{\alpha}|b_p|\tilde{a}_r\dot{\tilde{a}}_r + \frac{1}{\beta}|b_p|\tilde{a}_y\dot{\tilde{a}}_y \\ &= e\{-a_m e + b_p(\tilde{a}_r r + \tilde{a}_y y)\} + \frac{1}{\alpha}|b_p|\tilde{a}_r\dot{\hat{a}}_r + \frac{1}{\beta}|b_p|\tilde{a}_y\dot{\hat{a}}_y \end{aligned} \tag{2.20}$$

If we choose the following adaptation laws

$$\dot{\hat{a}}_r = -\alpha\,\mathrm{sgn}(b_p)er \tag{2.21}$$

$$\dot{\hat{a}}_y = -\beta\,\mathrm{sgn}(b_p)ey \tag{2.22}$$

then

$$\dot{V} = -a_m e^2 + b_p e(\tilde{a}_r r + \tilde{a}_y y) - |b_p|\tilde{a}_r\,\mathrm{sgn}(b_p)er - |b_p|\tilde{a}_y\,\mathrm{sgn}(b_p)ey \tag{2.23}$$

Since $b_p = |b_p|\mathrm{sgn}(b_p)$, $\dot{V}$ becomes $-a_m e^2$ which is negative definite. Thus, the signals e, $\tilde{a}_r$ and $\tilde{a}_y$ are bounded. Furthermore, because of the boundedness of $\dot{e}$, according to Barbalat's lemma [71], e will converge to zero. Since it is assumed that sign of b_p is known, the update laws (2.21) and (2.22) will give the desired control law. A general block diagram of the above described direct MRAC scheme is shown in Fig. 2.11 [71].

Consider the first-order linear system is given by

$$\dot{y} = 2y + 5u \tag{2.24}$$

To design of the adaptive controller of the form (2.15) using adaptation laws (2.21) and (2.22), assume that plant parameters are known as $a_p = -2$ and $b_p = 5$. The reference model parameters are chosen as $a_m = 3$ and $b_m = 3$. The initial values of the controller parameters $\hat{a}_r$ and $\hat{a}_y$ are chosen 0. The initial conditions of both the plant and model are also taken as 0. Using the control law (2.15) and update laws (2.21) and (2.22), the system is simulated for $r = 2$. The simulation results are shown in Fig. 2.12, Fig. 2.13, and Fig. 2.14 where Fig. 2.12 shows the tracking performance and Fig. 2.13 shows the evolution of the

adaptive parameters. It is noted that though the final parameters are not exactly the same as that of the desired ones, the tracking is achieved as time progress. Fig. 2.14 shows the control input signal.

Consider (2.13) consists an additional nonlinear term giving a first-order nonlinear system of the following form:

$$\dot{y} = -a_p y - c_p f(y) + b_p u \tag{2.25}$$

where $f(y)$ is a known nonlinear function. Then the control law in this case can be chosen as

$$u = \hat{a}_r r + \hat{a}_y y + \hat{a}_f f(y) \tag{2.26}$$

with the adaptation laws (2.21), (2.22) and an additional adaptation law:

$$\dot{\hat{a}}_f = -\gamma \operatorname{sgn}(b_p) e f(y) \tag{2.27}$$

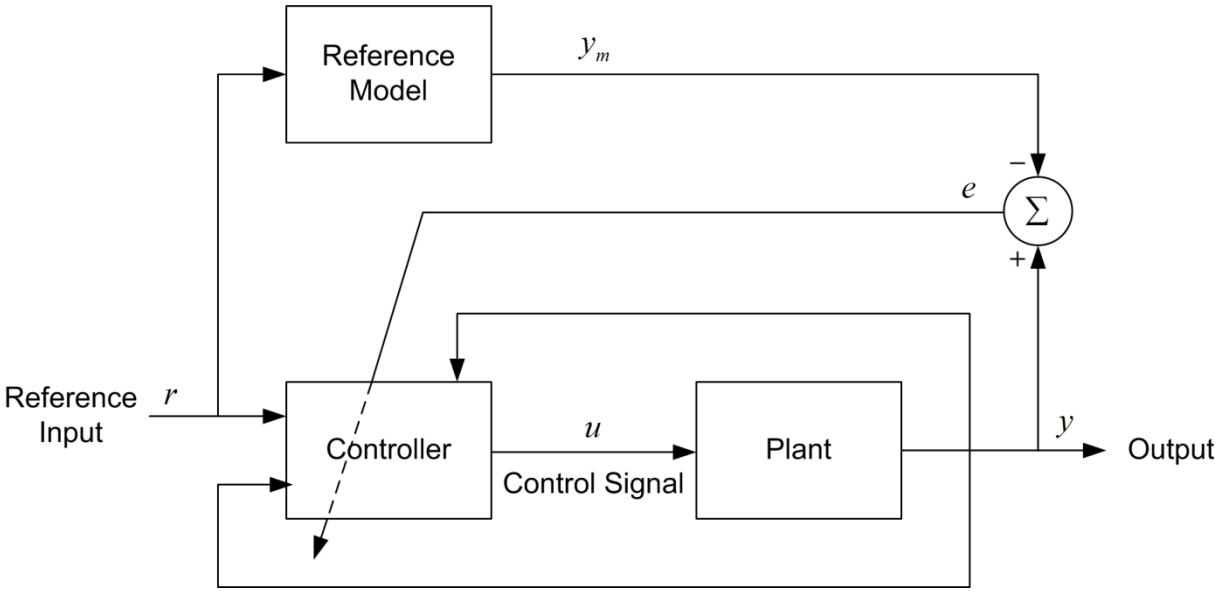


Fig. 2.11 General block diagram of direct MRAC scheme.

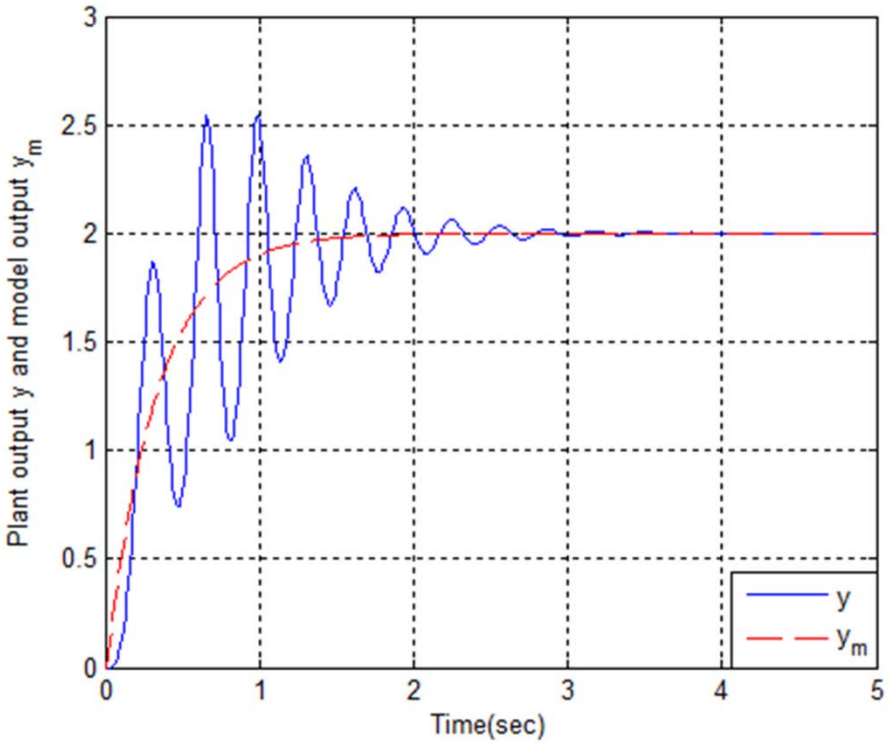


Fig. 2.12 Trajectory tracking.

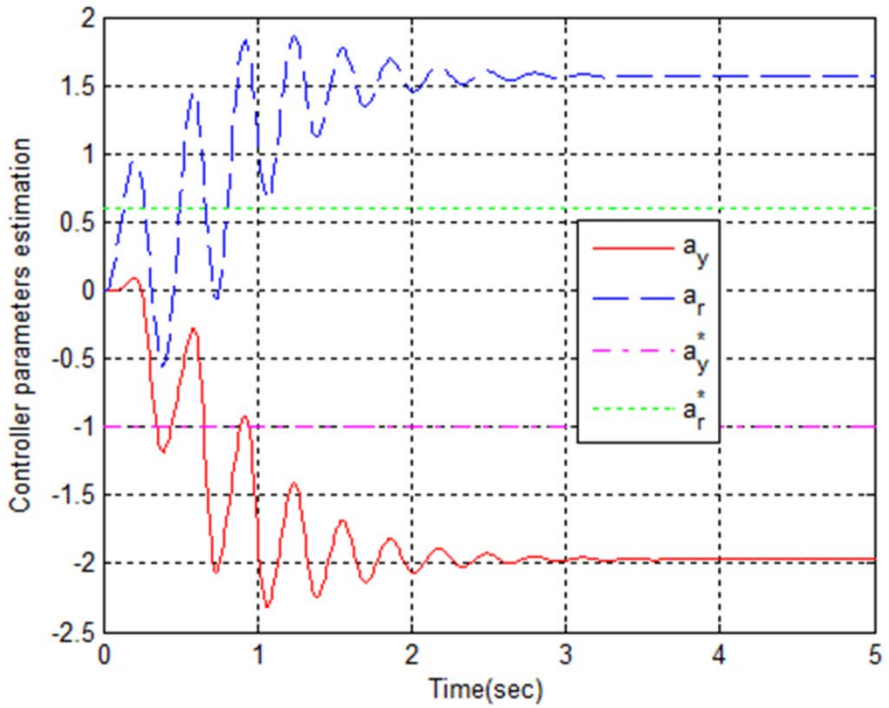


Fig. 2.13 Evolution of adaptive controller parameters.

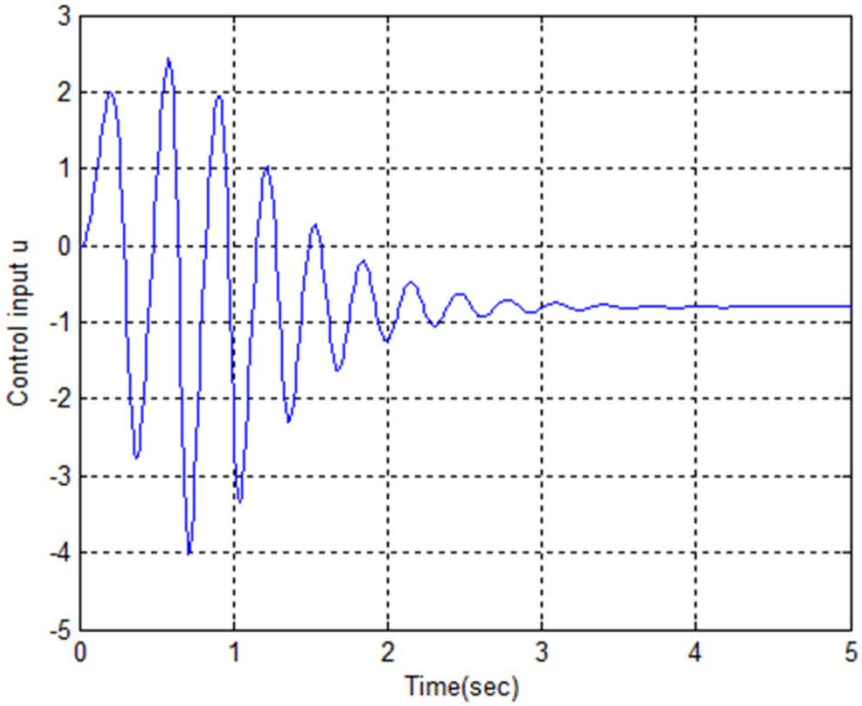


Fig. 2.14 Control input.

2.5 OPTIMAL CONTROL OF NONLINEAR INVERTED PENDULUM DYNAMICAL SYSTEM USING PID CONTROLLER AND LQR

The Inverted Pendulum (IP) is an inherently unstable system with highly nonlinear dynamics. This is a system which belongs to the class of under-actuated mechanical systems having fewer control inputs than degrees of freedom. This renders the control task more challenging making the inverted pendulum system a classical benchmark for the design, testing, evaluating and comparing of different classical & contemporary control techniques. Thus, the control of inverted pendulum has been a research interest in the field of control engineering. Due to its importance this is a choice of dynamic system to analyse its dynamic model and propose a control law. The aim of this case study is to stabilize the inverted pendulum such that the position of the cart on the track is controlled quickly and accurately so that the pendulum is always erected in its inverted position during such movements. Realistically, this simple mechanical system is representative of a class of attitude control problems whose goal is to maintain the desired vertically oriented position at all times [1, 2, 9]. There are many papers present which have taken the inverted pendulum-cart dynamical system for implementing the various control schemes [61, 62, 85, 86, 179-181].

The optimal control design methods are used for the optimal performance of the dynamical systems being controlled. There are various optimal control methods presented in the literature for linear and nonlinear dynamical systems [6, 9, 10, 42, 43]. The linear quadratic regulator (LQR) and linear quadratic Gaussian (LQG) which is a combination of a

linear quadratic estimator (LQE) (i.e. Kalman filter) and a LQR [42, 43, 47] are optimal control design methods for linear dynamical systems. The optimal control problems for nonlinear dynamical systems have been of research interest in control engineering. In recent trends even the various advance control approaches [7, 15, 43, 61, 62, 85, 86, 179-181] are developing and being tried for many dynamical systems control, the simplicity of control algorithms along with the fulfilment of control objectives is further desired. The neural network based control design requires a large data set collected from experiments for networks training & testing; the fuzzy control requires framing of rules which becomes complex for higher order systems; and the evolutionary computational techniques are slow in computation, these drawbacks put limit to their implementation even these techniques provide automation & intelligent features to the controlled systems. Also the algorithms of several adaptive, sliding mode and robust control approaches are comparatively complex even these have certain merits.

The objective & contribution of this section is to present a simple approach to control nonlinear dynamical systems. The simple control algorithms of LQR and PID control which are generally used for control of the linear dynamical systems have been used in this section to control the nonlinear inverted pendulum-cart dynamical system. Here the control objective is to control the system such that the cart reaches at a desired position and the inverted pendulum stabilizes in upright position. The modelling, simulation and performance analysis of optimal control of nonlinear inverted pendulum-cart dynamic system using PID controller and LQR have been presented for both cases of without & with disturbance input. The comprehensive performance investigation shows that the proposed control method is simple, effective, and robust.

2.5.1 Mathematical Modelling of Nonlinear Inverted Pendulum System

2.5.1.1 Inverted pendulum system equations

The free body diagram of an inverted pendulum mounted on a motor driven cart is shown in Fig. 2.15 [1, 2, 61, 85, 86, 179-181]. The system equations of this nonlinear dynamic system can be derived as follows. It is assumed here that the pendulum rod is mass-less, and the hinge is frictionless. In such assumption, the whole pendulum mass is concentrated in the centre of gravity (COG) located at the center of the pendulum ball. The cart mass and the ball point mass at the upper end of the inverted pendulum are denoted as *M* and *m*, respectively. There is an externally *x*-directed force on the cart, $u(t)$, and a gravity force acts on the point mass at all times. The coordinate system considered is shown in Fig. 2.15, where $x(t)$ represents the cart position, and $\theta(t)$ is the tilt angle referenced to the vertically upward direction.

A force balance on the system in the *x*-direction can be written as

$$M\frac{d^2}{dt^2}x + m\frac{d^2}{dt^2}x_G = u \tag{2.28}$$

where, the time-dependent centre of gravity (COG) of the point mass is given by the coordinates, (x_G, y_G). For the point mass assumed here, the location of the center of gravity of the pendulum mass is simply

$$x_G = x + l\sin\theta \qquad \text{and} \qquad y_G = l\cos\theta \tag{2.29}$$

where l is the pendulum rod length. Substituting (2.29) into (2.28) it is written as

$$(M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + ml\cos\theta\ddot{\theta} = u \tag{2.30}$$

In a similar way, a torque balance on the system is performed. Fig. 2.16 shows the force components acting on the system. The resultant torque balance can be written as

$$(F_x\cos\theta)l - (F_y\sin\theta)l = (mg\sin\theta)l \tag{2.31}$$

where, F_x , and F_y are the force components in *x* and *y* directions respectively may be determined as

$$F_x = m\frac{d^2}{dt^2}x_G = m\left[\ddot{x} - l\sin\theta\dot{\theta}^2 + l\cos\theta\ddot{\theta}\right] \tag{2.32}$$

$$F_y = m\frac{d^2}{dt^2}y_G = -m\left[l\cos\theta\dot{\theta}^2 + l\sin\theta\ddot{\theta}\right] \tag{2.33}$$

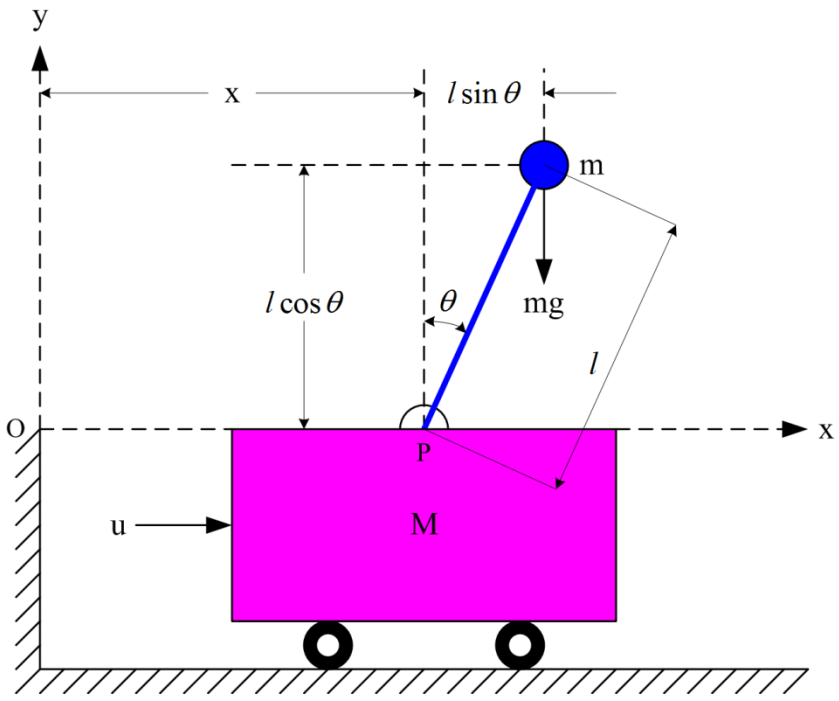


Fig. 2.15 Motor driven inverted pendulum-cart system.

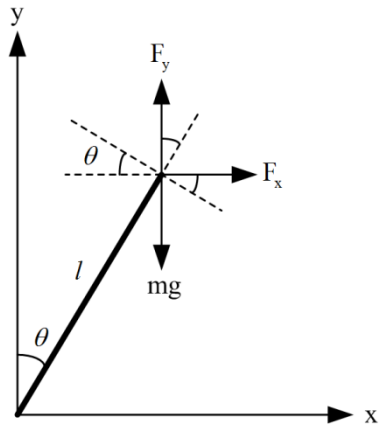


Fig. 2.16 Vector diagram for force components in torque balance.

Substituting (2.32) and (2.33) into (2.31) we have

$$m\ddot{x}\cos\theta + ml\ddot{\theta} = mg\sin\theta \tag{2.34}$$

Equations (2.30) and (2.34) are the defining equations for this system. These two equations are manipulated algebraically to have only a single second derivative term in each equation. Thus from (2.34), we have

$$ml\ddot{\theta} = mg\sin\theta - m\ddot{x}\cos\theta$$

and putting this into (2.30) gives

$$(M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + mg\cos\theta\sin\theta - m\ddot{x}\cos^2\theta = u$$

or

$$(M+m-m\cos^2\theta)\ddot{x} = u + ml\sin\theta\dot{\theta}^2 - mg\cos\theta\sin\theta \tag{2.35}$$

Similarly, from (2.34), we have

$$\ddot{x} = \frac{g\sin\theta - l\ddot{\theta}}{\cos\theta}$$

and putting this into (2.30), we have

$$\frac{(M+m)(g\sin\theta - l\ddot{\theta})}{\cos\theta} - ml\sin\theta\dot{\theta}^2 + ml\cos\theta\ddot{\theta} = u$$

or

$$(M+m)(g\sin\theta - l\ddot{\theta}) - ml\cos\theta\sin\theta\dot{\theta}^2 + ml\cos^2\theta\ddot{\theta} = u\cos\theta$$

and

$$\left(ml\cos^2\theta - (M+m)l\right)\ddot{\theta} = u\cos\theta - (M+m)g\sin\theta + ml\cos\theta\sin\theta\dot{\theta}^2 \tag{2.36}$$

Finally, by dividing the lead coefficients of (2.35) and (2.36) we may derive the system equations describing the cart position dynamics and the pendulum angle dynamics respectively. Thus we have

$$\ddot{x} = \frac{u + ml(\sin\theta)\dot{\theta}^2 - mg\cos\theta\sin\theta}{M+m-m\cos^2\theta} \tag{2.37}$$

$$\ddot{\theta} = \frac{u\cos\theta - (M+m)g\sin\theta + ml(\cos\theta\sin\theta)\dot{\theta}}{ml\cos^2\theta - (M+m)l} \tag{2.38}$$

Equations (2.37) and (2.38) represent a nonlinear system which is relatively complicated from a mathematical viewpoint. Following subsection presents the standard state space form of these two nonlinear equations.

2.5.1.2 Nonlinear system state space equations of inverted pendulum

For numerical simulation of the nonlinear model for the inverted pendulum-cart dynamic system, it is required to represent the nonlinear equations (2.37) and (2.38) into standard state space form,

$$\frac{d}{dt}x = f(x,u,t) \tag{2.39}$$

Considering the state variables as following:

$$x_1 = \theta \quad , \quad x_2 = \dot{\theta} = \dot{x}_1 \quad , \quad x_3 = x \quad , \quad x_4 = \dot{x} = \dot{x}_3 \tag{2.40}$$

Then, the final state space equation for the inverted pendulum system may be written as

$$\frac{d}{dt}x = \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{d}{dt}\begin{bmatrix} \theta \\ \dot{\theta} \\ x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} \tag{2.41}$$

where,

$$f_1 = x_2 \tag{2.42}$$

$$f_2 = \frac{u\cos x_1 - (M+m)g\sin x_1 + ml(\cos x_1 \sin x_1)x_2^2}{ml\cos^2 x_1 - (M+m)l} \tag{2.43}$$

$$f_3 = x_4 \tag{2.44}$$

$$f_4 = \frac{u + ml(\sin x_1)x_2^2 - mg\cos x_1 \sin x_1}{M + m - m\cos^2 x_1} \tag{2.45}$$

If both the pendulum angle θ and the cart position x are the variables of interest, then the output equation may be written as

$$y = Cx \text{ or } y = \begin{bmatrix} \theta \\ x \end{bmatrix} = Cx = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} \theta \\ \dot{\theta} \\ x \\ \dot{x} \end{bmatrix} \tag{2.46}$$

Equations (2.41) and (2.46) give a complete state space representation of the nonlinear inverted pendulum-cart dynamic system.

2.5.1.3 Linear system state space equations of inverted pendulum

Since the goal of this particular system is to keep the inverted pendulum in upright position around $\theta = 0$, the linearization might be considered about this upright equilibrium point. The linear model for the system around the upright stationary point is derived by simply linearization of the nonlinear system given in (2.41). Since the usual *A* and *B* matrices are zero for this case; and so every term is put into the nonlinear vector function, $f(x,u,t)$, then the linearized form for the system becomes

$$\frac{d}{dt}\delta x = J_x(x_0,u_0)\delta x + J_u(x_0,u_0)\delta u \tag{2.47}$$

where, the reference state is defined with the pendulum stationary and upright with no input force. Under these conditions, $x_0 = 0$, and $u_0 = 0$.

Since the nonlinear vector function is rather complicated, the components of the Jacobian matrices are determined systemically, term by term. The elements of the first second, third, and fourth columns of $J_x(x_0,u_0)$ are given by $\left.\frac{\partial f_i}{\partial x_1}\right|_{x_0,u_0}$, $\left.\frac{\partial f_i}{\partial x_2}\right|_{x_0,u_0}$, $\left.\frac{\partial f_i}{\partial x_3}\right|_{x_0,u_0}$, and $\left.\frac{\partial f_i}{\partial x_4}\right|_{x_0,u_0}$ respectively. Thus, combining all these separate terms gives

$$J_x(x_0,u_0) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{(M+m)g}{Ml} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{mg}{M} & 0 & 0 & 0 \end{bmatrix} \tag{2.48}$$

For the derivative of the nonlinear terms with respect to u, we have

$$J_u(x_0,u_0) = \begin{bmatrix} \partial f_1/\partial u \\ \partial f_2/\partial u \\ \partial f_3/\partial u \\ \partial f_4/\partial u \end{bmatrix}_{x_0,u_0} = \begin{bmatrix} 0 \\ \frac{\cos x_1}{ml\cos^2 x_1 - (M+m)l} \\ 0 \\ \frac{1}{M+m-m\cos^2 x_1} \end{bmatrix}_{x_0,u_0} = \begin{bmatrix} 0 \\ \frac{-1}{Ml} \\ 0 \\ \frac{1}{M} \end{bmatrix} \tag{2.49}$$

Finally, after all these manipulations (2.47) may be written explicitly as

$$\frac{d}{dt}\delta x = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{(M+m)g}{Ml} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{mg}{M} & 0 & 0 & 0 \end{bmatrix} \delta x + \begin{bmatrix} 0 \\ \frac{-1}{Ml} \\ 0 \\ \frac{1}{M} \end{bmatrix} \delta u \tag{2.50}$$

This is the open loop linearized model for the inverted pendulum with a cart force, $\delta u(t)$, (written in perturbation form). Thus, LTI system is in standard state space form. The equation (2.50) may be written in general as

$$\frac{d}{dt}\delta x = A\delta x + B\delta u \tag{2.51}$$

Equation (2.51) along with the output equation (2.46) represents the final linear model of the inverted pendulum-cart system. This is the simplified model which is used to study the system behavior in general and to design LQR.

2.5.1.4 Inverted pendulum system equations with disturbance input

The system equations of this nonlinear dynamic system with disturbance input can be derived as follows. Consider a disturbance input due to wind effects acting on the inverted pendulum in addition to force on the cart, $u(t)$. Let F_w represent the horizontal wind force on the pendulum point mass. With this additional force component, the force balance equation (2.28) becomes

$$M\frac{d^2}{dt^2}x + m\frac{d^2}{dt^2}x_G = u + F_w \tag{2.52}$$

which can be manipulated as to give

$$(M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + ml\cos\theta\ddot{\theta} = u + F_w \tag{2.53}$$

Similarly, the torque in the clockwise direction caused by the horizontal wind disturbance is $(F_w\cos\theta)l$. Adding the torque contribution of this term the torque balance equation (2.31) becomes

$$(F_x\cos\theta)l - (F_y\sin\theta)l = (mg\sin\theta)l + (F_w\cos\theta)l \tag{2.54}$$

which can be modified to give

$$m\ddot{x}\cos\theta + ml\ddot{\theta} = mg\sin\theta + F_w\cos\theta \tag{2.55}$$

Equations (2.53) and (2.55) are the defining equations for this system with a disturbance input.

The state space equation for inverted pendulum system with disturbance input is derived as same of equation (2.41) with following modification

$$f_2 = \frac{u\cos x_1 - (M+m)g\sin x_1 + ml(\cos x_1 \sin x_1)x_2^2 - \frac{M}{m}F_w \cos x_1}{ml\cos^2 x_1 - (M+m)l} \tag{2.56}$$

$$f_4 = \frac{u + ml(\sin x_1)x_2^2 - mg\cos x_1 \sin x_1 + F_w \sin^2 x_1}{M + m - m\cos^2 x_1} \tag{2.57}$$

The output equation of nonlinear inverted pendulum system with disturbance input remains same as equation (2.46).

The linearized model can also be developed as following:

$$\frac{d}{dt}\delta x = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{(M+m)g}{Ml} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{mg}{M} & 0 & 0 & 0 \end{bmatrix} \delta x + \begin{bmatrix} 0 \\ \frac{-1}{Ml} \\ 0 \\ \frac{1}{M} \end{bmatrix} \delta u + \begin{bmatrix} 0 \\ \frac{-1}{ml} \\ 0 \\ 0 \end{bmatrix} \delta F_w \tag{2.58}$$

This is the open loop linearized model for the inverted pendulum with a cart force, $\delta u(t)$, and a horizontal wind disturbance, $\delta F_w(t)$. The two inputs have been separated for convenience, thus the LTI system can be written as

$$\frac{d}{dt}\delta x = A\delta x + b_1\delta u + b_2\delta F_w \tag{2.59}$$

2.5.2 Control Methods

The following control methods are presented here to control the nonlinear inverted pendulum-cart dynamic system.

2.5.2.1 PID Control

To stabilize the inverted pendulum in upright position and to control the cart at desired position using PID control approach two PID controllers- angle PID controller, and cart PID controller have been designed for the two control loops of the system. The equations of PID control are given as following:

$$u_p = K_{pp}e_\theta(t) + K_{ip}\int e_\theta(t)dt + K_{dp}\frac{de_\theta(t)}{dt} \tag{2.60}$$

$$u_c = K_{pc}e_x(t) + K_{ic}\int e_x(t)dt + K_{dc}\frac{de_x(t)}{dt} \tag{2.61}$$

where, $e_\theta(t)$ and $e_x(t)$ are angle error and cart position error. Since the pendulum angle dynamics and cart position dynamics are coupled to each other so the change in any controller parameters affects both the pendulum angle and cart position which makes the tuning tedious. The tuning of controller parameters is done using trial & error method and observing the responses of SIMULINK model to be the optimal.

2.5.2.2 Optimal Control Design Using LQR

Optimal control refers to a class of methods that can be used to synthesize a control policy which results in best possible behavior with respect to the prescribed criterion (i.e. control policy which leads to maximization of performance). The main objective of optimal control is to determine control signals that will cause a process (plant) to satisfy some physical constraints and at the same time extremize (maximize or minimize) a chosen performance criterion (performance index (PI) or cost function). The optimal control problem is to find a control which causes the dynamical system to reach a target or follow a state variable (or trajectory) and at the same time extremize a PI which may take several forms [1, 2, 4, 6, 7, 9, 10, 42, 43].

Linear quadratic regulator (LQR) is one of the optimal control techniques, which takes into account the states of the dynamical system and control input to make the optimal control decisions. This is simple as well as robust [1, 2, 4, 6, 7, 9, 10, 42, 43].

After linearization of nonlinear system equations about the upright (unstable) equilibrium position having initial conditions as $x_0 = [0, 0, 0, 0]^T$, the linear state-space equation is obtained as

$$\dot{x} = Ax + Bu \tag{2.62}$$

where, $x = \left[\theta, \dot{\theta}, x, \dot{x}\right]^T$

The state feedback control $u = -Kx$ leads to

$$\dot{x} = (A - BK)x \tag{2.63}$$

where, *K* is derived from minimization of the cost function

$$J = \int \left(x^T Q x + u^T R u\right) dt \tag{2.64}$$

where, *Q* and *R* are positive semi-definite and positive definite symmetric constant matrices respectively.

The LQR gain vector *K* is given by

$$K = R^{-1} B^T P \tag{2.65}$$

where, *P* is a positive definite symmetric constant matrix obtained from the solution of matrix algebraic reccatti equation (ARE)

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \tag{2.66}$$

In the optimal control of nonlinear inverted pendulum dynamical system using PID controller & LQR approach, all the instantaneous states of the nonlinear system, pendulum angle θ, angular velocity $\dot{\theta}$, cart position x, and cart velocity $\dot{x}$ have been considered available for measurement which are directly fed to the LQR. The LQR is designed using the linear state-space model of the system. The optimal control value of LQR is added negatively with PID control value to have a resultant optimal control. The tuning of the PID controllers

which are used here either as PID control method or PID+LQR control methods is done by trial & error method and observing the responses achieved to be optimal.

2.5.3 Simulation Results and Analysis

The MATLAB-SIMULINK models for the simulation of modeling, analysis, and control of nonlinear inverted pendulum-cart dynamical system without & with disturbance input have been developed. The typical parameters of inverted pendulum-cart system setup are selected as [85, 180]: mass of the cart (*M*): 2.4 kg, mass of the pendulum (*m*): 0.23 kg, length of the pendulum (*l*): 0.36 m, length of the cart track (*L*): ± 0.5 m, friction coefficient of the cart & pole rotation is assumed negligible. The disturbance input parameters which has been taken in simulation are [205]: Band Limited White Noise Power = 0.001, Sample Time = 0.01, Seed = 23341.

After linearization the system matrices used to design LQR are computed as below:

$$A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 29.8615 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -0.9401 & 0 & 0 & 0 \end{bmatrix},\quad B=\begin{bmatrix} 0 \\ -1.1574 \\ 0 \\ 0.4167 \end{bmatrix},\quad C=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},\quad \text{and} \quad D=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

With the choice of

$$Q=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 500 & 0 \\ 0 & 0 & 0 & 250 \end{bmatrix},\quad \text{and} \quad R=1$$

The LQR gain vector is obtained as following:

$$K=\begin{bmatrix} -137.7896 & -25.9783 & -22.3607 & -27.5768 \end{bmatrix}$$

Here three control schemes have been implemented for optimal control of nonlinear inverted pendulum-cart dynamical system:

1. PID control method having two PIDs i.e. angle PID & cart PID
2. Two PIDs (i.e. angle PID & cart PID) with LQR control method
3. One PID (i.e. cart PID) with LQR control method.

The tuned PID controller parameters of these control schemes for cases of without & with disturbance input are given as in Table 2.1 & Table 2.2 respectively.

Table 2.1 PID controller parameters of control schemes: without disturbance input case

Control Schemes	Angle PID Control			Cart PID Control		
	K_{pp}	K_{ip}	K_{dp}	K_{pc}	K_{ic}	K_{dc}
PID	-40	0	-8	-1	0	-3
2 PID+LQR	1	1	1	1.5	-7.5	5
1 PID+LQR	---	---	---	1.5	-7.5	5

Table 2.2 PID controller parameters of control schemes: with disturbance input case

Control Schemes	Angle PID Control			Cart PID Control		
	K_{pp}	K_{ip}	K_{dp}	K_{pc}	K_{ic}	K_{dc}
PID	-40	0	-8	-1.25	0	-3.6
2 PID+LQR	1	1	1	1.5	-7.5	5
1 PID+LQR	---	---	---	1.5	-7.5	5

The SIMULINK models for control of nonlinear inverted pendulum system using PID control method for both cases of without & with disturbance input are shown in Fig. 2.17 and Fig. 2.19 respectively. The band limited white noise has been added as the disturbance input to the system. Here only pendulum angle θ and cart position x have been considered for the measurement. The reference angle has been set to 0 (rad), and reference cart position is set to 0.1 (m). The simulation results for both cases are shown in Fig. 2.18 and Fig. 2.20 respectively. It is observed that the pendulum stabilizes in vertically upright position after two small overshoots for case of without disturbance input, and it stabilizes upright with minor oscillations for case of with continuous disturbance input also. The cart position x reaches the desired position of 0.1 (m) quickly & smoothly for case of without disturbance input, and quickly with minor oscillations for case of with continuous disturbance input. The control input u is bounded for both cases in ranges [-0.1 0.1], and [-1 1] respectively. Thus simulation results justify the effectiveness & robustness of the PID control.

The SIMULINK models for optimal control of nonlinear inverted pendulum-cart system using two PID controllers (angle PID & cart PID) with LQR control method for both cases of without & with disturbance input are shown in Fig. 2.21 and Fig. 2.23 respectively. In this approach all the states of the system θ, $\dot{\theta}$, x and $\dot{x}$ are fed to LQR, which is designed using the linear state-space model of the system. Here also the angle θ & cart position x have been taken as variables of interest for control, and the band limited white noise has been added as the disturbance input to the system. The reference angle is set to 0 (rad), and the reference cart position has been set to 0.1 (m). The simulation results for both cases are shown in Fig. 2.22 and Fig. 2.24 respectively. Here responses of angle θ, angular velocity $\dot{\theta}$, cart position x, cart velocity $\dot{x}$, and control u have been plotted. It is observed that the pendulum stabilizes in vertically upright position quickly & smoothly after two minor undershoots and a minor overshoot for case of without disturbance input, and it stabilizes in vertically upright position with minute oscillations for case of with continuous disturbance input also. The angular velocity approaches 0 (rad/s) quickly for case of without disturbance input, and it oscillates by approx. +/-0.01 (rad/s) remaining at most in range approx +/-0.02

(rad/s) for case of with continuous disturbance input. The cart position x reaches smoothly the desired position of 0.1 (m) quickly in approx. 6 seconds, and the cart velocity reaches to zero for both cases. The control input u is bounded for both cases in ranges [-0.1 0.1], and [-1 1] respectively. The simulation results justify the effectiveness & robustness of the 2PID+LQR control.

The SIMULINK models for optimal control of nonlinear inverted pendulum-cart system using one PID controller (cart PID) with LQR control method for both cases of without & with disturbance input are shown in Fig. 2.25 and Fig. 2.27 respectively. This control method is similar to 2PID+LQR control method in all respect of control techniques but differs only in number of PID controllers used. Here only cart PID controller has been used, and angle PID controller has not been used. Here only cart position x has been taken as variable of interest for control. The reference cart position has been set to 0.1 (m). The desired angle to be zero is directly taken care of by state feedback control of LQR which is designed using the linear state-space model of the system with vertically upright position as reference. The band limited white noise has been added as the disturbance input to the system as same. The simulation results for both cases are shown in Fig. 2.26 and Fig. 2.28 respectively. Here also responses of angle θ, angular velocity $\dot{\theta}$, cart position x, cart velocity $\dot{x}$, and control u have been plotted. It is observed that the pendulum stabilizes in vertically upright position quickly & smoothly after two minor undershoots and a minor overshoot for case of without disturbance input, and it stabilizes in vertically upright position with minute oscillations for case of with continuous disturbance input also. The angular velocity approaches 0 (rad/s) quickly for case of without disturbance input, and it oscillates by approx. +/-0.01 (rad/s) remaining at most in range approx +/-0.02 (rad/s) for case of with continuous disturbance input. The cart position x reaches the desired position of 0.1 (m) quickly & smoothly in approx. 6 seconds for both cases. The cart velocity reaches to zero for case of without disturbance input, and it oscillates very near to zero for case of with continuous disturbance input. The control input u is bounded for both cases in ranges [-0.1 0.1], and [-1 1] respectively. The simulation results justify the effectiveness & robustness of the cart PID+LQR control.

The maximum absolute values of system states & control showing maximum absolute variations with respect to desired nominal values in simulation for both cases of without & with disturbance input are shown in Table 2.3 & Table 2.4 respectively.

Table 2.3 Maximum absolute values of system states & control for without disturbance input case

Control Schemes	θ	$\dot{\theta}$	x	$\dot{x}$	u
PID	0.0046	---	0.0976	---	0.1402
2 PID+LQR	0.0029	0.0070	0.1000	0.0330	0.1500
1 PID+LQR	0.0029	0.0067	0.1000	0.0331	0.1500

Table 2.4 Maximum absolute values of states & control for with disturbance input case

Control Schemes	θ	$\dot{\theta}$	x	$\dot{x}$	u
PID	0.0118	---	0.1481	---	1.3598
2 PID+LQR	0.0030	0.0209	0.0993	0.0441	1.3774
1 PID+LQR	0.0030	0.0204	0.1019	0.0439	1.3882

Comparing the results it is observed that the responses of both alternatives of PID+LQR control method are better than PID control, which are smooth & fast also. It is also observed that the responses of 2PID+LQR control and cart PID+LQR control are similar. Just the cart position response of 2PID+LQR control is smoother than cart PID+LQR control and so it is slightly better, which is due to the additional degree of freedom of control added by the angle PID controller. But the cart PID+LQR control has structural simplicity in its credit. The analysis of the performances of the control schemes of PID control, 2PID+LQR control, and cart PID+LQR control for the nonlinear inverted pendulum-cart dynamical system without & with disturbance input gives that these control schemes are effective & robust. The advantage of this simulation study is that, it demonstrates that, the proposed PID+LQR control approach is a simple, effective & robust technique for the optimal control of nonlinear dynamical systems.

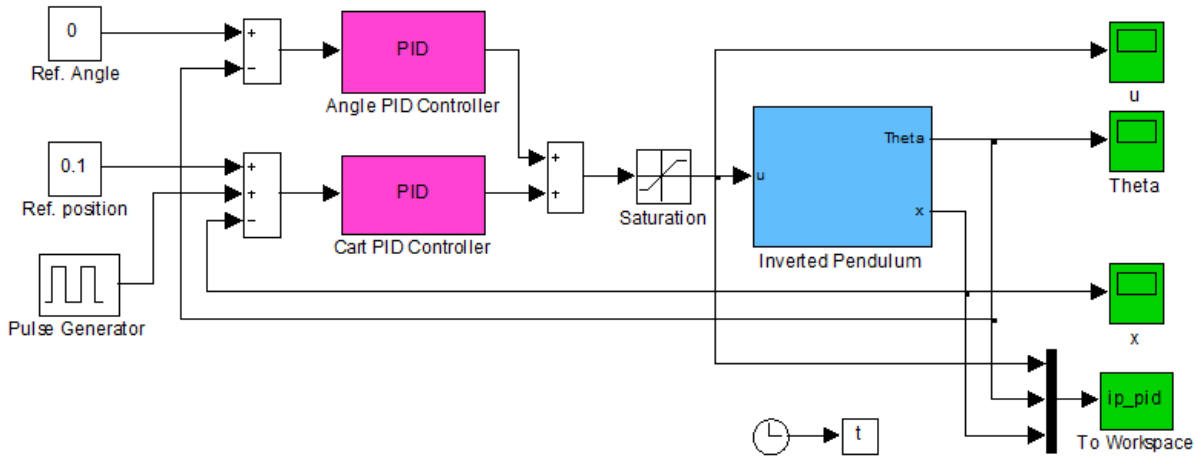


Fig. 2.17 PID control of nonlinear inverted pendulum system.

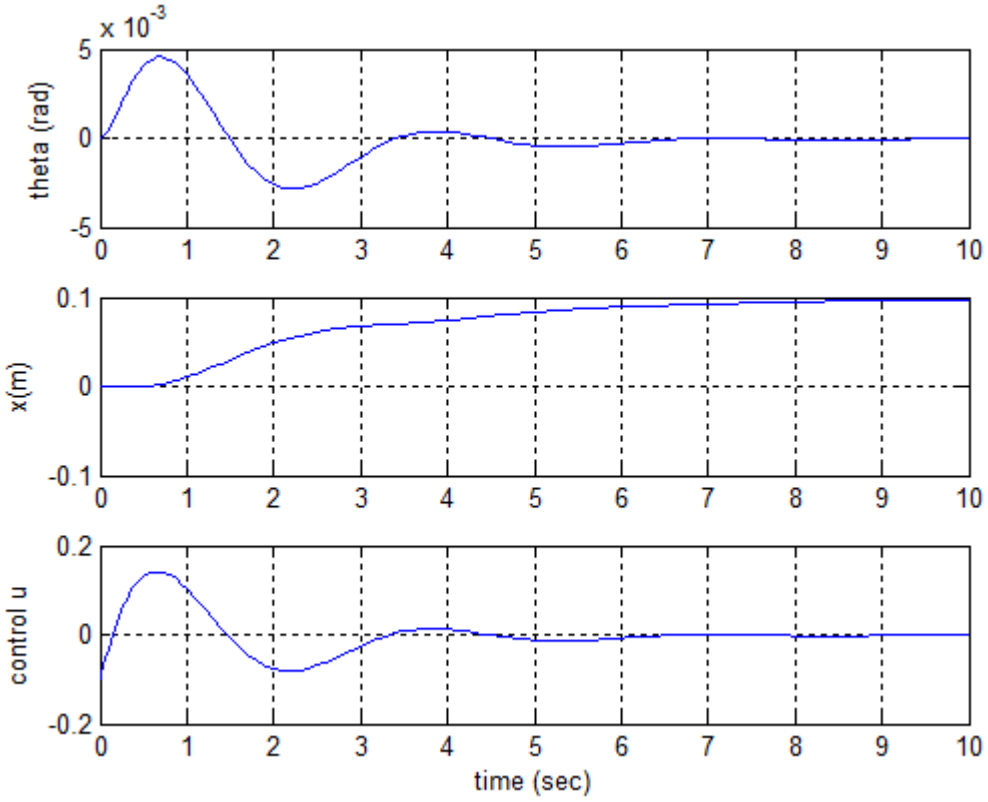


Fig. 2.18 Responses of pendulum angle θ, cart position x, and control force u of nonlinear inverted pendulum system with PID control.

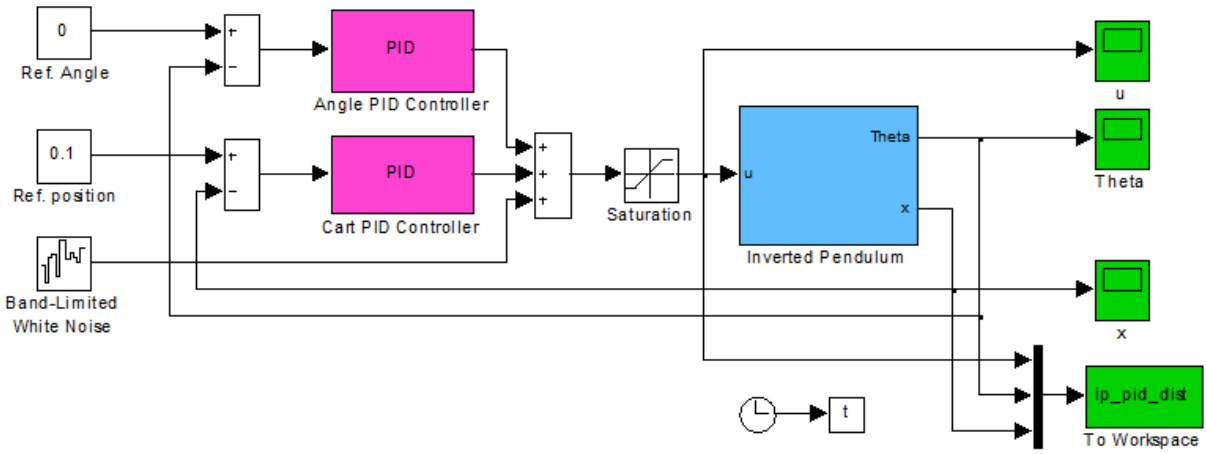


Fig. 2.19 PID control of nonlinear inverted pendulum system with disturbance input.

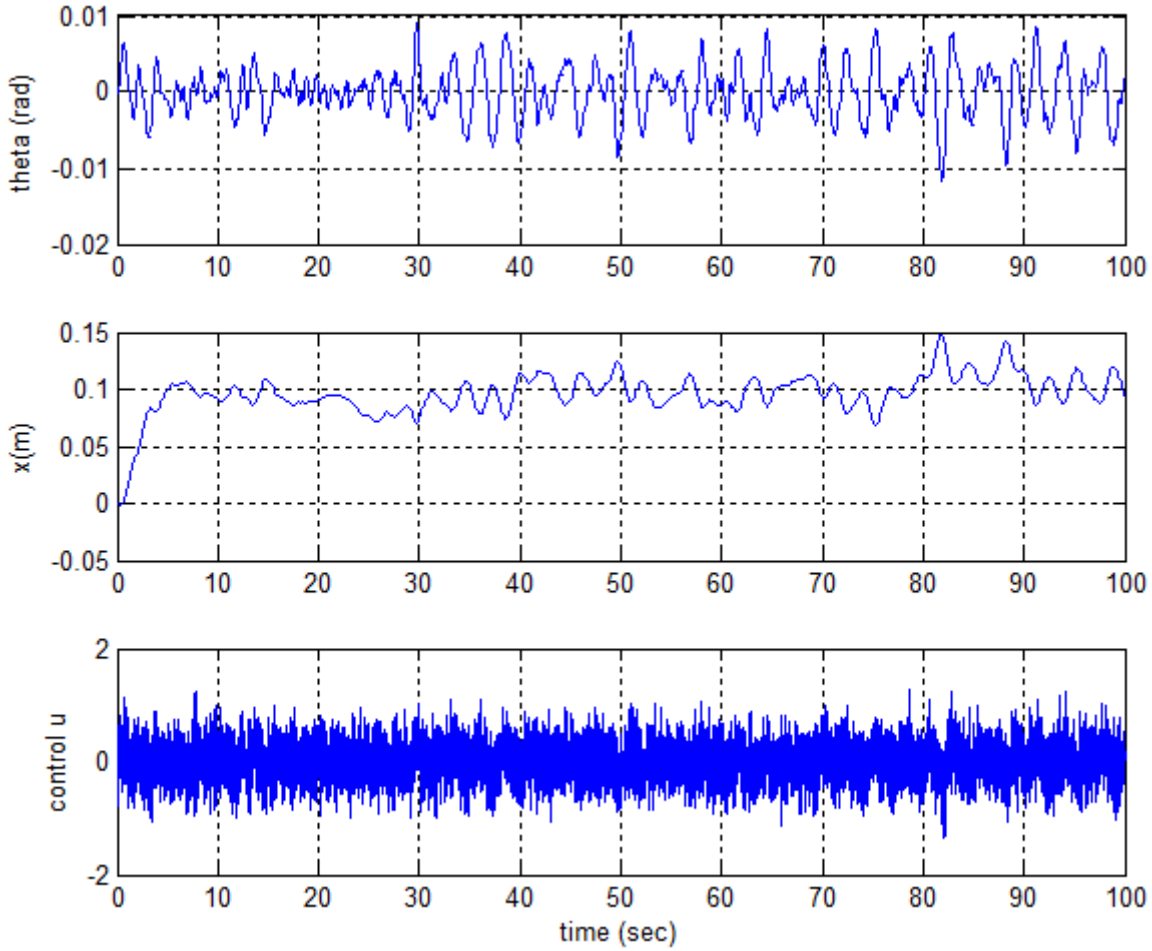


Fig. 2.20 Responses of pendulum angle θ, cart position x, and control force u of nonlinear inverted pendulum system with PID control with disturbance input.

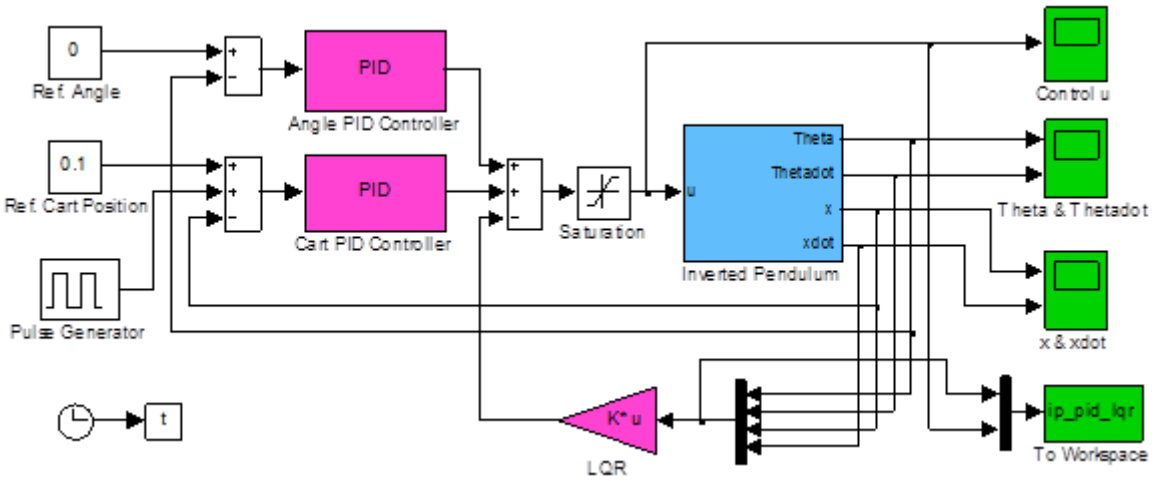


Fig. 2.21 Cart PID, angle PID & LQR control of nonlinear inverted pendulum system.

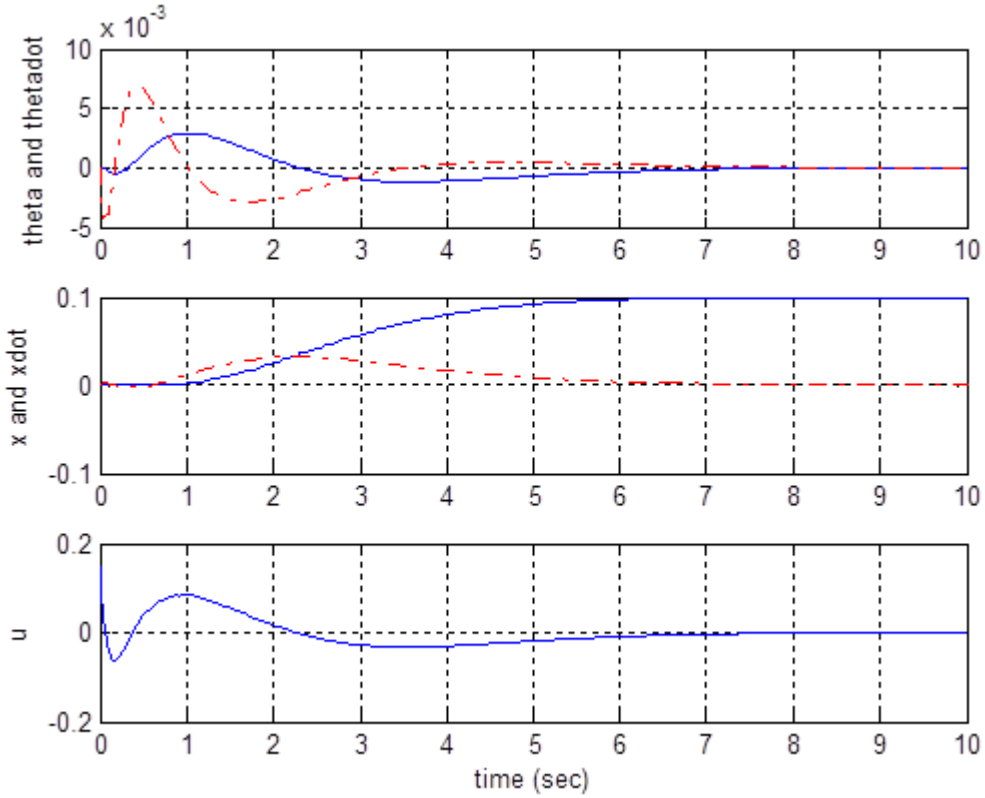


Fig. 2.22 Responses of pendulum angle θ (blue solid line), angular velocity $\dot{\theta}$ (red dashed line), cart position x (blue solid line), cart velocity $\dot{x}$ (red dashed line), and control force u of nonlinear inverted pendulum system with cart PID, angle PID & LQR control.

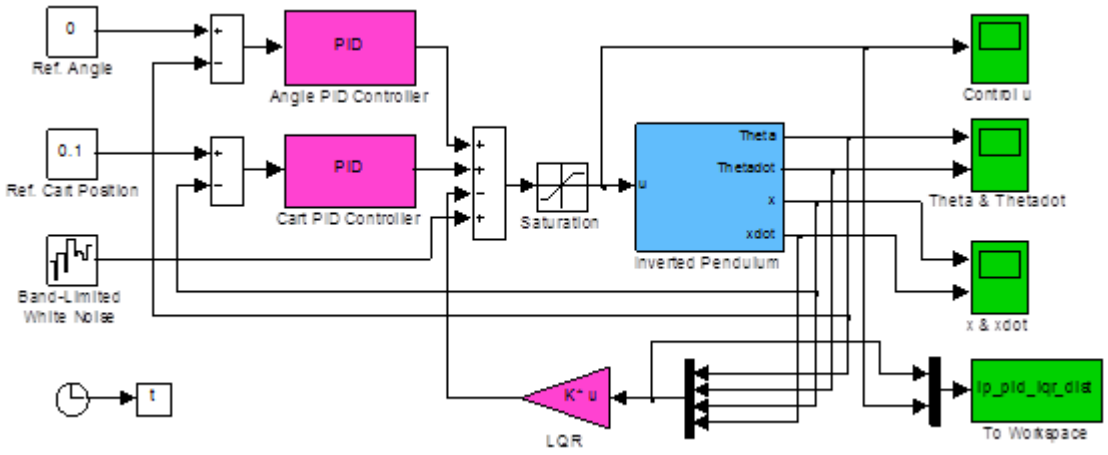


Fig. 2.23 Cart PID, angle PID & LQR control of nonlinear inverted pendulum system with disturbance input.

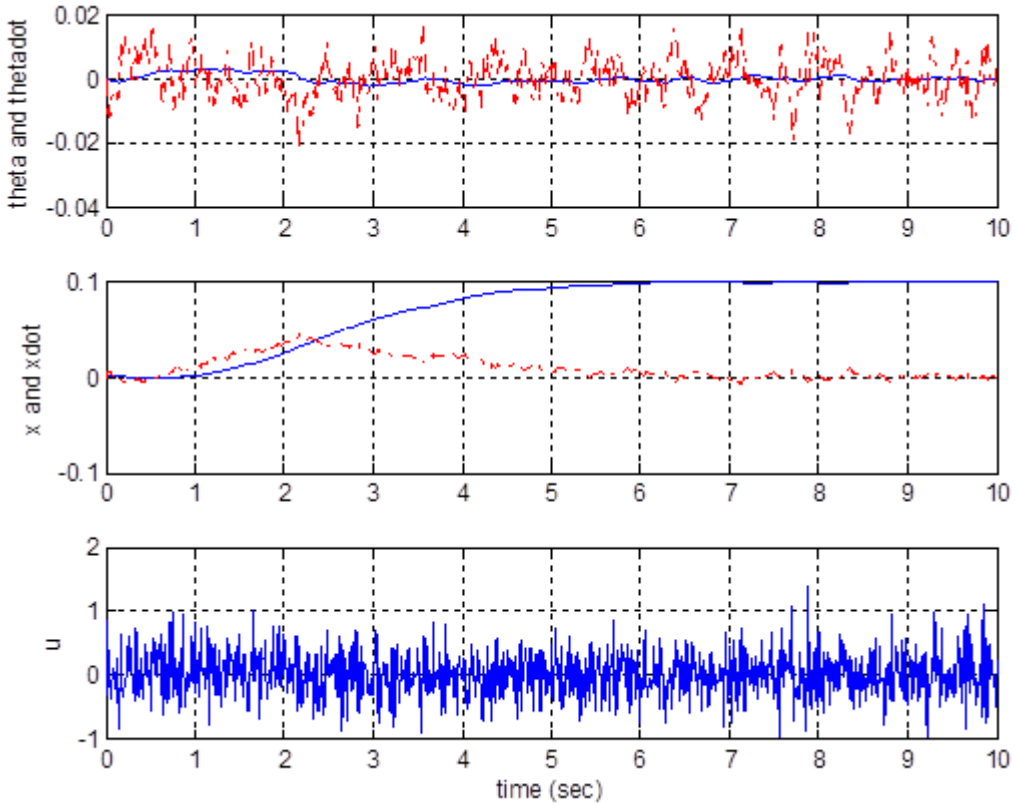


Fig. 2.24 Responses of pendulum angle θ (blue solid line), angular velocity $\dot{\theta}$ (red dashed line), cart position *x* (blue solid line), cart velocity $\dot{x}$ (red dashed line), and control force *u* of nonlinear inverted pendulum system with disturbance input using cart PID, angle PID & LQR control.

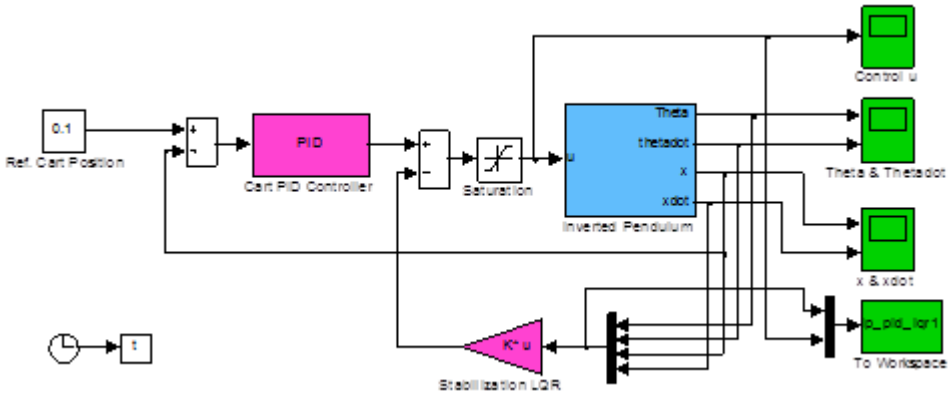


Fig. 2.25 Cart PID & LQR control of nonlinear inverted pendulum system.

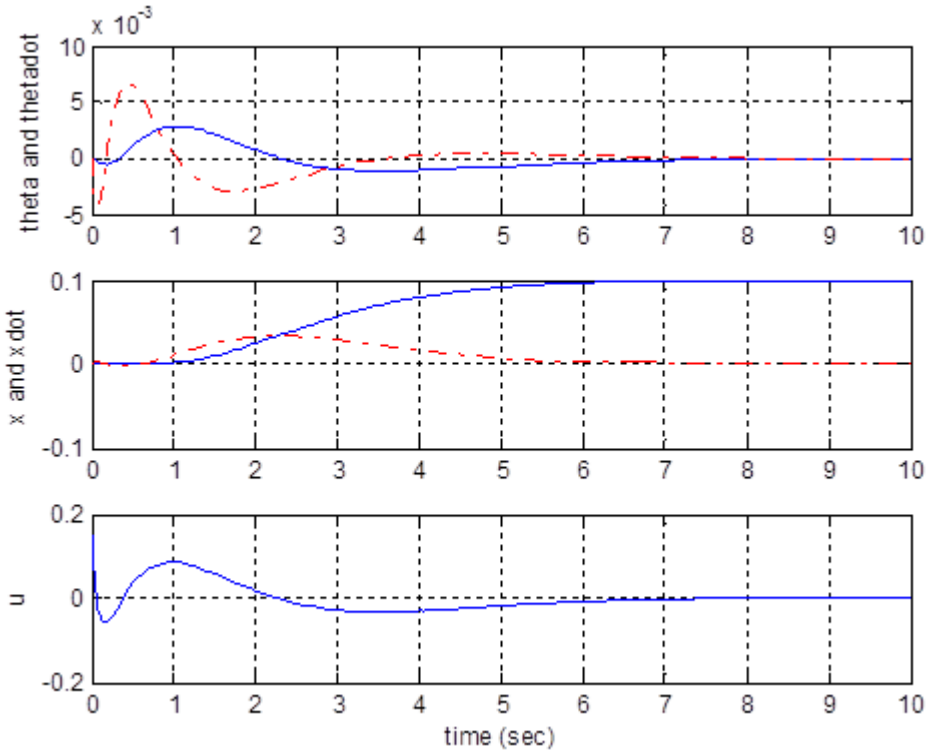


Fig. 2.26 Responses of pendulum angle θ (blue solid line), angular velocity $\dot{\theta}$ (red dashed line), cart position x (blue solid line), cart velocity $\dot{x}$ (red dashed line), and control force u of nonlinear inverted pendulum system with cart PID & LQR control.

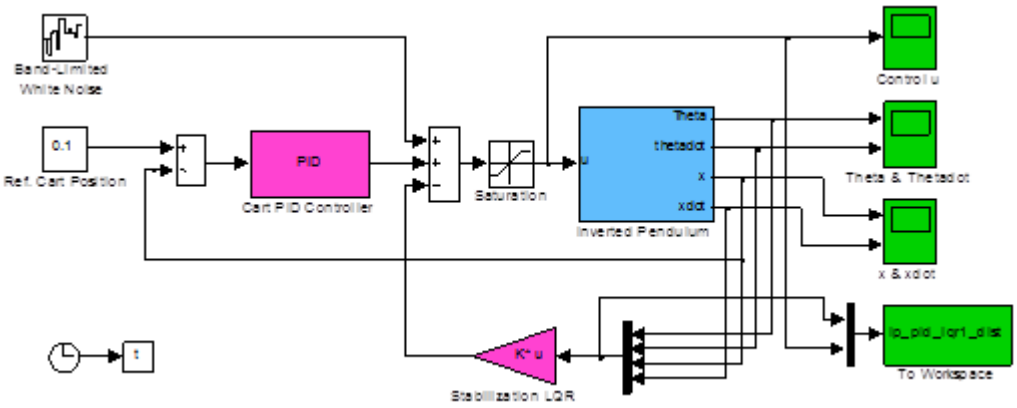


Fig. 2.27 Cart PID & LQR control of nonlinear inverted pendulum system with disturbance input.

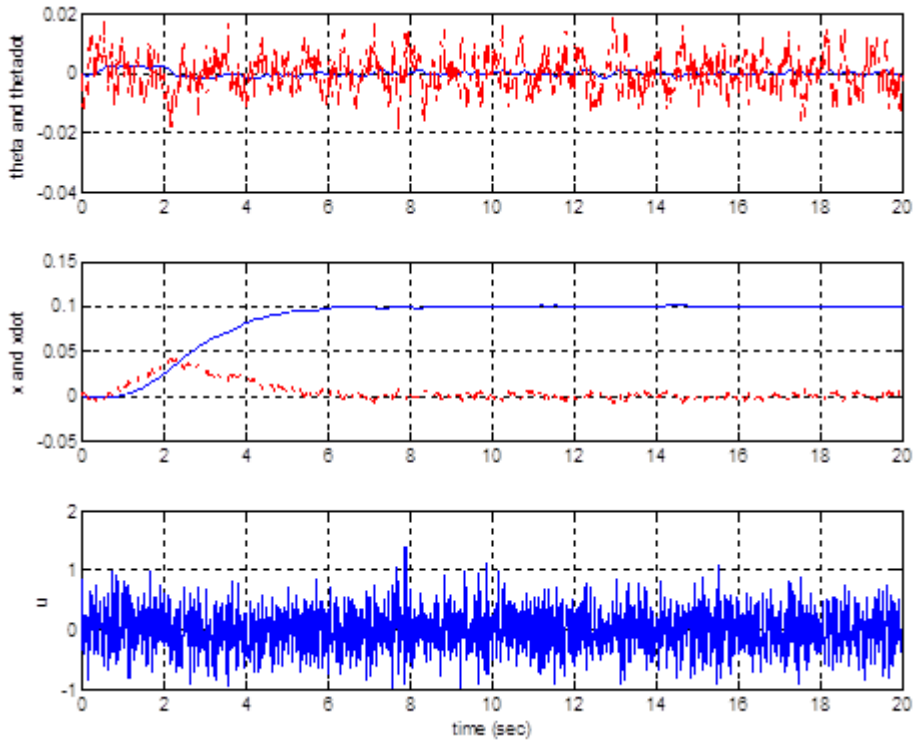


Fig. 2.28 Responses of pendulum angle θ (blue solid line), angular velocity $\dot{\theta}$ (red dashed line), cart position x (blue solid line), cart velocity $\dot{x}$ (red dashed line), and control force u of nonlinear inverted pendulum system with disturbance input using cart PID & LQR control.

2.5.4 Discussion

PID control, and LQR, an optimal control technique to make the optimal control decisions have been implemented to control the nonlinear inverted pendulum-cart system without & with continuous disturbance input. To compare the results of proposed PID+LQR control method, PID control method has been implemented. In the optimal control of nonlinear inverted pendulum dynamical system using PID controller and LQR approach all

the instantaneous states of the nonlinear system are considered available for measurement, which are directly fed to the LQR. The LQR is designed using the linear state-space model of the system. The optimal control value of LQR is added negatively with PID control value to have a resultant optimal control. The tuning of the PID controllers which are used here either as PID control method or PID+LQR control methods is done by trial & error method and observing the responses achieved to be optimal. The simulation results and performance analysis justify the comparative advantage of optimal control using LQR method. The pendulum stabilizes in upright position and cart reaches the desired position quickly & smoothly even under the continuous disturbance input such as wind force justify that the control schemes are effective & robust. The analysis of the responses of control schemes gives that the performance of proposed PID+LQR control method is better than PID control. The comparative performance investigation for this benchmark system establishes that the proposed PID+LQR control approach being simple, effective & robust control scheme for the optimal control of nonlinear dynamical systems. The performance investigation of this control approach with tuning of PID controller parameters using GA, and PSO instead of trial & error method may be done as a future scope of this work.

2.6 OPTIMAL CONTROL USING LINEAR QUADRATIC REGULATOR FOR AUTOMATIC GENERATION CONTROL OF TWO-AREA INTERCONNECTED POWER SYSTEM

The normal operation of interconnected power systems requires the balance of total power generation with total electrical load demand and associated system losses. The changes in normal operating point of a power system with time result in deviations of nominal system frequency and tie-line power. The frequency deviation is a direct and useful index to indicate the imbalance of power generation and demand. A prolonged off-normal frequency deviation has several undesirable effects such as directly affecting the power system operation, security, reliability and efficiency by damaging equipment, degrading load performance, overloading transmission lines, and triggering the protection devices. Automatic generation control (AGC) is a centralised real-time closed loop control process to balance the power generation and demand in electrical power systems at a minimum cost. The AGC system maintains the system frequency to specified nominal value and regulates the tie-line power flows between control areas at economic dispatch; and thus, minimizes the deviations of system frequency and tie-line power. The AGC system can include one or more control subsystems such as load frequency control (LFC), economic dispatch control, environmental dispatch control, security dispatch control and the like. Despite the load variations in different control areas the system frequency and tie-line power are maintained to their scheduled values, this function of AGC is commonly known as automatic load

frequency control (ALFC). The term LFC is synonymously used for the term AGC in literature. The function of AGC to allocate the power generation requirement among generating units to minimize (optimize) the incremental cost of delivered power is known as economic dispatch. The AGC system sends signals to the under-control generating units to realize its function, which performance depends on the way the generating units respond to the signals. The response characteristics of generating units depend on many factors such as type of unit, fuel, control strategy, and operating point. The energy control centres have the energy management system (EMS) consisting of the AGC, security control, supervisory control and data acquisition (SCADA) and load management to manage electric energy. Among the various function of AGC system, the LFC is the most important issue in power system [184, 186].

The critical review on recent philosophies of AGC is presented in [187]. The various approaches of AGC are presented in [91, 92, 183-185, 187].

The interconnected power systems operation and control is required to be with optimal performance for which optimal control design for AGC problem is needed. The linear quadratic regulator (LQR) provides an optimal control solution for linear dynamical systems. In this section the performance analysis of optimal control using LQR for AGC of two-area interconnected power system is presented. The comparative performance analysis of conventional AGC scheme using integral control, optimal control using LQR and optimal integral control using integral controller and LQR is also presented.

2.6.1 Mathematical Modelling for Automatic Generation Control of Two-Area Interconnected Power System

The monitoring and control of interconnected power system is divided between several control areas. A control area or region represents a coherent group of generating units and loads where all the generators respond in unison to changes in load or speed changer settings. The frequency is assumed to remain same throughout a control area in both static and dynamic conditions. The various control areas are interconnected through tie-lines [183, 184]. A sudden load disturbance in any area causes the deviations in frequencies of other areas and tie-lines powers. In order to maintain the system frequency at its nominal value and the tie-lines power flows between areas at its scheduled value, the automatic generation control (AGC) which is also known as the load frequency control (LFC) balances the power generation and demand in each control area. There are two variables of interest in this, frequency and tie-line power flows. The linear combination of weighted variations of these variables is represented by a single variable known as the area control error (ACE). As the AGC drives the ACE to zero, the deviations in system frequency and tie-line power becomes zero, and the power system operates in equilibrium.

The schematic diagram of frequency control mechanism of electrical power system is shown in Fig. 2.29 [186]. Based on the accepted frequency operating standards there are three operating conditions corresponding to the frequency variation ranges: primary control, secondary or supplementary control, and emergency control. Under the normal operation, the small initial frequency deviation is overcome by the governing mechanism of turbine-generator unit. Under this normal operating condition, the natural governor response is known as the primary control. The governor senses the machine speed, and adjusts the input valve setting to change the mechanical power output to track the load change and to restore frequency to a nominal value. Under the off-normal operation, the larger frequency deviation is overcome by the secondary control, which is the task of LFC or AGC. The AGC restores the nominal system frequency according to the available amount of power reserve. However, under the emergency conditions of serious load-generation imbalance and the associated rapid frequency changes following a significant fault, the AGC system may be unable to restore the system frequency via the secondary frequency control loop. Under this emergency operation, the emergency control and protection schemes, such as under-frequency load shedding (UFLS) are used to restore the normal operation of power system.

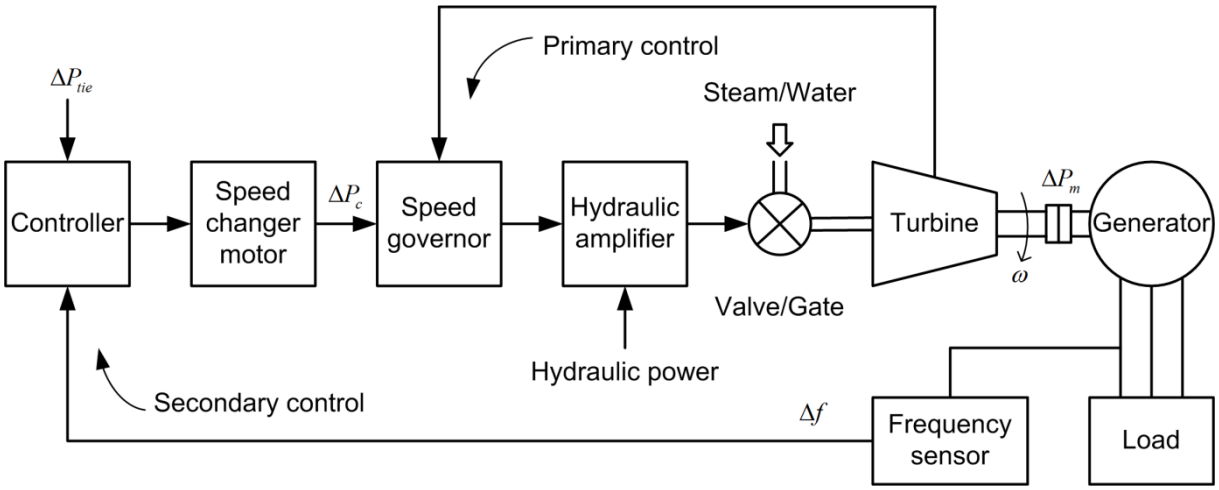


Fig. 2.29 Frequency control mechanism.

2.6.1.1 Two-Area Interconnected Power System Dynamic Equations

The power system is exposed to small changes in load under its normal operating state, and undergoes small deviations of system states, thus the power system dynamics can be represented by the linearized perturbation model around the operating point. The linearized model can be used to control design using the linear control theory. The block diagram of transfer function model for AGC of two-area interconnected power system is shown in Fig. 2.30 [184]. For convenience, each control area is represented by an equivalent generating unit consisting of equivalent governor, turbine and generator system. In each area the power is transported by tie-line, which is accounted for incremental power balance equation of each area.

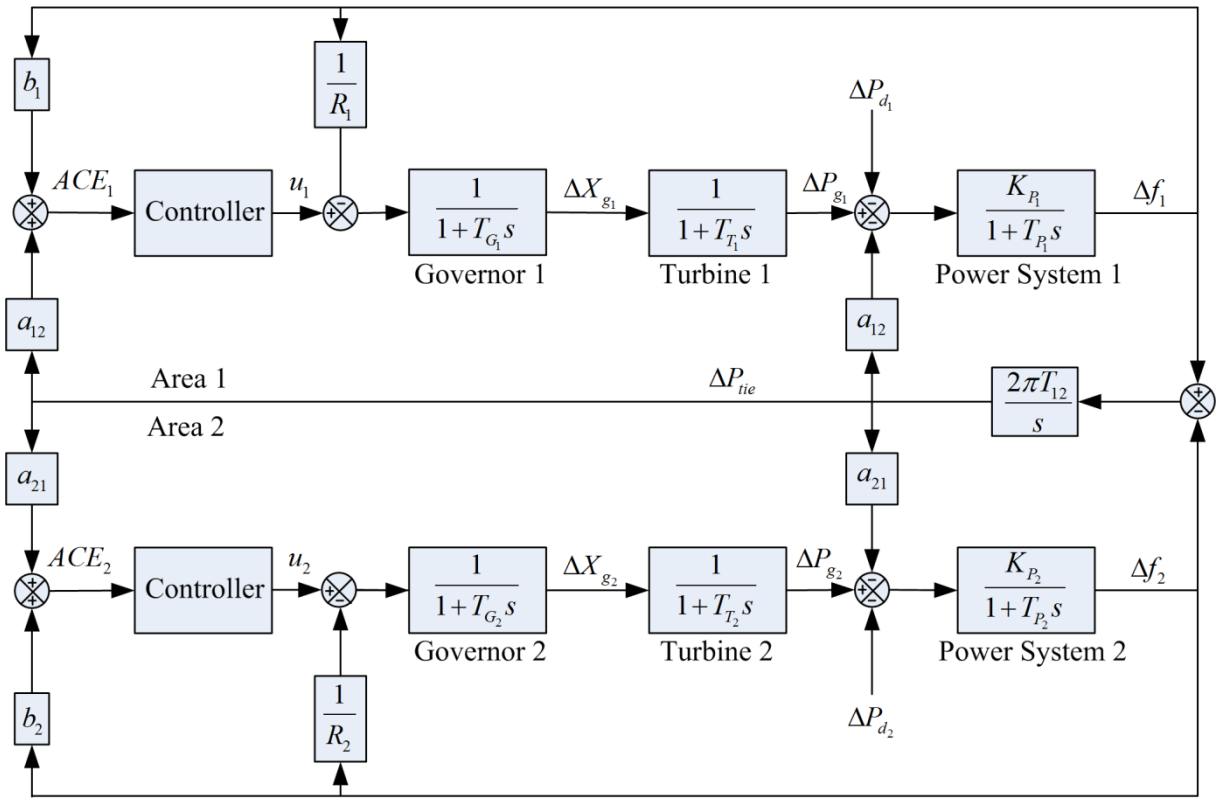


Fig. 2.30 Block diagram of two-area power system automatic generation control.

For AGC system design and analysis the linearized incremental model of two-area interconnected power system is given by the following dynamic equations [184].

For k^{th} control area, k=1, 2, the system dynamic equations are as following:

The incremental power balance equation which gives the power system generator-load dynamics is given by

$$\Delta \dot{f}_k(t) = -\frac{1}{T_{P_k}}\Delta f_k(t) + \frac{K_{P_k}}{T_{P_k}}\Delta P_{g_k}(t) - \frac{K_{P_k}}{T_{P_k}}\Delta P_{tie_k}(t) - \frac{K_{P_k}}{T_{P_k}}\Delta P_{d_k}(t) \tag{2.67}$$

The incremental power generation of turbine system is given by dynamic equation

$$\Delta \dot{P}_{g_k}(t) = -\frac{1}{T_{T_k}}\Delta P_{g_k}(t) + \frac{1}{T_{T_k}}\Delta X_{g_k}(t) \tag{2.68}$$

The speed governor system dynamics with incremental change in governor valve position is given by

$$\Delta \dot{X}_{g_k}(t) = -\frac{1}{R_k T_{G_k}}\Delta f_k(t) - \frac{1}{T_{G_k}}\Delta X_{g_k}(t) + \frac{1}{T_{G_k}}\Delta P_{c_k}(t) \tag{2.69}$$

The dynamic equation of incremental tie-line power can be written as

$$\Delta \dot{P}_{tie_{kj}}(t) = 2\pi T_{kj}\Delta f_k(t) - 2\pi T_{kj}\Delta f_j(t) \tag{2.70}$$

where $k \in J_2 = \{1, 2\}$, $j \in J_2 \setminus \{k\}$ and

Δf_k = the incremental frequency deviation (Hz);

ΔP_{g_k} = the incremental change in generator output (pu MW);

ΔX_{g_k} = the incremental change in governor valve position (pu MW);

ΔP_{c_k} = the incremental change in speed changer position (pu MW) which acts as the control input;

T_{G_k} = governor time constants (s);

T_{T_k} = turbine time constants (s);

T_{P_k} = power system generator-load model time constants (s), given as $T_{P_k} = \frac{2H_k}{D_k f_k^0}$;

K_{P_k} = power system generator-load model gains (Hz /pu MW), given as $K_{P_k} = \frac{1}{D_k}$;

H_k = power system generator-load inertia constants (pu MW/Hz);

D_k = load damping coefficient (pu MW/Hz);

f_k^0 = nominal system frequency (Hz);

R_k = speed regulation due to governor action (Hz./Pu MW)

ΔP_{d_k} = load disturbance (pu MW);

$\Delta P_{tie_{kj}}$ = the incremental tie-line power flowing from area *k* to *j* (pu MW);

T_{kj} = the electrical stiffness, also known as the synchronizing coefficient between area *k* and *j*, given by

$$T_{kj} = \frac{|V_k||V_j|}{P_{r_k} X_{kj}} \cos(\delta_k - \delta_j) \tag{2.71}$$

where

δ_k = the power angle of equivalent machine, and $\frac{d\delta_k}{dt} = 2\pi f_k$;

V_k = the area bus voltage;

X_{kj} = the tie-line reactance between area *k* and *j*;

P_{r_k} = the rated power of area *k*;

From (2.70) and (2.71) we have

$$\frac{\Delta P_{tie_{12}}}{\Delta P_{tie_{21}}} = -\frac{T_{12}}{T_{21}} = -\frac{P_{r_2}}{P_{r_1}} = -\frac{1}{a_{12}} \tag{2.72}$$

where a_{12} is the area 1 to area 2 size ratio constant,

and for generalization, we can write

$$\Delta P_{tie_{kj}} = a_{kj}\Delta P_{tie}\,,\ \Delta P_{tie_{12}} = a_{12}\Delta P_{tie}\,,\ \Delta P_{tie_{21}} = a_{21}\Delta P_{tie}\,,\ a_{12} = -a_{21} = 1 \tag{2.73}$$

where ΔP_{tie} = incremental tie-line power, and a_{kj} = area *k* to *j* size ratio constant;

The area control error (ACE) is the change in area frequency which when used in integral control loop the steady state frequency error becomes zero. In order to make steady state tie-line power error zero in a two-area control the integral control loops in each area be included to integrate the incremental tie-line power signal and fed it back to the speed changer. This is done by a single integrating block by defining ACE as a linear combination of incremental frequency and tie-lie power. The ACE is given by

$$ACE_k(t) = \Delta P_{tie_{kj}}(t) + b_k \Delta f_k(t) \tag{2.74}$$

where b_k = the frequency bias (pu MW/Hz);

The speed changer settings ΔP_{c_k} can be generated by the integrals of ACE_k, then by considering a variable $\Delta E_k(t) = \int ACE_k(t)dt$, from (2.74) the ACE dynamics is given by

$$\Delta \dot{E}_k(t) = \Delta P_{tie_{kj}}(t) + b_k \Delta f_k(t) \tag{2.75}$$

Thus, in the conventional integral control scheme of AGC the speed changer settings ΔP_{c_k} can be given by

$$\Delta P_{c_k}(t) = -K_{I_k} \int ACE_k(t)dt \tag{2.76}$$

where K_{I_k} = integral gain.

Equations (2.67) to (2.76) give the system dynamics of two-area interconnected power system for AGC system design and analysis.

The block diagram of transfer function model of AGC system using integral control for two-area interconnected power system is shown in Fig. 2.31 [184].

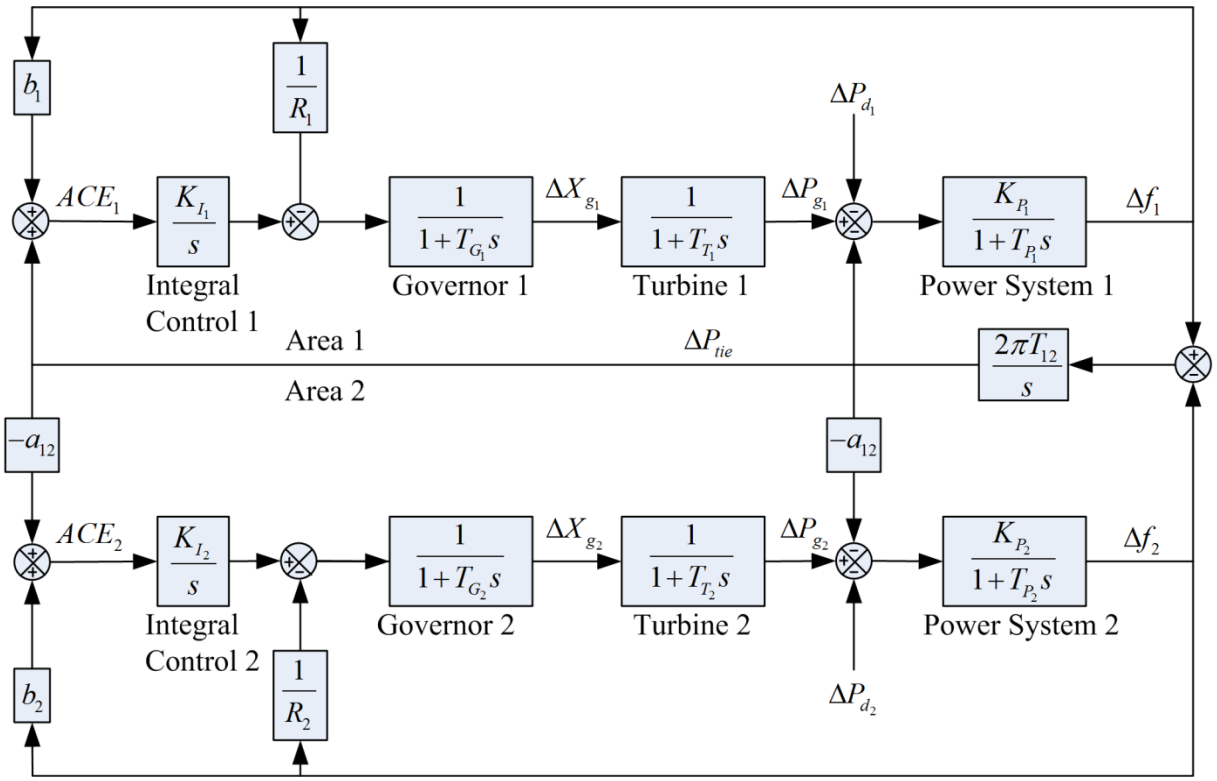


Fig. 2.31 Block diagram of AGC system using integral control for two-area interconnected power system.

2.6.1.2 State Space Model of Two-Area Interconnected Power System

The state space model of two-area interconnected power system can be obtained from the dynamic equations (2.67) to (2.76) considering the state variables as [184] $x_1(t) = \Delta f_1(t)$, $x_2(t) = \Delta P_{G_1}(t)$, $x_3(t) = \Delta X_{g_1}(t)$, $x_4(t) = \Delta f_2(t)$, $x_5(t) = \Delta P_{G_2}(t)$, $x_6(t) = \Delta X_{g_2}(t)$, $x_7(t) = \Delta P_{tie}(t)$, $x_8(t) = \Delta E_1(t) = \int ACE_1 dt$, $x_9(t) = \Delta E_2(t) = \int ACE_2 dt$, control inputs $u_1(t) = \Delta P_{c_1}(t)$, $u_2(t) = \Delta P_{c_2}(t)$, and load disturbances $w_1(t) = \Delta P_{d_1}(t)$, $w_2(t) = \Delta P_{d_2}(t)$.

Thus, the system equations of two-area interconnected power system are written as

$$\dot{x}_1(t) = -\frac{1}{T_{P_1}} x_1(t) + \frac{K_{P_1}}{T_{P_1}} x_2(t) - \frac{K_{P_1}}{T_{P_1}} x_7(t) - \frac{K_{P_1}}{T_{P_1}} w_1(t) \tag{2.77}$$

$$\dot{x}_2(t) = -\frac{1}{T_{T_1}} x_2(t) + \frac{1}{T_{T_1}} x_3(t) \tag{2.78}$$

$$\dot{x}_3(t) = -\frac{1}{R_1 T_{G_1}} x_1(t) - \frac{1}{T_{G_1}} x_3(t) + \frac{1}{T_{G_1}} u_1(t) \tag{2.79}$$

$$\dot{x}_4(t) = -\frac{1}{T_{P_2}} x_4(t) + \frac{K_{P_2}}{T_{P_2}} x_5(t) + \frac{a_{12} K_{P_2}}{T_{P_2}} x_7(t) - \frac{K_{P_2}}{T_{P_2}} w_2(t) \tag{2.80}$$

$$\dot{x}_5(t) = -\frac{1}{T_{T_2}} x_5(t) + \frac{1}{T_{T_2}} x_6(t) \tag{2.81}$$

$$\dot{x}_6(t) = -\frac{1}{R_2 T_{G_2}} x_4(t) - \frac{1}{T_{G_2}} x_6(t) + \frac{1}{T_{G_2}} u_2(t) \tag{2.82}$$

$$\dot{x}_7(t) = 2\pi T_{12} x_1(t) - 2\pi T_{12} x_4(t) \tag{2.83}$$

$$\dot{x}_8(t) = b_1 x_1(t) + x_7(t) \tag{2.84}$$

$$\dot{x}_9(t) = b_2 x_4(t) - a_{12} x_7(t) \tag{2.85}$$

Equations (2.77) to (2.85) can be written in state vector matrix form as

$$\dot{x} = Ax + Bu + Fw \tag{2.86}$$

where

$$x(t) = \left[x_1(t), x_2(t), ..., x_9(t)\right]^T$$
$$= \left[\Delta f_1(t), \Delta P_{g_1}(t), \Delta X_{g_1}(t), \Delta f_2(t), \Delta P_{g_2}(t), \Delta X_{g_2}(t), \Delta P_{tie}(t), \int ACE_1 dt, \int ACE_2 dt\right]^T$$

is the state vector, $u(t) = \begin{bmatrix} u_1(t) & u_2(t) \end{bmatrix}^T = \begin{bmatrix} \Delta P_{c_1}(t) & \Delta P_{c_2}(t) \end{bmatrix}^T$ is the control vector, and $w(t) = \begin{bmatrix} w_1(t) & w_2(t) \end{bmatrix}^T = \begin{bmatrix} \Delta P_{d_1}(t) & \Delta P_{d_2}(t) \end{bmatrix}^T$ is the load disturbance vector, and matrices *A*, *B*, and *F* are given as

$$A = \begin{bmatrix}
-\frac{1}{T_{P_1}} & \frac{K_{P_1}}{T_{P_1}} & 0 & 0 & 0 & 0 & -\frac{K_{P_1}}{T_{P_1}} & 0 & 0 \\
0 & -\frac{1}{T_{T_1}} & \frac{1}{T_{T_1}} & 0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{1}{R_1 T_{G_1}} & 0 & -\frac{1}{T_{G_1}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{T_{P_2}} & \frac{K_{P_2}}{T_{P_2}} & 0 & \frac{a_{12} K_{P_2}}{T_{P_2}} & 0 & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{T_{T_2}} & \frac{1}{T_{T_2}} & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{R_2 T_{G_2}} & 0 & -\frac{1}{T_{G_2}} & 0 & 0 & 0 \\
2\pi T_{12} & 0 & 0 & -2\pi T_{12} & 0 & 0 & 0 & 0 & 0 \\
b_1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & b_2 & 0 & 0 & -a_{12} & 0 & 0
\end{bmatrix}$$

$$B=\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{T_{G_1}} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{T_{G_2}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\quad F=\begin{bmatrix} -\frac{K_{P_1}}{T_{P_1}} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -\frac{K_{P_2}}{T_{P_2}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

In the conventional integral control scheme, the control inputs are determined by

$$u_1=-K_{I_1}x_8=-K_{I_1}\int ACE_1dt \tag{2.87}$$

$$u_2=-K_{I_2}x_9=-K_{I_2}\int ACE_2dt \tag{2.88}$$

Replacing (2.87) and (2.88) into (2.86), the state equation of AGC system using integral control can be written as

$$\dot{x}=A_w x+Fw \tag{2.89}$$

where

$$A_w=\begin{bmatrix} -\frac{1}{T_{P_1}} & \frac{K_{P_1}}{T_{P_1}} & 0 & 0 & 0 & 0 & -\frac{K_{P_1}}{T_{P_1}} & 0 & 0 \\ 0 & -\frac{1}{T_{T_1}} & \frac{1}{T_{T_1}} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{R_1T_{G_1}} & 0 & -\frac{1}{T_{G_1}} & 0 & 0 & 0 & 0 & -\frac{K_{I_1}}{T_{G_1}} & 0 \\ 0 & 0 & 0 & -\frac{1}{T_{P_2}} & \frac{K_{P_2}}{T_{P_2}} & 0 & \frac{a_{12}K_{P_2}}{T_{P_2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_{T_2}} & \frac{1}{T_{T_2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{R_2T_{G_2}} & 0 & -\frac{1}{T_{G_2}} & 0 & 0 & -\frac{K_{I_2}}{T_{G_2}} \\ 2\pi T_{12} & 0 & 0 & -2\pi T_{12} & 0 & 0 & 0 & 0 & 0 \\ b_1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & b_2 & 0 & 0 & -a_{12} & 0 & 0 \end{bmatrix}$$

Considering the frequency deviations of area 1 and area 2, and tie-line power deviation as output variables, the output equation can be written as

$$y=Cx \tag{2.90}$$

where $y(t) = \begin{bmatrix} y_1(t) & y_2(t) & y_3(t) \end{bmatrix}^T = \begin{bmatrix} \Delta f_1(t) & \Delta f_2(t) & \Delta P_{tie}(t) \end{bmatrix}^T$ is the output vector, and output matrix *C* is given as

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

2.6.2 Control Methods

The schemes by which ACEs are processed and control signals are generated give the control design methods for AGC problem. In this section, following three control methods of AGC for two-area interconnected power system are presented.

2.6.2.1 Integral Control

The integral control is the conventional control design approach for AGC system. The integral control makes the steady state error of system frequency of both areas and steady state tie-line power error to zero. The integral control of two-area interconnected power system for AGC problem is defined by (2.87) and (2.88) for area 1 and area 2 respectively. The control inputs are generated by weighted integrals of ACEs in each area. The integral gains are tuned for satisfactory response of area frequency deviations and tie-line power deviation.

2.6.2.2 Optimal Control Using LQR

The optimal control design for dynamic system defined by (2.86) is obtained by synthesizing the control vector *u* by a linear combination of all the states *x* (i.e. full state feedback) with feedback constants determined by an optimality criterion.

Since the state equation (2.86) contain an additional disturbance term F_w unlike the standard form of state equation. A constant disturbance vector *w* would drive some of the system states and the control vector *u* to constant steady state values; while the cost function used in optimal control design requires that the system state and control vectors have zero steady state values to have a minimum for the cost function.

Define system state and control vectors as the sum of transient and steady state terms as

$$x = x_{ts} + x_{ss} \tag{2.91}$$

$$u = u_{ts} + u_{ss} \tag{2.92}$$

where x_{ts} and u_{ts} are transient state terms, and x_{ss} and u_{ss} are steady state terms of system state vector and control vector respectively.

For a constant disturbance vector *w*, the steady state is reached when $\dot{x} = 0$ in (2.86), which gives

$$0 = Ax_{ss} + Bu_{ss} + Fw \tag{2.93}$$

Substituting (2.91) and (2.92) into (2.86), we have

$$\dot{x}_{ts} = A(x_{ts} + x_{ss}) + B(u_{ts} + u_{ss}) + Fw \tag{2.94}$$

Substituting (2.93) into (2.94), we have

$$\dot{x}_{ts} = Ax_{ts} + Bu_{ts} \tag{2.95}$$

Equation (2.95) represents the system model on terms of excursion of state and control vectors from their respective steady state values.

For the full state feedback control

$$u = -Kx \tag{2.96}$$

where *K* is the feedback gain matrix.

Substituting (2.91) and (2.92) into (2.96) we have

$$u_{ts} + u_{ss} = -K(x_{ts} + x_{ss}) \tag{2.97}$$

For a stable system both x_{ts} and u_{ts} go to zero, therefore

$$u_{ss} = -Kx_{ss} \tag{2.98}$$

Hence

$$u_{ts} = -Kx_{ts} \tag{2.99}$$

For constant values of disturbance inputs w_1 and w_2 the steady state values of state and control variables can be obtained by observing (2.77) to (2.95). These steady state values are

$x_{1ss} = x_{4ss} = x_{7ss} = 0$, $x_{2ss} = x_{3ss} = w_1$, $x_{5ss} = x_{6ss} = w_2$, $u_{1ss} = w_1$, $u_{2ss} = w_2$,

x_{8ss} =constant, x_{9ss} =constant

The values of x_{8ss} and x_{9ss} depend upon the feedback constants.

To transfer an arbitrary initial state $x_{ts}(0)$ to origin in infinite time (i.e. $x_{ts}(\infty) = 0$), using control u_{ts}, the state feedback gain matrix *K* is required to be determined for system (2.95). The linear quadratic regulator (LQR) gives an infinite horizon optimal control solution, and thus transfers the system states $x(t)$ from an arbitrary initial state $x(0)$ at $t = 0$ to origin in infinite time (i.e. $x(\infty) = 0$, at $t = \infty$) by minimizing a quadratic cost function or performance index (PI). The quadratic cost function is given by

$$J = \frac{1}{2}\int_0^\infty \left(x^T Qx + u^T Ru\right)dt \tag{2.100}$$

where *Q*, and *R* are positive semi-definite and positive definite real symmetric constant matrices respectively.

The optimal control design for system (2.95) using cost function (2.100) terms of excursions of state and control vectors lead to the optimal control of system (2.86) satisfying the steady state and transient state conditions, thus the quadratic cost function is determined by considering the excursions of states and control inputs. The excursions considered in design are:

1. Excursions of ACEs ($x_{7ts}+b_1x_{1ts}$; $-a_{12}x_{7ts}+b_2x_{4ts}$) about steady state values are minimized. The steady state values of ACEs are zero. The excursions of ACEs also represent the excursions of states x_1, x_4, and x_7.
2. Excursions of states x_2, x_3, x_5, x_6 (x_{2ts}; x_{3ts}; x_{5ts}; x_{6ts}) about steady state are minimized. The steady state values of these states are constant.
3. Excursions of $\int ACE\ dt$ (x_{8ts}; x_{9ts}) about the steady state values are minimized. The steady state values of $\int ACE\ dt$ are constant.
4. Excursions of control vector (u_{1ts}; u_{2ts}) about steady state values are minimized. The steady state value of control vector is constant. The objective of this minimization is to limit the control effort indirectly within the physical capability of components. For example, the steam valve cannot be opened more than a certain value without causing the boiler pressure to drop severely.

Thus considering the above criterion, the cost function (2.100) can be written as

$$J=\frac{1}{2}\int_0^\infty\left[\left(ACE_1^2+ACE_2^2\right)+\Delta P_{g_1}^2+\Delta X_{g_1}^2+\Delta P_{g_2}^2+\Delta X_{g_2}^2+\left(\int ACE_1dt\right)^2+\left(\int ACE_2dt\right)^2+k\left(u_1^2+u_2^2\right)\right]dt$$

or

$$J=\frac{1}{2}\int_0^\infty\left[b_1^2x_1^2+x_2^2+x_3^2+b_2^2x_4^2+x_5^2+x_6^2+(1+a_{12}^2)x_7^2+x_8^2+x_9^2+2b_1x_1x_7-2a_{12}b_2x_4x_7+k(u_1^2+u_1^2)\right]dt \tag{2.101}$$

where *k* is a real positive constant.

Comparing (2.100) and (2.101), we have

$$Q=\begin{bmatrix} b_1^2 & 0 & 0 & 0 & 0 & 0 & b_1 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & b_2^2 & 0 & 0 & -a_{12}b_2 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ b_1 & 0 & 0 & -a_{12}b_2 & 0 & 0 & (1+a_{12}^2) & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},\quad R=k\begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}$$

The infinite horizon optimal control solution by the minimization of cost function (2.100) is obtained by the solution of algebraic Riccati equation (ARE)

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \tag{2.102}$$

where *P* is a positive definite real symmetric constant matrix.

The matrix *P* gives the LQR gain matrix *K*, which is given by

$$K = R^{-1}B^T P \tag{2.103}$$

Using the LQR gain matrix *K*, the optimal control for the dynamic system (2.86) is obtained. Replacing $u = -Kx$ in (2.86) we have the stable system

$$\dot{x} = (A - BK)x + Fw \tag{2.104}$$

2.6.2.3 Optimal Control Using Integral Controller and LQR

The state dynamics of two-area interconnected power system automatic generation control with integral controllers (2.87) and (2.88) is defined by (2.89). The dynamic system (2.89) is in the standard form of state equation. The optimal performance of this system depends on the optimal values of load disturbance inputs and optimal values system states with integral control, and thus the optimal values of control inputs, that is the speed changer settings. The optimal control design using LQR for dynamic system (2.89) is obtained by minimizing the quadratic cost function, which is given by

$$J = \frac{1}{2}\int_0^\infty \left(x^T Q x + w^T R w\right) dt \tag{2.105}$$

Considering the constraints of states similar to the excursions of states in (2.101) and constraints of load disturbance input vector the cost function (2.39) is determined, which is given by

$$J = \frac{1}{2}\int_0^\infty \left[\left(ACE_1^2 + ACE_2^2\right) + \Delta P_{g_1}^2 + \Delta X_{g_1}^2 + \Delta P_{g_2}^2 + \Delta X_{g_2}^2 + \left(\int ACE_1 dt\right)^2 + \left(\int ACE_2 dt\right)^2 + k\left(w_1^2 + w_2^2\right)\right] dt \tag{2.106}$$

Comparing (2.105) and (2.106), we have *Q* and *R* matrices similar to obtained in (2.101).

By the minimization of cost function (2.106), the infinite horizon optimal control solution for dynamic system (2.89) is obtained by the solution of ARE

$$A_w^T P_w + P_w A_w - P_w F R^{-1} F^T P_w + Q = 0 \tag{2.107}$$

where P_w is a positive definite real symmetric constant matrix.

The LQR gain matrix K_w is obtained as

$$K_w = R^{-1}F^T P_w \tag{2.108}$$

Thus, the optimal control law is obtained as

$$w = -K_w x \tag{2.109}$$

Substituting the feedback control law (2.109) into (2.89), we have the asymptotically stable closed loop system

$$\dot{x} = (A_w - FK_w)x \tag{2.110}$$

2.6.3 Simulation Results and Analysis

The system parameters of two-area interconnected power system having identical generating units considered in simulation are [184]:

$f^0 = 50\,\text{Hz}, \quad R_1 = R_2 = 2.4\,\text{Hz/puMW}, \quad b_1 = b_2 = 0.425, \quad T_{G_1} = T_{G_2} = 0.08\,\text{s}, \quad T_{T_1} = T_{T_2} = 0.3\,\text{s},$

$T_{P_1} = T_{P_2} = 20\,\text{s}, \; K_{P_1} = K_{P_2} = 120\,\text{Hz/puMW}, \; T_{12} = 0.08674, \; a_{12} = -a_{21} = 1, K_{I_1} = K_{I_2} = 0.75$

For the dynamic system defined by (2.86) and (2.90) the system matrix *A*, input matrix *B*, load disturbance matrix *F*, output matrix *C*, and direct transmission matrix *D* used in control design and analysis are computed as below:

$$A = \begin{bmatrix} -0.05 & 6 & 0 & 0 & 0 & 0 & -6 & 0 & 0 \\ 0 & -3.3333 & 3.3333 & 0 & 0 & 0 & 0 & 0 & 0 \\ -5.2083 & 0 & -12.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.05 & 6 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3.3333 & 3.3333 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5.2083 & 0 & -12.5 & 0 & 0 & 0 \\ 0.545 & 0 & 0 & -0.545 & 0 & 0 & 0 & 0 & 0 \\ 0.425 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.425 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 12.5 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 12.5 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} -6 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -6 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

For the dynamic system defined by (2.89) and (2.90) for two-area power system AGC using the integral control the system matrix A_w is computed as below:

$$A_a = \begin{bmatrix} -0.05 & 6 & 0 & 0 & 0 & 0 & -6 & 0 & 0 \\ 0 & -3.3333 & 3.3333 & 0 & 0 & 0 & 0 & 0 & 0 \\ -5.2083 & 0 & -12.5 & 0 & 0 & 0 & 0 & -9.375 & 0 \\ 0 & 0 & 0 & -0.05 & 6 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3.3333 & 3.3333 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5.2083 & 0 & -12.5 & 0 & 0 & -9.375 \\ 0.545 & 0 & 0 & -0.545 & 0 & 0 & 0 & 0 & 0 \\ 0.425 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.425 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

The step responses of integral control are shown in Fig. 2.32 and Fig. 2.33. Fig. 2.32 shows the step responses of area frequency deviations and tie-line power deviation for a 1% load disturbance in area 1. Fig. 2.33 shows the step responses of area frequency deviations and tie-line power deviation for a 2% load disturbance in area 2. It is observed that deviations in area frequencies and tie-line power converge to zero slowly, and stabilizes the system by maintaining the balance of power generation and load.

For the two-area power system defined by (2.89) and (2.90) the optimal AGC is designed using LQR. With the choice of

$$Q = \begin{bmatrix} 0.1806 & 0 & 0 & 0 & 0 & 0 & 0.425 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1806 & 0 & 0 & -0.425 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0.425 & 0 & 0 & -0.425 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

the LQR gain matrix K is obtained as

$$K = \begin{bmatrix} 0.2097 & 0.5196 & 0.2690 & -0.0202 & -0.0392 & -0.0082 & -0.2482 & 0.5774 & 0.0000 \\ -0.0202 & -0.0392 & -0.0082 & 0.2097 & 0.5196 & 0.2690 & 0.2482 & 0.0000 & 0.5774 \end{bmatrix}$$

The eigen values of closed loop system matrix (A-BK) are obtained as

$-14.8871,\quad -14.8721,\quad -0.9173 + j3.2912,\quad -0.9173 - j3.2912,\quad -1.8789 + j2.1780,$

$-1.8789 - j2.1780,\ -2.2371,\ -0.4980,\ -0.4051$

The step responses of optimal control using LQR are shown in Fig. 2.34 and Fig. 2.35. Fig. 2.34 shows the step responses of area frequency deviations and tie-line power deviation for a 1% load disturbance in area 1. Fig. 2.35 shows the step responses of area frequency deviations and tie-line power deviation for a 2% load disturbance in area 2. It is observed that

deviations in area frequencies and tie-line power converge to zero faster than integral control, and stabilizes the system by maintaining the balance of power generation and load.

For the dynamic system defined by (2.89) and (2.90) of two-area power system with integral control the optimal AGC is designed using LQR with the same choice of quadratic cost function matrices Q and R.

The LQR gain matrix K_w obtained for this case is

$$K_w = \begin{bmatrix} -0.3738 & -0.2640 & -0.0147 & 0.0410 & 0.0112 & -0.0002 & -0.1643 & -0.5863 & 0.0083 \\ 0.0410 & 0.0112 & -0.0002 & -0.3738 & -0.2640 & -0.0147 & 0.1643 & 0.0083 & -0.5863 \end{bmatrix}$$

The eigen values of closed loop system matrix (A_w-FK) are obtained as

$$-13.1987,\quad -13.1869,\quad -1.4745+j3.4178,\quad -1.4745-j3.4178,\quad -1.8787+j2.5088,$$
$$-1.8787-j2.5088,\ -1.1196+j0.6543,\ -1.1196-j0.6543,\ -0.9238$$

The step responses of optimal control using integral controller and LQR are shown in Fig. 2.36 and Fig. 2.37. Fig. 2.36 shows the step responses of area frequency deviations and tie-line power deviation for a 1% load disturbance in area 1. Fig. 2.37 shows the step responses of area frequency deviations and tie-line power deviation for a 2% load disturbance in area 2. For simulation of responses for the case of load disturbances in both areas simultaneously, the SIMULINK model of optimal AGC of two-area interconnected power system using integral controller and LQR is shown in Fig. 2.38. Fig. 2.39 shows the step responses of area frequency deviations and tie-line power deviation for load disturbances in both areas as 1% in area 1 and 2% in area 2 respectively. It is observed that deviations in area frequencies and tie-line power converge to zero fast. The step responses obtained in this case are smoother and faster than both cases of integral control and optimal control using LQR. The optimal control using integral controller and LQR stabilizes the system by maintaining the balance of power generation and load.

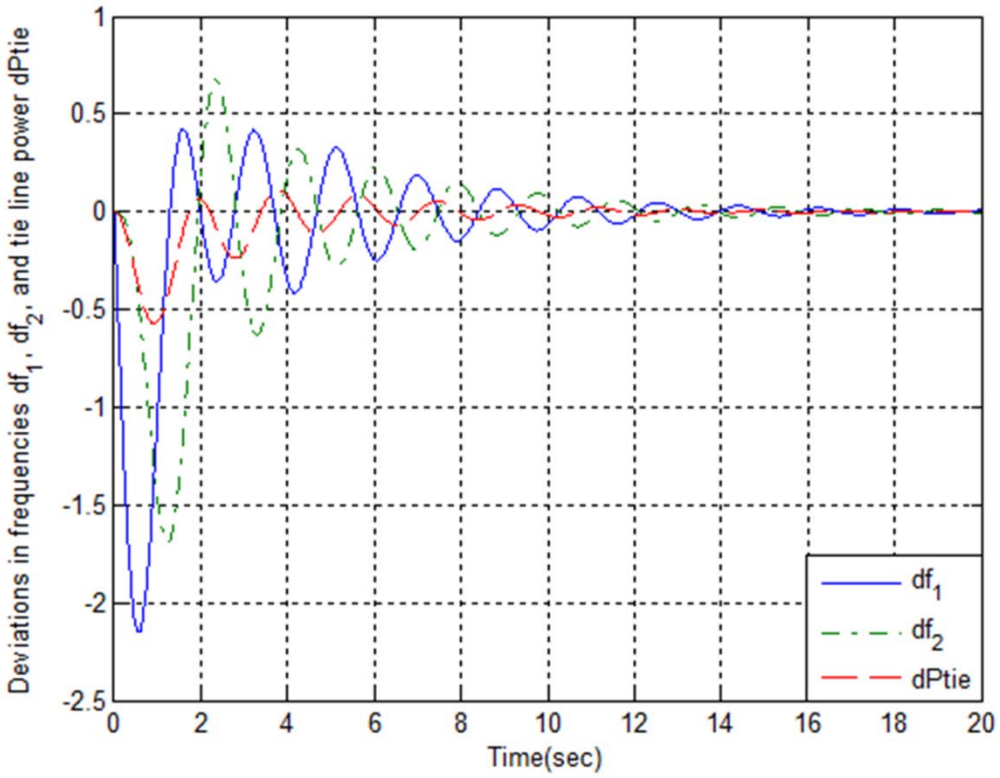


Fig. 2.32 Step responses of deviations in area frequencies and tie-line power for a 1% load disturbance in area 1 using integral control.

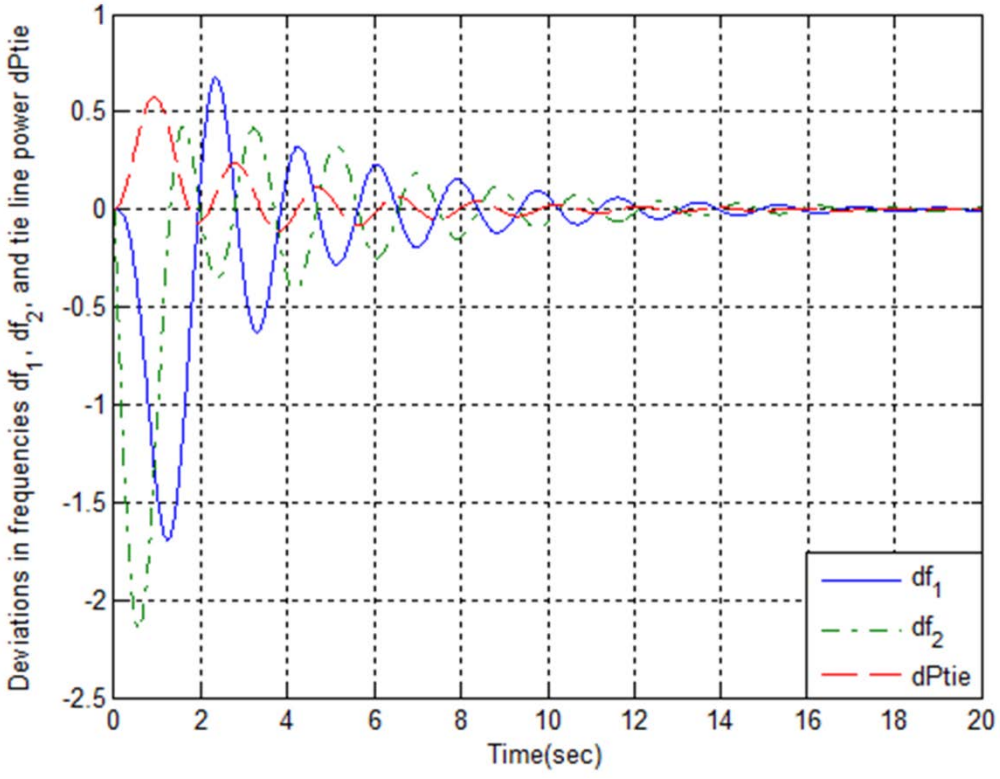


Fig. 2.33 Step responses of deviations in area frequencies and tie-line power for a 2% load disturbance in area 2 using integral control.

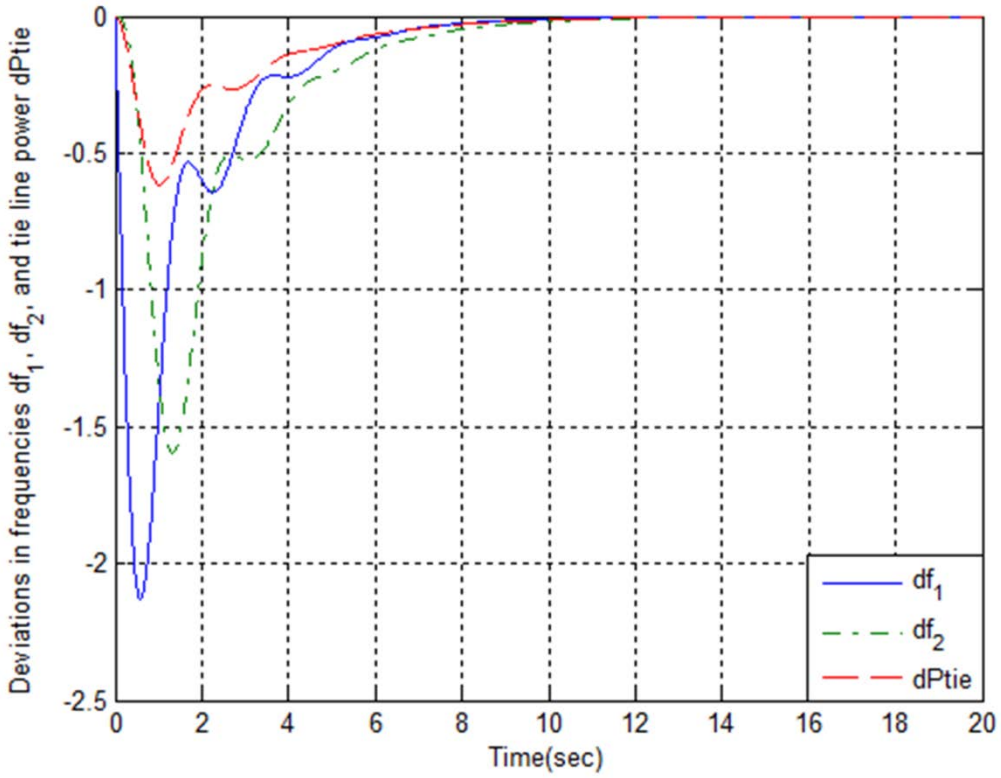


Fig. 2.34 Step responses of deviations in area frequencies and tie-line power for a 1% load disturbance in area 1 using LQR.

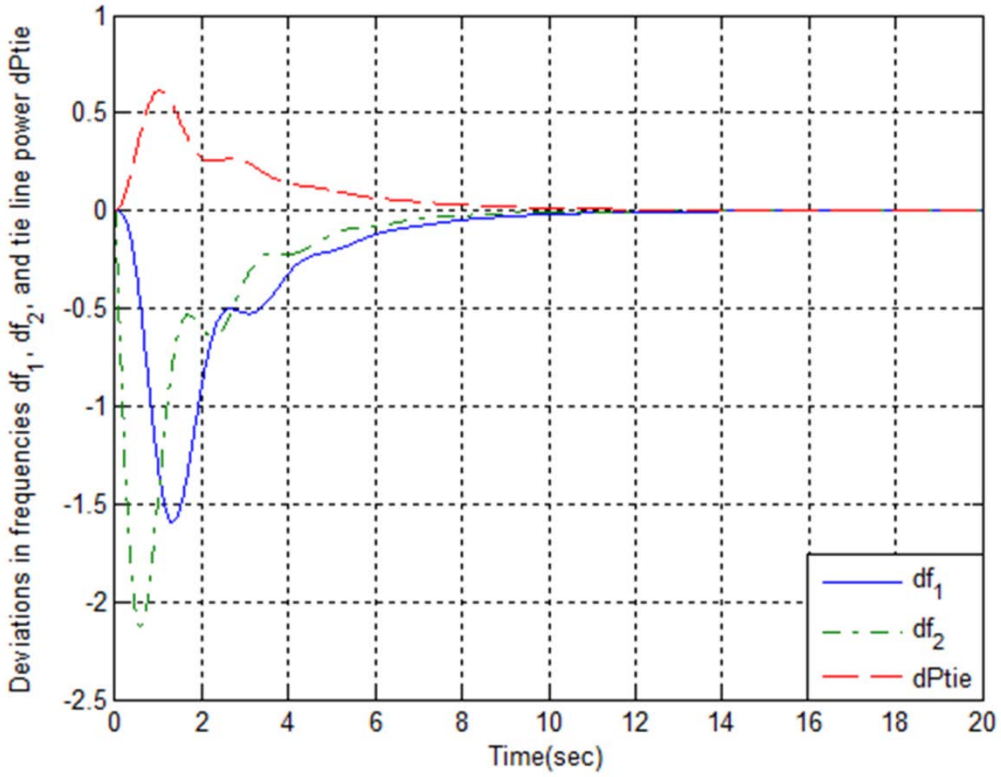


Fig. 2.35 Step responses of deviations in area frequencies and tie-line power for a 2% load disturbance in area 2 using LQR.

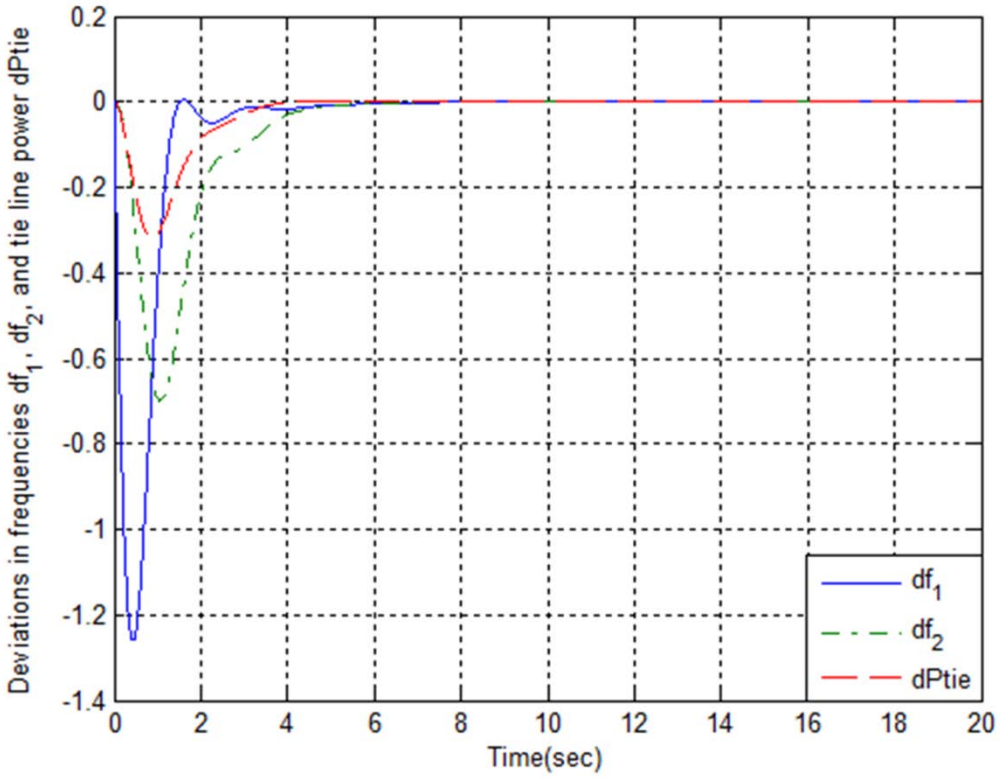


Fig. 2.36 Step responses of deviations in area frequencies and tie-line power for a 1% load disturbance in area 1 using integral controller and LQR.

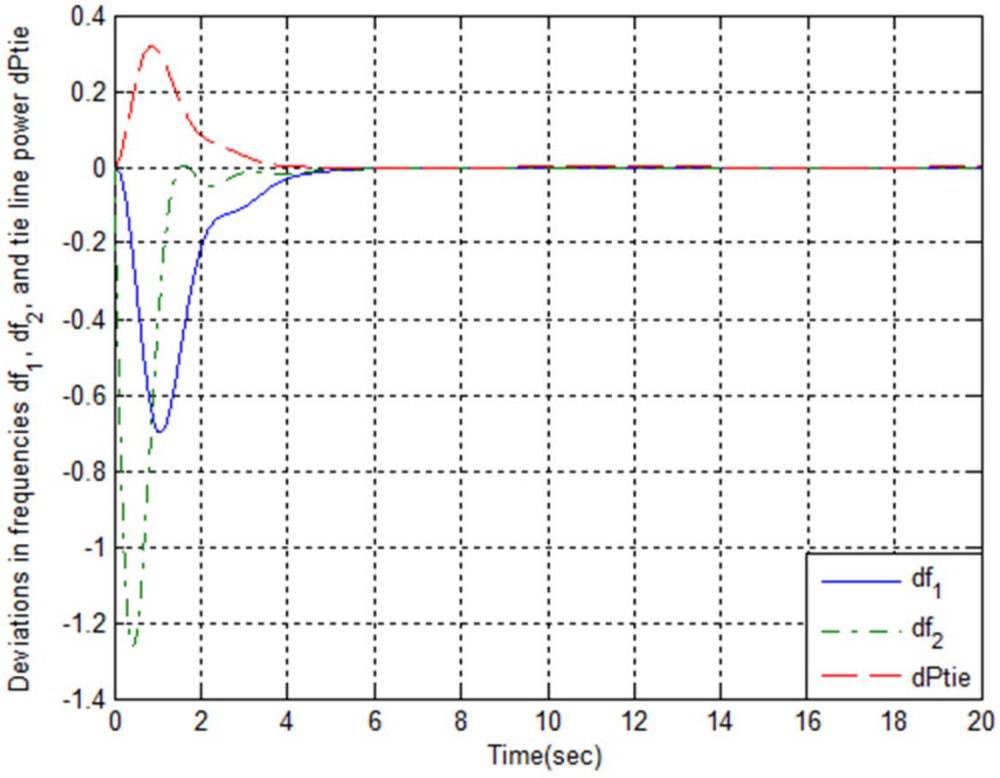


Fig. 2.37 Step responses of deviations in area frequencies and tie-line power for a 2% load disturbance in area 2 using integral controller and LQR.

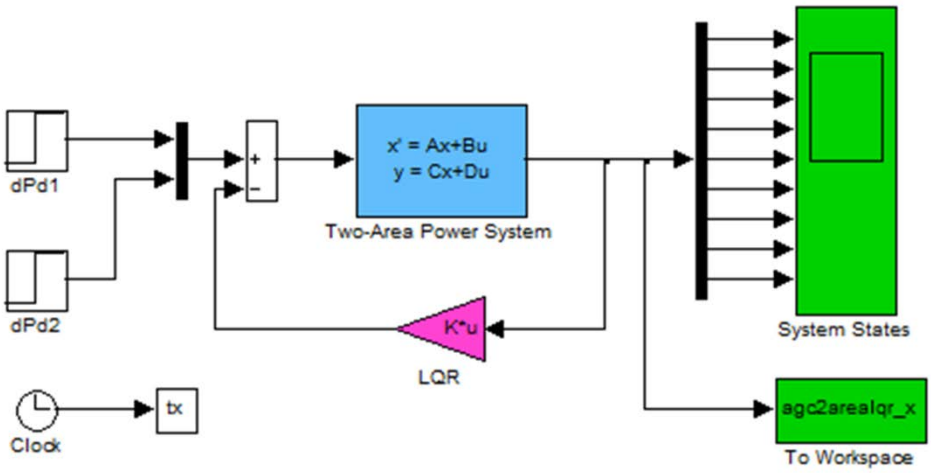


Fig. 2.38 Optimal AGC of two-area interconnected power system using integral controller and LQR.

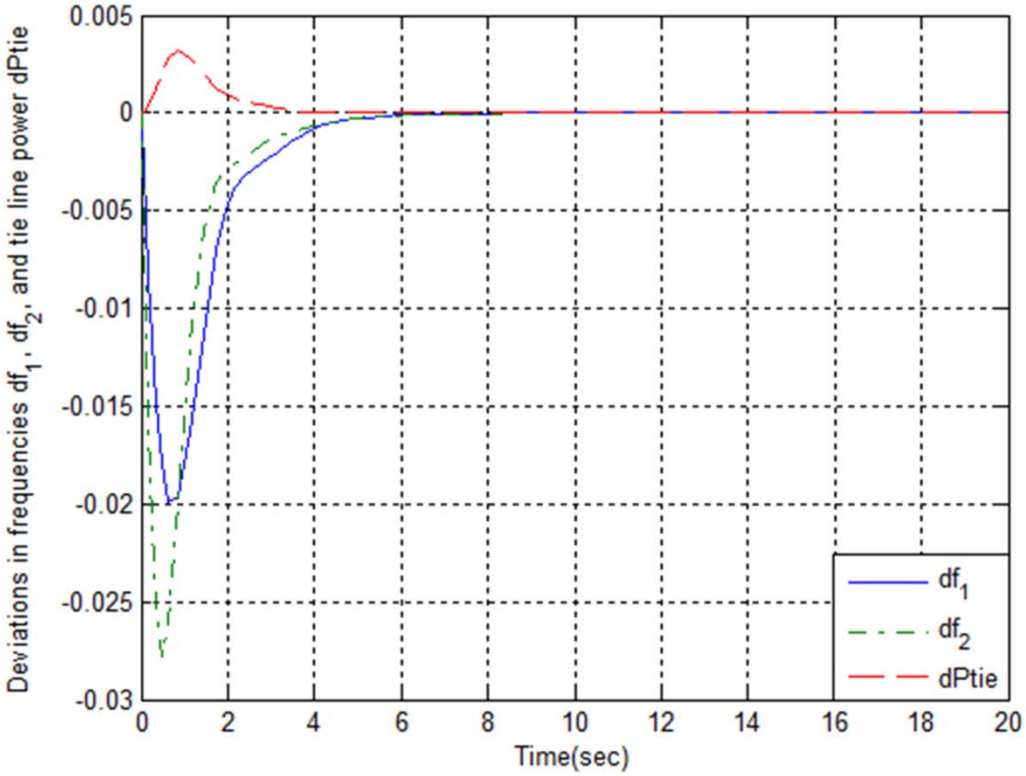


Fig. 2.39 Step responses of deviations in area frequencies and tie-line power for simultaneous 1% load disturbance in area 1 and 2% load disturbance in area 2 using integral controller and LQR.

2.6.4 Discussion

The control design and performance analysis of automatic generation control system for two-area interconnected power systems are presented. The control schemes of AGC system presented are conventional integral control, optimal control, and optimal control using integral controller and LQR. The optimal control design using integral controller and LQR

gives an optimal integral controller for AGC system. The LQR provides an infinite horizon optimal control solution for AGC system. The simulation results and analysis are presented for each of the cases of these control schemes. The step responses obtained in the case of optimal control using integral controller and LQR are smoother and faster than both the cases of integral control and optimal control using LQR. The comparative analysis of simulation results justify the effectiveness and robustness of optimal control using integral controller and LQR for automatic generation control of two-area interconnected power systems.

2.7 CONCLUSIONS

The optimal and adaptive control schemes of dynamical systems with applications are discussed in this chapter. The performance analyses of control schemes for applications to linear and nonlinear dynamical systems are presented. The indirect adaptive control and direct adaptive control for first order dynamical systems are discussed. The optimal control of nonlinear inverted pendulum dynamical system using PID controller and LQR is presented. The performance analyses of control schemes of PID control, 2PID+LQR control, 1PID+LQR control for nonlinear inverted pendulum-cart system for both cases of system models without and with disturbance input are presented. It is observed that PID+LQR control scheme provides optimal control of nonlinear inverted pendulum dynamical system, the responses of which are effective and robust. The optimal control of two-area interconnected power system for AGC problem is also presented. The performance analyses of control schemes of conventional integral control, optimal control using LQR, and optimal control using integral controller and LQR for AGC system of two-area interconnected power system are presented. It is observed that the integral controller and LQR control scheme provides the optimal control of AGC system. The infinite horizon optimal control solution using LQR design provides the effective and robust performance for dynamical systems. Thus, it is observed from the simulation results that the combined control schemes of PID+LQR provide better responses for dynamical systems control. These optimal control techniques require the system model to be completely known. The optimal control design schemes give the offline solution for the control problems. The adaptive control schemes require a priori knowledge of the system dynamics, and give online solution for the control problems.

CHAPTER 3

INTELLIGENT CONTROL OF DYNAMICAL SYSTEMS

This chapter discusses the intelligent control schemes of dynamical systems using neural networks and fuzzy systems. The theoretical background on the intelligent control schemes and their applications are briefly described for completeness of the topic under discussion. The performance analysis of neural network control scheme and fuzzy control scheme are presented for linear and nonlinear dynamical systems applications.

3.1 INTRODUCTION

The need for higher degree of autonomy of control systems for high-performance real-world applications under significant unanticipated uncertainties in the system and environment has led the need of alternative novel intelligent control methodologies. The intelligent control methodologies can give better control solution for dynamical systems under uncertainties than the available classical control methodologies. Emulating certain characteristics of intelligent biological systems the intelligent computational techniques such as neural networks, fuzzy logic, evolutionary computation, and machine learning have been developed. The intelligent control has emerged from the integration of automatic control design with intelligent computational techniques which have given novel solutions to the various control system problems.

Recently several researchers have tried to explore the intelligent computational techniques synergistically with various control design techniques such as PID control, optimal control, and adaptive control techniques by applying certain methodologies for certain applications [13, 14, 16, 29, 32, 35, 48, 49, 69, 71-73, 87, 107, 109, 116, 119]. Neural network control and fuzzy control have played a major role in the development of intelligent control schemes for various dynamical systems applications. The basic concepts of intelligent control and intelligent control techniques are briefly introduced in the following section. The implementation of neural network control and fuzzy control for certain dynamical systems control applications are presented in the following sections of this chapter respectively.

3.2 INTELLIGENT CONTROL

The intelligent control pertains to a control technique that possesses some sort of intelligence. The term “intelligent control” is loosely used to denote a control technique that can be accomplished using the “intelligence” of a human having the knowledge in the particular domain of control. If a system can be properly controlled by a human in the control loop, then that system would be a good candidate for intelligent control. [72]. A control methodology is an intelligent control methodology if it uses human-, animal-, or biologically

motivated techniques and procedures (e.g., forms of representation or decision making) to develop or implement a controller for a dynamical system. The intelligent control methodologies include for example, the neural control methodology which is motivated by low-level biological representation and decision making; the fuzzy control methodology which includes (1) the use of fuzzy sets and fuzzy logic for rule-based representation of a human's knowledge about how to control, (2) fuzzy inference for modelling human deductive processes, and (3) conventional or fuzzy processors for implementation; the biologically motivated genetic algorithms; learning control which incorporates learning theories into controllers; expert control which uses a rule-based expert system, etc. Many intelligent control methodologies result from the synthesis of several intelligent and conventional control methodologies. Thus, the intelligent controller can be defined as the physical device called a controller if is developed or implemented with (1) an intelligent control methodology or (2) conventional systems or control techniques to emulate or perform control functions that are normally performed by humans, animals, or biological systems. The control system that can overcome with three categories of difficulties in the control of complex systems as of computational complexity, nonlinearity, and uncertainty, qualify to be called intelligent control system. The control system will be qualitatively more intelligent as more it has the ability to deal with these difficulties [69]. The information abstraction and knowledge-based decision making that incorporates abstracted information are important in intelligent control. Unlike conventional control, intelligent control techniques possess capabilities of effectively dealing with incomplete information concerning the plant and its environment, and unexpected or unfamiliar conditions [72]. The intelligent controller can adapt its action to change its parameters under the conditions of changes in the system dynamics. In a control engineering perspective, it is not necessary to design every automatic control loop with intelligent methodology. The industrial processes have a hierarchy of control loops in which the topmost level has intelligent controllers whereas the lowest level has the fast acting, simplest and most robust controllers.

The area of intelligent control has been emerged from the integration of control methodologies with intelligent computational techniques. The computational procedures of neural networks (NNs), fuzzy logic (FL), and genetic algorithms (GA) belong to the class of "soft computing" techniques, which can be directly utilized in intelligent control, either separately or synergistically [72]. Fuzzy control has an impact in the control community due to the simple approach it provides to use heuristic control knowledge for nonlinear control problems. However, in the more complicated situations where the plant parameters are subject to perturbations, or when the dynamics of the system are too complex to be characterized reliably by an explicit mathematical model, the adaptive schemes are included for online operation and use adaptation heuristics to automatically determine the parameters

of the controller. Neural networks that make use of the organizational principles of human brains are widely known for the powerful abilities, such as learning and adaptation capabilities, fault tolerance, parallelism and generalization. The most useful property of neural networks is their ability to approximate any continuous function to a desired degree of accuracy through learning. The main issue in the neural network based control approaches is the choice of a satisfying neural network structure and tuning algorithm to improve the system performance. Genetic algorithm is the search & optimization approach, which is based on the principle of evolutional process of natural genetics and natural selection of the fittest members of a given population to breed the next generation.

The synergistic combination of terms of “intelligent control”, “optimal control”, and “adaptive control” have led to the synthesis of control schemes of “intelligent optimal control”, “intelligent adaptive control”, and “intelligent adaptive optimal control”.

3.2.1 Intelligent Control Systems

As per the Oxford dictionary, the word intelligence is derived from intellect, which is the faculty of knowing, reasoning and understanding. Thus, the intelligent behaviour is the ability to reason, plan and learn, which in turn requires access to knowledge. The intelligence in human beings is a creation of nature. The human beings learn from nature and try to imitate the process of cognition and intelligence into machines. The revolution in artificial intelligence (AI) is an attempt to replace the human intelligence with machine intelligence. The intelligent control system is the combination of AI techniques with those of control techniques to design autonomous systems that can sense, reason, plan, learn, and act in an intelligent manner. The intelligent control system should be able to achieve sustained desired behaviour under conditions of uncertainty such as uncertainty in plant models, unpredictable environmental changes, incomplete, inconsistent or unreliable sensor information, and actuator malfunctions [43].

As considered by Johnson and Picton (1995), the structure of an intelligent control system consists of following subsystems as shown in Fig. 3.1 [43].

1. Perception subsystem: This collects information from the plant and the environment, and processes it into a form suitable for the cognition subsystem. The essential elements of perception subsystem are: (a) sensor array which provides raw data about the plant and the environment, (b) signal processing which transforms information into a suitable form, and (c) data fusion which uses multidimensional data spaces to build representations of the plant and its environment. Here a key technology is pattern recognition.
2. Cognition subsystem: The cognition in an intelligent control system is concerned with the decision making process under conditions of uncertainty. The key activities of cognition subsystem are: (a) reasoning using (i) knowledge-based systems, and (ii)

fuzzy logic; (b) strategic planning using (i) optimum policy evaluation, (ii) adaptive search and genetic algorithms, and (iii) path planning; and (c) learning using (i) supervised learning in neural networks, (ii) unsupervised learning in neural networks, and (iii) adaptive learning.

3. Actuation subsystem: The actuators operate using signals received from the cognition subsystem in order to drive the plant to some desired states. In the event of actuator (or sensor) failure, an intelligent control system should be capable of being able to reconfigure its control strategy.

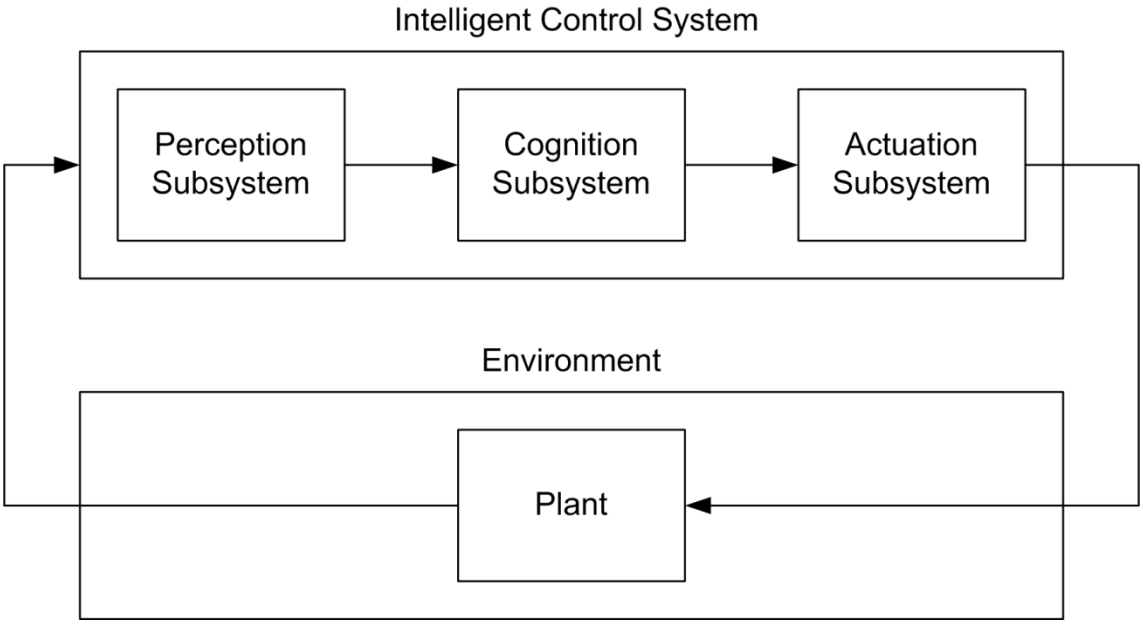


Fig. 3.1 Intelligent Control System Structure.

3.2.2 Intelligent Control Techniques

Motivated from intelligent biological systems emulating certain characteristics, the recent development in the area of computational intelligence (CI), a branch of artificial intelligence (AI), has led the development of different intelligent control paradigms. The intelligent computational techniques include artificial neural network (ANN), fuzzy logic theory (FL), evolutionary computational techniques such as genetic algorithm (GA), and swarm intelligence (SI) etc., machine learning such as support vector machine (SVM), and knowledge-based systems (KBS) or expert systems (ES) etc. The intelligent computational techniques have been integrated with control techniques resulting in intelligent control techniques which are classified based on the corresponding computational paradigms. The intelligent control techniques have given promising solutions for various control applications [7, 9, 10, 13-16, 29-43, 48, 49, 61-79, 81-99].

The neural network control techniques are developed using the concepts of the artificial neural network (ANN), often called the neural network (NN), which is the most generic form of AI for emulating the human thinking process. The human brain is comprised

of approximately 10 billion individual nerve cells known as biological neurons. The brain has various sensory and motor neurons. Each neuron is interconnected to many other neurons, forming a densely interconnected network called neural network. These massive interconnections provide an exceptionally large computing power and memory, and thus the capacity to remember, think, learn and reason. The neuron is the basic building block of central nervous system that processes and communicates information to and from various parts of the body. From information processing point of view a biological neuron has three major functional regions: the cell body, called soma, the axon, and the dendrites. The dendrites receive information from other neurons. The soma collects and combines the incoming information from other neurons. The axon transmits information to other neurons. The axon is a long fiber having tubular structure bounded by a typical cell membrane, serves as transmission line. The junction of an axon with dendrite of another neuron is called a synapse. The synapses provide memory for accumulating experience or knowledge. A single axon may be involved with hundreds of other synaptic connections. The information processing within the biological neuron involves two distinct operations: synaptic operation, and somatic operation. Synaptic operation provides a weight to the neural inputs. According to the past experience (knowledge or memory) stored in the synapse, the synaptic operation assigns a relative weight (significance) to each incoming signal. The somatic operation provides the aggregation, thresholding, and nonlinear activation to the inputs of dendrites. The soma produces an output signal if the weighted aggregation of neural inputs exceeds a certain threshold [65, 72]. The ANN tends to emulate the biological nervous system of the human brain in a very limited way by an electronic circuit or computer program. McCulloch and Pitts (1943) proposed a simple model of a neuron called artificial neuron [65]. Artificial neurons are only a modest resemblance to real things. They model approximately three of the processes that biological neurons perform out of at least 150 processes which are performed by biological neurons in the human brain. These functions of artificial neuron are: evaluation of input signals determining the strength of each one; calculation of sum of input signals and its comparison with threshold value; and determination of output signal using an activation function [7]. However inferior the ANN model of the biological nervous system, it tends to solve many important problems. The ANN has potential advantages for intelligent control, which are: ability to adapt and learn from the environment and can generalize from given training data to unseen data; ability to fault-tolerance due to its massive parallel structure; ability to approximate any nonlinear continuous function to a desired degree of accuracy; applicability to multivariable systems; ability to implement in real-time; and can be implemented using VLSI hardware which will further increase speed to neural computing [43, 72].

Basically, the neural network architecture is of two types: (i) feedforward, and (ii) feedback (recurrent) networks, and neural network learning algorithm is of two types: (i) supervised learning, and (ii) unsupervised learning. The back-propagation (BP) is the most popular training method for a multi-layer feedforward network; therefore, the standard network trained by this algorithm is often called the BP network. Rumelhart, Hinton, and Williams proposed the BP training method, although other researchers made contributions to it independently. Basically, it is a generalization of the delta learning rule developed by Widrow and Hoft (1960) for Adaline training. Hebb (1949) described a technique which became known as 'Hebbian' learning. Rosenblatt (1961), devised a single layer of neurons, called perceptron that was used for optical pattern recognition. Associative memories are faint imitations of the human brain's ability to associate patterns. An associative memory which belongs to the class of single layer feedforward or recurrent network architecture depending on its association capability, exhibits Hebbian learning. One of the first applications of this technology for control purposes was by Widrow and Smith (1964) who developed an ADaptive LINear Element (ADLINE) to stabilize and control an inverted pendulum. Kohonen (1988) and Anderson (1972) investigated similar areas, looking into 'associative' and 'interactive' memory, and also 'competitive learning'. The back propagation training algorithm was investigated by Werbos (1974) and further developed by Rumelhart (1986) and others, leading to the concept of multi-layer perceptrons (MLP) [43, 66, 72]. The neural networks literature is well developed. The detailed discussion on neural networks theory and applications can be referred in literature [65, 66]. Neural networks have been applied for various control, identification, and estimation applications [7, 10, 13-15, 43, 65, 66, 69-74, 76, 78, 79, 82, 87, 98].

The fuzzy control techniques are developed using the fuzzy logic (FL) theory emulating the human knowledge representation and reasoning process. The fuzzy logic theory has provided a mathematical strength to capture the uncertainties associated with human cognitive processes, such as thinking and reasoning. Fuzzy logic provides an inference morphology that enables approximate human reasoning capabilities to be applied to knowledge-based systems. The fuzzy logic was first proposed by Lotfi A. Zadeh (1965). The fuzzy logic systems can address the imprecision or vagueness in input-output descriptions of systems using fuzzy sets. The concept of fuzzy sets arose as an answer to the problems of paradoxes, uncertainties, and imprecision in real-world data, which could not be represented by crisp sets. Zadeh classified computing into “hard computing”, i.e. precise computation, and “soft computing”. Soft computing techniques such as fuzzy logic, neural network, genetic algorithms, probabilistic reasoning, chaos theory, and learning theory etc, are inspired by biological computational processes and nature’s problem solving strategies. [16, 43, 66, 72]. Fuzzy set theory provides a means for representing uncertainty. In general,

probability theory is the primary tool for analyzing uncertainty, which assumes that the uncertainty is a random process. However, not all uncertainty is random, and fuzzy set theory is used to model the kind of uncertainty associated with imprecision, vagueness and lack of information. Conventional set theory having sharp or crisp boundaries, distinguishes between those elements that are members of a set and those that are not. Fuzzy sets theory relates to classes of objects with un-sharp boundaries in which membership is a matter of degree. The degree of belongingness in the fuzzy sets is defined by the term “membership function”. The central concept of fuzzy set theory is that the membership function can have any value between 0 and 1. A crisp set is defined by the characteristic function that can have only two values {0,1}. A fuzzy set is defined by a membership function that can have an infinite number of values; any real number in the closed interval [0,1]. A fuzzy variable has values that are expressed by the linguistic values. A membership function (MF) is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. The input space is sometimes referred to as the universe of discourse.

The fuzzy logic control (FLC) system has four functional units: the fuzzification process, fuzzy rule-base, fuzzy inference mechanism, and defuzzification process. The fuzzification process is concerned with the mapping of crisp inputs to fuzzy inputs. The fuzzifier converts the crisp input to a linguistic variable using the membership functions and universes of discourse. The fuzzifier takes decisions about number of inputs, discretization or normalization of universe of discourse, fuzzy partition of input and output spaces, and choice of membership function of a primary fuzzy set. The common membership functions used in fuzzy control systems are: triangular, trapezoid, and Gaussian. The knowledge-base consists of information of linguistic variable definitions (data-base) and fuzzy rule-base (control-base). The fuzzy rule-base is a collection of antecedent-consequent linguistic rules in the form of ‘if-situation-then-action’, where both situation and action have suitable fuzzy representation. The fuzzy rules describe the control strategy which depends on the choice of process state (input) variables and control (output) variables. The rule-base is generated using a priori knowledge of either one or all of the sources which are: physical laws that govern the plant dynamics, data from existing controllers, and imprecise heuristic knowledge obtained from experienced experts. The fuzzy inference mechanism establishes a logical connection between input and output fuzzy sets. The fuzzy inference engine converts the fuzzy input to the fuzzy output using fuzzy rules. The defuzzification process is concerned with the mapping of inferred fuzzy output to non-fuzzy (crisp) output. The defuzzifier converts the fuzzy output to crisp output using membership functions. The various defuzzification methods are: centroid or centre of area (COA) or centre of gravity (COG), centre of average or weighted average, mean-max, and centre of sums. The COG is the most popular

defuzzification methods. Based on fuzzy inference systems used, the fuzzy logic control systems are basically classified as: Mamdani type FLC and Takagi-Sugeno type FLC. The FLC was first implemented by Mamdani and Assilian (1975). The fuzzy control literature is well developed. The detailed discussion on fuzzy logic control theory and applications can be referred in literature [66, 67]. The fuzzy control systems have been successfully implemented for various industrial control applications [66, 67].

The intelligent control with optimization techniques using evolutionary computational techniques are developed emulating the process of evolution in biological species and natural processes. The evolutionary computing strategies were introduced by Rechenberg (1960). The evolutionary computational algorithms include: genetic algorithms (Holland, 1975), simulated annealing (Kirkpatrik, et.al. 1983), cellular automata (Wolfram, 1983), evolution strategy (Kost, 1995), particle swarm optimization (Ebarhart and Kennedy, 1995), ant colony optimization (Dorigo and Caro, 1999), and random cost (Kost and Baumann, 1999). GA is an adaptive search and optimization algorithms. GA and evolutionary strategies are based on the principle of evolutional process of natural genetics and natural selection of the fittest members of a given population to breed the next generation. The simulated annealing mimics the cooling phenomenon of molten metals to constitute a search procedure. The collective behaviour that emerges from a group of social insects such as ants, bees, wasps, and termites has been known as swarm intelligence. The particle swarm optimization is a population based search algorithm mimicking the social behaviour of birds, bees or a school of fishes. The foraging of ants has led to a novel algorithm called ant colony optimization. Random cost method is a stochastic algorithm which moves as enthusiastically uphill as downhill. This method has no severe problems in escaping from a dead end and is able to find the optima. A cellular automaton is a discrete dynamical system that model complex behaviour of cells on a lattice structure. [43, 66, 72].

The intelligent control techniques are developed using hybrid intelligent systems. The hybrid intelligent control systems have given promising solutions for a wide variety of real-world control problems [15, 16, 66, 92, 100]. The hybrid intelligent systems are developed by synergistic combinations of fuzzy systems, neural networks, genetic algorithms, and expert systems etc. Every intelligent system has particular computational properties (e.g. ability to learn, explanation of decisions) that make them suited for particular problems and not for others. Each of them has their own merits and demerits. The synergistic combination of different intelligent systems containing properties of each of them makes them functionally more suitable for a wider class of applications with better performance on cost of increase in computational complexity. The hybrid intelligent systems commonly are: hybrid neuro-fuzzy systems, hybrid neuro-GA systems, hybrid fuzzy-GA systems, and hybrid neuro-fuzzy-GA systems. The adaptive neuro-fuzzy inference system (ANFIS) is one of the methods to

combine fuzzy logic and artificial neural networks. The ANFIS structure can be tuned automatically by a least-square estimation (for output membership functions) and a back-propagation algorithm (for output and input membership functions). The combined fusion of neural networks, fuzzy logic, and genetic algorithms technologies developed to explore the novel solutions of various problems.

3.3 INTELLIGENT CONTROL OF NONLINEAR PROCESS SYSTEM USING RADIAL BASIS FUNCTION NEURAL NETWORKS

The artificial neural networks (ANN) possessing the capabilities of generalization, function approximation, learning and adaptation etc. have played a significant role in the development of intelligent control systems. The neural networks (NN) are widely used for system identification and control applications. The feed-forward network is extensively used in intelligent control systems. Although there are many variants of feed-forward networks, multi-layer perceptrons (MLP) networks and radial basis function (RBF) networks have been used in many applications. RBF networks have gained considerable attention as an alternative to MLP trained by backpropagation algorithm. The features of simple architecture and linear response of radial basis function neural network make it more suitable for intelligent control applications. The complex industrial processes involve the automatic control system for processes. The automation of certain industrial processes needs the advanced controllers as the microcontroller-based liquid level control systems. Various advance control techniques for industrial processes are suggested in the process control literature. In this section the indirect adaptive control using radial basis function neural network for a nonlinear process system is presented. A surge tank liquid level control system is considered as an example of a nonlinear process system.

3.3.1 Mathematical Modelling of Surge Tank Process System

The industrial processes often involve flow of liquids through connecting pipes and tanks. The liquid level control of a surge tank system is an important part of the industrial processes. The flow of such processes is often turbulent and not laminar. Such systems are represented by nonlinear differential equations. However, these nonlinear differential equations can be linearized if the operation region is limited. In general, the surge tank process system has a nonlinear dynamics. The schematic diagram of a surge tank system is shown in Fig. 3.2 [71].

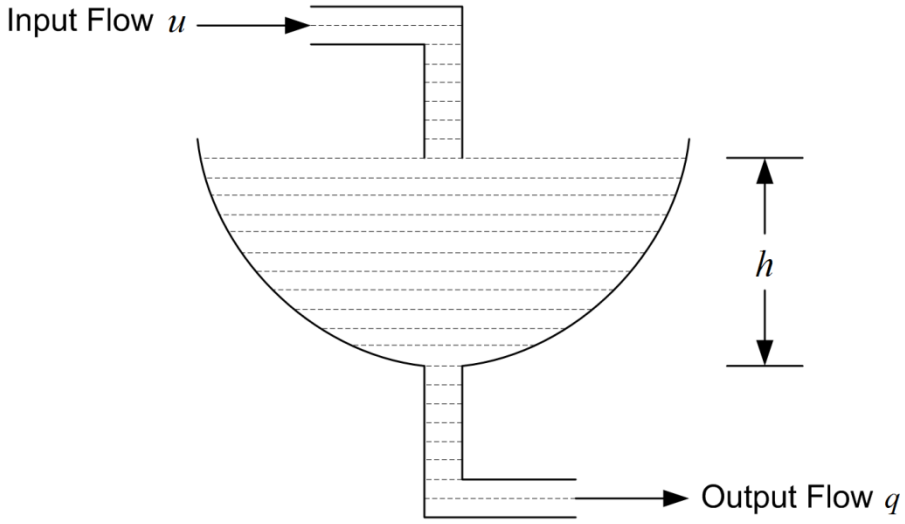


Fig. 3.2 Schematic diagram of a surge tank system.

Consider the surge tank is fed by the incompressible liquid with constant density of liquid at inlet and outlet. At steady-state the liquid level in the surge tank will be constant. The general liquid mass balance will give the dynamics of the surge tank liquid level system. The surge tank system dynamics is represented by the differential equation as

$$A(h(t))\frac{dh(t)}{dt} = u(t) - q(t) \quad \text{or}$$

$$\frac{dh(t)}{dt} = -\frac{\sqrt{2gh(t)}}{A(h(t))} + \frac{1}{A(h(t))}u(t) \tag{3.1}$$

where g is the acceleration due to gravity, $u(t)$ is the input flow in m^3/sec, $q(t)$ is the output flow in m^3/sec, $h(t)$ is the liquid level of the surge tank, the output of the system in m, and $A(h(t))$ is the cross-sectional area of tank in m^3, which is given by

$$A(h(t)) = \sqrt{ah^2(t) + b} \tag{3.2}$$

where a and b are parameters of tank. The input $u(t)$ can be both positive and negative, i.e. it can pull liquid out of the tank as well as fill it in the tank.

3.3.2 Neural Network Control Using Radial Basis Function Networks

The radial basis function neural network (RBFNN) has simple architecture with a single hidden layer, which is especially good in applications requiring locally tunable properties. This network uses radial basis functions to represent an input in terms of radial centres. Since the response of this network is linear in terms of weights, this network is very suitable for intelligent control applications. Thus, weight update rules for such networks within intelligent control paradigm become easy to derive through Lyapunov stability analysis. The RBFNN has been used for indirect adaptive control of a nonlinear affine system.

3.3.2.1 Radial Basis Function Neural Networks

The RBFNN architecture is shown in Fig. 3.3 [71]. The RBFNN consists of three layers: an input layer, a hidden layer and an output layer. The hidden units provide a set of functions that constitute an arbitrary basis for the input patterns.

1. Hidden units are known as radial centres. Each radial centres is represented by a vector $c_i,\ i=1,...,h$, where *h* is number of radial centres in the hidden layer.
2. The transformation from the input space to the hidden unit space is nonlinear whereas the transformation from the hidden unit space to the output space is linear.
3. Dimension of each centre for a *p* input network is *p*x1.

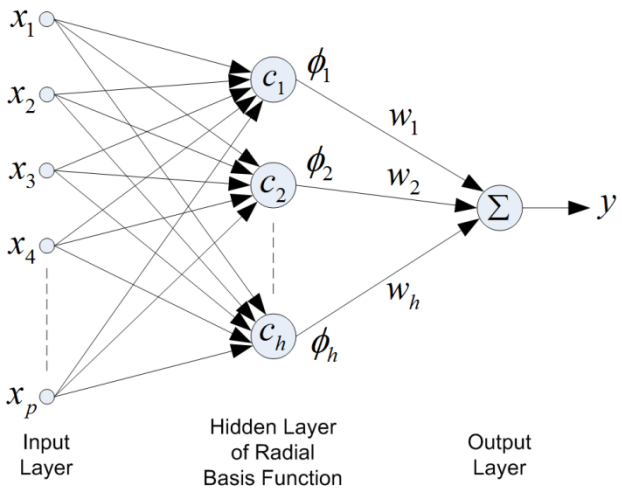


Fig. 3.3 Architecture of radial basis function neural network.

The RBFN uses radial basis functions (RBF) to represent an input in terms of radial centres. The RBF in hidden layer produces a significant non-zero response only when the input falls within a small localized region of the input space. Each hidden unit, known as radial centre, has its own receptive field in the input space, i.e. each centre is representative of one or some of the input patterns. This is called local representation of inputs, and the network is also known as localized receptive field network.

Consider for instance, an input vector *x* which lies in the receptive field for centre c_i. This would activate the hidden centre c_i and by a proper choice of weights, the target output is obtained. Suppose an input vector lies between two receptive field centres, then both those hidden units will be appreciably activated. The output will be a weighted average of the corresponding targets. The inputs are clustered around the centres and the output is linear in terms of weights w_i

$$y = \sum_{i=1}^{h} \phi_i w_i \tag{3.3}$$

The response of the i^{th} radial centre in RBFN usually expressed by the following expression:

$$\phi_i = \phi\left(\left\|x - c_i\right\|\right) \tag{3.4}$$

where $\left\|x - c_i\right\|$ is the Euclidean distance between *x* and c_i, and $\phi(.)$ is radial basis function. This justifies the name radial basis function. Function $\phi(.)$ can take various forms; the Gaussian form is more widely used. When RBF is Gaussian, each node produces an identical output for inputs within a fixed radial distance from the centre, i.e. they are radially symmetric.

A Gaussian basis function is typically parameterized by two parameters: the centre which defines its position, and a spread parameter that determines its shape. The spread parameter is equal to the standard deviation σ in case of a one-dimensional Gaussian function. In the case of a multivariable input vector x, the parameters that define the shape of the hyper-Gaussian function are elements of a covariance matrix Σ. With the selection of the same spread parameter σ for all components of the input vector, the covariance matrix $\Sigma = diag(\sigma^2)$.

The RBF (Gaussian) neuron is expressed by the following equation

$$\phi(x, c, \sigma) = \exp\left(-\frac{\left\|x - c\right\|^2}{2\sigma^2}\right) \tag{3.5}$$

where $x = [x_1, x_2, ...x_n]^T$ is the input vector, *c* is the centre and σ is the spread parameter of the Gaussian function. Unlike sigmoidal neuron, there are no connection weights between the input terminals and the RBF unit; the centre *c* and the spread parameter σ represent the weights.

In the RBF network, the output of each RBF node is the same for all input points *x* having the same Euclidean distance from the respective centres c_i, and decreases exponentially with the distance. Whereas, the output of each sigmoidal node in a MLP network saturates to the same value with increasing $\sum_i w_i x_i$. Thus, the activation responses of the nodes are of a local nature in the RBF and of a global nature in the MLP. This intrinsic difference has important repercussions for both the convergence speed and the generalization performance. In general, MLP learns slower than RBF. In contrast, MLP exhibits improved generalization properties, especially for the region that are not represented sufficiently in the training set.

The training of RBF network requires optimal selection of the centre c_i and weights w_i, $i = 1$ to h. This is a two-fold problem, unlike in a MLP network. As the different layers of network perform different tasks, both the layers are optimized using different techniques and

in different time scales. Several strategies are applicable depending on the way in which the radial centres are specified.

The RBFNN can be trained using gradient-descent algorithm. The gradient-descent algorithm is a supervised training by error correction learning to update the centres and weights of RBFNN. The update rule for centre learning is given as

$$c_{ij}(k+1) = c_{ij}(k) - \eta_1 \frac{\partial E(k)}{\partial c_{ij}(k)} \tag{3.6}$$

for $i = 1$ to *p* and $j = 1$ to *m*, where *k* is the sample instants.

Similarly the updates for weights are given as

$$w_i(k+1) = w_i(k) - \eta_2 \frac{\partial E(k)}{\partial w_i(k)} \tag{3.7}$$

where the cost function is

$$E(t) = \frac{1}{2}\sum (y_d - y)^2 . \tag{3.8}$$

where y is given by (3.3) with its desired value y_d, and radial basis function be taken as Gaussian given by (3.5). η_1 and η_2 are learning rates which determine the speed of convergence.

Differentiating *E* with respect to w_i yields

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial w_i} = -(y_d - y)\phi_i \tag{3.9}$$

Differentiating *E* with respect to c_{ij} yields

$$\frac{\partial E}{\partial c_{ij}} = \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial \phi_i} \times \frac{\partial \phi_i}{\partial c_{ij}} = -(y_d - y) \times w_i \times \frac{\partial \phi_i}{\partial z_i} \times \frac{\partial z_i}{\partial c_{ij}} \tag{3.10}$$

where $z = \left\| x - c_i \right\|$,

$$\frac{\partial \phi_i}{\partial z_i} = -\frac{z_i}{\sigma^2}\phi_i \tag{3.11}$$

$$\frac{\partial z_i}{\partial c_{ij}} = \frac{\partial}{\partial c_{ij}}\left(\sum (x_j - c_{ij})^2\right)^{1/2} = \frac{-(x_j - c_{ij})}{z_i} \tag{3.12}$$

After manipulation, the update rule for the centres is

$$c_{ij}(k+1) = c_{ij}(k) + \eta_1 (y_d - y) w_i \frac{\phi_i}{\sigma^2}(x_j - c_{ij}) \tag{3.13}$$

The update rule for the weights is

$$w_i(k+1) = w_i(k) + \eta_2 (y_d - y)\phi_i \tag{3.14}$$

The gradient-descent vector $\partial E/\partial c_{ij}$ exhibits a clustering effect. Unlike the supervised learning in a MLP network, there is no backpropagation of error in a RBF network.

3.3.2.2 Model Identification

Consider a continuous-time affine nonlinear system given by

$$\dot{x} = f(x) + g(x)u \tag{3.15}$$

where $x \in \mathcal{R}^n$ is the system state vector and $u \in \mathcal{R}^m$ is the input vector, $f(x)$ is *n*x1 vector, and $g(x)$ is an *n*x*m* matrix. It is assumed that the nonlinear functions $f(x)$ and $g(x)$ are unknown. Thus, it is required to perform system identification of (3.15) [71].

Before the parametric structure for the system identification can be selected, model (3.15) is expressed as

$$\dot{x} = -x + \overline{f}(x) + g(x)u \tag{3.16}$$

where $\overline{f}(x) = f(x) + x$. Let two RBF networks represent functions $\overline{f}$ and g, respectively. Then the neural network (NN) based model of the system (3.15) is given by

$$\dot{\hat{x}} = -\hat{x} + \hat{W}_1^T \phi_1(x) + \hat{W}_2^T \Phi_2(x)u \tag{3.17}$$

where $\hat{W}_1 \in \mathcal{R}^{L_1 \times n}$ and $\hat{W}_2 \in \mathcal{R}^{L_2 \times n}$ are weight matrices, $\phi_1 \in \mathcal{R}^{L_1 \times 1}$ and $\Phi_2 \in \mathcal{R}^{L_2 \times m}$ are the basis functions of the networks respectively. L_1 and L_2 are the number of hidden neurons in the networks. The NN-based model (3.17) will fairly approximate (3.15) if the appropriate update laws for $\hat{W}_1$ and $\hat{W}_2$ are derived such that the system response of (3.17) follows that of actual plant (3.15).

Let us assume that there exists two ideal weight matrices W_1 and W_2 such that $\overline{f}(x)$ and $g(x)$ can be written as

$$\overline{f}(x) = W_1^T \phi_1(x) \tag{3.18}$$

$$g(x) = W_2^T \Phi_2(x) \tag{3.19}$$

As it is not possible to find a NN that can exactly approximate a nonlinear function, the actual plant dynamics (3.16) in terms of ideal NN weights are given by

$$\dot{x} = -x + W_1^T \phi_1(x) + W_2^T \Phi_2(x)u \tag{3.20}$$

The error vector between the system states and model states is defined as

$$e = x - \hat{x} \tag{3.21}$$

Thus the error dynamics between the actual plant and the identified model can be derived by subtracting (3.17) from (3.20)

$$\begin{aligned}\dot{e} &= \dot{x} - \dot{\hat{x}} \\ &= -x + W_1^T \phi_1(x) + W_2^T \Phi_2(x) u + \hat{x} - \hat{W}_1^T \phi_1(x) - \hat{W}_2^T \Phi_2(x) u \\ &= -e + \tilde{W}_1^T \phi_1(x) + \tilde{W}_2^T \Phi_2(x) u\end{aligned} \tag{3.22}$$

To derive the weight update laws, let us consider a Lyapunov function candidate, V as

$$V = \frac{1}{2} e^T e + tr\left\{ \frac{1}{\alpha_1} \tilde{W}_1^T \tilde{W}_1 + \frac{1}{\alpha_2} \tilde{W}_2^T \tilde{W}_2 \right\} \tag{3.23}$$

where *tr* represents trace of a matrix. Taking the derivative of (3.23)

$$\dot{V} = e^T \dot{e} + tr\left\{ \frac{1}{\alpha_1} \tilde{W}_1^T \dot{\tilde{W}}_1 + \frac{1}{\alpha_2} \tilde{W}_2^T \dot{\tilde{W}}_2 \right\} \tag{3.24}$$

Since W_1 and W_2 are constant matrices, then (3.24) is written as

$$\begin{aligned}\dot{V} &= e^T \dot{e} + tr\left\{ -\frac{1}{\alpha_1} \tilde{W}_1^T \dot{\hat{W}}_1 - \frac{1}{\alpha_2} \tilde{W}_2^T \dot{\hat{W}}_2 \right\} \\ &= e^T \left(-e + \tilde{W}_1^T \phi_1(x) + \tilde{W}_2^T \Phi_2(x) u \right) + tr\left\{ -\frac{1}{\alpha_1} \tilde{W}_1^T \dot{\hat{W}}_1 - \frac{1}{\alpha_2} \tilde{W}_2^T \dot{\hat{W}}_2 \right\}\end{aligned} \tag{3.25}$$

Using the properties of trace, (3.25) can be written as

$$\dot{V} = -e^T e + tr\left\{ \tilde{W}_1^T \phi_1(x) e^T + \tilde{W}_2^T \Phi_2(x) u e^T - \frac{1}{\alpha_1} \tilde{W}_1^T \dot{\hat{W}}_1 - \frac{1}{\alpha_2} \tilde{W}_2^T \dot{\hat{W}}_2 \right\} \tag{3.26}$$

Thus, consider the update laws

$$\dot{\hat{W}}_1 = \alpha_1 \phi_1(x) e^T \tag{3.27}$$

$$\dot{\hat{W}}_2 = \alpha_2 \Phi_2(x) u e^T \tag{3.28}$$

Substituting (3.27) and (3.28) into (3.26), we have

$$\dot{V} = -e^T e \tag{3.29}$$

Equation (3.29) is negative definite. Since V is taken to be positive definite and $\dot{V}$ turns to be negative definite, error e will converge to 0. Thus the model states will match the system states as time progresses which serve the purpose of system identification.

3.3.2.3 Controller Design

Based on the system identified model a control law can be designed using the feedback linearization technique, if the nonlinear functions matrix g is invertible. Suppose that g is a square matrix and is invertible. Also suppose that system should follow a desired trajectory x_d. In feedback linearization technique [7, 71], the control law is designed such that the error dynamics becomes linear as well as stable. Here the objective is to design a control law for the identified model such that the model state vector tracks x_d. If the model is

an accurate representation of the system, the system states will also follow the same desired trajectory.

Using the certainty equivalence principle [71], consider the control law

$$u = \left(\hat{W}_2^T \Phi_2(x)\right)^{-1} \left[\hat{x} - \hat{W}_1^T \phi_1(x) + \dot{x}_d + K_v \hat{e} \right] \tag{3.30}$$

where x_d is the desired state vector, $\hat{e} = x_d - \hat{x}$, and K_v is a design parameter. Substituting (3.30) into (3.17), we have

$$\begin{aligned} \dot{\hat{x}} &= -\hat{x} + \hat{W}_1^T \phi_1(x) + \hat{W}_2^T \Phi_2(x) \left(\hat{W}_2^T \Phi_2(x)\right)^{-1} \left[\hat{x} - \hat{W}_1^T \phi_1(x) + \dot{x}_d + K_v \hat{e} \right] \\ &= \dot{x}_d + K_v \hat{e} \\ \Rightarrow \dot{\hat{e}} &= -K_v \hat{e} \end{aligned} \tag{3.31}$$

If we consider a Lyapunov function candidate as $V = \frac{1}{2} \hat{e}^T \hat{e}$, then $\dot{V}$ can be written as

$$\dot{V} = \hat{e}^T \dot{\hat{e}} = -\hat{e}^T K_v \hat{e} \tag{3.32}$$

that is negative definite. Thus, $\hat{e}$ will converge to zero with time, which serves the purpose of tracking.

When the matrix $g(x)$ is a rectangular matrix, the negative definiteness of $\dot{V}$ cannot be ensured because the use of pseudo-inverse of $g(x)$ will not exactly cancel the nonlinearity $f(x)$. In such cases, if the system is a SISO system, then mostly it can be modelled in a strict feedback form like

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 \\ \dot{\hat{x}}_2 &= \hat{x}_3 \\ &\cdots\cdots\cdots\cdots \\ \dot{\hat{x}}_n &= \hat{f}(x) + \hat{g}(x) u \end{aligned} \tag{3.33}$$

$$\hat{y} = \hat{x}_1 \tag{3.34}$$

where $\hat{f}(x)$ and $\hat{g}(x)$ are scalar functions.

Let us define the output tracking error as

$$\hat{e} = y_d - \hat{y} = x_{1d} - \hat{x}_1 \tag{3.35}$$

and another variable *r* a filtered tracking error given by

$$r = \hat{e}^{(n-1)} + \lambda_1 \hat{e}^{(n-2)} + \ldots + \lambda_{n-1} \hat{e} \tag{3.36}$$

Using the certainty equivalence principle, if the control law is chosen as

$$u = \frac{1}{g(x)} \left[-\hat{f}(x) + k_v r + \lambda_1 \hat{e}^{(n-1)} + \ldots + \lambda_{n-1} \hat{e}^{(1)} + \dot{x}_{nd} \right] \tag{3.37}$$

then, we have

$$\dot{r} = -k_v r \tag{3.38}$$

which is a linear and stable dynamics. Considering a Lyapunov function $V = \frac{1}{2} r^2$, we can show that r converges to zero with time. $\lambda_1, ..., \lambda_{n-1}$ are design parameters which can be chosen such that the dynamics given by (3.36) is stable.

3.3.3 Simulation Results and Analysis

In this section, the application of indirect adaptive control using RBFNN with feedback linearization technique is presented for control of a nonlinear process system. A liquid level control system for a surge tank is considered as an example of a nonlinear process system.

Define the surge tank liquid level control system output as $h(t)$ and model output as $h_m(t)$, then from (3.17) we have

$$\dot{h}_m(t) = -h_m(t) + \hat{W}_1^T \phi_1(h) + \hat{W}_2^T \phi_2(h) u \tag{3.39}$$

where $\hat{W}_1$ and $\hat{W}_2$ are two weight vectors, ϕ_1 and ϕ_2 are Gaussian functions.

The simulation parameters considered are: system parameters *a*=1; *b*=2; number of neurons in hidden layer of networks for function $\hat{f}(x)$ is 30, and for function $\hat{g}(x)$ is 20 respectively and thus the sizes of weight vectors respectively; learning parameters $\alpha_1 = 0.8$, $\alpha_2 = 0.5$; acceleration due to gravity $g = 9.8$ m/s^2; For system identification the training data generated for a simulation time of 50 seconds with sample time $T = 0.01$sec with a sinusoidal input $u = 10 + 5\sin(t)$ of magnitude between 5 and 10. The initial liquid level considered is 8 m. The system and model dynamics are evolved while updating the weights using the update laws (3.17) and (3.18), where $e = h - h_m$. For neural network control the initial liquid level considered is 8 m, desired liquid level $h_d = 5$m, design parameter $k_v = 4$, and remaining parameters are considered same as for the model identification for a simulation time of 20 seconds. In the simulation the centres and spread of Gaussian functions are considered as $c_i = 5 + 15 * rand(n,1)$, and $\sigma = 1$ respectively for RBFNN in both cases, where n is the number of radial centres in hidden layer of RBFNN. The simulation results obtained are shown in Fig. 3.4Fig. 3.5, and Fig. 3.6. Fig. 3.4 shows the system output and model output for system identification. It is observed that the identified system model is fairly approximating the actual system. Fig. 3.5 shows tracking of desired liquid level $h_d = 5$m based on identified model, showing the system output, model output, and desired output. It is observed that the model output tracks the desired trajectory. The liquid level of 4.2352 m is obtained for the actual system whereas the liquid level of 5.0000 m is obtained for the system model using RBFNN. The control input is shown in Fig. 3.6. Using the control law

(3.30) the control input is generated to achieve the desired liquid level. The obtained control input is the input flow of 9.1104 m^3/sec to achieve the desired liquid level in the surge tank.

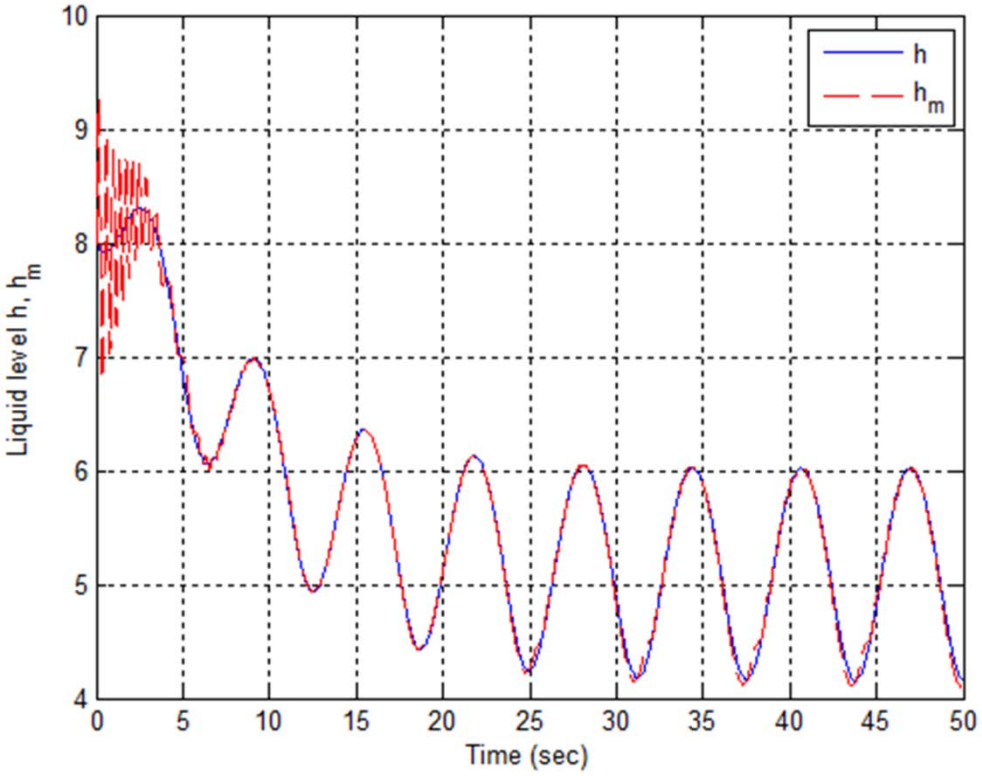


Fig. 3.4 System output and model output for system identification.

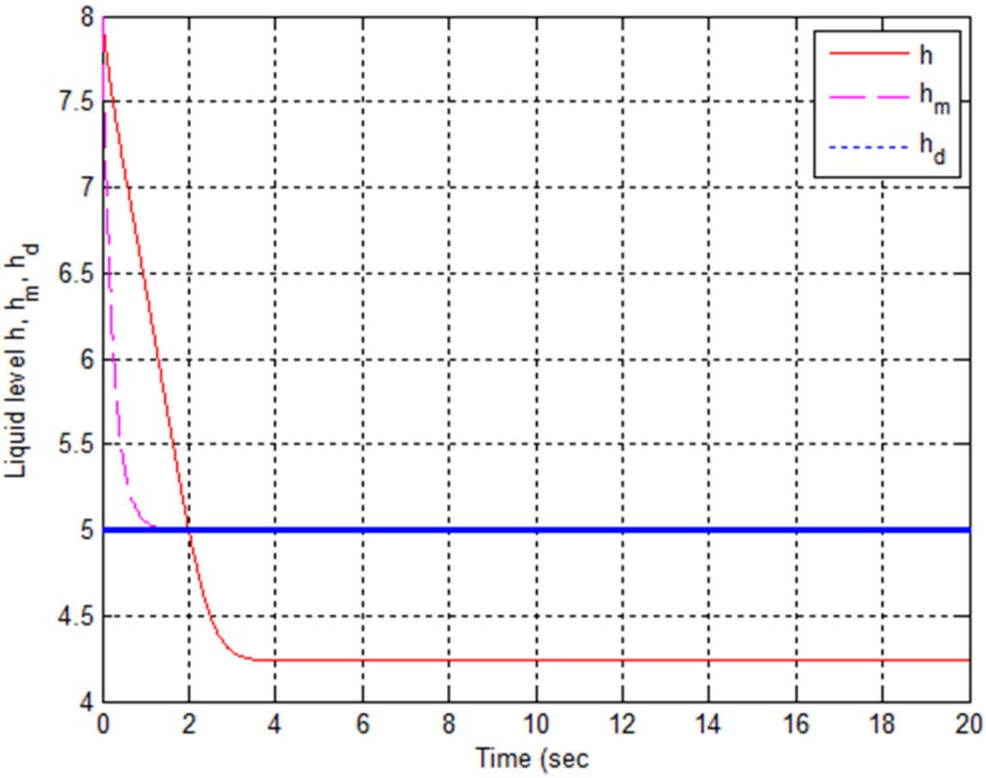


Fig. 3.5 Tracking of desired liquid level.

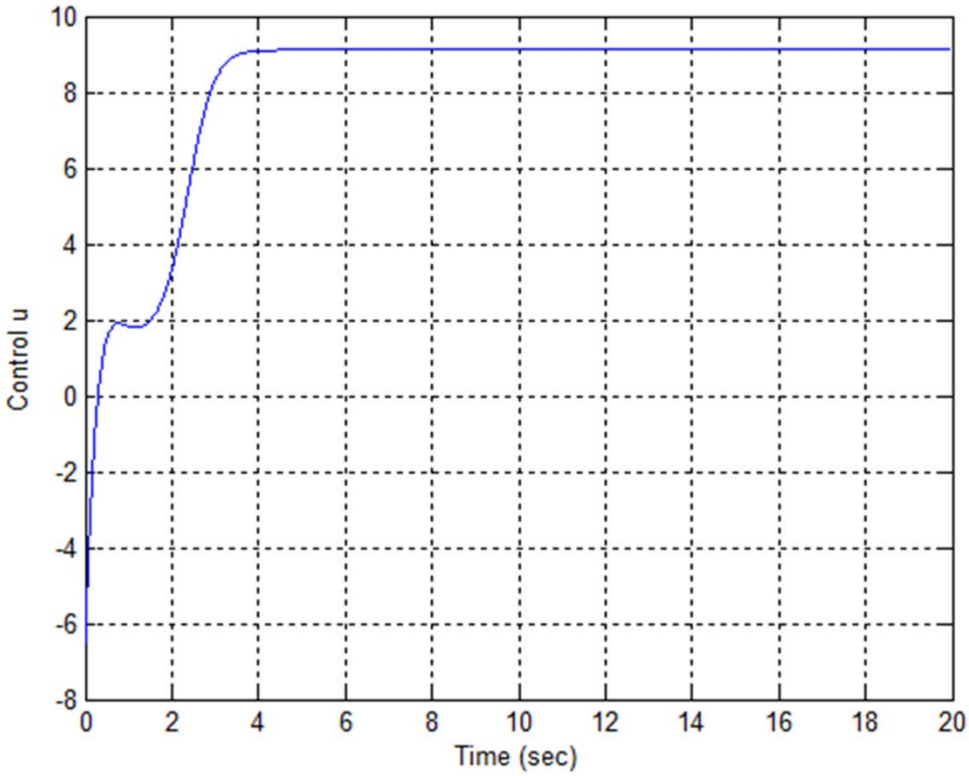


Fig. 3.6 Control input.

3.3.4 Discussion

The indirect adaptive control using RBFNN gives an intelligent control solution for a nonlinear process system. The feedback linearization technique for affine nonlinear system control gives the linear and stable error dynamics, cancelling the nonlinearity in the system. The response of neural network control using RBF Networks is fast.

3.4 INTELLIGENT CONTROL OF NONLINEAR INVERTED PENDULUM DYNAMICAL SYSTEM USING FUZZY LOGIC SYSTEMS

Fuzzy control is an intelligent control technique that uses the human expert knowledge to make control decisions. It provides a simple approach to use heuristic control knowledge for nonlinear control problems. Recently fuzzy logic systems have been widely used for various applications [9, 15, 43, 61, 62, 67, 68, 85, 86]. In recent trends even the various advance control approaches are developing and being tried for many dynamical systems control, the comparative performance analysis of fuzzy control methods using Mamdani and Takagi-Sugeno-Kang (TSK) fuzzy inference systems (FIS) is desired. The nonlinear inverted pendulum-cart dynamical system is used as a benchmark control problem for this comparative study with control objectives of cart position control at a desired position and stabilization of inverted pendulum in the vertically upright position. The comprehensive performance investigation for the control of nonlinear inverted pendulum-cart system for both cases of without & with disturbance input is also desired.

This section comprehensively presents the intelligent control of nonlinear inverted pendulum-cart dynamical system using Mamdani FIS and TSK FIS. The objectives & contributions of this section is to present the modelling, control design and comparative performance analysis of Mamdani FIS and TSK FIS based intelligent control methods for nonlinear inverted pendulum-cart dynamic system for both cases of without & with continuous disturbance input. The comparative performance analysis of these fuzzy control methods have also been done with PID control method. The simulation results and performance analysis justify the comparative advantages of fuzzy control methods. The performance analysis shows that the proposed control methods are simple, effective, and robust. The mathematical modelling of nonlinear inverted pendulum system as discussed in section 2.5.2 is referred in this section for fuzzy control design.

3.4.1 Fuzzy Control Using Mamdani and TSK Fuzzy Inference Systems

In the implementation of fuzzy control the mathematical model equations of systems are not needed but the expert knowledge of the system behaviour is required to be known, that is the information which is gathered by equations. Fuzzy logic systems address the imprecision or vagueness in input–output descriptions of systems using fuzzy sets. Fuzzy logic may be employed to represent a set of fuzzy rules, the knowledge of a human controlling a plant. This is the process of knowledge representation. Then, a rule of inference in fuzzy logic may be used according to this fuzzy knowledge base to make control decisions for a given set of plant observations. This task concerns “knowledge processing”. In this sense, fuzzy logic in intelligent control serves to represent and process the control knowledge of a human in a given plant. The fuzzy models categorized into three types according to expressions of the consequent part are:

1. Mamdani Fuzzy Inference System (Ebrahim Mamdani, 1975)

   $y = A$ where, A is a fuzzy number (3.40)

2. Takagi-Sugeno-Kang (TSK) Fuzzy Inference System (T-S-K, 1985)

   $$y = a_0 + \sum a_i x_i \tag{3.41}$$

   where, a_i is a constant and x_i is the input variable.

3. Simplified model

   $y = c$ where, c is a constant. (3.42)

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned. Basically Mamdani type FIS and Takagi-Sugeno-Kang (i.e. Sugeno type) FIS are generally used in practice. Both are similar in many respects. The schemes of operation of fuzzy rules in Mamdani FIS, and TSK FIS are shown in Figs. 3, and 4 respectively . The first two parts of the fuzzy inference process, fuzzifying the inputs and

applying the fuzzy operator, are exactly the same. These two types of inference systems vary somewhat in the way outputs are determined. The main difference between Mamdani FIS and Sugeno FIS is that the Sugeno FIS output membership functions are either linear or constant.

Fuzzy rules of Mamdani fuzzy inference systems:

If x is (A linguistic term) and y is (B linguistic term) Then z is (C linguistic term)

where A, B, & C are fuzzy sets

Examples: R1: if X is small and Y is small then Z is small

R2: if X is small and Y is large then Z is large

Fuzzy rules of TSK fuzzy inference systems:

If x is A and y is B Then z = f(x, y)

where A & B are fuzzy sets, and f(x, y) is crisp function very often a polynomial function w.r.t. x and y

Examples: R1: if X is small and Y is small then z = -x +y +1

R2: if X is small and Y is large then z = -y +3

Advantages of the Mamdani & Sugeno Methods:

Mamdani method is intuitive, widespread acceptance, and well suited to human input whereas Sugeno method is computationally efficient, works well with linear techniques (e.g., PID control), works well with optimization and adaptive techniques, has guaranteed continuity of the output surface, and well suited to mathematical analysis.

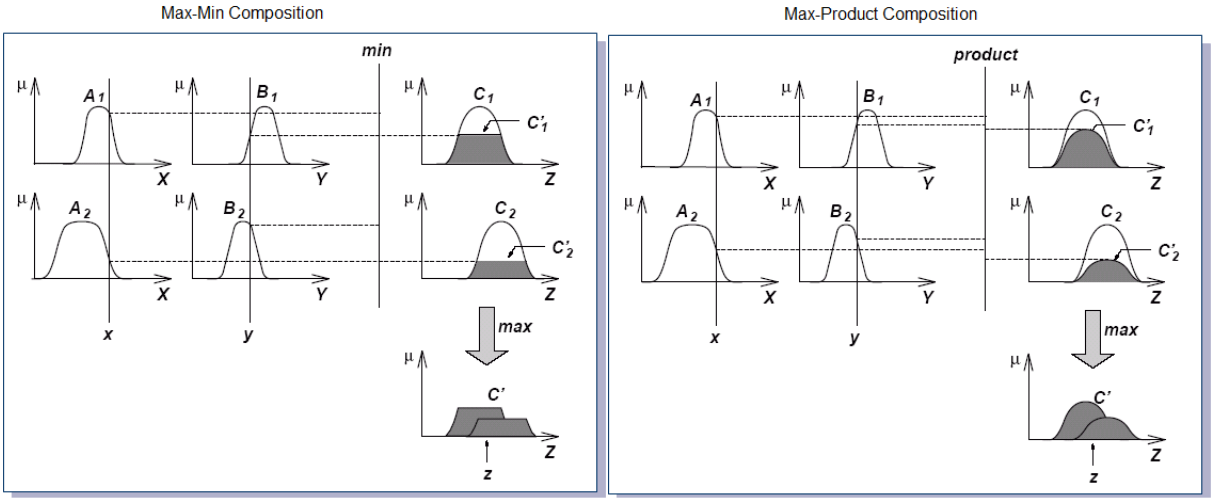


Fig. 3. Scheme of operation of fuzzy rules in Mamdani fuzzy inference system.

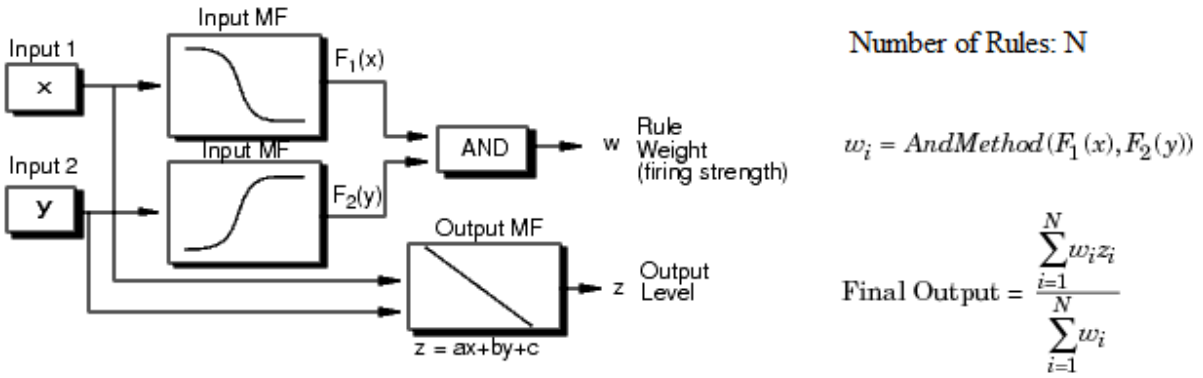


Fig. 4. Scheme of Operation of Fuzzy Rules in TSK Fuzzy Inference System.

Here two fuzzy control approaches using Mamdani FIS and TSK FIS for nonlinear plant model of inverted pendulum-cart dynamic system have been implemented. In Mamdani type FIS two PD-fuzzy logic controllers (FLC) - angle FLC, and cart FLC have been used for the two control loops of the system. The inputs to FLCs are errors in pendulum angle & angular velocity, and errors in cart position & cart velocity respectively. The output is control force. In the PD type fuzzy control using Mamdani type FIS 49 rules with 7 triangular membership functions (NL, NM, NS, Z, PS, PM, PL) for each variable have been used for angle FLC & cart FLC both. In the fuzzy control approach using TSK type (commonly known as Sugeno type) FIS direct fuzzy control approach has been used with 2 gbell membership functions (N &P) for each input variable and 16 linear membership functions (u1 to u16) for output variable. The rule bases for pendulum angle control, and cart position control using Mamdani type FIS are given in Tables 1 and 2 respectively. The rule base for direct fuzzy control using TSK type FIS is given in Table 3.

Table 3.1 Rule base for pendulum angle FLC using Mamdani FIS

Control Force u		Error in angular velocity Thetadot (ethetadot)						
		NL	NM	NS	Z	PS	PM	PL
Error in angle Theta (etheta)	NL	NL	NL	NL	NL	NM	NS	Z
	NM	NL	NL	NL	NM	NS	Z	PS
	NS	NL	NL	NM	NS	Z	PS	PM
	Z	NL	NM	NS	Z	PS	PM	PL
	PS	NM	NS	Z	PS	PM	PL	PL
	PM	NS	Z	PS	PM	PL	PL	PL
	PL	Z	PS	PM	PL	PL	PL	PL

Table 3.2 Rule base for cart FLC using Mamdani FIS

Control Force u		Error in Cart velocity xdot (exdot)						
		NL	NM	NS	Z	PS	PM	PL
Error in cart position x (ex)	NL	NL	NL	NL	NL	NM	NS	Z
	NM	NL	NL	NL	NM	NS	Z	PS
	NS	NL	NL	NM	NS	Z	PS	PM
	Z	NL	NM	NS	Z	PS	PM	PL
	PS	NM	NS	Z	PS	PM	PL	PL
	PM	NS	Z	PS	PM	PL	PL	PL
	PL	Z	PS	PM	PL	PL	PL	PL

Table 3.3 Rule base for direct FLC using TSK FIS

1. If (theta is N) & (thetadot is N) & (x is N) & (xdot is N) then (u is u1)
2. If (theta is N) &(thetadot is N) & (x is N) & (xdot is P) then (u is u2)
3. If (theta is N) & (thetadot is N) & (x is P) & (xdot is N) then (u is u3)
4. If (theta is N) & (thetadot is N) & (x is P) & (xdot is P) then (u is u4)
5. If (theta is N) & (thetadot is P) & (x is N)& (xdot is N) then (u is u5)
6. If (theta is N) & (thetadot is P) & (x is N) & (xdot is P) then (u is u6)
7. If (theta is N) & (thetadot is P) & (x is P) & (xdot is N) then (u is u7)
8. If (theta is N) & (thetadot is P) & (x is P) & (xdot is P) then (u is u8)
9. If (theta is P) & (thetadot is N) & (x is N) & (xdot is N) then (u is u9)
10. If (theta is P) & (thetadot is N) & (x is N) & (xdot is P) then (u is u10)
11. If (theta is P) & (thetadot is N) & (x is P) & (xdot is N) then (u is u11)
12. If (theta is P) & (thetadot is N) & (x is P) & (xdot is P) then (u is u12)
13. If (theta is P) & (thetadot is P) & (x is N) & (xdot is N) then (u is u13)
14. If (theta is P) & (thetadot is P) & (x is N) & (xdot is P) then (u is u14)
15. If (theta is P) & (thetadot is P) & (x is P) & (xdot is N) then (u is u15)
16. If (theta is P) & (thetadot is P) & (x is P) & (xdot is P) then (u is u16)

3.4.2 Simulation Results and Analysis

The MATLAB-SIMULINK models for the simulation of modeling, analysis, and control of nonlinear inverted pendulum-cart dynamical system without & with continuous disturbance input have been developed. The typical parameters of inverted pendulum-cart system setup are selected as [85, 180]: mass of the cart (*M*): 2.4 kg, mass of the pendulum (*m*): 0.23 kg, length of the pendulum (*l*): 0.36 m, length of the cart track (*L*): ± 0.5 m, friction coefficient of the cart & pole rotation is assumed negligible. The disturbance input parameters which has

been taken in simulation are [205] : Band Limited White Noise Power = 0.001, Sample Time = 0.01, Seed = 23341.

The SIMULINK models for fuzzy control of inverted pendulum system for linear plant model without & with disturbance input using Mamdani type FIS are shown in Fig. 3.7 and Fig. 3.9 respectivey. Here the linearized state-space model has been used, and only pendulum angle theta and cart position x have been considered for the measurement. The band limited white noise has been added as the disturbance input to the system. The reference angle is set to 0 and the desired cart position has been set to 0.1 (m). In the fuzzy control using Mamdani type FIS In the fuzzy control approach using Mamdani type FIS the universe of discourses for linear plant model of inverted pendulum are: angle FLC: etheta=[-0.3 0.3], ethetadot=[-1 1], u=[-1 1], and cart FLC: ex=[-1 1], exdot=[-1 1], u [-1 1]. The change of sign of gains with linear & nonlinear plant models are just due to considered change of sign in error detectors. The tuned gains of PD-FLCs for linear plant model are: angle FLC: K_p=-40, K_d=-8, and cart FLC: K_p=-1, K_d=-4. Both angle & cart FLCs used 49 rules with 7 triangular MFs (NL, NM, NS, Z, PS, PM, PL) for each variable. The simulation results for both cases are shown in Fig. 3.8 and Fig. 3.10 respectively. It is observed here that the pendulum stabilizes vertically with minor oscillations in range approx +/-0.002 (rad), and the cart reaches the desired position of 0.1 (m) quickly with minor oscillations in both cases. The control input u is bounded in range of [-0.1 0.1] for without disturbance input case and in range of [-1 1] for with disturbance input case. The simulation results justify the effectiveness of the PD-FLC for the linearized model case.

The SIMULINK models for fuzzy control of nonlinear inverted pendulum system using Mamdani FIS of PD type for both cases of without & with disturbance input are shown in Fig. 3.11 and Fig. 3.13 respectively. The universe of discourses for PD-FLCs are: Angle FLC: etheta=[-0.3 0.3], ethetadot=[-1 1], u=[-1 1], and Cart FLC: ex=[-1 1], exdot=[-1 1], u=[-1 1]. The tuned gains of PD-FLCs for both cases are: Angle FLC: K_{pp}=40, K_{dp}=8, and Cart FLC: K_{pc}=1.25, K_{dc}=3.6. Both angle & cart FLCs used 49 rules with 7 triangular MFs (NL, NM, NS, Z, PS, PM, PL) for each variable. The simulation results for both cases are shown in Fig. 3.12 and Fig. 3.14 respectively. Here all the state variables θ, $\dot{\theta}$, x, $\dot{x}$ and control u have been plotted. It is observed here that the pendulum stabilizes in vertically upright position quickly with minute oscillations, and the angular velocity oscillates in approx range +/-0.002 (rad/s) for case of without disturbance input and in approx. range +/-0.02 (rad/s) for case of with disturbance input. The cart reaches the desired position of 0.1 (m) quickly with minor steady state error, and the cart velocity oscillates very near to zero. The control input u is bounded for both cases in approx. ranges [-0.1 0.1] and [-1 1] respectively. During initial transients the control input u is also bounded in approx. range +/-0.12N for case of without

disturbance input. The simulation results justify the effectiveness & robustness of the PD-FLC.

The SIMULINK models for direct fuzzy control of nonlinear inverted pendulum system using TSK type FIS for both cases of without & with disturbance input are shown in Fig. 3.15 and Fig. 3.17 respectively. The universe of discourses of TSK FLC are: theta=[-0.3 0.3], thetadot=[-1 1], x=[-0.5 0.5], xdot=[-0.5 0.5], u=[-10 10]. In the direct fuzzy control approach using TSK Type FIS 16 rules with 2 gbell MFs (N &P) for each input variable, and 16 linear MFs for output variable have been used. The simulation results are shown in Fig. 3.16 and Fig. 3.18 respectively. Here also all the state variables θ, $\dot{\theta}$, x, $\dot{x}$ and control u have been plotted. It is observed here that the pendulum stabilizes in vertically upright position quickly & smoothly for case of without disturbance input and with acceptable minute oscillations for case of with disturbance input. The angular velocity approaches zero for case of without disturbance input and oscillates minutely in range approx +/-0.02 (rad/s) for case of with disturbance input. The cart reaches the desired position of 0.1 (m) quickly & smoothly after initial small transients, and the cart velocity approaches zero for case of without disturbance input and oscillates very near to zero for case of with disturbance input. The control input u is bounded for both cases in approx. ranges [-0.1 0.1] and [-1 1] respectively. During initial transients the control input u is also bounded in approx. range +/-0.5N for case of without disturbance input. The simulation results justify the effectiveness & robustness of the TSK-FLC.

The analysis of the responses of PID control scheme which is presented in section 2.5.4, and the responses of Mamdani type PD-fuzzy control, and TSK type direct fuzzy control for nonlinear inverted pendulum system for both cases of without & with continuous disturbance input gives that these control schemes are effective & robust. It is observed here that the responses of Mamdani PD-fuzzy control are not as smooth as PID control & TSK fuzzy control but pendulum reaches near upright position faster than both. The responses of TSK fuzzy control are fast, and smoother than Mamdani fuzzy control and PID control both. Overall the performance of fuzzy control methods especially of TSK fuzzy control is better than PID control.

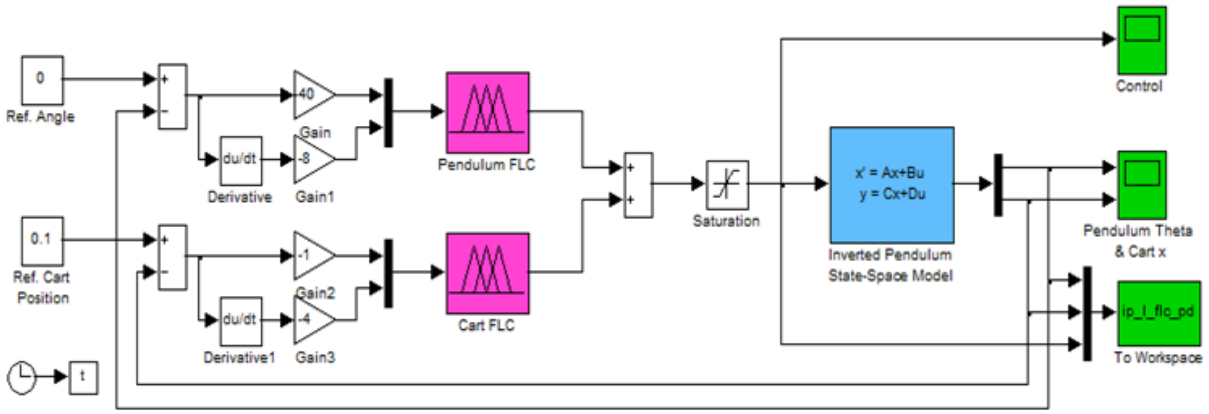


Fig. 3.7 PD-fuzzy control of linearized model of inverted pendulum system using Mamdani type FIS.

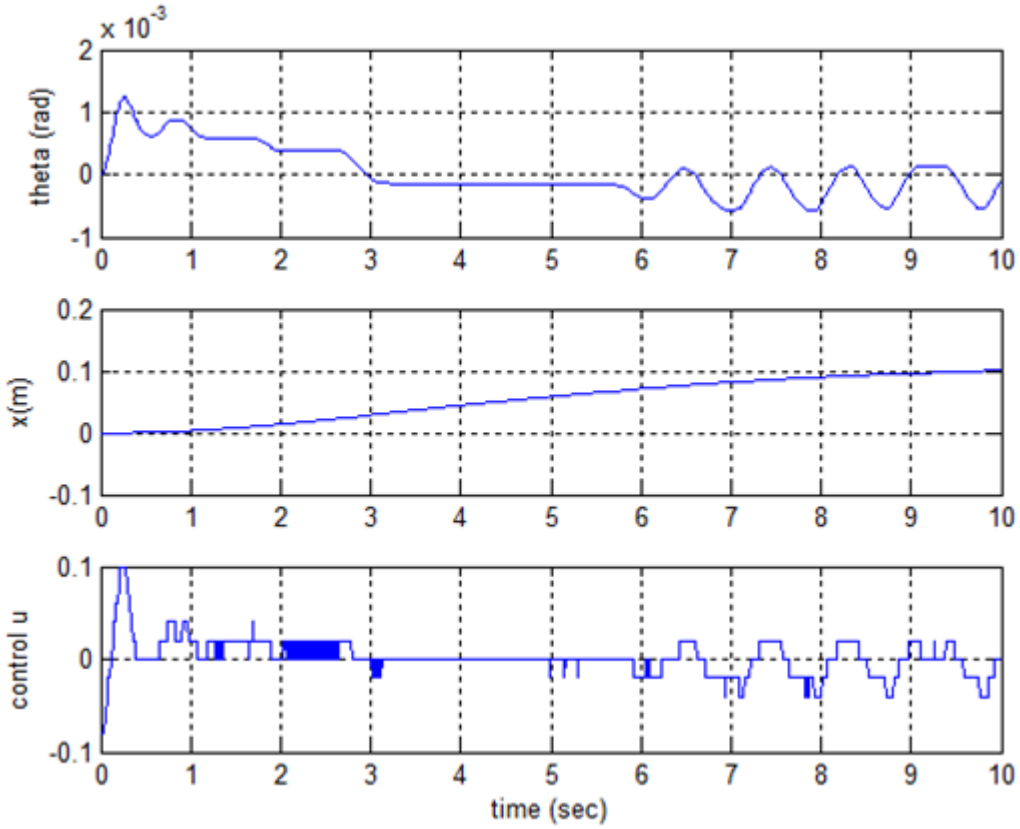


Fig. 3.8 Responses of pendulum angle theta, cart position x, and control force u of linearized model of inverted pendulum system with fuzzy control using Mamdani FIS.

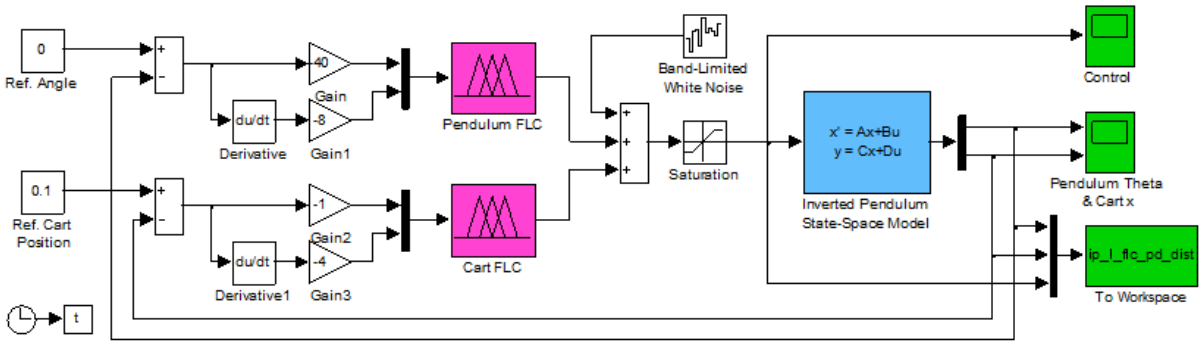


Fig. 3.9 PD-fuzzy control of linearized model of inverted pendulum system with disturbance input using Mamdani FIS.

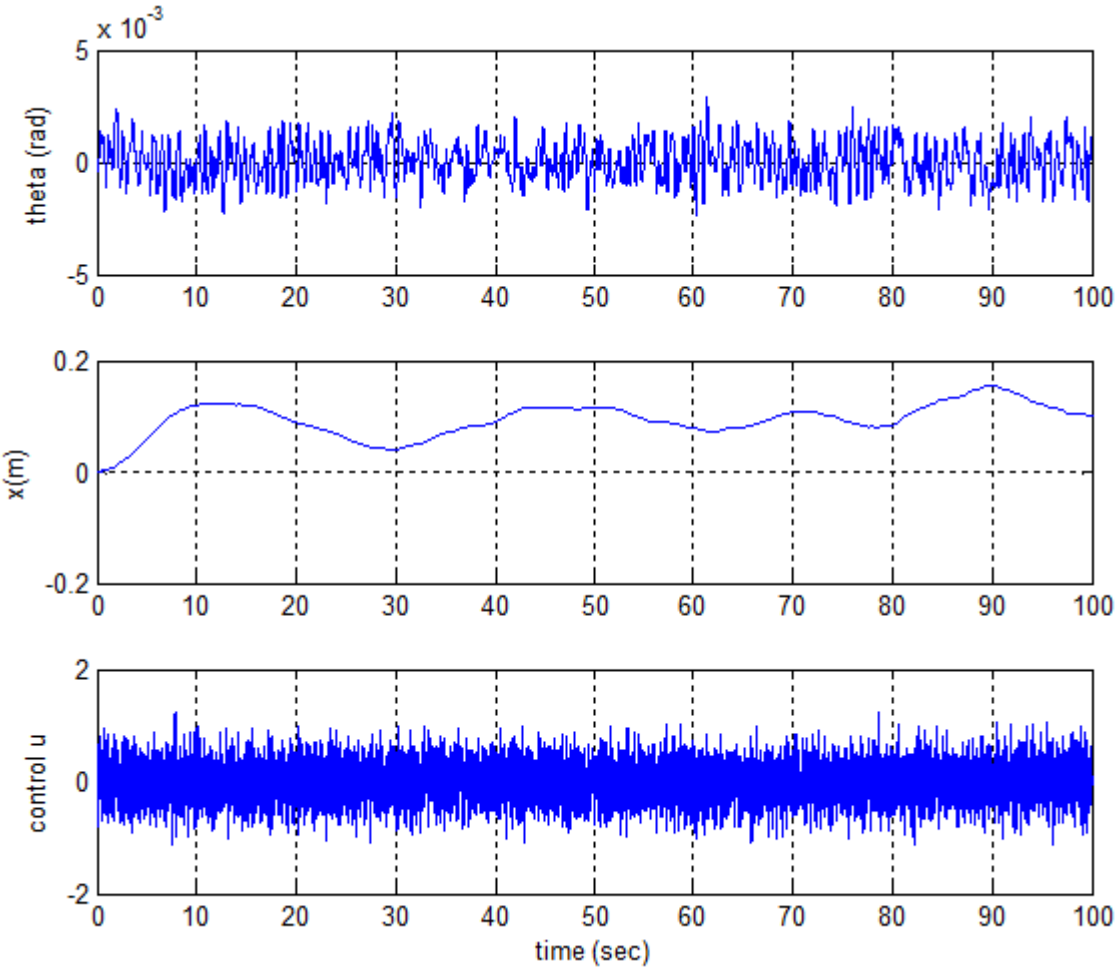


Fig. 3.10 Responses of pendulum angle theta, cart position x, and control force u of linearized model of inverted pendulum system with fuzzy control using Mamdani FIS.

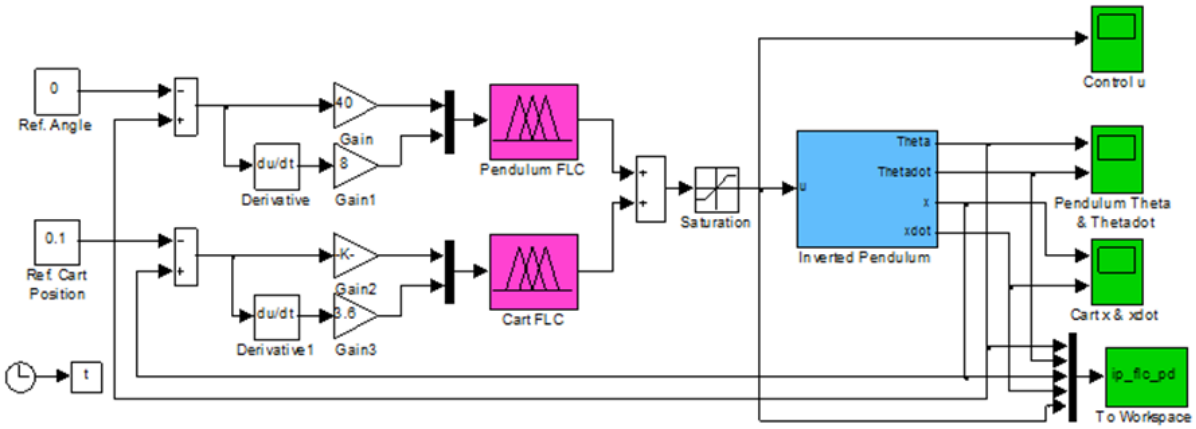


Fig. 3.11 PD-fuzzy control of nonlinear inverted pendulum system using Mamdani type FIS.

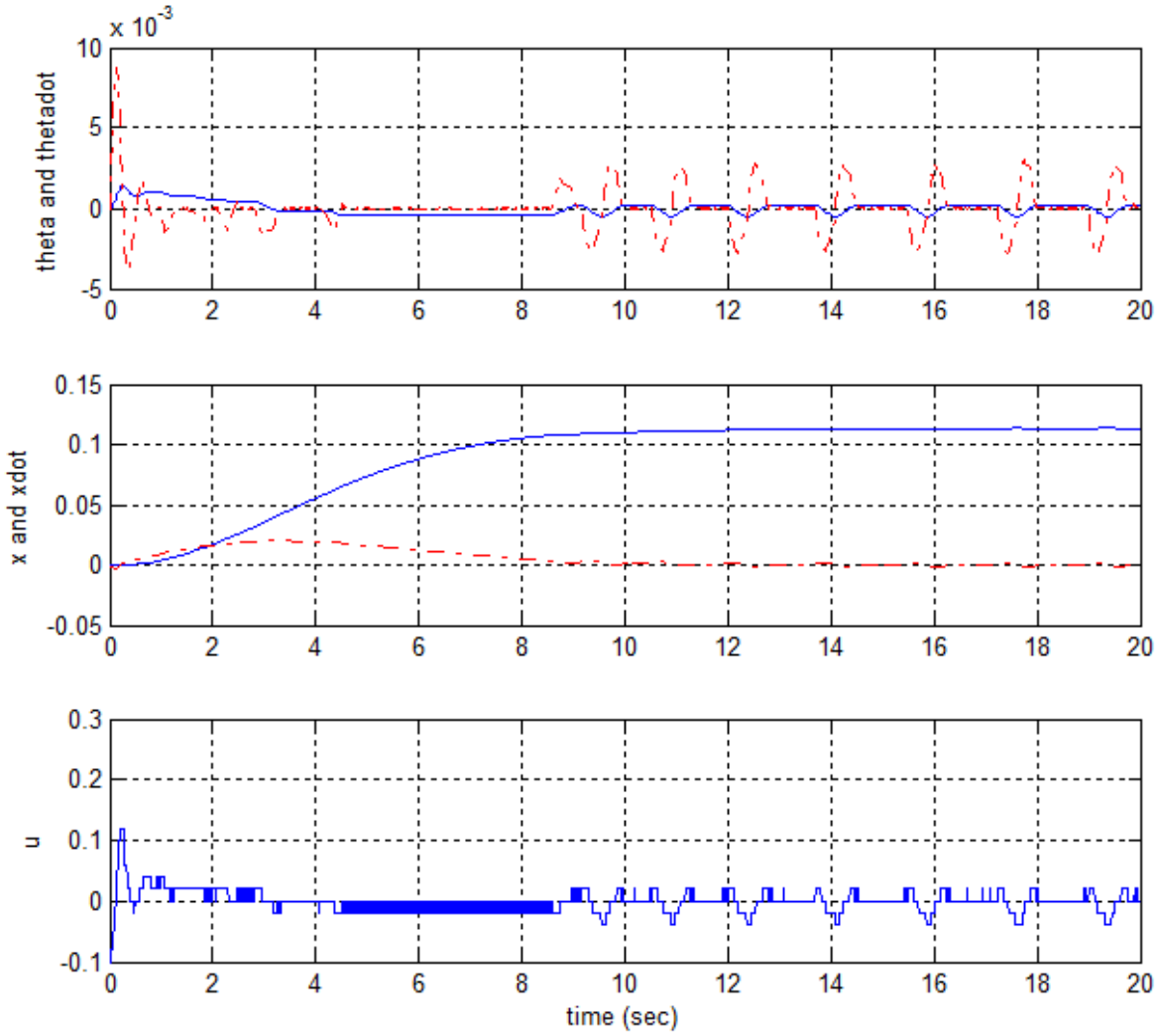


Fig. 3.12 Responses of pendulum angle theta & angular velocity $\dot{\theta}$, cart position x & cart velocity $\dot{x}$, and control force u of nonlinear inverted pendulum system with fuzzy control using Mamdani FIS.

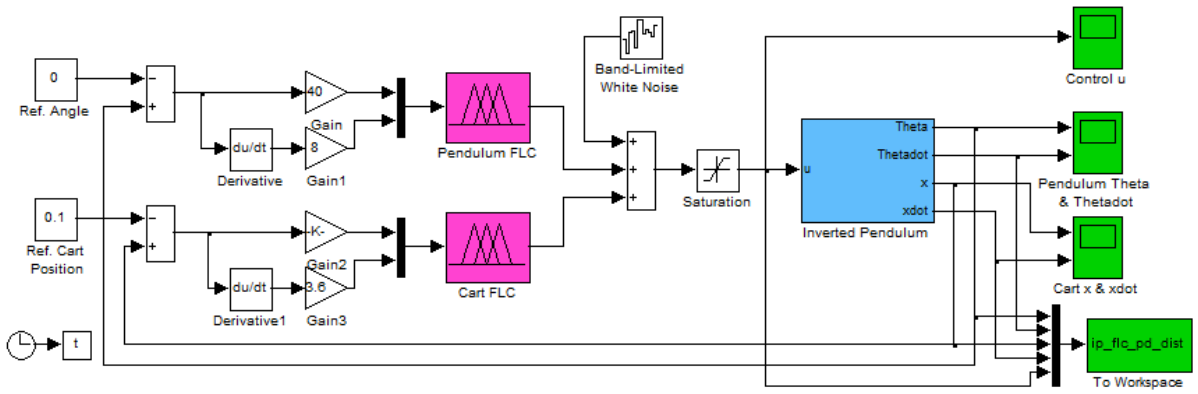


Fig. 3.13 PD-fuzzy control of nonlinear inverted pendulum system with disturbance input using Mamdani FIS.

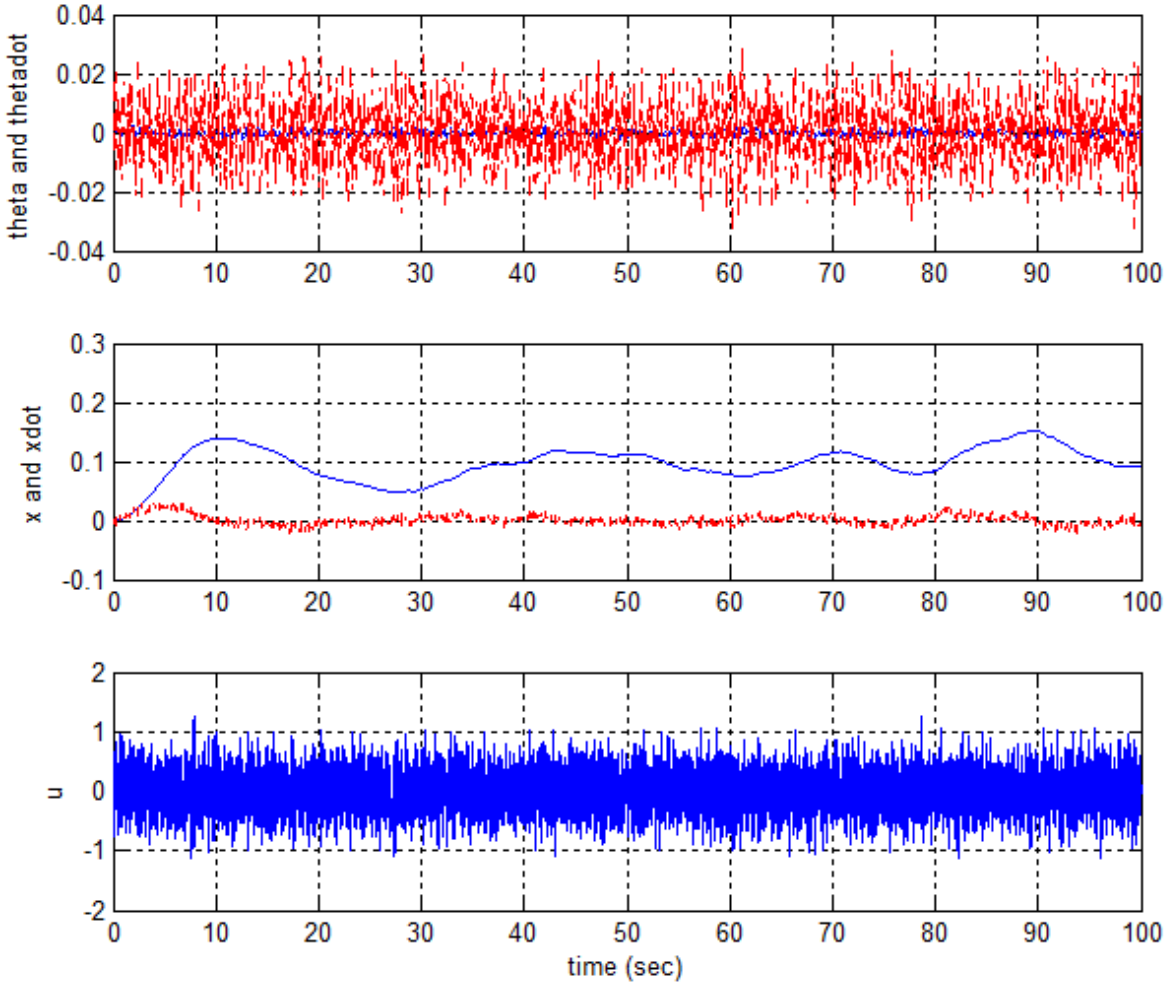


Fig. 3.14 Responses of PD-fuzzy control of nonlinear inverted pendulum system with disturbance input using Mamdani FIS.

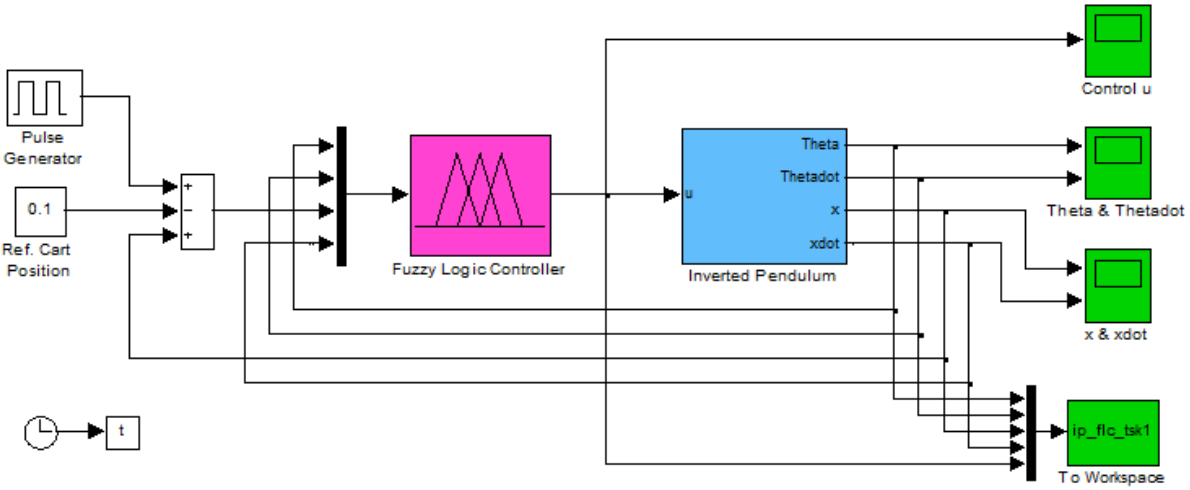


Fig. 3.15 Direct fuzzy control of nonlinear inverted pendulum system using TSK type FIS.

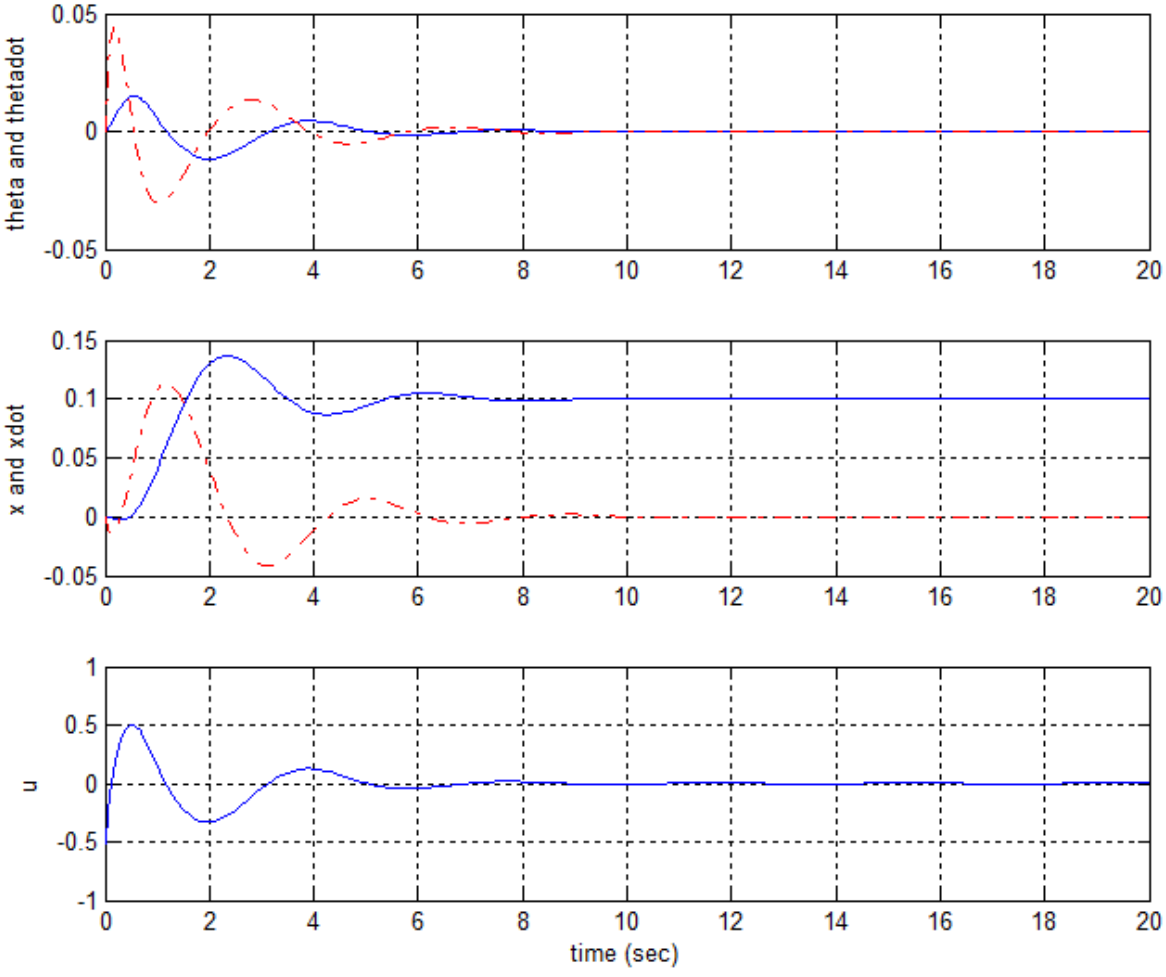


Fig. 3.16 Responses of pendulum angle theta & angular velocity thetadot, cart position x & cart velocity xdot, and control force u of nonlinear inverted pendulum system with fuzzy control using TSK FIS.

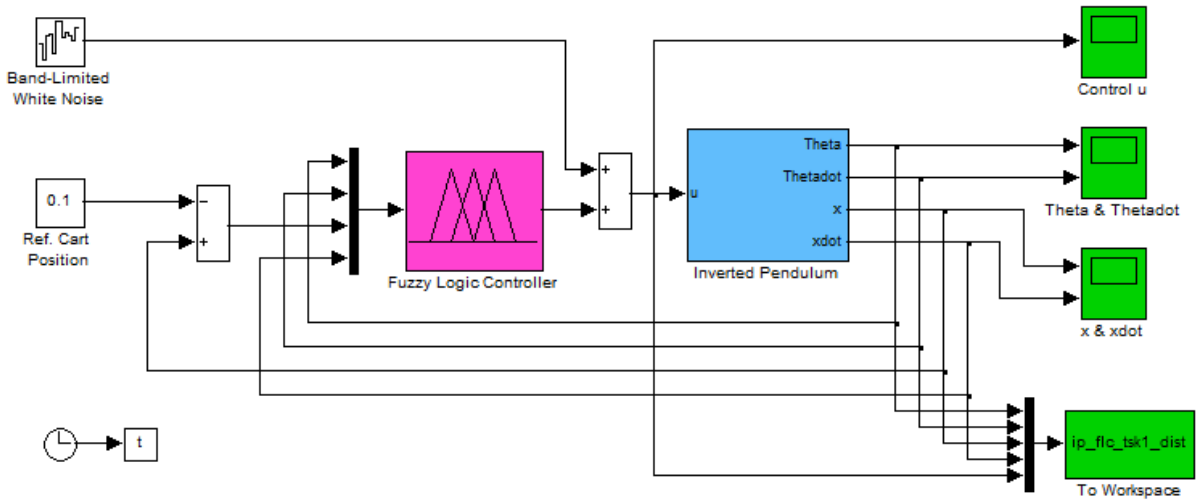


Fig. 3.17 Direct fuzzy control of nonlinear inverted pendulum system with disturbance input using TSK FIS.

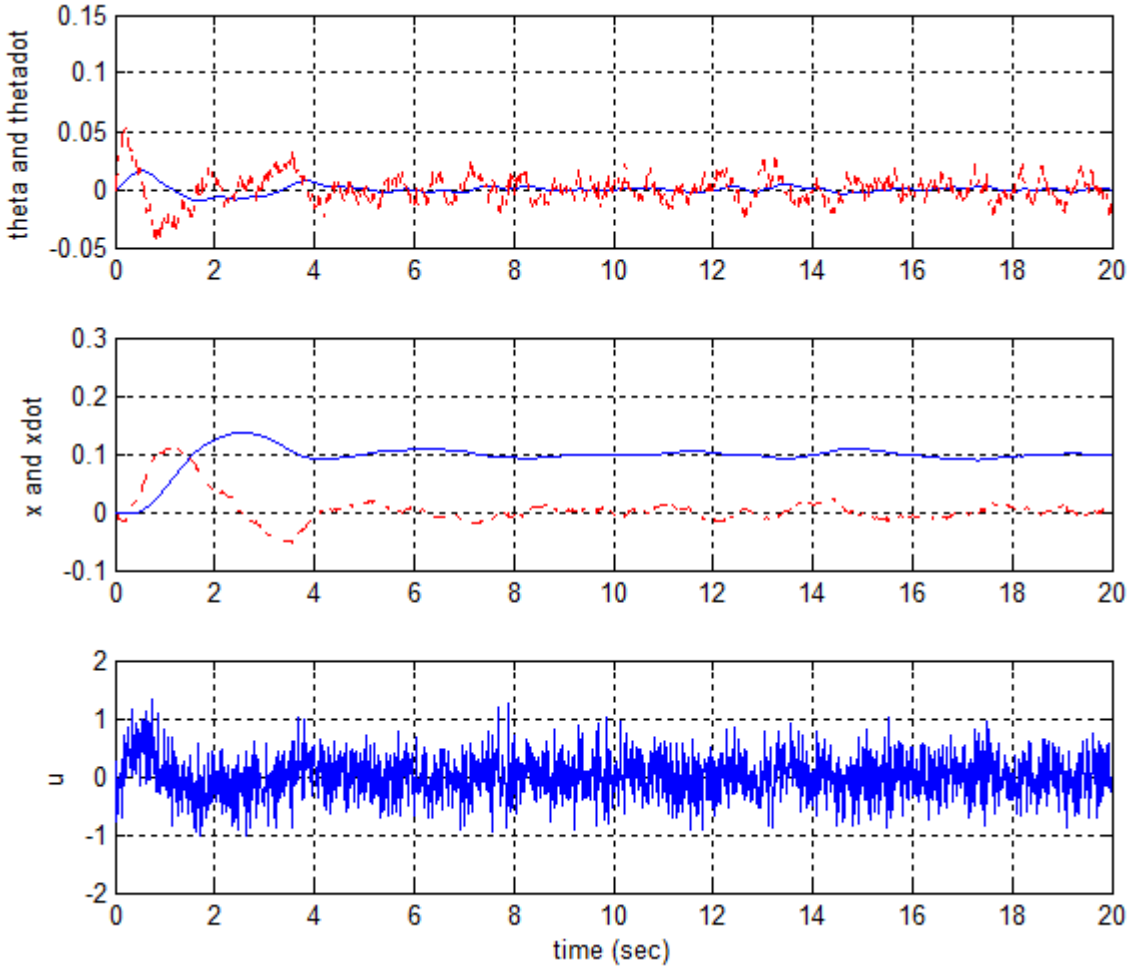


Fig. 3.18 Responses of fuzzy control of nonlinear inverted pendulum system with disturbance input using TSK FIS.

3.4.3 Discussion

Fuzzy control has been implemented using Mamdani and TSK fuzzy inference systems to control the nonlinear inverted pendulum dynamical system for both cases of without & with continuous disturbance input. The comparative performance analysis of fuzzy control methods with PID control has also been presented. The simulation results and analysis justify the comparative advantages of fuzzy control methods. The pendulum stabilizes in vertically upright position and cart approaches the desired position even under the continuous disturbance input such as wind force justify that the control schemes are effective & robust. The analysis of the responses of the control schemes gives that the performance of PD-FLC using Mamdani type FIS is better than PID controller, and the performance of TSK FLC is better than both. The response of direct fuzzy control using TSK FIS is more smooth & fast than both PID control and Mamdani type PD-fuzzy control.

3.5 FUZZY-PI BASED AUTOMATIC GENERATION CONTROL OF TWO-AREA INTERCONNECTED NONLINEAR POWER SYSTEM

Automatic generation control (AGC) which ensures the convergence of deviations in system frequency and tie-line power flows from nominal values to zero, is one of the important control problems in interconnected power system design and operation. AGC is becoming more significant today due to the increasing size, changing structure, emerging renewable energy sources and new uncertainties, environmental constraints, and complexity of power systems. AGC markets require the increased intelligence and flexibility to ensure that they are capable of maintaining a generation-load balance, following serious disturbances. The operation and control of future power systems will be challenging due to its increased complexity with ancillary services and energy markets. The intelligent AGC will be able to provide effective solution to this problem [186].

Due to features of simplicity, robustness, and reliability, fuzzy logic is used to provide control solutions for a wide range of applications including power system control and operation. The conventional control schemes are essentially based on linearized mathematical models of the controlled systems. Fuzzy control methods synthesize the controller based on measurements, long-term experiences, and knowledge of domain experts/operators. Fuzzy control provides the simple approach to use the heuristic control knowledge for nonlinear control problems.

In this section the performance analysis of fuzzy-proportional-integral (fuzzy-PI) controller for AGC of two-area interconnected nonlinear power system is presented. The comparative performance analysis with conventional AGC scheme using integral control is also presented.

3.5.1 Mathematical Modelling for Automatic Generation Control of Two-Area Interconnected Nonlinear Power System

Electrical power systems are complex nonlinear dynamic systems. Since a power system is only exposed to small changes in load demand during its normal operation, a linearized model is sufficient to represent the dynamic behavior of power system around the operating point. On the basis of this linearized model the control laws are developed using the linear control theory. However, in case of a sudden large change in load demand due to frequent switching of large load units in deregulated operations, large overshoot or prolonged oscillation of governor valve position occurs. In the presence of such nonlinearities and parametric uncertainties the system becomes highly nonlinear, and thus the control design based on linearized model may not be effective. Thus the consideration of nonlinearities is important in the synthesis and analysis of AGC system. The linearized mathematical model for AGC system of two-area interconnected power system is discussed in section 2.6.1. In this section the nonlinearities due to governor valve position limits is considered in the mathematical model of two-area interconnected nonlinear power system for AGC problem.

3.5.1.1 Two-Area Interconnected Nonlinear Power System Dynamic Equations

In the case of large change in $\Delta P_{d_k}(t)$ for k[th] control area, the dynamic equations (2.1) to (2.4) will not appropriately represent the power system due to the governor valve position limits. The large change in $\Delta P_{d_k}(t)$ require the large change in speed changer setting $\Delta P_{c_k}(t)$ to regulate frequency deviations $\Delta f_k(t)$, but the governor output $\Delta X_{g_k}(t)$ will not change beyond a specified limit. There are several ways of modelling governor valve position limits, out of which the practical piston-like steam valve structure can be modelled as a limiter with cut-off actions. The governor valve position limit is represented by a nonlinear function $\eta_k(\Delta X_{g_k}(t))$ with maximum limit $\Delta X_{gM_k}(t)$, and the minimum limit $\Delta X_{gm_k}(t)$. This governor valve position limits nonlinearity function is shown in Fig. 3.19 [99]. Including the governor valve position limits nonlinearity the block diagram of two-area interconnected nonlinear power system is shown is Fig. 3.20.

Considering the nonlinearities of governor valve position limits in the two-area interconnected nonlinear power system the dynamic equations for k[th] control area are written as [99]:

$$\Delta \dot{f}_k(t) = -\frac{1}{T_{P_k}}\Delta f_k(t) + \frac{K_{P_k}}{T_{P_k}}\Delta P_{g_k}(t) - \frac{K_{P_k}}{T_{P_k}}\Delta P_{tie_{kj}}(t) - \frac{K_{P_k}}{T_{P_k}}\Delta P_{d_k}(t) \tag{3.43}$$

$$\Delta \dot{P}_{g_k}(t) = -\frac{1}{T_{T_k}}\Delta P_{g_k}(t) + \frac{1}{T_{T_k}}\eta_k(\Delta X_{g_k}(t)) \tag{3.44}$$

$$\Delta \dot{X}_{g_k}(t) = -\frac{1}{R_k T_{G_k}} \Delta f_k(t) - \frac{1}{T_{G_k}} \Delta X_{g_k}(t) + \frac{1}{T_{G_k}} \Delta P_{c_k}(t) \tag{3.45}$$

$$\Delta \dot{P}_{tie_{kj}}(t) = 2\pi T_{kj} \Delta f_k(t) - 2\pi T_{kj} \Delta f_j(t) \tag{3.46}$$

$$\Delta \dot{E}_k(t) = ACE_k(t) = a_{kj} \Delta P_{tie_{kj}}(t) + b_k \Delta f_k(t) \tag{3.47}$$

where $k \in J_2 = \{1, 2\}$, $j \in J_2 \setminus \{k\}$ for two-area power system, $\eta_k(\Delta X_{g_k})$ is the nonlinear function of governor valve position limits, and the remaining variables and constants have same meaning as in section 2.6.1. Also the remaining descriptions of dynamic equations for two-area interconnected nonlinear power system are same as in section 2.6.1.

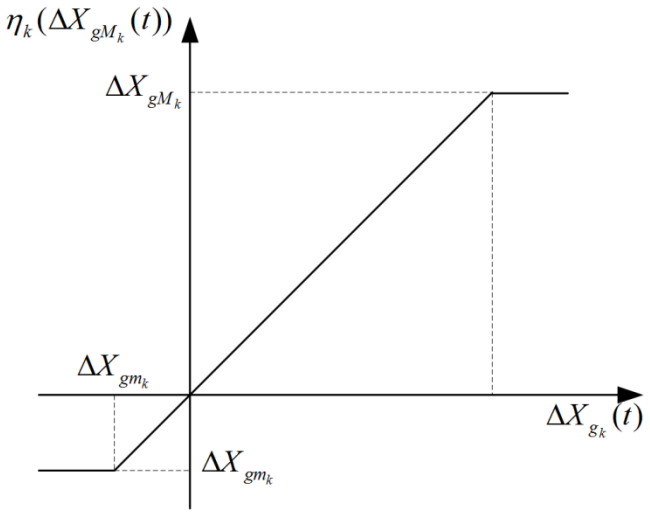


Fig. 3.19 Nonlinearity function of governor valve position limits.

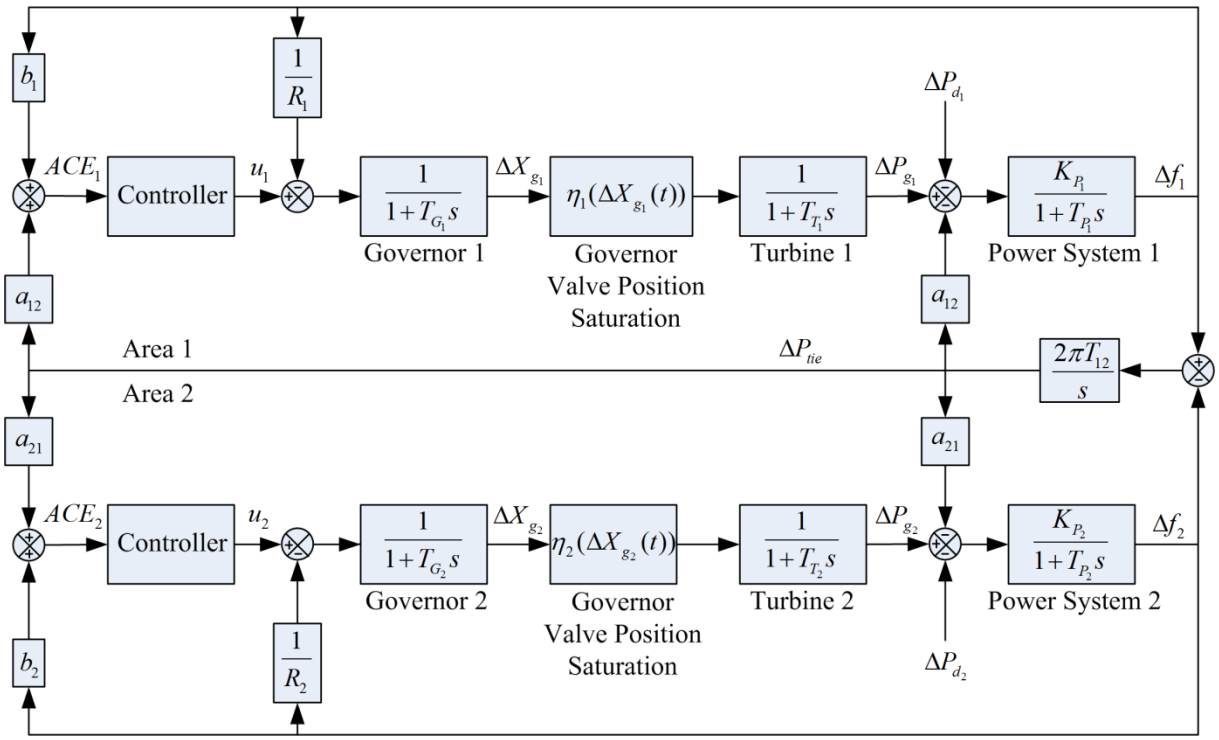


Fig. 3.20 Block diagram of two-area interconnected nonlinear power system.

3.5.1.2 State Space Model of Two-Area Interconnected Nonlinear Power System

The state space model of two-area interconnected nonlinear power system can be obtained from the dynamic equations (3.43) to (3.47) by writing these equations for area 1 and area 2, and considering the variables as in section 2.6.1.2.

The state space model of two-area interconnected nonlinear power system is given as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \\ \dot{x}_5(t) \\ \dot{x}_6(t) \\ \dot{x}_7(t) \\ \dot{x}_8(t) \\ \dot{x}_9(t) \end{bmatrix} = \begin{bmatrix} -\frac{1}{T_{P_1}}x_1(t)+\frac{K_{P_1}}{T_{P_1}}x_2(t)-\frac{K_{P_1}}{T_{P_1}}x_7(t)-\frac{K_{P_1}}{T_{P_1}}w_1(t) \\ -\frac{1}{T_{T_1}}x_2(t)+\frac{1}{T_{T_1}}\eta_1(x_3(t)) \\ -\frac{1}{R_1T_{G_1}}x_1(t)-\frac{1}{T_{G_1}}x_3(t)+\frac{1}{T_{G_1}}u_1(t) \\ -\frac{1}{T_{P_2}}x_4(t)+\frac{K_{P_2}}{T_{P_2}}x_5(t)+\frac{a_{12}K_{P_2}}{T_{P_2}}x_7(t)-\frac{K_{P_2}}{T_{P_2}}w_2(t) \\ -\frac{1}{T_{T_2}}x_5(t)+\frac{1}{T_{T_k}}\eta_2(x_6(t)) \\ -\frac{1}{R_2T_{G_2}}x_4(t)-\frac{1}{T_{G_2}}x_6(t)+\frac{1}{T_{G_2}}u_2(t) \\ 2\pi T_{12}x_1(t)-2\pi T_{12}x_4(t) \\ b_1x_1(t)+x_7(t) \\ b_2x_4(t)-a_{12}x_7(t) \end{bmatrix}$$

or

$$\dot{x}(t) = f(x(t), u(t), w(t)) \tag{3.48}$$

where

$$x(t) = \left[x_1(t), x_2(t), ..., x_9(t)\right]^T$$
$$= \left[\Delta f_1(t), \Delta P_{g_1}(t), \Delta X_{g_1}(t), \Delta f_2(t), \Delta P_{g_2}(t), \Delta X_{g_2}(t), \Delta P_{tie}(t), \int ACE_1 dt, \int ACE_2 dt\right]^T$$

is the state vector, $u(t) = \left[u_1(t) \quad u_2(t)\right]^T = \left[\Delta P_{c_1}(t) \quad \Delta P_{c_2}(t)\right]^T$ is the control vector, and $w(t) = \left[w_1(t) \quad w_2(t)\right]^T = \left[\Delta P_{d_1}(t) \quad \Delta P_{d_2}(t)\right]^T$ is the load disturbance vector; and $f(\bullet)$ is a vector valued nonlinear affine function.

3.5.2 Control Methods

For AGC system design and analysis for two-area interconnected nonlinear power system the conventional integral control method and the fuzzy-PI control method are presented in this section.

3.5.2.1 Integral Control

The integral control for AGC system of two-area interconnected nonlinear power system is defined similar to the integral control for the linear system defined by (2.87) and (2.88) for area 1 and area 2 respectively, generating the control inputs by weighted integrals of ACEs in each area. The weights of integrals of ACEs are integral gains, which are tuned for satisfactory response of deviations of area frequencies and tie-line power.

3.5.2.2 Fuzzy-PI Control

The fuzzy logic controller (FLC) for AGC system takes ACE and rate of change of ACE as input variables which are processed using fuzzy inference system to synthesize the control signal. The fuzzy-PI control scheme is the combination of fuzzy control and PI control schemes. In fuzzy-PI control the output signal of FLC is fed as the input signal to the PI controller to synthesize the control signal for the AGC system. The schematic diagram of fuzzy-PI controller is shown in Fig. 3.21. The dynamic equation of fuzzy-PI control signal given by

$$u(t) = k_P u_{flc}(t) + k_I \int_0^t u_{flc}(\tau) d\tau \tag{3.49}$$

where

$u(t) = \Delta P_c(t)$ is the control signal,

$u_{flc}(t)$ = the output signal of FLC,

k_P = proportional gain,

k_I = integral gain

The block diagram of AGC system using fuzzy-PI controller for two-area interconnected nonlinear power system is shown in Fig. 3.20. Each control area has a fuzzy-PI controller. In this fuzzy control approach Mamdani type fuzzy inference system (FIS) is used. The input variables ACEs and d/dt(ACEs), and output variable u_{flc} are represented by 7 triangular membership functions (MFs) of fuzzy sets negative-large (NL), negative-medium (NM), negative-small (NS), zero (Z), positive-small (PS), positive-medium (PM), and positive-large (PL) for each variable. The output signal of FLC is synthesized using 49 fuzzy rules of Mamdani type. The fuzzy rule-base is constructed using the dynamic relation between ACE and ΔP_c. The fuzzy rule-base of Mamdani FIS is given in Table 3.5. The centre-of-gravity (COG) defuzzyfication method is used to obtain the crisp value of output signal of FLC.

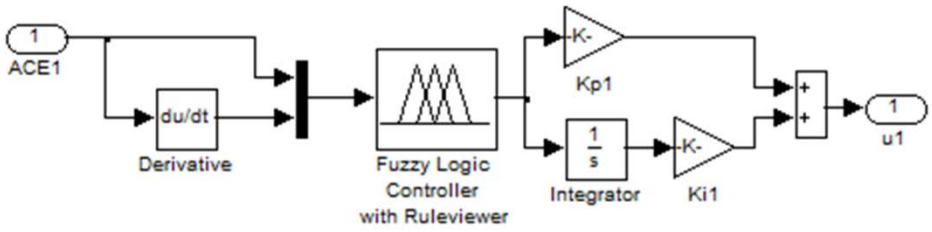


Fig. 3.21 Schematic diagram of fuzzy-PI controller.

Table 3.4 Fuzzy rule-base using Mamdani FIS for AGC of two-area power system.

Fuzzy control signal u_{flc}		Rate of change of area control error d/dt(ACE)						
		NL	NM	NS	Z	PS	PM	PL
Area control error ACE	NL	PL	PL	PL	PM	PM	PS	Z
	NM	PL	PM	PM	PM	PS	Z	NS
	NS	PL	PM	PS	PS	Z	NS	NM
	Z	PM	PM	PS	Z	NS	NM	NM
	PS	PM	PS	Z	NS	NS	NM	NL
	PM	PS	Z	NS	NM	NM	NM	NL
	PL	Z	NS	NM	NM	NL	NL	NL

3.5.3 Simulation Results and Analysis

In this section the simulation results and performance analysis of fuzzy-PI control for AGC of two-area interconnected power system are presented for both cases of linear and nonlinear system using MATLAB-SIMULINK models. The comparative performance analysis of fuzzy-PI control with conventional integral control is also presented. The integral gains of integral controllers and PI gains of fuzzy-PI controllers are tuned by trial and error method by observing the system responses to be optimal.

Considering the identical generating units in the two-area interconnected nonlinear power system, the system parameters taken in simulation are [184]:

$f^0 = 50$ Hz, $R_1 = R_2 = 2.4$ Hz/puMW, $b_1 = b_2 = 0.425$, $T_{G_1} = T_{G_2} = 0.08$ s, $T_{T_1} = T_{T_2} = 0.3$ s, $T_{P_1} = T_{P_2} = 20$ s, $K_{P_1} = K_{P_2} = 120$ Hz/puMW, $T_{12} = 0.08674$, $a_{12} = -a_{21} = 1$.

3.5.3.1 Linear System Case

For simulation of performance of integral control for linear model of AGC system of two-area interconnected power system with load disturbances in both areas simultaneously, the SIMULINK model is shown in Fig. 3.22. The step load disturbances considered in both control areas are: $\Delta P_{d_1} = 0.01$ pu in area 1, and $\Delta P_{d_2} = 0.02$ pu area 2. The tuned integral gains are: $K_{I_1} = K_{I_2} = -0.75$. The step responses of deviations in area frequencies and tie-line power for load disturbances in both areas as 1% in area 1 and 2% in area 2 respectively, using integral controller are shown in Fig. 3.23. It is observed that the deviations in area frequencies and tie-line power converge to zero slowly after vanishing initial oscillations.

For simulation of performance of fuzzy-PI control for linear model of AGC system of two-area interconnected power system with load disturbances in both areas simultaneously, the SIMULINK model is shown in Fig. 3.24. Both control areas have similar fuzzy-PI controllers with just differences in tuning of PI gains. The FLC designed using Mamdani FIS. The FLC used 49 fuzzy rules with 7 triangular MFs (NL, NM, NS, Z, PS, PM, PL) for each variable. The defuzzification method used is centroid. The universe of discourses for variables are: ACE=[-0.025 0.025], d/dt(ACE)=[-0.1 0.1], u_{flc}=[-1 1]. The tuned PI gains are: $k_{P_1} = 0.20$, $k_{I_1} = 0.60$, $k_{P_2} = 0.25$, $k_{I_2} = 0.60$. The step load disturbances considered in both control areas are: $\Delta P_{d_1} = 0.01$ pu in area 1, and $\Delta P_{d_2} = 0.02$ pu area 2. The step responses of deviations in area frequencies and tie-line power for load disturbances in both areas as 1% in area 1 and 2% in area 2 respectively, using fuzzy-PI controller are shown in Fig. 3.25. It is observed that the deviations in area frequencies and tie-line power converge to zero quickly after vanishing small initial transients. The responses are smooth and very fast.

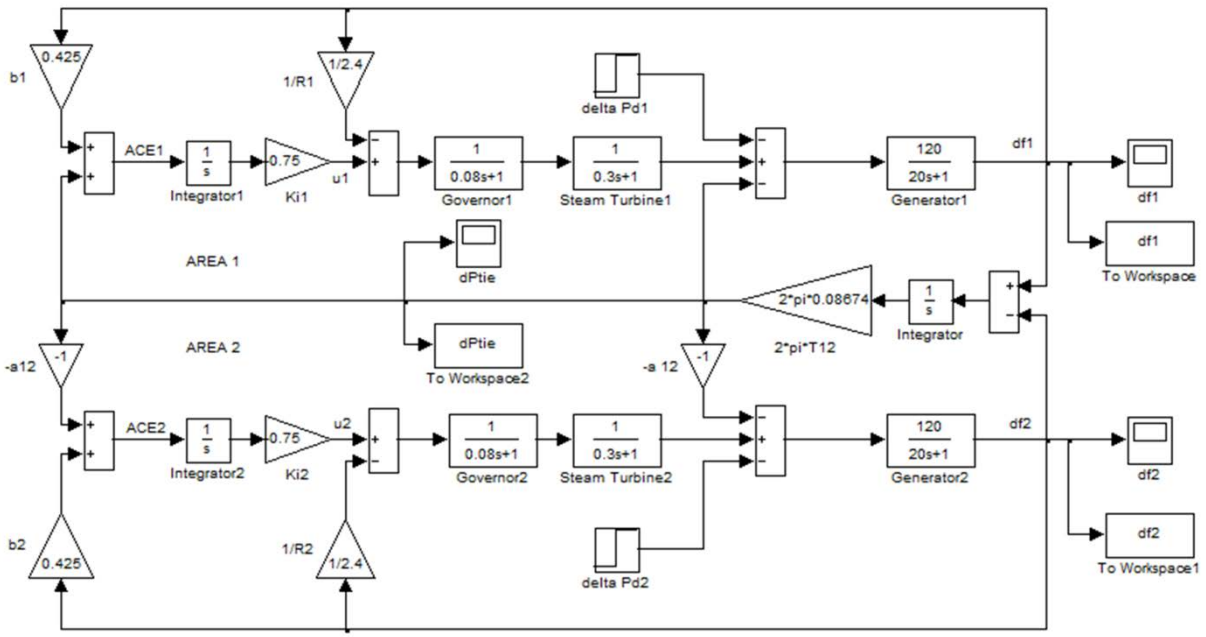


Fig. 3.22 SIMULINK model of integral control for linear model of AGC system of two-area interconnected power system.

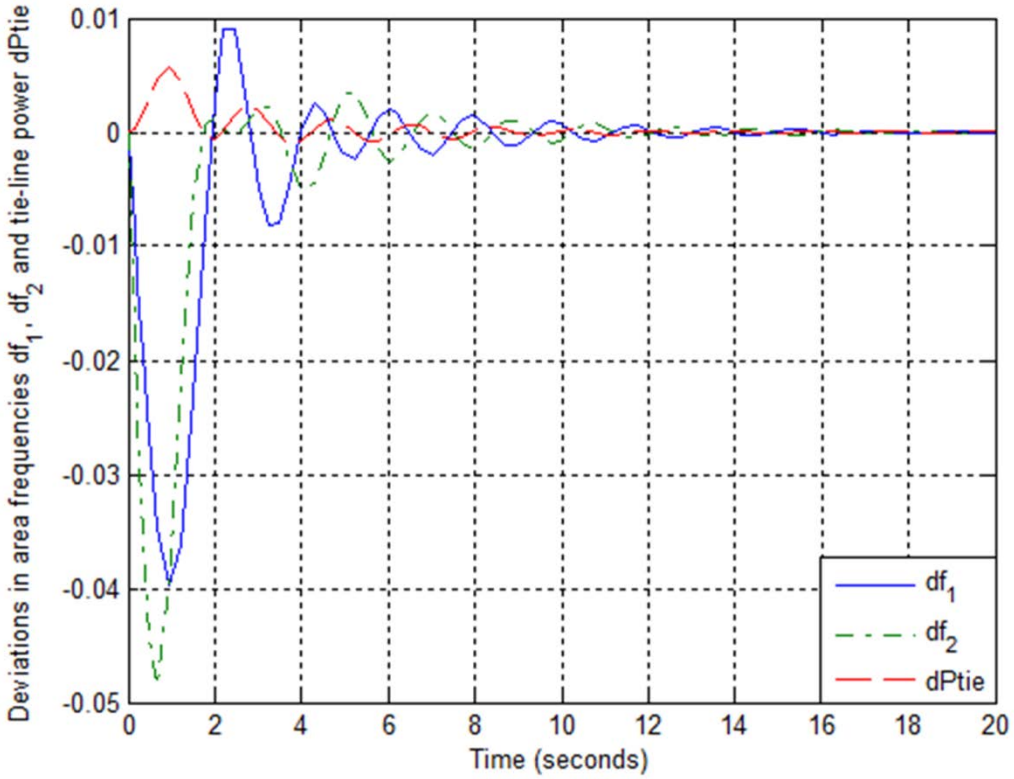


Fig. 3.23 Step responses of integral control for linear model of AGC system of two-area interconnected power system with load disturbances of 1% in area 1 and 2% in area 2.

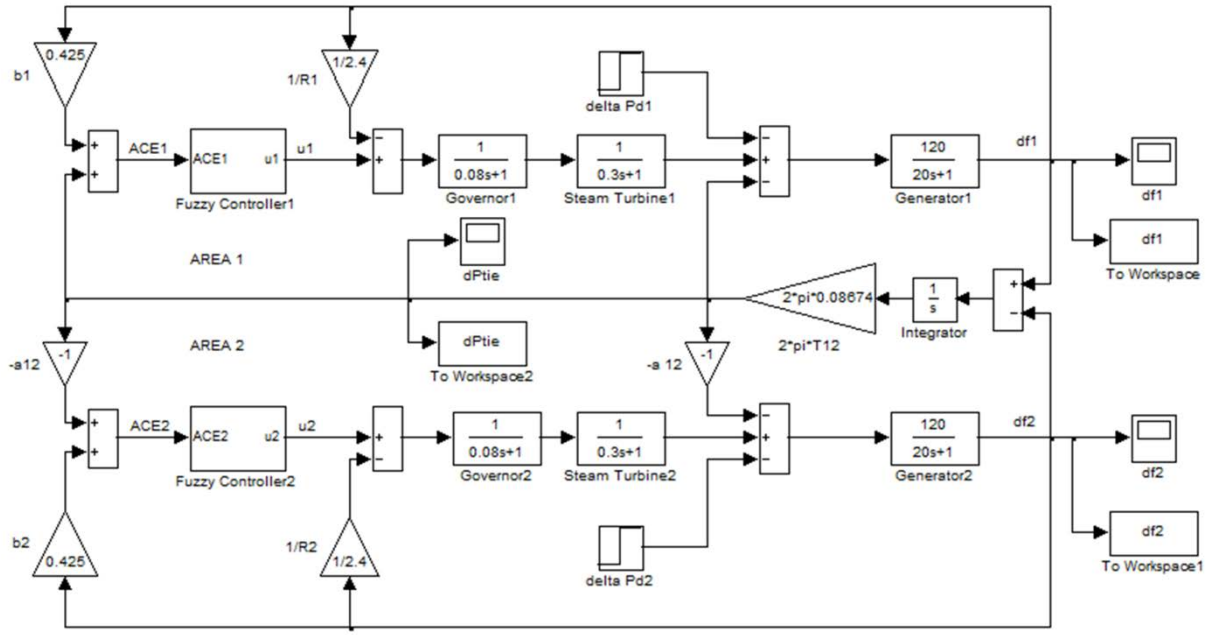


Fig. 3.24 SIMULINK model of fuzz-PI control for linear model of AGC system of two-area interconnected power system.

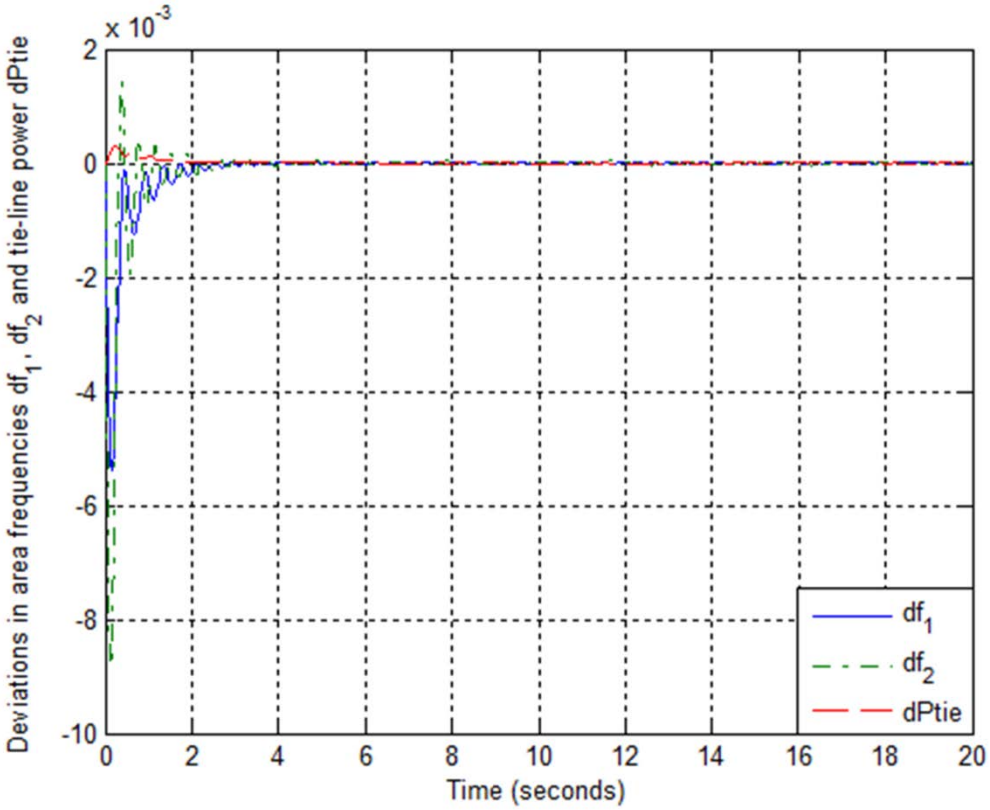


Fig. 3.25 Step responses of fuzz-PI control for linear model of AGC system of two-area interconnected power system with load disturbances of 1% in area 1 and 2% in area 2.

3.5.3.2 Nonlinear System Case

The SIMULINK model for simulation of performance of integral control for AGC system of two-area interconnected nonlinear power system with simultaneous load disturbances in both control areas is shown in Fig. 3.26. The step load disturbances considered in both control areas are: $\Delta P_{d_1} = 0.01$ pu in area 1, and $\Delta P_{d_2} = 0.02$ pu area 2. The tuned integral gains are: $K_{I_1} = K_{I_2} = -0.60$. The step responses of deviations in area frequencies and tie-line power for load disturbances in both areas as 1% in area 1 and 2% in area 2 respectively, using integral controller for nonlinear system case are shown in Fig. 3.27. It is observed that the deviations in area frequencies and tie-line power converge to zero slowly after vanishing initial oscillations.

The SIMULINK model for simulation of performance of fuzzy-PI control for AGC system of two-area interconnected nonlinear power system with simultaneous load disturbances in both control areas is shown in Fig. 3.28. In this nonlinear system case also, both control areas have similar fuzzy-PI controllers with just differences in tuning of PI gains. The FLC designed using Mamdani FIS. The FLC used 49 fuzzy rules with 7 triangular MFs (NL, NM, NS, Z, PS, PM, PL) for each variable. The defuzzification method used is centroid. The universe of discourses for variables are: ACE=[-0.025 0.025], d/dt(ACE)=[-0.1 0.1], u_{flc}=[-1 1]. The tuned PI gains are: $k_{P_1} = 0.20$, $k_{I_1} = 0.60$, $k_{P_2} = 0.30$, $k_{I_2} = 0.68$. The step load disturbances considered in both control areas are: $\Delta P_{d_1} = 0.01$ pu in area 1, and $\Delta P_{d_2} = 0.02$ pu area 2. The step responses of deviations in area frequencies and tie-line power for load disturbances in both areas as 1% in area 1 and 2% in area 2 respectively, using fuzzy-PI controller for nonlinear system case are shown in Fig. 3.29. It is observed that the deviations in area frequencies and tie-line power converge to zero quickly after vanishing small initial transients. The responses are smooth and very fast.

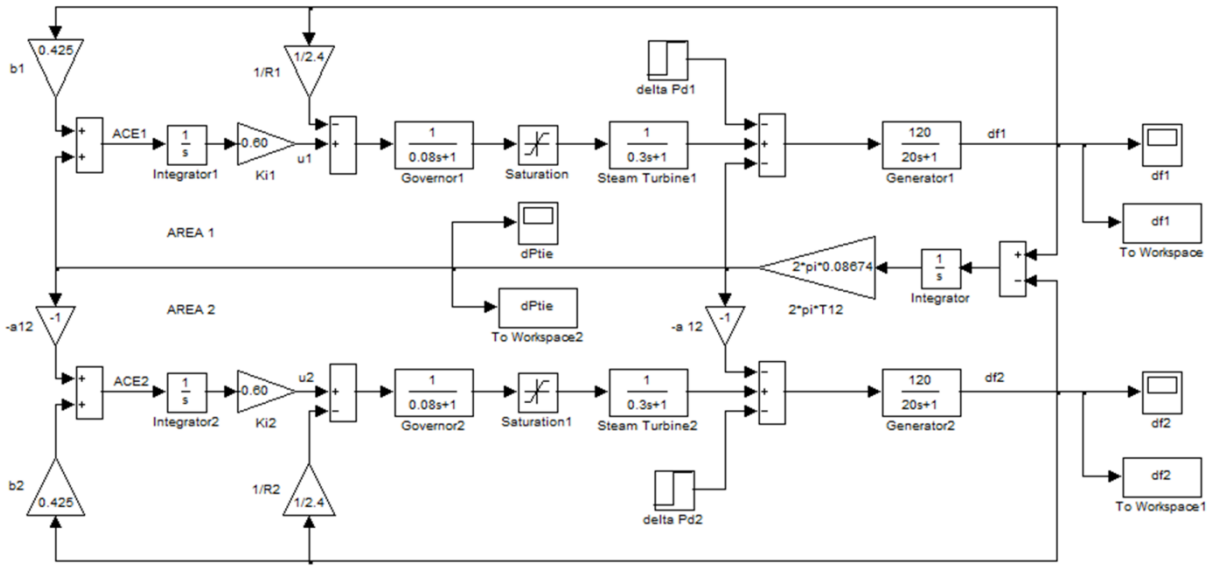


Fig. 3.26 SIMULINK model of integral control for nonlinear model of AGC system of two-area interconnected power system.

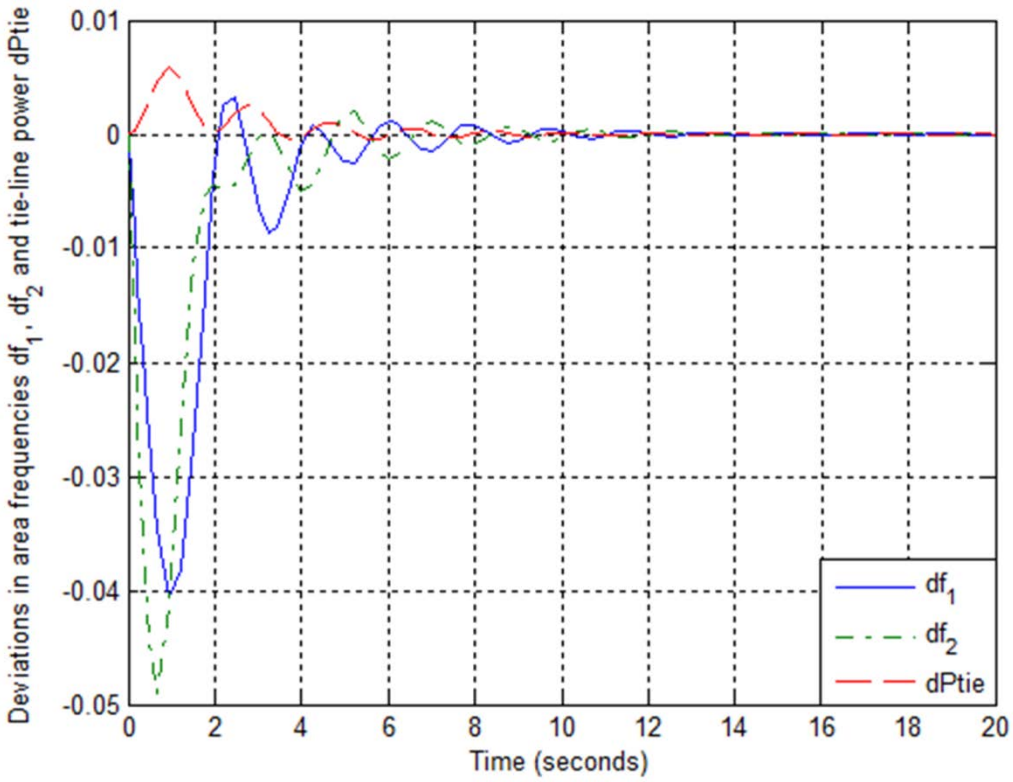


Fig. 3.27 Step responses of integral control for nonlinear model of AGC system of two-area interconnected power system with load disturbances of 1% in area 1 and 2% in area 2.

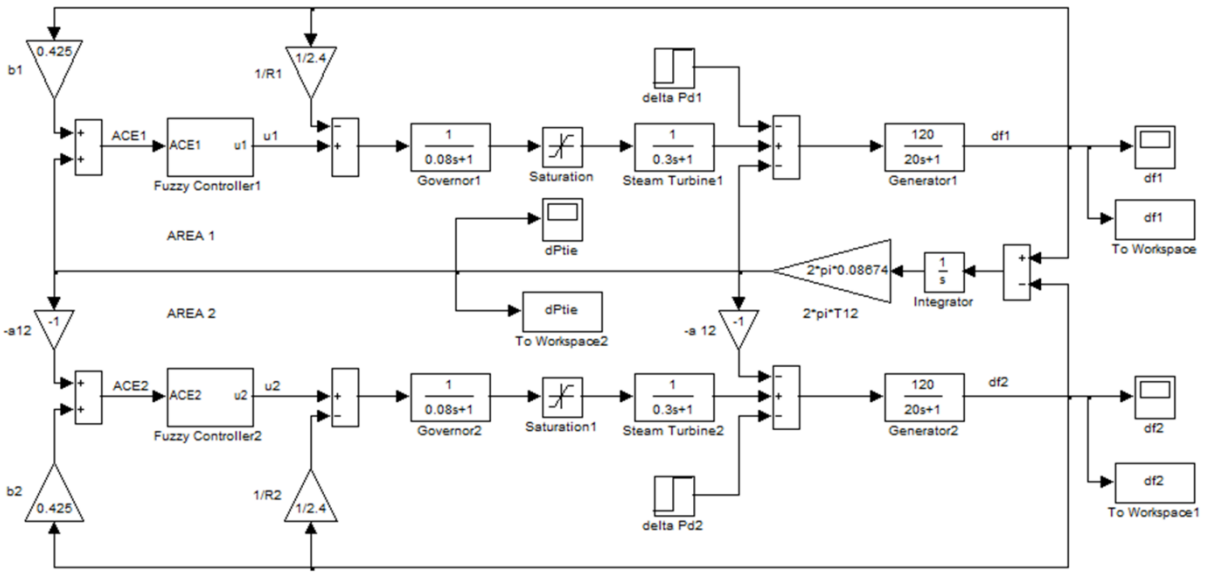


Fig. 3.28 SIMULINK model of fuzz-PI control for nonlinear model of AGC system of two-area interconnected power system.

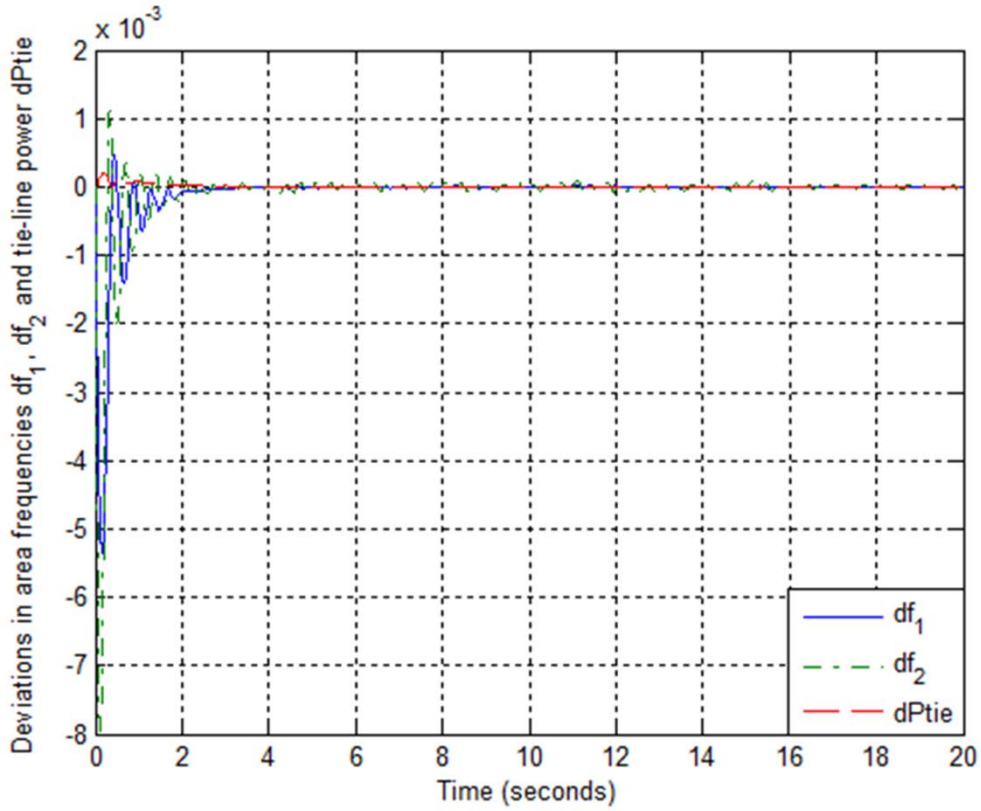


Fig. 3.29 Step responses of fuzz-PI control for nonlinear model of AGC system of two-area interconnected power system with load disturbances of 1% in area 1 and 2% in area 2.

The analysis of responses of both the integral control scheme and fuzzy-PI control scheme using Mamdani FIS for AGC system of two-area interconnected power system for both cases of linear system and nonlinear system, it is observed that the deviations in area frequencies and tie-line power converge to zero, stating that the balance of power generation

and load demand is maintained in both control areas, and the system frequency is maintained at specified value. It is observed that the integral controller give similar responses in both cases of linear or nonlinear AGC system models with changes in integral gains tuning. Similarly it is observed that the fuzzy-PI controller also give similar responses in both cases of linear or nonlinear AGC system models with changes in PI gains tuning. The comparative performance analysis of fuzzy-PI control with integral control gives that the responses of fuzzy-PI control are smoother and much faster than integral control. Both control schemes are effective & robust. Thus, it is justified that the performance of fuzzy-PI control scheme is better than the integral control scheme.

3.5.4 Discussion

The modelling and control design for automatic generation control of two-area interconnected nonlinear power system considering the governor valve position limits nonlinearity is discussed. The performance analysis of fuzzy-PI control scheme for AGC system of two-area interconnected nonlinear power system is presented. For the comparative performance analysis the integral control scheme is also presented. Also both the fuzzy-PI control scheme and integral control scheme are implemented for both cases of linear and nonlinear system models for the comparative performance analysis. For both control schemes and both cases of system models, the simulation results are presented for AGC system of two-area interconnected power systems with simultaneous load disturbances in both control areas. The FLC is designed using Mamdani FIS with uniform triangular membership functions. The simulation results justify the effectiveness and robustness of both control schemes. From the responses of both control schemes, it is observed that the deviations in area frequencies and tie-line power converge to zero. Thus, the balance of power generation and load demand is maintained in both control areas, and the system frequency is maintained at specified value. By the comparative performance analysis between fuzzy-PI control scheme and integral control scheme, it is observed that the responses of fuzzy-PI controller are smoother and much faster than integral controller, and thus, the performance of fuzzy-PI control scheme is better than the integral control scheme. Fuzzy-PI control scheme provide a simple and intelligent control method for AGC system of two-area interconnected nonlinear power systems.

3.6 CONCLUSIONS

The intelligent control is the integration of automatic control design with intelligent computational techniques. The intelligent control schemes of dynamical systems with applications using neural networks and fuzzy logic systems are discussed in this chapter. The performance analysis of neural network control scheme using radial basis function neural networks for a nonlinear process system is presented. The RBFNNs are used in

indirect adaptive control structure for this application. The neural network controller using RBFNNs gives fast tracking performance for the nonlinear process system. The intelligent control of nonlinear inverted pendulum dynamical system using fuzzy logic systems is presented. The performance analyses of fuzzy control schemes using Mamdani FIS, and TSK FIS for nonlinear inverted pendulum system for both cases of system models without and with disturbance input are presented. The comparative performance analysis between both fuzzy control schemes and PID control is also presented for this application. From the analysis of simulation results, it is observed that the response of direct FLC using TSK FIS is faster and smoother than both PD-FLC using Mamdani FIS and PID control schemes. The response of PD-FLC is better than PID control, whereas the response of TSK-FLC is better than both of these. The intelligent control using fuzzy-PI control scheme using Mamdani FIS for AGC system of two-area interconnected nonlinear power system is also presented. The conventional integral control scheme is also presented for the comparative performance analysis. Both of these control schemes are implemented for both cases of linear and nonlinear system models. From the analysis of simulation results, it is observed that the fuzzy-PI controller give smoother and much faster response than integral controller, and thus it is a simple and better control method. The neural network control and fuzzy control schemes provide the better control solution than the conventional control schemes. The neural networks and fuzzy logic systems based control schemes provide the intelligent automation for dynamical systems. The neural networks systems require the input-output data measurements for their implementation, whereas the fuzzy logic systems require the heuristics domain expert knowledge for representation of system dynamics in terms of linguistic variables. These control design techniques require a priori knowledge of the system dynamics. These control design techniques can give the online solution for the dynamical systems control problems.

CHAPTER 4

ADAPTIVE OPTIMAL CONTROL USING POLICY ITERATION TECHNIQUE FOR LTI SYSTEMS

This chapter presents a comprehensive performance analysis of adaptive optimal control of continuous-time linear time-invariant (LTI) dynamical systems with adaptive critic scheme using policy iteration (PI) technique. The convergence analysis of PI algorithm and its online implementation for adaptive optimal control of continuous-time LTI systems are discussed. The application of control scheme is presented for certain examples of continuous-time LTI systems.

4.1 INTRODUCTION

The performance of the controlled systems is desired to be optimal which should be valid also when applied in the real situation. The optimal control design techniques require the knowledge of system model, specification of performance indices, boundary conditions and physical constraints for design solution and synthesizing the control policy. The optimal control is an offline design approach which needs the complete knowledge of the system dynamics. The optimal control designed with approximate system model will not give optimal performance when applied in the real situation, as it will not be sensitive to changes in system dynamics. Thus, for optimal performance of dynamic system in real situation, the adaptation of control parameters is desired. Adaptive control has the objective of maintaining consistent performance of systems which have known structure but unknown constant or slowly time-varying parameter values. The adaptive control design techniques require a priori knowledge of system dynamics for design solution and synthesizing the control policy. The adaptive control is an online design approach which is able to deal with uncertainties, is generally not optimal in the sense of minimizing a formal performance function as specified for the optimal control. Thus, to have both features of control design, it is desired to design online adaptive optimal control. The adaptive optimal control is designed either by adding optimality features to adaptive control (e.g. the adaptation of control parameters is done by seeing the desired performance improvement reflected by an optimality criterion functional) or by adding adaptive features to optimal control (e.g. optimal control policy is improved relative to the adaptation of parameters of system model).

There are basically two ways of solving the associated optimal control problem; one is Pontryagin's minimum principle and the other is Bellman's dynamic programming (DP) [7, 42, 71]. However, the solution of Hamilton-Jacobi-Bellman (HJB) equation associated with DP has a computational complexity. The infinite horizon optimal control design using linear quadratic regulator (LQR) by offline solution of algebraic Riccati equation (ARE) and HJB equation require the complete knowledge of the system dynamics.

Policy iteration (PI) technique is a computational intelligence technique that belongs to reinforcement learning (RL) algorithms [123-126]. The online PI technique provides an online adaptive optimal control solution for an infinite horizon problem subject to the real-time dynamics of a continuous-time system. Based on actor-critic structure, PI algorithms consist of two-step iteration: policy evaluation and policy improvement. Instead of solving HJB equation by direct approach, the PI algorithm starts by evaluating the cost of a given initial admissible control policy, which is often accomplished by solving a nonlinear Lyapunov equation. This updated cost is then used to obtain an updated improved control policy which will have a lower associated cost. This is often accomplished by minimizing a Hamiltonian function with respect to the updated cost [124, 125, 155, 156, 158-160, 166]. This is the so-called `greedy policy' with respect to the updated cost [158]. These two steps of policy evaluation and policy improvement are repeated until the policy improvement step no longer changes the actual policy and thus converging to the optimal control. It is noted that the infinite horizon cost can be evaluated only in the case of admissible and stabilizing control policies. Admissibility is in fact a condition for the control policy which is used to initialize the algorithm [158]. PI algorithm requires an initial stabilizing control policy, but value iteration (VI) algorithm does not require an initial stabilizing control policy ing control policy, but VI does not require an initial stabilizing control policy [124, 125, 158]. The PI technique based adaptive critic scheme performs adaptive optimal control without using complete knowledge of the system dynamics. The knowledge of systems internal dynamics (i.e. matrix A) is not needed for evaluation of cost or the update of control policy; only the knowledge of input matrix B is required for updating the control policy. Thus this control scheme becomes partially model-free. The online PI algorithm solves the optimal control problem along a single state trajectory, which does not require knowledge of the system internal dynamics, and thus can be viewed as a direct adaptive optimal control technique. Unlike the regular adaptive controllers which rely on online identification of system dynamics followed by model based controller design, the PI method relies on identification of cost function associated with a given control policy followed by policy improvement in sense of minimizing the identified cost.

The Adaptive optimal control using PI technique is presented for linear systems in [155-158, 161, 162, 166, 169], and for nonlinear systems using neural networks in actor-critic configuration in [123, 158-165, 167]. Even though adaptive critic is successfully implemented in several real-life problems; the performance investigation of the adaptive critic control scheme with practical applications is not explored much in the literature. Even certain recent papers are appeared on adaptive critic designs and PI technique with certain applications, the comprehensive analysis with practical applications is much desired.

In this chapter the performance analysis of adaptive optimal control using PI technique for continuous-time LTI systems is presented comprehensively. The application of PI based control scheme is implemented for certain general and practical examples of LTI systems- general LTI SISO system, higher-order LTI system- a mechanical system, load frequency control (LFC) of power system, automatic voltage regulator (AVR) of power system, and DC motor speed control system.

4.2 INFINITE HORIZON OPTIMAL CONTROL OF CONTINUOUS-TIME LTI SYSTEMS

The infinite horizon optimal control (i.e. linear quadratic regulator (LQR)) problem using policy iteration technique for continuous-time LTI systems without using the complete knowledge of system internal dynamics is presented in this section [71, 144, 155, 156].

Consider the continuous-time linear time-invariant dynamical system described by

$$\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \tag{4.1}$$

where $\boldsymbol{x}(t) \in \mathcal{R}^n$, $u(t) \in \mathcal{R}^m$ and $(A,\ B)$ is stabilizable, subject to the optimal control problem

$$u^*(t) = \arg \min_{\substack{u(t) \\ t_0 \le t \le \infty}} V(t_0, x(t_0), u(t)) \tag{4.2}$$

where the infinite horizon quadratic cost function to be minimized is expressed as

$$V(x(t_0), t_0) = \int_{t_0}^{\infty} (x^T(\tau)Qx(\tau) + u^T(\tau)Ru(\tau))d\tau \tag{4.3}$$

with $Q \ge 0,\ R > 0$ and $(Q^{1/2}, A)$ detectable.

The solution of this optimal control problem, determined by Bellman's optimality principle, is given by

$$\boldsymbol{u}(t) = -K\boldsymbol{x}(t) \qquad \text{with } K = R^{-1}B^T P \tag{4.4}$$

where the matrix *P* is the unique positive definite solution of the Algebraic Riccati Equation (ARE)

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \tag{4.5}$$

Equation (4.4) gives a stabilizing closed loop controller determined from the unique positive semi-definite solution of ARE under the detectability condition. Here to solve (4.5), both system matrix A and control input matrix B must be known i.e. complete knowledge of the system dynamics is required. Due to this reason, the development of algorithms that will converge to solution of optimization problem without performing prior system identification and using explicit models of system dynamics is needed from the control systems point of view.

4.3 CONTINUOUS-TIME ADAPTIVE CRITICS

In adaptive critic design (ACD), (4.4) and (4.3) or (4.5) are represented by two parametric function approximation networks namely action network (actor) and critic network (critic) respectively. The action network which provides control signals represents the relationship between state and input. The critic network which learns the desired performance index for some performance index/cost function represents the relationship between state and costate vector. The critic evaluates the performance of actor and the actor is improved based on the feedback from the critic network. These two functional networks approximating HJB equation which leads to ARE for linear systems, successively adapt to determine the optimal control solution for a system. In general, ACD uses incremental optimization combined with a parametric structure to efficiently approximate the optimal cost and control. In ACD a short-term cost metric is optimized that ensures optimization of the cost over all future times. ACDs function as supervised learning systems and reinforcement learning systems [124-126, 143].

4.4 POLICY ITERATION TECHNIQUE

In this section, online policy iteration technique which gives optimal control solution of the LQR problem, without using knowledge of the system internal dynamics (i.e. system matrix A) is presented. It gives an adaptive controller which converges to the state feedback optimal controller. The policy iteration technique [124-126, 155, 156, 158-160, 166]] is based on an actor-critic structure, consists of two-step iteration- critic update and actor update. For a given stabilizing controller critic computes the associated infinite horizon cost. The actor computes the control policy and is represented by its parameters (i.e. feedback controller gain) [155, 156].

4.4.1 Policy Iteration Algorithm

Let a stabilizing gain *K* for (1), under the assumption that (A, B) is stabilizable, such that $\dot{\boldsymbol{x}} = (A - BK)\boldsymbol{x}$ is a stable closed loop system. Then the corresponding infinite horizon quadratic cost is given by

$$V(\boldsymbol{x}(t)) = \int_t^{\infty} \boldsymbol{x}^T(\tau)(Q + K^T RK)\boldsymbol{x}(\tau)d\tau = \boldsymbol{x}^T(t)P\boldsymbol{x}^T(t) \tag{4.6}$$

where *P* is the real symmetric positive definite solution of the Lyapunov matrix equation

$$(A - BK)^T P + P(A - BK) = -(K^T RK + Q) \tag{4.7}$$

and $V(\boldsymbol{x}(t))$ serves as a Lyapunov function for (1) with controller gain *K*. The cost function (4.6) can be written as

$$V(\boldsymbol{x}(t)) = \int_t^{t+T} \boldsymbol{x}^T(\tau)(Q + K^T RK)\boldsymbol{x}(\tau)d\tau + V(\boldsymbol{x}(t+T)) \tag{4.8}$$

Based on (4.8), denoting $\boldsymbol{x}(t)$ with $\boldsymbol{x}_t$, with parameterization $V(\boldsymbol{x}_t) = \boldsymbol{x}_t^T P \boldsymbol{x}_t$ and considering an initial stabilizing control gain K_1, following two-step online policy iteration algorithm can be implemented:

1. Policy evaluation

$$\boldsymbol{x}_t^T P_i \boldsymbol{x}_t = \int_t^{t+T} \boldsymbol{x}_\tau^T (Q + K_i^T R K_i) \boldsymbol{x}_\tau d\tau + \boldsymbol{x}_{t+T}^T P_i \boldsymbol{x}_{t+T} \tag{4.9}$$

2. Policy improvement

$$K_{i+1} = R^{-1} B^T P_i \tag{4.10}$$

Equations (4.9) and (4.10) formulate a new policy iteration algorithm. It is important to note that this algorithm does not require system matrix *A* for its solution, only control input matrix *B* must be known for updating *K*.

4.4.2 Convergence Analysis

The convergence of policy iteration algorithm is discussed in this subsection referring the lemmas, remarks and theorems in [155, 156].

Let $A_i = A - BK_i$, then for the system $\dot{\boldsymbol{x}} = A_i \boldsymbol{x}$, a Lyapunov function may be $V_i(\boldsymbol{x}_t) = \boldsymbol{x}_t^T P_i \boldsymbol{x}_t$, $\forall \boldsymbol{x}_t$, and

$$\frac{d(\boldsymbol{x}_t^T P_i \boldsymbol{x}_t)}{dt} = \boldsymbol{x}_t^T (A_i^T P_i + P_i A_i) \boldsymbol{x}_t = -\boldsymbol{x}_t^T (K_i^T R K_i + Q) \boldsymbol{x}_t \tag{4.11}$$

then, $\forall T > 0$, from (4.11) we may have

$$\int_t^{t+T} \boldsymbol{x}_t^T (Q + K_i^T R K_i) \boldsymbol{x}_t d\tau = -\int_t^{t+T} \frac{d(\boldsymbol{x}_t^T P_i \boldsymbol{x}_t)}{dt} d\tau = \boldsymbol{x}_t^T P_i \boldsymbol{x}_t - \boldsymbol{x}_{t+T}^T P_i \boldsymbol{x}_{t+T} \tag{4.12}$$

which is same as (4.9). From (4.11) the Lyapunov equation is

$$A_i^T P_i + P_i A_i = -(K_i^T R K_i + Q) \tag{4.13}$$

For a stabilizing control policy K_i the matrix A_i is stable and $K_i^T R K_i + Q > 0$ then there exists a unique solution of the Lyapunov equation (4.13), $P_i > 0$. Thus if A_i is asymptotically stable, the solution of (4.9) is the unique solution of (4.13), and thus both (4.9) and (4.13) are equivalent. Although the same solution is obtained whether solving (4.13) or (4.9), (4.9) can be solved without using any knowledge on the system matrix *A*. Thus, PI algorithm (4.9) & (4.10) is equivalent to iterating between (4.13) & (4.10), without using knowledge of the system internal dynamics, if A_i is stable at each iteration.

Let the control policy K_i is stabilizing with the associated cost $V_i(\boldsymbol{x}_t) = \boldsymbol{x}_t^T P_i \boldsymbol{x}_t$. For the state trajectories generated while using the controller K_{i+1}, take the positive definite cost

function $V_i(\boldsymbol{x}_t)$ as a Lyapunov function candidate. Taking the derivative of $V_i(\boldsymbol{x}_t)$ along the trajectories generated by K_{i+1} one obtains

$$\begin{aligned}\dot{V}_i(\boldsymbol{x}_t) &= \boldsymbol{x}_t^T[P_i(A-BK_{i+1})+(A-BK_{i+1})^T P_i]\boldsymbol{x}_t \\ &= \boldsymbol{x}_t^T[P_i(A-BK_i)+(A-BK_i)^T P_i]\boldsymbol{x}_t + \boldsymbol{x}_t^T[P_iB(K_i-K_{i+1})+(K_i-K_{i+1})^T B^T P_i]\boldsymbol{x}_t\end{aligned} \quad (4.14)$$

Using the policy update given by (4.10) and completing the squares, the second term can be written as

$$\boldsymbol{x}_t^T[K_{i+1}^T R(K_i-K_{i+1})+(K_i-K_{i+1})^T RK_{i+1}]\boldsymbol{x}_t = \boldsymbol{x}_t^T[-(K_i-K_{i+1})^T R(K_i-K_{i+1})-K_{i+1}^T RK_{i+1}+K_i^T RK_i]\boldsymbol{x}_t$$

Using (4.13) the first term in the equation can be written as $-\boldsymbol{x}_t^T[K_i^T RK_i+Q]\boldsymbol{x}_t$ and summing up the two terms one obtains

$$\dot{V}_i(\boldsymbol{x}_t) = -\boldsymbol{x}_t^T[(K_i-K_{i+1})^T R(K_i-K_{i+1})]\boldsymbol{x}_t - \boldsymbol{x}_t^T[Q+K_{i+1}^T RK_{i+1}]\boldsymbol{x}_t \quad (4.15)$$

Thus, under the initial assumptions from the problem setup $Q \geq 0, R > 0$, $V_i(\boldsymbol{x}_t)$ is a Lyapunov function proving that the updated control policy $\boldsymbol{u} = -K_{i+1}\boldsymbol{x}$ is stabilizing with K_{i+1} given by (4.10), and thus, if (4.10) is used for updating the control policy then the new control policy will be stabilizing. Thus it is concluded that if the initial control policy given by K_1 is stabilizing, then all policies obtained using the iteration (4.9)-(4.10) will be stabilizing policies.

Let $Ric(P_i)$ be the matrix valued function defined as

$$Ric(P_i) = A^T P_i + P_i A - P_i BR^{-1}B^T P_i + Q \quad (4.16)$$

and let Ric'_{P_i} denote the Fréchet derivative of $Ric(P_i)$ taken with respect to P_i. The matrix function Ric'_{P_i} evaluated at a given matrix M will thus be

$$Ric'_{P_i}(M) = (A-BR^{-1}B^T P_i)^T M + M(A-BR^{-1}B^T P_i).$$

Equations (4.11) and (4.10) can be compactly written as

$$A_i^T P_i + P_i A_i = -(P_{i-1}BR^{-1}B^T P_{i-1} + Q) \quad (4.17)$$

Subtracting $A_i^T P_{i-1} + P_{i-1}A_i$ on both sides gives

$$A_i^T(P_i - P_{i-1}) + (P_i - P_{i-1})A_i = -(P_{i-1}A + A^T P_{i-1} - P_{i-1}BR^{-1}B^T P_{i-1} + Q) \quad (4.18)$$

which is Newton's method

$$P_i = P_{i-1} - (Ric'_{P_{i-1}})^{-1} Ric(P_{i-1}) \quad (4.19)$$

Thus, the iteration between (4.9) and (4.10) is equivalent to Newton's method formulation (4.19) by use of the introduced notations $Ric(P_i)$ and Ric'_{P_i}. Newton's method, i.e. iteration (4.13) & (4.10), conditioned by an initial stabilizing policy will converge to the solution of ARE. And, if the initial policy is stabilizing, all the subsequent control policies will be stabilizing. This proven equivalence between (4.13) & (4.10), and (4.9) & (4.10), shows

that the online policy iteration algorithm will converge to the solution of the optimal control problem (4.2) with the infinite horizon quadratic cost (4.3) without using knowledge of the internal dynamics of the controlled system (4.1). Thus, under the assumptions of stabilizability of $(A,\ B)$ and detectability of $(Q^{1/2}, A)$, with $Q \geq 0, R > 0$ in the cost index (4.2), the policy iteration (4.9) & (4.10), conditioned by an initial stabilizing controller, converges to the optimal control solution given by (4.3) where the matrix *P* satisfies the ARE (4.4).

Thus the only requirement for convergence to the optimal controller consists in an initial stabilizing policy that will guarantee a finite value for the cost $V_1(\boldsymbol{x}_t) = \boldsymbol{x}_t^T P_1 \boldsymbol{x}_t$. Under the assumption that the system to be controlled is stabilizable and implementation of an optimal state feedback controller is possible and desired. It is reasonable to assume that a stabilizing (though not optimal) state feedback controller is available to begin the iteration. In fact in many cases the system to be controlled is itself stable such that the initial controller can be chosen as zero.

4.5 ADAPTIVE OPTIMAL CONTROL USING POLICY ITERATION TECHNIQUE

This section presents the online implementation of policy iteration algorithm based adaptive optimal control without using knowledge of the system internal dynamics. The implementation of PI algorithm only needs knowledge of *B* matrix which explicitly appears in (4.10). The system matrix *A* is not required for computation of either of two steps of PI algorithm, as that information is embedded in the states $\boldsymbol{x}(t)$ and $\boldsymbol{x}(t+T)$ which are observed online.

Associated with the policy K_i, to find the critic parameters (matrix P_i) of the cost function in (4.9), the term $\boldsymbol{x}^T(t) P_i \boldsymbol{x}(t)$ is written as

$$\boldsymbol{x}^T(t) P_i \boldsymbol{x}(t) = \overline{p}_i^T \overline{\boldsymbol{x}}(t) \tag{4.20}$$

where $\overline{\boldsymbol{x}}(t)$ denotes the Kronecker product quadratic polynomial basis vector with the elements $\{x_i(t) x_j(t)\}_{i=1,n;\, j=i,n}$ and $p = v(P)$ with $v(\bullet)$ a vector valued matrix function that acts on symmetric matrices and returns a column vector by stacking the elements of the diagonal and upper triangular part of the symmetric matrix into a vector where the off-diagonal elements are taken as $2P_{ij}$ [155, 156]. Using (4.20), (4.9) is rewritten as

$$\overline{p}_i^T (\overline{\boldsymbol{x}}(t) - \overline{\boldsymbol{x}}(t+T)) = \int_t^{t+T} \boldsymbol{x}^T(\tau)(Q + K_i^T R K_i)\boldsymbol{x}(\tau) d\tau \tag{4.21}$$

In this equation $\overline{p}_i$ is the vector of unknown parameters and $\overline{\boldsymbol{x}}(t) - \overline{\boldsymbol{x}}(t+T)$ acts as a regression vector. The right hand side target function is denoted by $d(\overline{\boldsymbol{x}}(t), K_i)$ (also known as the reinforcement on the time interval $[t, t+T]$),

$$d(\bar{\boldsymbol{x}}(t), K_i) \equiv \int_t^{t+T} \boldsymbol{x}^T(\tau)(Q + K_i^T R K_i)\boldsymbol{x}(\tau)d\tau$$

is measured based on the system states over the time interval $[t, t+T]$. System (4.1) is augmented by introducing a new state $V(t)$ defined as $\dot{V}(t) = \boldsymbol{x}^T(t)Q\boldsymbol{x}(t) + \boldsymbol{u}^T(t)R\boldsymbol{u}(t)$, and so the value of $d(\bar{\boldsymbol{x}}(t), K_i)$ can be measured by taking two measurements of this newly introduced system state since $d(\bar{\boldsymbol{x}}(t), K_i) = V(t+T) - V(t)$. This new state signal is the output of an analog integration block having as inputs the quadratic terms $\boldsymbol{x}^T(t)Q\boldsymbol{x}(t)$ and $\boldsymbol{u}^T(t)\mathrm{R}\boldsymbol{u}(t)$ which can also be obtained using an analog processing unit.

The parameter vector $\bar{\mathrm{p}}_i$ of the function $V_i(\boldsymbol{x}_t)$ (i.e. critic), which will then yield the matrix P_i, is found by minimizing, in the least-squares sense, the error between the target function, $d(\bar{\boldsymbol{x}}(t), K_i)$, and the parameterized left hand side of (4.21). Evaluating the right hand side of (4.21) at $N \geq n(n+1)/2$ (the number of independent elements in the matrix P_i) points $\bar{\boldsymbol{x}}^i$ in the state space, over the same time interval *T*, the least-squares solution is obtained as

$$\bar{p}_i = (\boldsymbol{X}\boldsymbol{X}^T)^{-1}\boldsymbol{X}\boldsymbol{Y} \tag{4.22}$$

where

$$\boldsymbol{X} = [\bar{\boldsymbol{x}}_\Delta^1 \;\; \bar{\boldsymbol{x}}_\Delta^2 \;\; ... \;\; \bar{\boldsymbol{x}}_\Delta^N],\; \bar{\boldsymbol{x}}_\Delta^i = \bar{\boldsymbol{x}}^i(t) - \bar{\boldsymbol{x}}^i(t+T),\; \boldsymbol{Y} = [d(\bar{\boldsymbol{x}}^1, K_i) \;\; d(\bar{\boldsymbol{x}}^2, K_i) \;\; ...d(\bar{\boldsymbol{x}}^N, K_i)]^T$$

The least-squares problem can be solved in real-time after a sufficient number of data points are collected along a single state trajectory, under the regular presence of an excitation requirement. Alternatively, (4.22) can be solved also using recursive estimation algorithms (e.g. gradient descent algorithms or the recursive least squares (RLS) algorithm) in which case a persistence of excitation condition is required. Due to this reason there are no real issues related to the algorithm becoming computationally expensive with the increase of the state space dimension [155, 156].

In the case in which the cost function (4.9) is solved for in a single step (e.g. using a method such as the exact least-squares described by (4.22)), the online algorithm has the same quadratic convergence speed as Newton's method. For the case in which the solution of (4.9) is obtained iteratively, the convergence speed of the online algorithm will decrease. In this case at each step in the PI algorithm (which involves solving (4.9) & (4.10)) a recursive gradient descent algorithm, which most often has exponential convergence, will be used for solving (4.10). Thus it is resolved that the convergence speed of the online algorithm will depend on the chosen technique for solving (4.9). Even the convergence property of online algorithm is not affected by the value of sample time *T*; it affects the excitation condition necessary in the setup of a numerically well posed least squares problem

and obtaining the least squares solution (4.22). More precisely, assuming without loss of generality that the matrix $\boldsymbol{X}$ in (4.22) is square, and letting $\varepsilon > 0$ be a desired lower bound on the determinant of $\boldsymbol{X}$, then the chosen sampling time T must satisfy

$$T > \frac{a\varepsilon}{\prod_{I=1}^{n} \left| \lambda_I (A_c) \right|} \tag{4.23}$$

where λ_I denotes the eigenvalues of closed loop system and $a > 0$ is a scaling factor. From this point of view a minimal insight relative to the system dynamics would be required for choosing the sampling time T [155, 156].

The online PI algorithm requires only measurements of states at discrete moments in time, t and $t+T$, as well as knowledge of observed cost over time interval $[t, t+T]$, which is $d(\bar{\boldsymbol{x}}(t), K_i)$. Therefore, knowledge of system matrix A is not required for the cost evaluation or the control policy update, and only the matrix B is required for the control policy update, using (4.10), which makes the tuning algorithm only partially model-free. The PI algorithm converges to optimal control solution measuring cost along a single state trajectory, provided that there is enough initial excitation in the system. Since the algorithm iterates only on stabilizing policies which will make the system states go to zero, sufficient excitation in the initial state of the system is necessary. In the case that excitation is lost prior to obtaining the convergence (system reaches the equilibrium point) a new experiment needs to be conducted having as a starting point the last policy from the previous experiment. In this case, the control policy is updated at time $t+T$, after observing the state $\boldsymbol{x}(t+T)$ and it is used for controlling the system during the time interval $[t+T, t+2T]$. The critic stops updating the control policy when the difference between the system performances evaluated at two consecutive steps crosses below a designer specified limit, i.e. the algorithm has converged to the optimal controller. Also in the case that this error is bigger than this specified limit the critic again starts tuning the actor parameters to obtain an optimal control policy. If there is a sudden change in system dynamics described by the matrix A as long as the present controller is stabilizing for the new matrix A, the algorithm will converge to the solution to the corresponding new ARE. Thus the algorithm is suitable for online implementation from the control theory point of view.

Fig. 4.1 [155, 156] shows the schematic block diagram of adaptive optimal control with actor-critic structure for LTI system. Since the system is augmented with an extra state $V(t)$ that is part of the adaptive critic control scheme thus this controller is actually a dynamic controller with the cost state. This adaptive optimal controller has a hybrid structure with a continuous-time internal state followed by a sampler and discrete-time update rule. The application of the proposed control scheme is presented in the following section 5.

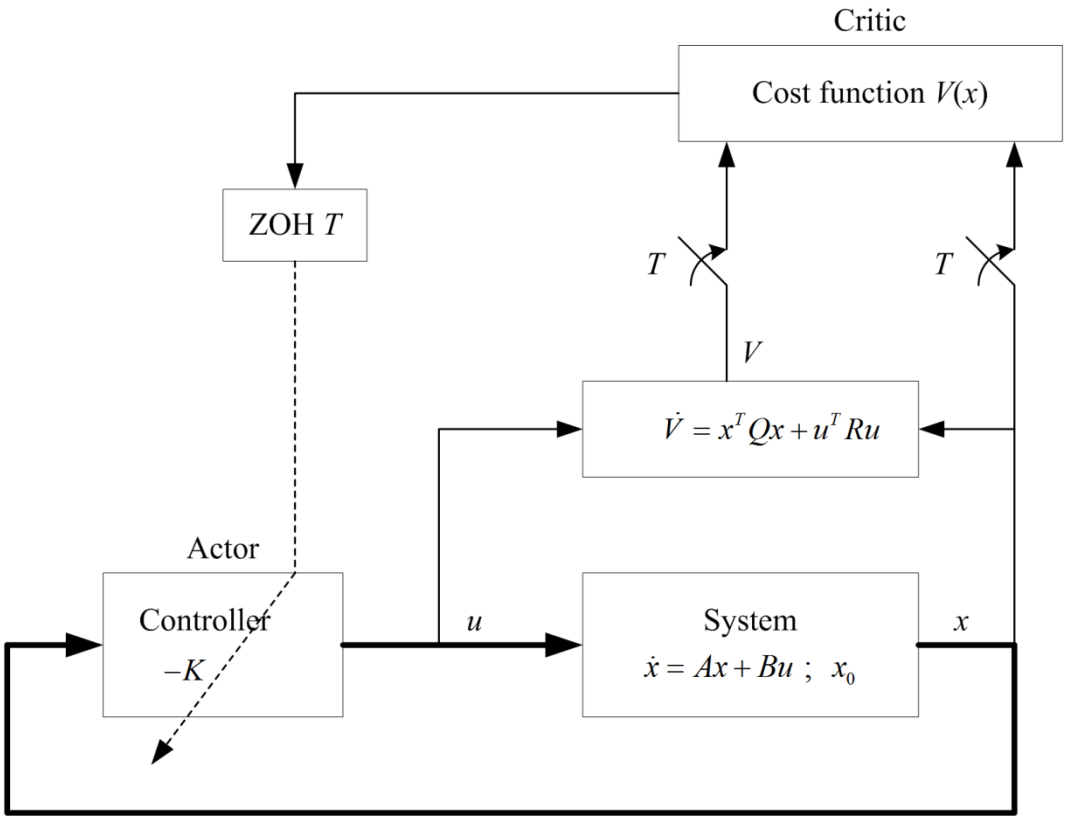


Fig. 4.1 Adaptive optimal control with actor-critic structure.

4.6 SIMULATION RESULTS AND ANALYSIS

This section presents the systems modeling, simulation results and performance analysis to demonstrate the application of adaptive optimal control using online PI technique considering general and practical examples of LTI systems- general LTI SISO system, higher-order LTI system- a mechanical system, load frequency control (LFC) of power system, automatic voltage regulator (AVR) of power system, and DC motor speed control system.

4.6.1 General LTI SISO System

Consider a general LTI SISO dynamical system described by the transfer function [7]

$$G(s) = \frac{Y(s)}{U(s)} = \frac{s+3}{s^3 + 9s^2 + 24s + 20} \tag{4.24}$$

The state space model of this system is given as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -20 & -24 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \quad \text{and} \quad y = \begin{bmatrix} 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

For system (4.24), in implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = [0.1, 0.2, 0.1, 0]$;

$P = [0\ 0\ 0; 0\ 0\ 0;\ 0\ 0\ 0]$. The length of the simulation in samples is taken 60, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = \begin{bmatrix} 2.2818 & 1.5196 & 0.0250 \\ 1.5196 & 2.2882 & 0.0840 \\ 0.0250 & 0.0840 & 0.0647 \end{bmatrix}$$

$$P = \begin{bmatrix} 2.2818 & 1.5196 & 0.0250 \\ 1.5196 & 2.2882 & 0.0840 \\ 0.0250 & 0.0840 & 0.0647 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK = [0.0250 \quad 0.0840 \quad 0.0647]$$

$$K = [0.0250 \quad 0.0840 \quad 0.0647]$$

The eigenvalues of closed loop system are obtained as

-1.9662 + 0.1893i, -1.9662 - 0.1893i, -5.1323

The simulation responses using PI technique for LTI system (4.24) are shown in Figs. 2 to 6. Fig. 4.2 shows the system state trajectories which converge towards the equilibrium point. Fig. 4.3 shows the control signal trajectory which also converges towards zero. Fig. 4.4 shows the evolution of closed loop poles of the system during simulation. Fig. 4.5 shows the convergence of critic parameters of matrix *P* towards optimal values. Fig. 4.6 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update.

Simulation with change in system parameter is also done at sample *k*=21; (i.e. *t*=1.05 seconds), such that *A*(3,1)=-22, then the solution is obtained as

$$RicP = \begin{bmatrix} 2.5057 & 1.6406 & 0.0227 \\ 1.6406 & 2.3494 & 0.0890 \\ 0.0227 & 0.0890 & 0.0652 \end{bmatrix}$$

$$P = \begin{bmatrix} 2.2818 & 1.5196 & 0.0250 \\ 1.5196 & 2.2882 & 0.0840 \\ 0.0250 & 0.0840 & 0.0647 \end{bmatrix}$$

$$RicK = [0.0227 \quad 0.0890 \quad 0.0652]$$

$$K = [0.0250 \quad 0.0840 \quad 0.0647]$$

The eigenvalues of closed loop system are obtained as

-1.8771 + 0.7899i, -1.8771 - 0.7899i, -5.3106

Fig. 4.7 shows the evolution of closed loop poles of the system during simulation with change in system parameter at sample *k*=21. The remaining responses obtained are same as above in this case. Fig. 4.8 presents the closed loop response of LTI system (4.24) using both approaches of LQR and adaptive critic (AC) using PI technique. It remains exactly the same also for case with change in system parameters. It is observed here that the adaptive optimal controller using PI technique gives the similar response as of standard LQR.

It is observed from the above simulation results that critic parameter matrix *P* and actor parameter *K* obtained using PI technique are converging adaptively to optimal values and are of same values of *RicP* and *RicK* respectively that obtained from LQR approach. Also in case of change in the system parameter in real situation the controller adapts it and converges to same optimal values. Thus the actor *K* and critic *P* parameters remain unchanged. The analysis of the above simulation results demonstrates that the proposed adaptive optimal control scheme is partially model-free, effective & robust.

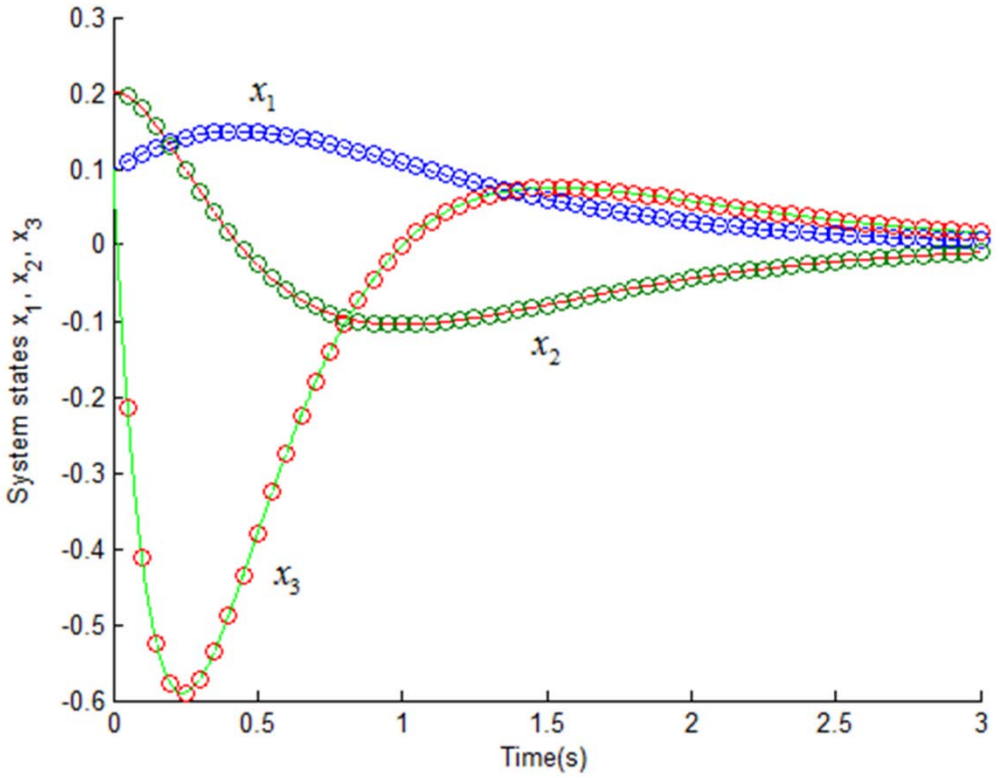


Fig. 4.2 System states.

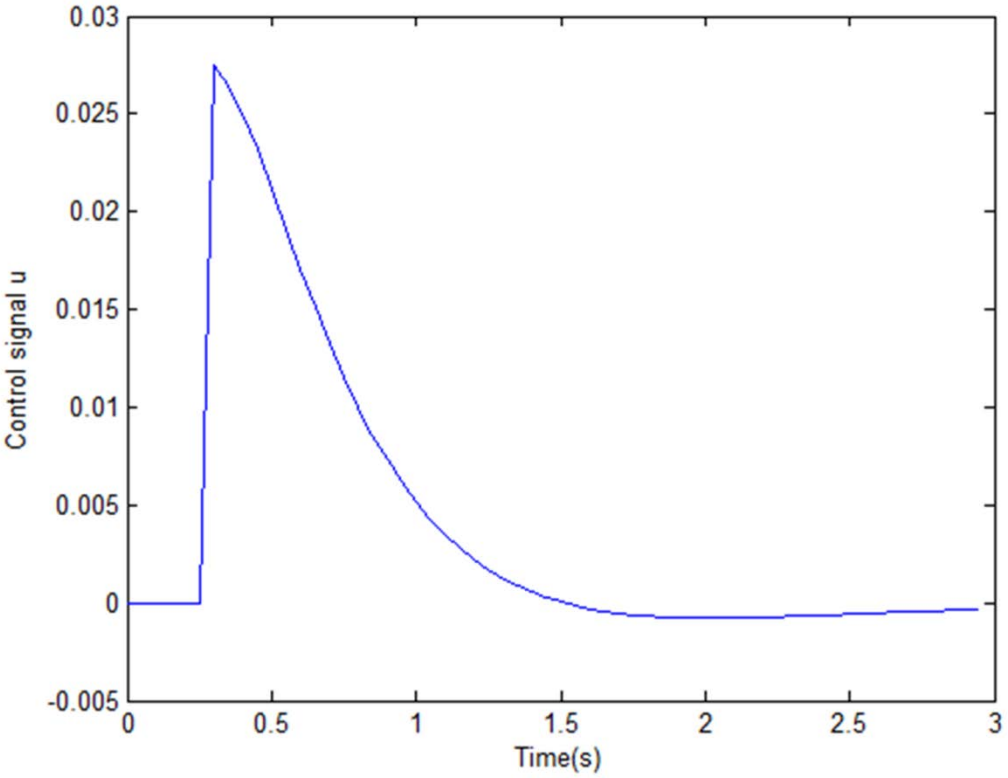


Fig. 4.3 Control signal.

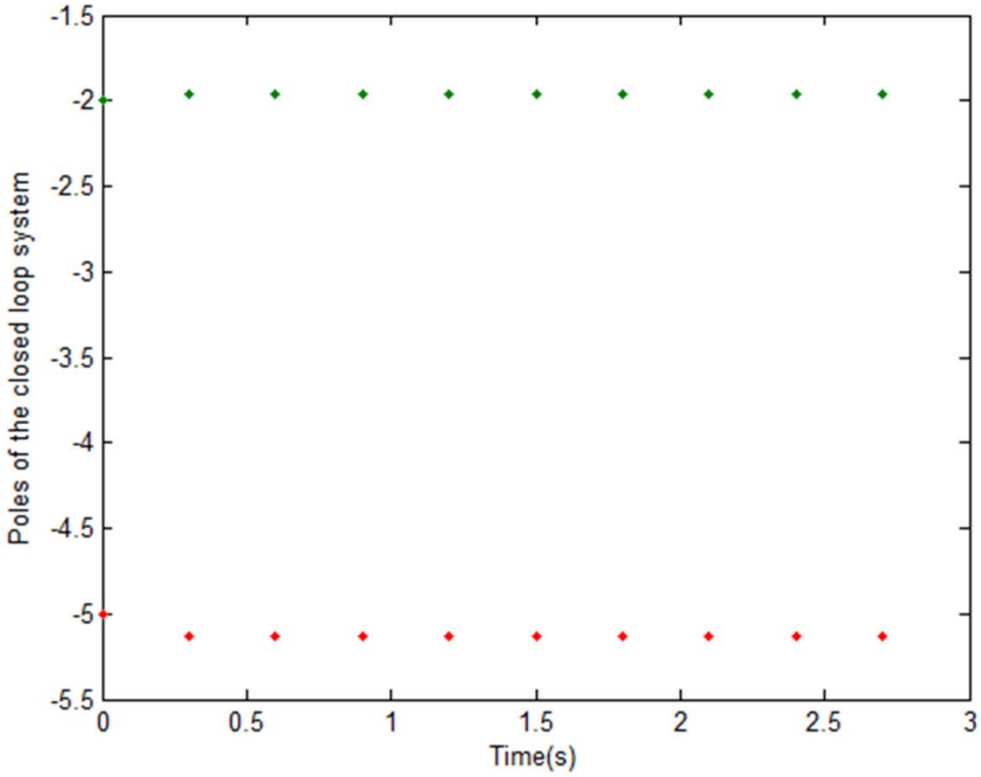


Fig. 4.4 Evolution of poles of closed loop system.

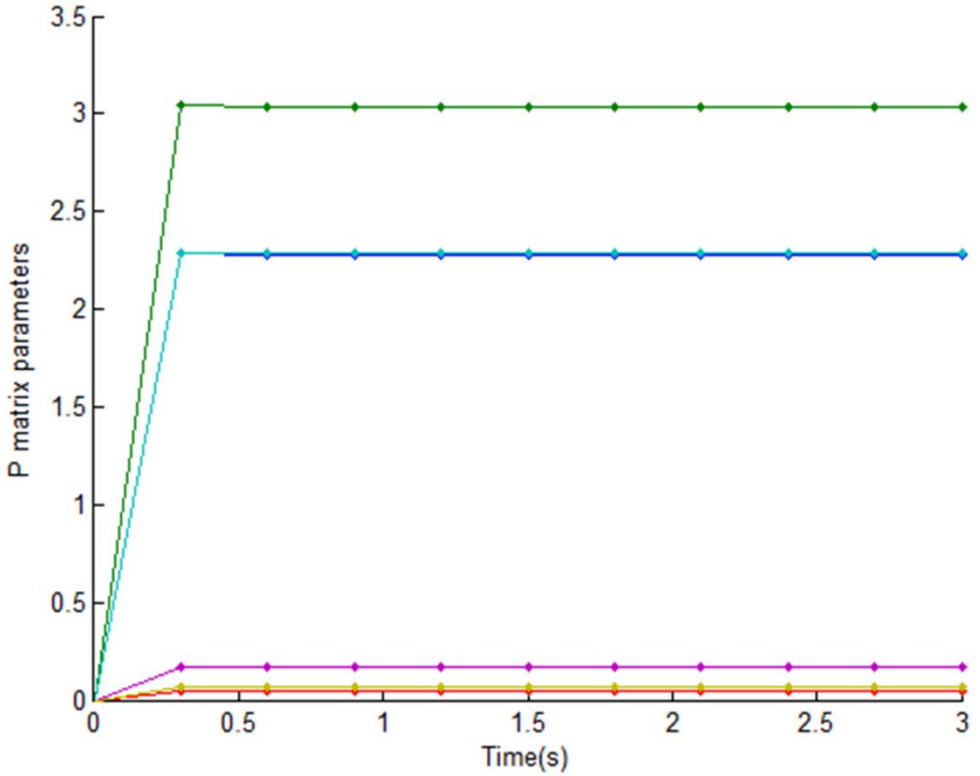


Fig. 4.5 Critic parameters.

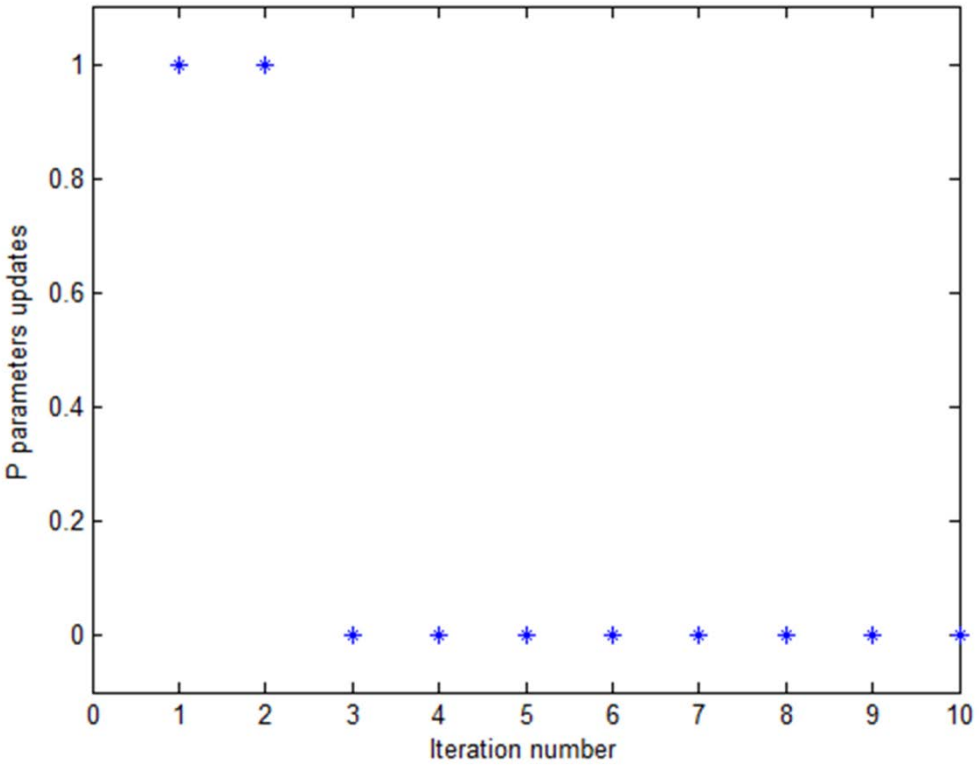


Fig. 4.6 Updating of critic parameters.

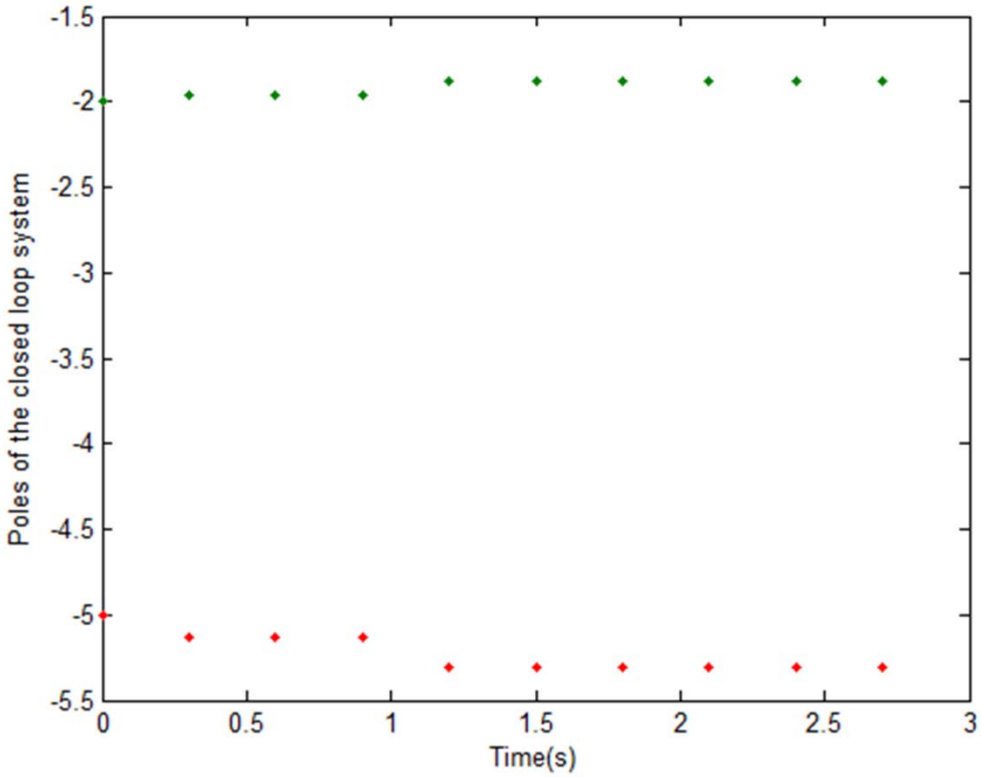


Fig. 4.7 Evolution of poles of closed loop system with change in system parameters at k=21.

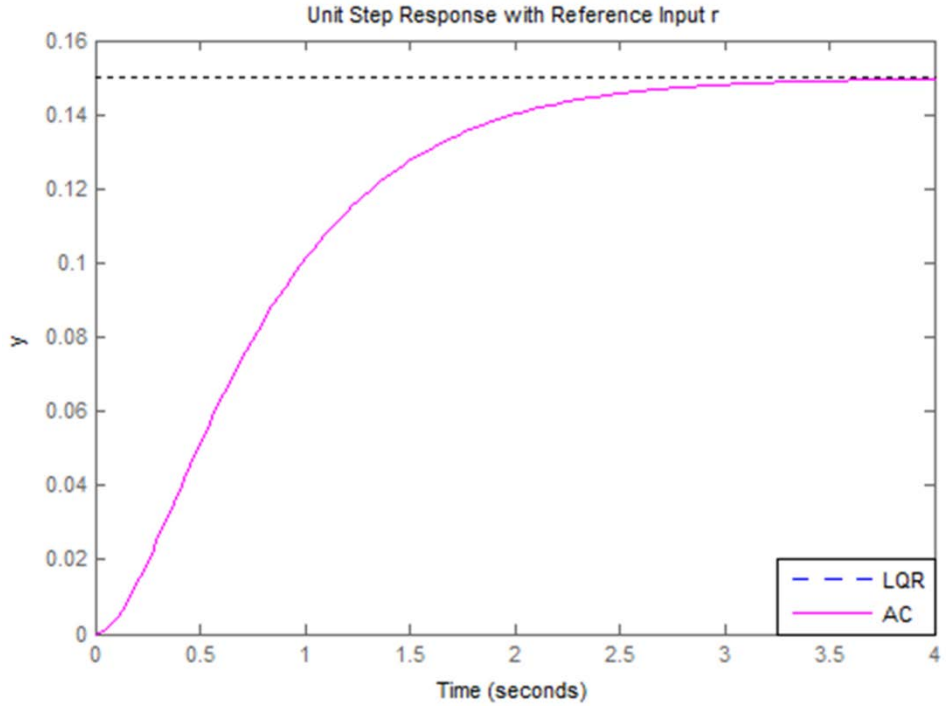


Fig. 4.8 Unit step response of closed loop system.

4.6.2 General Higher Order LTI System- A Mechanical System

Consider a mechanical system shown in Fig. 4.9 as a practical example of higher order LTI system [1]. The system is at rest for t<0, At t=0, a step force f of α newton is applied to mass m_2. The force $f = \alpha u$, where u is a step force of 1 newton. The displacements y_1 and y_2 are measured from the respective equilibrium positions of the carts before f is applied. Assume the system parameters: $m_1 = 10$ kg, $m_2 = 20$ kg, $b = 20$ N-s/m, $k_1 = 30$ N/m, $k_2 = 60$ N/m, and $\alpha = 10$. The dynamic equations for this system are written as

$$m_1\ddot{y}_1 = -k_1 y_1 - k_2(y_1 - y_2) - b(\dot{y}_1 - \dot{y}_2) \tag{4.25}$$

$$m_2\ddot{y}_2 = -k_2(y_2 - y_1) - b(\dot{y}_2 - \dot{y}_1) + \alpha u \tag{4.26}$$

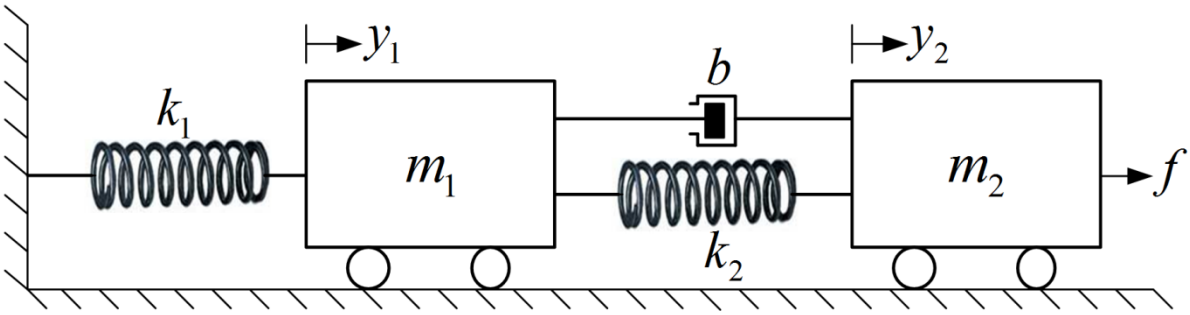


Fig. 4.9 Mechanical system

Considering the state variables as $x_1 = y_1$, $x_2 = \dot{y}_1$, $x_3 = y_2$, $x_4 = \dot{y}_2$, and the output variables as y_1, y_2, the state space model of the system is written as

$$\begin{aligned} \dot{X} &= AX + Bu \\ y &= CX + Du \end{aligned} \tag{4.27}$$

where,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1 + k_2}{m_1} & -\frac{b}{m_1} & \frac{k_2}{m_1} & \frac{b}{m_1} \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{m_2} & \frac{b}{m_2} & -\frac{k_2}{m_2} & -\frac{b}{m_2} \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{\alpha}{m_2} \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

At the above system parameters values we have

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -9 & -2 & 6 & 2 \\ 0 & 0 & 0 & 1 \\ 3 & 1 & -3 & -1 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.5 \end{bmatrix}$$

For this higher order LTI system which is defined by (4.27), in implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as

$x_0 = [0.5, 0.2, 0.1, 0, 0]$; $P = \begin{bmatrix} 0 & 0 & 0 & 0; & 0 & 0 & 0 & 0; & 0 & 0 & 0 & 0; & 0 & 0 & 0 & 0 \end{bmatrix}$.

The length of the simulation in samples is taken 120, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = \begin{bmatrix} 2.3433 & -0.2224 & -0.2824 & -0.8069 \\ -0.2224 & 0.5692 & 0.5492 & 0.9811 \\ -0.2824 & 0.5492 & 3.2418 & 1.2046 \\ -0.8069 & 0.9811 & 1.2046 & 2.7331 \end{bmatrix}$$

$$P = \begin{bmatrix} 2.3443 & -0.2224 & -0.2827 & -0.8072 \\ -0.2224 & 0.5693 & 0.5492 & 0.9813 \\ -0.2827 & 0.5492 & 3.2425 & 1.2048 \\ -0.8072 & 0.9813 & 1.2048 & 2.7337 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$RicK = \begin{bmatrix} -0.4034 & 0.4906 & 0.6023 & 1.3666 \end{bmatrix}$, $K = \begin{bmatrix} -0.4036 & 0.4907 & 0.6024 & 1.3669 \end{bmatrix}$

The eigenvalues of closed loop system are obtained as

-1.4796 + 2.9795i, -1.4796 - 2.9795i, -0.3621 + 0.9042i, -0.3621 - 0.9042i

The simulation responses using PI technique for higher order LTI system (4.27) are shown in Figs. 4.10 to 4.14. Fig. 4.10 shows the system state trajectories which converge towards the equilibrium point. Fig. 4.11 shows the control signal trajectory which also converges towards zero. Fig. 4.12 shows the evolution of closed loop poles of the system during simulation. Fig. 4.13 shows the convergence of critic parameters of matrix *P* towards optimal values. Fig. 4.14 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update. Fig. 4.15 presents the closed loop response of higher order LTI system (4.27) using both approaches of LQR and adaptive critic (AC) using PI technique. It is observed here that the adaptive optimal controller using PI technique gives the similar responses as of standard LQR.

Simulation with change in system parameter is also done at sample *k*=41; (i.e. *t*=2.05 seconds), as $k_1 = 40$ N/m, such that *A*(2,1)=-10, then the solution is obtained as

$$RicP = \begin{bmatrix} 2.6885 & -0.2612 & -0.3383 & -0.9959 \\ -0.2612 & 0.4903 & 0.5106 & 0.8273 \\ -0.3383 & 0.5106 & 3.3735 & 1.1342 \\ -0.9959 & 0.8273 & 1.1342 & 2.5046 \end{bmatrix}$$

$$P = \begin{bmatrix} 2.3443 & -0.2224 & -0.2827 & -0.8072 \\ -0.2224 & 0.5693 & 0.5492 & 0.9813 \\ -0.2827 & 0.5492 & 3.2425 & 1.2048 \\ -0.8072 & 0.9813 & 1.2048 & 2.7337 \end{bmatrix}$$

and the actor gains *RicK* and *K* respectively are obtained as

$$RicK = \begin{bmatrix} -0.4980 & 0.4136 & 0.5671 & 1.2523 \end{bmatrix},\ K = \begin{bmatrix} -0.4036 & 0.4907 & 0.6024 & 1.3669 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-1.4546 + 3.1006i, -1.4546 - 3.1006i, -0.3872 + 1.0133i, -0.3872 - 1.0133i

The simulation responses using PI technique for the case of change in system parameter are similar as above. In this case Fig. 4.16 shows the evolution of closed loop poles adapting the controller. For the case of change in parameters of system (4.27) Fig. 4.17 shows the closed loop response using both approaches of LQR and adaptive critic (AC) using PI technique. It is observed here also that the adaptive optimal controller using PI technique gives the similar responses as of standard LQR. The controller performs adapting the change in system parameters.

It is observed in the above simulation results that critic parameter matrix *P* and actor parameter *K* obtained using PI technique are converging adaptively to optimal values and are mostly of same values of *RicP* and *RicK* respectively that obtained from LQR approach. Also in case of change in the system parameter in real situation the controller adapts it and converges to same optimal values. Thus the actor *K* and critic *P* parameters remain unchanged. The analysis of the above simulation results demonstrates that the proposed adaptive optimal control scheme is partially model-free, effective & robust.

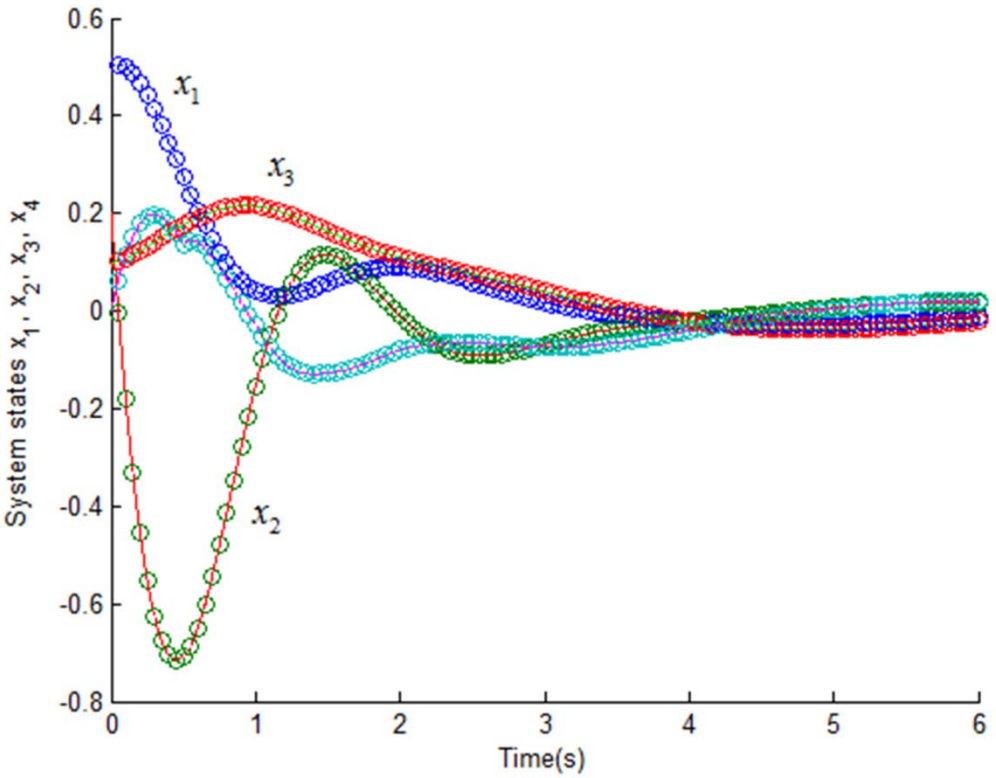


Fig. 4.10 System states.

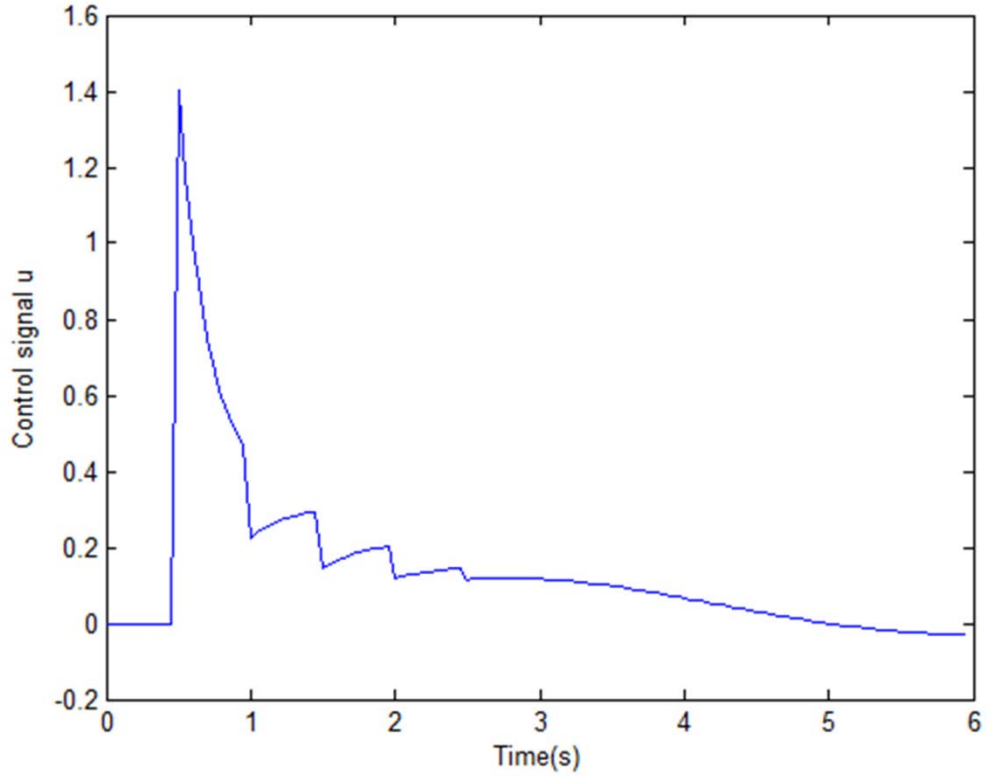


Fig. 4.11 Control signal.

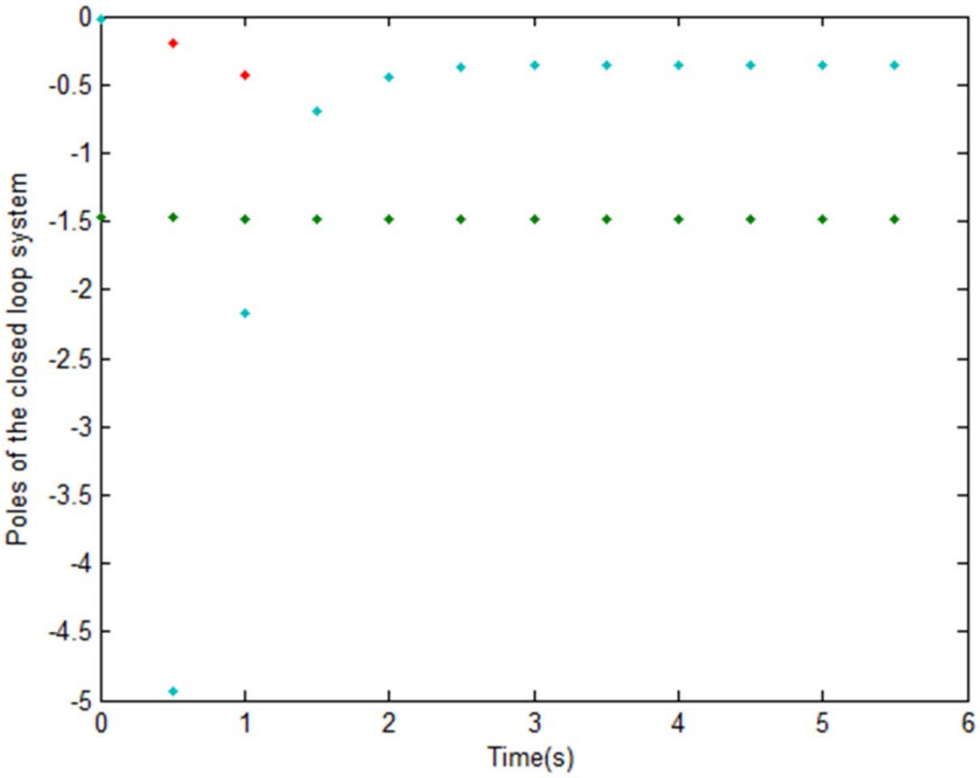


Fig. 4.12 Evolution of poles of closed loop system.

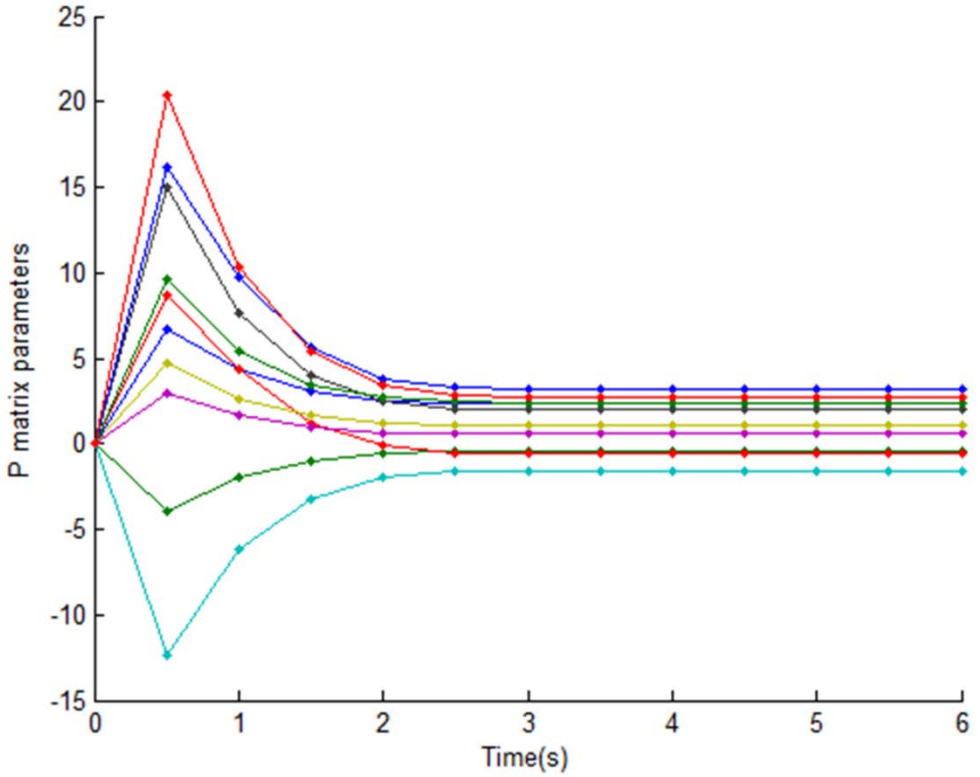


Fig. 4.13 Critic parameters.

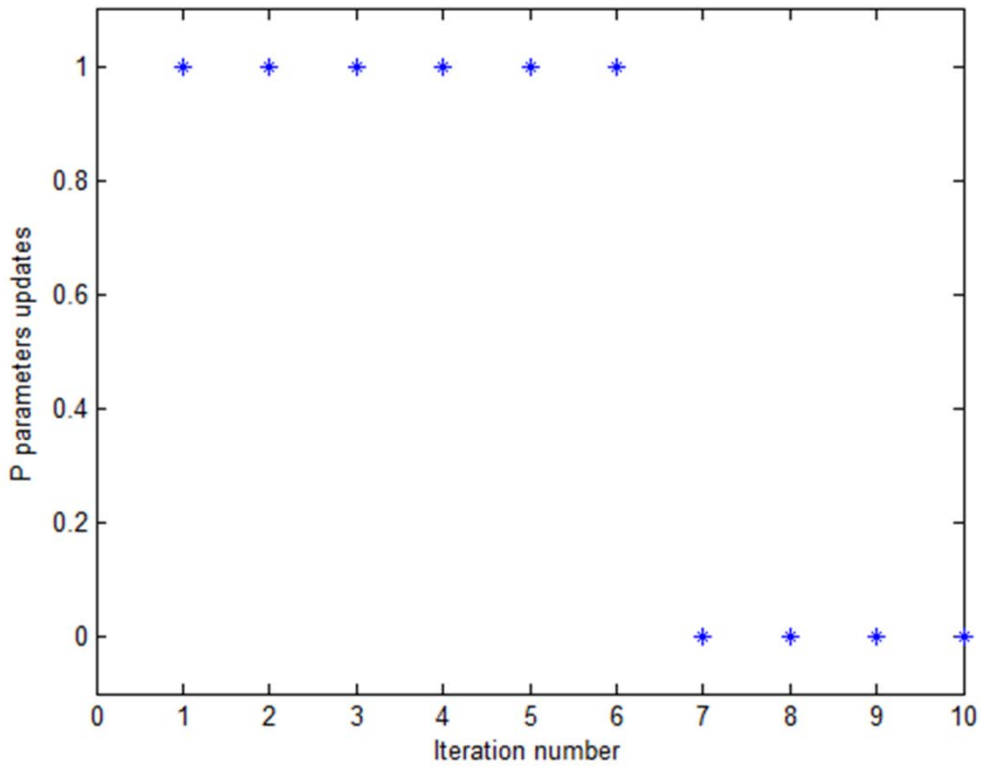


Fig. 4.14 Updating of critic parameters.

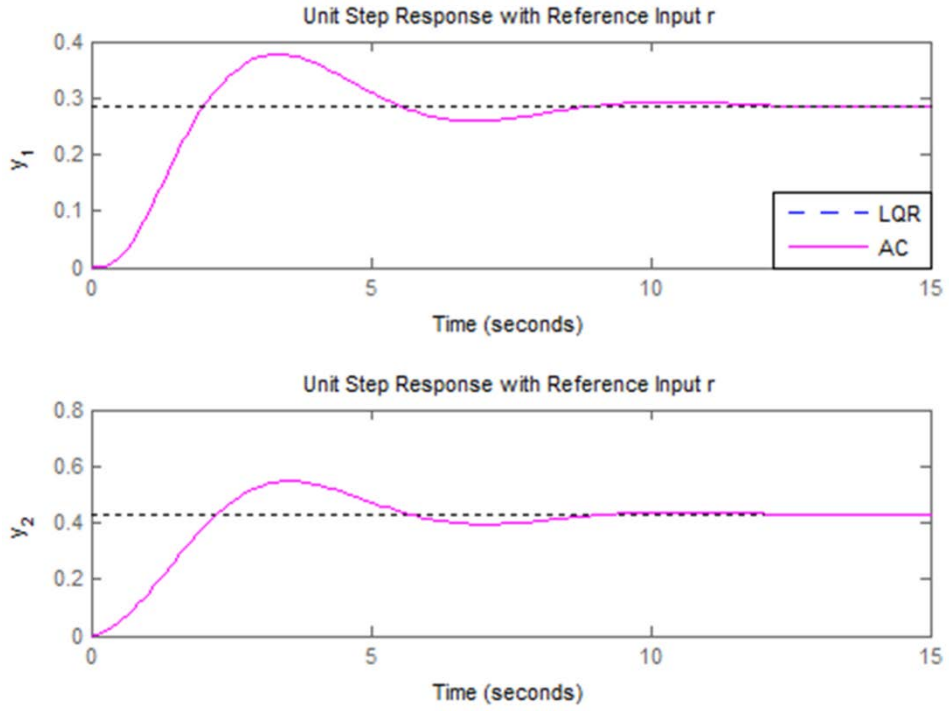


Fig. 4.15 Unit step responses of closed loop system.

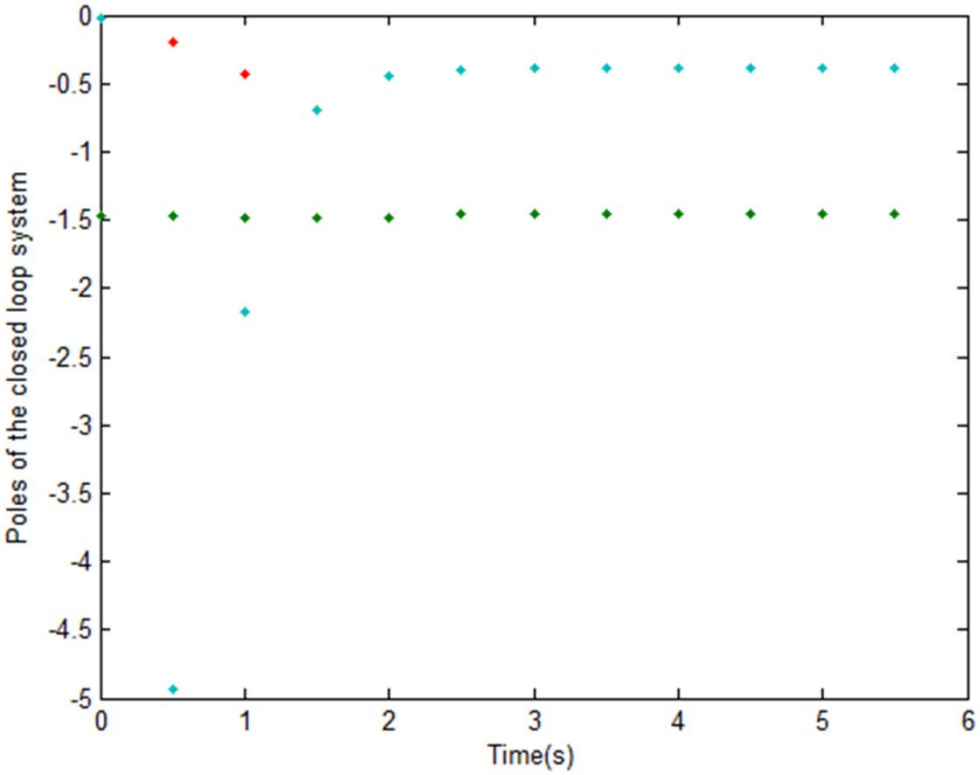


Fig. 4.16 Evolution of closed loop poles with change in system parameters at k=41.

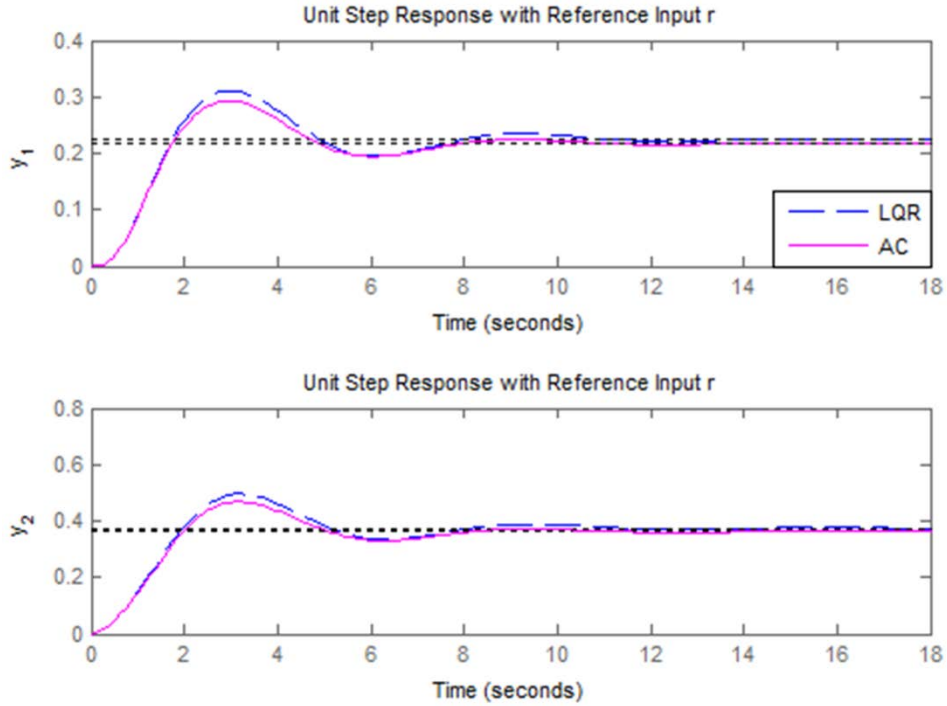


Fig. 4.17 Unit step responses of closed loop system with change in system parameters at k=41.

4.6.3 Load Frequency Control System

The active power and frequency control is known as load frequency control (LFC), which plays an important function of power system operation where the main objective is to regulate the output power of each generator at prescribed levels while keeping the frequency fluctuations with in predetermined limits [182-185]. In the large scale interconnected power systems the system frequency and the inter-area tie-line power are required to be near to the scheduled values as possible. LFC is required to be robust to the unknown external disturbances and system model and parameter uncertainties. Many control strategies for power systems load frequency control have been presented in literature [26, 28, 98, 99, 182-185, 188-198].

In this section the incremental linear system models of LFC without and with integral control both are considered to investigate the performance of proposed control scheme also under the system's structural change. The performance is also investigated for change in system parameters at certain instant of time to demonstrate that the proposed approach is partially model-free.

4.6.3.1 LFC System model without integral control

The functional block diagram of single-area power system load frequency control without integral control model is shown in Fig. 4.18 [26, 28, 182]

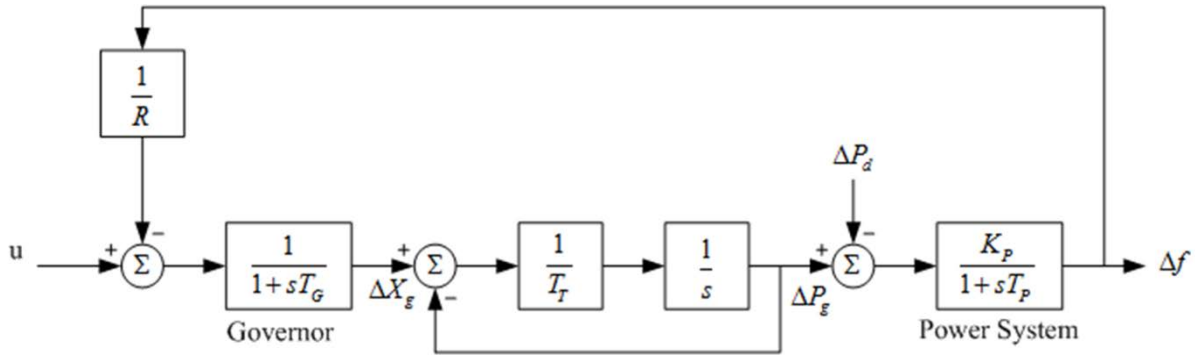


Fig. 4.18 Block diagram of load frequency control of power system.

The state space model of the system is derived as following

$$\dot{x} = Ax(t) + Bu(t) + F\Delta P_d(t) \tag{4.28}$$

$$y = \Delta f(t) = Cx(t) \tag{4.29}$$

where state vector $x(t) = [\Delta f(t) \quad \Delta P_g(t) \quad \Delta X_g(t)]^T$; $\Delta f(t)$ is incremental frequency deviation (Hz); $\Delta P_g(t)$ is incremental change in generator output (p.u. MW); $\Delta X_g(t)$ is incremental change in governor value position (p.u. MW); $\Delta P_d(t)$ is load disturbance (p.u. MW); R is

speed regulation due to governor action (Hz. p.u. MW^{-1}); T_G is governor time constants (s); T_T is turbine time constants (s); T_P is plant model constants (s); K_P is plant gain; and

$$A=\begin{bmatrix} -\frac{1}{T_P} & \frac{K_P}{T_P} & 0 \\ 0 & -\frac{1}{T_T} & \frac{1}{T_T} \\ -\frac{1}{RT_G} & 0 & -\frac{1}{T_G} \end{bmatrix},\ B=\begin{bmatrix} 0 \\ 0 \\ \frac{1}{T_G} \end{bmatrix},\ F=\begin{bmatrix} -\frac{K_P}{T_P} \\ 0 \\ 0 \end{bmatrix},\ C=\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

The range of LFC system parameters [26, 28, 188-192] is

$$\frac{1}{T_P}\in[0.033,0.1],\qquad \frac{K_P}{T_P}\in[4,12],\qquad \frac{1}{T_T}\in[2.564,4.762],\qquad \frac{1}{T_G}\in[9.615,17.857],$$

$$\frac{1}{RT_G}\in[3.081,10.639]$$

Considering the values of system parameters around the above range, let we have

$$A=\begin{bmatrix} -0.0665 & 11.5 & 0 \\ 0 & -2.5 & 2.5 \\ -9.5 & 0 & -13.7360 \end{bmatrix},\ B=\begin{bmatrix} 0 \\ 0 \\ 13.7360 \end{bmatrix},\ F=\begin{bmatrix} -11.5 \\ 0 \\ 0 \end{bmatrix}$$

The LFC system transfer function is given by

$$G(s)=\frac{\Delta f(s)}{u(s)}=\frac{394.9}{s^3+16.3s^2+35.42s+275.4}$$

For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0=[0,0.1,0,0]$; $P=\begin{bmatrix}0\ 0\ 0;0\ 0\ 0;\ 0\ 0\ 0\end{bmatrix}$. The length of simulation in samples is taken 60, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP=\begin{bmatrix} 0.3600 & 0.5313 & 0.0367 \\ 0.5313 & 1.5662 & 0.1690 \\ 0.0367 & 0.1690 & 0.0500 \end{bmatrix},\quad P=\begin{bmatrix} 0.3673 & 0.5357 & 0.0367 \\ 0.5357 & 1.5967 & 0.1733 \\ 0.0367 & 0.1733 & 0.0507 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK=\begin{bmatrix}0.5044 & 2.3212 & 0.6867\end{bmatrix},\ K=\begin{bmatrix}0.5040 & 2.3811 & 0.6962\end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-19.9774, -2.9441 + 3.9285i, -2.9441 - 3.9285i

Simulation with change in system parameter is also done at sample k=21; (i.e. t=1.05 seconds), such that *A*(2,2)=-4, and *A*(2,3)=4; then the solution is obtained as

$$RicP = \begin{bmatrix} 0.3006 & 0.3466 & 0.0370 \\ 0.3466 & 0.7568 & 0.1243 \\ 0.0370 & 0.1243 & 0.0532 \end{bmatrix}, \quad P = \begin{bmatrix} 0.3673 & 0.5357 & 0.0367 \\ 0.5357 & 1.5967 & 0.1733 \\ 0.0367 & 0.1733 & 0.0507 \end{bmatrix}$$

and the actor gains *RicK* and *K* respectively are obtained as

$$RicK = \begin{bmatrix} 0.5077 & 1.7077 & 0.7305 \end{bmatrix}, \; K = \begin{bmatrix} 0.5040 & 2.3811 & 0.6962 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-5.4251 + 4.1489i, -5.4251 - 4.1489i, -16.5154

Figs. 4.19 to 4.24 shows the simulation responses of LFC model (4.28) & (4.29) with ACD using PI technique. Fig. 4.19 shows system state trajectories which converge towards the equilibrium point. Fig. 4.20 shows control signal trajectory which also converge towards zero. Fig. 4.21 shows evolution of closed loop poles of the system during simulation. Fig. 4.22 shows convergence of critic parameters of matrix *P* towards optimal values. Fig. 2.23 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update. The simulation responses for the case of change in system parameter are similar as above and just the evolution of poles changes adapting the controller. The evolution of poles of the closed loop system for this case is shown in Fig.4.24.

Fig. 4.25 presents the closed loop response of LFC system without integral control model using both approaches of LQR and adaptive critic (AC). It is observed that the closed loop response with unit step load disturbance $\Delta P_d(t)$ have steady state error. The adaptive optimal controller using ACD gives similar responses as of standard LQR. Fig. 4.26 shows closed loop unit step response of load frequency for LFC system without integral control model and with change in system parameters using both approaches of LQR and adaptive critic (AC). It is observed here that the closed loop response with unit step load disturbance $\Delta P_d(t)$ is unaffected by change in system parameters and remains at same value as before but using LQR it is affected by change in system parameters. Thus the controller using ACD adapts for change in system internal dynamics.

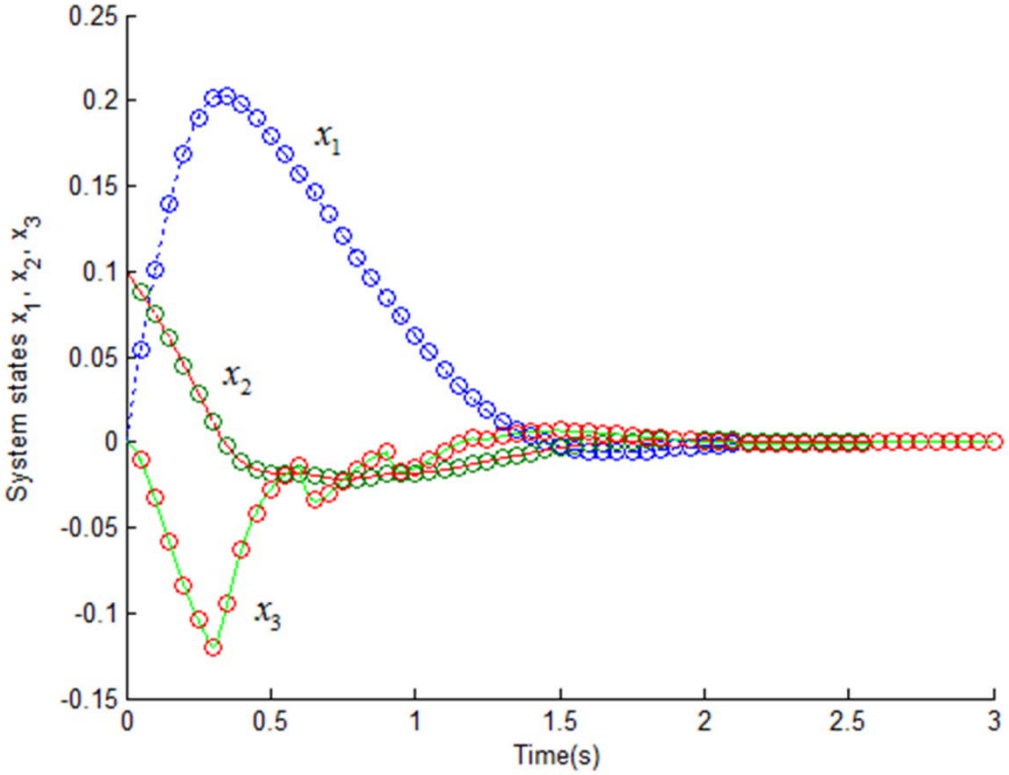


Fig. 4.19 System states

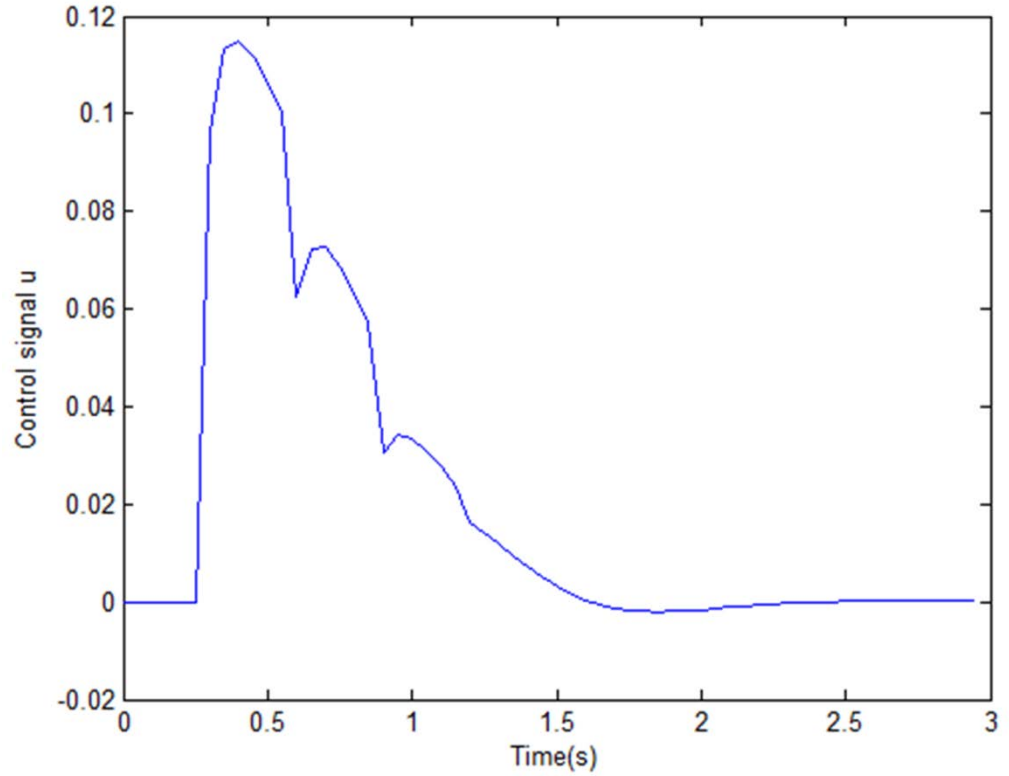


Fig. 4.20 Control signal

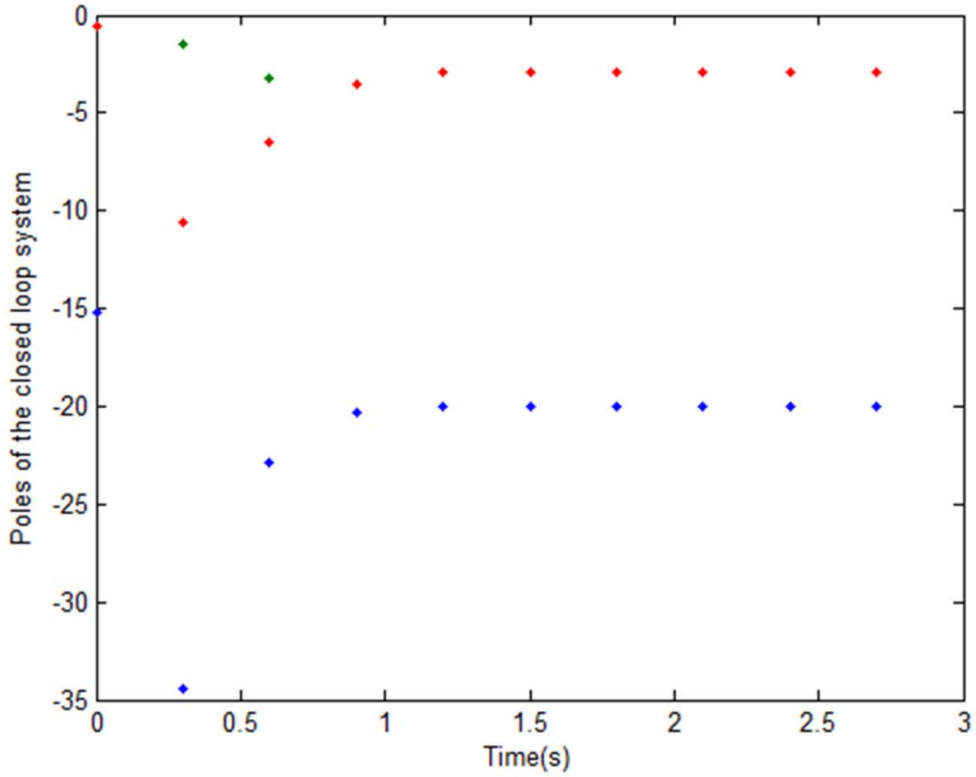


Fig. 4.21 Evolution of poles of closed loop system

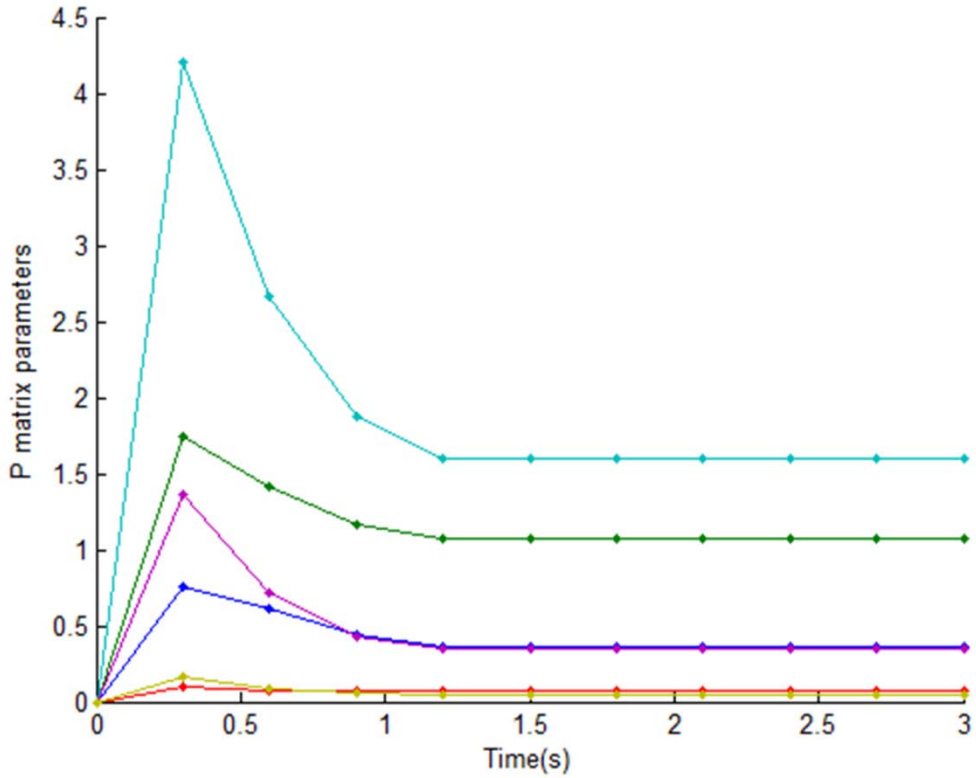


Fig. 4.22 Critic parameters

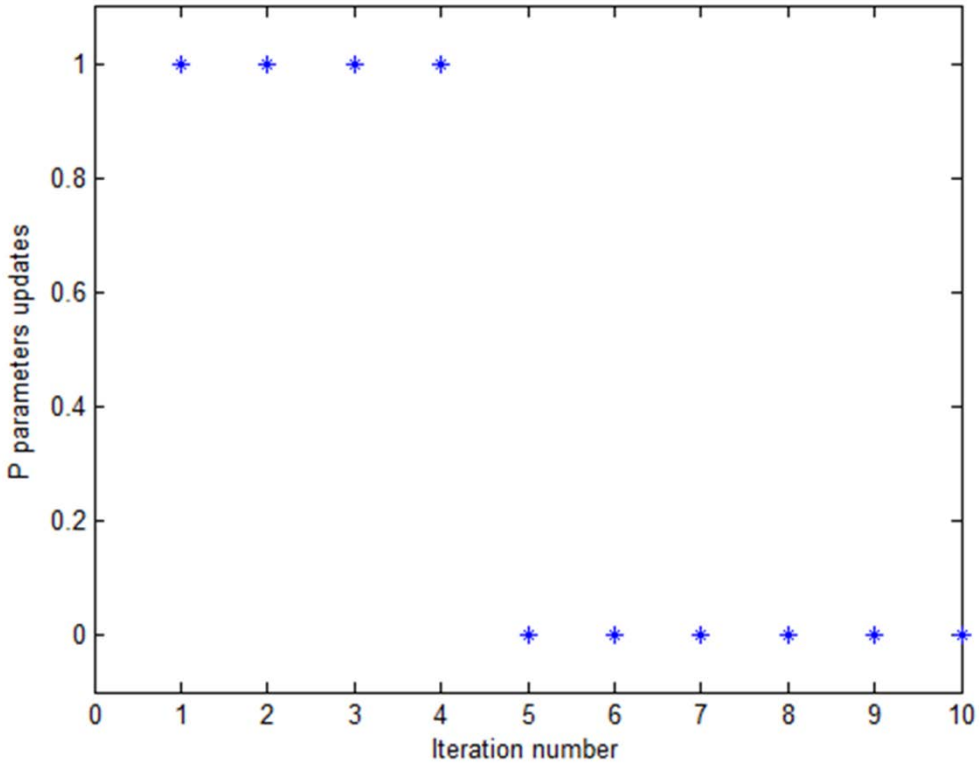


Fig. 4.23 Updating of critic parameters

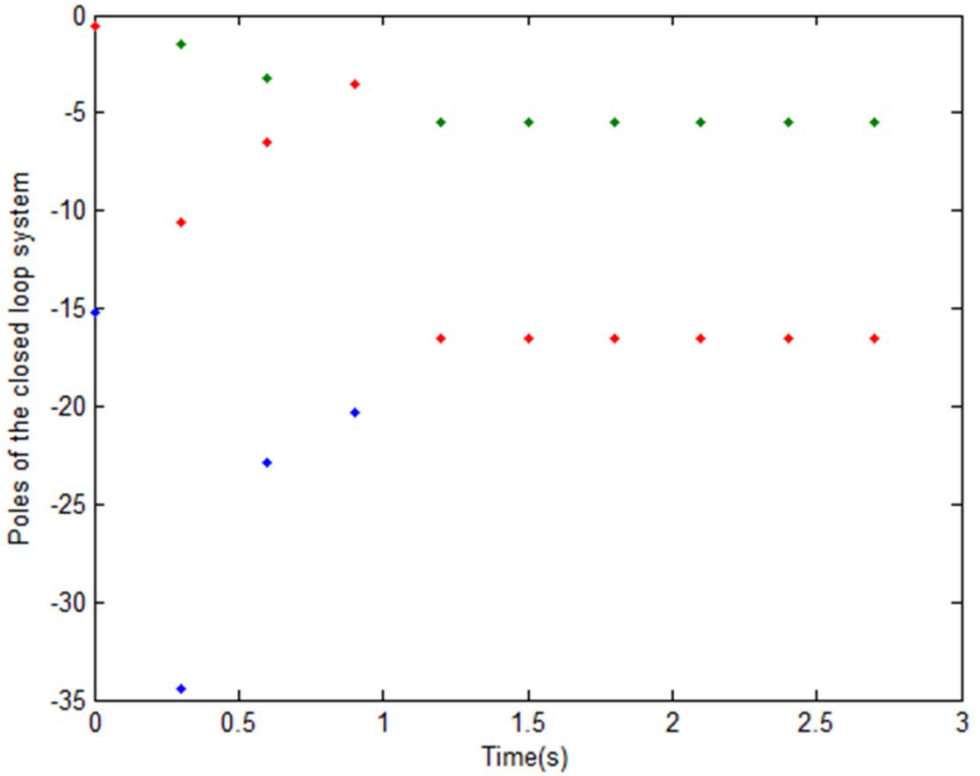


Fig. 4.24 Evolution of poles of closed loop system with change in system parameter at sample k=21 (i.e. t=1.05 seconds), A(2,2)=-4, and A(2,3)=4.

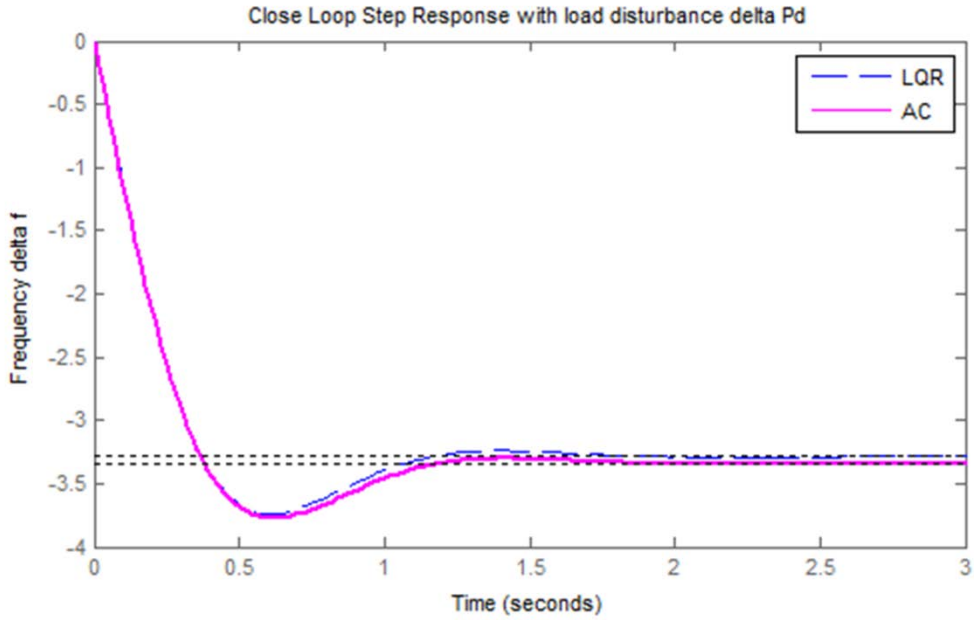


Fig. 4.25 Closed loop unit step response of load frequency for LFC system without integral control model.

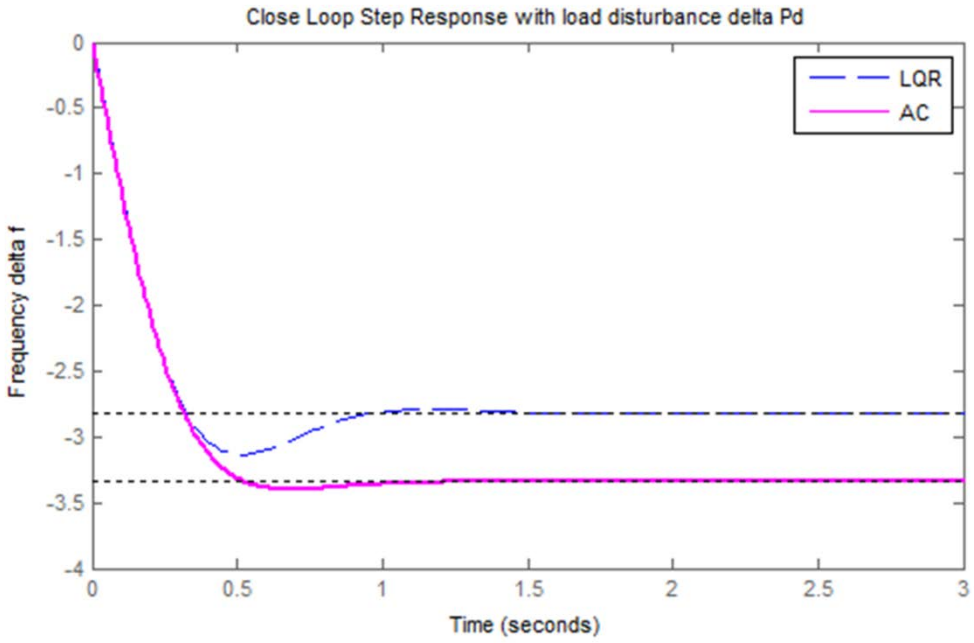


Fig. 4.26 Closed loop unit step response of load frequency for LFC system without integral control model and with change in system parameters at sample k=21 (i.e. t=1.05 seconds), A(2,2)=-4, A(2,3)=4.

4.6.3.2 LFC System model with integral control

The functional block diagram of single-area power system load frequency control with integral control model is shown in Fig. 4.27 [189-192].

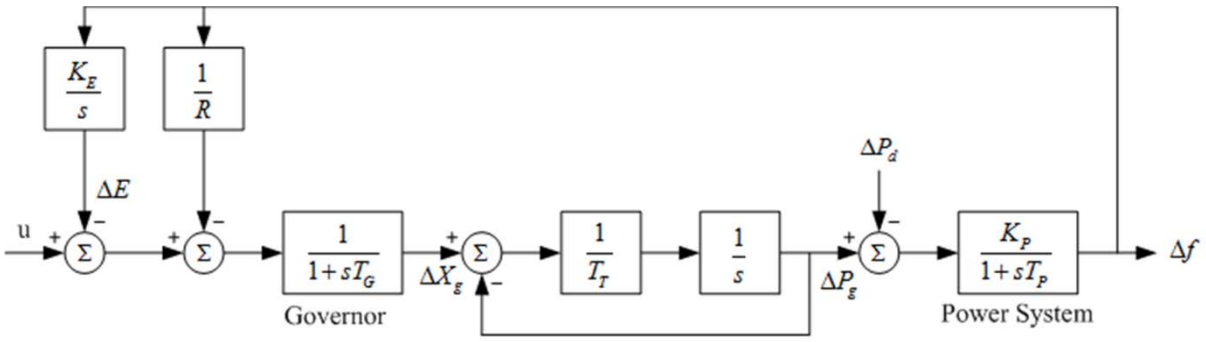


Fig. 4.27 Block diagram of load frequency control of power system with integral control.

Introducing an integral control of $\Delta f(t)$ in the LFC system dynamics, the state space model of the system is derived by modifying (4.28) & (4.29) as following [189-192]. The state $\Delta E(t)$, incremental change in integral control is included in the state vector $x(t)$ as, $x(t) = [\Delta f(t) \quad \Delta P_g(t) \quad \Delta X_g(t) \quad \Delta E(t)]^T$, and may be defined by

$$\Delta E(t) = K_E \int_0^t \Delta f(t) d\tau \tag{4.30}$$

to insure the regulation property of $\Delta f(t)$, i.e.

$$\Delta \dot{E}(t) = K_E \Delta f(t) \tag{4.31}$$

where, K_E is integral control gain; and

$$A = \begin{bmatrix} -\frac{1}{T_P} & \frac{K_P}{T_P} & 0 & 0 \\ 0 & -\frac{1}{T_T} & \frac{1}{T_T} & 0 \\ -\frac{1}{RT_G} & 0 & -\frac{1}{T_G} & -\frac{1}{T_G} \\ K_E & 0 & 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{T_G} \\ 0 \end{bmatrix}, \; F = \begin{bmatrix} -\frac{K_P}{T_P} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

Considering the values of system parameters around the above range, let we have

$$A = \begin{bmatrix} -0.0665 & 11.5 & 0 & 0 \\ 0 & -2.5 & 2.5 & 0 \\ -9.5 & 0 & -13.7360 & -13.7360 \\ 0.6 & 0 & 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 13.7360 \\ 0 \end{bmatrix}, \; F = \begin{bmatrix} -11.5 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Under this case the LFC system transfer function is given by

$$G(s) = \frac{\Delta f(s)}{u(s)} = \frac{394.9s + 2.677 \times 10^{-29}}{s^4 + 16.3s^3 + 35.42s^2 + 275.4s + 236.9}$$

For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = [0, 0.1, 0, 0, 0]$;

$P = [0\ 0\ 0\ 0;\ 0\ 0\ 0\ 0;\ 0\ 0\ 0\ 0;\ 0\ 0\ 0\ 0]$. The length of simulation in samples is taken 100, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = \begin{bmatrix} 0.4600 & 0.6911 & 0.0519 & 0.4642 \\ 0.6911 & 1.8668 & 0.2002 & 0.5800 \\ 0.0519 & 0.2002 & 0.0533 & 0.0302 \\ 0.4642 & 0.5800 & 0.0302 & 2.2106 \end{bmatrix},$$

$$P = \begin{bmatrix} 0.5428 & 0.7621 & 0.0552 & 0.5619 \\ 0.7621 & 2.2278 & 0.2504 & 0.6393 \\ 0.0552 & 0.2504 & 0.0610 & 0.0302 \\ 0.5619 & 0.6393 & 0.0302 & 2.3280 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4,5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK = [0.7135 \quad 2.7499 \quad 0.7323 \quad 0.4142],$$

$$K = [0.7587 \quad 3.4394 \quad 0.8372 \quad 0.4142]$$

The eigenvalues of closed loop system are obtained as

-20.1027, -3.4914 + 3.3266i, -3.4914 - 3.3266i, -0.7168

Simulation with change in system parameter is also done at sample k=81; (i.e. t=4.05 seconds), such that *A*(4,1)=0.8, then the solution is obtained as

$$RicP = \begin{bmatrix} 0.4981 & 0.7486 & 0.0573 & 0.4837 \\ 0.7486 & 1.9694 & 0.2106 & 0.5884 \\ 0.0573 & 0.2106 & 0.0544 & 0.0302 \\ 0.4837 & 0.5884 & 0.0302 & 1.7894 \end{bmatrix},$$

$$P = \begin{bmatrix} 0.5428 & 0.7621 & 0.0552 & 0.5619 \\ 0.7621 & 2.2278 & 0.2504 & 0.6393 \\ 0.0552 & 0.2504 & 0.0610 & 0.0302 \\ 0.5619 & 0.6393 & 0.0302 & 2.3280 \end{bmatrix}$$

and the actor gains *RicK* and *K* respectively are obtained as

$$RicK = [0.7869 \quad 2.8934 \quad 0.7473 \quad 0.4142],$$

$$K = [0.7587 \quad 3.4394 \quad 0.8372 \quad 0.4142]$$

The eigenvalues of closed loop system are obtained as

-20.0826, -1.0625, -3.3286 + 3.1399i, -3.3286 - 3.1399i

Figs. 4.28 to 4.33 show simulation responses of LFC with integral control model with ACD using PI technique. Fig. 4.28 shows system state trajectories which converge towards the equilibrium point. Fig. 4.29 shows control signal trajectory which also converge towards zero. Fig. 4.30 shows evolution of closed loop poles of the system during simulation. Fig. 4.31 shows convergence of critic parameters of matrix *P* towards optimal values. Fig. 4.32 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update. The simulation responses for the case of change in system parameter are similar as above and just the evolution of poles changes adapting the controller. Fig. 4.33 shows evolution of closed loop poles in this case.

Fig. 4.34 presents the closed loop response of LFC system with integral control model with unit step using both approaches of LQR and adaptive critic (AC). It is observed that there is no steady state error in the closed loop responses which is due to including integral control in LFC model. The adaptive optimal controller using ACD gives similar responses as of standard LQR. Fig. 4.35 presents the closed loop response of load frequency with unit step load disturbance $\Delta P_d(t)$ for case of change in system parameters. It is observed here also that the response of ACD remains exactly the same as before and is not affected by change in system parameter. The controller performs adapting the change in system parameters. This demonstrates that the proposed control scheme is effective & robust.

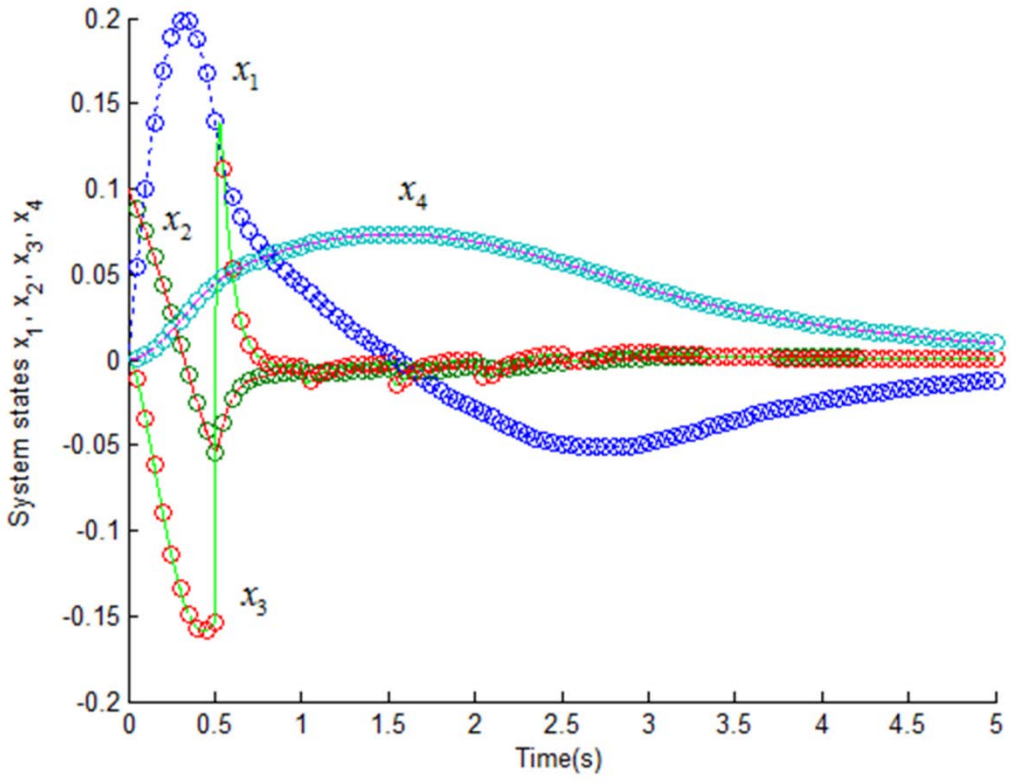


Fig. 4.28 System states

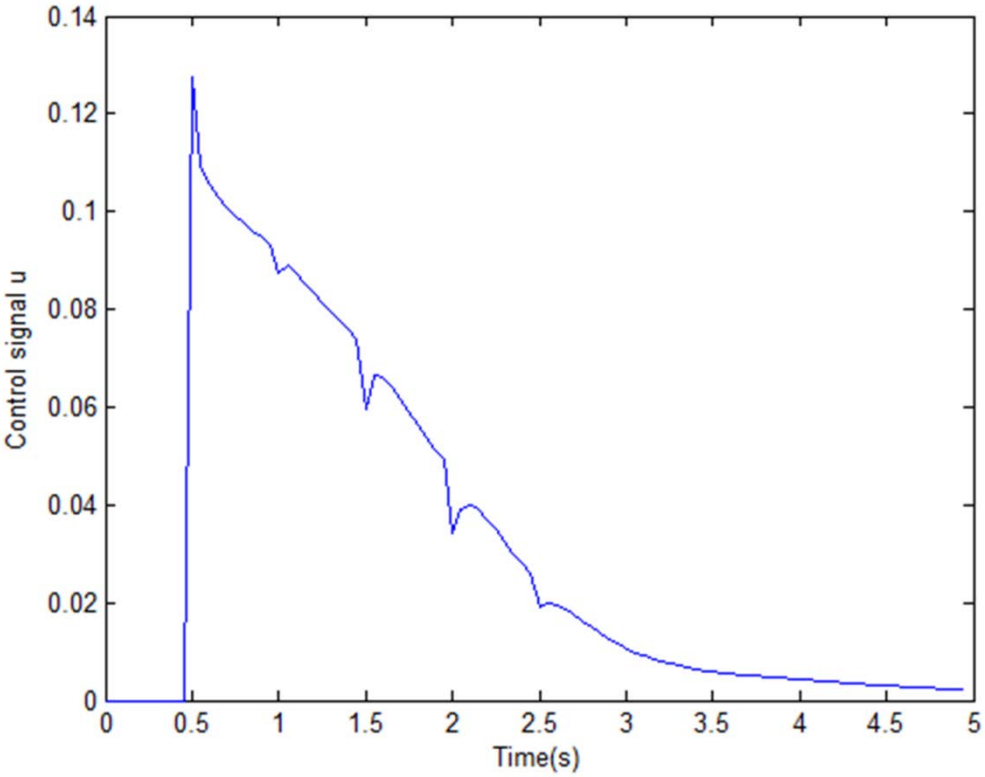


Fig. 4.29 Control signal

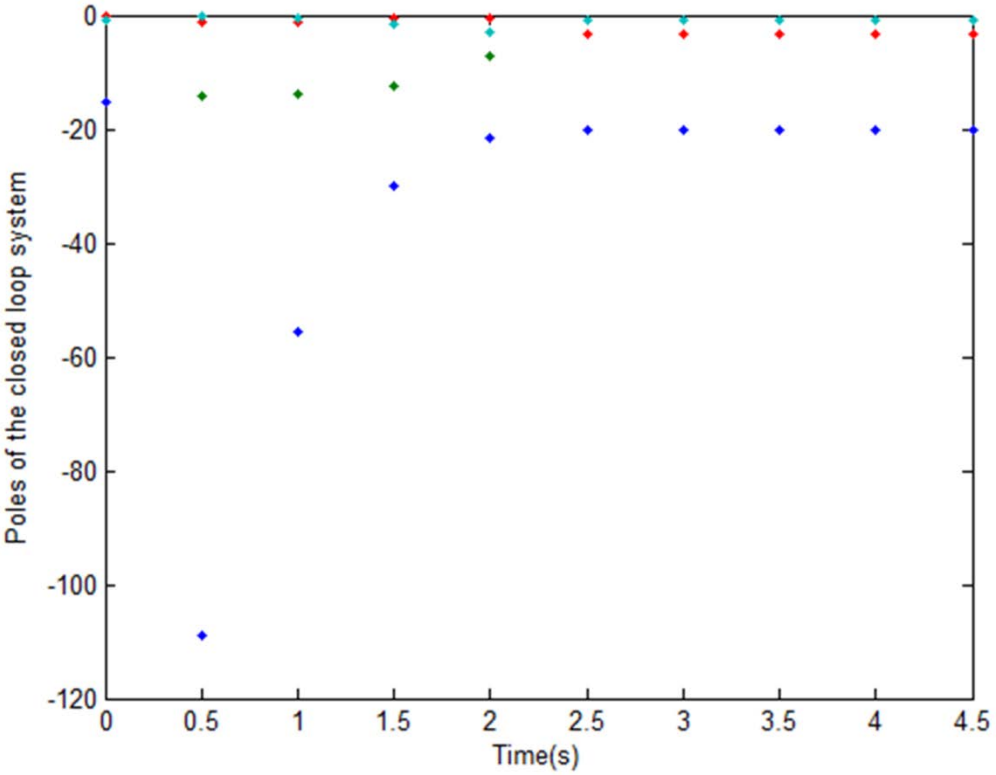


Fig. 4.30 Evolution of poles of closed loop system

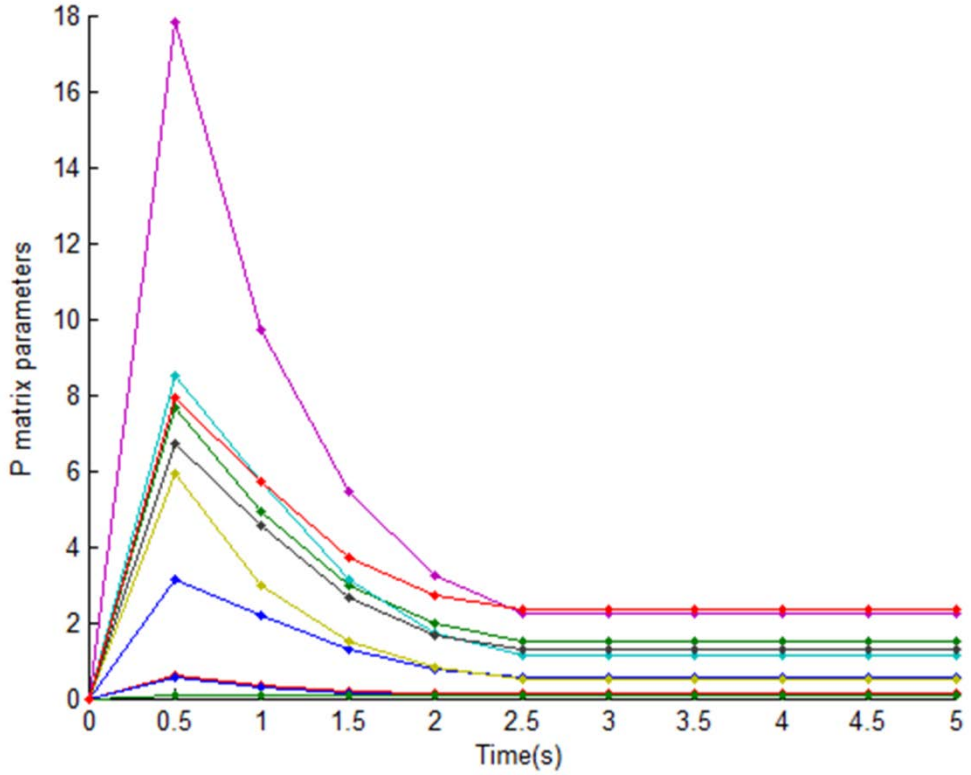


Fig. 4.31 Critic parameters

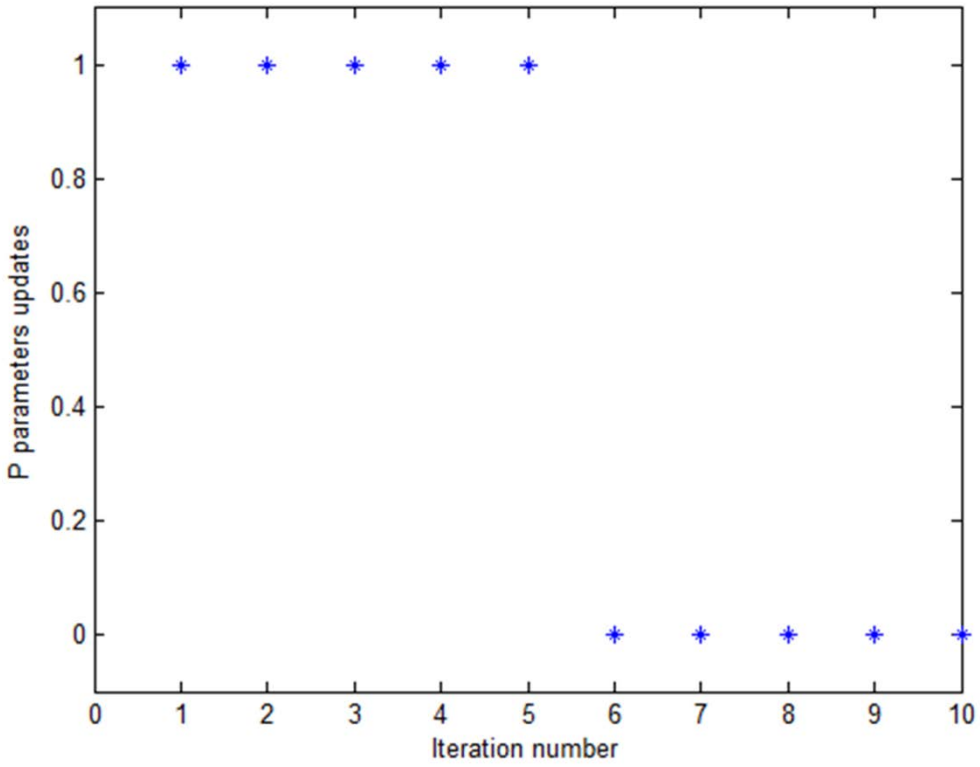


Fig. 4.32 Updating of critic parameters

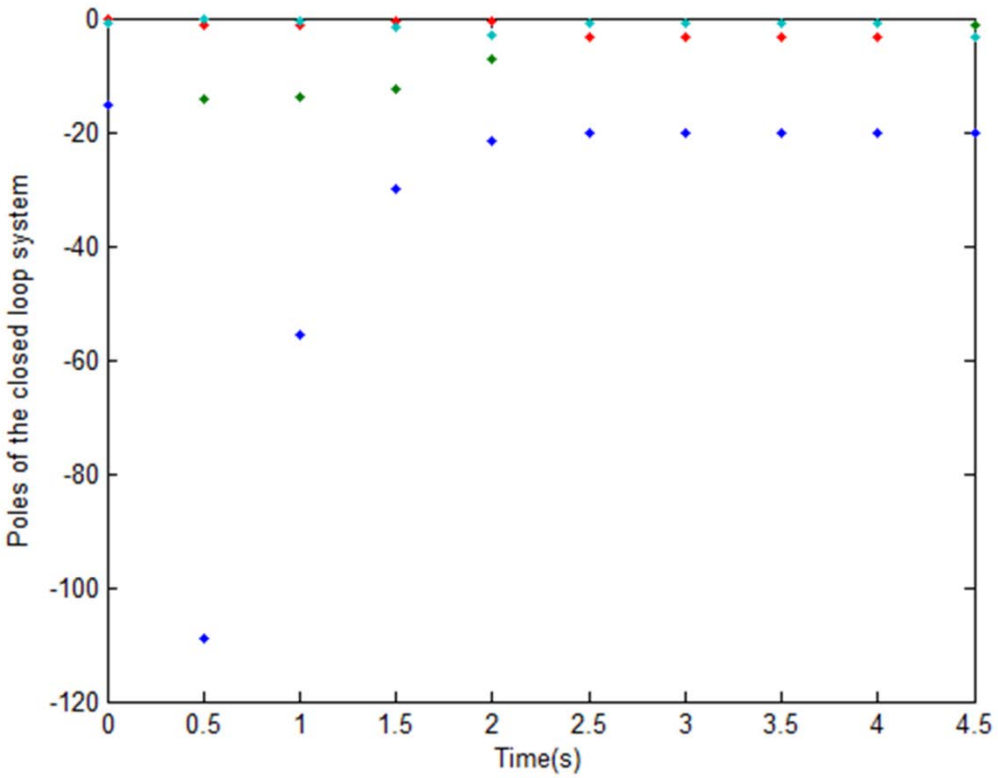


Fig. 4.33 Evolution of poles of closed loop system with change in system parameter at sample k=81 (i.e. t=4.05 seconds), A(4,1)=0.8.

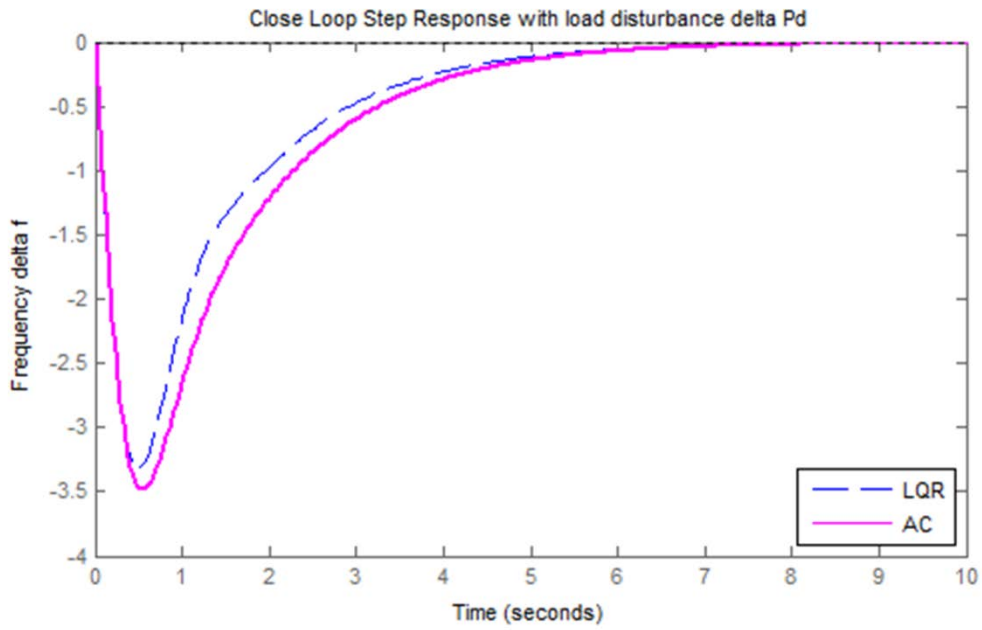


Fig. 4.34 Closed loop unit step response of load frequency for LFC system with integral control model.

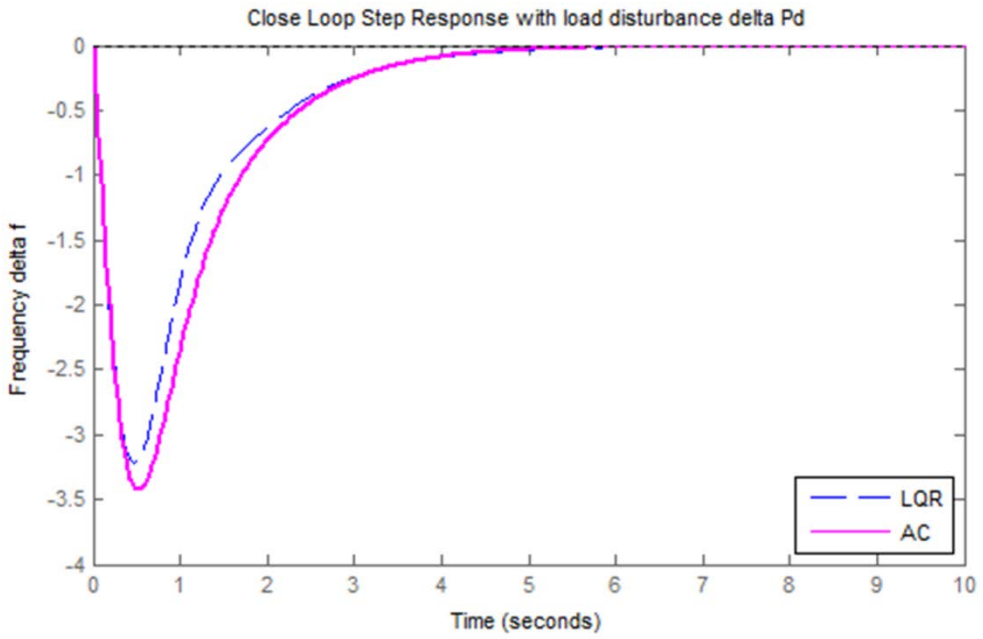


Fig. 4.35 Closed loop unit step response of load frequency for LFC system with integral control model and with change in system parameter at sample k=81 (i.e. t=4.05 seconds), A(4,1)=0.8.

In the above simulation results it is observed that critic parameter matrix *P* and actor parameter *K* obtained from control scheme of ACD using PI technique are converging adaptively to optimal values and *RicP* and *RicK* respectively are mostly of same values to that obtained from LQR approach. Also in case of change in system parameter in real situation the controller adapts it and converges to same optimal values. Thus the actor *K* and critic *P* parameters remain unchanged.

Analyzing the simulation results obtained for LFC system in all the above mentioned cases of system models without & with integral control, and changes in system parameters applying adaptive critic control using online PI technique, it is established that this adaptive critic control scheme provides a promising adaptive optimal control solution for dynamical systems without complete knowledge of the system dynamics. The structural change introduced in system dynamics by including integral control is augmenting the system behavior such as of its credit that removing the steady state error in closed loop responses. The structural change in system will not be adapted by the proposed controller but it will adapt the change in system parameters in real situation at any moment of time. Thus this technique is partially model-free, effective & robust.

4.6.4 Automatic Voltage Regulator System

The industrial and domestic electrical appliances are designed to operate at a certain voltage and frequency rating and thus their performance is dependent on the quality of power supply. The performance of equipments is adversely affected and possibly may cause them

damage if there is prolonged operation of the equipments outside the allowable range of voltages. Thus the terminal voltage of all the equipments in the system should be maintained within acceptable limits. Automatic voltage regulator (AVR) is an essential system in an electric power system which maintains the magnitude of terminal voltage of a synchronous generator at a specified level by its field excitation control. The incremental model of an AVR system is a continuous-time linear time-invariant (LTI) dynamical system.

The basic concept of AVR system for electric power system is discussed in [182-185]. Recently the control design using various approaches for automatic voltage regulator system and power system stabilizers has attracted researchers [36-41, 175, 176, 199-203]. The conventional and recent control schemes for AVR system present in the literature are generally off-line and not giving adaptive and optimal control solution at the same time in real situation. Thus the investigation of adaptive optimal control solution for AVR system is desired for automation of voltage control in electric power system.

Automatic voltage regulator (AVR) maintains the generator terminal voltage and controls the reactive power flow by controlling generator field excitation. A simple AVR system comprises four main components, namely amplifier, exciter, generator, and comparator & sensor. Since the time constant of sensor is normally very small thus it may be neglected for a simplified mathematical model of AVR system of a single-area power system. Fig. 36 shows the functional block diagram of a simple AVR system neglecting the sensor dynamics [182-185]. Fig. 37 shows the functional block diagram of an AVR system including sensor dynamics [36, 37, 39-41, 200]. Considering the major time constant and ignoring the saturation or other nonlinearities, and deviated from a normal state, the linearized incremental mathematical modelling of AVR system is presented as following [36-41, 182-185].

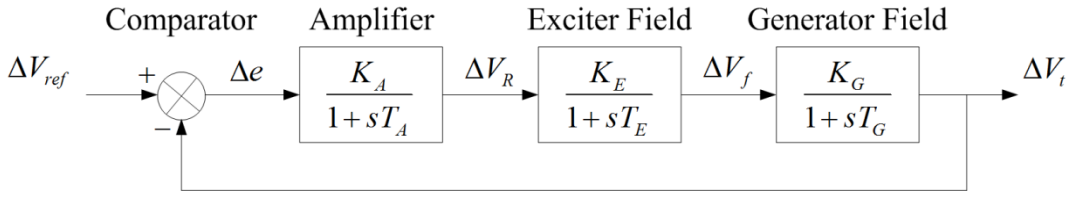


Fig. 4.36 Block diagram of automatic voltage regulator neglecting sensor dynamics.

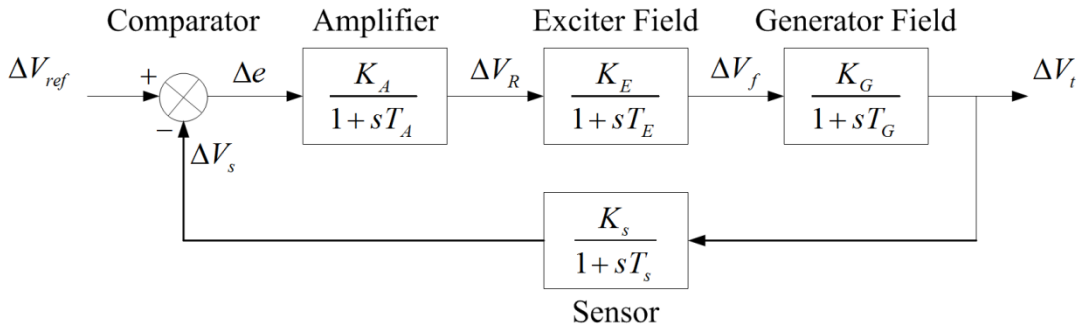


Fig. 4.37 Block diagram of automatic voltage regulator including sensor dynamics.

(a) Amplifier model

$$G_A(s) = \frac{\Delta V_R(s)}{\Delta e(s)} = \frac{K_A}{1 + sT_A} \tag{4.32}$$

where nominally $10 < K_A < 400$ and a small time constant is in the range $0.02 < T_A < 0.1$.

(b) Exciter model

$$G_E(s) = \frac{\Delta V_f(s)}{\Delta V_R(s)} = \frac{K_E}{1 + sT_E} \tag{4.33}$$

where nominally $10 < K_E < 400$ and a time constant is in the range $0.5 < T_E < 1$.

(c) Generator model

$$G_G(s) = \frac{\Delta E(s)}{\Delta V_f(s)} \cong \frac{\Delta V_t(s)}{\Delta V_f(s)} = \frac{K_G}{1 + sT_G} \tag{4.34}$$

where nominally $0.7 < K_G < 1$ and a time constant is in the range $1 < T_G < 2$. These constants vary depending on the load.

(d) Sensor model

$$H_s(s) = \frac{\Delta V_s(s)}{\Delta V_t(s)} = \frac{K_s}{1 + sT_s} \tag{4.35}$$

where nominally $K_s = 1$, and the time constant is in the range $0.001 < T_s < 0.06$.

The AVR system parameters consider are [39, 41, 182-185]: T_A = 0.1 sec, T_E = 0.4 sec, T_G = 1 sec, T_s = 0.01 sec, K_A = 10, K_E = 1, K_G = 1, and K_s = 1.

4.6.4.1 AVR system neglecting sensor dynamics

The open loop transfer function is given by

$$G(s) = \frac{\Delta V_t(s)}{\Delta e(s)} = \frac{K_A K_E K_G}{(1 + sT_A)(1 + sT_E)(1 + sT_G)} \tag{4.36}$$

The closed loop transfer function of AVR system neglecting the sensor dynamics is written as

$$G_{CL}(s) = \frac{K_A K_E K_G}{(1 + sT_A)(1 + sT_E)(1 + sT_G) + K_A K_E K_G} \tag{4.37}$$

Analysis of (4.37) gives that the static error decreases with increased loop gain as for p% static error the loop gain $K_A K_E K_G > \frac{100}{p} - 1$. For stability compensation a series phase lead compensator $G_s(s) = 1 + sT_c$ can be included in the system.

The state space model of AVR system neglecting sensor dynamics in phase variable companion form may be written from transfer function model (4.37) as following:

$$\dot{X} = AX + Bu \tag{4.38}$$

$$y = CX + Du \tag{4.39}$$

where, $X = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$, $y = x_1 = \Delta V_t$, $u = \Delta V_{ref}$, and

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\dfrac{1+K_A K_E K_G}{T_A T_E T_G} & -\left(\dfrac{1}{T_A T_E}+\dfrac{1}{T_E T_G}+\dfrac{1}{T_G T_A}\right) & -\left(\dfrac{1}{T_A}+\dfrac{1}{T_E}+\dfrac{1}{T_G}\right) \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

$$C = \begin{bmatrix} \dfrac{K_A K_E K_G}{T_A T_E T_G} & 0 & 0 \end{bmatrix}, \; D = 0 .$$

For the given system parameters the closed loop transfer function of AVR system (4.37) is obtained as

$$G_{CL}(s) = \frac{250}{s^3 + 13.5s^2 + 37.5s + 275} \tag{4.40}$$

and the state space model of AVR system is obtained as

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -275 & -37.5 & -13.5 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} 250 & 0 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 0 \end{bmatrix}$$

In implementation of PI algorithm for system (4.40), the initial conditions for states and cost function, and critic parameters are taken as $x_0 = \begin{bmatrix} 0.1 & 0.05 & 0.04 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 0\ 0\ 0; 0\ 0\ 0; 0\ 0\ 0 \end{bmatrix}$. The length of the simulation in samples is taken 120, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = \begin{bmatrix} 167.9165 & 22.5745 & 0.0018 \\ 22.5745 & 11.3630 & 0.6104 \\ 0.0018 & 0.6104 & 0.0820 \end{bmatrix}, \; P = \begin{bmatrix} 167.9185 & 22.5747 & 0.0018 \\ 22.5747 & 11.3631 & 0.6104 \\ 0.0018 & 0.6104 & 0.0820 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK = \begin{bmatrix} 0.0018 & 0.6104 & 0.0820 \end{bmatrix}, \; K = \begin{bmatrix} 0.0018 & 0.6104 & 0.0820 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-12.3013, -0.6404 + 4.6846i, -0.6404 - 4.6846i.

The simulation responses using PI technique for LTI system (4.40) are shown in Figs. 4.38 to 4.42. Fig. 4.38 shows the system state trajectories which converge towards the equilibrium point. Each circle “o” on state trajectories shows the modification of initial

conditions to new states values during simulation in each sample. Fig. 4.39 shows the control signal trajectory which also converges towards zero. Fig. 4.40 shows the evolution of closed loop poles of the system during simulation. Fig. 4.41 shows the convergence of critic parameters of matrix *P* towards optimal values. Fig. 4.42 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update.

Simulation is also done for the case of change in system parameter which changes the elements of system matrix *A* during simulation. Consider the change in system parameter as T_G = 1.25 seconds at sample *k*=41; (i.e. *t*=2.05 seconds) then for this case system transfer function is obtained as

$$G_{CL}(s) = \frac{200}{s^3 + 13.3s^2 + 35s + 220} \tag{4.41}$$

, and thus the system state model is changed such that *A*(3,1) = -220; *A*(3,2) = -35; *A*(3,3) = -13.3; and $C = \begin{bmatrix} 200 & 0 & 0 \end{bmatrix}$. In this case the unique positive definite solution of *RicP* and *P* using LQR and PI technique respectively are obtained as

$$RicP = \begin{bmatrix} 103.1486 & 16.0069 & 0.0023 \\ 16.0069 & 8.8044 & 0.4685 \\ 0.0023 & 0.4685 & 0.0726 \end{bmatrix},\ P = \begin{bmatrix} 167.9185 & 22.5747 & 0.0018 \\ 22.5747 & 11.3631 & 0.6104 \\ 0.0018 & 0.6104 & 0.0820 \end{bmatrix}$$

, and the actor gains of LQR design *RicK* , and actor *K* by adaptive critic scheme using PI respectively are obtained as

$$RicK = \begin{bmatrix} 0.0023 & 0.4685 & 0.0726 \end{bmatrix},\ K = \begin{bmatrix} 0.0018 & 0.6104 & 0.0820 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-11.9427, -0.7196 + 4.2313i, -0.7196 - 4.2313i.

Fig. 4.43 shows the evolution of closed loop poles of the system during simulation with change in system parameter at sample *k*=41. The remaining responses obtained are same as above in this case. Fig. 4.44 presents the closed loop response of LTI system (4.41) using both approaches of LQR and adaptive critic (AC) using PI technique by replacing $u = -Kx + r$ in state equations where $r = V_{ref}$ is a unit step input and *K* is actor gains *RicK* and *K* respectively. It remains exactly the same also for case with change in system parameters. It is observed here that the adaptive optimal controller using PI technique gives the similar response as of standard LQR.

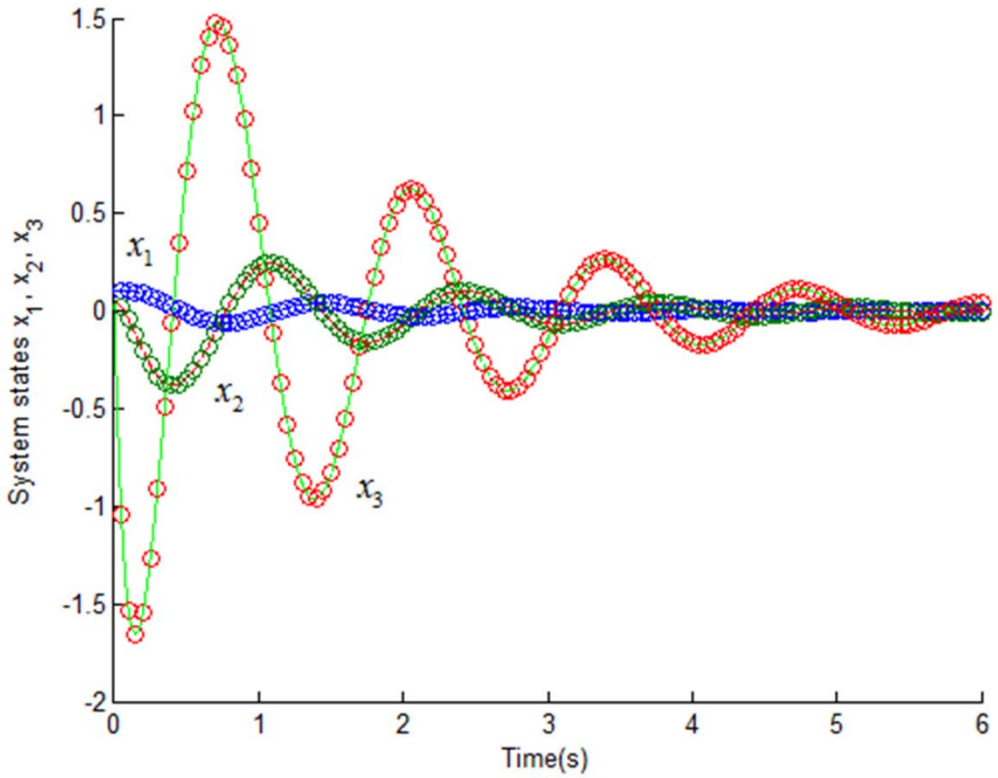


Fig. 4.38 System states.

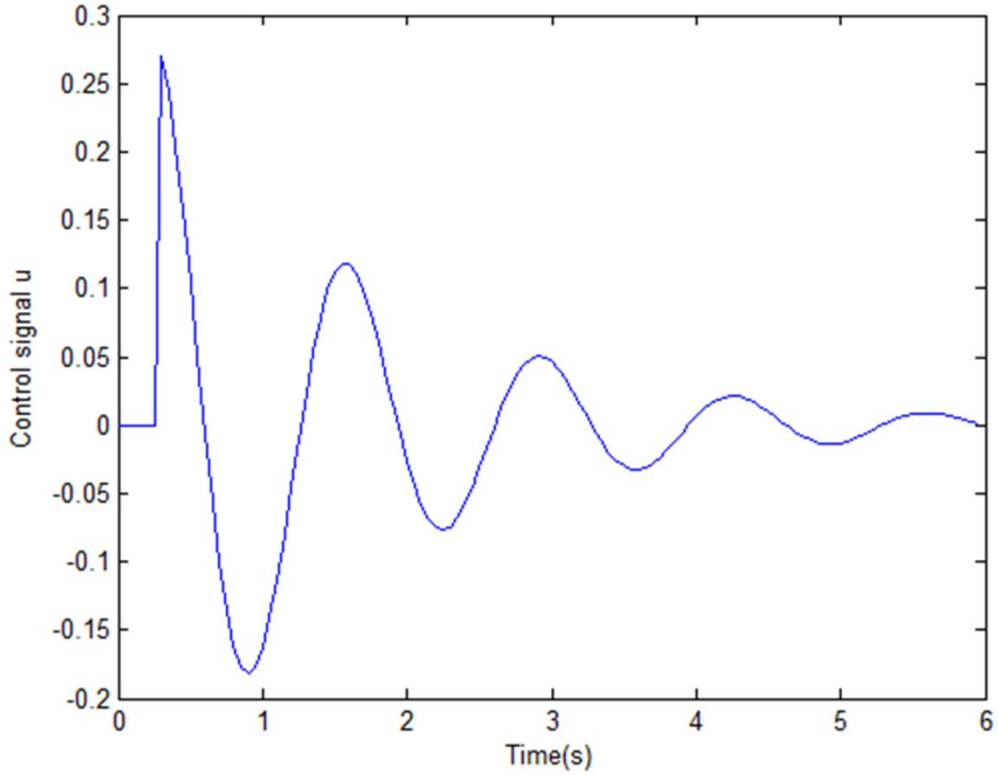


Fig. 4.39 Control signal.

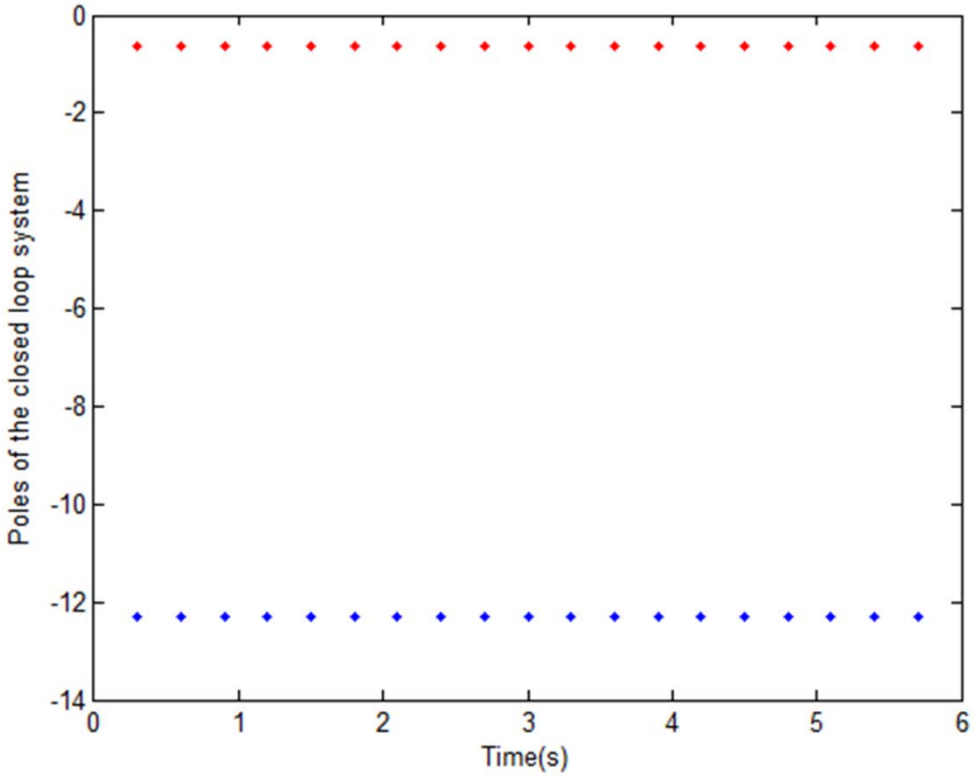


Fig. 4.40 Evolution of poles of closed loop system.

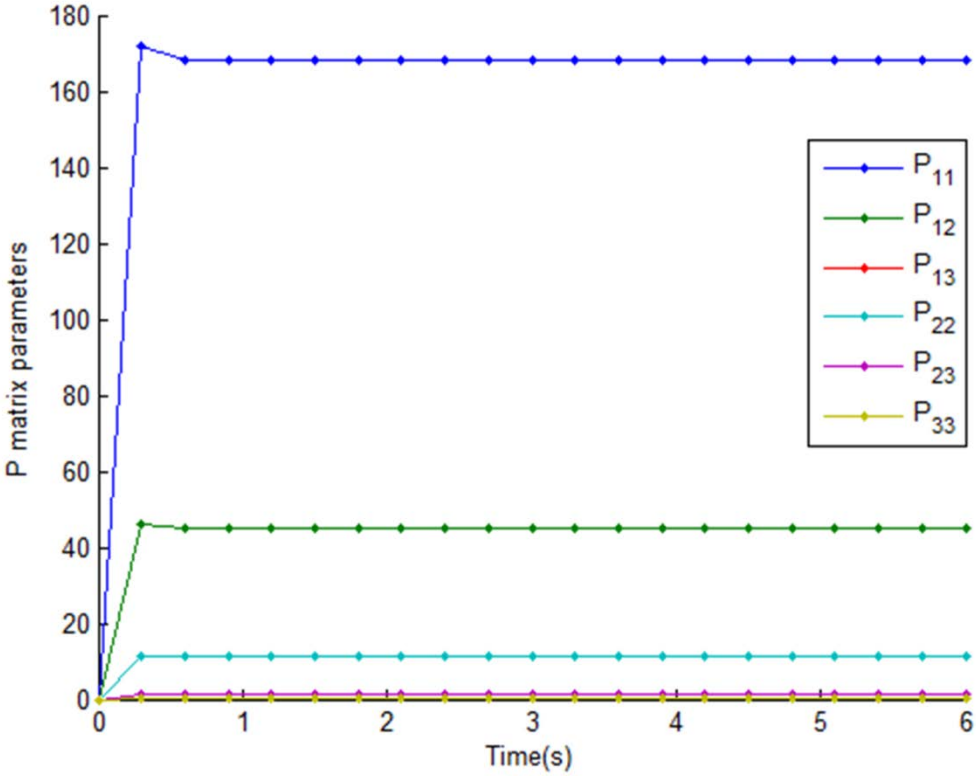


Fig. 4.41 Critic parameters.

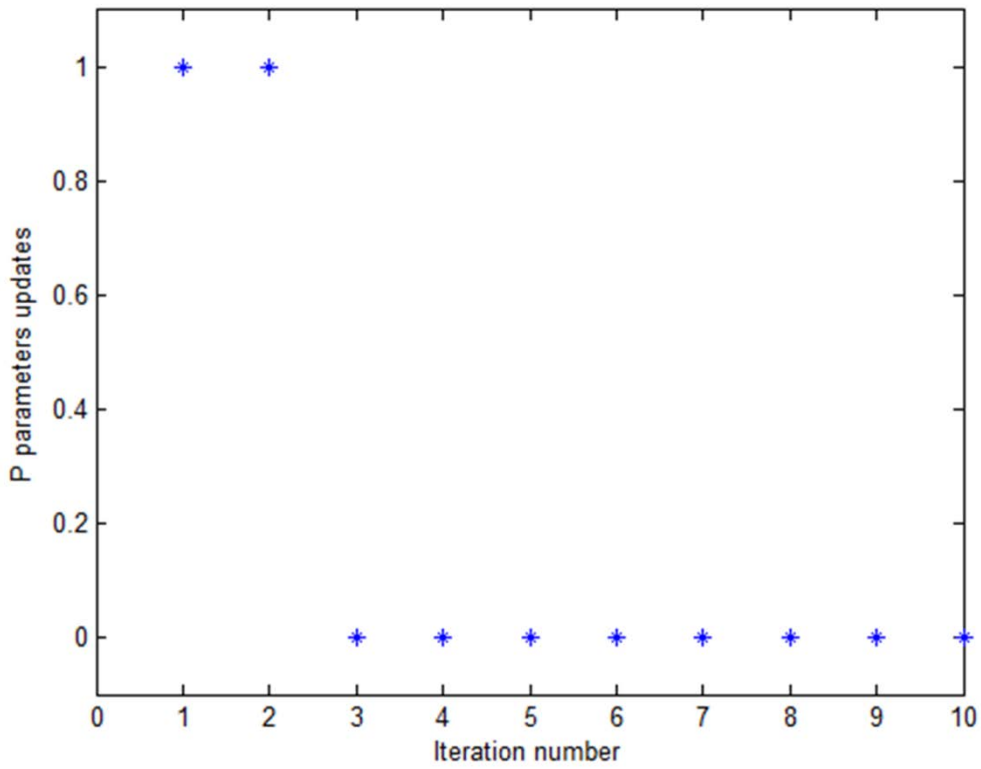


Fig. 4.42 Updating of critic parameters.

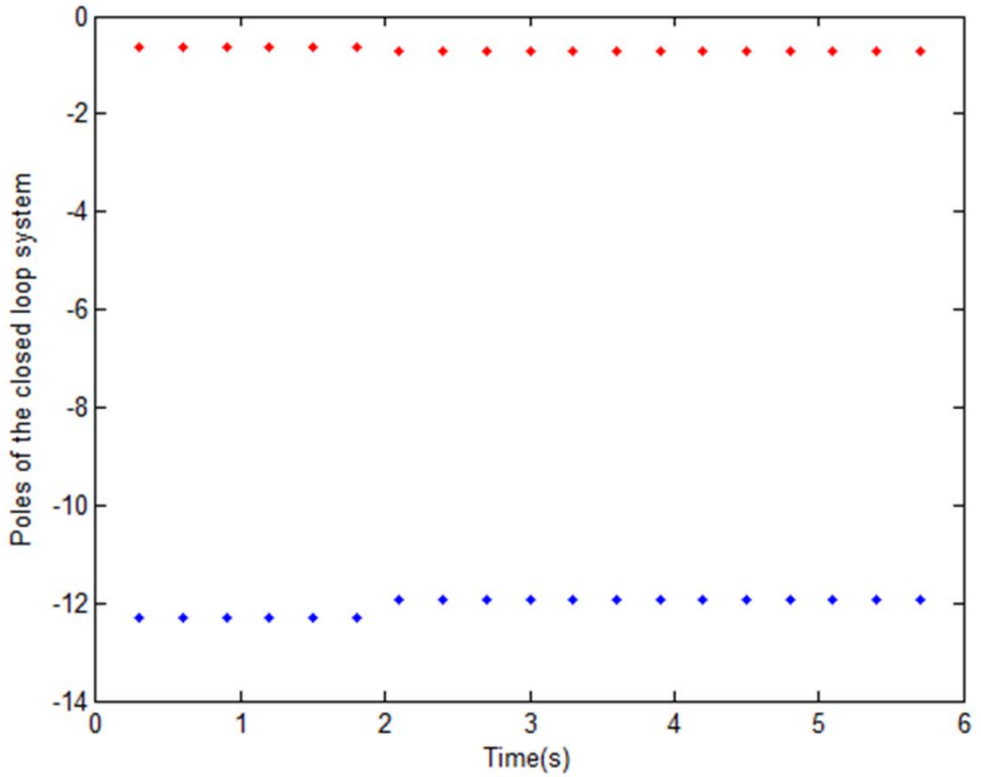


Fig. 4.43 Evolution of poles of closed loop system with change in system parameters at k=41.

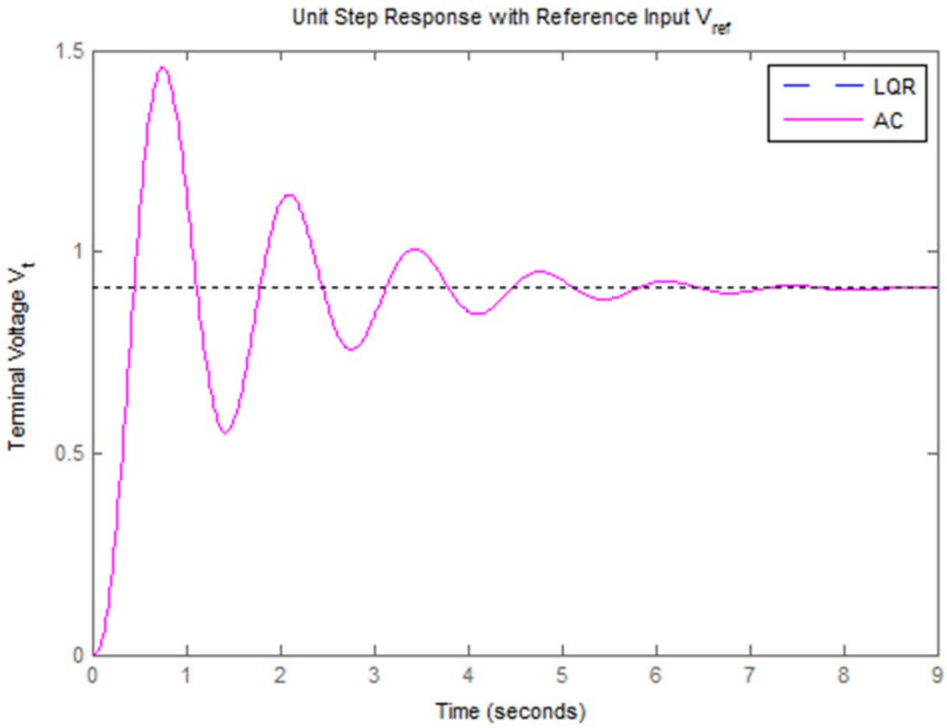


Fig. 4.44 Unit step response of closed loop system.

4.6.4.2 AVR system including sensor dynamics

The closed loop transfer function of AVR system including sensor dynamics is written as

$$G_{CL}(s)=\frac{K_A K_E K_G(1+sT_s)}{(1+sT_A)(1+sT_E)(1+sT_G)(1+sT_s)+K_A K_E K_G K_s} \tag{4.42}$$

Including the sensor dynamics in the AVR system modelling, a pole and a zero are added.

The state space model of AVR system including sensor dynamics in phase variable companion form may be written from transfer function model (4.42) similar as (4.38) and (4.39) with following modification:

$X=\begin{bmatrix}x_1 & x_2 & x_3 & x_4\end{bmatrix}^T$, and

$$A=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -a_4 & -a_3 & -a_2 & -a_1\end{bmatrix},\quad B=\begin{bmatrix}0\\0\\0\\1\end{bmatrix},\quad C=\begin{bmatrix}\frac{K_A K_E K_G}{T_A T_E T_G T_s} & \frac{K_A K_E K_G}{T_A T_E T_G} & 0 & 0\end{bmatrix},\quad D=0,$$

where, $a_1=\left(\frac{1}{T_A}+\frac{1}{T_E}+\frac{1}{T_G}+\frac{1}{T_s}\right)$, $a_2=\left(\frac{1}{T_A T_E}+\frac{1}{T_A T_G}+\frac{1}{T_A T_s}+\frac{1}{T_E T_G}+\frac{1}{T_E T_s}+\frac{1}{T_G T_s}\right)$,

$a_3=\left(\frac{1}{T_A T_E T_G}+\frac{1}{T_A T_E T_s}+\frac{1}{T_A T_G T_s}+\frac{1}{T_E T_G T_s}\right)$, and $a_4=\frac{1+K_A K_E K_G K_s}{T_A T_E T_G T_s}$.

In the both cases of neglecting and including sensor dynamics AVR system is a continuous-time single-input single-output (SISO) linear time-invariant (LTI) system.

For the given system parameters the transfer function of AVR system (4.42) is obtained as $G_{CL}(s) = \dfrac{250s + 25000}{s^4 + 113.5s^3 + 1387.5s^2 + 3775s + 27500}$ (4.43)

, and the state space model of AVR system is obtained as

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -27500 & -3775 & -1387.5 & -113.5 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},\ C = \begin{bmatrix} 25000 & 250 & 0 & 0 \end{bmatrix},\ D = \begin{bmatrix} 0 \end{bmatrix}$$

In implementation of PI algorithm for system (4.43), the initial conditions for states and cost function, and critic parameters are taken as $x_0 = \begin{bmatrix} 0.1 & 0.05 & 0.04 & 0.02 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 0\ 0\ 0\ 0; 0\ 0\ 0\ 0; 0\ 0\ 0\ 0 \end{bmatrix}$. The length of the simulation in samples is taken 120, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = 1.0 \times 10^3 \begin{bmatrix} 6.5160 & 0.8940 & 0.1290 & 0.0000 \\ 0.8940 & 0.3225 & 0.0446 & 0.0002 \\ 0.1290 & 0.0446 & 0.0100 & 0.0000 \\ 0.0000 & 0.0002 & 0.0000 & 0.0000 \end{bmatrix},$$

$$P = 1.0 \times 10^3 \begin{bmatrix} 6.5164 & 0.8940 & 0.1290 & -0.0000 \\ 0.8940 & 0.3225 & 0.0446 & 0.0002 \\ 0.1290 & 0.0446 & 0.0100 & 0.0000 \\ -0.0000 & 0.0002 & 0.0000 & 0.0000 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK = \begin{bmatrix} 0.0000 & 0.2369 & 0.0325 & 0.0047 \end{bmatrix},$$

$$K = \begin{bmatrix} -0.0001 & 0.2370 & 0.0325 & 0.0047 \end{bmatrix}$$

The eigenvalues of closed loop system in this case are obtained as

-99.9763, -12.4887, -0.5199 + 4.6642i, and -0.5199 - 4.6642i.

The simulation responses using PI technique for LTI system (4.43) are shown in Figs. 4.45 to 4.49. Fig. 4.45 shows the system state trajectories which converge towards the equilibrium point. Each circle “o” on state trajectories shows the modification of initial conditions to new states values during simulation in each sample. Fig. 4.46 shows the control

signal trajectory which also converges towards zero. Fig. 4.47 shows the evolution of closed loop poles of the system during simulation. Fig. 4.48 shows the convergence of critic parameters of matrix *P* towards optimal values. Fig. 4.49 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update.

Simulation is also done for the case of change in system parameter which changes the elements of system matrix *A* during simulation. Consider the change in system parameter as T_G = 1.25 seconds at sample *k*=41; (i.e. *t*=2.05 seconds) then for this case the system transfer function is obtained as

$$G_{CL}(s) = \frac{200s + 20000}{s^4 + 113.3s^3 + 1365s^2 + 3520s + 22000} \tag{4.44}$$

and the system state model is changed such that *A*(4,1) = -22000; *A*(4,2) = -3520; *A*(4,3) = -1365; *A*(4,4) = -113.3; and *C* = [20000 200 0 0]. In this case also the unique positive definite solution of *RicP* and *P* using LQR and PI technique respectively are obtained as

$$RicP = 1.0 \times 10^3 \begin{bmatrix} 3.5299 & 0.5643 & 0.1021 & 0.0000 \\ 0.5643 & 0.2072 & 0.0345 & 0.0002 \\ 0.1021 & 0.0345 & 0.0091 & 0.0000 \\ 0.0000 & 0.0002 & 0.0000 & 0.0000 \end{bmatrix},$$

$$P = 1.0 \times 10^3 \begin{bmatrix} 6.5164 & 0.8940 & 0.1290 & -0.0000 \\ 0.8940 & 0.3225 & 0.0446 & 0.0002 \\ 0.1290 & 0.0446 & 0.0100 & 0.0000 \\ -0.0000 & 0.0002 & 0.0000 & 0.0000 \end{bmatrix}$$

, and the actor gains of LQR design *RicK*, and actor *K* by adaptive critic scheme using PI respectively are obtained as

$$RicK = \begin{bmatrix} 0.0000 & 0.1604 & 0.0256 & 0.0046 \end{bmatrix},\ K = \begin{bmatrix} -0.0001 & 0.2370 & 0.0325 & 0.0047 \end{bmatrix}$$

The eigenvalues of closed loop system in this case are obtained as

-99.9821, -12.0977, -0.6125 + 4.2206i, and -0.6125 - 4.2206i.

Fig. 4.50 shows the evolution of closed loop poles of the system during simulation with change in system parameter at sample *k*=41. The remaining responses obtained are same as above in this case. Fig. 4.51 presents the closed loop response of LTI system (4.44) using both approaches of LQR and adaptive critic (AC) using PI technique by replacing $u = -Kx + r$ in state equations where $r = V_{ref}$ is a unit step input and *K* is actor gains *RicK* and *K* respectively. It remains exactly the same also for case with change in system parameters. It is observed here that the adaptive optimal controller using PI technique gives the similar response as of standard LQR.

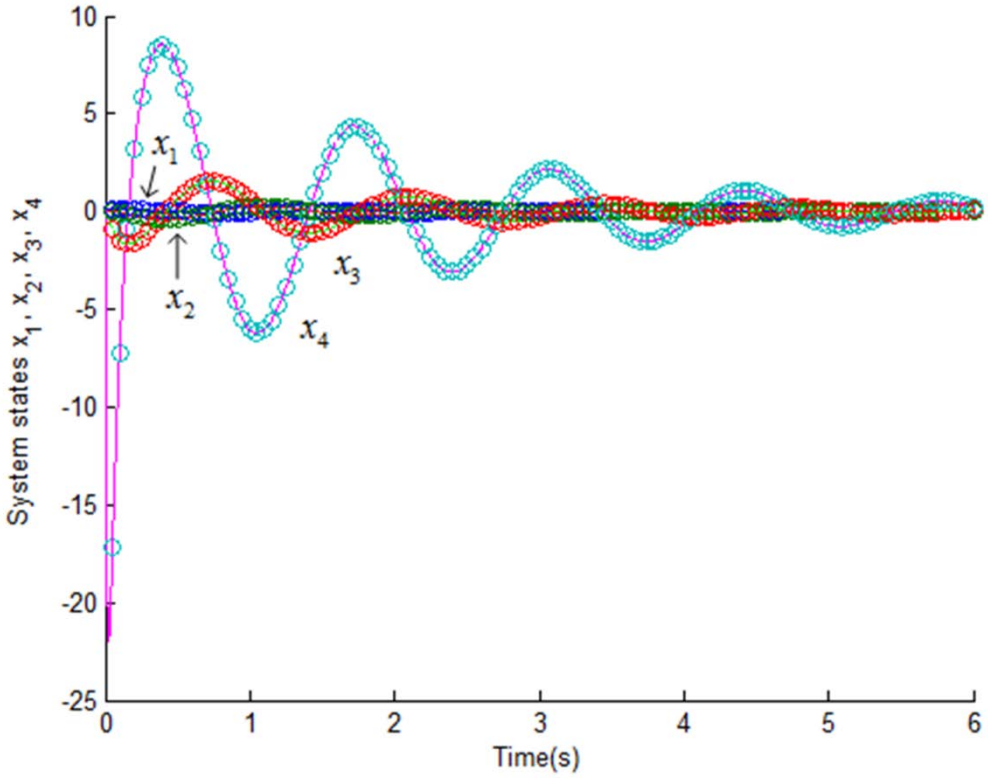


Fig. 4.45 System states.

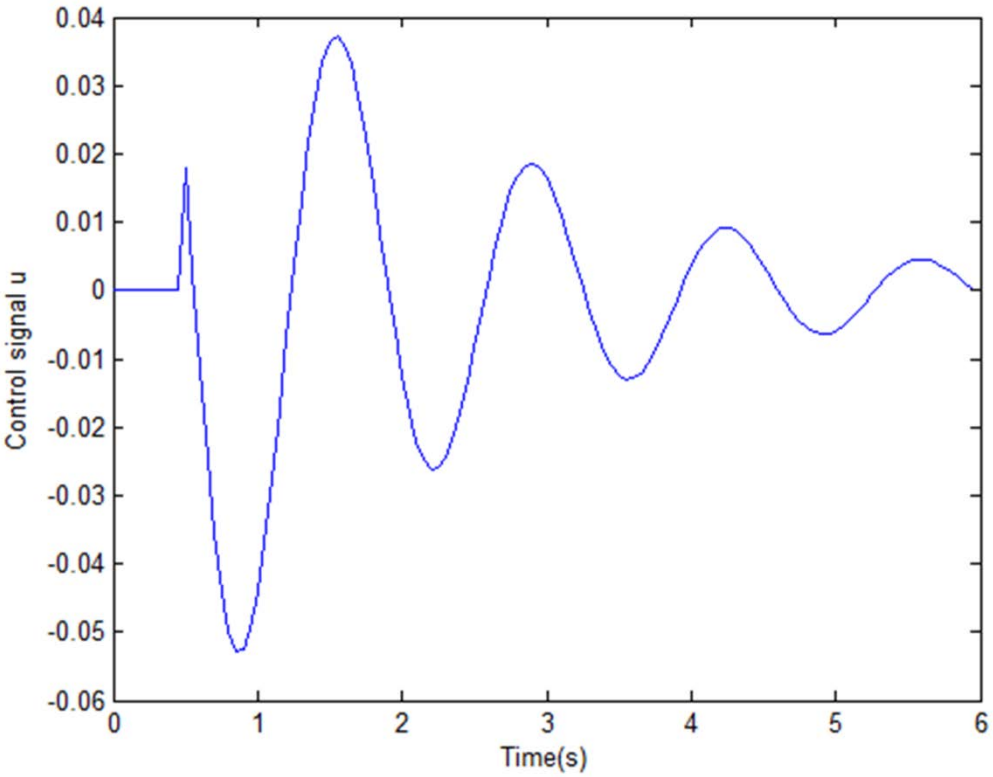


Fig. 4.46 Control signal.

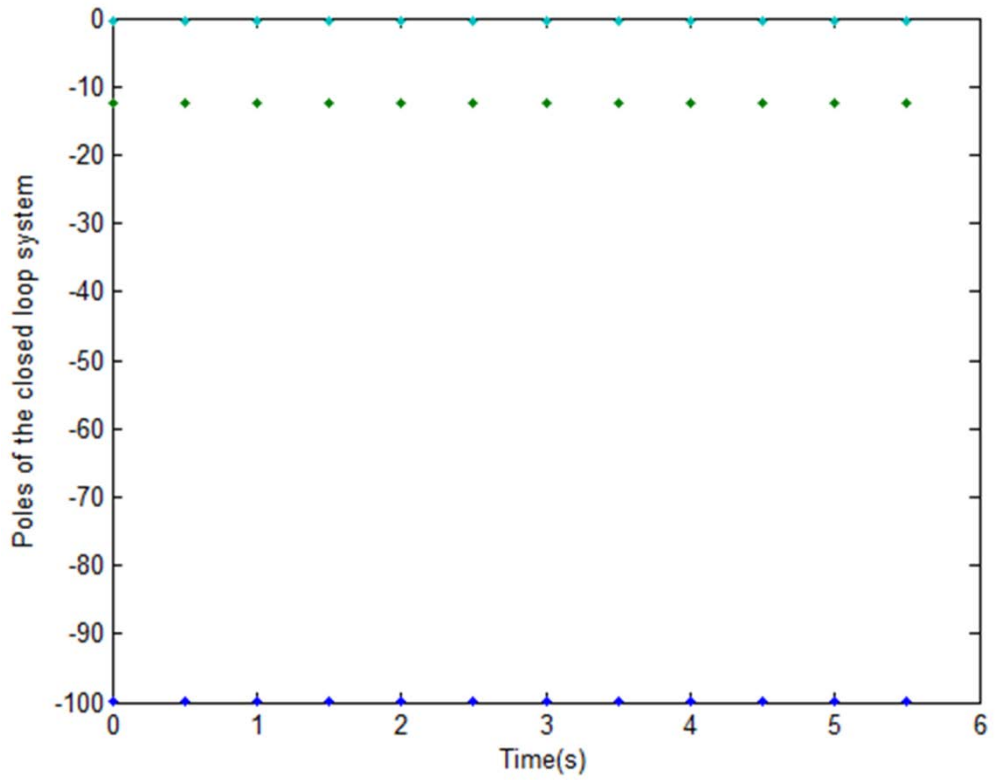


Fig. 4.47 Evolution of poles of closed loop system.

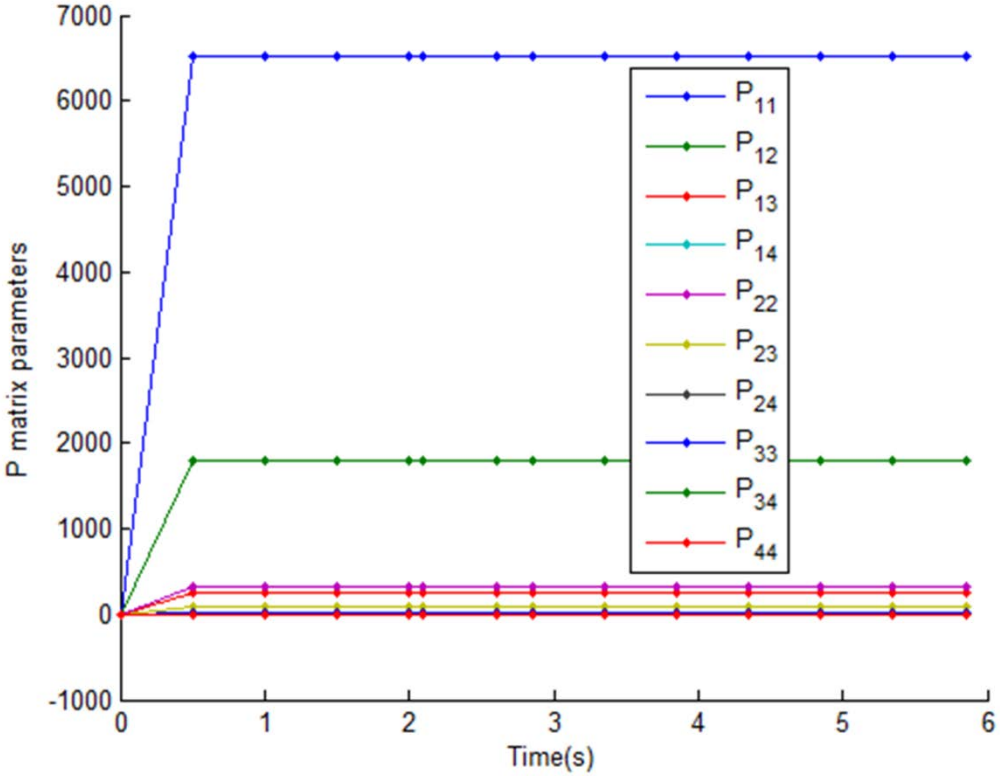


Fig. 4.48 Critic parameters.

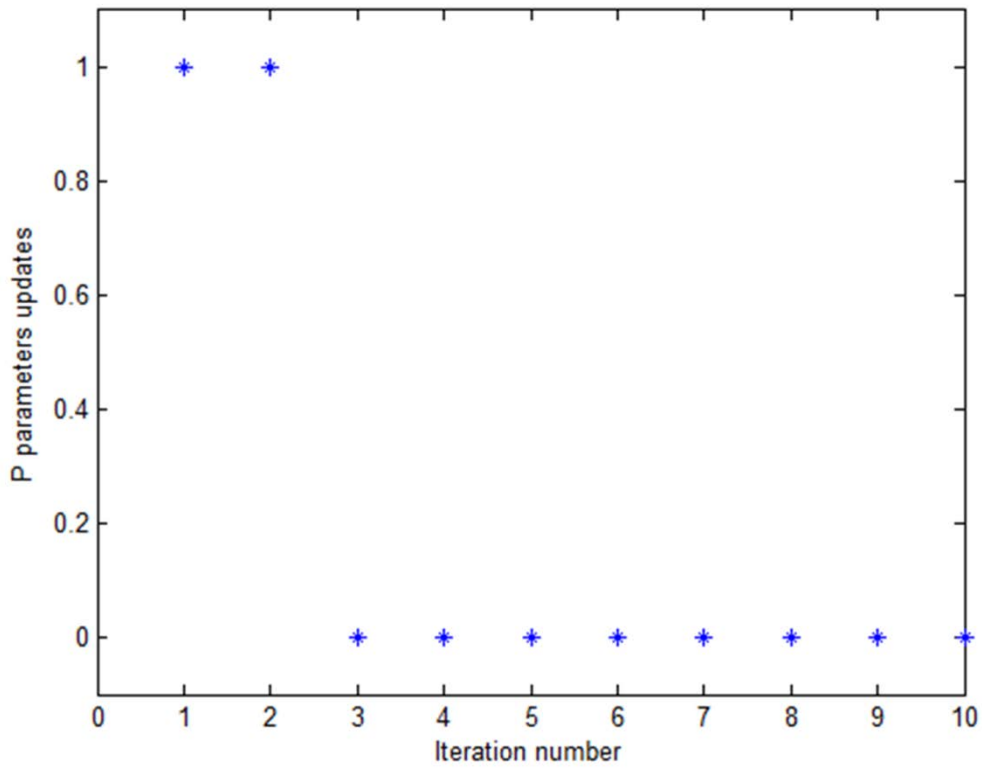


Fig. 4.49 Updating of critic parameters.

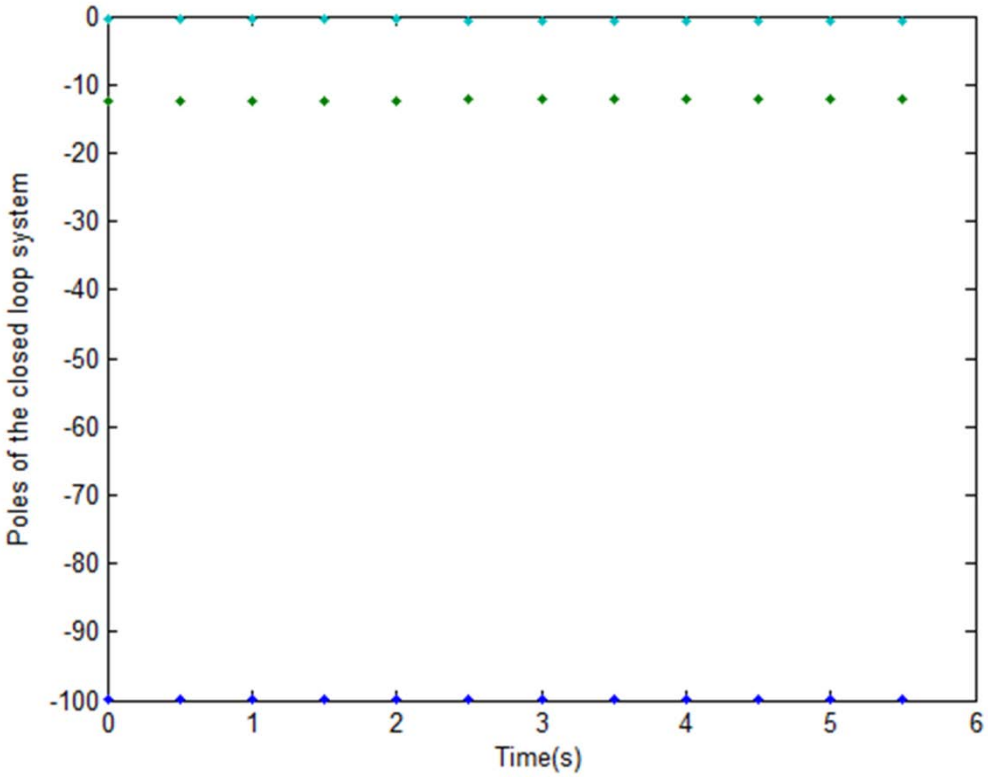


Fig. 4.50 Evolution of poles of closed loop system with change in system parameters at k=41.

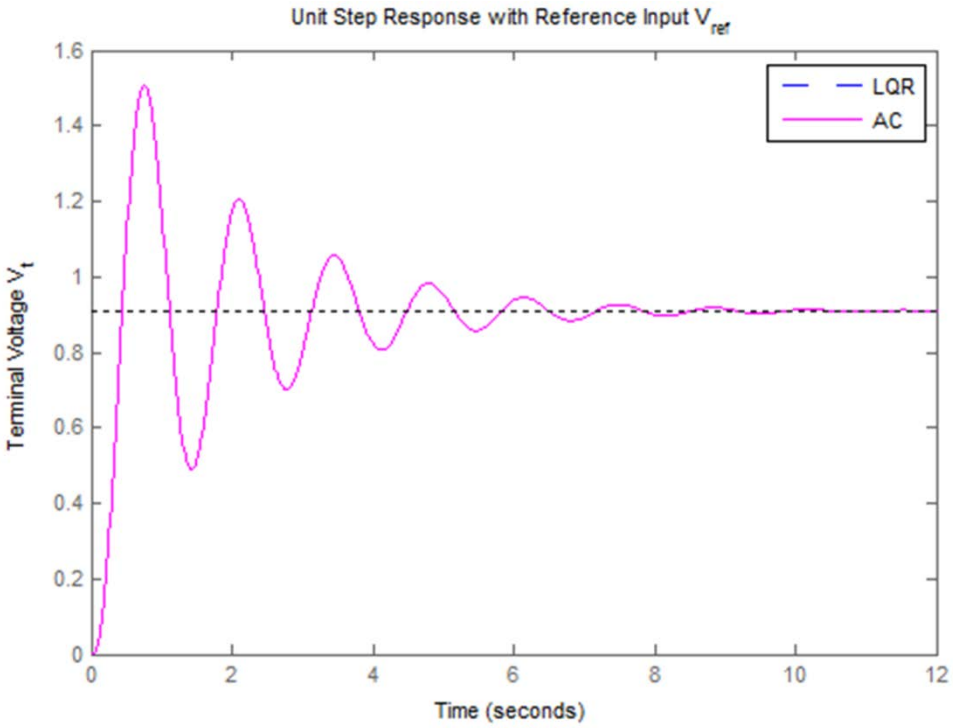


Fig. 4.51 Unit step response of closed loop system.

It is observed from the above simulation results for AVR system for both cases that critic parameter matrix *P* and actor parameter *K* obtained from PI based adaptive critic scheme are converging adaptively to optimal values and are mostly of same values of *RicP* and *RicK* respectively that obtained from LQR approach. Also in case of change in the system parameter in real situation the controller adapts it and converges to same optimal values. Thus the actor *K* and critic *P* parameters remain unchanged.

Analyzing the simulation results obtained for AVR system in both the cases of system models neglecting and including sensor dynamics, and changes in system parameters applying adaptive critic design using online PI technique, it is established that this proposed control scheme provide a promising adaptive optimal control solution for dynamical systems without complete knowledge of the system dynamics. The effect of system modelling uncertainties has been demonstrated by analyzing the simulation results for AVR system for both cases of neglecting and including sensor dynamics. It is observed that the simulation results obtained in both cases are similar. Thus it inference that the system model may be simplified by neglecting subsystems whose time constant is very small without considerably affecting the system characteristics and thus simplifying control design and its performance. The structural change introduced in system dynamics by including sensor dynamics is augmenting the system behaviour such as of its credit in closed loop response. The structural change in system will not be adapted by the proposed controller and in that case control algorithm has to be modified according to order of system dynamics. The proposed

controller adapts the change in system parameters in real situation at any moment of time. Thus this technique is partially model-free, effective & robust.

4.6.5 DC Motor Speed Control System

Consider the linearized state space model of speed control system of a separately excited dc motor with load disturbance model [7], for both of system models without and with integral compensator. Investigation of performance is also carried for change in system parameters at certain instant of time to demonstrate that the proposed control approach is partially model-free.

4.6.5.1 DC motor speed control system model without integral compensator

Consider the speed control system of armature controlled DC motor. The applied voltage V_a controls the angular velocity ω of the shaft, and T_L is the load disturbance torque. The block diagram of DC motor speed control system is shown in Fig. 2. The physical constants of dc motor speed control system model considered here in simulation are [204]: armature resistance, $R_a = 2$ Ohms; armature inductance, $L_a = 0.5$ Henrys; Torque constant, K_T = 0.1; Back EMF constant, K_b = 0.1; Viscous friction constant, *b* = *0.2* (Newton-m)/(rad/sec); and Moment of inertia, $J = 0.02$ kg.m^2/s^2.

The state space model of the system is derived as

$$\dot{x} = Ax + Bu + FT_L \tag{4.45}$$

$$y = \omega = Cx \tag{4.46}$$

where state vector $x = \begin{bmatrix} \omega & i_a \end{bmatrix}^T$, control input $u = V_a$,

$$A = \begin{bmatrix} -\frac{b}{J} & \frac{K_T}{J} \\ -\frac{K_b}{L_a} & -\frac{R_a}{L_a} \end{bmatrix},\ B = \begin{bmatrix} 0 \\ \frac{1}{L_a} \end{bmatrix},\ F = \begin{bmatrix} -\frac{1}{J} \\ 0 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

At the given nominal values of system parameters, we have

$$A = \begin{bmatrix} -10 & 5 \\ -0.2 & -4 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 2 \end{bmatrix},\ F = \begin{bmatrix} -50 \\ 0 \end{bmatrix}$$

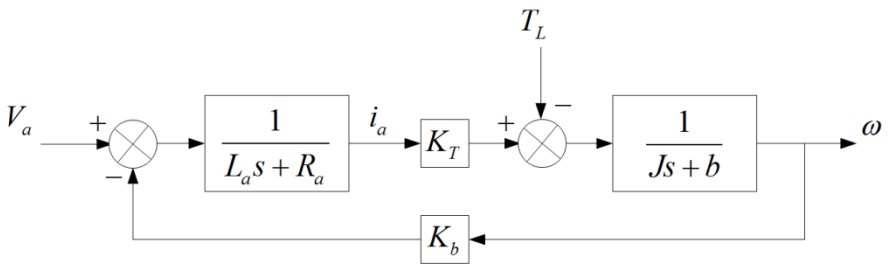


Fig. 4.52 Block diagram of DC motor speed control system.

For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = [0.8, 0.4, 0]$; $P = \begin{bmatrix} 0 & 0; & 0 & 0 \end{bmatrix}$. The length of simulation in samples is taken 30, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = \begin{bmatrix} 0.0496 & 0.0152 \\ 0.0152 & 0.1349 \end{bmatrix}, \quad P = \begin{bmatrix} 0.0496 & 0.0152 \\ 0.0152 & 0.1350 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK = \begin{bmatrix} 0.0304 & 0.2698 \end{bmatrix}, \quad K = \begin{bmatrix} 0.0304 & 0.2699 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-9.7497, -4.7901

Simulation with change in system parameters is also done at sample k=21; (i.e. t=1.05 seconds), such that *b*=0.3, *J*=0.025, *b*/*J* = 12, and 1/*J* = 40, and so the parameter *A*(1,1) = -12, then the solution is obtained as

$$RicP = \begin{bmatrix} 0.0415 & 0.0110 \\ 0.0110 & 0.1302 \end{bmatrix}, \quad P = \begin{bmatrix} 0.0496 & 0.0152 \\ 0.0152 & 0.1350 \end{bmatrix}$$

and the actor gains *RicK* and *K* respectively are obtained as

$$RicK = \begin{bmatrix} 0.0219 & 0.2605 \end{bmatrix}, \quad K = \begin{bmatrix} 0.0304 & 0.2699 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-11.8209, -4.7189

The simulation responses of adaptive optimal control using PI technique for DC motor speed control system model (4.45) & (4.46) are shown in Fig. 4.53 to 4.57. Fig. 4.53 shows system state trajectories which converge towards the equilibrium point. Fig. 4.54 shows control signal trajectory which also converge towards zero. Fig. 4.55 shows evolution of closed loop poles of the system during simulation. Fig. 4.56 shows convergence of critic parameters of matrix *P* towards optimal values. Fig. 4.57 shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update. Fig. 4.58 shows evolution of closed loop poles for case of change in system parameters during simulation. Rest of responses remains similar as of Fig. 4.53 to 4.57.

Fig. 4.59 presents closed loop response of DC motor speed control system without integral compensator model using both approaches of LQR and adaptive critic (AC). It is

observed that the closed loop responses with unit step reference input ω_r , and unit step load disturbance torque ΔT_L have- steady state error.

Fig. 4.60 shows closed loop response with load disturbance torque T_L=-0.1 N-m between the time period t=[4, 8] seconds. Figs. 4.61 and 4.62 show closed loop responses as above for the case of change in system parameters during simulation. Here it is observed that the adaptive optimal controller using PI based adaptive critic scheme gives similar responses as of standard LQR.

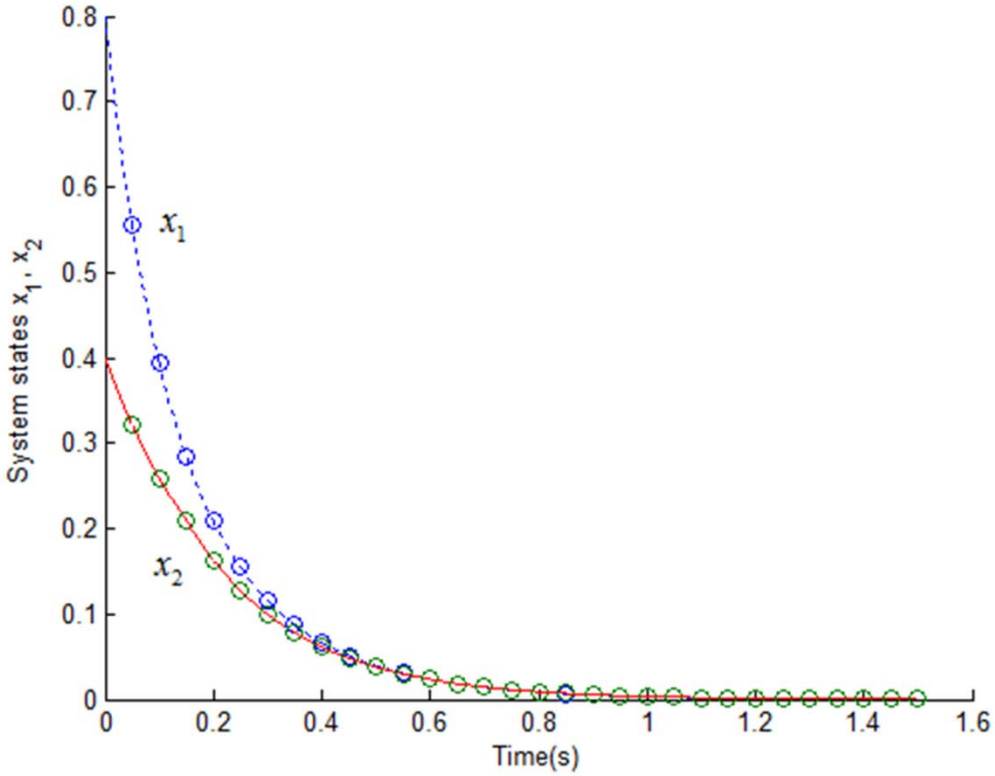


Fig. 4.53 System states

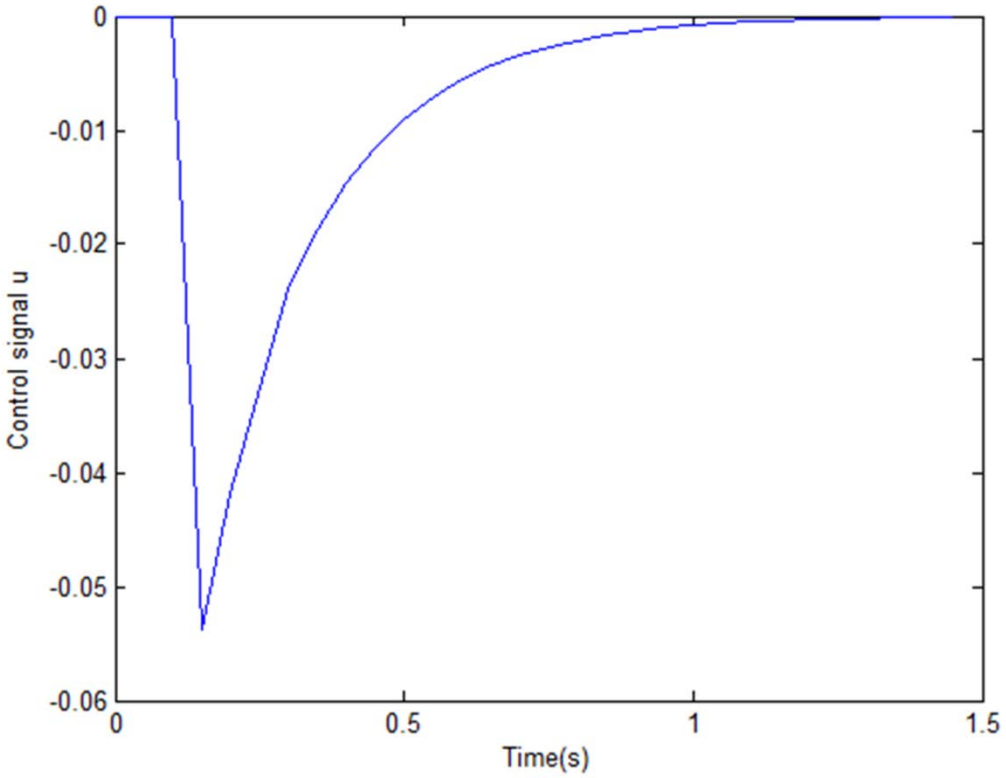


Fig. 4.54 Control signal

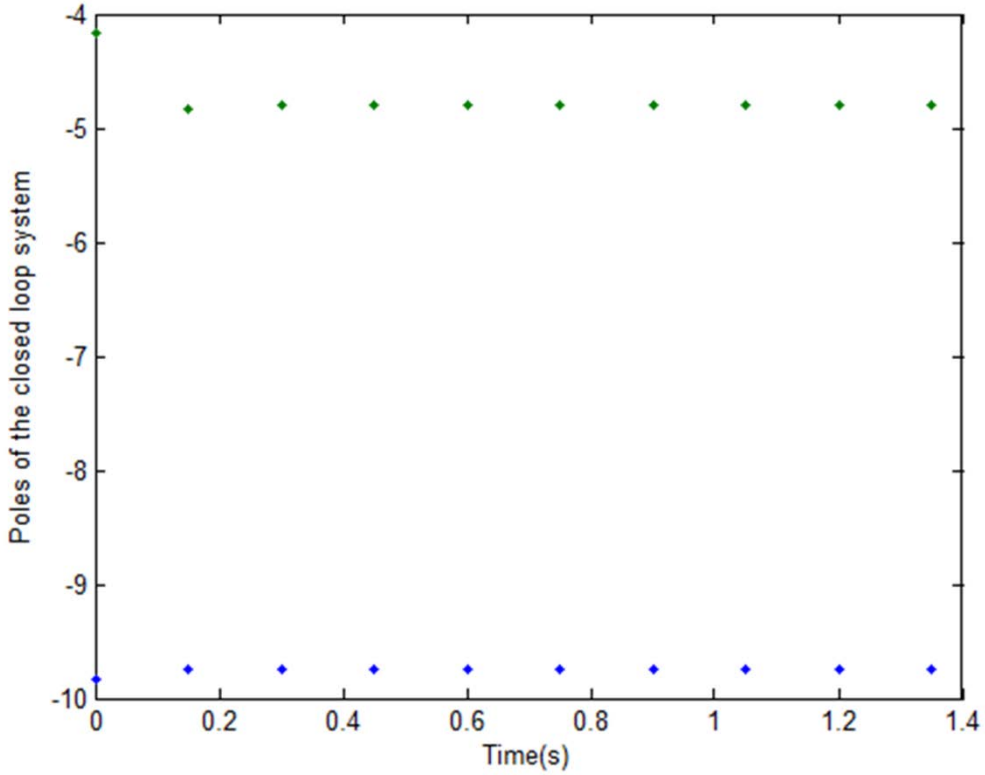


Fig. 4.55 Evolution of poles of closed loop system

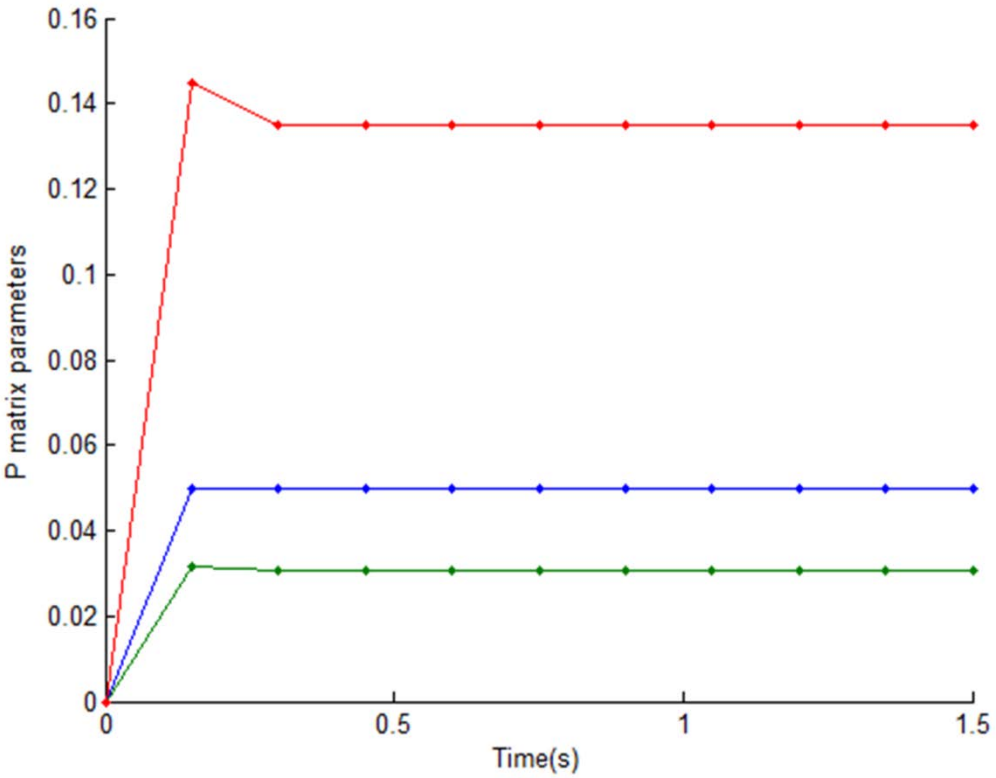


Fig. 4.56 Critic parameters

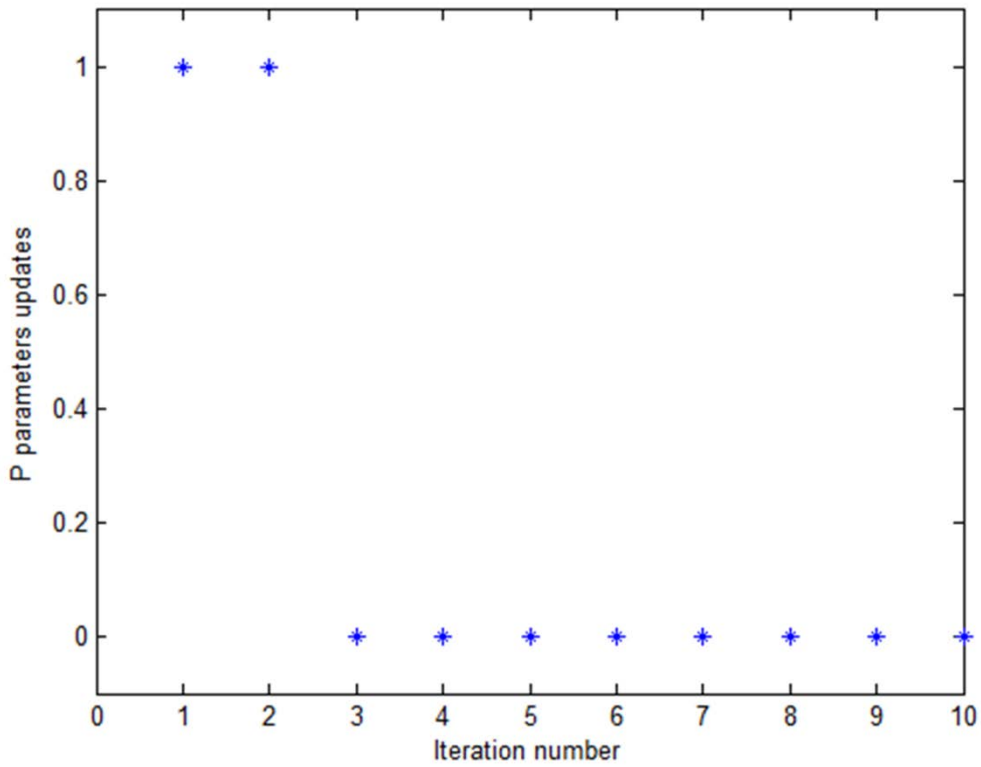


Fig. 4.57 Updating of critic parameters

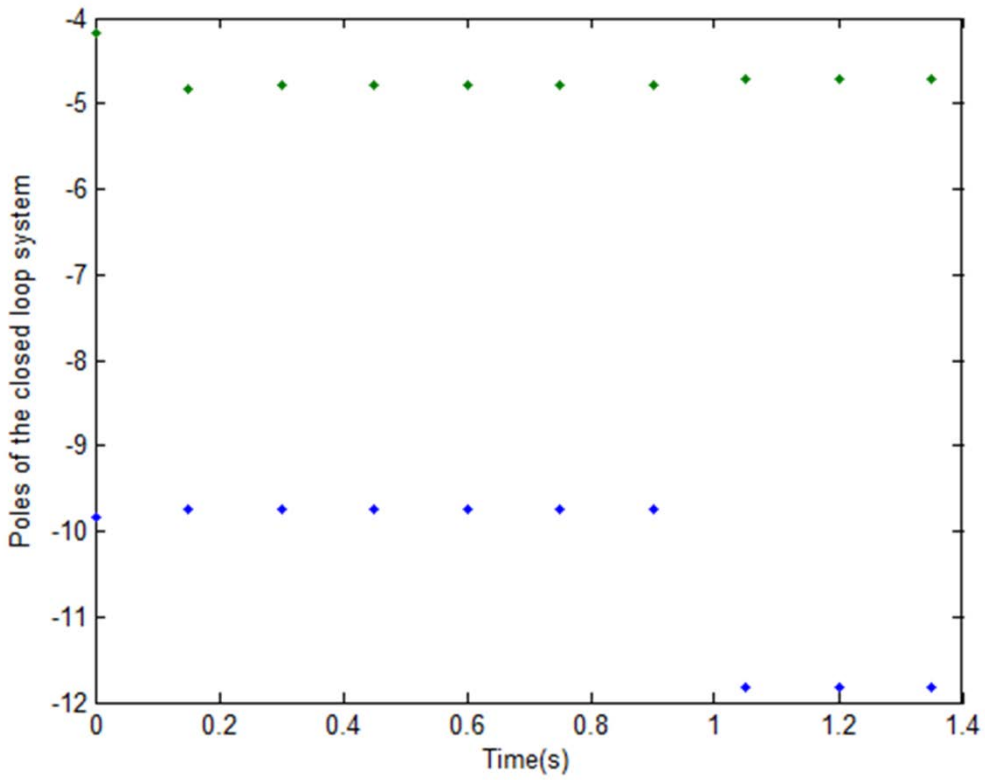


Fig. 4.58 Evolution of poles of closed loop system with change in system parameter at sample k=21 (i.e. t=1.05 seconds), A(1,1)=-12.

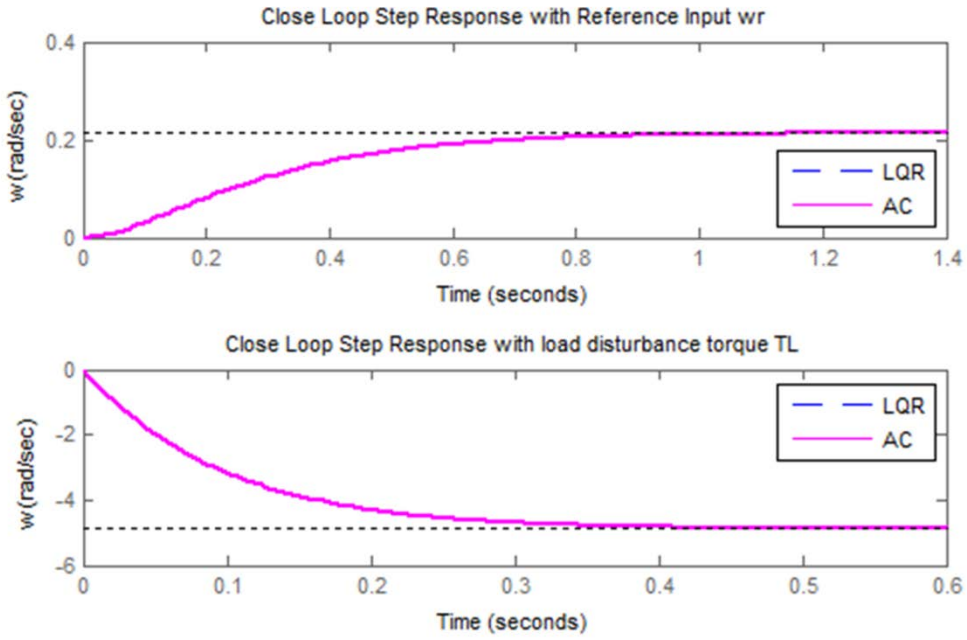


Fig. 4.59 Closed loop unit step response of DC motor speed.

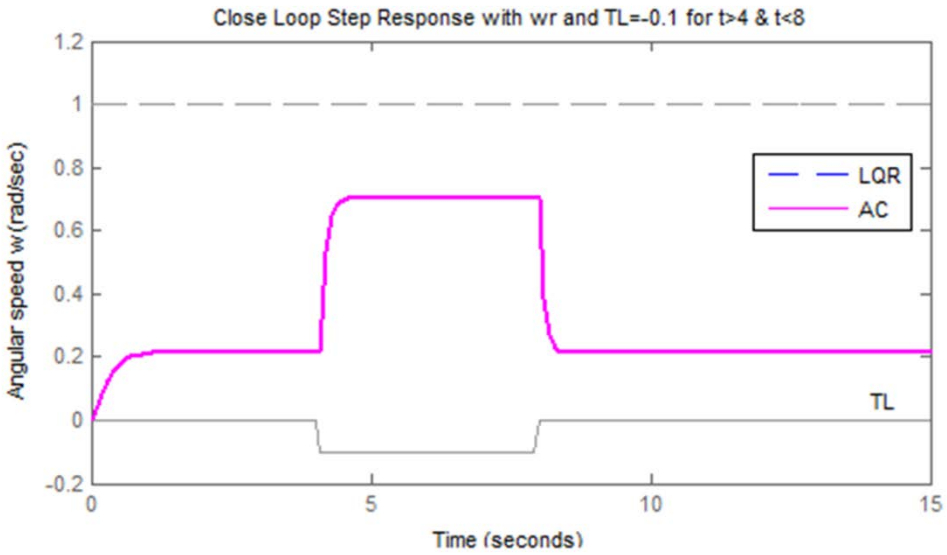


Fig. 4.60 DC motor speed response with load disturbance for time t=4 to 8 seconds.

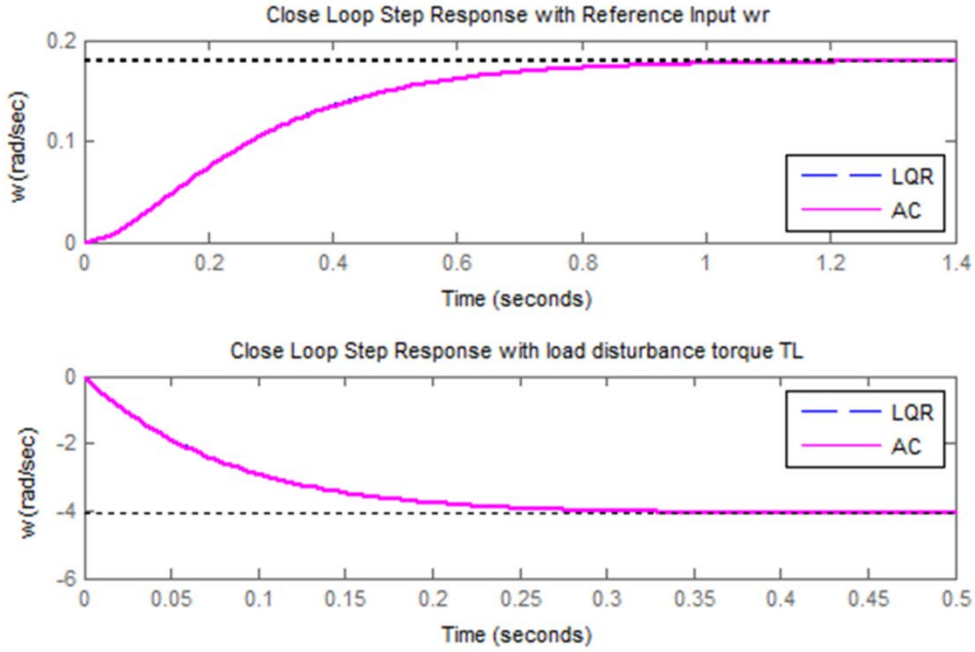


Fig. 4.61 Closed loop unit step response of DC motor speed for system model with change in system parameters.

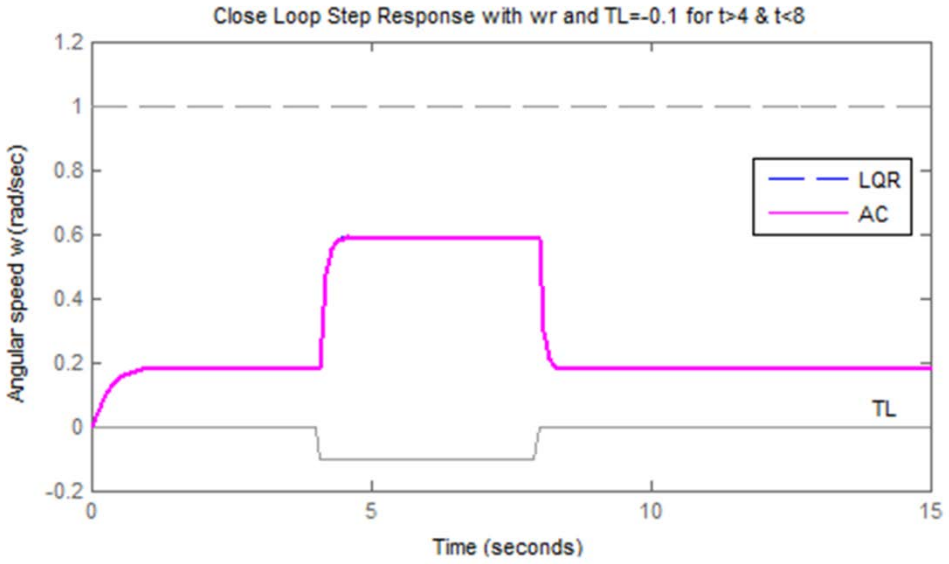


Fig. 4.62 DC motor speed response with change in system parameters and with load disturbance for time t=4 to 8 seconds.

4.6.5.2 DC motor control system model with integral compensator

Fig. 4.63 shows the block diagram of DC motor speed control system with integral compensator for speed. The augmented system state space model with integral control dynamics to remove the steady-state error, the third state variable is taken as $E = \int_0^t (\omega_r - \omega)dt$, and then $V_a = u + K_I E$, which makes 3rd column of the system matrix *A* having "non-all-elements-zero" column. The value of K_I is chosen by root locus design for a stable closed-loop system.

Thus, we have the state space model of the system with integral compensator by modifying (4.45) & (4.46) with state vector $x = \begin{bmatrix} \omega & i_a & E \end{bmatrix}^T$, disturbance input vector $w = \begin{bmatrix} T_L & \omega_r \end{bmatrix}^T$,

$$A = \begin{bmatrix} -\frac{b}{J} & \frac{K_T}{J} & 0 \\ -\frac{K_b}{L_a} & -\frac{R_a}{L_a} & \frac{K_I}{L_a} \\ -1 & 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ \frac{1}{L_a} \\ 0 \end{bmatrix},\ F = \begin{bmatrix} -\frac{1}{J} & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Thus, with above parameters values the system matrices are

$$A = \begin{bmatrix} -10 & 5 & 0 \\ -0.2 & -4 & 10 \\ -1 & 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix},\ F = \begin{bmatrix} -50 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

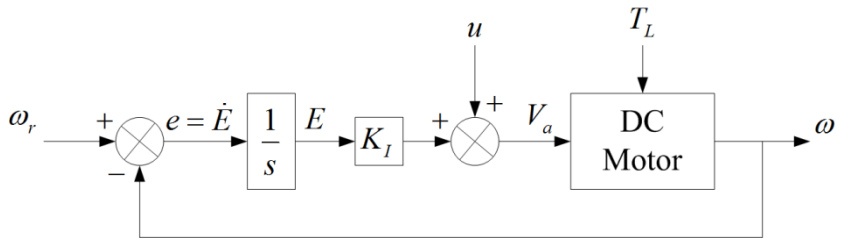


Fig. 4.63 Block diagram of DC motor speed control system with integral compensator.

For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = [0.8, 0.4, 0.2, 0]$; $P = [0\ 0\ 0;\ 0\ 0\ 0;\ 0\ 0\ 0]$. The length of simulation in samples is taken 60, and sample time *T*=0.05 seconds. The cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The unique positive definite solution of ARE (4.5), denoted here by matrix *RicP*, and adaptive optimal critic matrix *P* of adaptive critic scheme using PI in (4.9) & (4.10) with (4.22) respectively are obtained as

$$RicP = \begin{bmatrix} 0.0840 & 0.0301 & -0.3479 \\ 0.0301 & 0.1512 & -0.0495 \\ -0.3479 & -0.0495 & 3.7961 \end{bmatrix},\ P = \begin{bmatrix} 0.0840 & 0.0301 & -0.3481 \\ 0.0301 & 0.1513 & -0.0495 \\ -0.3481 & -0.0495 & 3.7982 \end{bmatrix}$$

, and the actor gains of LQR design by (4.4) & (4.5) denoted here by *RicK* , and actor *K* by adaptive critic scheme using PI in (4.10) respectively are obtained as

$$RicK = [0.0602 \quad 0.3024 \quad -0.0990],\ K = [0.0602 \quad 0.3025 \quad -0.0990]$$

The eigenvalues of closed loop system are obtained as

-10.5444, -2.0303 + 0.8448i, -2.0303 - 0.8448i

Simulation with change in system parameters is also done at sample k=21; (i.e. t=1.05 seconds), such that *b*=0.3, *J*=0.025, *b*/*J* = 12, and 1/*J* = 40, and so the parameter *A*(1,1) = -12, then the solution is obtained as

$$RicP = \begin{bmatrix} 0.0687 & 0.0220 & -0.3300 \\ 0.0220 & 0.1424 & -0.0495 \\ -0.3300 & -0.0495 & 4.1942 \end{bmatrix},\ P = \begin{bmatrix} 0.0840 & 0.0301 & -0.3481 \\ 0.0301 & 0.1513 & -0.0495 \\ -0.3481 & -0.0495 & 3.7982 \end{bmatrix}$$

and the actor gains *RicK* and *K* respectively are obtained as

$$RicK = [0.0440 \quad 0.2847 \quad -0.0990],\ K = [0.0602 \quad 0.3025 \quad -0.0990]$$

The eigenvalues of closed loop system are obtained as

-12.3281, -1.4775, -2.7994

The simulation responses of adaptive optimal control using PI technique for DC motor speed control system model with integral compensator are shown in Fig. 9. Fig. 9(a) shows system state trajectories which converge towards the equilibrium point. Fig. 9(b) shows control signal trajectory which also converge towards zero. Fig. 9(c) shows evolution of

closed loop poles of the system during simulation. Fig. 9(d) shows convergence of critic parameters of matrix *P* towards optimal values. Fig. 9(e) shows *P* parameters updating with iteration, here * at one indicate update, and * at zero indicate no update. The simulation responses of adaptive optimal control using PI based adaptive critic scheme for DC motor speed control system model with integral compensator for the case of change in system parameter are similar as above and just the evolution of poles changes adapting the controller. Fig. 9(f) shows evolution of closed loop poles in the case of change in system parameter.

Fig.10 presents the closed loop responses of DC motor speed control system with integral compensator model using both approaches of LQR and adaptive critic (AC). It is observed that the closed loop responses with unit step reference input ω_r , and unit step load disturbance torque ΔT_L have no steady state error. Fig. 11 shows closed loop response with load disturbance torque T_L=-0.1 N-m between the time period t=[4, 8] seconds. Here also it is observed that the adaptive optimal controller using PI based adaptive critic scheme gives similar responses as of standard LQR.

Fig. 12 shows the closed loop responses of DC motor speed control system with integral compensator model using both approaches of LQR and adaptive critic (AC) for the case of change in parameters. The controller performs adapting the change in system parameters. Fig. 13 shows closed loop response with load disturbance torque T_L=-0.1 N-m between the time period t=[4, 8] seconds. It is observed in this case that the disturbance rejection is much better than the case of system model without integral compensator. The inclusion of integral compensator in the system model removes the steady state error and also improves the disturbance rejection capability of controlled system. It is also observed here that the adaptive optimal controller using PI based adaptive critic scheme gives similar responses as of standard LQR. This demonstrates that the proposed control scheme is effective & robust.

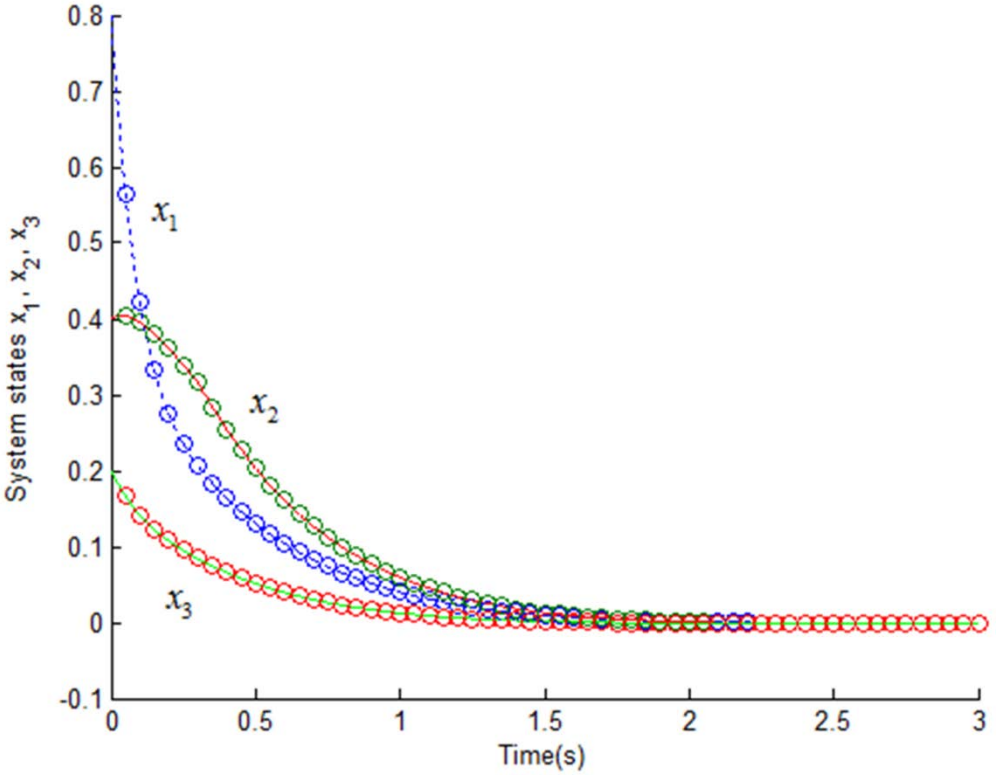


Fig. 4.64 System states

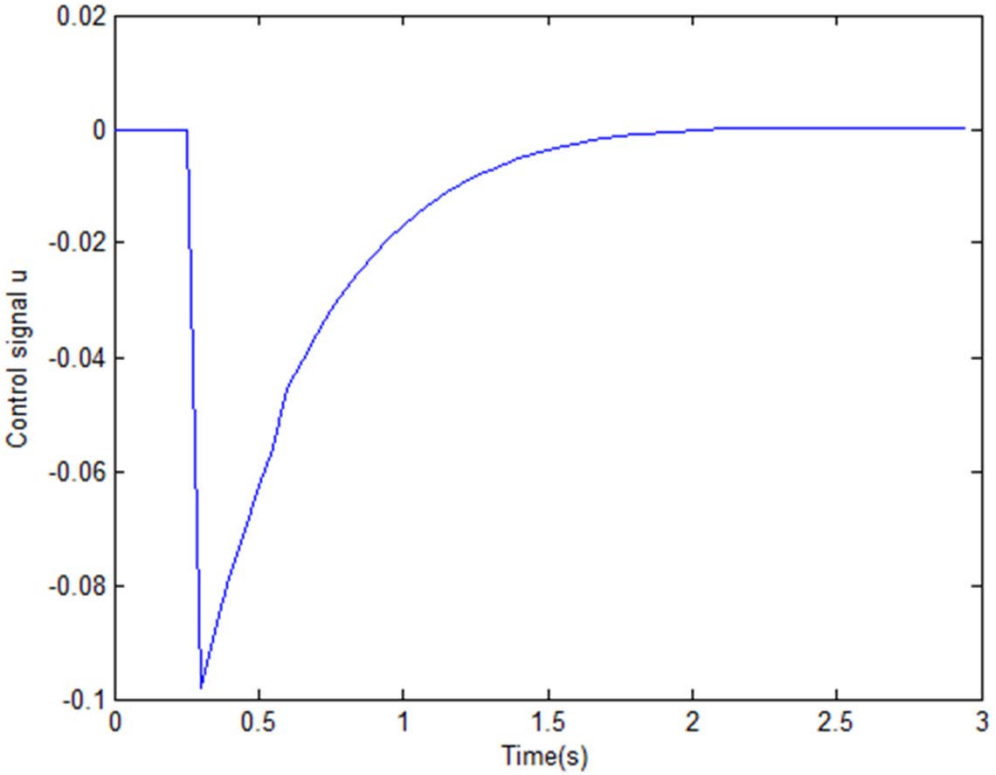


Fig. 4.65 Control signal

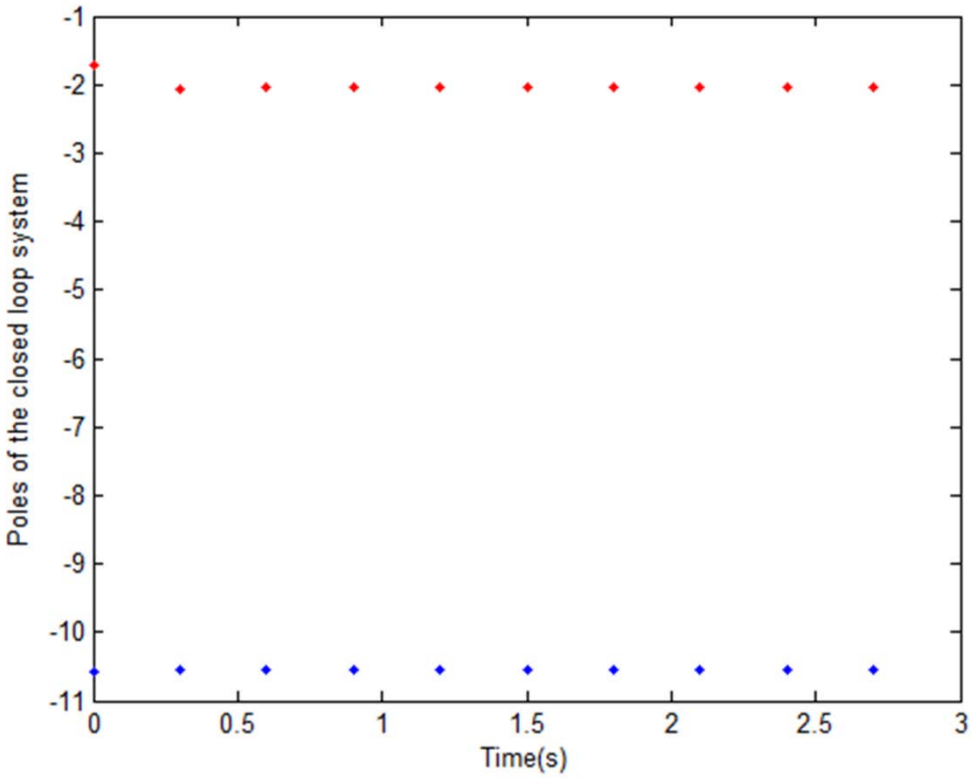


Fig. 4.66 Evolution of poles of closed loop system

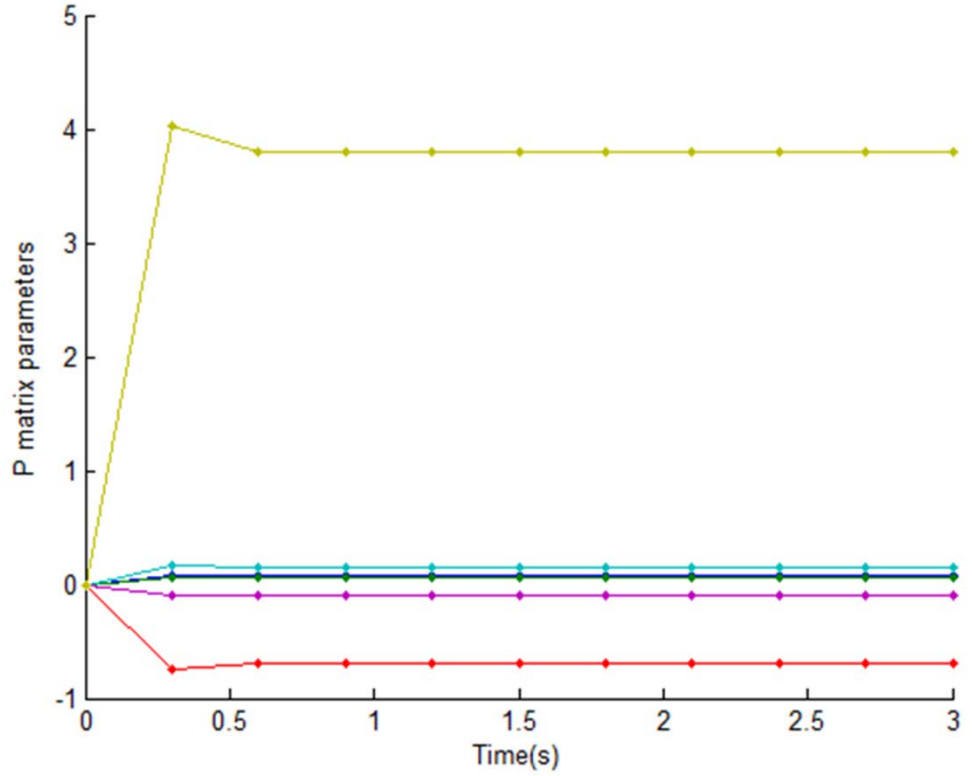


Fig. 4.67 Critic parameters

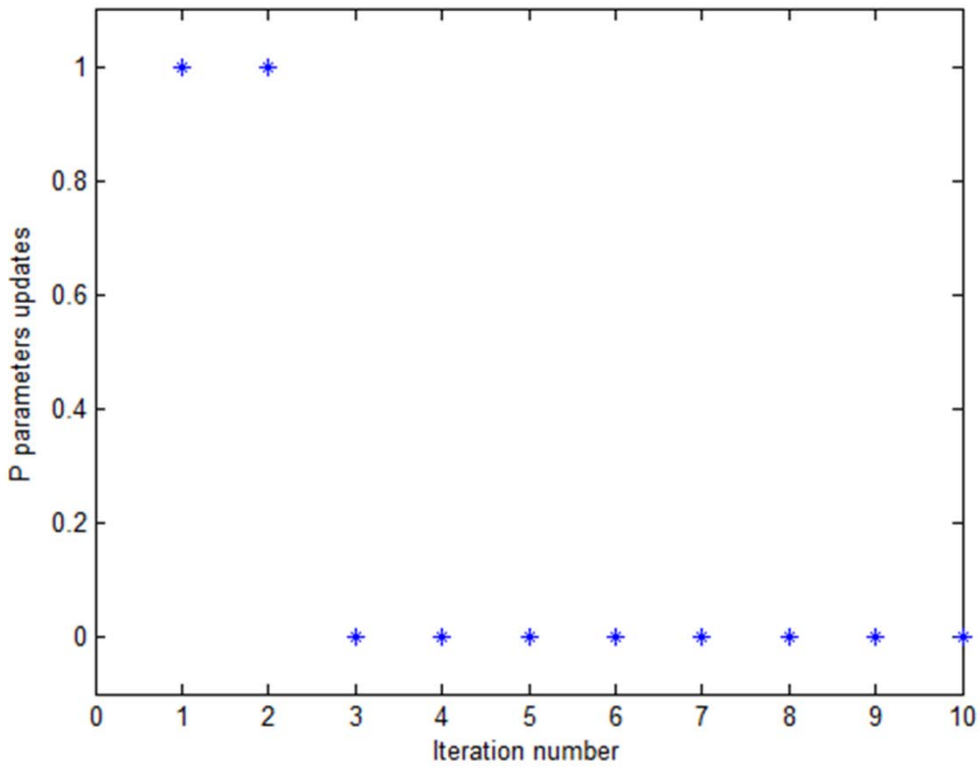


Fig. 4.68 Updating of critic parameters

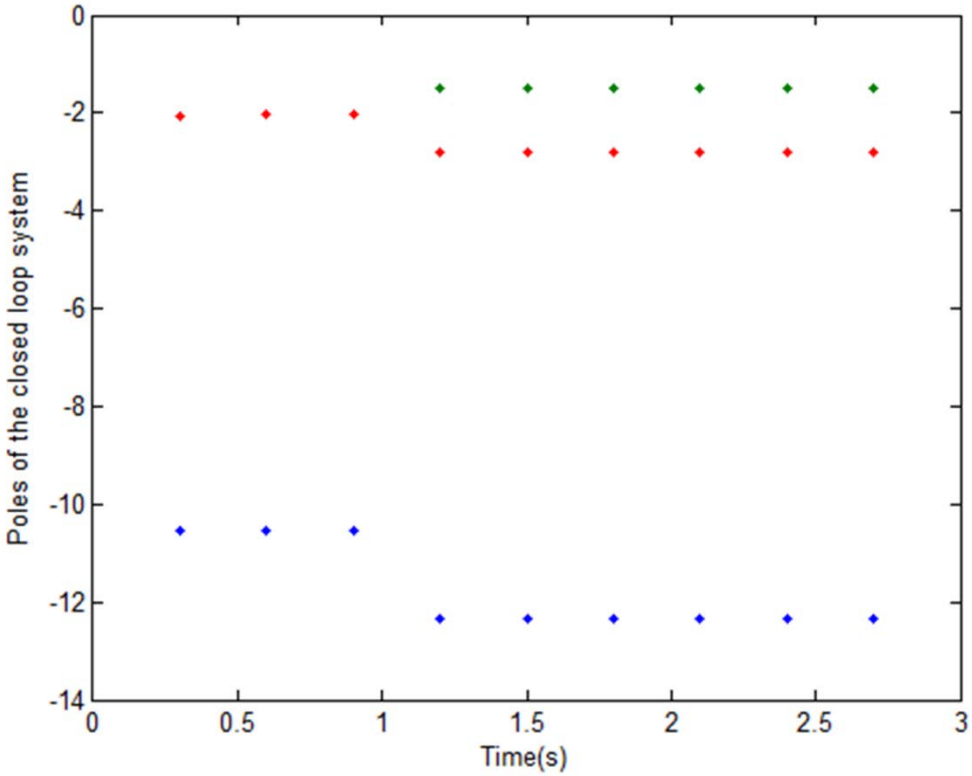


Fig. 4.69 Evolution of poles of closed loop system with change in system parameter at sample k=21 (i.e. t=1.05 seconds), A(1,1)=-12.

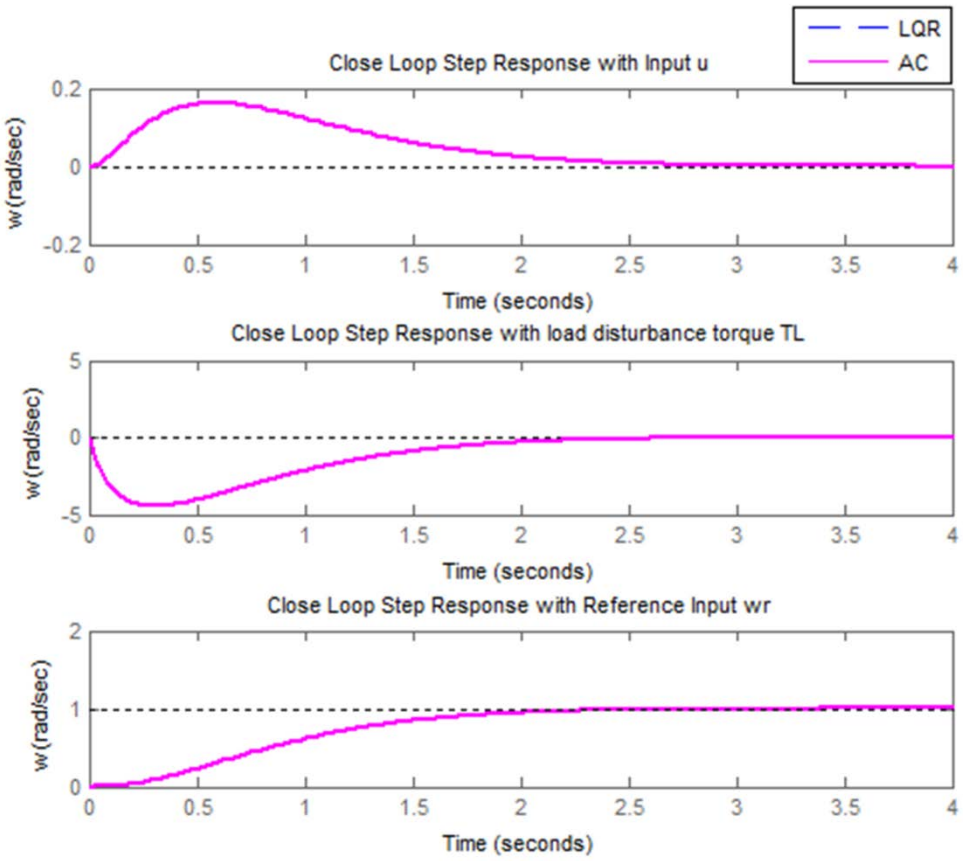


Fig. 4.70 Closed loop unit step response of DC motor speed control for system model with integral compensator.

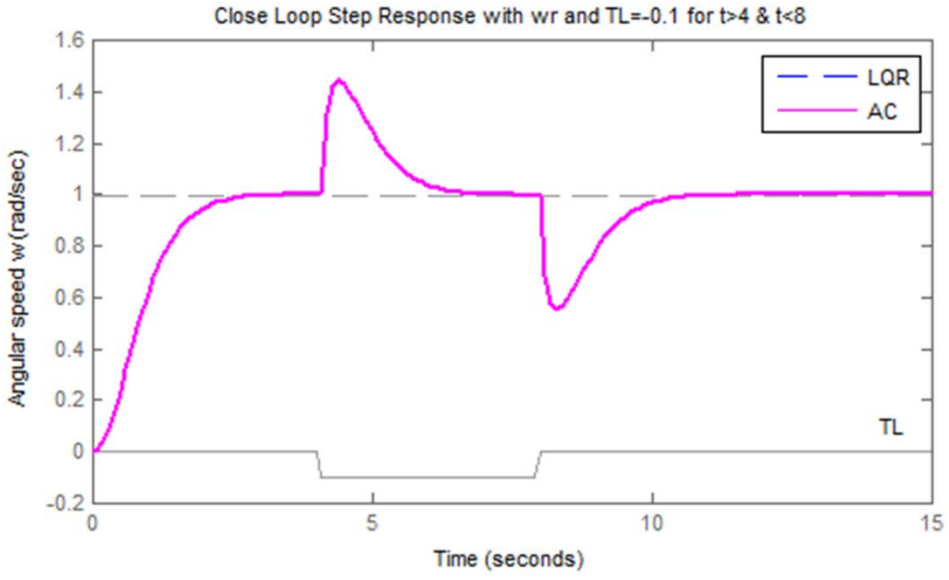


Fig. 4.71 DC motor speed response with load disturbance for time t=4 to 8 seconds for system model with integral compensator.

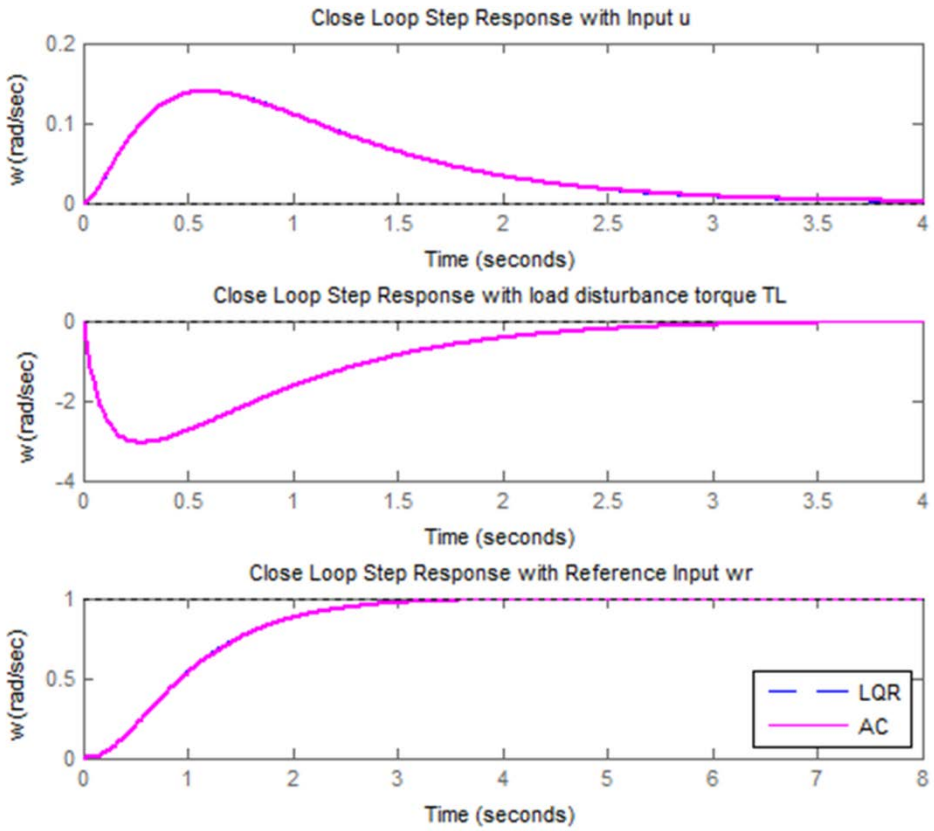


Fig. 4.72 Closed loop unit step response of DC motor speed for system model with integral compensator with change in system parameters.

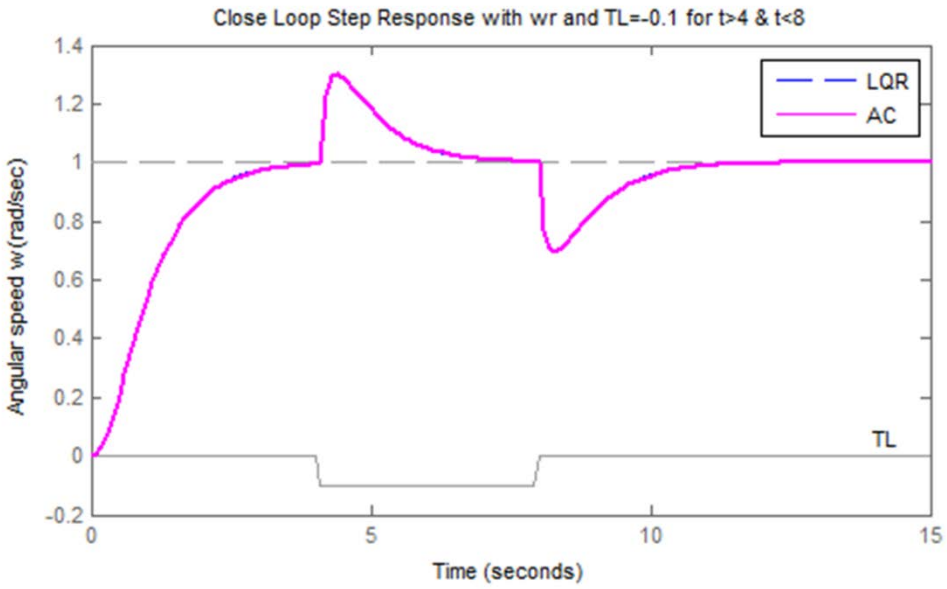


Fig. 4.73 DC motor speed response with load disturbance for time t=4 to 8 seconds for system model with integral compensator and change in system parameter.

In the above simulation results for both cases it is observed that critic parameter matrix *P* and actor parameter *K* obtained from adaptive critic scheme using PI technique are converging adaptively to optimal values and values of *RicP* and *RicK* respectively are mostly

same to that obtained from LQR approach. Also in case of change in system parameter in real situation the controller adapts it and converges to same optimal values. Thus the actor *K* and critic *P* parameters remain unchanged.

Analyzing the simulation results obtained for both cases of DC motor control system models without & with integral compensator, and changes in system parameters applying adaptive critic scheme using online PI technique, it is established that the proposed control scheme provide a promising adaptive optimal control solution for dynamical systems without complete knowledge of the system dynamics. The structural change introduced in system dynamics by including integral compensator is augmenting the system behavior such as of its credit that removing the steady state error in closed loop responses. The structural change in system will not be adapted by the proposed controller but it will adapt the change in system parameters in real situation at any moment of time. Thus, this control scheme is partially model-free, effective & robust.

4.7 CONCLUSIONS

The comprehensive performance analysis of adaptive optimal control using policy iteration technique for continuous-time LTI systems has been presented in this chapter. The infinite horizon control using online PI technique based adaptive critic scheme gives an adaptive optimal control solution to the real-time dynamics of a continuous-time LTI system. The infinite horizon optimal control solution using ARE requires the complete knowledge of the system dynamics, which is not required using PI technique based adaptive critic scheme. Online PI algorithm does not require the knowledge of system's internal dynamics (i.e. matrix *A*) for evaluation of cost or the update of control policy; only the knowledge of input matrix *B* is required for updating the control policy. Thus, adaptive optimal control scheme using online policy iteration technique is partially model-free. The convergence of the proposed algorithm, under the condition of initial stabilizing controller, to the solution of the state feedback optimal control problem has been established. To demonstrate the applications of control scheme, the systems modeling, simulation results, and performance analysis are presented for certain general and practical example of LTI systems- a general LTI SISO system, a 4^{th} order mechanical system, LFC system without and with integral control, AVR system neglecting and including sensor dynamics, and DC motor speed control system without and with integral compensator, and also for all the examples with change in system parameters at certain instant of time. The simulation results justify the effectiveness & robustness of the online adaptive optimal control scheme using PI technique.

CHAPTER 5

ADAPTIVE OPTIMAL CONTROL USING POLICY ITERATION TECHNIQUE FOR AFFINE NONLINEAR SYSTEMS

In this chapter the adaptive critic design using policy iteration technique with neural network approximation of cost function for adaptive optimal control of continuous-time affine nonlinear dynamical systems is presented. The application of control scheme is implemented for both general and practical examples of affine nonlinear systems.

5.1 INTRODUCTION

Adaptive critic (AC) based control design has evolved as a powerful and promising technique to solve optimal control problems for nonlinear systems. Adaptive critic designs (ACD) which have potential of replicating critical aspects of human intelligence such as the ability to cope with a large number of variables in parallel, in real-time, and in a noisy nonlinear non-stationary environment, have proved to provide brain-like intelligence on a very small scale but good enough for solving engineering problems. The concept of ACD utilizes two parametric structures known as the actor and the critic to efficiently approximate the optimal cost and control using incremental optimization of a short-term cost metric that ensures optimization of the cost over all future times. The actor parameterizes the control policy. The critic approximates a value-related cost function and captures the effect that the control law will have on the future cost which describes the performance of control system. At any given time, the critic provides guidance to improve the control policy, and the actor to update the critic.

The infinite horizon optimal control design using HJB equation requires the complete knowledge of system dynamics with an offline solution. The policy iteration (PI) technique which is based on actor-critic structure consists of two-step iteration: policy evaluation and policy improvement. The online PI algorithm solves the optimal control problem along a single state trajectory, does not require the knowledge of system internal state dynamics, and thus giving a direct adaptive optimal control scheme. The online PI technique provides an adaptive optimal control solution for an infinite horizon problem subject to the real-time dynamics of a continuous-time system. Using neural network approximation of cost function for online implementation, the PI algorithm gives a high-level intelligent control.

Adaptive optimal control using various approaches has been presented recently in [122-178]. The ACD for various control applications applying certain approaches in both discrete-time and continuous-time frameworks are described recently in certain papers [122-126, 128-130, 132, 135-178]. Adaptive optimal control using PI technique is presented for linear systems in [155-158, 161, 162, 166, 169], and for nonlinear systems using neural networks in actor-critic configuration in [123, 158-165, 167]. The performance analysis of PI

based adaptive critic control scheme with applications to practical affine nonlinear systems is desired.

In this chapter the performance analysis of adaptive optimal control using PI technique with neural network approximation of cost function for continuous-time affine nonlinear systems is presented considering state regulation problems of certain general and practical examples of affine nonlinear systems- single link manipulator, inverted pendulum-cart system, and Vander Pol's oscillator system.

5.2 INFINITE HORIZON OPTIMAL CONTROL OF CONTINUOUS-TIME AFFINE NONLINEAR SYSTEMS

Consider a continuous-time affine nonlinear system given by

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \tag{5.1}$$

with $x(t) \in \mathcal{R}^n$, $f(x(t)) \in \mathcal{R}^n$, $g(x(t) \in \mathcal{R}^{n\times m}$ and the input $u(t) \in U \subset \mathcal{R}^m$, and the boundary conditions are as $x(t_0) = x_0$ fixed and $x(t_f)$ free.

It is assumed here that $f(0) = 0$, such that $f(x) + g(x)u$ is Lipschitz continuous on a set $\Omega \subseteq \mathcal{R}^n$ which contains the origin, and that the dynamical system is stabilizable on Ω, i.e. there exists a continuous control function $u(t) \in U$ such that the system is asymptotically stable on Ω.

It is required here to find a control input that minimizes the performance index given by

$$V(x(t_0), t_0) = \phi(x(T), T) + \int_{t_0}^{T} L(x(\tau), u(\tau))d\tau \tag{5.2}$$

The utility function *L* is given by

$$L(x,u) = Q(x) + u^T R u \tag{5.3}$$

with $Q(x)$ being positive definite, i.e. $\forall x \neq 0,\ Q(x) > 0$ and $x = 0 \Rightarrow Q(x) = 0$, and $R \in \mathcal{R}^{m\times m}$ a positive definite matrix.

Let us define a scalar function $V^*(x^*(t), t)$ as the optimal value of the performance index *V* for an initial state $x^*(t)$ at time *t*, i.e.

$$V^*(x^*(t), t) = \phi(x(T), T) + \int_{t}^{T} L(x^*(\tau), u^*(\tau), \tau)d\tau \tag{5.4}$$

At the boundary conditions it is given that

$$V^*(x^*(T), T) = \phi(x^*(T), T) \tag{5.5}$$

It is noticeable here that although the global asymptotic stability is guaranteed in a case of linear dynamical system, it is generally difficult to guarantee in a case of general continuous-time nonlinear dynamical system due to its non-smooth nature at the points of discontinuities of $\dot{x}$, and of the gradient of cost function. Thus, considering here to the case

in which asymptotic stability is desired and the cost function is continuously differentiable in a region $\Omega \subseteq \boldsymbol{\mathcal{R}}^n$.

The infinite horizon integral cost function associated with the control input $\{u(\tau); \tau \geq t_0\}$ is defined as

$$V^u(x(t_0)) = \int_{t_0}^{\infty} L(x(\tau), u(\tau)) d\tau \tag{5.6}$$

where $x(\tau)$ denotes the solution of (5.1) for initial condition $x(t_0) \in \Omega$ and input $\{u(\tau); \tau \geq t_0\}$. For the cost function which explicitly depend on time the optimal value function $V^*(x^*, t)$ must satisfy the HJB equation given by

$$\frac{\partial V^*}{\partial t} + \min_{\boldsymbol{u}} H\left(\boldsymbol{x}, \frac{\partial V^*}{\partial \boldsymbol{x}}, \boldsymbol{u}, t\right) = 0 \tag{5.7}$$

where the second term of (5.7) is the Hamiltonian functional. It provides the solution to the optimal control problem for general nonlinear dynamical systems.

Definition 1: Admissible (Stabilizing) Policy [159, 160]: A control policy $\mu(x)$ is defined as admissible with respect to (5.6) on Ω, denoted by $\mu \in \Psi(\Omega)$ if $\mu(x)$ is continuous on Ω, $\mu(0) = 0$, $\mu(x)$ stabilizes (5.1) on Ω and $V(x_0)$ is finite $\forall x_0 \in \Omega$.

The cost function associated with any admissible control policy $\mu \in \Psi(\Omega)$ is given by

$$V^{\mu}(x(t)) = \int_{t}^{\infty} L(x(\tau), \mu(x(\tau))) d\tau \tag{5.8}$$

where $V^{\mu}(x)$ is continuously differentiable belonging to the class C^1 functions. A C^1 function is exactly a function whose derivative exits and is of class C^0 that consist of all continuous functions. The infinitesimal version of (5.8) is written as

$$0 = L(x, \mu(x)) + \lambda^T (f(x) + g(x)\mu(x)) \ , \ V^{\mu}(0) = 0 \tag{5.9}$$

where $\lambda = \dfrac{\partial V^{\mu}}{\partial x}$ is the costate vector which denotes the gradient of the cost function V^{μ} with respect to *x*, as the cost function does not depend explicitly on time. Equation (5.9) is a Lyapunov equation for nonlinear systems which can be solved for the cost function $V^{\mu}(x)$ for the given controller $\mu(x) \in \Psi(\Omega)$. If $V^{\mu}(x)$ satisfies (5.9), with $L(x, \mu(x)) \geq 0$, then $V^{\mu}(x)$ is a Lyapunov function for the system (5.1) with an admissible control policy $\mu(x)$.

The optimal control problem can now be formulated: Given the continuous-time system (5.1), the set $u \in \Psi(\Omega)$ of admissible control policies, and the infinite horizon cost functional (5.6), find an admissible control policy such that the cost index (5.6) associated with the system (5.1) is minimized.

Defining the Hamiltonian of the problem

$$H(x,u,\lambda) = L(x(t),u(t)) + \lambda^T(f(x(t)) + g(x(t))u(t)) \tag{5.10}$$

the optimal cost function $V^*(x)$ satisfies the HJB equation

$$0 = \arg\min_{u\in\Psi(\Omega)}\left[H(x,u,\lambda^*)\right] \tag{5.11}$$

This Hamiltonian, which does not depend explicitly on time and corresponds to the infinite horizon optimal control problem, must be identically zero when evaluated on an extremal trajectory. Also every admissible policy, $\mu(x)$, and associated cost, $V^\mu(x)$, with $V^\mu(0) = 0$, satisfy $H(x,\mu,\lambda) = 0$. Thus this equation can be used to solve for the cost associated with any given admissible policy.

Assuming the existence of the unique solution of (5.11), the optimal control function for the given problem is

$$u^*(x) = -\frac{1}{2}R^{-1}g^T(x)\lambda^* \tag{5.12}$$

Substituting this optimal control policy in the Hamiltonian we obtain the formulation of the HJB equation in terms of λ^*:

$$0 = Q(x) + \lambda^{*T}f(x) - \frac{1}{4}\lambda^{*T}G(x)\lambda^* \quad , \quad V^*(0) = 0 \tag{5.13}$$

where $G(x) = g(x)R^{-1}g^T(x)$.

This is a necessary and sufficient condition for the optimal value function. For the linear system case, considering a quadratic cost functional, the equivalent of this HJB equation is the well known Riccati equation. In order to find the optimal control solution for the problem the HJB equation (5.13) is required to be solved for the cost function and then substitute the solution in (5.12) to obtain the optimal control. However, the analytical offline solution of the HJB equation is difficult to obtain in most cases. It also requires the complete knowledge of the system dynamics (i.e. the functions $f(x),\ g(x)$ should be known).

5.3 CONTINUOUS-TIME ADAPTIVE CRITICS

In adaptive critic design (ACD) for affine nonlinear systems, the equations (5.12) and (5.13) are represented by two parametric function approximation networks known as action network (*actor*) and critic network (*critic*) respectively. The action network which provides control signals represents the relationship between state and input. The critic network which learns the desired performance index for some performance index/cost function represents the relationship between state and costate vector. The critic evaluates the performance of actor and the actor is improved based on the feedback from the critic network. These two functional networks approximating HJB equation successively adapt to determine the optimal control solution for a system. In general, ACD uses incremental optimization combined with a

parametric structure to efficiently approximate the optimal cost and control. In ACD a short-term cost metric is optimized that ensures optimization of the cost over all future times.

In general, neural networks are used as the parametric structures in ACD [122, 124, 125, 128-130, 133, 140, 142-144, 147-149, 158-160, 168, 172-174, 207]. Other choices for parametric function approximations in ACD are support vector machines (SVM) [147], particle swarm optimization (PSO) [149],radial basis function neural network [148], and fuzzy systems [178]. Single network adaptive critic (SNAC) using neural networks [172-174], and Takagi-Sugeno fuzzy systems [178] are simplified options for parametric structures in ACD. ACDs function as supervised learning systems and reinforcement learning systems [124, 143]. The policy iteration algorithm [123-126, 155-167], and value iteration algorithm [124, 125, 168] are other options of ACD for solving online optimal control problem.

5.4 POLICY ITERATION TECHNIQUE

The policy iteration technique [123-126, 155-167] is based on an actor-critic structure, consists of two-step iteration- critic update and actor update. For a given stabilizing controller the critic computes the associated infinite horizon cost, and the actor computes the control policy. The online policy iteration algorithm adapts to solve the infinite horizon problem without using knowledge of the system internal dynamics (i.e. function $f(x)$). It gives an adaptive controller which converges to the state feedback optimal controller. Policy iteration algorithm for solving the HJB equation to find the infinite horizon optimal control of nonlinear dynamical systems is presented here. The convergence of the algorithm is discussed briefly in this section and for its detail the lemmas, remarks, and theorems may be referred in [159, 160].

5.4.1 Policy Iteration Algorithm

Let $u^{(0)}(x(t) \in \Psi(\Omega)$ be an admissible control policy such that the closed loop system is asymptotically stable on Ω, and T>0 such that as $x(t) \in \Omega$ also $x(t+T) \in \Omega$ (the existence of such T >0 is guaranteed by the admissibility of $\mu^{(0)}(.)$ on Ω), then the infinite horizon cost function (5.8) can be written as

$$V^{\mu}(x(t)) = \int_{t}^{t+T} L(x(\tau), \mu(x(\tau)))d\tau + V^{\mu}(x(t+T)) \tag{5.14}$$

Based on (5.14) and (5.11), considering an initial admissible control policy $u^{(0)}(x(t)$, the policy iteration algorithm is derived. This algorithm iterates between two steps

1. Policy evaluation: solve for $V^{\mu(i)}(x(t))$ using

$$V^{\mu(i)}(x(t)) = \int_{t}^{t+T} L(x(\tau), \mu^{(i)}(x(\tau)))d\tau + V^{\mu(i)}(x(t+T)) \tag{5.15}$$

with $V^{\mu(i)}(0) = 0$

and

2. Policy improvement: update the control policy using

$$\mu^{(i+1)} = \arg\min_{u \in \Psi(\Omega)} \left[H(x, u, \lambda^{(i)}) \right] \tag{5.16}$$

which explicitly is

$$\mu^{(i+1)}(x) = -\frac{1}{2} R^{-1} g^T(x) \lambda^{(i)} \tag{5.17}$$

The policy iteration algorithm converges to the optimal control policy $\mu^* \in \Psi(\Omega)$ with corresponding cost function $V^*(x_0) = \min_\mu \left(\int_0^\infty L(x(\tau), \mu(x(\tau)))d\tau \right)$.

Equations (5.15) and (5.17) give a formulation for the Policy Iteration algorithm to solve the optimal control problem without any knowledge of system internal dynamics $f(x)$. The knowledge of system's internal state dynamics, i.e. $f(x)$, is not required for evaluation of cost in (5.15) or update of control policy in (5.17), and only the knowledge of input-to-state dynamics, i.e. $g(x)$, is required for the policy update in (5.17). The information regarding system internal dynamics $f(x)$ is embedded in the states $x(t)$ and $x(t+T)$ which are sampled online. This online algorithm is motivated by the success of the online adaptive critic techniques proposed by computational intelligence researchers [127]. In the spirit of reinforcement learning algorithms, the integral term in (5.15) can be addressed as the reinforcement over the time interval $[t, t+T]$.

Policy iteration algorithm can be viewed as an actor-critic system. In this interpretation, the policy evaluation step is viewed as the work of a *critic*, who evaluates the performance of the current policy, $\mu^{(i)}$ i.e. generates an estimate of the value function $V^{\mu(i)}$ from states and reinforcement supplied by the environment as inputs. The policy improvement step is viewed as the work of an *actor*, who takes into account the latest evaluation of the critic, i.e. the estimate of the value function, and acts out the improved policy $\mu^{(i+1)}$.

Equation (5.15) is a discretized version of (5.8) and it can be viewed as a Lyapunov equation for nonlinear systems, may be referred also as

$$\begin{aligned} &LE(V^{\mu(i)}(x(t))) \triangleq \int_t^{t+T} L(x(\tau), \mu^{(i)}(x(\tau)))d\tau + V^{\mu(i)}(x(t+T)) - V^{\mu(i)}(x(t)) \\ &V^{\mu(i)}(0) = 0 \end{aligned} \tag{5.18}$$

5.4.2 Convergence Analysis

The convergence analysis of policy iteration algorithm is discussed in this subsection referring the lemmas, remarks and theorems in [159, 160].

Since $\mu^{(i)} \in \Psi(\Omega)$ then $V^{\mu(i)}(x(t)) \in C^1(\Omega)$, defined as $V^{\mu(i)}(x(t)) = \int_t^\infty L(x(\tau), \mu^{(i)}(x(\tau)))d\tau$, is a Lyapunov function for the system $\dot{x}(t) = f(x(t)) + g(x(t))\mu^{(i)}(x(t))$, then (9) is written in iteration form as

$$\begin{aligned} 0 &= L(x, \mu^{(i)}(x)) + (\lambda^{(i)})^T (f(x) + g(x)\mu^{(i)}(x)) \\ V^{\mu^{(i)}}(0) &= 0 \end{aligned} \tag{5.19}$$

Then, $V^{\mu^{(i)}} \in C^1(\Omega)$ satisfies

$$(\lambda^{(i)})^T (f(x) + g(x)\mu^{(i)}(x)) = -L(x(t), \mu^{(i)}(x(t))) \tag{5.20}$$

with $L(x(t), \mu^{(i)}(x(t)))$; $x(t) \neq 0$. Integrating (5.20) over the time interval $[t, t+T]$, we obtain (5.15). Let us assume that there exists another cost function $V \in C^1(\Omega)$ which satisfies (5.23) with the end condition $V(0) = 0$. This cost function also satisfies $\dot{V}(x(t)) = -L(x(t), \mu^{(i)}(x(t)))$. Thus, subtracting this from (5.15) we obtain

$$\begin{aligned} \left(\frac{d[V(x(t)) - V^{\mu^{(i)}}(x(t))]^T}{dx} \right) \dot{x} = \left(\frac{d[V(x(t)) - V^{\mu^{(i)}}(x(t))]^T}{dx} \right) \\ \times \left(f(x(t)) + g(x(t))\mu^{(i)}(x(t)) \right) = 0 \end{aligned} \tag{5.21}$$

which must hold for any *x* on the system trajectories generated by the stabilizing policy $\mu^{(i)}$. Thus $V(x(t)) = V^{\mu^{(i)}}(x(t)) + c$. As this relation must hold also for $x(t) = 0$ then $V(0) = V^{\mu^{(i)}}(0) + c \Rightarrow 0 = c$ and thus $V(x(t)) = V^{\mu^{(i)}}(x(t))$ i.e. Equation (5.15) has a unique solution which is equal to the unique solution of (5.19). Thus, if $\mu^{(i)} \in \Psi(\Omega)$ and $V^{\mu(i)}(x(t)) \in C^1(\Omega)$ satisfy (5.15), then the new control policy $\mu^{(i+1)}$, determined based on (5.16), is admissible for the system (5.1) [159, 160]. Even though the same solution is obtained from either (5.19) or (5.15), solving (5.19) requires $f(x)$, but solving (5.15) does not require any knowledge of the system internal dynamics $f(x)$.Thus, it shows that the algorithm (5.15) and (5.17) is equivalent to iterating between (5.19) and (5.17), without using knowledge of the system internal dynamics $f(x)$. As iterating on (5.19) and (5.17), conditioned by an initial admissible policy $\mu^{(0)}(x)$, all the subsequent control policies will be admissible and the iteration (5.19) and (5.17) will converge to the solution of the HJB equation, the equivalent of this the policy iteration (5.15) and (5.17) converges uniformly to the optimal control solution on the trajectories originating in Ω, i.e. $\forall \varepsilon > 0 \; \exists i_0 : \forall i \geq i_0$

$$\sup_{x \in \Omega} \left| V^{\mu(i)}(x) - V^*(x) \right| < \varepsilon \;, \quad \sup_{x \in \Omega} \left| \mu^{(i)}(x) - \mu^*(x) \right| < \varepsilon \tag{5.22}$$

Thus, the equivalence between (5.15) and (5.19) concludes that this online adaptive optimal control algorithm will converge to the solution of the optimal control problem (5.6), on Ω, without using knowledge of the internal dynamics of the controlled system (5.1). The positive definite solution of (5.15) is guaranteed when the system has stabilizable dynamics and when the performance functional satisfies zero state observability (i.e. observability of the system state through the cost function) [159, 160].

5.4.3 Neural Network Approximation of Cost Function for Online Implementation of PI Algorithm

For implementation of PI algorithm given by (5.15) and (5.17), which does not need the knowledge of internal state dynamics, $f(x)$, and only need the knowledge of input-to-state dynamics, $g(x)$, the real-time dynamics of continuous-time system is used by measuring online sampled states $x(t)$ and $x(t+T)$ which are involved in policy evaluation (5.15). Thus, for online implementation, the cost function approximation using parametric function approximation network is used to determine the optimal control solution for a system.

Due to their universal approximation property the neural networks are natural choice to approximate smooth functions on compact sets. Thus, for solving (5.15) for any $x \in \Omega$, the cost function approximation using a neural network is considered here. Thus, consider that the cost function $V^{\mu(i)}(x)$ can be approximated, for $x \in \Omega$, by

$$V_h^{\mu(i)}(x) = \sum_{j=1}^{h} w_j^{\mu(i)} \phi_j(x) = (\mathbf{w}_h^{\mu(i)})^T \varphi_h(x) \tag{5.23}$$

This can be seen as a neural network where *h* is the number of neurons on the hidden layer, $\varphi_h(x)$ is the vector of activation functions with neuron activation functions $\phi_j(x) \in C^1(\Omega)$, $\phi_j(0) \in 0$. $w_j^{\mu(i)}$ denote the weights of the output layer and $\mathbf{w}_h^{\mu(i)}$ is the weight vector. The output layer neuron has a linear activation function. The weights of the hidden layer are all equal to one and will not be changed during the training procedure [159, 160].

Consider an infinite set of linearly independent activation functions $\{\phi_j(x)\}_1^{\infty}$, such that $\phi_j(x) \in C^1(\Omega)$, $\phi_j(0) \in 0$, $j = \overline{1,\infty}$, which satisfy the completeness property (i.e. any function $f(x) \in C^1(\Omega)$, $f(0) \in 0$ can be represented as a linear combination of a subset of $\{\phi_j(x)\}_1^{\infty}$, then the exact solution of (5.15) can be expressed as

$$V^{\mu(i)}(x) = \sum_{j=1}^{\infty} c_j^{\mu(i)} \phi_j(x) = (\mathbf{c}_{\infty}^{\mu(i)})^T \varphi_{\infty}(x) \tag{5.24}$$

where $\varphi_\infty(x)$ is the vector of activation functions and $\mathbf{c}_\infty^{\mu^{(i)}}$ denotes the weight vector.

Using the neural network approximate description for the cost function, (5.23) and (5.24) can be written as

$$(\mathbf{w}_h^{\mu^{(i)}})^T \varphi_h(x(t)) = \int_t^{t+T} L(x, \mu^{(i)}(x))d\tau + (\mathbf{w}_h^{\mu^{(i)}})^T \varphi_h(x(t+T)) \tag{5.25}$$

Thus using this neural network approximation for the cost function, (5.26) will have the residual error

$$\delta_h^{\mu^{(i)}}(x(t),T) = \int_t^{t+T} L(x, \mu^{(i)}(x))d\tau + (\mathbf{w}_h^{\mu^{(i)}})^T [\varphi_h(x(t+T)) - \varphi_h(x(t))] \tag{5.26}$$

From the perspective of temporal difference learning methods this error can be viewed as a temporal difference residual error for continuous-time systems [159, 160]. The method of weighted residuals in the least-squares sense is used to determine the parameters of the neural network approximating the cost function $V_h^{\mu^{(i)}}$. Thus, in order to tune the parameters $\mathbf{w}_h^{\mu^{(i)}}$ of the cost function approximation $V_h^{\mu^{(i)}}$, minimize the objective function

$$J = \int_\Omega (\delta_h^{\mu^{(i)}}(x,T))^T \delta_h^{\mu^{(i)}}(x,T)dx \tag{5.27}$$

From (5.27) we have

$$\int_\Omega \frac{d\delta_h^{\mu^{(i)}}(x,T)}{d\mathbf{w}_h^{\mu^{(i)}}} \delta_h^{\mu^{(i)}}(x,T)dx = 0 \tag{5.28}$$

Using the inner product notation for the Lebesgue integral, (5.28) can be written as

$$\left\langle \frac{d\delta_h^{\mu^{(i)}}(x,T)}{d\mathbf{w}_h^{\mu^{(i)}}}, \delta_h^i(x,T) \right\rangle_\Omega = 0 \tag{5.29}$$

which is

$$\begin{aligned} &\left\langle [\varphi_h(x(t+T)) - \varphi_h(x(t))], [\varphi_h(x(t+T)) - \varphi_h(x(t))]^T \right\rangle_\Omega \mathbf{w}_h^{\mu^{(i)}} \\ &+ \left\langle [\varphi_h(x(t+T)) - \varphi_h(x(t))], \int_t^{t+T} L(x(\tau), \mu^{(i)}(x(\tau)))d\tau \right\rangle_\Omega \end{aligned} \tag{5.30}$$

with $\Phi = \left\langle [\varphi_h(x(t+T)) - \varphi_h(x(t))], [\varphi_h(x(t+T)) - \varphi_h(x(t))]^T \right\rangle_\Omega$ being invertible, we obtain

$$\mathbf{w}_h^{\mu^{(i)}} = -\Phi^{-1} \left\langle [\varphi_h(x(t+T)) - \varphi_h(x(t))], \int_t^{t+T} L(x(\tau), \mu^{(i)}(x(\tau)))d\tau \right\rangle_\Omega \tag{5.31}$$

For the proof of convergence of solution with cost function approximation, the theorems, lemmas, corollaries, facts and definitions may be referred in [159, 160].

Thus, using the tuned parameters $\mathbf{w}_h^{\mu^{(i)}}$ the cost function $V_h^{\mu^{(i)}}$ associated with the control policy $\mu^{(i)}$, can be calculated, and the control policy can be updated. The updated control policy will thus be

$$\mu^{(i+1)}(x) = -\frac{1}{2} R^{-1} g^T(x) \nabla \varphi_h^T(x) \mathbf{w}_h^{\mu^{(i)}} \tag{5.32}$$

Equation (5.32) gives the output of the *actor* neural network. The controller (*actor*) can be seen as a neural network, similar to the one approximating the *critic*, which has the same weight parameters as the *critic* neural network, but whose activation functions are the gradients of those in the *critic*.

In this implementation the activation function vector $\varphi_h(x)$ is considered to be the Kronecker product quadratic polynomial basis vector with the elements $\{x_i(t).x_j(t)\}_{i=1,n;\, j=1,n}$, where n is the number of system state variables. For application of nonlinear inverted pendulum system the activation function vector contains 10 quadratic elements, for single-link manipulator 3 quadratic elements, and also for Vander Pol's oscillator 3 quadratic elements are considered.

5.5 ADAPTIVE OPTIMAL CONTROL USING POLICY ITERATION TECHNIQUE

The implementation issues of online policy iteration technique based adaptive critic scheme for adaptive optimal control of continuous-time input-affine nonlinear systems is described in this section. The adaptive optimal control scheme with actor-critic structure for affine nonlinear systems is shown in Fig. 5.1 [159, 160]. For implementation of policy iteration algorithm given by (5.15) and (5.17), the real-time dynamics of continuous-time system is used by measurements of states $x(t)$ and $x(t+T)$ sampled online along state trajectories which determine the information of system internal state dynamics, $f(x)$, the knowledge that is not needed by PI algorithm, and only needed the knowledge of input-to-state dynamics, $g(x)$. In this control system structure the online adaptation scheme uses the dynamic information $\dot{V} = Q(x) + u^T R u$ of cost function (state) $V(t)$ for a control policy $\mu(x)$. In this implementation $V(t)$ is reset to zero at the beginning of each sample interval $[t, t+T]$, then the measurement $V(t+T)$ gives the reinforcement over time interval $[t, t+T]$ required to implement the policy evaluation step in (5.15), i.e. $V(t+T)$ gives the integral reinforcement term in (5.15). Thus, this is a dynamical controller whose memory is exactly the value $V(t)$ of using the current policy. Taking measurements at specific time values about the system states *x*, and the augmented system state *V*, (i.e. $x(t)$, $x(t+T)$ and $V(t+T) - V(t)$), the critic is able to evaluate the infinite horizon continuous-time performance of the system associated with a given control policy described in terms of the actor

parameters. The critic learns the cost function associated with a certain control behavior based on a computed temporal difference (TD) error signal, given by $V(t+T)-V(t)$. After observing the system state $x(t_{i+1})$ at time t_{i+1}, the control policy is updated to $\mu_h^{(i+1)}(x)$ (i.e. a new control policy at time t_{i+1}), which is used for controlling the system during the time interval $[t_{i+1}, t_{i+1}+T]$. This makes the PI algorithm suitable for online implementation. Since online PI algorithm does not require the system internal state dynamics $f(x)$ for the evaluation of the cost function or the update of the control policy, and only requires the input-to-state dynamics $g(x)$ for the update of the control policy, thus the control scheme based on PI technique becomes partially model free. This control system structure involves the continuous-time dynamics of the system, and the discrete-time sampled data for the policy evaluation and policy update steps at discrete moments in time. Thus, this is a hybrid continuous-time / discrete-time adaptive control structure. This control scheme gives the optimal control solution and performs online adaptation, is termed as optimal adaptive control. This gives a direct adaptive optimal control scheme. The adaptive optimal control using policy iteration method relies on identification of the cost function associated with a given control policy followed by policy improvement in the sense of minimizing the identified cost, whereas the regular adaptive controllers rely on online identification of the system dynamics followed by model based controller design. This technique involves computations at a supervisory level based on discrete-time data measured from the system to update parameters of both actor and critic. By using neural networks to parameterize actor and critic for online implementation, this control scheme becomes a high-level intelligent control scheme.

Using neural network approximation the solution of cost function, given by (5.25), can be obtained in real time, after a sufficient number of data points are collected along state trajectories in the region of interest Ω, and applying recursive least squares (RLS) algorithm which requires a persistence of excitation condition. Enough excitation must be present in the system for successful application of the algorithm. As PI algorithm iterates only on stabilizing controllers, if system state reached the equilibrium point, the data measured from the system can no longer be used in the adaptive algorithm; in that case the system must be again excited to the previously considered initial state and a new experiment needs to be conducted having as starting point the last policy obtained in the previous experiment. The convergence of iterations is determined by a threshold specified by control designer. The error between the system performances evaluated at two consecutive steps to be below threshold is considered as the criteria to stop the iterations [159, 160].

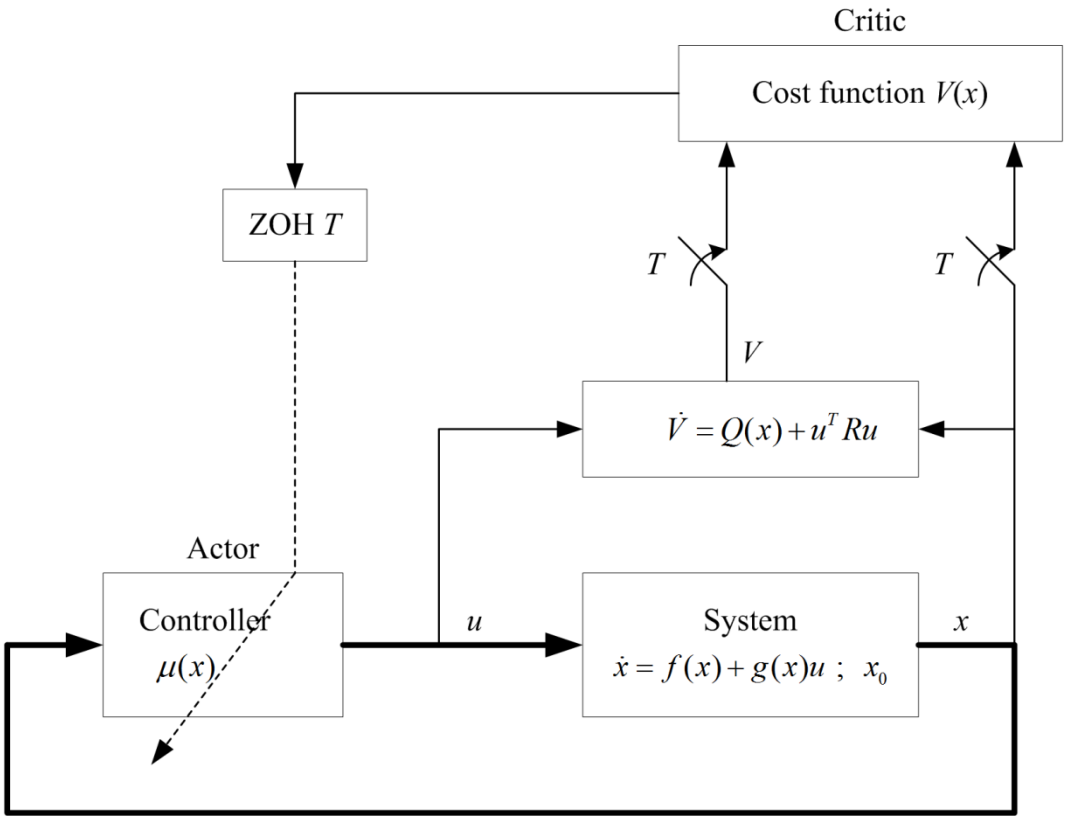


Fig. 5.1 Adaptive optimal control with actor-critic structure.

5.6 SIMULATION RESULTS AND ANALYSIS

In this section, the application of ACD using online PI technique with neural network approximation of cost function for adaptive optimal control is implemented for five examples of affine nonlinear systems- two general systems, and three practical systems- single-link manipulator, inverted pendulum system, and Vander Pol oscillator. The simulation results and performance analysis are presented to demonstrate the effectiveness of the control scheme.

5.6.1 General Affine Nonlinear System Example 1

Consider a general input-affine nonlinear system which is having stronger nonlinearities and quadratic cost function. This nonlinear system is given by the equations

$$\dot{x}_1 = -x_1 + x_2 \tag{5.33}$$

$$\dot{x}_2 = f(x) + g(x)u \tag{5.34}$$

where $f(x) = -\frac{1}{2}(x_1 + x_2) + \frac{1}{2}x_2(2 + \sin(2x_1))^2$, and $g(x) = 2 + \sin(2x_1)$.

For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken here also as $x_0 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 0 & 3 & 3/2 \end{bmatrix}$. The length of the simulation in samples is taken 150, number of operations 30, and sample time T=0.1 seconds.

The adaptive optimal critic matrix *P* of adaptive critic scheme using PI which gives the quadratic cost critic parameters at 5th iteration is obtained as

$$P = \begin{bmatrix} 0.5 & 0 & 1 \end{bmatrix}.$$

Thus the control policy is obtained as

$$u = -\frac{1}{2} R^{-1}(2 + \sin(2x_1)) P \begin{bmatrix} 0 \\ x_1 \\ 2x_2 \end{bmatrix} = -x_2(2 + \sin(2x_1))$$

The simulation responses of adaptive optimal control using PI technique for affine nonlinear system (5.33) & (5.34) are shown in Figs. 5.2 to 5.4. Fig. 5.2 shows the system state trajectories which converge towards the equilibrium point at origin. Fig. 5.3 shows the control signal trajectory which also converge towards zero. Fig. 5.4 shows the convergence of critic parameters of matrix P towards optimal values.

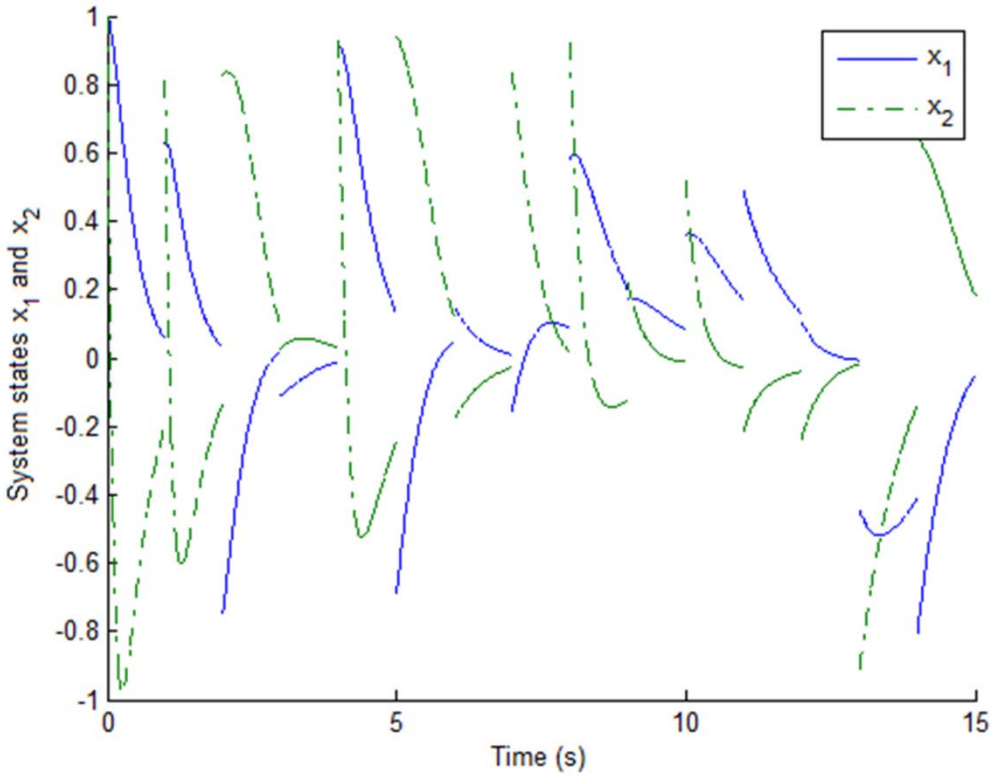


Fig. 5.2 System states

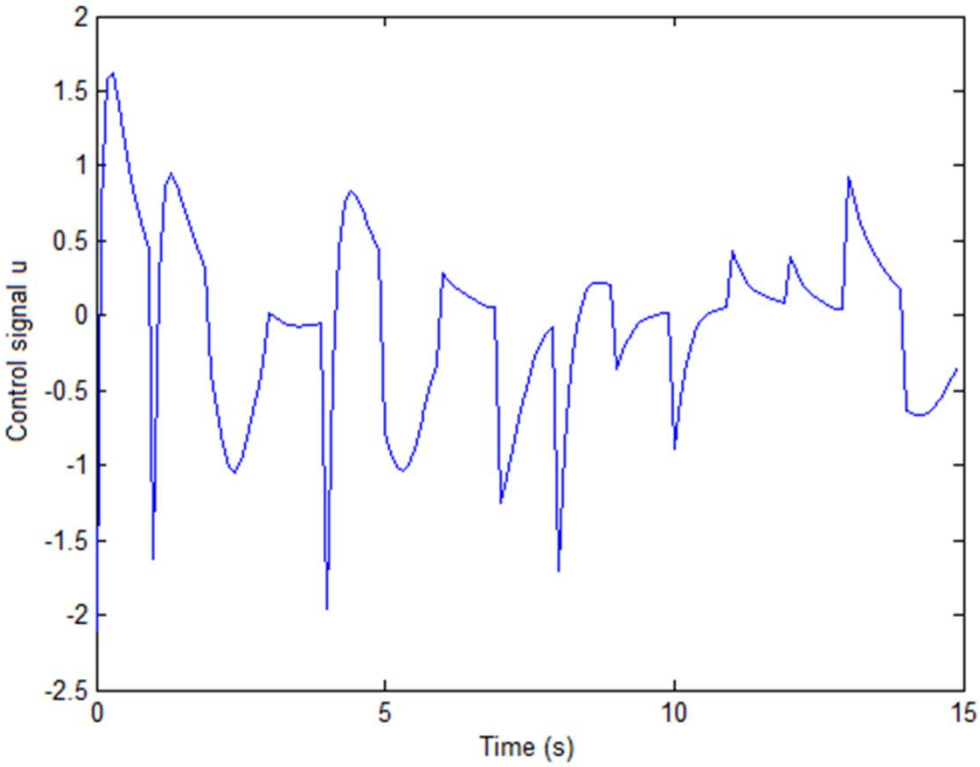


Fig. 5.3 Control signal

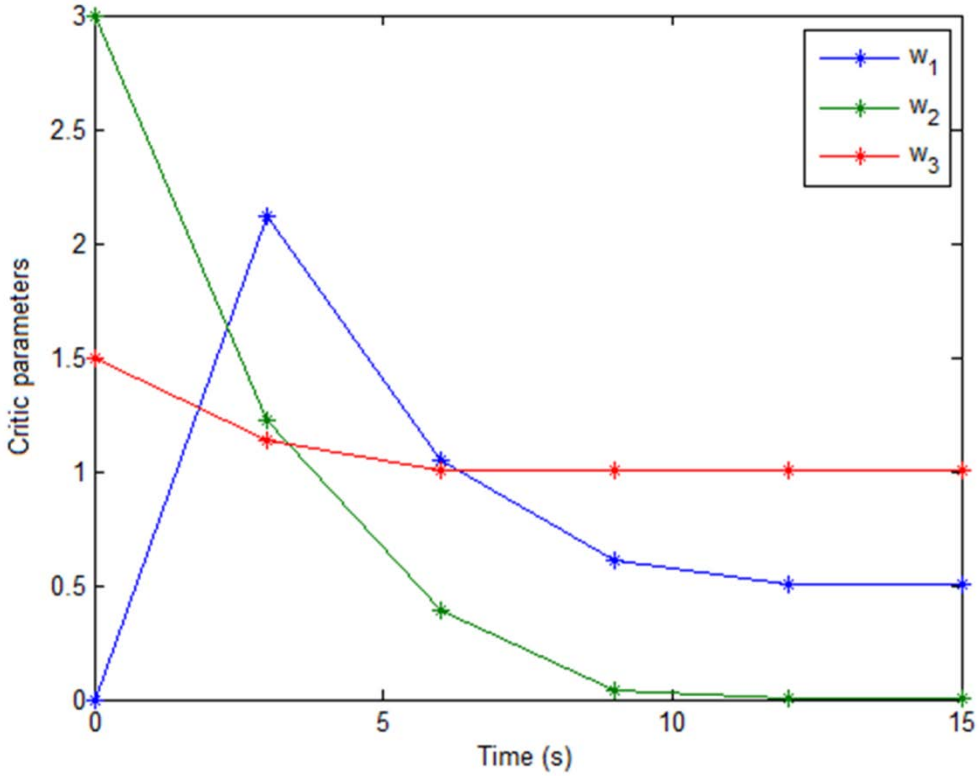


Fig. 5.4 Critic parameters

5.6.2 General Affine Nonlinear System Example 2

Consider a general input-affine nonlinear system which is having stronger nonlinearities and quadratic cost function. This nonlinear system is described by the dynamic equations

$$\dot{x}_1 = -x_1 + x_2 \tag{5.35}$$

$$\dot{x}_2 = f(x) + g(x)u \tag{5.36}$$

where $f(x) = -\frac{1}{2}(x_1 + x_2) + \frac{1}{2}x_2(2 + \cos(2x_1))^2$, and $g(x) = 2 + \cos(2x_1)$.

Considering the state regulation problem for this system defined by (5.35) & (5.36) with quadratic cost function. For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 0 & 3 & 3/2 \end{bmatrix}$. The length of the simulation in samples is taken 150, number of operations 30, and sample time *T*=0.1 seconds. Using neural network approximation of cost function the adjusted optimal weights $\mathbf{w}_L^{\mu^{(i)}}$ at 5th iteration during simulation give the adaptive optimal critic parameters *P*, which is obtained as

$$P = \begin{bmatrix} 0.5 & 0 & 1 \end{bmatrix}$$

Thus the control policy is obtained as

$$u = -\frac{1}{2}R^{-1}(2 + \cos(2x_1))P\begin{bmatrix} 0 \\ x_1 \\ 2x_2 \end{bmatrix} = -x_2(2 + \cos(2x_1))$$

The simulation responses of adaptive optimal control using PI technique for affine nonlinear system defined by (5.35) & (5.36) are shown in Figs. 5.5 to 5.7. Fig. 5.5 shows the system state trajectories showing x_1 and x_2 which converge towards the equilibrium point at origin. Fig. 5.6 shows the control signal trajectory *u* which also converge towards zero. Fig. 5.7 shows the convergence of critic parameters of matrix *P* towards optimal values.

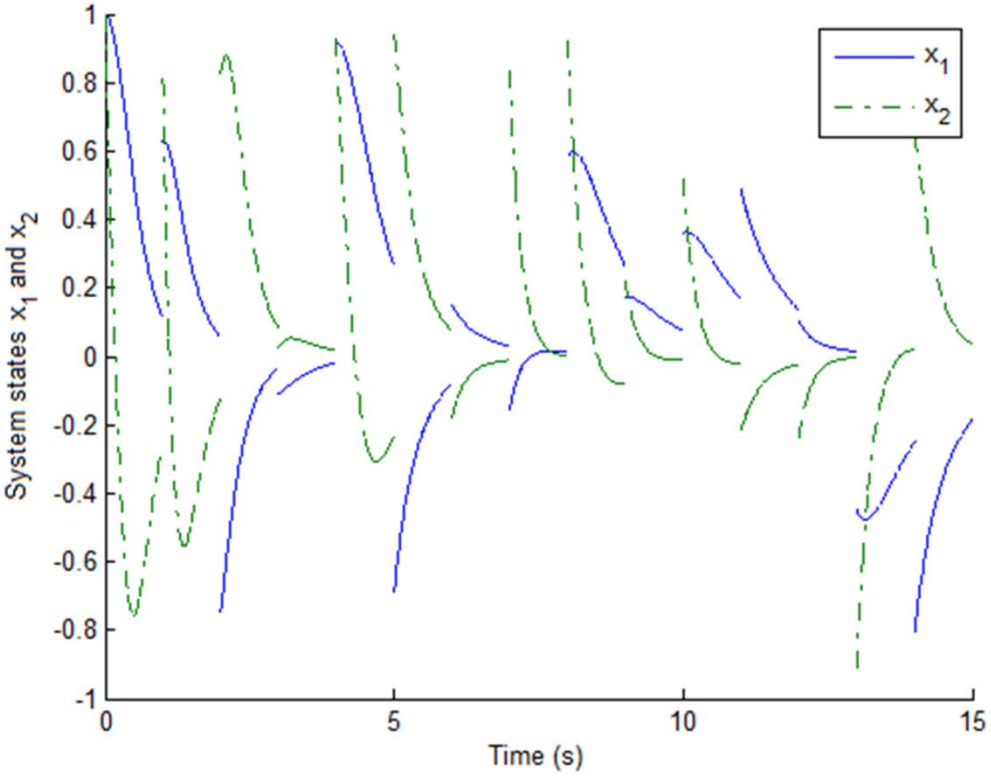


Fig. 5.5 Nonlinear system states trajectories

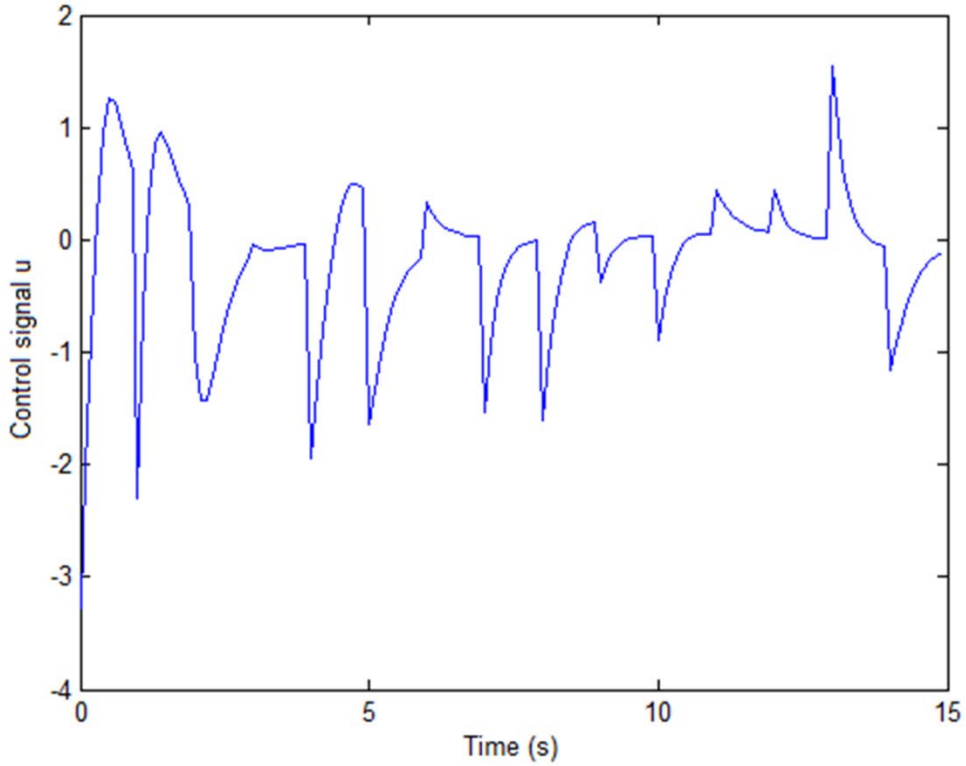


Fig. 5.6 Control signal

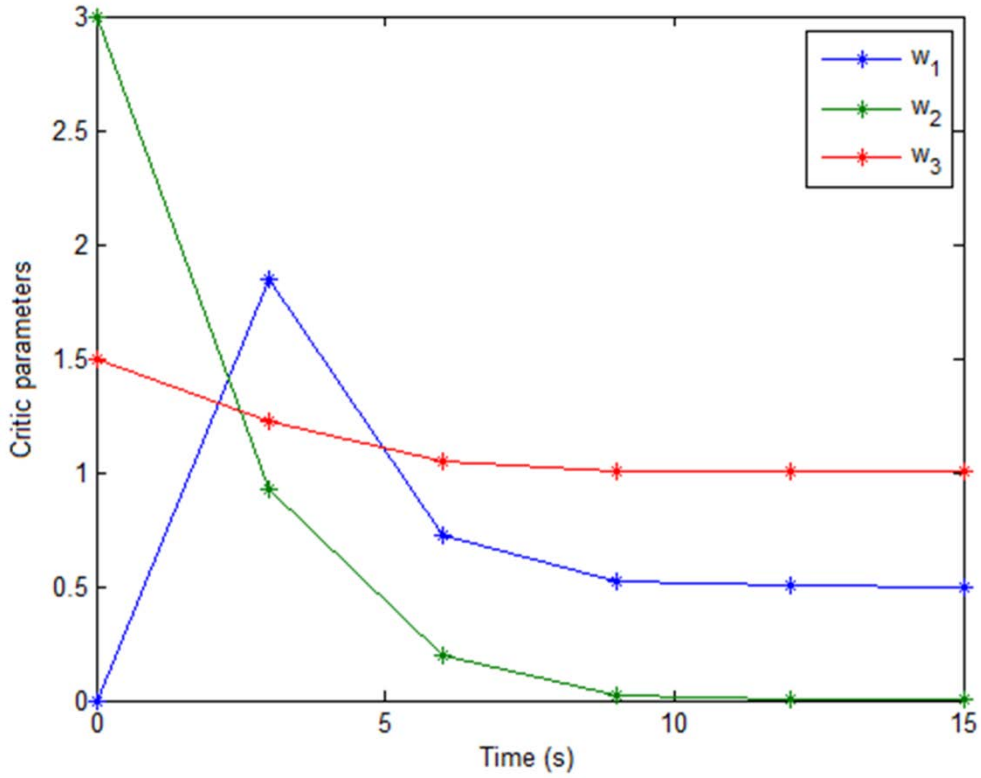


Fig. 5.7 Critic parameters

5.6.3 Single-Link Manipulator System

Consider the single-link manipulator which is also a benchmark nonlinear system problem. The dynamic system equations of a single-link manipulator [71, 178] are given as

$$\frac{d\theta}{dt} = \omega \tag{5.37}$$

$$\frac{d\omega}{dt} = -\frac{g}{l}\sin\theta + \frac{1}{ml^2}T_a \tag{5.38}$$

where θ is angular displacement, ω is angular velocity, and T_a is the actuating torque. Let acceleration due to gravity g=9.81m/s^2; link length l=1m, and link mass m=1kg. Considering $\theta = x_1$, and $\omega = x_2$ as the state variables , and $T_a = u$ is the control input, then dynamic system equations are written as

$$\dot{x}_1 = x_2 \tag{5.39}$$

$$\dot{x}_2 = -g\sin(x_1) + u \tag{5.40}$$

For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = \begin{bmatrix}1 & 1 & 0\end{bmatrix}$; $P = \begin{bmatrix}1 & 0 & 1/2\end{bmatrix}$. The quadratic cost function matrices are taken as $Q = [1\ \ 0; 0\ \ 1]$, $R = 1$. The length of the simulation in samples is taken 150, number of operations 30, and sample time T=0.1 seconds. At each iteration

data is measured along state trajectory defined from different initial conditions chosen randomly in the region of interest Ω.

Using neural networks approximation of cost function the adjusted optimal weights $\mathbf{w}_L^{\mu^{(i)}}$ at 5th iteration during simulation give the adaptive optimal critic parameters *P*, which is obtained as

$$P = \begin{bmatrix} 10.3212 & 0.0988 & 1.0487 \end{bmatrix}$$

Thus the control policy is obtained as

$$u = -\frac{1}{2} R^{-1} P \begin{bmatrix} 0 \\ x_1 \\ 2x_2 \end{bmatrix} = -0.0494 x_1 - 1.0487 x_2$$

The simulation responses of ACD using PI technique for single-link manipulator (5.39) & (5.40) are shown in Figs. 5.8 to 5.10. Fig. 5.8 shows the system state trajectories which converge towards the equilibrium point at origin. Fig. 5.9 shows the control signal trajectory which also converge towards zero. Fig. 5.10 shows the convergence of critic parameters of matrix *P* towards optimal values.

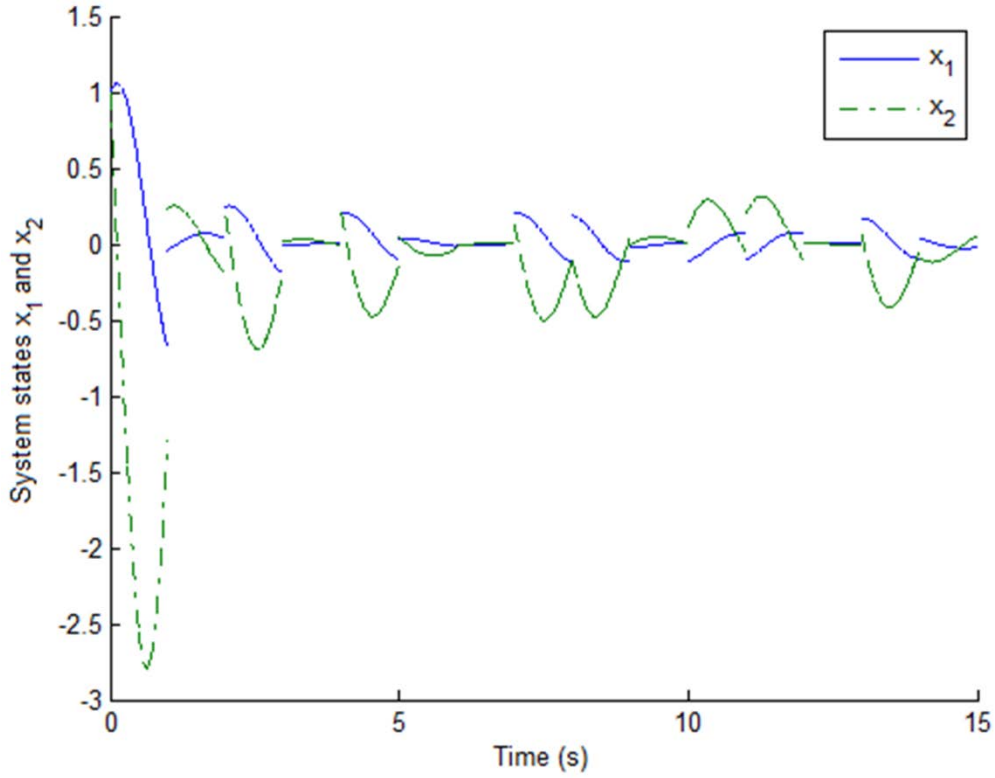


Fig. 5.8 Single-link manipulator system states trajectories- angular displacement x_1 and angular velocity x_2.

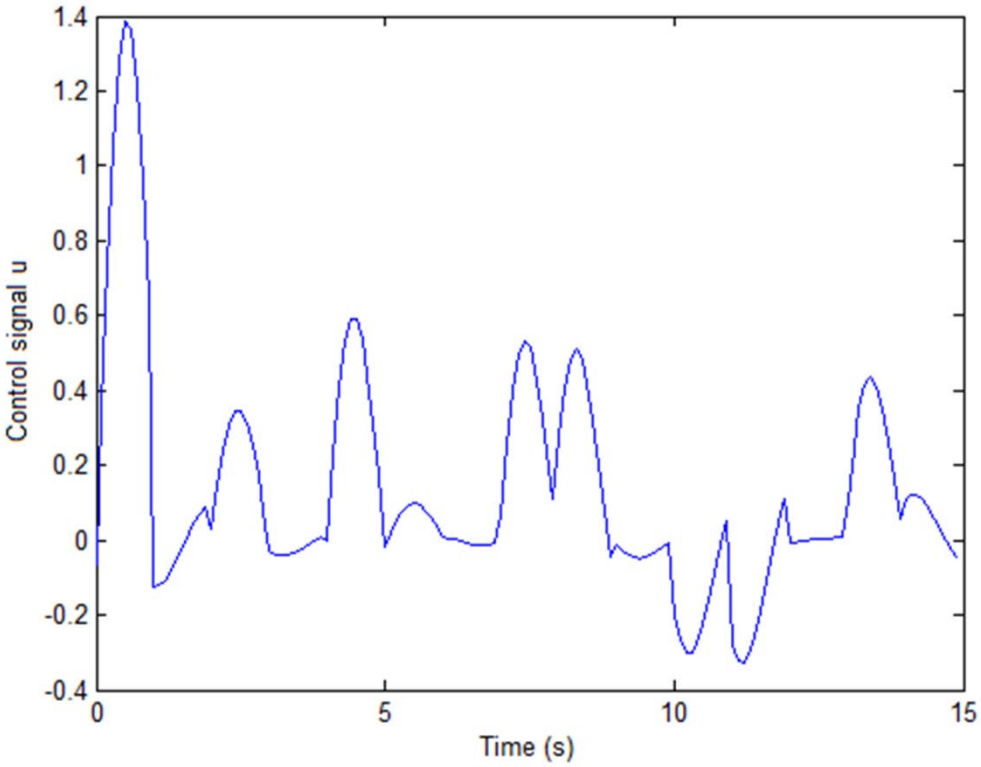


Fig. 5.9 Control signal u with adaptive optimal control using PI technique.

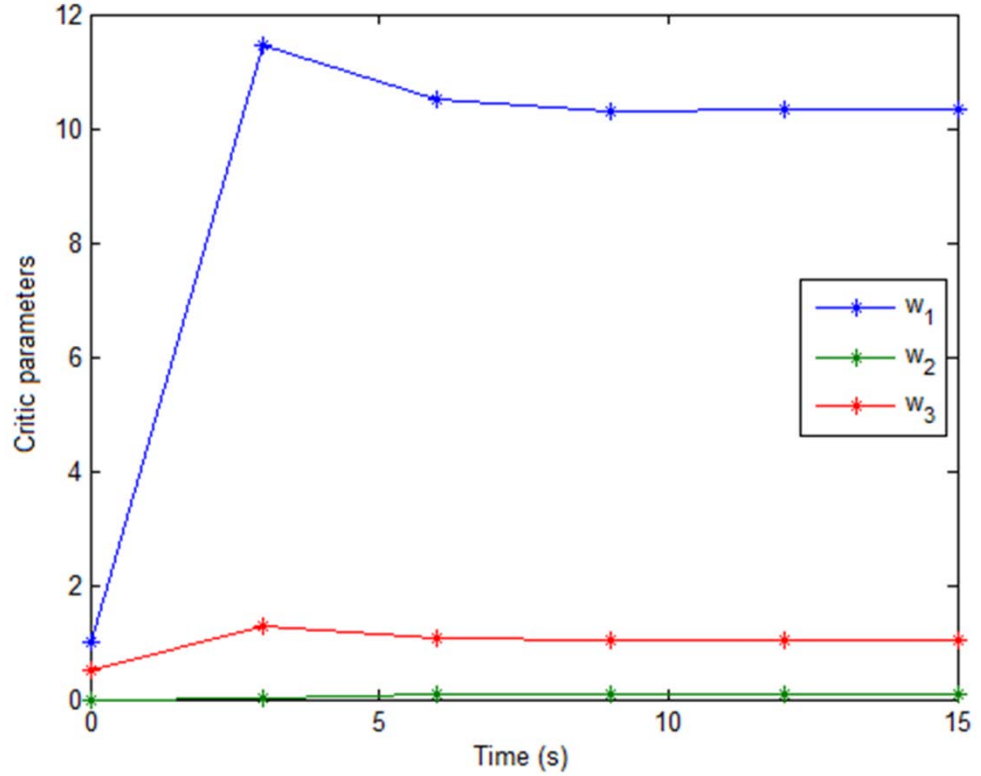


Fig. 5.10 Critic parameters.

5.6.4 Nonlinear Inverted Pendulum System

Consider the nonlinear inverted pendulum-cart dynamical system which represents the control input affine system. The dynamic equations of inverted pendulum-cart system are given as [1, 2, 180]

$$\ddot{x} = \frac{u + ml(\sin\theta)\dot{\theta}^2 - mg\cos\theta\sin\theta}{M + m - m\cos^2\theta} \tag{5.41}$$

$$\ddot{\theta} = \frac{u\cos\theta - (M+m)g\sin\theta + ml(\cos\theta\sin\theta)\dot{\theta}}{ml\cos^2\theta - (M+m)l} \tag{5.42}$$

where $x(t)$ is the cart position and $\theta(t)$ is the pendulum tilt angle referenced to the vertically upward direction, *M* is mass of cart, *m* is the pendulum ball point mass, *u* is an externally *x*-directed force on the cart, and *g* is the acceleration due to gravity.

Considering $x_1 = \theta$, $x_2 = \dot{\theta} = \dot{x}_1$, $x_3 = x$, $x_4 = \dot{x} = \dot{x}_3$ as state variables, (5.41) and (5.42) may be written in state space form

$$\frac{d}{dt}\boldsymbol{x} = \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{d}{dt}\begin{bmatrix} \theta \\ \dot{\theta} \\ x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{bmatrix} u \tag{5.43}$$

where,

$$f_1 = x_2 \tag{5.44}$$

$$f_2 = \frac{(M+m)g\sin x_1 + ml(\cos x_1 \sin x_1)x_2^2}{ml\cos^2 x_1 - (M+m)l} \tag{5.45}$$

$$f_3 = x_4 \tag{5.46}$$

$$f_4 = \frac{ml(\sin x_1)x_2^2 - mg\cos x_1 \sin x_1}{M + m - m\cos^2 x_1} \tag{5.47}$$

$$g_1 = 0 \tag{5.48}$$

$$g_2 = \frac{\cos x_1}{ml\cos^2 x_1 - (M+m)l} \tag{5.49}$$

$$g_3 = 0 \tag{5.50}$$

$$g_4 = \frac{1}{M + m - m\cos^2 x_1} \tag{5.51}$$

For the simulation the typical parameters of inverted pendulum-cart system setup are considered as [180]: mass of the cart (*M*): 2.4 kg, mass of the pendulum (*m*): 0.23 kg, length of the pendulum (*l*): 0.36 m, length of the cart track (*L*): ± 0.5 m, friction coefficient of the cart & pole rotation is assumed negligible. The acceleration due to gravity g=9.81 m/s^2.

For implementation of PI algorithm the initial conditions for system states and cost function, and the initial critic parameters are considered as $\boldsymbol{x}_0 = \begin{bmatrix} -\pi & \pi & 0.4 & 0.8 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$. The quadratic cost function matrices are taken as $Q = [1\ 0\ 0\ 0;\ 0\ 1\ 0\ 0;\ 0\ 0\ 1\ 0;\ 0\ 0\ 0\ 1]$, $R = 1$. The length of the simulation in samples is taken 1200, number of operations 20, and sample time *T*=0.05 seconds. At each iteration data is measured along state trajectory defined from different initial conditions chosen randomly in the region of interest Ω.

Using neural networks approximation of cost function the adjusted optimal weights $\mathbf{w}_L^{\mu(i)}$ at 60th iteration during simulation give the adaptive optimal critic parameters *P*, which is obtained as

$$P = [26.7512 \quad 0.1662 \quad 2.1849 \quad 5.2627 \quad -0.8502 \quad 0.2337 \quad 0.5275 \quad 2.4927 \quad 5.3840 \quad 6.8071]$$

Thus the control policy is obtained as

$$u = -\frac{1}{2} R^{-1} \begin{bmatrix} 0 \\ \dfrac{\cos x_1}{0.0828\cos^2 x_1 - 0.9468} \\ 0 \\ \dfrac{1}{2.63 - 0.23\cos^2 x_1} \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 & 0 & 0 \\ x_2 & x_1 & 0 & 0 \\ x_3 & 0 & x_1 & 0 \\ x_4 & 0 & 0 & x_1 \\ 0 & 2x_2 & 0 & 0 \\ 0 & x_3 & x_2 & 0 \\ 0 & x_4 & 0 & x_2 \\ 0 & 0 & 2x_3 & 0 \\ 0 & 0 & x_4 & x_3 \\ 0 & 0 & 0 & 2x_4 \end{bmatrix}^T P^T$$

The simulation responses are shown in Figs. 5.11 to 5.13. Fig. 5.11 shows the system state trajectories showing θ, $\dot{\theta}$, x, and $\dot{x}$ which converge towards the equilibrium point at origin. Fig. 5.12 shows the control signal trajectory *u* which also converge towards zero. Fig. 5.13 shows the convergence of critic parameters of matrix *P* towards optimal values. It is observed here that the states of pendulum-cart system are regulated quickly with bounded control input *u*. The inverted pendulum stabilizes in vertically upright position quickly.

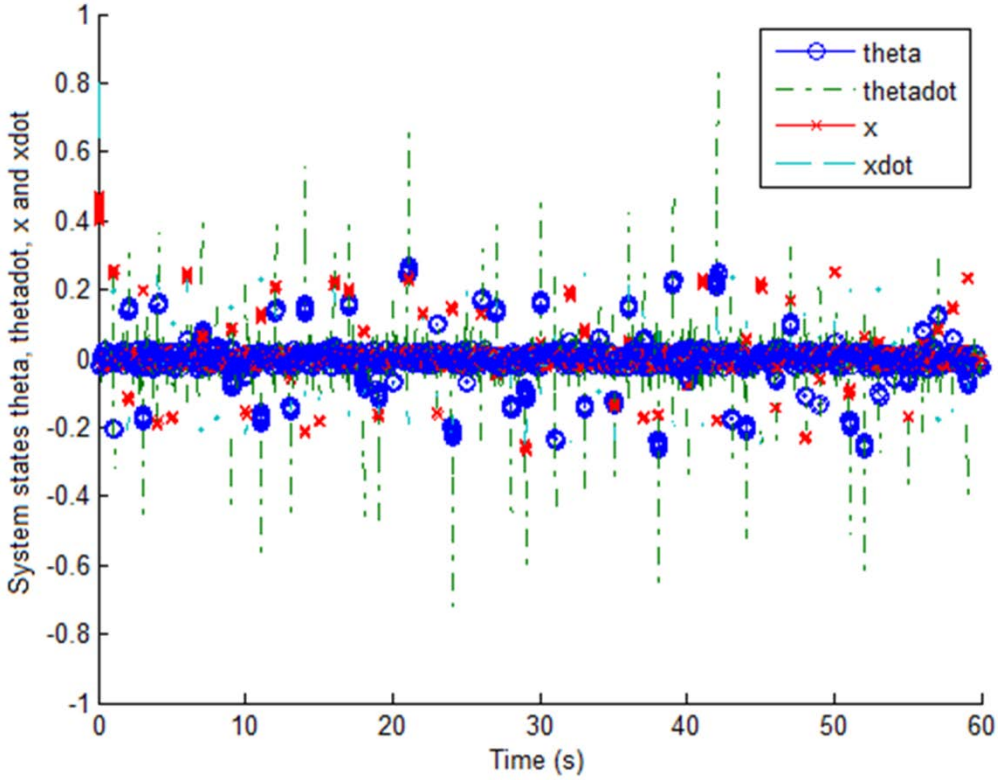


Fig. 5.11 Nonlinear inverted pendulum system states trajectories- pendulum angle θ, angular velocity $\dot{\theta}$, cart position x, cart velocity $\dot{x}$.

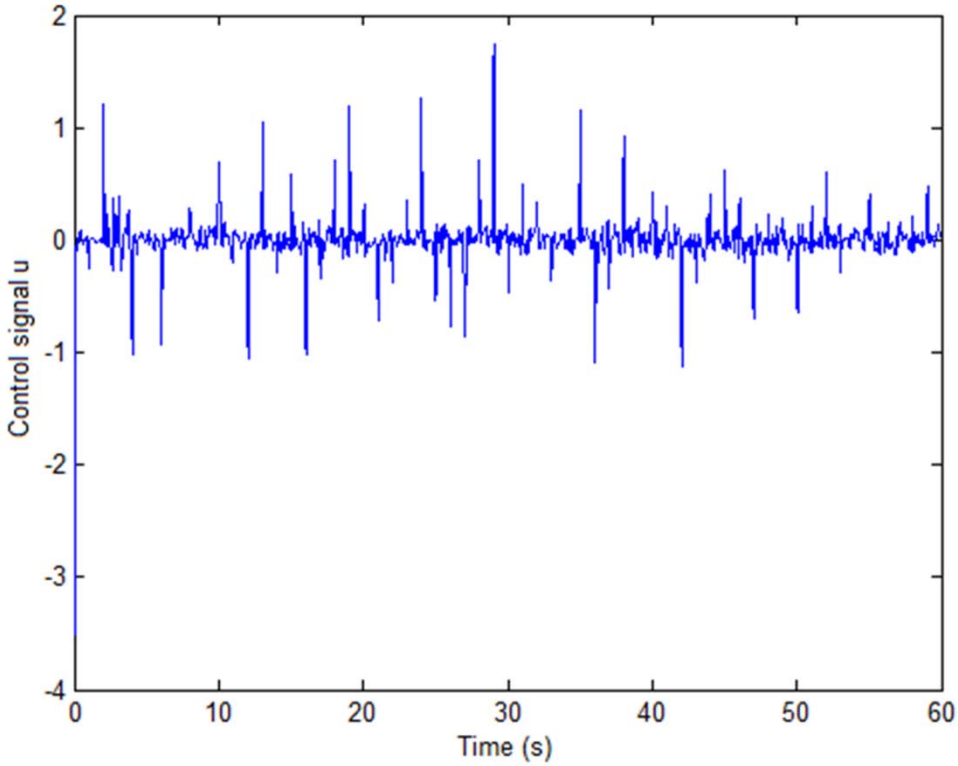


Fig. 5.12 Control force *u* with adaptive optimal control using PI technique.

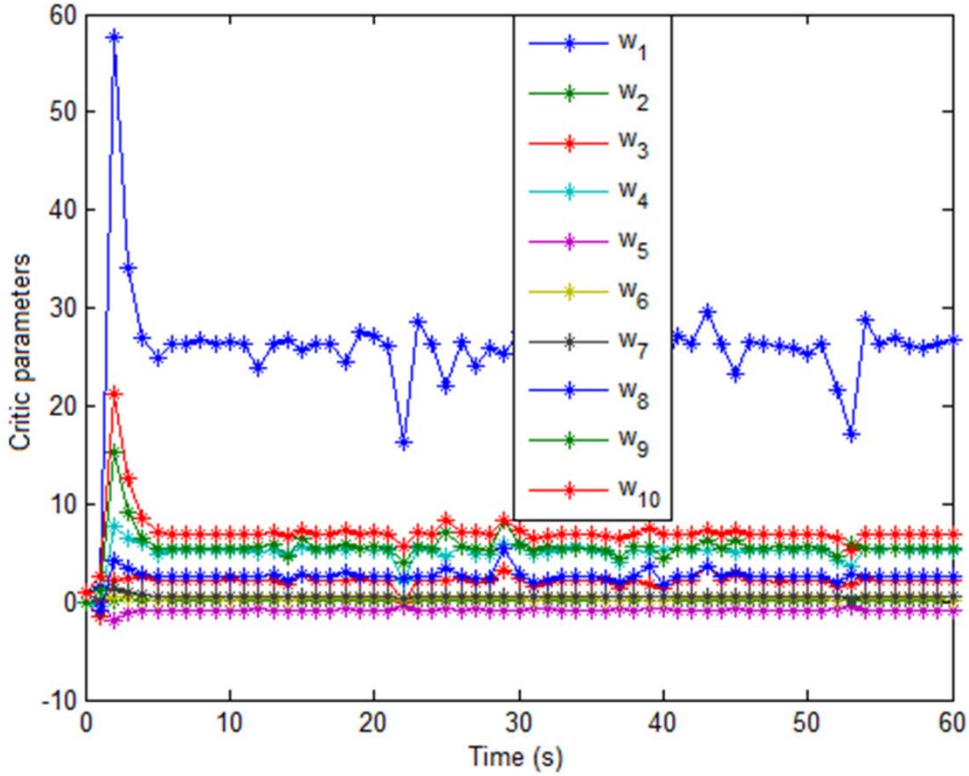


Fig. 5.13 Critic parameters.

5.6.5 Nonlinear Vander Pol Oscillator System

Consider the control design for Vander Pol oscillator, a benchmark nonlinear system problem. The homogeneous system for this problem has an unstable equilibrium point at the origin ($x_1 = x_2 = 0$), and a stable limit cycle. These properties make it a nontrivial regulator problem in the sense that without applying appropriate control, the states starting from any non-zero initial condition will never go to zero, and rather develop towards the limit cycle [173, 178]. The system dynamics of a Vander Pol oscillator is given by

$$\dot{x}_1 = x_2 \tag{5.52}$$

$$\dot{x}_2 = -x_1 + \alpha(1 - x_1^2)x_2 + (1 + x_1^2 + x_2^2)u \tag{5.53}$$

The state regulation problem is considered for system defined by (5.52) and (5.53) with quadratic cost function. Consider parameter $\alpha = 0.3$. For implementation of PI algorithm the initial conditions for states and cost function, and critic parameters are taken as $x_0 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$; $P = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$. The length of the simulation in samples is taken 900, number of operations 30, and sample time *T*=0.05 seconds. At each iteration data is measured along state trajectory defined from different initial conditions chosen randomly in the region of interest Ω. Using neural network approximation of cost function the adjusted optimal weights $\mathbf{w}_L^{\mu(i)}$ at 30th iteration during simulation give the adaptive optimal critic parameters *P*, which is obtained as

$$P = \begin{bmatrix} 2.1453 & 0.7862 & 1.6628 \end{bmatrix}$$

Thus the control policy is obtained as

$$u = -\frac{1}{2} R^{-1} \left(1 + x_1^2 + x_2^2\right) P \begin{bmatrix} 0 \\ x_1 \\ 2x_2 \end{bmatrix}$$

$$= -(1 + x_1^2 + x_2^2)(0.3931 x_1 + 1.6628 x_2)$$

The simulation responses of adaptive optimal control using PI technique for Vander Pol oscillator system are shown in Figs. 5.14 to 5.16. Fig. 5.14 shows the system state trajectories showing x_1 and x_2 which converge towards the equilibrium point at origin. Fig. 5.15 shows the control signal trajectory *u* which also converge towards zero. Fig. 5.16 shows the convergence of critic parameters of matrix *P* towards optimal values.

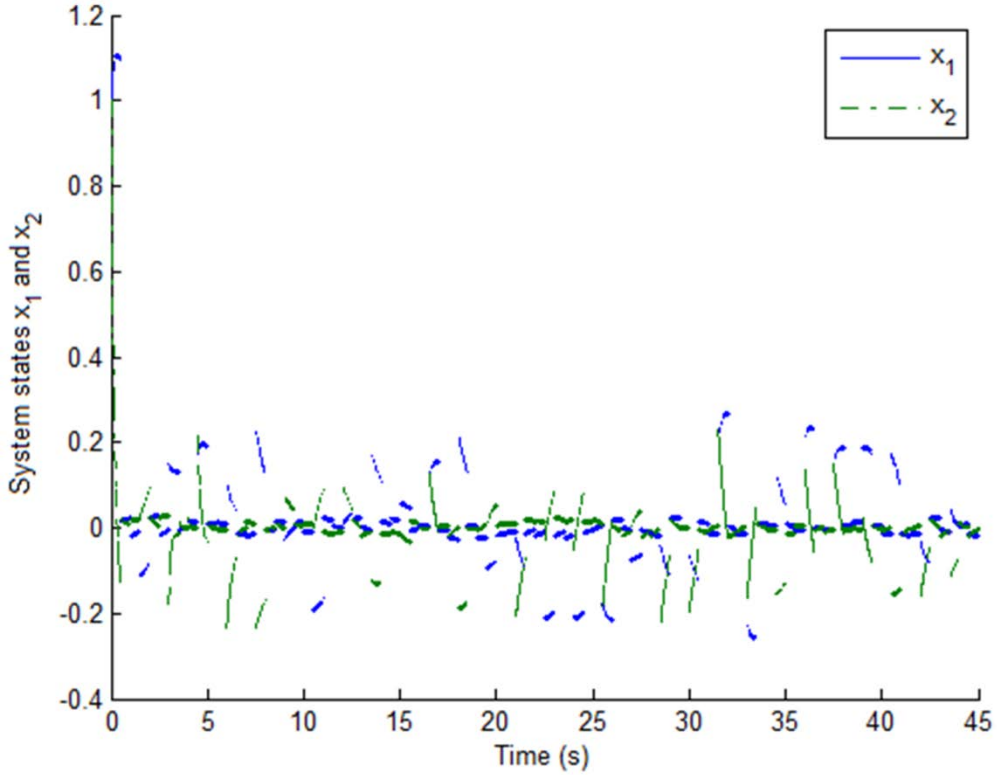


Fig. 5.14 Vander Pol oscillator system states trajectories

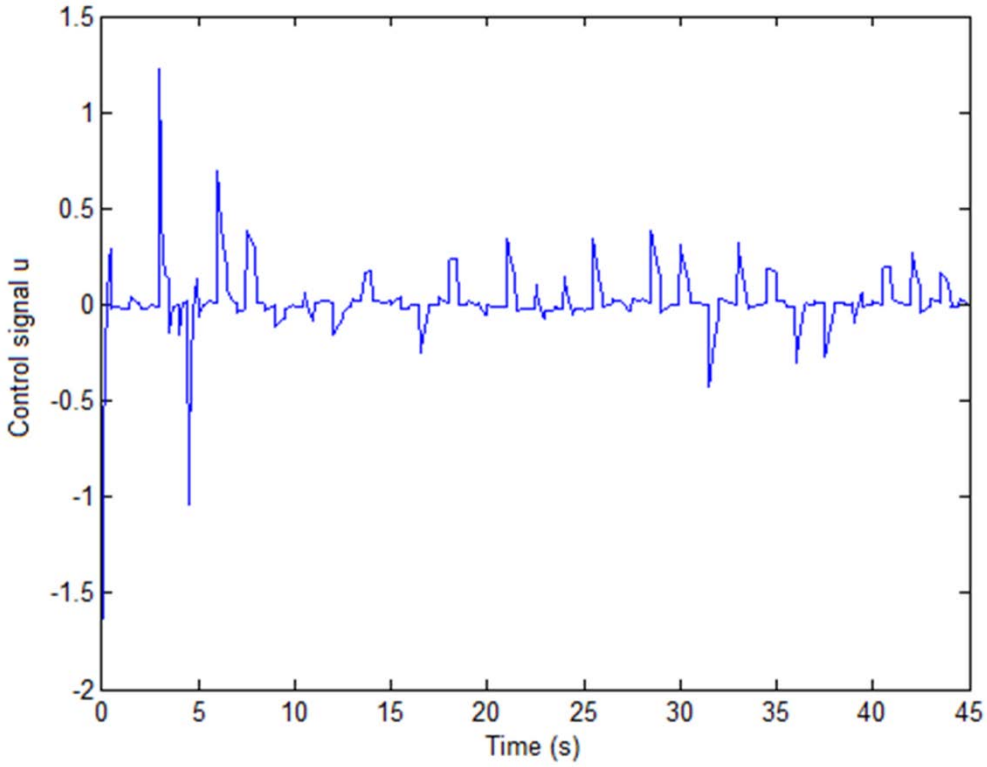


Fig. 5.15 Control signal

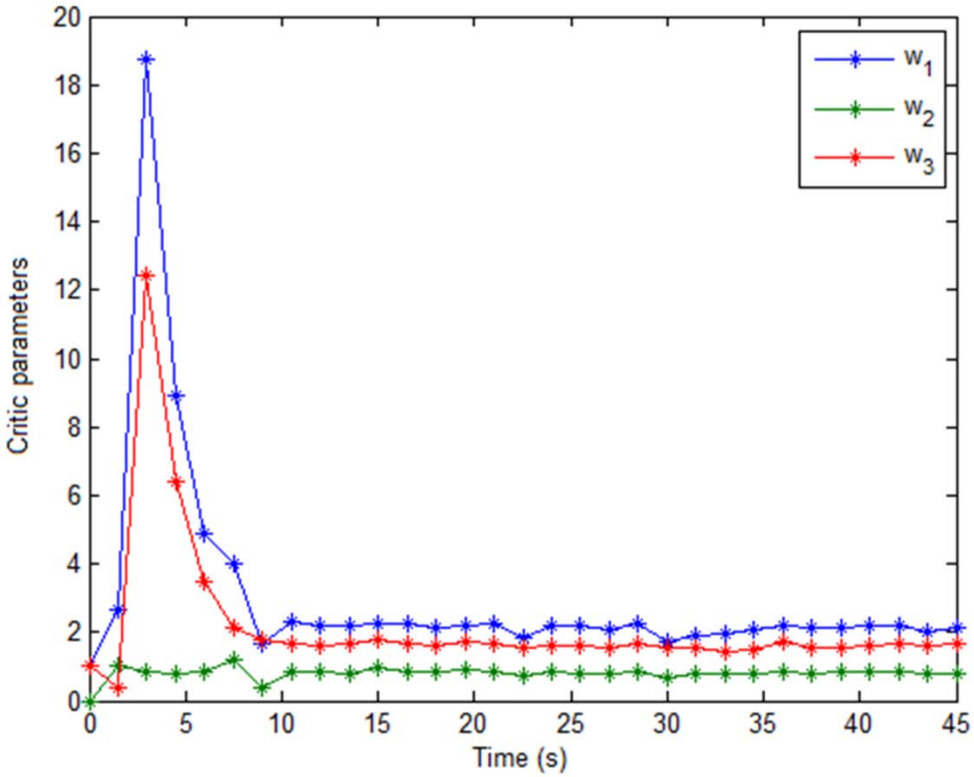


Fig. 5.16 Critic parameters

From the analysis of simulation results obtained for all the above examples of input-affine nonlinear systems, it is observed that the system states converge towards the equilibrium point at origin, and the control signal remains bounded converging towards zero.

The cost function approximation neural networks weights $\mathbf{w}_L^{\mu^{(i)}}$ are adjusted to the optimal values which give the critic parameters matrix *P* converging adaptively to optimal values and thus the control policy is adaptive optimal. The online PI algorithm does not require the system internal state dynamics extracting that information from real-time dynamics by online measurement of data along state trajectory defined from different initial conditions chosen randomly in the region of interest Ω at each iteration, thus providing a partially model-free approach. The online PI algorithm requires an initial stabilizing controller for converging to the optimal solution. The online implementation using neural network approximation of cost function, the recursive least squares algorithm requires a persistence of excitation condition. The adaptive critic design using policy iteration technique provides adaptive optimal control solution for affine nonlinear systems. Thus, the simulation results and performance analysis demonstrate the effectiveness of online policy iteration technique based adaptive critic control scheme.

5.7 CONCLUSIONS

This chapter presents adaptive critic design using online policy iteration technique for continuous-time affine nonlinear dynamical systems. ACD involves two parametric function approximation networks known as a critic and an actor to efficiently approximate the optimal cost and control. Based on actor-critic structure the online PI algorithm consists of two-step iteration: policy evaluation and policy improvement requires an initial stabilizing control policy to converge towards state feedback optimal control. The infinite horizon optimal solution using HJB requires the complete knowledge of the system dynamics, which becomes partially model free using the online policy iteration technique based adaptive critic scheme. The online PI algorithm solves online the continuous-time optimal control problem without using the knowledge of system internal dynamics, the information which is extracted from real-time dynamics by online measurement of sampled states along state trajectory. The knowledge of internal state dynamics (i.e. $f(x)$) is not needed for evaluation of cost or the update of control policy; and only the knowledge of input-to-state dynamics (i.e. $g(x)$) is required for updating the control policy. Thus, the ACD using online PI technique gives an infinite horizon adaptive optimal control solution for continuous-time affine nonlinear systems. In this chapter the neural network approximation of cost function is used for online implementation of PI algorithm with the Kronecker product quadratic polynomial basis vector considered to be the activation function vector. In this section the application of control scheme is implemented considering the state regulation problem for two general and three practical examples of input-affine nonlinear systems- a single-link manipulator, and inverted pendulum system, and Vander Pol oscillator. The simulation results and performance analysis are presented which justify the effectiveness of the proposed control scheme.

CHAPTER 6

INTELLIGENT ADAPTIVE OPTIMAL CONTROL USING SYNCHRONOUS POLICY ITERATION TECHNIQUE FOR CONTINUOUS-TME DYNAMICAL SYSTEMS

This chapter describes the intelligent adaptive optimal control of continuous-time dynamical systems using synchronous policy iteration technique with neural networks approximation of cost function and control policy. The application of control scheme is presented for selected examples of both continuous-time LTI systems and affine nonlinear systems.

6.1 INTRODUCTION

The optimal control problems of continuous-time nonlinear systems can be obtained by the solution of nonlinear Lyapunov equations using neural networks which are trained to be approximate solution. The nonlinear optimal control solution can be obtained by using the policy iteration (PI) algorithm that is built on an actor-critic structure which involves two neural networks (NN) known as critic network and actor network. The critic is trained to approximate the solution of Lyapunov equation at the policy evaluation step, and the actor is trained to approximate the control policy at the policy improvement step. Based on discrete-time information measured from the system, the sequential updates of critic and actor are done in the PI algorithm (i.e. while one is tuned the other one remains constant). In the generalized policy iteration (GPI) algorithm either one or both of the policy evaluation and policy improvement steps are not required to complete before the next step is started. At each step the cost of control policy is not completely evaluated, but only updated the current cost estimate towards that value. The policy improvement step is initiated before the policy evaluation step converges with the assumption that the iterative solution of each of the two steps is approached. This approach is termed as optimistic PI [154]. The synchronous policy iteration algorithm is a variant of GPI. In the synchronous PI algorithm the tuning of both actor and critic neural networks are done simultaneously based on the continuous-time measurement from the environment. The synchronous PI algorithm requires the complete knowledge of system dynamics for solution, and also requires a certain persistence of excitation (PE) condition for convergence [161]. The online synchronous PI technique provides online adaptive optimal control solution for continuous-time linear and nonlinear systems. The synchronous PI algorithm with tuning algorithms for both actor and critic networks, and convergence proof are proposed recently in certain papers [154, 161] In this chapter the applications of synchronous PI technique using neural networks approximation of cost function and control policy for intelligent adaptive optimal control of continuous-time LTI systems and affine nonlinear systems are presented. For the implementation of online synchronous PI based control scheme, two practical examples of LTI systems- LFC system and AVR system are considered. And three general systems and two practical systems

examples of affine nonlinear systems- single-link manipulator and Vander Pol oscillator are considered for applications of control scheme. The concepts of infinite horizon optimal control of continuous-time systems, and continuous-time adaptive critics are same as discussed in sections 5.2 and 5.3 respectively. The formulation of infinite horizon optimal control of continuous-time affine nonlinear systems as discussed in section 5.2 is referred as same for the mathematical formulations in the following sections of this chapter.

6.2 SYNCHRONOUS POLICY ITERATION TECHNIQUE

The synchronous policy iteration technique is an iterative method which belongs to the reinforcement learning technique, tunes the actor and critic networks concurrently at the same time. The synchronous PI algorithm, neural networks approximation of value function, the tuning and convergence of actor and critic networks, and adaptive optimal control using synchronous PI technique for continuous-time systems are discussed in the following sub-sections.

6.2.1 Synchronous Policy Iteration Algorithm

Same to the PI algorithm the synchronous PI algorithm consists of two iterative steps [158]:

1. Policy evaluation: given $\mu^{(i)}(x)$, solve for the value $V^{\mu^{(i)}}(x(t))$ using nonlinear Lyapunov equation

$$\begin{aligned} &0 = L(x, \mu^{(i)}(x)) + (\lambda^{(i)})^T (f(x) + g(x)\mu^{(i)}(x)), \\ &V^{\mu^{(i)}}(0) = 0 \end{aligned} \tag{6.1}$$

2. Policy improvement: update the control policy using Hamiltonian functional

$$\mu^{(i+1)} = \arg\min_{u \in \Psi(\Omega)} \left[H(x, u, \lambda^{(i)}) \right] \tag{6.2}$$

which explicitly is

$$\mu^{(i+1)}(x) = -\frac{1}{2} R^{-1} g^T(x) \lambda^{(i)} \tag{6.3}$$

where the variables in above equations have same meaning as discussed in section 5.2 and 5.4. The PI algorithm requires an admissible policy $\mu^{(0)}(x(t)) \in \Psi(\Omega)$ to ensure its convergence. The convergence proofs of synchronous PI algorithms are presented for continuous-time systems in [158, 161].

6.2.2 Neural Network Approximation of Value Function

In the standard PI algorithm, the critic and actor are updated alternatively by solving (6.1) and (6.3). In actor-critic structure the cost function $V^{\mu^{(i)}}(x(t))$, and the control policy $\mu^{(i+1)}(x)$ are approximated by neural networks (NN) at each step of PI algorithm, known as

critic NN and actor NN respectively which are tuned alternatively in the least square sense to solve the value function (6.1) and the control policy (6.3) respectively. Thus, while one neural network is being tuned, the other is held constant, whereas in synchronous PI algorithm these two networks are tuned simultaneously in real-time to guarantee convergence to the control policy as well as stability during the training process [158, 161].

Let the assumptions that the solution of nonlinear Lyapunov equation is smooth (i.e. $V(x) \in C^1(\Omega)$), and positive definite (i.e. $Q(x) > 0$, $x \in \Omega - \{0\}$; $Q(0) = 0$) which is guaranteed for the stabilizable system dynamics if performance functional satisfies zero-state observability. With such assumptions, there exists a complete independent basis set $\{\phi_1(x)\}$ such that the solution $V(x)$ and its gradient $\lambda = \frac{\partial V(x)}{\partial x}$ are uniformly approximated according to Weirstrass higher-order approximation theorem, and there exist coefficients c_i, such that [158].

$$V(x) = \sum_{i=1}^{\infty} c_i \phi_i(x) = \sum_{i=1}^{N} c_i \phi_i(x) + \sum_{i=N+1}^{\infty} c_i \phi_i(x)$$

$$V(x) \equiv C_1^T \varphi_1(x) + \sum_{i=N+1}^{\infty} c_i \phi_i(x) \tag{6.4}$$

$$\frac{\partial V(x)}{\partial x} = \sum_{i=1}^{\infty} c_i \frac{\partial \phi_i(x)}{\partial x} = \sum_{i=1}^{N} c_i \frac{\partial \phi_i(x)}{\partial x} + \sum_{i=N+1}^{\infty} c_i \frac{\partial \phi_i(x)}{\partial x} \tag{6.5}$$

where $\varphi_1(x) = [\phi_1(x) \;\; \phi_2(x) \; ... \; \phi_N(x)]^T : \mathcal{R}^n \to \mathcal{R}^N$. The last terms in (6.4) and (6.5) converge to zero as $N \to \infty$.

Since the critic NN is based on value function approximation (VFA) in Sobolev norm (i.e. approximation of value function $V(x)$, and its gradient) [158], thus, consider that it has weights W_1 such that the value $V(x)$ is approximated by a neural network as

$$V(x) = W_1^T \varphi_1(x) + \varepsilon(x) \tag{6.6}$$

where $\varphi_1(x) : \mathcal{R}^n \to \mathcal{R}^N$ is the activation functions vector, *N* is the number of neurons in the hidden layer, and $\varepsilon(x)$ is the neural networks approximation error which is bounded by a constant on a compact set. Selecting the activation functions $\{\phi_i(x) : i = 1, N\}$ to provide a complete basis set $\{\phi_i(x) : i = 1, \infty\}$ such that $V(x)$ and its derivative are uniformly approximated.

Taking the derivative of value function $V(x)$ with respect to *x,* we have

$$\frac{\partial V(x)}{\partial x} = \left(\frac{\partial \varphi_1^T(x)}{\partial x} \right)^T W_1 + \frac{\partial \varepsilon(x)}{\partial x} = \nabla \varphi_1^T(x) W_1 + \nabla \varepsilon \tag{6.7}$$

According to the Weirstrass higher-order approximation theorem a complete basis set exists if $V(x)$ is sufficiently smooth, and thus, as the number of hidden-layer neurons $N \to \infty$, the approximation errors $\varepsilon \to 0$, $\nabla\varepsilon \to 0$ uniformly. Also these errors are bounded by constants on a compact set.

Using the neural networks approximation of value function, considering a fixed control policy *u*(*t*), the nonlinear Lyapunov equation, and thus the Hamiltonian becomes

$$H(x,u,W_1) = W_1^T \nabla\varphi_1 (f + gu) + Q(x) + u^T R u = \varepsilon_H \tag{6.8}$$

where the residual error due to the function approximation error is given by

$$\varepsilon_H = -(\nabla\varepsilon)^T (f + gu) \tag{6.9}$$

This residual error is bounded on a compact set under the Lipschitz assumption on the system dynamics. The detail discussion on convergence of solution can be referred in [154, 158].

The effect of approximation error on HJB equation is given by

$$W_1^T \nabla\varphi_1 f - \frac{1}{4} W_1^T \nabla\varphi_1 g R^{-1} g^T \nabla\varphi_1^T W_1 + Q(x) = \varepsilon_{HJB} \tag{6.10}$$

where the residual error due to function approximation error is given by

$$\varepsilon_{HJB} = -\nabla\varepsilon^T f + \frac{1}{2} W_1^T \nabla\varphi_1 g R^{-1} g^T \nabla\varepsilon + \frac{1}{4} \nabla\varepsilon^T g R^{-1} g^T \nabla\varepsilon \tag{6.11}$$

As the number of hidden layer neurons increases the error ε_{HJB} converges uniformly to zero (i.e. $\forall\varepsilon > 0$, $\exists N(\varepsilon) : \sup_{x\in\Omega} \left\| \varepsilon_{HJB} \right\| < \varepsilon$).

6.2.3 Tuning and Convergence Analysis of Online Synchronous PI Algorithm

The standard PI algorithms for continuous-time systems which are offline methods require the complete knowledge of system dynamics (i.e. $f(x)$ and $g(x)$) to obtain the solution. Consistent with the online learning mechanism in the mammal brain, the online synchronous PI algorithm is an adaptive learning algorithm that uses simultaneous continuous-time tuning for the actor and critic neural networks for continuous-time systems. The synchronous PI algorithm also require the complete knowledge of the system dynamics, yet can approximately solve the optimal control problem online using real-time measurements of closed-loop signals. The synchronous PI algorithm is consistent with adaptive control schemes which first design an observer for system state and unknown dynamics, and then design feedback control using this observer. The value function approximation using neural networks is like the observer design for value function. The following subsections briefly present the synchronous tuning and convergence of critic NN and actor NN respectively. For the detail discussion on tuning and convergence analysis for

critic NN and actor NN of synchronous PI algorithm, the assumptions, lemmas, theorems, definitions, and remarks may be referred in [154, 158, 161].

6.2.3.1 Critic Neural Networks

The critic NN weights W_1, that provide the best approximate solution for (6.8) are unknown. Thus the output of the critic NN is

$$\hat{V}(x) = \hat{W}_1^T \varphi_1(x) \tag{6.12}$$

where $\hat{W}_1$ are the current estimated values of the critic NN weights. Since $\varphi_1(x) : \mathbb{R}^n \to \mathbb{R}^N$ is the activation functions vector with *N* the number of neurons in the hidden layer, then the approximate Hamiltonian is given by

$$H(x, \hat{W}_1, u) = \hat{W}_1^T \nabla \varphi_1 (f + gu) + Q(x) + u^T R u = e_1 \tag{6.13}$$

Define the critic weight estimation error as

$$\tilde{W}_1 = W_1 - \hat{W}_1 \tag{6.14}$$

then

$$e_1 = -\tilde{W}_1^T \nabla \varphi_1 (f + gu) + \varepsilon_H \tag{6.15}$$

For a given admissible control policy *u*, it is desired to select $\hat{W}_1$ such that the squared residual error is minimized. The squared residual error is given by

$$E_1 = \frac{1}{2} e_1^T e_1 \tag{6.16}$$

then $\hat{W}_1(t) \to W_1$, and . $e_1 \to \varepsilon_H$.

Consider the normalized gradient descent algorithm as the tuning law for the critic NN weights as

$$\dot{\hat{W}}_1 = -a_1 \frac{\partial E_1}{\partial \hat{W}_1} = -a_1 \frac{\sigma_1}{(\sigma_1^T \sigma_1 + 1)^2} \left[\sigma_1^T \hat{W}_1 + Q(x) + u^T R u \right] \tag{6.17}$$

where $\sigma_1 = \nabla \varphi_1 (f + gu)$. This is a modified Levenberg-Marquardt algorithm in which $(\sigma_1^T \sigma_1 + 1)^2$ is used for normalization instead of $(\sigma_1^T \sigma_1 + 1)$. This is required in convergence proofs [154, 157, 158, 161].

From (6.8) we have

$$Q(x) + u^T R u = -W_1^T \nabla \varphi_1 (f + gu) + e_H \tag{6.18}$$

Let the notations

$$\bar{\sigma}_1 = \frac{\sigma_1}{(\sigma_1^T \sigma_1 + 1)} \ , \qquad m_s = (\sigma_1^T \sigma_1 + 1) \tag{6.19}$$

Substituting (6.18) in (6.17) and with (6.19), the critic weight estimation error dynamics is obtained as

$$\dot{\tilde{W}}_1 = -a_1\bar{\sigma}_1\bar{\sigma}_1^T\tilde{W}_1 + a_1\bar{\sigma}_1\frac{\varepsilon_H}{m_s} \tag{6.20}$$

From (6.20) with defining output *y*, the error dynamical system is defined as

$$\begin{aligned}\dot{\tilde{W}}_1 &= -a_1\bar{\sigma}_1\bar{\sigma}_1^T\tilde{W}_1 + a_1\bar{\sigma}_1\frac{\varepsilon_H}{m_s}\\ y &= \bar{\sigma}_1^T\tilde{W}_1\end{aligned} \tag{6.21}$$

Even though the critic tuning algorithms of the form (6.17) are traditionally used, when the convergence of critic weights can be guaranteed is generally not known. To guarantee the convergence of $\hat{W}_1$ to W_1, the persistence of excitation (PE) assumption and associated technical lemmas are proposed in [154, 157, 158, 161].

Persistence of Excitation Assumption [154, 157, 158, 161]: Let the signal $\bar{\sigma}_1$ be persistently exciting over the interval $[t, t+T]$, i.e. there exist constants $\beta_1 > 0, \beta_2 > 0, T > 0$ such that, for all *t*,

$$\beta_1 I \le S_0 \equiv \int_t^{t+T}\bar{\sigma}_1(\tau)\bar{\sigma}_1^T(\tau)d\tau \le \beta_2 I \tag{6.22}$$

The PE assumption is required in adaptive control with system identification using recursive least squares (RLS) algorithm. The PE is required to identify the critic parameters to approximate $V(x)$. The PE condition (6.22) is equivalent to the uniform complete observability (UCO) of the error dynamics system (6.21). According to UCO that gives the bounded-input-bounded-output (BIBO) condition, the state $\tilde{W}_1(t)$ is bounded. The lemmas, theorems, and remarks for proof of convergence of critic NN can be referred in [154, 157, 158, 161]:.

The error dynamics system (6.21) is exponentially stable. In fact if $\varepsilon_H = 0$ then $\tilde{W}(t)$ decays exponentially, and $\left\|\tilde{W}(kT)\right\| \le e^{-\alpha kT}\left\|\tilde{W}(0)\right\|$ with decay constant

$$\alpha = -\frac{1}{T}\ln\left(\sqrt{1-2a_1\beta_3}\right) \tag{6.23}$$

This can be proved by taking the Lyapunov function $L = \frac{1}{2}\tilde{W}_1^T a_1^{-1}\tilde{W}_1$ and PE condition.

Let $\left\|\varepsilon_H\right\| \le \varepsilon_{\max}$ and $\left\|y\right\| \le y_{\max}$ then $\left\|\tilde{W}_1\right\|$ converges exponentially to the residual set

$$\tilde{W}_1(t) \le \frac{\sqrt{\beta_2 T}}{\beta_1}\left\{y_{\max} + \delta\beta_2 a_1\left(\varepsilon_{\max} + y_{\max}\right)\right\} \tag{6.24}$$

where δ is a positive constant of the order of 1. Equation (6.24) can be proved using the state solution of error dynamics system (6.21) considering $\dot{\tilde{W}}_1(t) = a_1\bar{\sigma}_1 u$, $\dot{x} = B(t)u(t)$, $y = C^T(t)x(t)$, where $x(t) = \tilde{W}_1$, $B = a_1\bar{\sigma}_1$, $C = \bar{\sigma}_1$ that be PE, and setting $u = -y + \frac{\varepsilon_H}{m_s}$, then,

$$\|u\| \le \|y\| + \left\|\frac{\varepsilon_H}{m_s}\right\| \le y_{\max} + \varepsilon_{\max}. \tag{6.25}$$

Following Theorem 6.1 shows that the critic tuning law (6.17) indeed guarantees the boundedness of the output of error dynamic system [154, 157, 158, 161]:.

Theorem 6.1 [154, 157, 158, 161]**:** Let $u(t)$ be any admissible bounded control input. Let tuning for the critic NN be provided by (6.17) and assume that $\bar{\sigma}_1$ is persistently exciting. Let the residual error in (6.8) be bounded $\|\varepsilon_H\| \le \varepsilon_{\max}$. Then the critic parameter error is practically bounded by

$$\tilde{W}_1(t) \le \frac{\sqrt{\beta_2 T}}{\beta_1}\left\{\left[1 + 2\delta\beta_2 a_1\right]\varepsilon_{\max}\right\} \tag{6.26}$$

Proof: Consider the Lyapunov function candidate as following:

$$L(t) = \frac{1}{2} tr\left\{\tilde{W}_1^T a_1^{-1}\tilde{W}_1\right\} \tag{6.27}$$

The derivative of *L* is given by

$$\dot{L} = -tr\left\{\tilde{W}_1^T \frac{\sigma_1}{m_s^2}\left[\sigma_1^T\tilde{W}_1 - \varepsilon_H\right]\right\}$$

$$\dot{L} = -tr\left\{\tilde{W}_1^T \frac{\sigma_1\sigma_1^T}{m_s^2}\tilde{W}_1\right\} + tr\left\{\tilde{W}_1^T \frac{\sigma_1}{m_s}\frac{\varepsilon_H}{m_s}\right\}$$

$$\dot{L} \le -\left\|\frac{\sigma_1^T}{m_s}\tilde{W}_1\right\|^2 + \left\|\frac{\sigma_1^T}{m_s}\tilde{W}_1\right\|\left\|\frac{\varepsilon_H}{m_s}\right\|$$

$$\dot{L} = -\left\|\frac{\sigma_1^T}{m_s}\tilde{W}_1\right\|\left[\left\|\frac{\sigma_1^T}{m_s}\tilde{W}_1\right\| - \left\|\frac{\varepsilon_H}{m_s}\right\|\right]$$

Therefore $\dot{L} \le 0$ if

$$\left\|\frac{\sigma_1^T}{m_s}\tilde{W}_1\right\| > \varepsilon_{\max} > \left\|\frac{\varepsilon_H}{m_s}\right\| \tag{6.28}$$

since $\|m_s\| \ge 1$. If (6.28) satisfies $L(t)$ decreases that provides an effective practical bound for $\left\|\bar{\sigma}_1^T\tilde{W}_1\right\|$.

The estimation error dynamics (6.21) with its output bounded effectively by $\|y\| < \varepsilon_{\max}$, (6.24) shows the exponential convergence to the residual set

$$\tilde{W}_1(t) \le \frac{\sqrt{\beta_2 T}}{\beta_1}\left\{\left[1 + 2\delta\beta_2 a_1\right]\varepsilon_{\max}\right\} \tag{6.29}$$

This completes the proof.

Since $N \to \infty$, $\varepsilon_H \to 0$ uniformly, thus $\varepsilon_{\max}$ decreases with increasing the number of hidden layer neurons. For the given control policy $u(t)$, the critic NN weights $\hat{W}_1$ converge to the actual unknown weights W_1, which solve the Hamiltonian equation (6.8). Thus, (6.12) converges close to the actual value function of the current control policy. Thus, this theorem shows that the tuning algorithm (6.17) is effective.

6.2.3.2 Actor Neural Networks

In online synchronous PI algorithm the weights of both actor and critic neural networks are tuned at the same time. The justified form of actor NN is desired to be determined. Let the solution $V(x) \in C^1(\Omega)$ to the nonlinear Lyapunov equation (6.1) for a given admissible policy *u*(*t*) is given by (6.4). Then by substituting (6.5) into (6.3) the policy improvement step in PI is given as

$$u(x) = -\frac{1}{2}R^{-1}g^T(x)\sum_{i=1}^{\infty} c_i \frac{\partial \phi_i(x)}{\partial x} \tag{6.30}$$

for some unknown coefficients c_i. Let the least square solution to (6.8) be W_1, then control policy with unknown critic weights W_1, is given by

$$u_1(x) = -\frac{1}{2}R^{-1}g^T(x)\nabla V_1(x) = -\frac{1}{2}R^{-1}g^T(x)\nabla \varphi_1^T(x)W_1 \tag{6.31}$$

Thus, the control policy in the form of an action NN that computes the control input in the structured form can be defined as

$$u_2(x) = -\frac{1}{2}R^{-1}g^T(x)\nabla \varphi_1^T \hat{W}_2 \tag{6.32}$$

where $\hat{W}_2$ denotes the current known values of the actor NN weights.

Define the actor NN estimation error as

$$\tilde{W}_2 = W_1 - \hat{W}_2 \tag{6.33}$$

Definition 6.1 [158]: The equilibrium point $x_e = 0$ of system $\dot{x} = f(x) + g(x)u$ is said to be uniformly ultimately bounded (UUB) if there exists a compact set $S \subset \mathcal{R}^n$ so that for all $x_0 \in S$ there exists a bound *B* and a time $T(B, x_0)$ such that $\|x(t) - x_e\| \le B$ for all $t \ge t_0 + T$.

Assumption 6.1 [158]**:** It is assumed that

(a) $f(\bullet)$ is Lipschitz continuous, and $g(\bullet)$ is bounded by a constant, $\|f(x)\| < b_f \|x\|$, $\|g(x)\| < b_g$.

(b) The NN approximation error and its gradient are bounded on a compact set containing Ω so that $\|\varepsilon\| < b_\varepsilon$, $\|\nabla\varepsilon\| < b_{\varepsilon_x}$.

(c) The NN activation functions and their gradients are bounded so that $\|\varphi_1(x)\| < b_\varphi$, $\|\nabla\varphi_1(x)\| < b_{\varphi_x}$.

The following theorem which is proposed in [158, 161] provides the tuning laws for the actor and critic NNs that guarantee convergence to the optimal controller along with closed-loop stability.

Theorem 6.2 [158, 161]**:** Let the system dynamics be given by $\dot{x} = f(x) + g(x)u$, $x(0) = x_0$, the critic NN be given by (6.12) and the control input be given by actor NN (6.32). Let tuning for the critic NN be provided by

$$\dot{\hat{W}}_1 = -a_1 \frac{\sigma_2}{(\sigma_2^T\sigma_2 + 1)^2}\left[\sigma_2^T\hat{W}_1 + Q(x) + u_2^T R u_2\right] \tag{6.34}$$

where $\sigma_2 = \nabla\varphi_1(f + gu_2)$, and assume that $\bar{\sigma}_2 = \dfrac{\sigma_2}{(\sigma_2^T\sigma_2 + 1)}$ is persistently exciting. Let the actor NN be tuned as

$$\dot{\hat{W}}_2 = -a_2\left[\left(F_2\hat{W}_2 - F_1\bar{\sigma}_2^T\hat{W}_1\right) - \frac{1}{4}\bar{D}_1(x)\hat{W}_2 m^T(x)\hat{W}_1\right] \tag{6.35}$$

where

$$\bar{D}_1(x) \equiv \nabla\varphi_1(x)g(x)R^{-1}g^T(x)\nabla\varphi_1^T(x)$$

$$m \equiv \frac{\sigma_2}{(\sigma_2^T\sigma_2 + 1)^2}$$

and $F_1 > 0$ and $F_2 > 0$ are tuning parameters. Then there exists an N_0, such that for number of hidden layer units $N > N_0$, the closed-loop system state, the critic parameter error $\tilde{W}_1 = W_1 - \hat{W}_1$ and the actor parameter error $\tilde{W}_2 = W_1 - \hat{W}_2$ are UUB, and (6.26) holds asymptotically so that convergence of $\hat{W}_1$ to the approximate optimal critic value W_1 is obtained.

Proof: The convergence proof is based on Lyapunov analysis considering the Lyapunov function as

$$L(t) = V(x) + \frac{1}{2} tr\left(\tilde{W}_1^T a_1^{-1} \tilde{W}_1\right) + \frac{1}{2}\left(\tilde{W}_2^T a_2^{-1} \tilde{W}_2\right) \tag{6.36}$$

With the chosen tuning laws it can be shown that the errors $\tilde{W}_1$ and $\tilde{W}_2$ are UUB and the convergence is obtained.

The derivative of Lyapunov function is given by

$$\begin{aligned} \dot{L}(x) &= \dot{V}(x) + \tilde{W}_1^T a_1^{-1} \dot{\tilde{W}}_1 + \tilde{W}_2^T a_2^{-1} \dot{\tilde{W}}_2 \\ &= \dot{L}_V(x) + \dot{L}_1(x) + \dot{L}_2(x) \end{aligned} \tag{6.37}$$

By expanding each term in (6.37) and further deriving it is proved that if *L* exceeds a certain bound, then, $\dot{L}$ is negative, which guarantees the convergence and closed-loop stability. After further derivation of (6.37) and analysis it can be demonstrated that the state and weights are UUB. The detail convergence proof of theorem 6.2 is referred in [158].

6.3 ADAPTIVE OPTIMAL CONTROL USING SYNCHRONOUS POLICY ITERATION TECHNIQUE

This section presents the implantation issues of adaptive optimal control using online synchronous PI technique for continuous-time linear time-invariant (LTI) systems and input-affine nonlinear systems. The adaptive optimal control using actor-critic structure for LTI systems is shown in Fig. 6.1, and for affine nonlinear systems is shown in Fig. 6.2 [158, 161]. Two neural networks- critic NN and actor NN are used for online implementation of synchronous PI algorithm. For the LTI systems application the formulation of (6.1), (6.2) and (6.3) is modified for $\dot{x} = Ax + Bu$, by replacing $f(x) = Ax$, and $g(x) = B$. The critic *P* and actor *K* are obtained from optimal weights of critic NN and actor NN respectively.

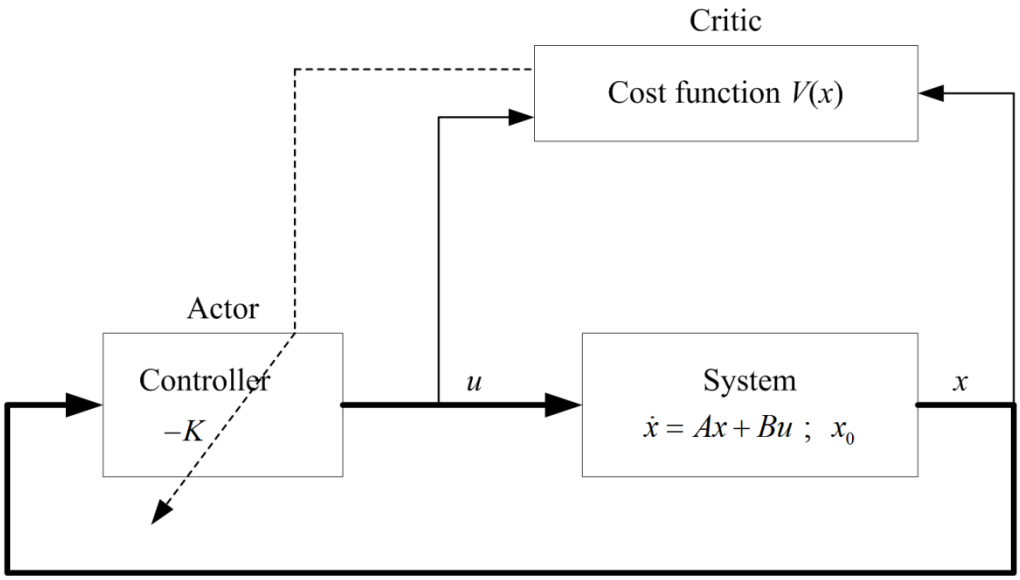


Fig. 6.1 Adaptive optimal control with actor-critic structure using synchronous PI trchnique for LTI systems.

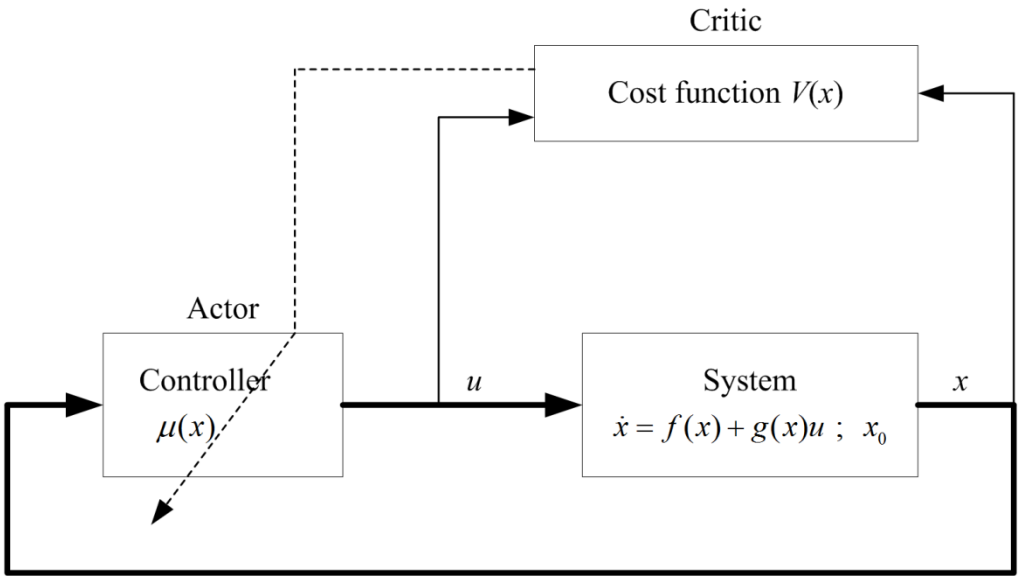


Fig. 6.2 Adaptive optimal control with actor-critic structure using synchronous PI trchnique for affine nonlinear systems.

In this control system structure the critic and the actor functions are approximated by neural networks. The critic NN and the actor NN are tuned synchronously in the synchronous PI algorithm to obtain the solution of (6.1) and (6.3), which are the policy evaluation and policy improvement steps of algorithm respectively. The online synchronous PI algorithm uses the continuous-time data measured from the system for online concurrent tuning of the critic NN and the actor NN. The tuning laws of the critic and actor can be interpreted as the extra state dynamics of the closed loop actor-critic control system. The equilibrium points of the extra states are described by the dynamics of the critic and actor parameters respectively characterizing the optimal critic and optimal actor. The implementation of synchronous PI algorithm requires the complete knowledge (i.e. internal state dynamics $f(x)$, and control input-to-state dynamics $g(x)$) of system dynamics. The convergence of synchronous PI algorithm requires the persistence of excitation (PE) condition on the signals that are used for the training of the critic NN and actor NN. The PE condition is explicitly embedded in the adaptation laws by adding a small probing noise to the control signal. On the convergence, the PE condition is no longer required, and the probing noise signal is turned off. The learning constants and tuning parameters are suitably selected for exponential convergence and the same values of critic and actor weights. In this implementation the activation functions vector $\varphi(x)$ is selected from system state variables considering generally in the sense of Kronecker product quadratic polynomial basis vector.

The online synchronous PI technique provides online adaptive optimal control solution with infinite horizon cost for affine-input continuous-time systems with known dynamics. Using neural networks approximations for cost function and control policy in the online implementation, the synchronous PI technique based adaptive critic scheme provides an

intelligent adaptive optimal control solution for continuous-time time-invariant dynamical systems.

The applications of this control scheme are presented for both continuous-time LTI systems and affine nonlinear systems. The simulation results and analysis for LTI systems and affine nonlinear systems are presented in the following sections respectively.

6.4 SIMULATION RESULTS AND ANALYSIS FOR LTI SYSTEMS

In this section, the application of online synchronous PI technique based adaptive critic scheme with neural networks approximation of critic and actor functions for adaptive optimal control is implemented for two practical examples of LTI systems- load frequency control (LFC) of power system, and automatic voltage regulator (AVR) of power system. LFC system models without and with integral control; and AVR system models neglecting and including sensor dynamics are considered for implementation of control scheme and performance analysis. The simulation results and performance analysis are presented to demonstrate the effectiveness of the control scheme. The performance of adaptive critic control scheme using synchronous PI technique is compared with LQR approach.

The quadratic cost function is considered for these examples. The Kronecker product quadratic polynomial basis vector is considered as the critic NN activation functions vector $\varphi_1(x)$. For convergence the PE condition ensured by adding a small probing noise given by

$$n(t) = e^{-0.05t}\sin^2(t)\cos(t)$$

The probing noise is turned off after convergence is reached.

6.4.1 Load Frequency Control of Power System

Consider the single-area power system load frequency control problem as discussed in section 4.6.3.

6.4.1.1 LFC System Model without Integral Control

As discussed in section 4.6.3.1 the state space model of LFC system without integral control is given by

$$A = \begin{bmatrix} -0.0665 & 11.5 & 0 \\ 0 & -2.5 & 2.5 \\ -9.5 & 0 & -13.7360 \end{bmatrix},\quad B = \begin{bmatrix} 0 \\ 0 \\ 13.7360 \end{bmatrix},\quad F = \begin{bmatrix} -11.5 \\ 0 \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},\ D = 0$$

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0 = \begin{bmatrix} 0 & 0.1 & 0 \end{bmatrix}$, and both the critic and actor NNs weight vectors W_1 and W_2 are initialized with elements of unity values. The learning constants are taken as $a_1 = 50$, $a_2 = 1$.

The quadratic cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The critic NN activation function vector is selected as

$$\varphi_1(x) = \begin{bmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_2^2 & x_2x_3 & x_3^2 \end{bmatrix}^T$$

Let the critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1 = W_c = \begin{bmatrix} W_{c1} & W_{c2} & W_{c3} & W_{c4} & W_{c5} & W_{c6} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} & W_{a4} & W_{a5} & W_{a6} \end{bmatrix}^T$$

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \le 150$. After 200 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.7383 & 1.0215 & 0.0569 & 1.0321 & 0.7407 & -0.0493 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 0.7383 & 1.0215 & 0.0569 & 1.0321 & 0.7407 & -0.0493 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2} R^{-1} \begin{bmatrix} 0 \\ 0 \\ 13.7360 \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 & 0 \\ x_2 & x_1 & 0 \\ x_3 & 0 & x_1 \\ 0 & 2x_2 & 0 \\ 0 & x_3 & x_2 \\ 0 & 0 & 2x_3 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control $u = -R^{-1}B^T Px$.

The unique positive definite solution of ARE denoted here by matrix *RicP* is obtained as

$$RicP = \begin{bmatrix} 0.3600 & 0.5313 & 0.0367 \\ 0.5313 & 1.5662 & 0.1690 \\ 0.0367 & 0.1690 & 0.0500 \end{bmatrix}$$

and from critic NN weights W_c the adaptive optimal critic matrix *P* of adaptive critic scheme using synchronous PI is obtained as

$$P = \begin{bmatrix} 0.7383 & 0.5107 & 0.0285 \\ 0.5107 & 1.0321 & 0.3703 \\ 0.0285 & 0.3703 & -0.0493 \end{bmatrix}$$

The actor gains of LQR design *RicK*, and actor *K* of adaptive critic scheme using synchronous PI respectively are obtained as

$$RicK = \begin{bmatrix} 0.5044 & 2.3212 & 0.6867 \end{bmatrix} \quad , \quad K = \begin{bmatrix} 0.3909 & 5.0871 & -0.6769 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-2.5140, -2.2453 +13.0352i, -2.2453 -13.0352i

Fig. 6.3 shows the system state trajectories which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.4 and 6.5. Fig. 6.6 presents the closed loop response of LFC system without integral control model using both approaches of LQR and adaptive critic (AC). It is observed that the closed loop response with unit step load disturbance have steady state error.

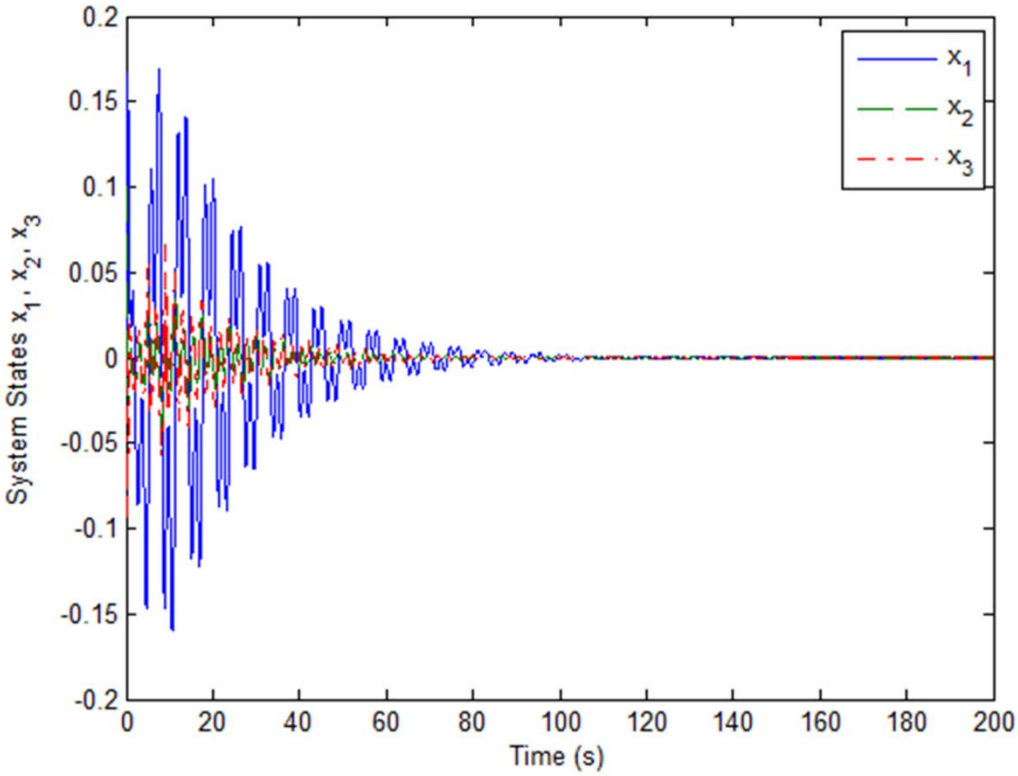


Fig. 6.3 System states trajectories.

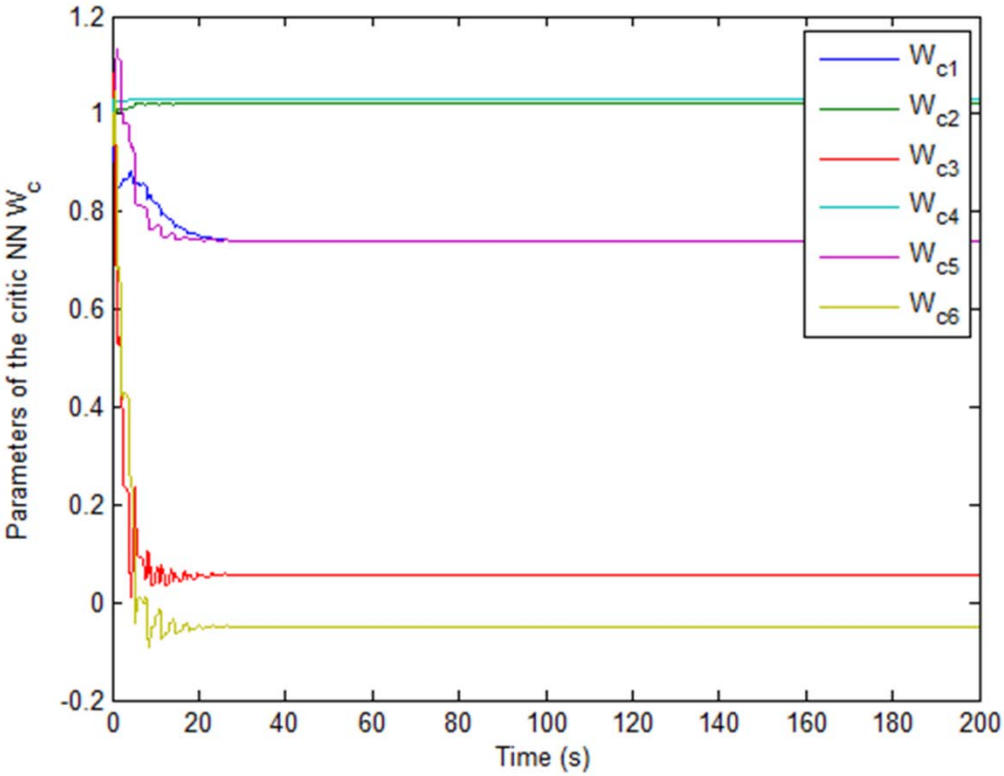


Fig. 6.4 Convergence of critic parameters.

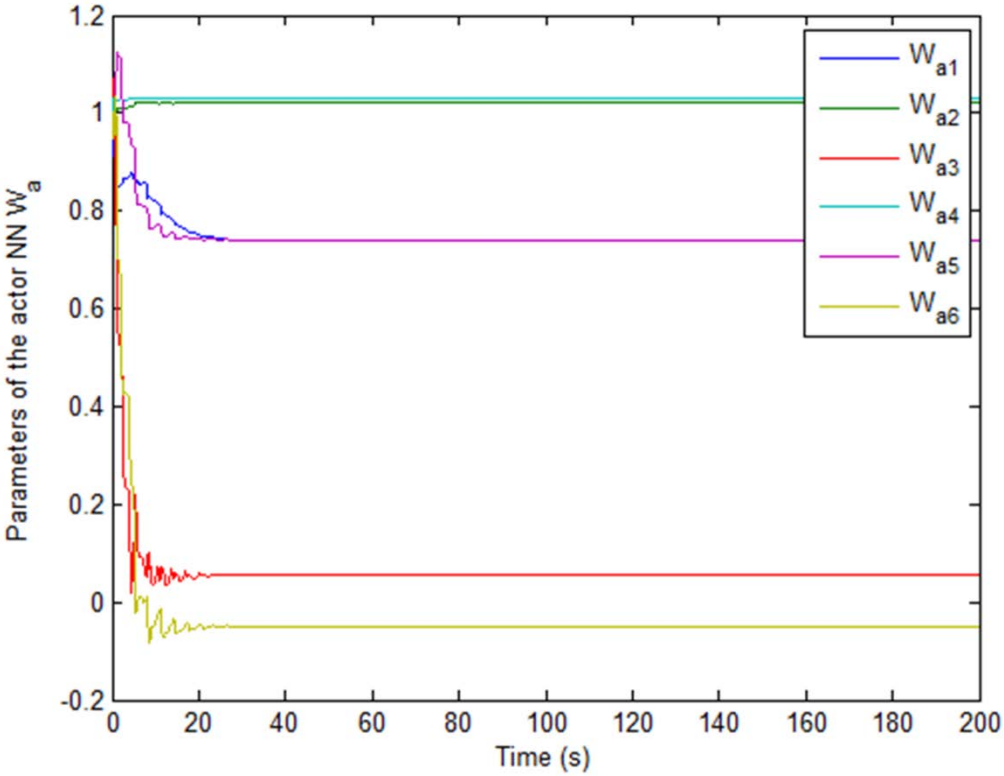


Fig. 6.5 Convergence of actor parameters.

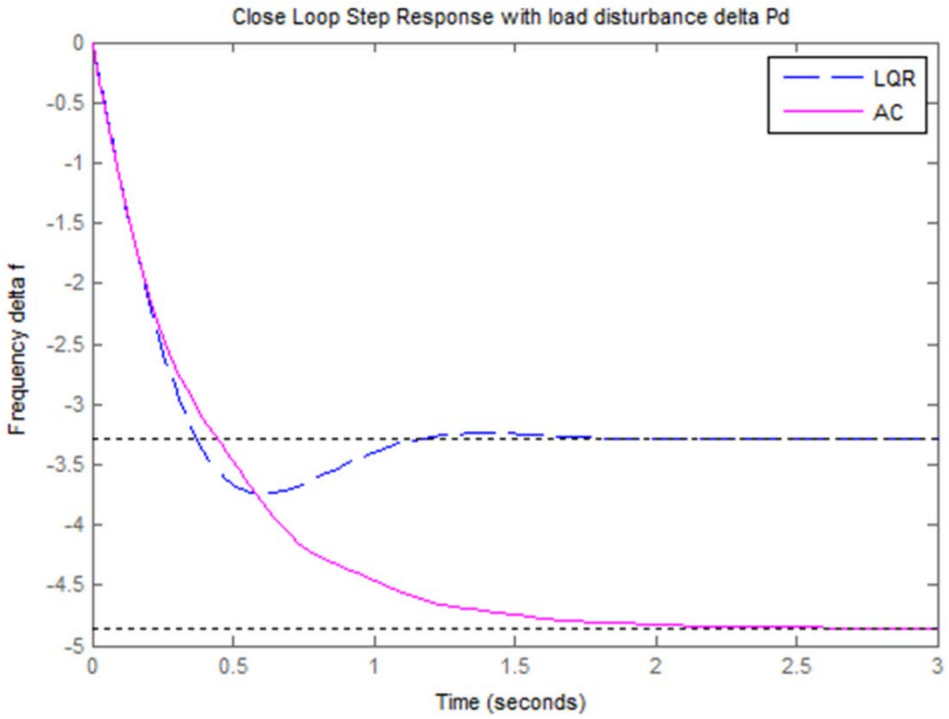


Fig. 6.6 Closed loop step response with load disturbance of LFC system model without integral control.

6.4.1.2 LFC System Model with Integral Control

As discussed in section 4.6.3.2 the state space model of LFC system with integral control is given by

$$A=\begin{bmatrix} -0.0665 & 11.5 & 0 & 0 \\ 0 & -2.5 & 2.5 & 0 \\ -9.5 & 0 & -13.7360 & -13.7360 \\ 0.6 & 0 & 0 & 0 \end{bmatrix},\ B=\begin{bmatrix} 0 \\ 0 \\ 13.7360 \\ 0 \end{bmatrix},\ F=\begin{bmatrix} -11.5 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$C=\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix},\ D=0$$

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0=\begin{bmatrix} 0 & 0.1 & 0 & 0 \end{bmatrix}$, and both the critic and actor NNs weight vectors W_1 and W_2 are initialized with elements of unity values. The learning constants are taken as $a_1=50$, $a_2=10$. The quadratic cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The critic NN activation function vector is selected as

$$\varphi_1(x)=\begin{bmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 & x_2^2 & x_2x_3 & x_2x_4 & x_3^2 & x_3x_4 & x_4^2 \end{bmatrix}^T$$

Let the critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1=W_c=\begin{bmatrix} W_{c1} & W_{c2} & W_{c3} & W_{c4} & W_{c5} & W_{c6} & W_{c7} & W_{c8} & W_{c9} & W_{c10} \end{bmatrix}^T$$

$$\hat{W}_2=W_a=\begin{bmatrix} W_{a1} & W_{a2} & W_{a3} & W_{a4} & W_{a5} & W_{a6} & W_{a7} & W_{a8} & W_{a9} & W_{a10} \end{bmatrix}^T$$

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \le 150$. After 200 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.6925 & 1.0202 & 0.1021 & 0.9960 & 1.0355 & 0.8005 & 0.9952 & 0.0179 & 0.2468 & 1.0107 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 0.6925 & 1.0202 & 0.1021 & 0.9960 & 1.0355 & 0.8005 & 0.9952 & 0.0179 & 0.2468 & 1.0107 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2} R^{-1} \begin{bmatrix} 0 \\ 0 \\ 13.7360 \\ 0 \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 & 0 & 0 \\ x_2 & x_1 & 0 & 0 \\ x_3 & 0 & x_1 & 0 \\ x_4 & 0 & 0 & x_1 \\ 0 & 2x_2 & 0 & 0 \\ 0 & x_3 & x_2 & 0 \\ 0 & x_4 & 0 & x_2 \\ 0 & 0 & 2x_3 & 0 \\ 0 & 0 & x_4 & x_3 \\ 0 & 0 & 0 & 2x_4 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control $u = -R^{-1}B^T Px$.

The unique positive definite solution of ARE denoted here by matrix *RicP* is obtained as

$$RicP = \begin{bmatrix} 0.4600 & 0.6911 & 0.0519 & 0.4642 \\ 0.6911 & 1.8668 & 0.2002 & 0.5800 \\ 0.0519 & 0.2002 & 0.0533 & 0.0302 \\ 0.4642 & 0.5800 & 0.0302 & 2.2106 \end{bmatrix}$$

and from critic NN weights W_c the adaptive optimal critic matrix P of adaptive critic scheme using synchronous PI is obtained as

$$P = \begin{bmatrix} 0.6925 & 0.5101 & 0.0511 & 0.4980 \\ 0.5101 & 1.0355 & 0.4002 & 0.4976 \\ 0.0511 & 0.4002 & 0.0179 & 0.1234 \\ 0.4980 & 0.4976 & 0.2468 & 1.0107 \end{bmatrix}$$

The actor gains of LQR design *RicK*, and actor *K* of adaptive critic scheme using synchronous PI respectively are obtained as

$$RicK = \begin{bmatrix} 0.7135 & 2.7499 & 0.7323 & 0.4142 \end{bmatrix},$$

$$K = \begin{bmatrix} 0.7014 & 5.4975 & 0.2455 & 1.6953 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-8.4497 +10.5385i, -8.4497 -10.5385i, -1.3876 + 1.2549i, -1.3876 - 1.2549i

Fig. 6.7 shows the LFC system state trajectories which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.8 and 6.9. Fig. 6.10 presents the closed loop response of LFC system with integral control model using both approaches of LQR and adaptive critic (AC). It is observed here that there is no steady state error in the closed loop response with unit step load disturbance which is due to including integral control in LFC model. The adaptive optimal controller using synchronous PI based adaptive critic scheme gives faster response than LQR approach.

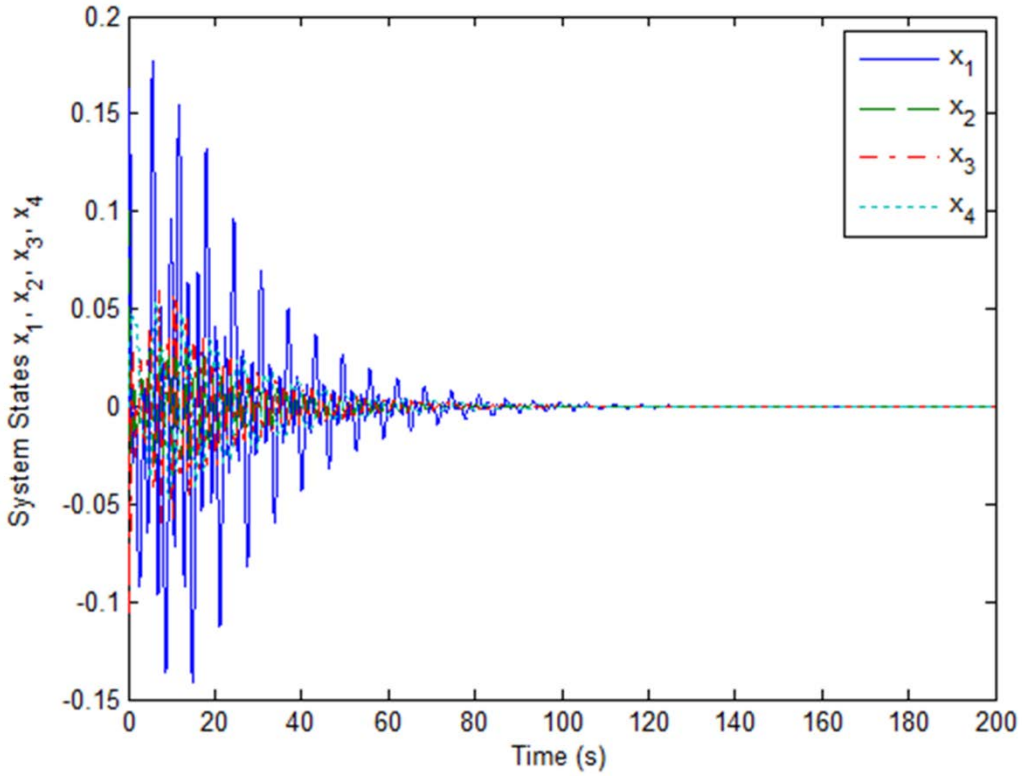


Fig. 6.7 System states trajectories.

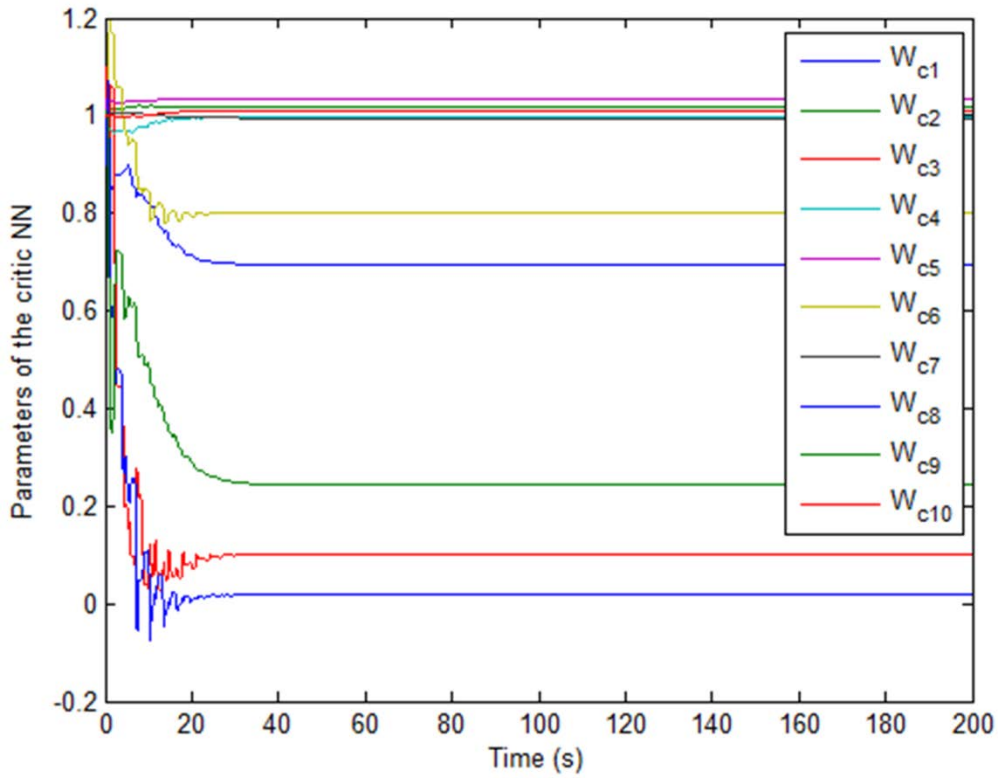


Fig. 6.8 Convergence of critic parameters.

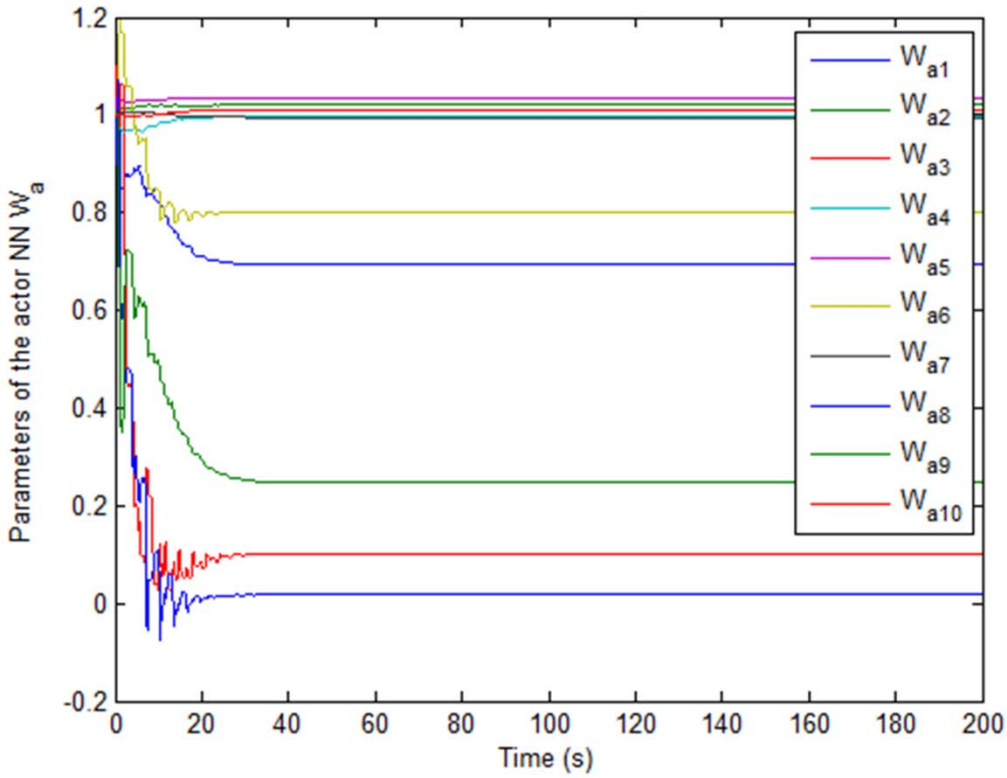


Fig. 6.9 Convergence of actor parameters.

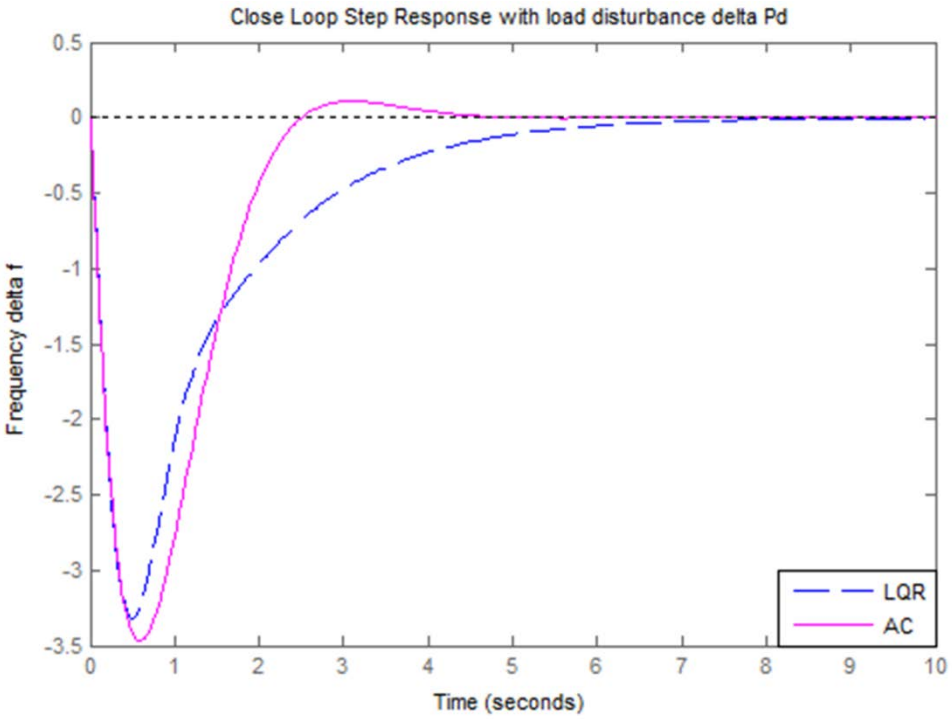


Fig. 6.10 Closed loop step response with load disturbance of LFC system with integral control model.

6.4.2 Automatic Voltage Regulator of Power System

Consider the power system automatic voltage regulator design problem as discussed in section 4.6.4.

6.4.2.1 AVR System Model Neglecting Sensor Dynamics

As discussed in section 4.6.4.1 consider the state space model of AVR system neglecting sensor dynamics given by

$$A=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -275 & -37.5 & -13.5 \end{bmatrix},\quad B=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\quad C=\begin{bmatrix} 250 & 0 & 0 \end{bmatrix},\quad D=\begin{bmatrix} 0 \end{bmatrix}$$

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0=\begin{bmatrix} 0.1 & 0.05 & 0.04 \end{bmatrix}$, and both the critic and actor NNs weight vectors W_1 and W_2 are initialized with elements of unity values. The learning constants are taken as $a_1=50$, $a_2=1$. The quadratic cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The critic NN activation function vector is selected as

$$\varphi_1(x)=\begin{bmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_2^2 & x_2x_3 & x_3^2 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1=W_c=\begin{bmatrix} W_{c1} & W_{c2} & W_{c3} & W_{c4} & W_{c5} & W_{c6} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} & W_{a4} & W_{a5} & W_{a6} \end{bmatrix}^T$$

which are tuned synchronously using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \leq 150$. After 200 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.9962 & 1.0533 & 1.0664 & 1.1610 & -0.3366 & 0.1060 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 0.9962 & 1.0533 & 1.0664 & 1.1610 & -0.3366 & 0.1060 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2} R^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 & 0 \\ x_2 & x_1 & 0 \\ x_3 & 0 & x_1 \\ 0 & 2x_2 & 0 \\ 0 & x_3 & x_2 \\ 0 & 0 & 2x_3 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control $u = -R^{-1}B^T Px$.
The unique positive definite solution of ARE denoted here by matrix *RicP* is obtained as

$$RicP = \begin{bmatrix} 167.9165 & 22.5745 & 0.0018 \\ 22.5745 & 11.3630 & 0.6104 \\ 0.0018 & 0.6104 & 0.0820 \end{bmatrix}$$

and from critic NN weights W_c the adaptive optimal critic matrix *P* of adaptive critic scheme using synchronous PI is obtained as

$$P = \begin{bmatrix} 0.9962 & 0.5267 & 0.5332 \\ 0.5267 & 1.1610 & -0.1683 \\ 0.5332 & -0.1683 & 0.1060 \end{bmatrix}$$

The actor gains of LQR design *RicK*, and actor *K* of adaptive critic scheme using synchronous PI respectively are obtained as

$$RicK = \begin{bmatrix} 0.0018 & 0.6104 & 0.0820 \end{bmatrix} \quad , \quad K = \begin{bmatrix} 0.5332 & -0.1683 & 0.1060 \end{bmatrix}$$

The eigenvalues of closed loop system are obtained as

-12.3879, -0.6090 + 4.6767i, -0.6090 - 4.6767i

Fig. 6.11 shows the state trajectories of AVR system neglecting sensor dynamics which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.12 and 6.13. Fig. 6.14 presents the closed loop response of AVR system neglecting sensor dynamics model using both approaches of LQR and adaptive critic (AC) by replacing $u = -Kx + r$ in state equation,

where $r = V_{ref}$ is a unit step input; and *K* is actor gains *RicK* and *K* respectively. It is observed that the adaptive optimal controller using synchronous PI based adaptive critic scheme gives the similar response as of LQR approach.

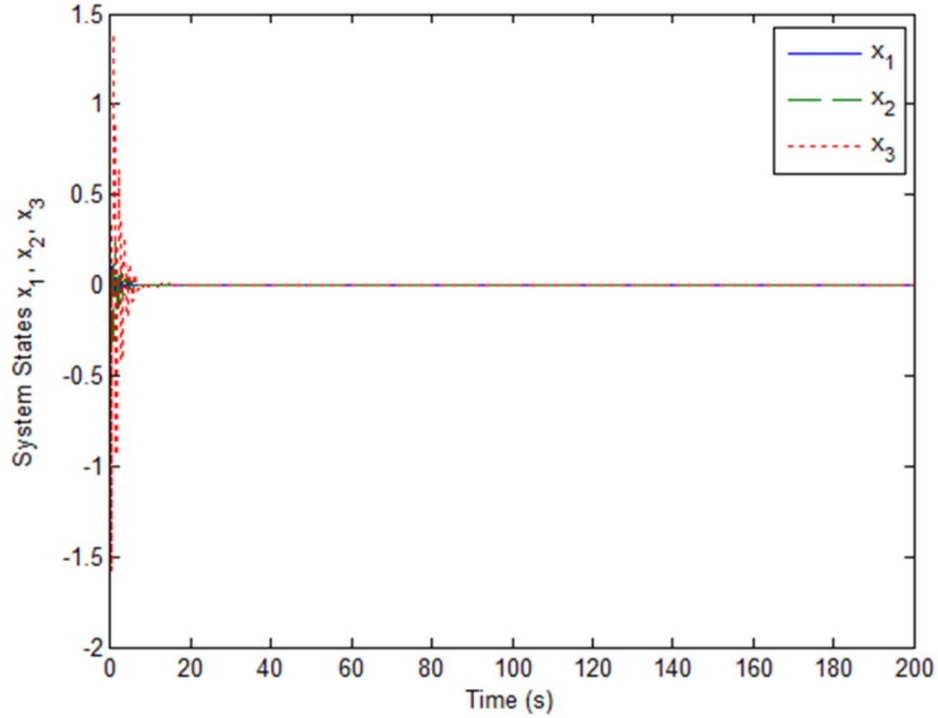


Fig. 6.11 System states trajectories.

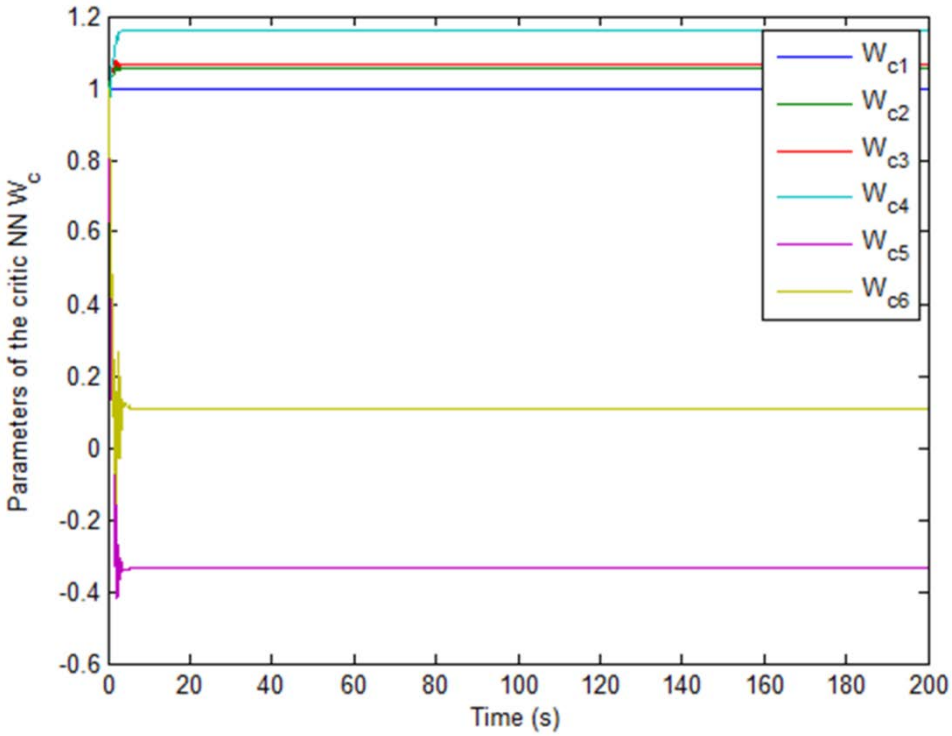


Fig. 6.12 Convergence of critic parameters.

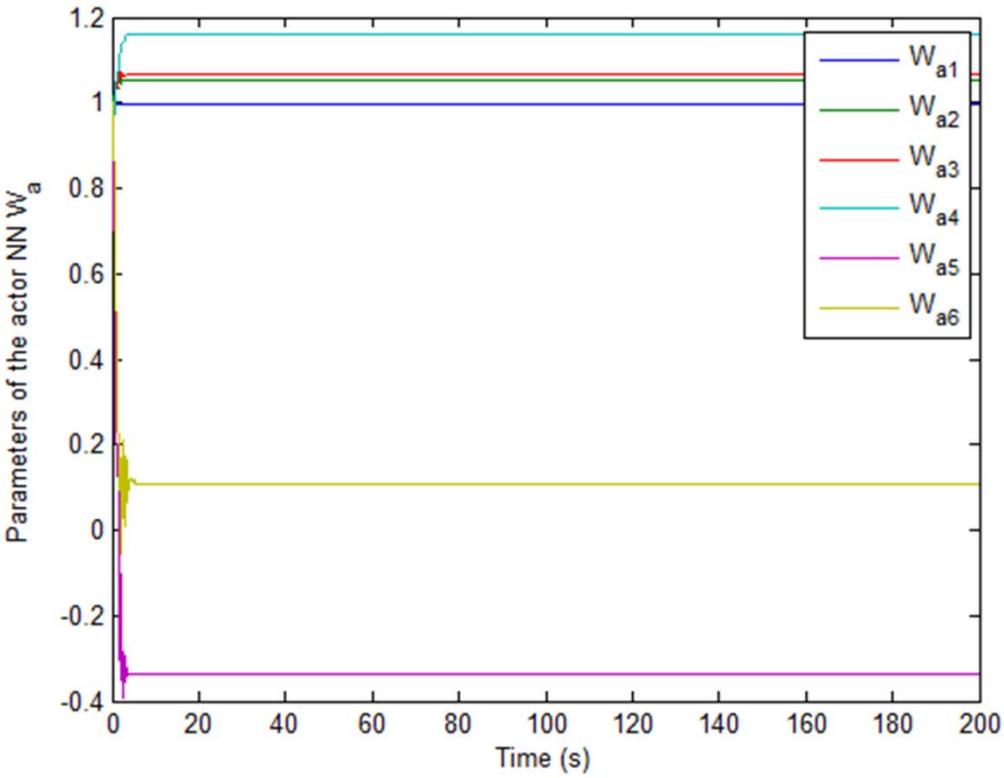


Fig. 6.13 Convergence of actor parameters.

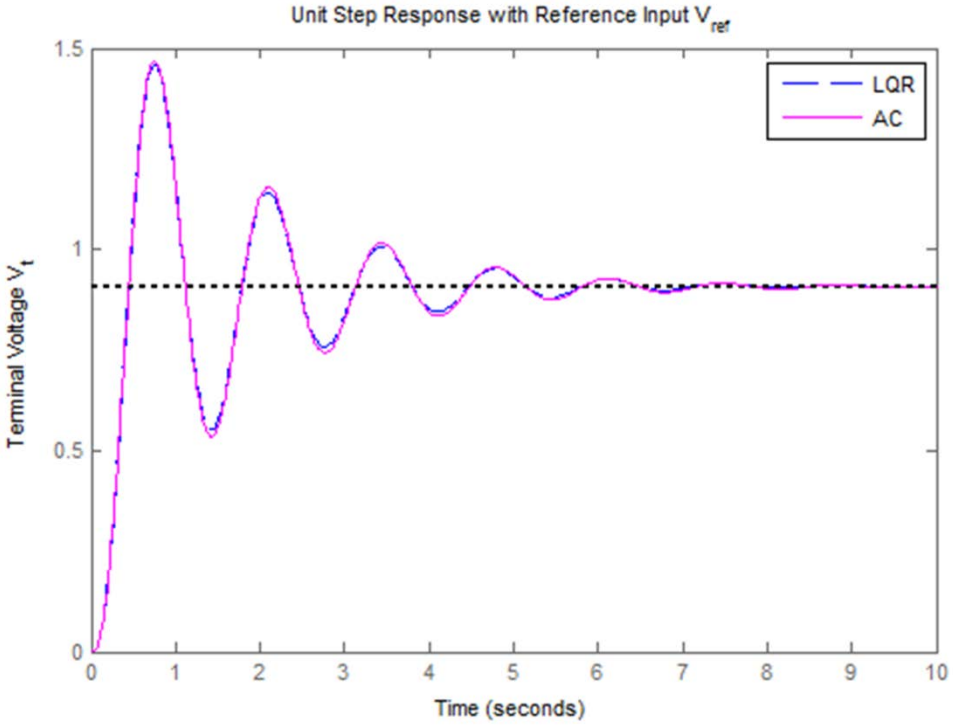


Fig. 6.14 Closed loop step response with V_{ref} of AVR system model neglecting sensor dynamics.

6.4.2.2 AVR System Model Including Sensor Dynamics

As discussed in section 4.6.4.2 consider the state space model of AVR system including sensor dynamics given by

$$A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -27500 & -3775 & -1387.5 & -113.5 \end{bmatrix} \quad , \quad B=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad ,$$

$$C=\begin{bmatrix} 25000 & 250 & 0 & 0 \end{bmatrix} \ , \ D=\begin{bmatrix} 0 \end{bmatrix}$$

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0=\begin{bmatrix} 0.1 & 0.05 & 0.04 & 0.02 \end{bmatrix}$, and both the critic and actor NNs weight vectors W_1 and W_2 are initialized with elements of unity values. The learning constants are taken as $a_1=50$, $a_2=10$. The quadratic cost function parameters *Q* and *R* are taken as identity matrices of appropriate dimensions. The critic NN activation function vector is selected as

$$\varphi_1(x)=\begin{bmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 & x_2^2 & x_2x_3 & x_2x_4 & x_3^2 & x_3x_4 & x_4^2 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1=W_c=\begin{bmatrix} W_{c1} & W_{c2} & W_{c3} & W_{c4} & W_{c5} & W_{c6} & W_{c7} & W_{c8} & W_{c9} & W_{c10} \end{bmatrix}^T$$

$$\hat{W}_2=W_a=\begin{bmatrix} W_{a1} & W_{a2} & W_{a3} & W_{a4} & W_{a5} & W_{a6} & W_{a7} & W_{a8} & W_{a9} & W_{a10} \end{bmatrix}^T$$

which are tuned synchronously using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \le 150$. After 200 seconds the critic NN and actor NN weights are converged respectively to

$$W_c=\begin{bmatrix} 1.0006 & 0.9981 & 0.9884 & 1.0540 & 0.9937 & 1.0477 & 1.0957 & 1.2055 & -0.2882 & 0.0922 \end{bmatrix}^T$$

$$W_a=\begin{bmatrix} 1.0006 & 0.9981 & 0.9884 & 1.0540 & 0.9937 & 1.0477 & 1.0957 & 1.2055 & -0.2882 & 0.0922 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x)=-\frac{1}{2}R^{-1}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 & 0 & 0 \\ x_2 & x_1 & 0 & 0 \\ x_3 & 0 & x_1 & 0 \\ x_4 & 0 & 0 & x_1 \\ 0 & 2x_2 & 0 & 0 \\ 0 & x_3 & x_2 & 0 \\ 0 & x_4 & 0 & x_2 \\ 0 & 0 & 2x_3 & 0 \\ 0 & 0 & x_4 & x_3 \\ 0 & 0 & 0 & 2x_4 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control $u = -R^{-1}B^T Px$.

The unique positive definite solution of ARE denoted here by matrix *RicP* is obtained as

$$RicP = 1.0 \times 10^3 \begin{bmatrix} 6.5160 & 0.8940 & 0.1290 & 0.0000 \\ 0.8940 & 0.3225 & 0.0446 & 0.0002 \\ 0.1290 & 0.0446 & 0.0100 & 0.0000 \\ 0.0000 & 0.0002 & 0.0000 & 0.0000 \end{bmatrix}$$

and from critic NN weights W_c the adaptive optimal critic matrix *P* of adaptive critic scheme using synchronous PI is obtained as

$$P = \begin{bmatrix} 1.0006 & 0.4991 & 0.4942 & 0.5270 \\ 0.4991 & 0.9937 & 0.5239 & 0.5478 \\ 0.4942 & 0.5239 & 1.2055 & -0.1441 \\ 0.5270 & 0.5478 & -0.2882 & 0.0922 \end{bmatrix}$$

The actor gains of LQR design *RicK*, and actor *K* of adaptive critic scheme using synchronous PI respectively are obtained as

$$RicK = \begin{bmatrix} 0.0000 & 0.2369 & 0.0325 & 0.0047 \end{bmatrix} ,$$

$$K = \begin{bmatrix} 0.5270 & 0.5478 & -0.2882 & 0.0922 \end{bmatrix}$$

The eigenvalues of closed loop system in this case are obtained as

1.0e+002 *[-1.0008, -0.1247, -0.0052 + 0.0466i, -0.0052 - 0.0466i]

Fig. 6.15 shows the state trajectories of AVR system including sensor dynamics which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.16 and 6.17. Fig. 6.18 presents the closed loop response of AVR system including sensor dynamics model using both approaches of LQR and adaptive critic (AC) by replacing $u = -Kx + r$ in state equation, where $r = V_{ref}$ is a unit step input; and *K* is actor gains *RicK* and *K* respectively. It is observed in this case also that the adaptive optimal controller using synchronous PI based adaptive critic scheme gives the similar response as of LQR approach.

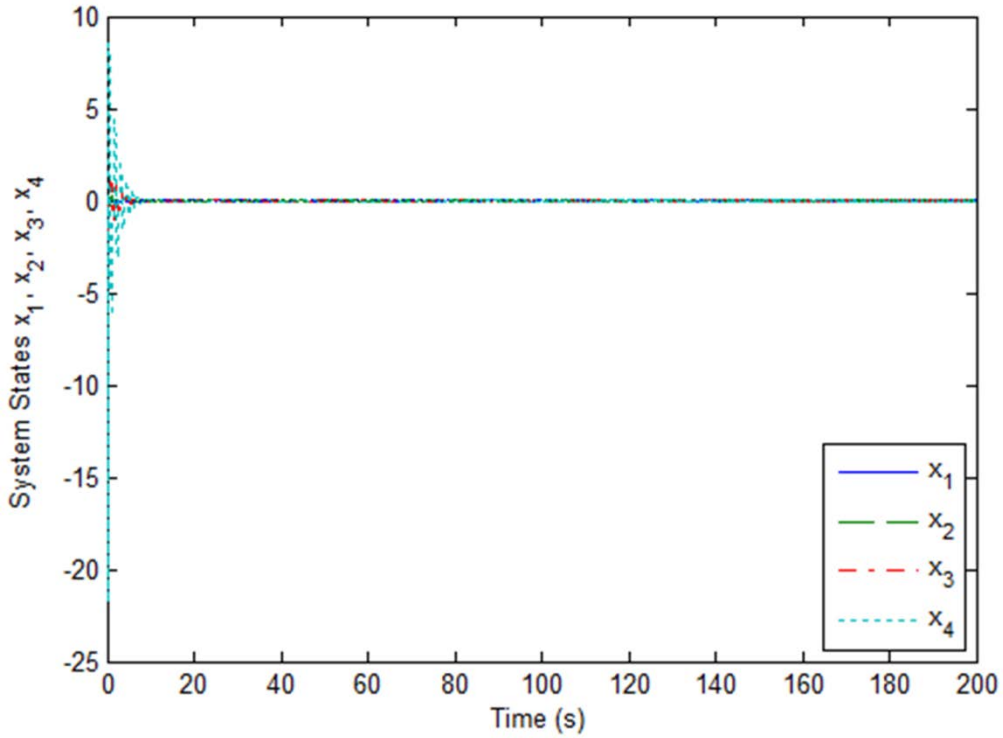


Fig. 6.15 System states trajectories.

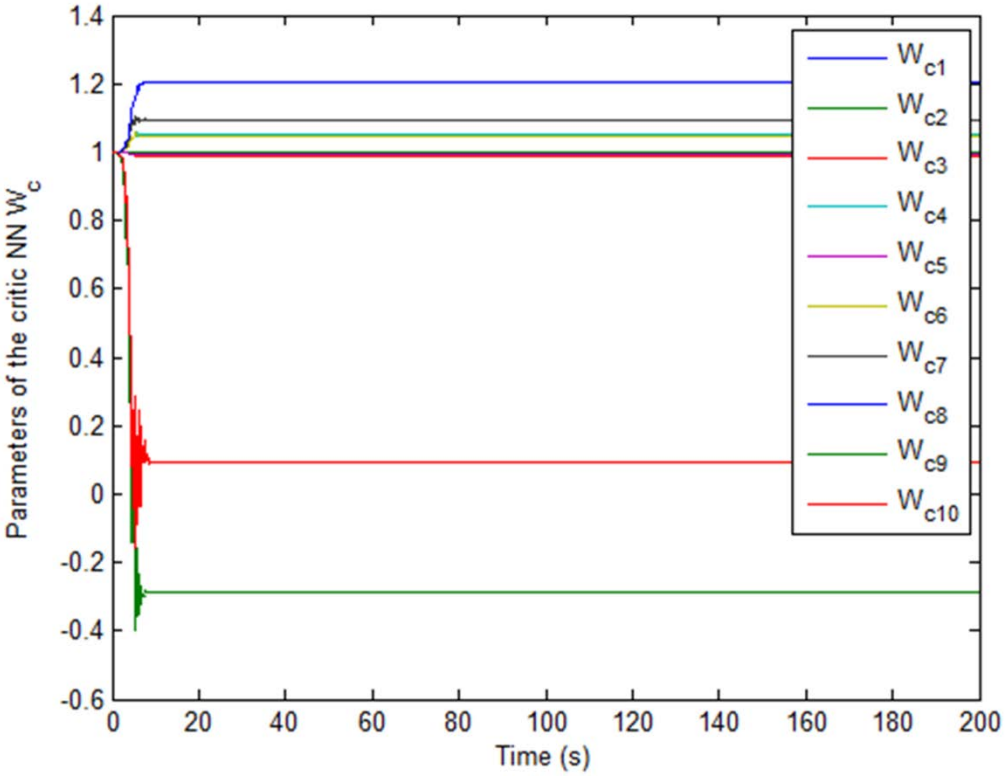


Fig. 6.16 Convergence of critic parameters.

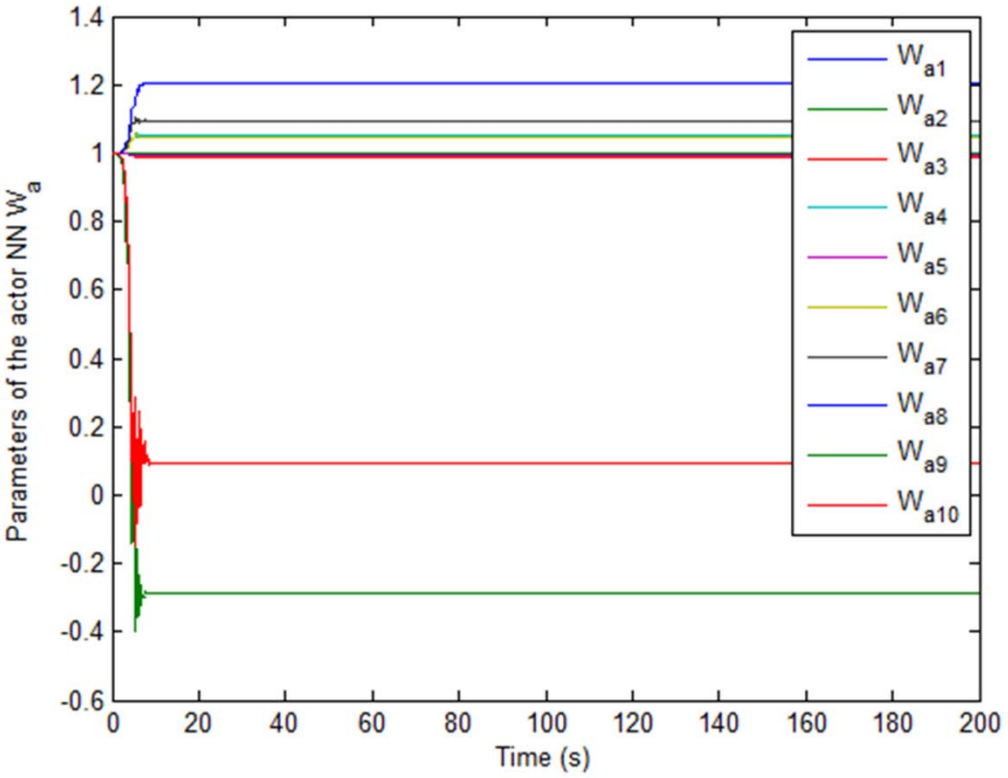


Fig. 6.17 Convergence of actor parameters.

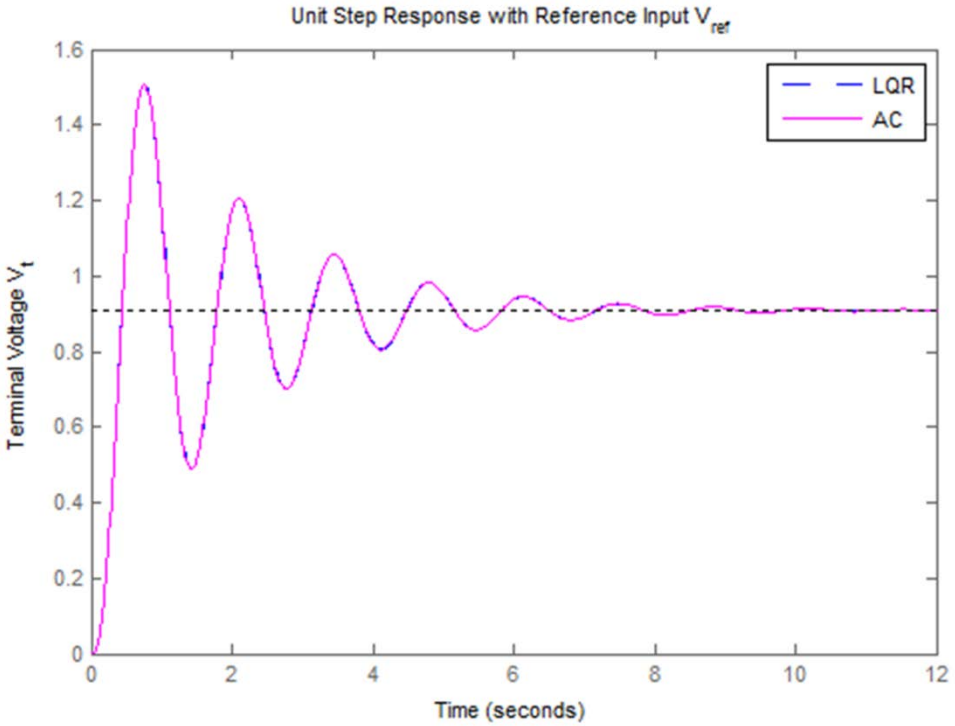


Fig. 6.18 Closed loop step response with V_{ref} of AVR system model including sensor dynamics.

6.5 SIMULATION RESULTS AND ANALYSIS FOR AFFINE NONLINEAR SYSTEMS

In this section, the application of adaptive critic control scheme using online synchronous PI technique with neural networks approximation of cost function and control policy is implemented for five examples of continuous-time time-invariant affine nonlinear systems- three general systems, and two practical systems- single-link manipulator, and Vander Pol oscillator. The simulation results and performance analysis are presented to demonstrate the effectiveness of the control scheme. The infinite horizon cost function as defined in section 5.2 is considered for the optimal control problems of these examples of affine nonlinear systems. The infinite horizon cost function $V = \int_0^\infty L(x(\tau), u(\tau))d\tau$, where $L(x,u) = Q(x) + u^T Ru$, is minimized for optimal control solution.

For the implementation of synchronous PI technique, the critic NN activation functions vector $\varphi_1(x)$ is considered as the Kronecker product quadratic polynomial basis vector. For convergence of critic NN and actor NN, the PE condition is ensured by adding a small probing noise to control signal. The probing noise is turned off after convergence is reached. The probing noise is considered as

$$n(t) = e^{-0.05t}\sin^2(t)\cos(t)$$

6.5.1 General Affine Nonlinear System Example 1

Consider a general continuous-time input-affine nonlinear system having stronger nonlinearities described by the following equations

$$\dot{x}_1 = -x_1 + x_2 \tag{6.38}$$

$$\dot{x}_2 = f(x) + g(x)u \tag{6.39}$$

where $f(x) = -x_1^3 - x_2^3 + \frac{1}{4}x_2(2+\cos(4x_1))^2$, and $g(x) = 2 + \cos(4x_1)$.

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0 = \begin{bmatrix} 1 & -1 \end{bmatrix}$, and both the critic NN and actor NN weight vectors are randomly initialized with each element respectively as $W_{10} = 2*rand$, $W_{20} = 3*rand$. The learning constants are taken as $a_1 = 50$, $a_2 = 1$. The infinite horizon cost function is considered as by defining $Q(x) = x_1^2 + x_2^2 + x_1^4 + x_2^4$, and $R = 1$.

The optimal value function is

$$V^*(x) = \frac{1}{4}x_1^4 + \frac{1}{2}x_2^2$$

and the optimal control signal is

$$u^*(x) = -\frac{1}{2}(2+\cos(4x_1)x_2$$

Consider the critic NN activation function vector as

$$\varphi_1(x) = \begin{bmatrix} x_1^2 & x_2^2 & x_1^4 & x_2^4 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1 = W_c = \begin{bmatrix} W_{c1} & W_{c2} & W_{c3} & W_{c4} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} & W_{a4} \end{bmatrix}^T$$

which are tuned at the same time using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \le 80$. After 100 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.3571 & 0.7007 & 0.8986 & 0.1634 \end{bmatrix}$$

$$W_a = \begin{bmatrix} 0.3571 & 0.7007 & 0.8986 & 0.1634 \end{bmatrix}$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2}R^{-1}\begin{bmatrix} 0 \\ 2+\cos(4x_1) \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 \\ 0 & 2x_2 \\ 4x_1^3 & 0 \\ 0 & 4x_2^3 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control policy.

Fig. 6.19 shows the state trajectories of nonlinear system (6.38) and (6.39) which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.20 and 6.21. Fig. 6.22 shows the 3-D mesh-plot of optimal value function. Fig. 6.23 shows the 3-D mesh-plot of neural network approximated value function using online synchronous PI technique. It is observed that the neural network approximated value function by online tuning using synchronous PI is similar to the optimal value function.

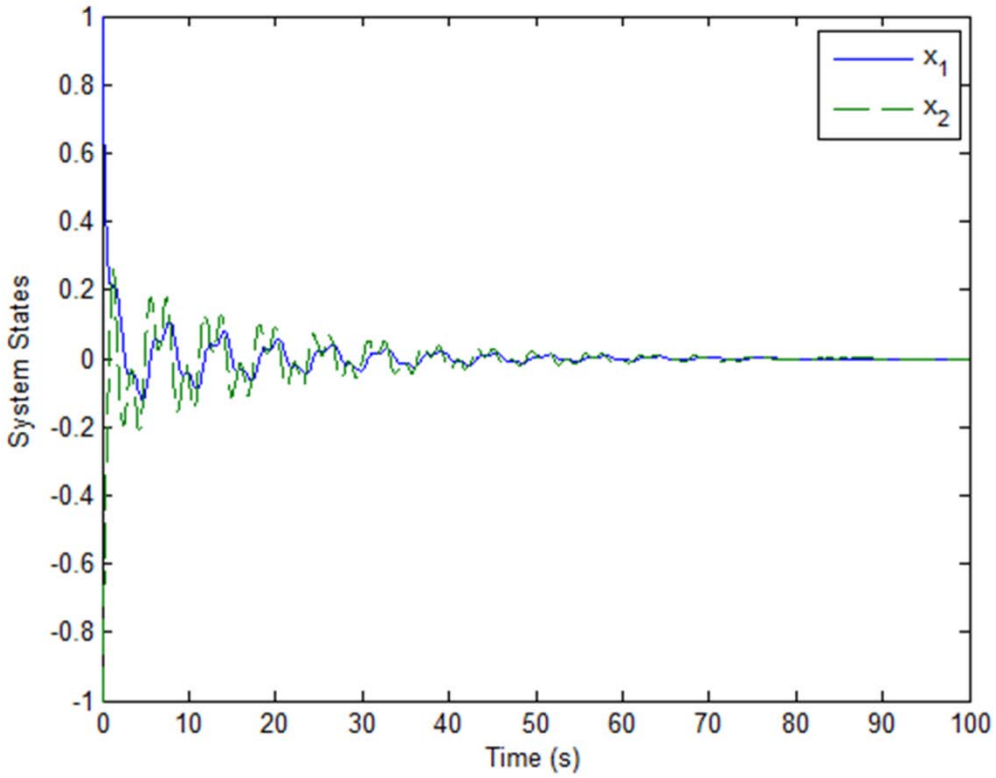


Fig. 6.19 System states trajectories.

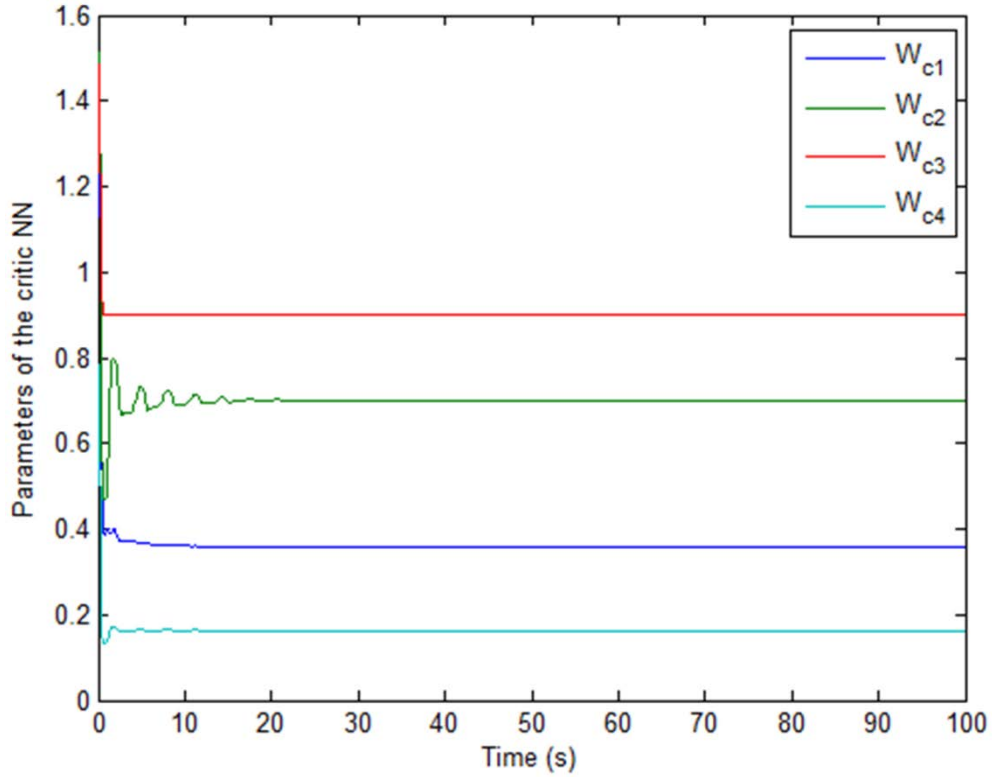


Fig. 6.20 Convergence of critic parameters.

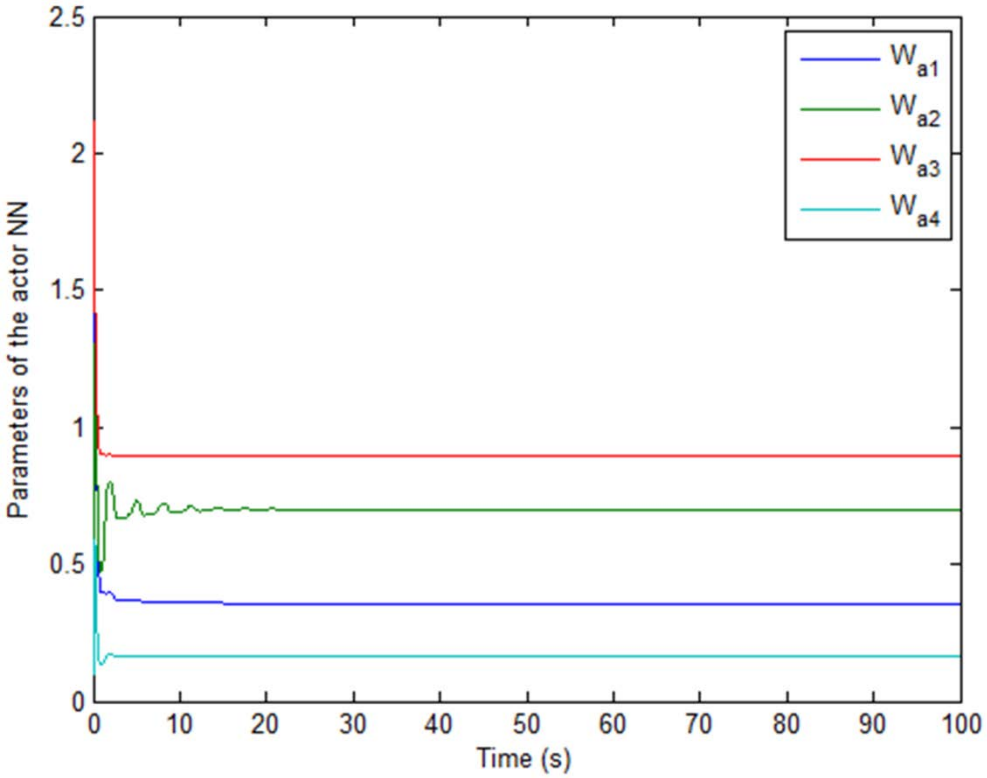


Fig. 6.21 Convergence of actor parameters.

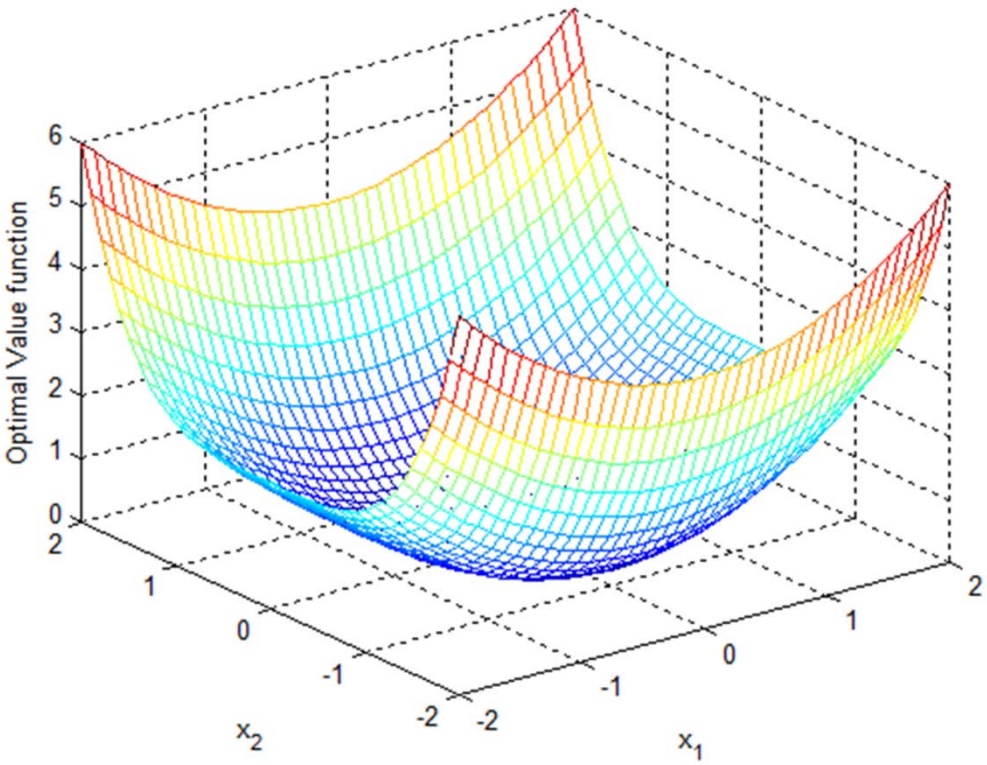


Fig. 6.22 3-D plot of optimal value function V*.

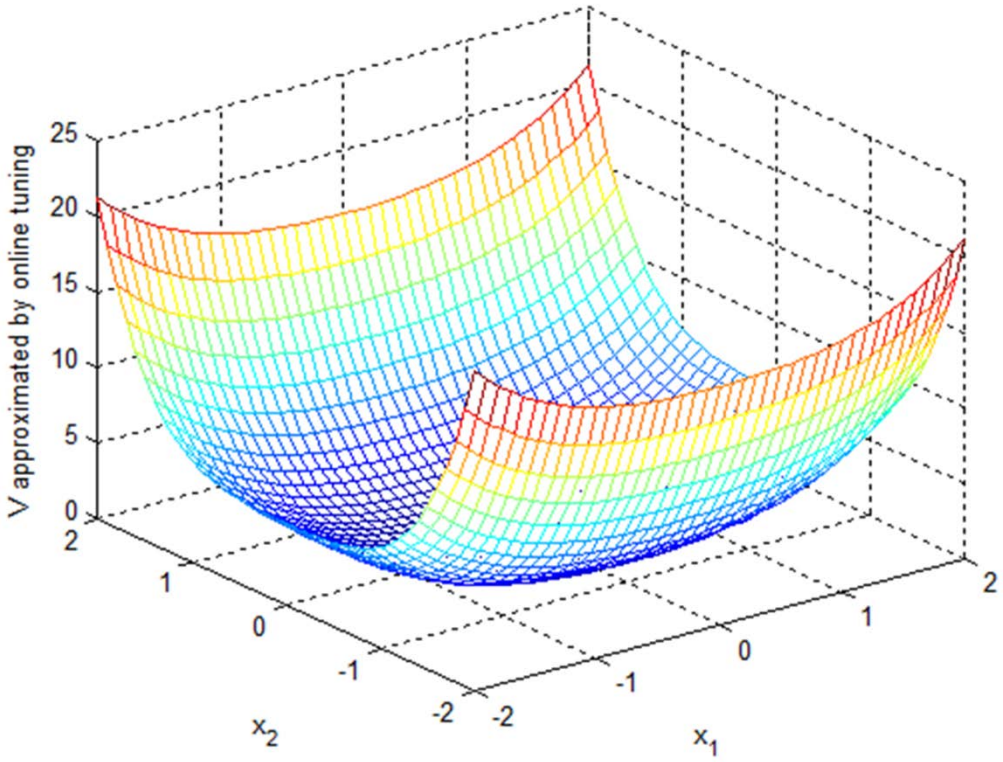


Fig. 6.23 3-D plot of neural network approximated value function by online tuning.

6.5.2 General Affine Nonlinear System Example 2

Consider a general continuous-time input-affine nonlinear system which is having stronger nonlinearities and quadratic cost function. This nonlinear system is described by the following equations

$$\dot{x}_1 = -x_1 + x_2 \tag{6.40}$$

$$\dot{x}_2 = f(x) + g(x)u \tag{6.41}$$

where $f(x) = -\frac{1}{2}(x_1 + x_2) + \frac{1}{2}x_2(2 + \sin(2x_1))^2$, and $g(x) = 2 + \sin(2x_1)$.

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0 = \begin{bmatrix} 1 & -1 \end{bmatrix}$, and both the critic NN and actor NN weight vectors are randomly initialized with each element respectively as $W_{10} = 2*rand$, $W_{20} = 3*rand$. The learning constants are taken as $a_1 = 5$, $a_2 = 1$. The infinite horizon cost function is considered as by defining $Q(x) = x_1^2 + x_2^2$, and $R = 1$.

The optimal value function is

$$V^*(x) = \frac{1}{2}x_1^2 + x_2^2$$

and the optimal control signal is

$$u^*(x) = -\frac{1}{2}(2 + \sin(2x_1)x_2$$

Consider the critic NN activation function vector as

$$\varphi_1(x) = \begin{bmatrix} x_1^2 & x_1x_2 & x_2^2 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1 = W_c = \begin{bmatrix} W_{c1} & W_{c2} & W_{c3} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} \end{bmatrix}^T$$

which are tuned at the same time using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \leq 80$. After 100 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.7151 & 0.1091 & 0.8276 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 0.7151 & 0.1091 & 0.8276 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2}R^{-1}\begin{bmatrix} 0 \\ 2 + \sin(2x_1) \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control policy.

Fig. 6.24 shows the state trajectories of nonlinear system (6.40) and (6.41) which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.25 and 6.26. Fig. 6.27 shows the 3-D mesh-plot of optimal value function. Fig. 6.28 shows the 3-D mesh-plot of neural network approximated value function using online synchronous PI technique. Also for this example of nonlinear system, it is observed that the neural network approximated value function by online tuning using synchronous PI is similar to the optimal value function.

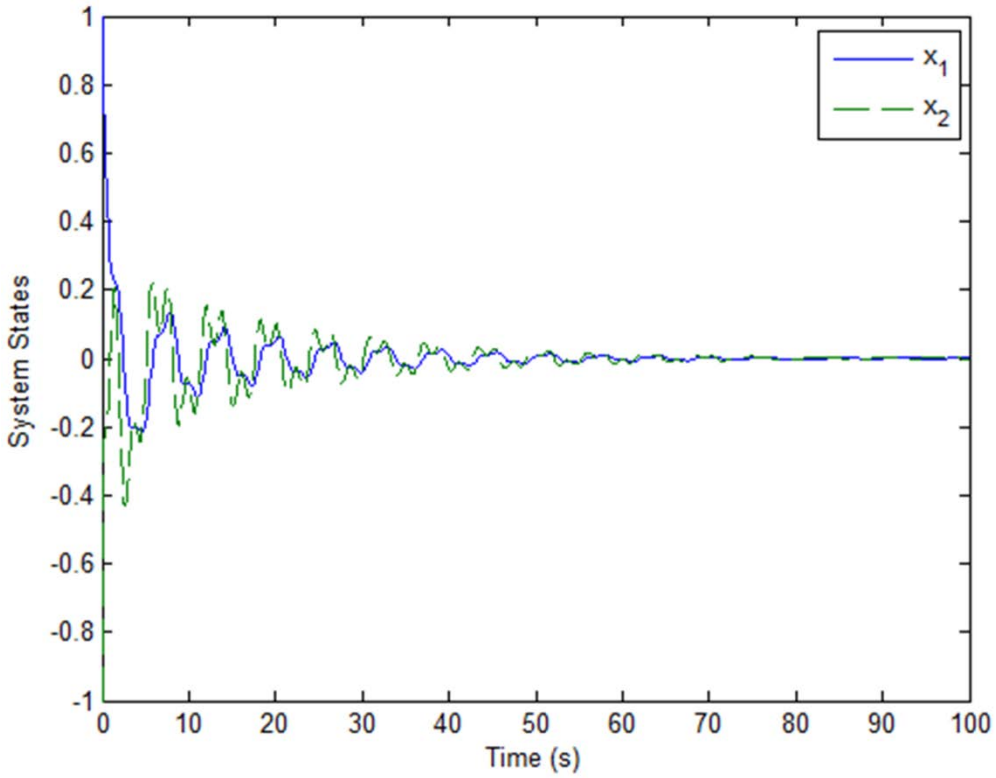


Fig. 6.24 System states trajectories.

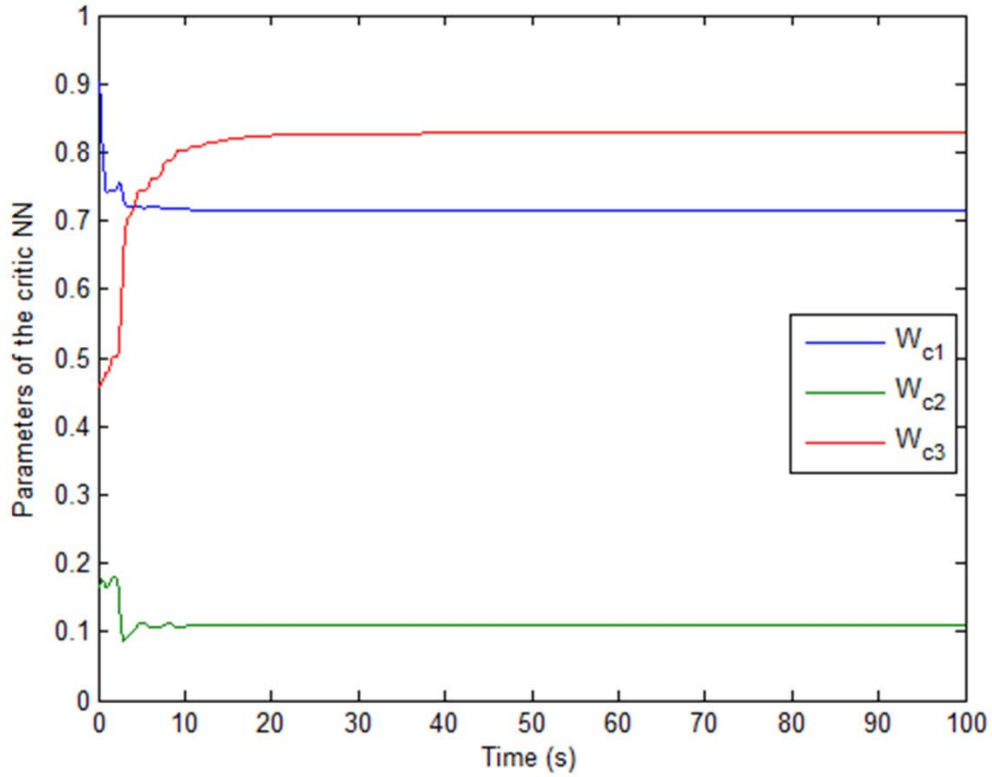


Fig. 6.25 Convergence of critic parameters.

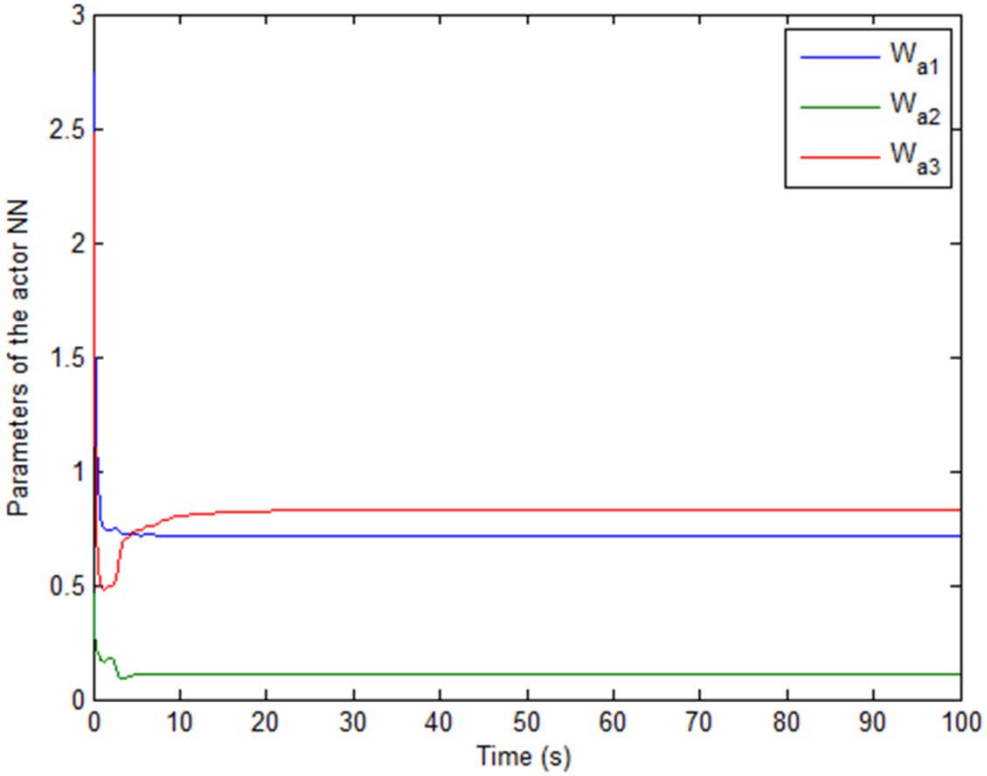


Fig. 6.26 Convergence of actor parameters.

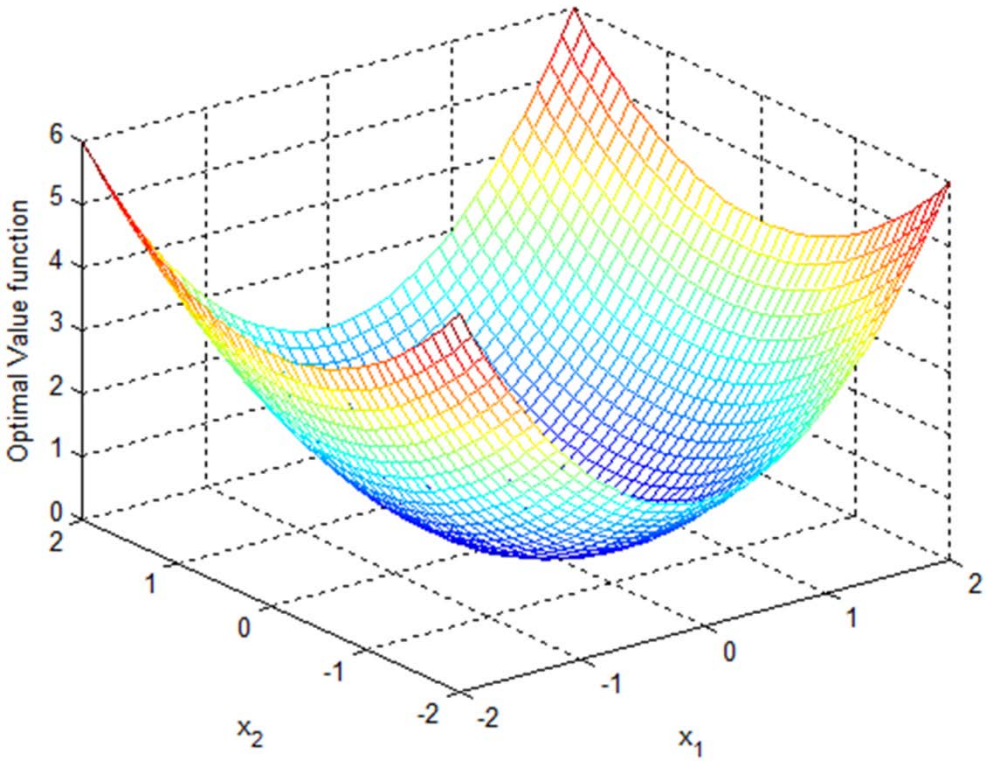


Fig. 6.27 3-D plot of optimal value function V*.

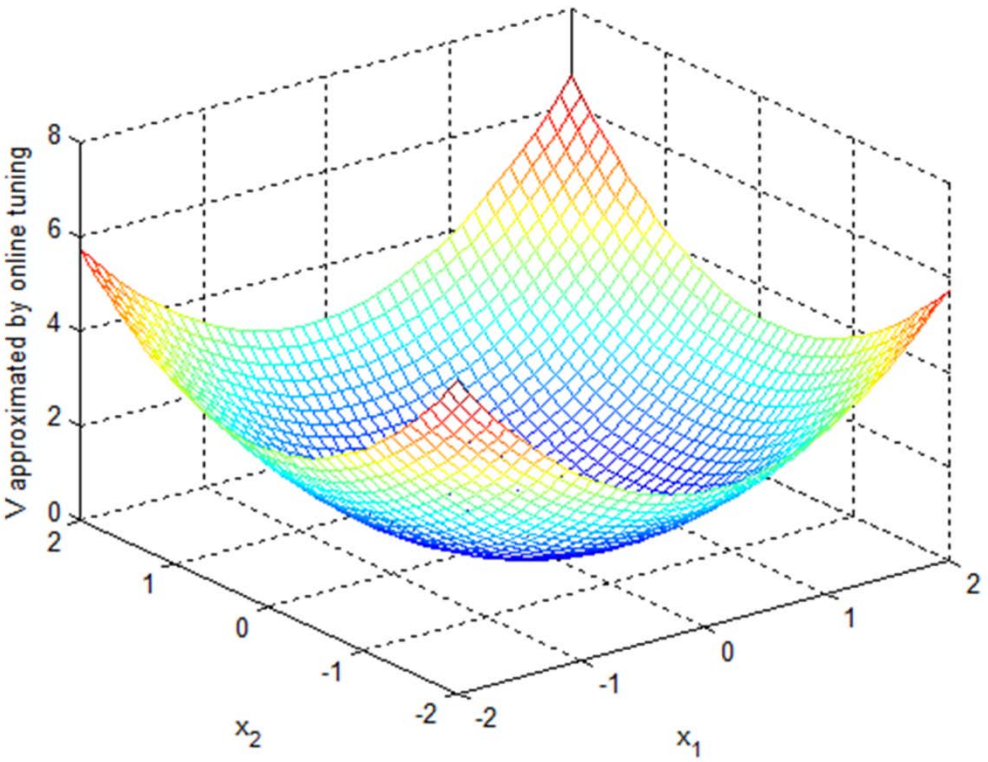


Fig. 6.28 3-D plot of neural network approximated value function by online tuning.

6.5.3 General Affine Nonlinear System Example 3

Consider a general continuous-time input-affine nonlinear system having stronger nonlinearities and quartic cost function. This nonlinear system is described by the following equations [160]

$$\dot{x}_1 = -x_1 + x_2 + 2x_2^3 \tag{6.42}$$

$$\dot{x}_2 = f(x) + g(x)u \tag{6.43}$$

where $f(x) = -\frac{1}{2}(x_1 + x_2) + \frac{1}{2}x_2(1 + 2x_2^2)\sin^2(x_1)$, and $g(x) = \sin(x_1)$.

For implementation of synchronous PI algorithm for this nonlinear system, the initial conditions of states are taken as $x_0 = \begin{bmatrix} 1 & -1 \end{bmatrix}$, and both the critic NN and actor NN weight vectors are randomly initialized with each element respectively as $W_{10} = 2 * rand$, $W_{20} = 3 * rand$. The learning constants are taken as $a_1 = 50$, $a_2 = 1$. The infinite horizon cost function is considered as by defining $Q(x) = x_1^2 + x_2^2 + 2x_2^4$, and $R = 1$.

The optimal value function is

$$V^*(x) = \frac{1}{2}x_1^2 + x_2^2 + x_2^4$$

and the optimal control signal is

$$u^*(x) = -\sin(x_1)(x_2 + 2x_2^3)$$

Consider the critic NN activation function vector as

$$\varphi_1(x) = \begin{bmatrix} x_1^2 & x_1x_2 & x_2^2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1 = W_c = \begin{bmatrix} W_{c1} & W_{c2} & W_{c3} & W_{c4} & W_{c5} & W_{c6} & W_{c7} & W_{c8} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} & W_{a4} & W_{a5} & W_{a6} & W_{a7} & W_{a8} \end{bmatrix}^T$$

which are tuned at the same time using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \le 80$. After 100 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.4021 & -0.0522 & 0.6481 & 0.3281 & 0.6601 & 0.8436 & -0.1613 & 0.2469 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 0.4021 & -0.0522 & 0.6481 & 0.3281 & 0.6601 & 0.8436 & -0.1613 & 0.2469 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2}R^{-1}\begin{bmatrix} 0 \\ \sin(x_1) \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \\ 4x_1^3 & 0 \\ 3x_1^2x_2 & x_1^3 \\ 2x_1x_2^2 & 2x_2x_1^2 \\ x_2^3 & 3x_1x_2^2 \\ 0 & 4x_2^3 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control policy.

Fig. 6.29 shows the state trajectories of nonlinear system (6.42) and (6.43) which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.30 and 6.31. Fig. 6.32 shows the 3-D mesh-plot of optimal value function. Fig. 6.33 shows the 3-D mesh-plot of neural network approximated value function using online synchronous PI technique. For this example of nonlinear system, it is observed that the neural network approximated value function by online tuning using synchronous PI is somewhat similar to the optimal value function.

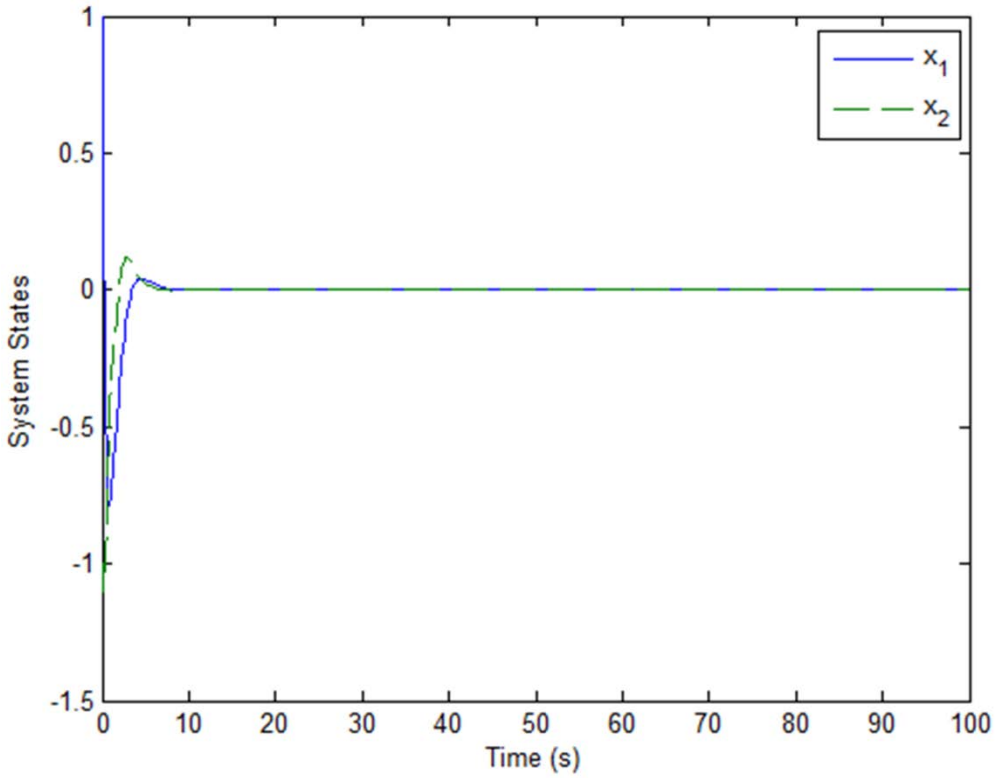


Fig. 6.29 System states trajectories.

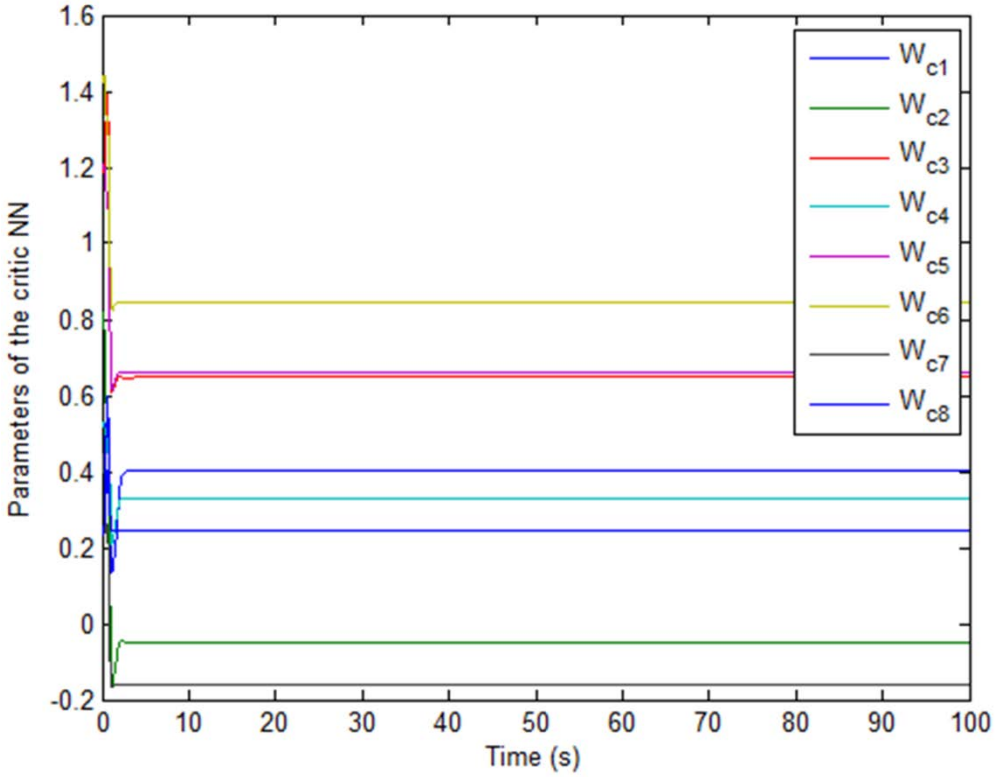


Fig. 6.30 Convergence of critic parameters.

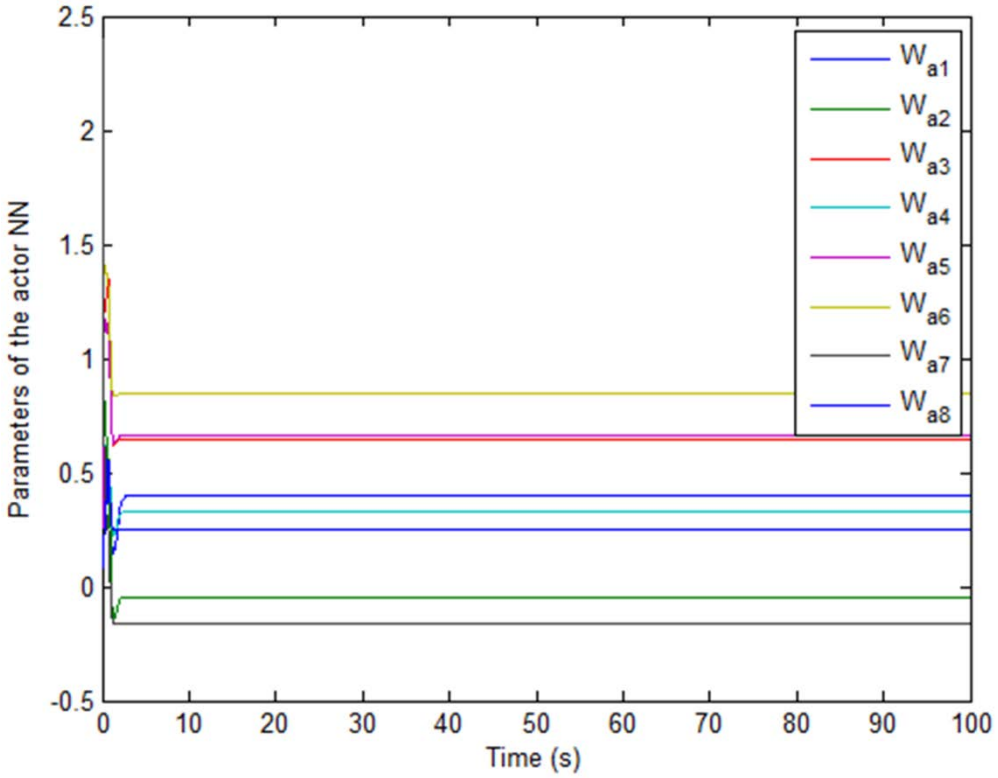


Fig. 6.31 Convergence of actor parameters.

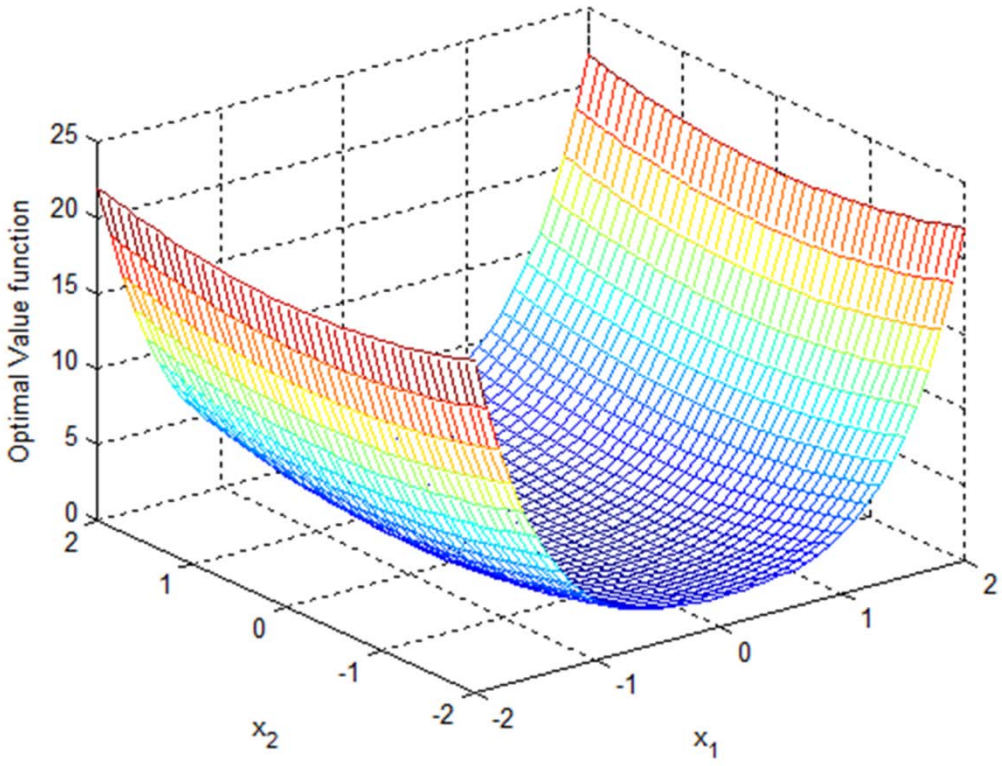


Fig. 6.32 3-D plot of optimal value function V*.

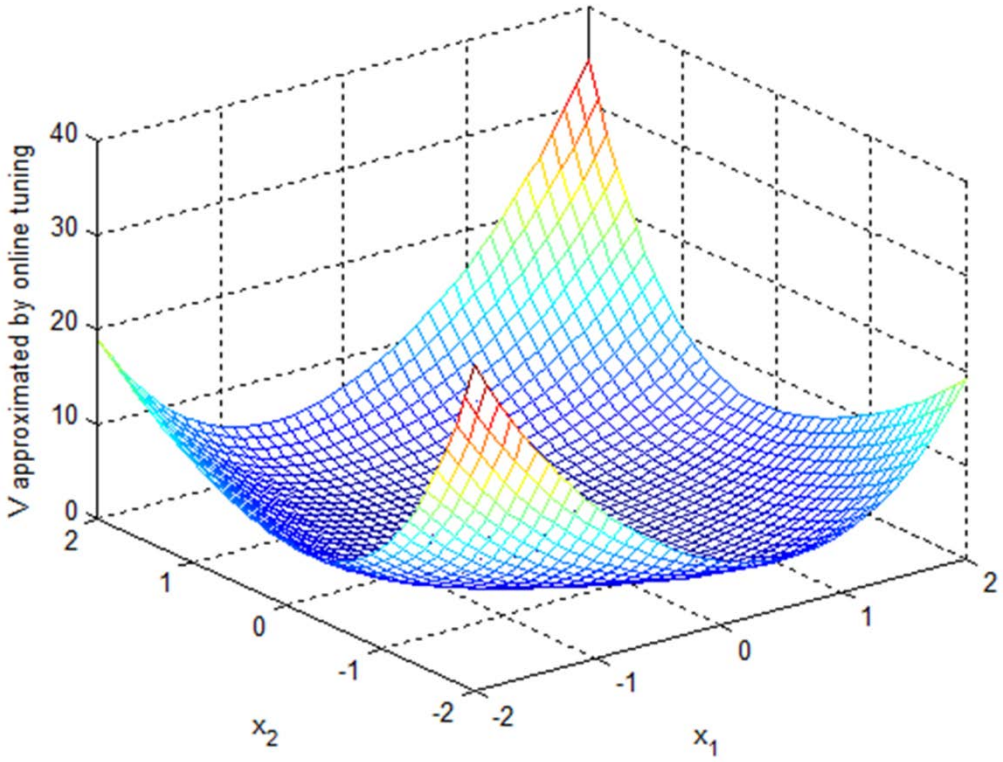


Fig. 6.33 3-D plot of neural network approximated value function by online tuning.

6.5.4 Single-Link Manipulator System

Consider the single-link manipulator system described by dynamic equations [71]

$$\dot{x}_1 = x_2 \tag{6.44}$$

$$\dot{x}_2 = -g\sin(x_1) + u \tag{6.45}$$

where $x_1 = \theta$ is angular displacement, $x_2 = \omega$ is angular velocity, $u = T_a$ is the actuating torque, and g is the acceleration due to gravity. The system's physical parameters are considered as link length *l*=1m, and link mass *m*=1kg.

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0 = \begin{bmatrix} \pi & -\pi \end{bmatrix}$, and both the critic NN and actor NN weight vectors are randomly initialized with each element respectively as $W_{10} = 2*rand$, $W_{20} = 3*rand$. The learning constants are taken as $a_1 = 2$, $a_2 = 1$. The infinite horizon quadratic cost function is considered as by defining $Q(x) = x_1^2 + x_2^2$, and $R = 1$.

The optimal value function is

$$V^*(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

and the optimal control signal is

$$u^*(x) = -\frac{1}{2}\sin(x_1)x_2$$

Consider the critic NN activation function vector as

$$\varphi_1(x) = \begin{bmatrix} x_1^2 & x_1 x_2 & x_2^2 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1 = W_c = \begin{bmatrix} W_{c1} & W_{c2} & W_{c3} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} \end{bmatrix}^T$$

which are tuned at the same time using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \leq 80$. After 100 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 1.1852 & 0.0044 & 0.9833 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 1.1852 & 0.0044 & 0.9833 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2} R^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control policy.

Fig. 6.34 shows the state trajectories of nonlinear system (6.44) and (6.45) which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.35 and 6.36. Fig. 6.37 shows the 3-D mesh-plot of optimal value function. Fig. 6.38 shows the 3-D mesh-plot of neural network approximated value function using online synchronous PI technique. It is observed that the neural network approximated value function by online tuning using synchronous PI is similar to the optimal value function.

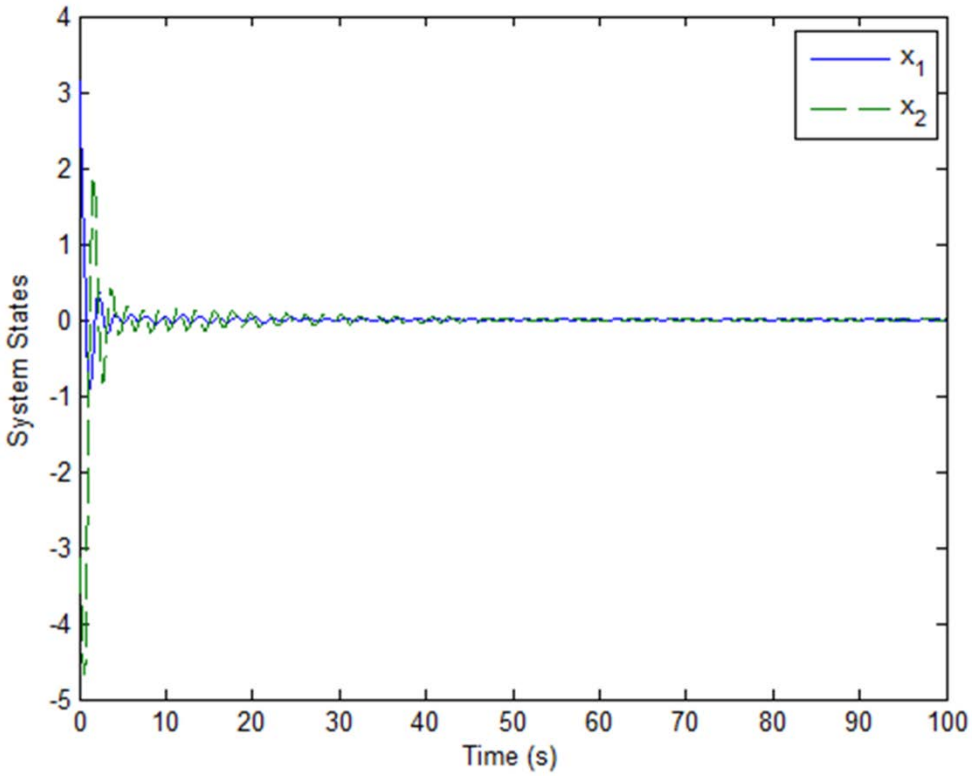


Fig. 6.34 System states trajectories.

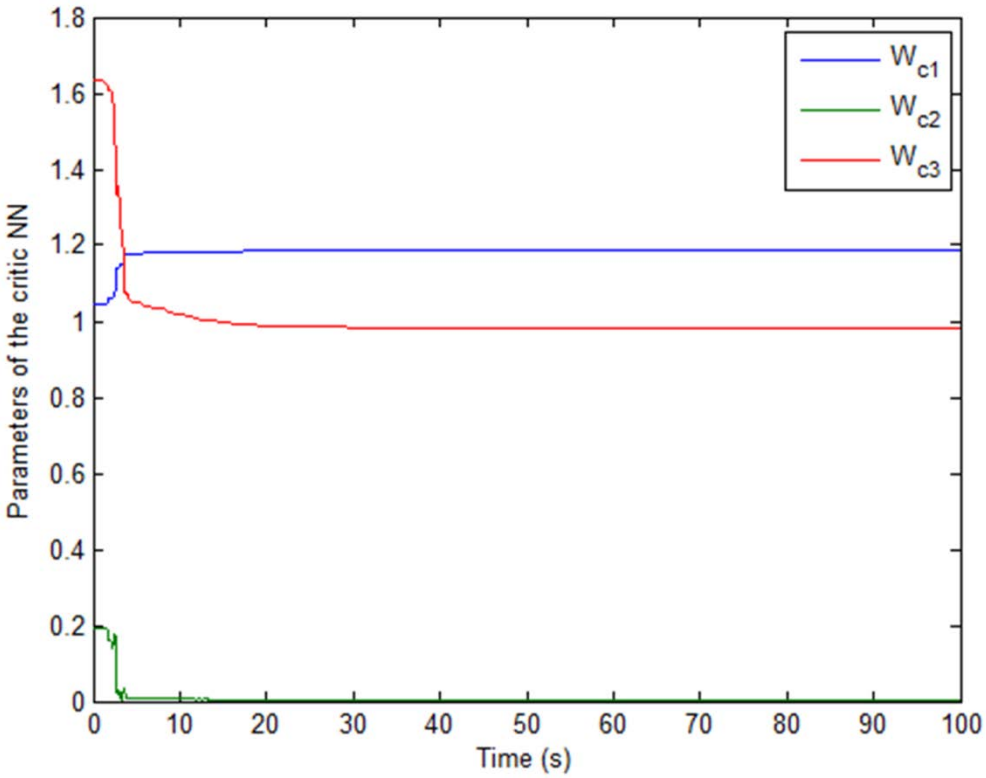


Fig. 6.35 Convergence of critic parameters.

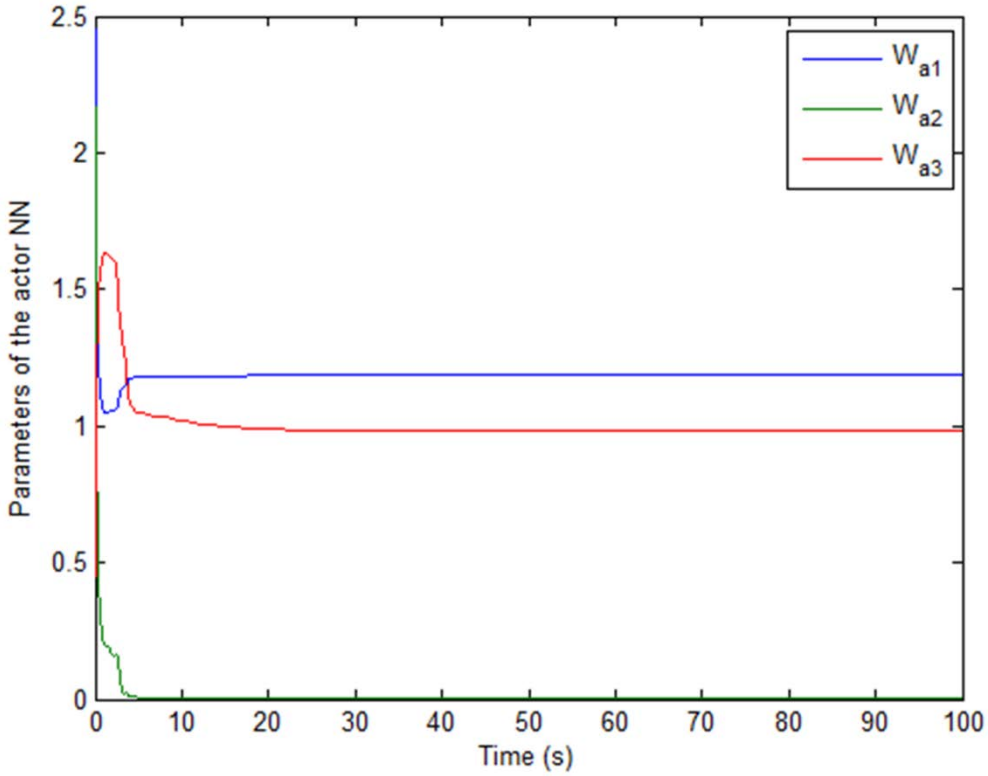


Fig. 6.36 Convergence of actor parameters.

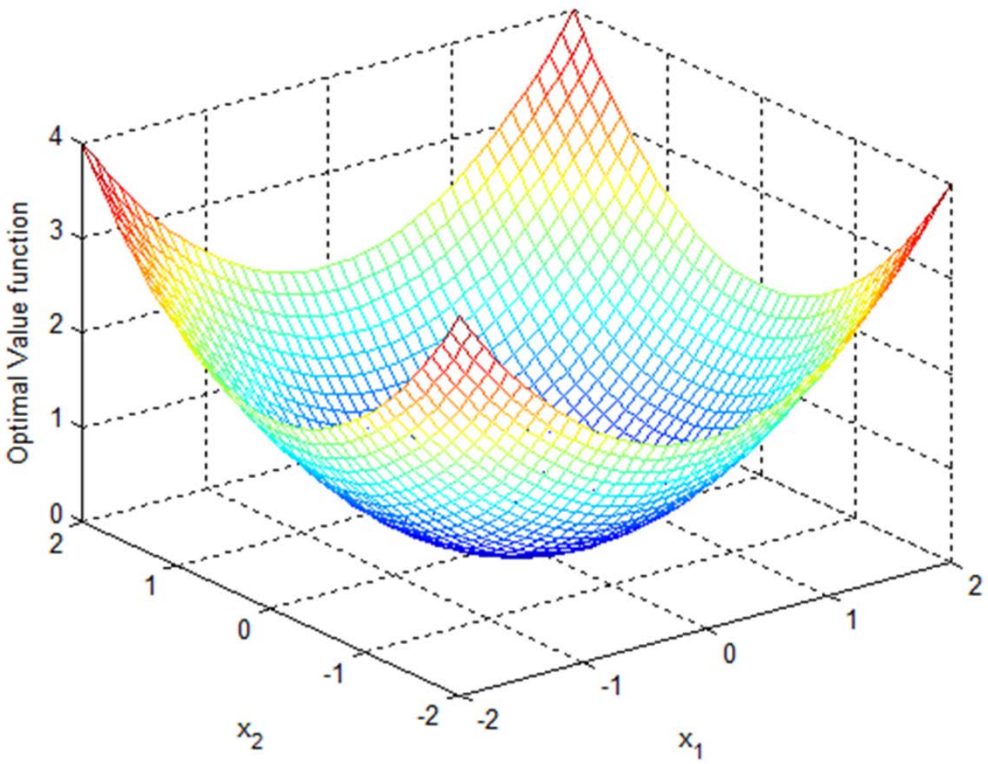


Fig. 6.37 3-D plot of optimal value function V*.

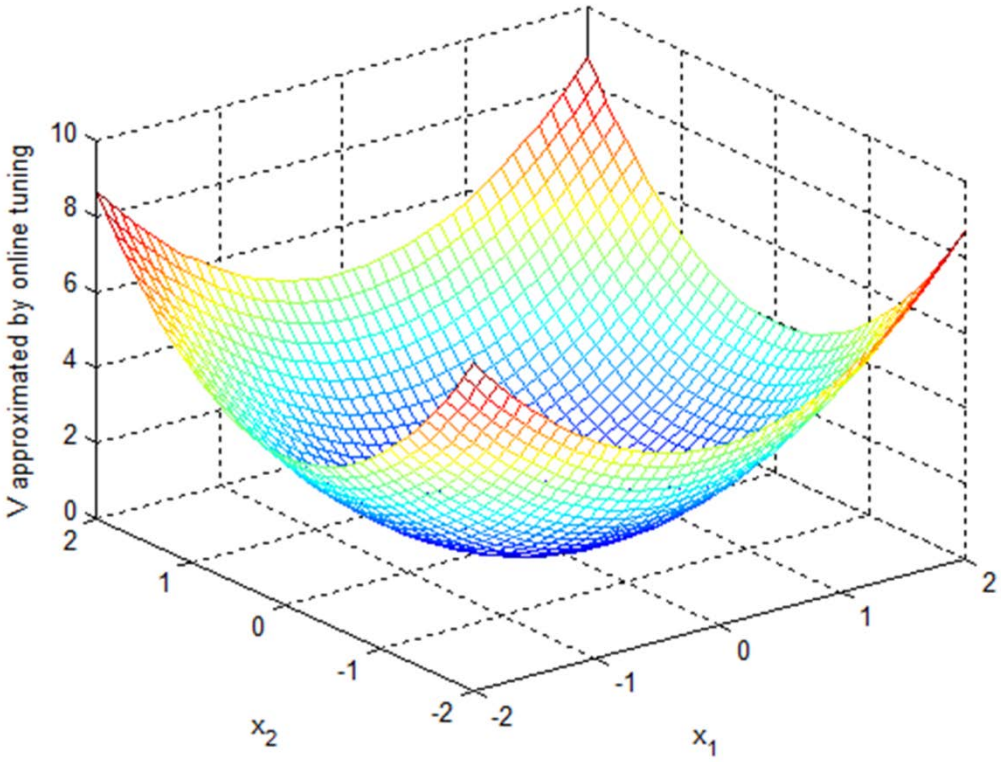


Fig. 6.38 3-D plot of neural network approximated value function by online tuning.

6.5.5 Vander Pol Oscillator System

Consider the state regulation problem of Vander Pol oscillator system described by dynamic equations [173]

$$\dot{x}_1 = x_2 \tag{6.46}$$

$$\dot{x}_2 = -x_1 + \alpha(1-x_1^2)x_2 + (1+x_1^2+x_2^2)u \tag{6.47}$$

with considering parameter $\alpha = 0.3$.

For implementation of synchronous PI algorithm the initial conditions of states are taken as $x_0 = \begin{bmatrix} 1 & -1 \end{bmatrix}$, and both the critic NN and actor NN weight vectors are randomly initialized with each element respectively as $W_{10} = 2*rand$, $W_{20} = 3*rand$. The learning constants are taken as $a_1 = 5$, $a_2 = 1$. The infinite horizon quadratic cost function is considered as by defining $Q(x) = x_1^2 + x_2^2$, and $R = 1$.

The optimal value function is

$$V^*(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

and the optimal control signal is

$$u^*(x) = -\frac{1}{2}(1+x_1^2+x_2^2)x_2$$

Consider the critic NN activation function vector as

$$\varphi_1(x) = \begin{bmatrix} x_1^2 & x_1 x_2 & x_2^2 \end{bmatrix}^T$$

The critic NN and actor NN weights are denoted respectively as

$$\hat{W}_1 = W_c = \begin{bmatrix} W_{c1} & W_{c2} & W_{c3} \end{bmatrix}^T$$

$$\hat{W}_2 = W_a = \begin{bmatrix} W_{a1} & W_{a2} & W_{a3} \end{bmatrix}^T$$

which are tuned at the same time using synchronous PI algorithm.

For the PE condition a small probing noise $n(t)$ is added to control input during simulation for time $t \leq 80$. After 100 seconds the critic NN and actor NN weights are converged respectively to

$$W_c = \begin{bmatrix} 0.8754 & 1.0159 & 0.5505 \end{bmatrix}^T$$

$$W_a = \begin{bmatrix} 0.8754 & 1.0159 & 0.5505 \end{bmatrix}^T$$

Thus, from (6.32) the actor NN is given by

$$\hat{u}_2(x) = -\frac{1}{2} R^{-1} \begin{bmatrix} 0 \\ 1 + x_1^2 + x_2^2 \end{bmatrix}^T \begin{bmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \end{bmatrix}^T W_a$$

which gives the neural network approximate solution of optimal control policy.

Fig. 6.39 shows the state trajectories of nonlinear system (6.46) and (6.47) which converge towards the equilibrium point zero. The convergence of parameters of critic NN and actor NN to optimal values are shown respectively in Figs. 6.40 and 6.41. Fig. 6.42 shows the 3-D mesh-plot of optimal value function. Fig. 6.43 shows the 3-D mesh-plot of neural network approximated value function using online synchronous PI technique.

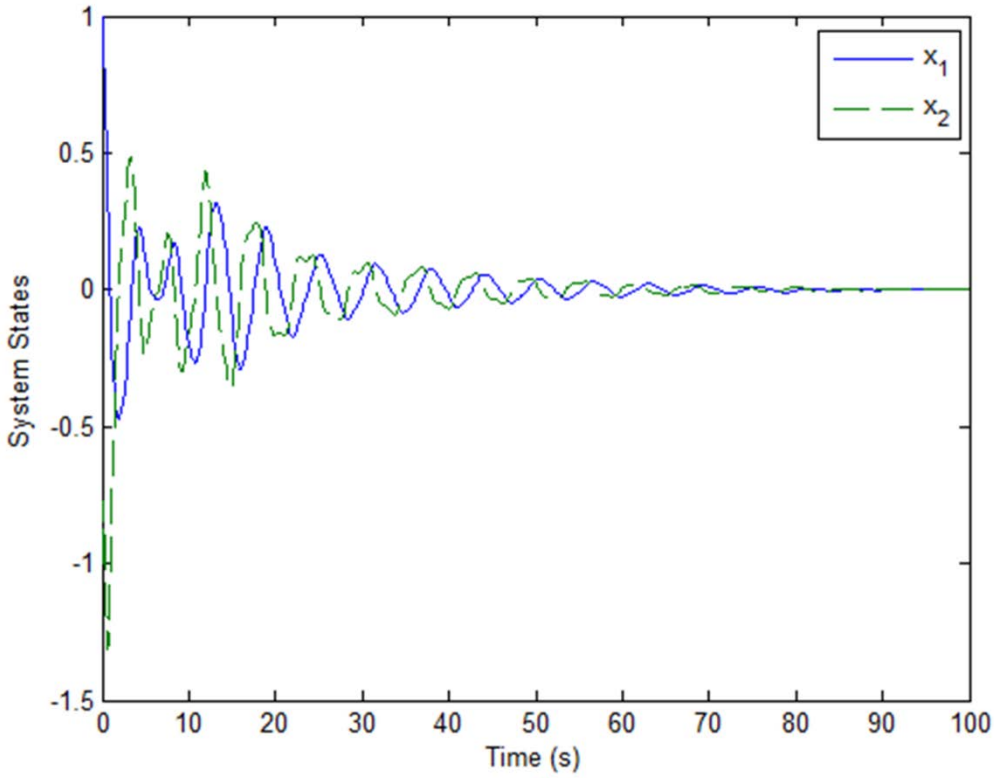


Fig. 6.39 System states trajectories.

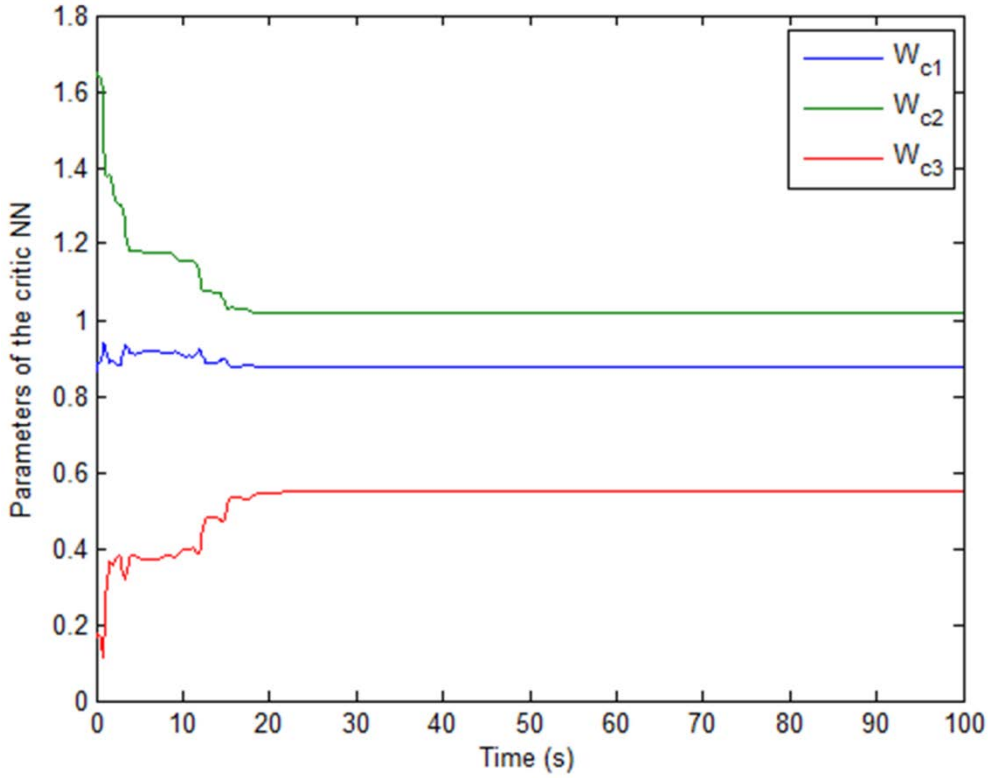


Fig. 6.40 Convergence of critic parameters.

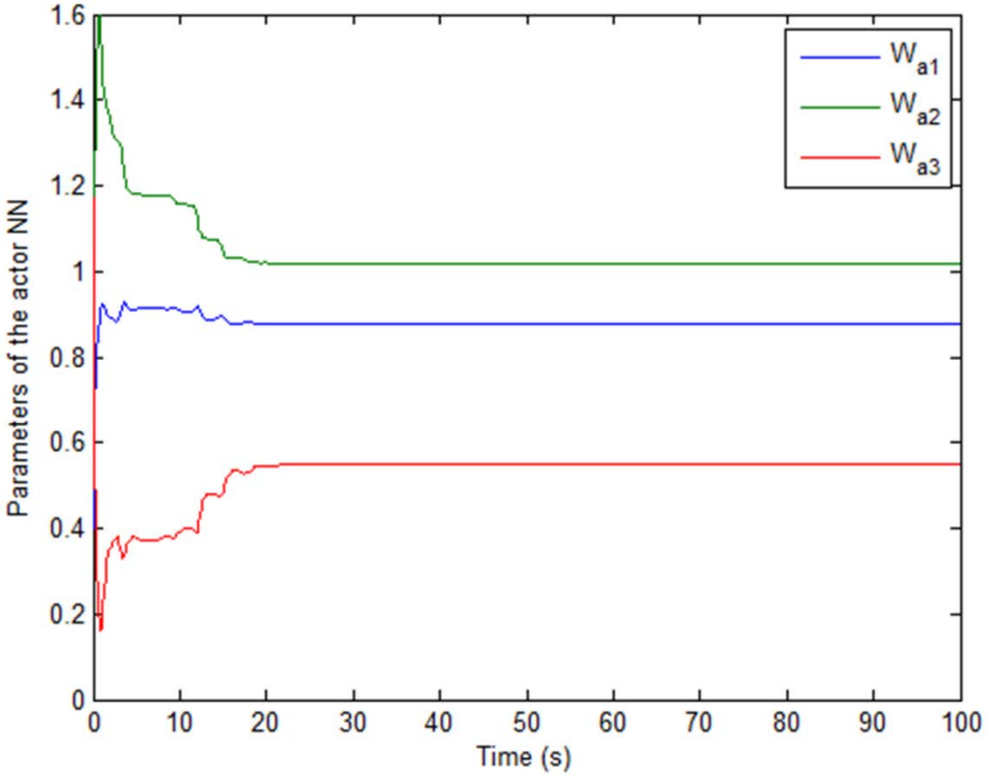


Fig. 6.41 Convergence of actor parameters.

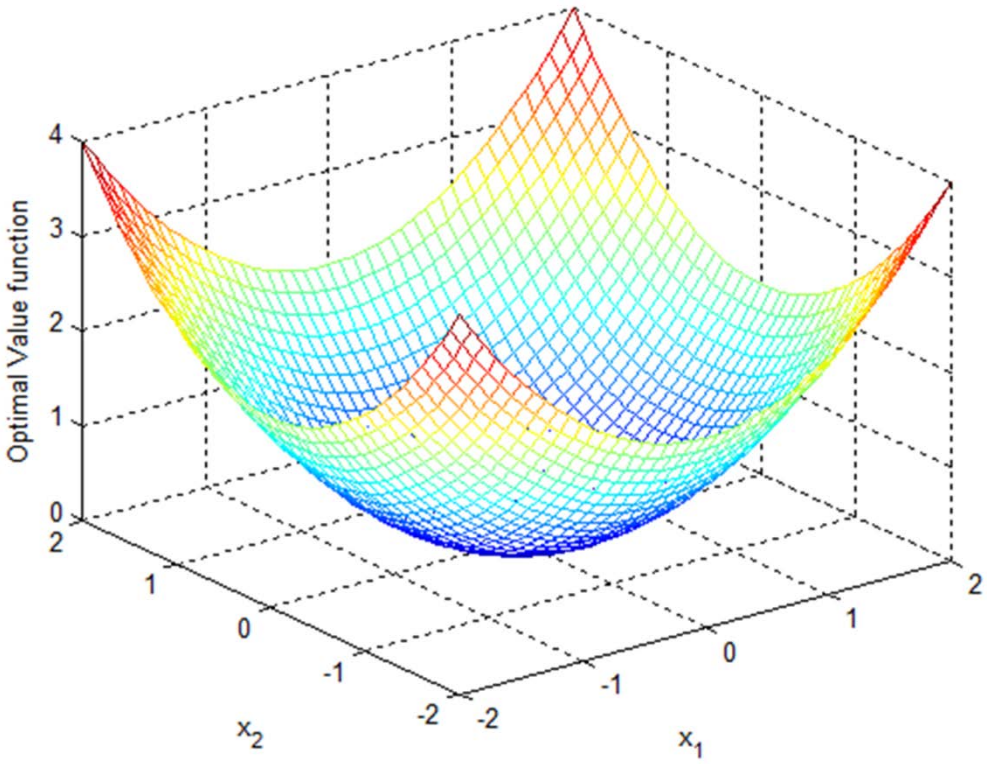


Fig. 6.42 3-D plot of optimal value function V*.

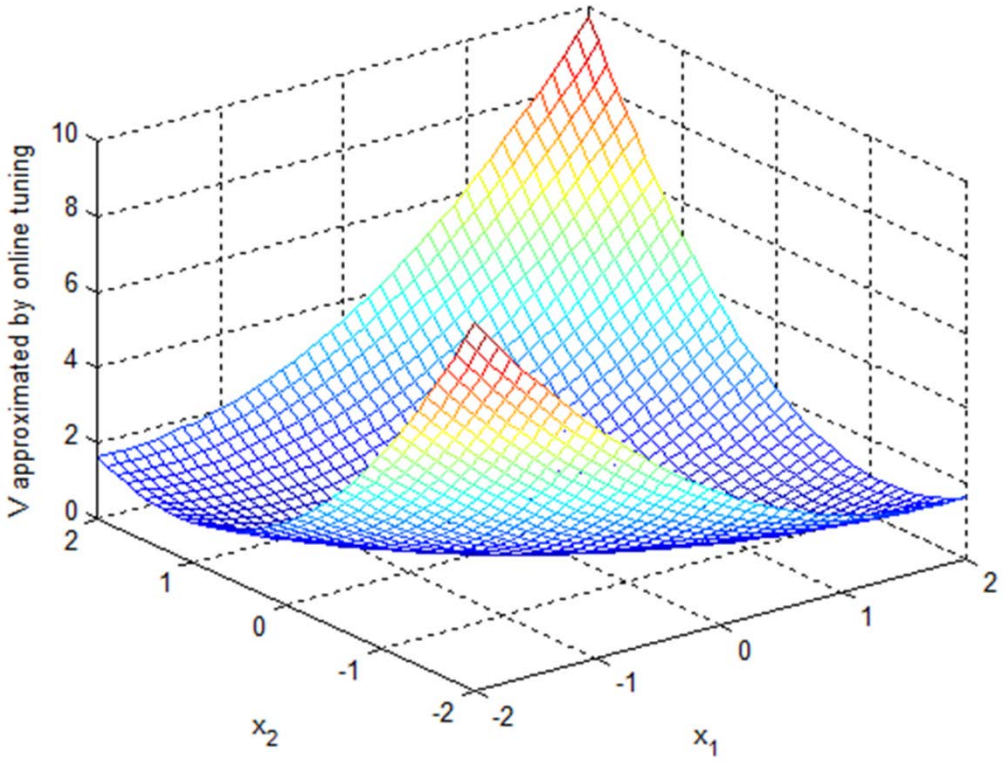


Fig. 6.43 3-D plot of neural network approximated value function by online tuning.

6.6 CONCLUSIONS

The synchronous PI algorithm is built on an actor-critic structure consisting of two step iteration: policy evaluation and policy improvement. These two steps are concurrently performed in synchronous PI, whereas in PI, these two steps are sequentially completed. The synchronous PI algorithm uses two neural networks known as critic NN and actor NN which approximate the value function and control policy respectively. In online synchronous PI algorithm, both the critic NN and actor NN are tuned simultaneously based on continuous-time data measurement from the system. The convergence of synchronous PI algorithm explicitly requires a PE condition which is ensured by adding a small probing noise to control signal. For evaluation of cost or update of control policy, the synchronous PI algorithm requires the complete knowledge of system dynamics whereas the PI algorithm requires only input-to-state dynamics. The online synchronous PI technique based adaptive critic scheme provides an infinite horizon online adaptive optimal control solution. Using neural networks to parameterize critic and actor for online implementation, the synchronous PI based control scheme provides a high-level intelligent control. Thus, the synchronous PI technique based adaptive critic scheme provides an intelligent adaptive optimal control.

In this chapter the application of online synchronous PI based control scheme is implemented for continuous-time time-invariant dynamical systems optimal control problems. Examples of both continuous-time LTI systems and input-affine nonlinear systems are considered for implementation of the control scheme. The LTI systems- LFC system models

without and with integral control both, and AVR system models neglecting and including sensor dynamics both are considered for implementation of synchronous PI based control scheme and performance analysis also with structural change in system dynamics. The performance of synchronous PI based adaptive critic control scheme is compared with LQR approach. It is observed that the synchronous PI based adaptive critic control scheme gives the similar response as of LQR approach. The input-affine nonlinear systems- general nonlinear system examples, single-link manipulator and Vander Pol's oscillator systems are considered for implementation of control scheme and performance analysis. The simulation results and performance analysis justify the effectiveness of synchronous PI based control scheme also for affine nonlinear systems control applications.

CHAPTER 7

CONCLUSIONS AND FUTURE SCOPE OF WORK

This chapter concludes the research work. The conclusions and future scope of works are presented in this chapter.

7.1 CONCLUSIONS

Since the conclusions are given at the end of each chapter. This section presents the conclusions of research work at a glance.

In this research work the performance investigations of various control schemes of PID control, optimal control, adaptive control, intelligent control, intelligent adaptive control, adaptive optimal control, and intelligent adaptive optimal control of dynamical systems considering certain linear and nonlinear dynamical systems applications are presented. Intelligent adaptive optimal control has been emerged from the integration of adaptive control and optimal control methodologies with intelligent computational techniques. Intelligent adaptive optimal control is a viable recent approach for the high performance applications with stringent design specifications. In this research work the performance investigation of intelligent adaptive optimal control of dynamical systems is presented. The applications of control schemes for dynamical systems control are implemented considering certain examples of linear and nonlinear dynamical systems to attempt this research investigation.

The performance of controlled systems is desired to be optimal which should be valid also when applied in the real situation. Adaptive control which is able to deal with uncertainties is generally not optimal. Adaptive control schemes require a priori knowledge of the system dynamics, and give online solution for the control problems. Optimal control is offline, and needs the knowledge of system dynamics for its design. Thus, to have both features of control design, it is desired to design online adaptive optimal control. The intelligent control techniques also require a priori knowledge of the system dynamics for control design; can give the online solution for the dynamical systems control problems.

Policy Iteration (PI) technique provides a direct adaptive optimal control of dynamical systems. The infinite horizon optimal solution using Hamilton-Jacobi-Bellman (HJB) equation and algebraic Riccati equation (ARE) which gives linear quadratic regulator (LQR) require the complete knowledge of the system dynamics. Also these techniques give offline solution. The online PI technique solves HJB equation by direct approach by starting with evaluation of the cost of a given initial admissible (stabilizing) control policy to converge towards state feedback optimal control. Based on actor-critic structure the PI algorithm consists of two-step iteration: policy evaluation and policy improvement. The online PI algorithm solves online the continuous-time optimal control problem without using the knowledge of system internal dynamics subject to the real-time dynamics of continuous-time system. The knowledge of

internal state dynamics is not needed for evaluation of cost or the update of control policy; and only the knowledge of input-to-state dynamics is required for updating the control policy. Thus, it provides a partially model-free approach. Using neural networks to parameterize actor and critic for online implementation, the PI based control scheme becomes a high-level intelligent control scheme with online adaptive optimal control solution. In PI algorithm the critic and actor are sequentially updated taking other one constant at policy evaluation and policy improvement steps. In online synchronous PI algorithm both critic NN and actor NN are tuned synchronously at the same time. The online synchronous PI algorithm needs the complete knowledge of system dynamics. It solves the optimal control problem online using real-time measurements of closed-loop signals. The persistence of excitation (PE) condition is explicitly required for convergence of synchronous PI algorithm. Using neural networks approximations for critic and actor both for online implementation it provides a high-level intelligent control. Thus, the online synchronous PI technique based control scheme gives an online intelligent adaptive optimal control solution for continuous-time dynamical systems.

In this research work the various research objectives have been considered by accomplishing which the research findings are summarized as following:

1. Optimal control of nonlinear inverted pendulum dynamical system using PID controller & LQR.
   - The inverted pendulum stabilizes in the upright position and cart reaches the desired position quickly & smoothly even under the continuous disturbance input.
   - The performance of proposed PID+LQR control method is better than PID control.
   - PID+LQR control scheme can be used for optimal control of nonlinear dynamical systems.
   - The proposed PID+LQR control approach is simple, effective & robust control scheme for the optimal control of nonlinear dynamical systems.
2. Intelligent control of nonlinear inverted pendulum dynamical system using Mamdani and TSK fuzzy inference systems.
   - The inverted pendulum stabilizes in vertically upright position and cart approaches the desired position even under the continuous disturbance input.
   - The performance of PD-FLC using Mamdani type FIS is better than PID controller.
   - The performance of direct fuzzy control using TSK FIS is better than both the PD-FLC using Mamdani type FIS and PID controller.
   - The response of direct fuzzy control using TSK FIS is more smooth & fast than both Mamdani type PD-fuzzy control and PID control.
3. Optimal control design using LQR for automatic generation control of two-area interconnected power system.
   - The deviations in area frequencies and tie-line power converge to zero.

- The step responses obtained in the case of optimal control using integral controller and LQR are smoother and faster than both the cases of integral control and optimal control using LQR for AGC system of two-area interconnected power system.
- The performance of optimal control using integral control and LQR is better than both integral control scheme and LQR method.
- The infinite horizon optimal control solution using LQR design provides the effective and robust performance for dynamical systems.

4. Intelligent control using fuzzy-PI controller for automatic generation control of two-area interconnected nonlinear power system.
   - The deviations in area frequencies and tie-line power converge to zero for both cases of linear and nonlinear system models of AGC system of two-area interconnected nonlinear power system.
   - The fuzzy-PI controller using Mamdani type FIS gives smoother and much faster response than integral controller, and thus it is a simple and better control method.
5. Intelligent control of process system using radial basis function neural networks.
   - The response of indirect adaptive control using RBFNN is fast for nonlinear process system.
6. Adaptive optimal control using policy iteration technique for LTI systems.
   - The online policy iteration (PI) technique based adaptive critic scheme provides an online infinite horizon adaptive optimal control of continuous-time LTI dynamical systems.
   - The applications of PI based adaptive critic control scheme for various general and practical examples of continuous-time LTI systems- general LTI SISO system, a 4^{th} order mechanical system, load frequency control of power system, automatic voltage regulator of power system, and DC motor speed control system; and the comprehensive performance analysis justify the effectiveness of this control scheme.
   - The online PI algorithm does not require the knowledge of system's internal dynamics (i.e. matrix *A*) for evaluation of cost or the update of control policy; only the knowledge of input matrix *B* is required for updating the control policy.
   - The adaptive optimal control scheme using PI technique is partially model-free, effective & robust.
   - The information of internal state dynamics is extracted from real-time data measurements along state trajectory at discrete moments of time.
   - The PI algorithm requires an initial stabilizing control policy.
   - The critic parameter matrix P and actor parameter matrix K obtained from adaptive critic scheme using PI technique converge adaptively to the optimal values.

- The critic parameter matrix P and actor parameter matrix K obtained from adaptive critic scheme using PI technique are mostly of same values to that obtained from LQR approach.
- Also in case of change in system parameter in real situation the controller adapts it and converges to same optimal values, and thus the actor K and critic P parameters remain unchanged.
- The structural change introduced in system dynamics by including integral control for LFC system, integral compensator for DC motor control system, and sensor dynamics for AVR system augments the system behavior such as of its credit to that particular system.
- Integral control in LFC system and integral compensator in DC motor control system remove the steady state error in closed loop responses.
- The closed loop step response of adaptive critic control scheme is similar to the LQR approach.
- The structural change in system is not adapted by the PI based adaptive critic controller.
- The PI based adaptive critic controller adapts the change in system parameters in real situation at any moment of time.
- The application of PI technique gives a novel adaptive optimal control design for a 4^{th} order mechanical system.
- The application of PI technique gives a novel adaptive optimal control design for AVR system.
- The application of PI technique gives a novel adaptive optimal control design for DC motor control system.
- The novel modeling of DC motor control system with integral compensator gives a stable open-loop system matrix A.

7. Adaptive optimal control using policy iteration technique for affine nonlinear systems.
   - The online PI technique based adaptive critic scheme provides an online infinite horizon adaptive optimal control of continuous-time affine nonlinear dynamical systems.
   - The applications of PI based adaptive critic control scheme considering the state regulation problem for certain general affine nonlinear systems and certain practical examples of affine nonlinear systems- single-link manipulator, inverted pendulum-cart system, Vander Pol oscillator system, and the comprehensive performance analysis justify the effectiveness of this control scheme.
   - The online PI algorithm does not require the knowledge of system's internal dynamics (i.e. $f(x)$) for evaluation of cost or the update of control policy; only the

knowledge of input-to-state dynamics (i.e. $g(x)$) is required for updating the control policy.

- The adaptive optimal control scheme using PI technique is partially model-free, effective & robust.
- The online PI algorithm requires an initial stabilizing controller for converging to the optimal solution.
- Using neural network approximation of cost function for online implementation of PI algorithm gives a high-level intelligent control.
- The system states converge towards the equilibrium point at origin, and the control signal remains bounded converging towards zero.
- Neural networks weights are adjusted to the optimal values which give the critic parameters converging adaptively to optimal values and thus the control policy is adaptive optimal.

8. Intelligent adaptive optimal control using synchronous policy iteration technique for LTI systems.
   - The applications of online synchronous PI based adaptive critic control scheme for practical examples of continuous-time LTI systems- load frequency control of power system without and with integral control, and automatic voltage regulator of power system neglecting and including sensor dynamics and comprehensive performance analysis justify the effectiveness of control scheme.
   - The closed loop step response of synchronous PI based adaptive critic control scheme is similar to the LQR approach.
   - The structural change in system is not adapted also by the synchronous PI based adaptive critic controller.
9. Intelligent adaptive optimal control using synchronous policy iteration technique for affine nonlinear systems.
   - The applications of online synchronous PI based control scheme also for affine nonlinear systems considering the state regulation problems for certain general affine nonlinear systems of stronger nonlinearities and practical examples of affine nonlinear systems- single-link manipulator, and Vander Pol oscillator, and performance analysis justify the effectiveness control scheme.
   - The online synchronous PI algorithm requires the complete knowledge of system dynamics for evaluation of cost or update of control policy.
   - The synchronous PI based control scheme is model-based whereas the PI based control scheme is partially model-free.
   - Both critic NN and actor NN are tuned simultaneously based on continuous-time data measurement from the system.

- The neural networks weights of both critic and actor are converged to same optimal values.
- The neural network approximated value function by online tuning using synchronous PI is almost similar to the optimal value function.
- The convergence of synchronous PI algorithm explicitly requires a PE condition which is ensured by adding a small probing noise to control signal.
- The online synchronous PI based adaptive critic control scheme provides an online infinite horizon adaptive optimal control.
- Using neural networks approximations of both critic and actor functions for online implementation, the online synchronous PI based adaptive critic scheme provides an intelligent adaptive optimal control for continuous-time dynamical systems.

7.2 FUTURE SCOPE OF WORKS

The intelligent adaptive optimal control is an emerging recent research area having a vide scope of work based on the complexity of the control problems with constraints and control algorithms. Related to the research objectives undertaken in this research work some of the future scope of research works is listed as under:

1. The performance investigation of PID+LQR control approach with tuning of PID controller parameters using GA, and PSO instead of trial & error method for inverted pendulum-cart system.
2. The real-time experimental investigation of PID+LQR control scheme.
3. The real-time experimental investigation of fuzzy control schemes.
4. Optimal control design using integral controller and LQR for AGC system of multi-area interconnected nonlinear power system.
5. Performance investigation of fuzzy-PID controller for multi-area interconnected nonlinear power systems.
6. The real-time experimental investigation of PI based adaptive critic control schemes for continuous-time linear and nonlinear dynamical systems.
7. Extension of application of PI technique to LFC system of multi-area interconnected power system problem.
8. Implementation of PI technique for non-affine nonlinear systems.
9. Performance investigation of PI based control scheme for tracking control problems.
10. Since PI algorithm requires an initial stabilizing control policy. The investigations may be carried out for non-requirement of initial stabilizing control policy.
11. The investigation to find the possibility of convergence of synchronous PI algorithm without explicit need of persistence of excitation condition.

12. Since the synchronous PI algorithm requires the complete knowledge of system dynamics, the further investigations can be made to modify the algorithm for partial model-free approach at par with PI algorithm.

PUBLICATIONS FROM THE RESEARCH WORK

International Journals

[1] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Optimal Control of Nonlinear Inverted Pendulum Dynamical System using PID Controller and LQR: Performance Analysis without & with Disturbance Input", International Journal of Automation and Computing (Springer) (**Accepted - In Press**).

[2] **Lal Bahadur Prasad**, Hari Om Gupta, and Barjeev Tyagi, "Intelligent Control of Nonlinear Inverted Pendulum Dynamical System using Mamdani & TSK Fuzzy Inference Systems: Performance Analysis without & with Disturbance Input", Journal of Control Engineering and Applied Informatics. (Under Review).

[3] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Adaptive Critic Design Using Policy Iteration Technique for LTI Systems: A Comprehensive Performance Analysis", Journal of Control, Automation and Electrical Systems (Springer) (Revision Under Review).

[4] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Adaptive Optimal Control using Policy Iteration Technique for LTI Systems: A Performance Analysis", Automatica (Elsevier) (Communicated).

[5] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Application of Policy Iteration Technique Based Adaptive Optimal Control Design for Automatic Voltage Regulator of Power System", International Journal of Electrical Power & Energy Systems, 2014 (Elsevier) (**In Press**) http://dx.doi.org/10.1016/j.ijepes.2014.06.057.

[6] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Application of Policy Iteration Technique for Adaptive Optimal Control of LTI Systems", IETE Journal of Research (Taylor & Francis) (**Accepted - In Press**).

[7] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Adaptive Critic Design using Policy Iteration Technique for Affine Nonlinear Systems: A Performance Analysis", Applied Soft Computing (Elsevier) (Under Review).

[8] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Application of Policy Iteration Technique for Adaptive Optimal Control of Affine Nonlinear Systems", International Journal of Automation and Computing (Springer) (Communicated).

[9] **Lal Bahadur Prasad**, Hari Om Gupta, and Barjeev Tyagi, "Adaptive Optimal Control using Policy Iteration Technique for Affine Nonlinear Systems", International Journal of Dynamics and Control (Springer) (Communicated).

[10] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, "Fuzzy-PI Based Automatic Generation Control of Nonlinear Power System", Electric Power Systems Research (Elsevier) (Communicated).

International Conferences

[11] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, “Optimal Control of Nonlinear Inverted Pendulum Dynamical System with Disturbance Input using PID Controller & LQR”, Proc. of 2011 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2011), at Holiday Inn, Bath Ferringhi, Penang, Malaysia, on dt. 25-27 Nov, 2011 (IEEE Conf. ID 19475, Print ISBN: 978-1-4577-1640-9, IEEE Xplore DOI: 10.1109/ICCSCE.2011.6190585), pp. 540-545. (Citations-8)

[12] **Lal Bahadur Prasad**, Hari Om Gupta, and Barjeev Tyagi, “Intelligent Control of Nonlinear Inverted Pendulum Dynamical System using Mamdani and TSK Fuzzy Inference Systems”, Proc. of 2011 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC 2011), at CIT, Kanyakumari, India, on dt. 15-18 Dec, 2011 (IEEE Xplore: CFB1120J-ART ISBN: 978-1-61284-1, Print Version: CFB1120J-PRT ISBN: 978-1-61284-766-5), pp. 106-111.

[13] **Lal Bahadur Prasad**, Hari Om Gupta, and Barjeev Tyagi, “Intelligent Control of Nonlinear Inverted Pendulum Dynamical System with Disturbance Input using Fuzzy Logic Systems”, Proc. of 2011 IEEE International Conference on Recent Advancements in Electrical, Electronics and Control Engineering (IConRAEeCE’11), at Mepco Schelnk Engineering College, Sivakasi, India, on dt. 15-17 Dec, 2011, (IEEE Conf. ID: 19064, IEEE Xplore DOI: 10.1109/ICONRAEeCE.2011.6129799), pp. 136-141. (Citations-4)

[14] **Lal Bahadur Prasad**, Barjeev Tyagi, and Hari Om Gupta, “Modelling & Simulation for Optimal Control of Nonlinear Inverted Pendulum Dynamical System using PID Controller & LQR”, Proc. of 6th Asia Modelling Symposium (AMS2012) / 6th Asia International Conference on Mathematical Modelling and Computer Simulation, at Bali, Indonesia, on dt. 29-31 May, 2012, (IEEE Conf. ID: 20807, IEEE CSDL & IEEE Xplore DOI: 10.1109/AMS.2012.21), pp. 138-143. (Citations-6)

[15] **Lal Bahadur Prasad**, Hari Om Gupta, and Barjeev Tyagi, “Adaptive Optimal Control of Nonlinear Inverted Pendulum System using Policy Iteration Technique”, Proc. of ACODS-2014: Advances in Control and Optimization of Dynamical Systems, IIT Kanpur, 13-15 March, 2014, 3 (1), pp. 1138-1145 (IFAC-PapersOnline) (Elsevier).

BIBLIOGRAPHY

[1] **K. Ogata**, *System Dynamics* 4ed. New Delhi: Pearson Education (Singapore) Pvt. Ltd., 2004.

[2] **K. Ogata**, *Modern Control Engineering*, 4 ed. New Delhi: Pearson Education (Singapore) Pvt. Ltd., 2005.

[3] **K. Ogata**, *Discrete-Time Control Systems*, 2 ed. New Delhi: Pearson Education, Inc., 2005.

[4] **I. J. Nagrath and M. Gopal**, *Control Systems Engineering*, 2 ed. New Delhi: Wiley Eastern Ltd., 1992.

[5] **M. Gopal**, *Control Systems: Principles and Design*, 2 ed. New Delhi: Tata McGraw Hill Publishing Company Ltd., 2002.

[6] **M. Gopal**, *Modern Control System Theory*, 2 ed. New Delhi: New Age International (P) Ltd., Publishers, 2004.

[7] **M. Gopal**, *Digital Control and State Variable Methods: Conventional and Intelligent Control Systems*, 4 ed. New Delhi: Tata McGraw Hill Education Pvt. Ltd., 2012.

[8] **D. Roy Choudhury**, *Modern Control Engineering*, 1 ed. New Delhi: PHI (P) Ltd., 2005.

[9] **Ajit K. Mandal**, *Introduction to Control Engineering: Modeling, Analysis and Design*, 1 ed. New Delhi: New Age International (P) Ltd., Publishers, 2000.

[10] **M. N. Bandyopadhyay**, *Control Engineering: Theory and Practice*, 1 ed. New Delhi: PHI (P) Ltd., 2004.

[11] **M. Gevers**, "Identification for control," Annual Reviews in Control, vol. 20, no. 0, pp. 95-106, 1996.

[12] **G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval**, "Past, present and future of nonlinear system identification in structural dynamics," Mechanical Systems and Signal Processing, vol. 20, no. 3, pp. 505-592, 2006.

[13] **Kumpati S. Narendra and Kannan Parthasarathy**, "Identification and Control of Dynamical Systems Using Neural Networks," IEEE Transactions on Neural Networks, vol. 1, no. 1, pp. 4-27, March 1990.

[14] **S. Srivastava, M. Singh, M. Hanmandlu, and A. N. Jha**, "New fuzzy wavelet neural networks for system identification and control," Applied Soft Computing, vol. 6, no. 1, pp. 1-17, 2005.

[15] **Patricia Melin and Oscar Castillo**, *Modelling, Simulation and Control of Non-linear Dynamical Systems: An Intelligent Approach Using Soft Computing and Fractal Theory*, 1 ed. London: Taylor & Francis Inc., 2002.

[16] **M. A. Denaï, F. Palis, and A. Zeghbib**, "Modeling and control of non-linear systems using soft computing techniques," Applied Soft Computing, vol. 7, no. 3, pp. 728-738, 2007.

[17] **L. Fortuna, G. Nunnari, and A. Gallo**, *Model Order Reduction Techniques with Applications in Electrical Engineering*, 1 ed. London: Springer-Verlag London Ltd., 1992.

[18] **S. K. Nagar and S. K. Singh**, "An algorithmic approach for system decomposition and balanced realized model reduction," Journal of the Franklin Institute, vol. 341, no. 7, pp. 615-630, 2004.

[19] **A. K. Sahani and S. K. Nagar**, "Design of digital controllers for multivariable systems via time-moments matching," Computers & Electrical Engineering, vol. 24, no. 5, pp. 335-347, 1998.

[20] **B. Bandyopadhyay, A. G. E. Abera, S. Janardhanan, and V. Sreeram**, "Sliding mode control design via reduced order model approach," International Journal of Automation and Computing, vol. 4, no. 4, pp. 329-334, 2007/10/01 2007.

[21] **S. K. Nagar, J. Pal, and J. D. Sharma**, "A computer-aided method for digital controller design via optimization," Computers & Electrical Engineering, vol. 16, no. 2, pp. 79-85, 1990.

[22] **J. Pal, S. K. Nagar, and J. D. Sharma**, "Digital controller design for systems with transport lag," International Journal of Systems Science, vol. 23, no. 12, pp. 2385-2392, 1992/12/01 1992.

[23] **K. J. Åström and T. Hägglund**, "The future of PID control," Control Engineering Practice, vol. 9, no. 11, pp. 1163-1175, 2001.

[24] **Kiam Heong Ang, Gregory Chong, and Yun Li**, "PID Control System Analysis, Design, and Technology," IEEE Transactions on Control Systems Technology, vol. 13, no. 4, pp. 559- 576, July 2005.

[25] **N. Hohenbichler**, "All stabilizing PID controllers for time delay systems," Automatica, vol. 45, no. 11, pp. 2678-2684, 2009.

[26] **D. G. Padhan and S. Majhi**, "A new control scheme for PID load frequency controller of single-area and multi-area power systems," ISA Transactions, vol. 52, no. 2, pp. 242-251, 2013.

[27] **D. Valério and J. S. da Costa**, "Tuning of fractional PID controllers with Ziegler–Nichols-type rules," Signal Processing, vol. 86, no. 10, pp. 2771-2784, 2006.

[28] **M. I. Alomoush**, "Load frequency control and automatic generation control using fractional-order controllers," Electrical Engineering, vol. 91, no. 7, pp. 357-368, 2010/03/01 2010.

[29] **Z. Zhen-Yu, M. Tomizuka, and S. Isaka**, "Fuzzy gain scheduling of PID controllers," Systems, Man and Cybernetics, IEEE Transactions on, vol. 23, no. 5, pp. 1392-1398, 1993.

[30] **H. Bao-Gang, G. K. I. Mann, and R. G. Gosine**, "A systematic study of fuzzy PID controllers-function-based evaluation approach," IEEE Transactions on Fuzzy Systems, vol. 9, no. 5, pp. 699-712, 2001.

[31] **V. Kumar and A. P. Mittal**, "Parallel fuzzy P+fuzzy I+fuzzy D controller: Design and performance evaluation," International Journal of Automation and Computing, vol. 7, no. 4, pp. 463-471, 2010/11/01 2010.

[32] **C.-F. Hsu, G.-M. Chen, and T.-T. Lee**, "Robust intelligent tracking control with PID-type learning algorithm," Neurocomputing, vol. 71, no. 1–3, pp. 234-243, 2007.

[33] **H. Delavari, A. Ranjbar Noiey, and S. Minagar**, "Artificial Intelligent Controller for a DC Motor," in *Advances in Computer Science and Engineering*. vol. 6, H. Sarbazi-Azad, B. Parhami, S.-G. Miremadi, and S. Hessabi, Eds.: Springer Berlin Heidelberg, 2009, pp. 842-846.

[34] **Z. Jun, L. Ping, and W. Xue-Song**, "Intelligent PID controller design with adaptive criterion adjustment via least squares support vector machine," in *Chinese Control and Decision Conference, 2009, CCDC '09*, 2009, pp. 7-12.

[35] **J.-S. Chiou, S.-H. Tsai, and M.-T. Liu**, "A PSO-based adaptive fuzzy PID-controllers," Simulation Modelling Practice and Theory, vol. 26, no. 0, pp. 49-59, 2012.

[36] **G. Zwe-Lee**, "A particle swarm optimization approach for optimum design of PID controller in AVR system," IEEE Transactions on Energy Conversion, vol. 19, no. 2, pp. 384-391, 2004.

[37] **V. Mukherjee and S. P. Ghoshal**, "Intelligent particle swarm optimized fuzzy PID controller for AVR system," Electric Power Systems Research, vol. 77, no. 12, pp. 1689-1698, 2007.

[38] **V. Mukherjee and S. P. Ghoshal**, "Comparison of intelligent fuzzy based AGC coordinated PID controlled and PSS controlled AVR system," International Journal of Electrical Power & Energy Systems, vol. 29, no. 9, pp. 679-689, 2007.

[39] **A. Chatterjee, V. Mukherjee, and S. P. Ghoshal**, "Velocity relaxed and craziness-based swarm optimized intelligent PID and PSS controlled AVR system," International Journal of Electrical Power & Energy Systems, vol. 31, no. 7–8, pp. 323-333, 2009.

[40] **I. Pan and S. Das**, "Chaotic multi-objective optimization based design of fractional order PIλDμ controller in AVR system," International Journal of Electrical Power & Energy Systems, vol. 43, no. 1, pp. 393-407, 2012.

[41] **I. Pan and S. Das**, "Frequency domain design of fractional order PID controller for AVR system using chaotic multi-objective optimization," International Journal of Electrical Power & Energy Systems, vol. 51, no. 0, pp. 106-118, 2013.

[42] **Frank L. Lewis**, *Optimal Control*. New York: John Wiley & Sons, Inc., 1986.

[43] **Ronald S. Burns**, *Advanced Control Engineering*, 1 ed. Oxford: Butterworth-Heinemann, 2001.

[44] **K. Zhou, J. C. Doyle, and K. Glover**, *Robust and optimal control* vol. 272: Prentice Hall New Jersey, 1996.

[45] **R. C. Loxton, K. L. Teo, V. Rehbock, and K. F. C. Yiu**, "Optimal control problems with a continuous inequality constraint on the state and the control," Automatica, vol. 45, no. 10, pp. 2250-2257, 2009.

[46] **J. Fisher and R. Bhattacharya**, "Linear quadratic regulation of systems with stochastic parameter uncertainties," Automatica, vol. 45, no. 12, pp. 2831-2841, 2009.

[47] **K. S. Grewal, R. Dixon, and J. Pearson**, "LQG controller design applied to a pneumatic stewart-gough platform," International Journal of Automation and Computing, vol. 9, no. 1, pp. 45-53, 2012/02/01 2012.

[48] **N. Kumaresan and P. Balasubramaniam**, "Optimal control for stochastic linear quadratic singular system using neural networks," Journal of Process Control, vol. 19, no. 3, pp. 482-488, 2009.

[49] **M. A. Abido**, "Optimal design of power-system stabilizers using particle swarm optimization," IEEE Transactions on Energy Conversion, vol. 17, no. 3, pp. 406-413, 2002.

[50] **V. V. Chalam**, *Adaptive Control Systems: Techniques and Applications*. New York: Marcel Dekker Inc., 1987.

[51] **Shankar Sastry and Marc Bodson**, *Adaptive Control: Stability, Convergence, and Robustness*. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1989.

[52] **K. J. Astrom and B. Wittenmark**, *Adaptive Control*, 2 ed. New Delhi: Pearson Education, Inc., 2006.

[53] **K. J. Åström**, "Theory and applications of adaptive control—A survey," Automatica, vol. 19, no. 5, pp. 471-486, 1983.

[54] **Karl Johan Astrom**, "Adaptive Feedback Control," Proceedings of the IEEE, vol. 75, no. 2, pp. 185-217, February 1987.

[55] **B. D. O. Anderson and A. Dehghani**, "Challenges of adaptive control–past, permanent and future," Annual Reviews in Control, vol. 32, no. 2, pp. 123-135, 2008.

[56] **A.-V. Duka, S. E. Oltean, and M. Dulau**, "Model Reference Adaptive vs. Learning Control for the Inverted Pendulum. A Comparative Case Study," Journal of Control Engineering and Applied Informatics, vol. 9, no. 3, 4, pp. 67-75, 2007.

[57] **S. Janardhanan and B. Bandyopadhyay**, "Output Feedback Sliding-Mode Control for Uncertain Systems Using Fast Output Sampling Technique," IEEE Transactions on Industrial Electronics, vol. 53, no. 5, pp. 1677-1682, 2006.

[58] **S. Janardhanan and B. Bandyopadhyay**, "Discrete sliding mode control of systems with unmatched uncertainty using multirate output feedback," IEEE Transactions on Automatic Control, vol. 51, no. 6, pp. 1030-1035, 2006.

[59] **S. Janardhanan and B. Bandyopadhyay**, "Output feedback discrete-time sliding mode control for time delay systems," IEE Proceedings - Control Theory and Applications, vol. 153, no. 4, pp. 387-396, 2006.

[60] **B. Bandyopadhyay and D. Fulwani**, "High-Performance Tracking Controller for Discrete Plant Using Nonlinear Sliding Surface," IEEE Transactions on Industrial Electronics, vol. 56, no. 9, pp. 3628-3637, 2009.

[61] **C.-W. Tao, J.-S. Taur, C. Wang, and U. Chen**, "Fuzzy hierarchical swing-up and sliding position controller for the inverted pendulum–cart system," Fuzzy Sets and Systems, vol. 159, no. 20, pp. 2763-2784, 2008.

[62] **W. M. Bessa, A. S. de Paula, and M. A. Savi**, "Chaos control using an adaptive fuzzy sliding mode controller with application to a nonlinear pendulum," Chaos, Solitons & Fractals, vol. 42, no. 2, pp. 784-791, 2009.

[63] **O. Kaynak, K. Erbatur, and M. Ertugnrl**, "The fusion of computationally intelligent methodologies and sliding-mode control-a survey," IEEE Transactions on Industrial Electronics, vol. 48, no. 1, pp. 4-17, 2001.

[64] **Y. Xinghuo and O. Kaynak**, "Sliding-Mode Control With Soft Computing: A Survey," IEEE Transactions on Industrial Electronics, vol. 56, no. 9, pp. 3275-3285, 2009.

[65] **J. M. Zurada**, *Introduction to Artificial Neural Systems*. St. Paul: West Publishing Company, 1992.

[66] **S. Rajasekaran and G. A. Vijayalakshmi Pai**, *Neural Networks, Fuzzy Logic, and Genetic Algorithms: Synthesis and Applications*, 1 ed. New Delhi: PHI (P) Ltd., 2003.

[67] **Kevin M. Passino and Stephen Yurkovich**, *Fuzzy Control*, 1 ed. California: Addison-Wesley Longman, Inc., 1998.

[68] **F. Gang**, "A Survey on Analysis and Design of Model-Based Fuzzy Control Systems," IEEE Transactions on Fuzzy Systems, vol. 14, no. 5, pp. 676-697, 2006.

[69] **M. M. Gupta and N. K. Shinha**, *Intelligent Control Systems: Theory and Applications*: IEEE Press, 1996.

[70] **A. Zilouchian and M. Jamshidi**, "Intelligent Control Systems using Soft Computing Methodologies," Boca Raton: CRC Press LLC, 2001.

[71] **Laxmidhar Behera and Indrani Kar**, *Intelligent Systems and Control: Principles and Applications*, 1 ed. New Delhi: Oxford University Press, 2009.

[72] **Lakhmi C. Jain and Clarence W. de Silva**, *Intelligent Adaptive Control: Industrial Applications*, 1 ed. Boca Raton: CRC Press LLC, 1999.

[73] **Yung C. Shin and Chengying Xu**, *Intelligent Systems: Modeling, Optimization, and Control*. Boca Raton: CRC Press, Taylor & Francis Group, LLC 2009.

[74] **W. Yu**, *Recent Advances in Intelligent Control Systems*: Springer, 2009.

[75] **K. J. Astrom and Thomas J. McAvoy**, "Intelligent control," J. Proc. Cont., vol. 2, no. 3, pp. 115-127, 1992.

[76] **T. I. Liu, E. J. Ko, and J. Lee**, "Intelligent control of dynamic systems," Journal of the Franklin Institute, vol. 330, no. 3, pp. 491-503, 1993.

[77] **T. L. Ward, P. A. S. Ralston, and K. E. Stoll**, "Intelligent control of machines and processes," Computers & Industrial Engineering, vol. 19, no. 1–4, pp. 205-209, 1990.

[78] **D. M. V. Kumar**, "Intelligent controllers for automatic generation control," in *TENCON'98, 1998 IEEE Region 10 International Conference on Global Connectivity in Energy, Computer, Communication and Control*, 1998, pp. 557-574.

[79] **R.-J. Wai and J.-M. Chang**, "Intelligent control of induction servo motor drive via wavelet neural network," Electric Power Systems Research, vol. 61, no. 1, pp. 67-76, 2002.

[80] "Fuzzy Inference Systems," in *MATLAB Fuzzy Logic Toolbox User's Guide*: Mathworks Inc.

[81] **C. Kai-Yuan and Z. Lei**, "Fuzzy Reasoning as a Control Problem," IEEE Transactions on Fuzzy Systems, vol. 16, no. 3, pp. 600-614, 2008.

[82] **K. M. Saridakis and A. J. Dentsoras**, "Soft computing in engineering design – A review," Advanced Engineering Informatics, vol. 22, no. 2, pp. 202-221, 2008.

[83] **S. J. Ovaska, A. Kamiya, and C. YangQuan**, "Fusion of soft computing and hard computing: computational structures and characteristic features," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 36, no. 3, pp. 439-448, 2006.

[84] **R.-E. Precup, M. L. Tomescu, S. Preitl, J. K. Tar, and A. S. Paul**, "Stability analysis approach for fuzzy logic control systems with Mamdani type fuzzy logic controllers," Journal of Control Engineering and Applied Informatics, vol. 9, no. 1, pp. 3-10, 2007.

[85] **G. Ray, S. K. Das, and B. Tyagi**, "Stabilization of Inverted Pendulum via Fuzzy Control," IE(I) Journal-EL, vol. 88, pp. 58-62, Sept. 2007.

[86] **L. Yanmei, C. Zhen, X. Dingyu, and X. Xinhe**, "Real-time controlling of inverted pendulum by fuzzy logic," in *Automation and Logistics, 2009. ICAL '09. IEEE International Conference on*, 2009, pp. 1180-1183.

[87] **B. K. Bose**, "Neural network applications in power electronics and motor drives—An introduction and perspective," IEEE Transactions on Industrial Electronics, vol. 54, no. 1, pp. 14-33, 2007.

[88] **C.-T. Chen and S.-T. Peng**, "Intelligent process control using neural fuzzy techniques," Journal of Process Control, vol. 9, no. 6, pp. 493-503, 1999.

[89] **L. Aguilar, P. Melin, and O. Castillo**, "Intelligent control of a stepping motor drive using a hybrid neuro-fuzzy ANFIS approach," Applied Soft Computing, vol. 3, no. 3, pp. 209-219, 2003.

[90] **A. Soundarrajan and S. Sumathi**, "Fuzzy-based intelligent controller for power generating systems," Journal of Vibration and Control, vol. 17, no. 8, pp. 1265-1278, July 1, 2011 2011.

[91] **S. H. Hosseini and A. H. Etemadi**, "Adaptive neuro-fuzzy inference system based automatic generation control," Electric Power Systems Research, vol. 78, no. 7, pp. 1230-1239, 2008.

[92] **S. R. Khuntia and S. Panda**, "Simulation study for automatic generation control of a multi-area power system by ANFIS approach," Applied Soft Computing, vol. 12, no. 1, pp. 333-341, 2012.

[93] **H. F. Ho, Y. K. Wong, and A. B. Rad**, "Robust fuzzy tracking control for robotic manipulators," Simulation Modelling Practice and Theory, vol. 15, no. 7, pp. 801-816, 2007.

[94] **P. J. Fleming and R. C. Purshouse**, "Evolutionary algorithms in control systems engineering: a survey," Control Engineering Practice, vol. 10, no. 11, pp. 1223-1241, 2002.

[95] **Y. del Valle, G. K. Venayagamoorthy, S. Mohagheghi, J. C. Hernandez, and R. G. Harley**, "Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems," IEEE Transactions on Evolutionary Computation, vol. 12, no. 2, pp. 171-195, 2008.

[96] **F.-J. Lin, L.-T. Teng, and M.-H. Yu**, "Radial Basis Function Network Control With Improved Particle Swarm Optimization for Induction Generator System," IEEE Transactions on Power Electronics, vol. 23, no. 4, pp. 2157-2169, 2008.

[97] **A. Demiroren and E. Yesil**, "Automatic generation control with fuzzy logic controllers in the power system including SMES units," International Journal of Electrical Power & Energy Systems, vol. 26, no. 4, pp. 291-305, 2004.

[98] **Y. Oysal**, "A comparative study of adaptive load frequency controller designs in a power system with dynamic neural network models," Energy Conversion and Management, vol. 46, no. 15–16, pp. 2656-2668, 2005.

[99] **H. J. Lee, J. B. Park, and Y. H. Joo**, "Robust load-frequency control for uncertain nonlinear power systems: A fuzzy logic approach," Information Sciences, vol. 176, no. 23, pp. 3520-3537, 2006.

[100] **P. Melin and O. Castillo**, "Adaptive intelligent control of aircraft systems with a hybrid approach combining neural networks, fuzzy logic and fractal theory," Applied Soft Computing, vol. 3, no. 4, pp. 353-362, 2003.

[101] **I. F. Chung, C.-J. Lin, and C.-T. Lin**, "A GA-based fuzzy adaptive learning control network," Fuzzy Sets and Systems, vol. 112, no. 1, pp. 65-84, 2000.

[102] **Y.-G. Leu, C.-M. Hong, and H.-J. Zhon**, "GA-Based Adaptive Fuzzy-Neural Control for a Class of MIMO Systems," in *Advances in Neural Networks – ISNN 2007*. vol. 4491, D. Liu, S. Fei, Z.-G. Hou, H. Zhang, and C. Sun, Eds.: Springer Berlin Heidelberg, 2007, pp. 45-53.

[103] **L. Behera, S. Chaudhury, and M. Gopal**, "Neuro-adaptive hybrid controller for robot-manipulator tracking control," Control Theory and Applications, IEE Proceedings -, vol. 143, no. 3, pp. 270-275, 1996.

[104] **P. M. Patre, S. Bhasin, Z. D. Wilcox, and W. E. Dixon**, "Composite Adaptation for Neural Network-Based Controllers," IEEE Transactions on Automatic Control, vol. 55, no. 4, pp. 944-950, 2010.

[105] **Chi-Hsu Wang, Tsung-Chih Lin, Tsu-Tian Lee, and Han-Leih Liu**, "Adaptive hybrid intelligent control for uncertain nonlinear dynamical systems," IEEE Transactions onSystems, Man, and Cybernetics, Part B: Cybernetics, vol. 32, no. 5, pp. 583-597, 2002.

[106] **C. Yeong-Chan**, "Intelligent robust control for uncertain nonlinear time-varying systems and its application to robotic systems," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 35, no. 6, pp. 1108-1119, 2005.

[107] **I. Kar and L. Behera**, "Direct adaptive neural control for affine nonlinear systems," Applied Soft Computing, vol. 9, no. 2, pp. 756-764, 2009.

[108] **K. Nouri, R. Dhaouadi, and N. Benhadj Braiek**, "Adaptive control of a nonlinear dc motor drive using recurrent neural networks," Applied Soft Computing, vol. 8, no. 1, pp. 371-382, 2008.

[109] **S. J. Yoo, J. B. Park, and Y. H. Choi**, "Indirect adaptive control of nonlinear dynamic systems using self recurrent wavelet neural networks via adaptive learning rates," Information Sciences, vol. 177, no. 15, pp. 3074-3098, 2007.

[110] **W. G. Seo, B. H. Park, and J. S. Lee**, "Intelligent learning control for a class of nonlinear dynamic systems," Control Theory and Applications, IEE Proceedings -, vol. 146, no. 2, pp. 165-170, 1999.

[111] **Y.-S. Yang and X.-F. Wang**, "Adaptive H∞ tracking control for a class of uncertain nonlinear systems using radial-basis-function neural networks," Neurocomputing, vol. 70, no. 4–6, pp. 932-941, 2007.

[112] **L. Cheng, Z.-G. Hou, and M. Tan**, "Adaptive neural network tracking control for manipulators with uncertain kinematics, dynamics and actuator model," Automatica, vol. 45, no. 10, pp. 2312-2318, 2009.

[113] **W. Liu, G. K. Venayagamoorthy, and D. C. Wunsch II**, "Design of an adaptive neural network based power system stabilizer," Neural Networks, vol. 16, no. 5, pp. 891-898, 2003.

[114] **S.-C. Tong and Y.-M. Li**, "Adaptive backstepping output feedback control for SISO nonlinear system using fuzzy neural networks," International Journal of Automation and Computing, vol. 6, no. 2, pp. 145-153, 2009/05/01 2009.

[115] **X. Gong, Z.-C. Hou, C.-J. Zhao, Y. Bai, and Y.-T. Tian**, "Adaptive backstepping sliding mode trajectory tracking control for a quad-rotor," International Journal of Automation and Computing, vol. 9, no. 5, pp. 555-560, 2012/10/01 2012.

[116] **G. K. Venayagamoorthy and R. G. Harley**, "Intelligent optimal control of excitation and turbine systems in power networks," in *IEEE Power Engineering Society General Meeting*, 2006, pp. 1-8.

[117] **Y. Becerikli, A. F. Konar, and T. Samad**, "Intelligent optimal control with dynamic neural networks," Neural Networks, vol. 16, no. 2, pp. 251-259, 2003.

[118] **J. V. da Fonseca Neto, I. S. Abreu, and F. N. da Silva**, "Neural–Genetic Synthesis for State-Space Controllers Based on Linear Quadratic Regulator Design for Eigenstructure Assignment," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 40, no. 2, pp. 266-285, 2010.

[119] **T. Cheng, F. L. Lewis, and M. Abu-Khalaf**, "A neural network solution for fixed-final time optimal control of nonlinear systems," Automatica, vol. 43, no. 3, pp. 482-490, 2007.

[120] **C.-S. Chen**, "Quadratic optimal neural fuzzy control for synchronization of uncertain chaotic systems," Expert Systems with Applications, vol. 36, no. 9, pp. 11827-11835, 2009.

[121] **Y. H. Kim, F. L. Lewis, and D. M. Dawson**, "Intelligent optimal control of robotic manipulators using neural networks," Automatica, vol. 36, no. 9, pp. 1355-1364, 2000.

[122] **N. S. Bhuvaneswari, G. Uma, and T. R. Rangaswamy**, "Adaptive and optimal control of a non-linear process using intelligent controllers," Applied Soft Computing, vol. 9, no. 1, pp. 182-190, 2009.

[123] **D. Vrabie, F. Lewis, and D. Levine**, "Neural Network-Based Adaptive Optimal Controller – A Continuous-Time Formulation," in *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*. vol. 15, D.-S. Huang, D. Wunsch, II, D. Levine, and K.-H. Jo, Eds.: Springer Berlin Heidelberg, 2008, pp. 276-285.

[124] **F. L. Lewis and D. Vrabie**, "Reinforcement learning and adaptive dynamic programming for feedback control," IEEE Circuits and Systems Magazine, vol. 9, no. 3, pp. 32-50, 2009.

[125] **F. L. Lewis and K. G. Vamvoudakis**, "Optimal adaptive control for unknown systems using output feedback by reinforcement learning methods," in *8th IEEE International Conference on Control and Automation (ICCA)*, 2010, pp. 2138-2145.

[126] **W. C. Wong and J. H. Lee**, "A reinforcement learning-based scheme for direct adaptive optimal control of linear stochastic systems," Optimal Control Applications and Methods, vol. 31, no. 4, pp. 365-374, 2010.

[127] **J. J. Murray, C. J. Cox, G. G. Lendaris, and R. Saeks**, "Adaptive dynamic programming," IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 32, no. 2, pp. 140-153, 2002.

[128] **F.-Y. Wang, H. Zhang, and D. Liu**, "Adaptive Dynamic Programming: An Introduction," IEEE Computational Intelligence Magazine, vol. 4, no. 2, pp. 39-47, 2009.

[129] **Z. Huang, J. Ma, and H. Huang**, "An approximate dynamic programming method for multi-input multi-output nonlinear system," Optimal Control Applications and Methods, vol. 34, no. 1, pp. 80-95, 2013.

[130] **S. Ferrari and R. F. Stengel**, "An adaptive critic global controller," in *Proceedings of the 2002 American Control Conference*, 2002, pp. 2665-2670 vol.4.

[131] **D. P. Bertsekas**, "Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC," European Journal of Control, vol. 11, no. 4–5, pp. 310-334, 2005.

[132] **D. Vrabie, M. Abu-Khalaf, F. L. Lewis, and W. Youyi**, "Continuous-Time ADP for Linear Systems with Partially Unknown Dynamics," in *Approximate Dynamic Programming and Reinforcement Learning, 2007. ADPRL 2007. IEEE International Symposium on*, 2007, pp. 247-253.

[133] **A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf**, "Discrete-Time Nonlinear HJB Solution Using Approximate Dynamic Programming: Convergence Proof," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 38, no. 4, pp. 943-949, 2008.

[134] **J. M. Lee and J. H. Lee**, "An approximate dynamic programming based approach to dual adaptive control," Journal of Process Control, vol. 19, no. 5, pp. 859-864, 2009.

[135] **J. H. Lee and W. Wong**, "Approximate dynamic programming approach for process control," Journal of Process Control, vol. 20, no. 9, pp. 1038-1048, 2010.

[136] **H. Nosair, Y. Yang, and J. M. Lee**, "Min–max control using parametric approximate dynamic programming," Control Engineering Practice, vol. 18, no. 2, pp. 190-197, 2010.

[137] **S. Ferrari, J. E. Steck, and R. Chandramohan**, "Adaptive Feedback Control by Constrained Approximate Dynamic Programming," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 38, no. 4, pp. 982-987, 2008.

[138] **G. K. Venayagamoorthy, R. G. Harley, and D. C. Wunsch**, "Comparison of heuristic dynamic programming and dual heuristic programming adaptive critics for neurocontrol of a turbogenerator," IEEE Transactions on Neural Networks, vol. 13, no. 3, pp. 764-773, 2002.

[139] **J.-W. Park, R. G. Harley, G. K. Venayagamoorthy, and G. Jang**, "Dual heuristic programming based nonlinear optimal control for a synchronous generator," Engineering Applications of Artificial Intelligence, vol. 21, no. 1, pp. 97-105, 2008.

[140] **Ding Wang, Derong Liu, and Qinglai Wei**, "Finite-horizon neuro-optimal tracking control for a class of discrete-time nonlinear systems using adaptive dynamic programming approach," Neurocomputing, vol. 78, no. 1, pp. 14-22, 2012.

[141] **Z.-P. Jiang and Y. Jiang**, "Robust adaptive dynamic programming for linear and nonlinear systems: An overview," European Journal of Control, vol. 19, no. 5, pp. 417-425, 2013.

[142] **D. V. Prokhorov, R. A. Santiago, and D. C. Wunsch II**, "Adaptive critic designs: A case study for neurocontrol," Neural Networks, vol. 8, no. 9, pp. 1367-1372, 1995.

[143] **D. V. Prokhorov and D. C. Wunsch-II**, "Adaptive critic designs," IEEE Trans. on Neural Nets., vol. 8, no. 5, pp. 997-1007, September 1997.

[144] **T. Hanselmann, L. Noakes, and A. Zaknich**, "Continuous-time adaptive critics," IEEE Trans. on Neural Nets., vol. 18, no. 3, pp. 631-647, 2007.

[145] **Silvia Ferrari and R. F. Stengel**, "Online adaptive critic flight control," Journal of Guidance, Control, and Dynamics, vol. 27, no. 5, September-October 2004.

[146] **S. Ferrari and R. F. Stengel**, "Model based adaptive critic designs," in *Handbook of learning and Approximate Dynamic Programming*, Jennie Si, A. G. Barto, W. B. Powell, and D. W. II, Eds.: IEEE Press, 2004, pp. 64-94.

[147] **A. K. Deb, Jayadeva, M. Gopal, and S. Chandra**, "SVM-Based Tree-Type Neural Networks as a Critic in Adaptive Critic Designs for Control," Neural Networks, IEEE Transactions on, vol. 18, no. 4, pp. 1016-1030, 2007.

[148] **C.-K. Lin**, "Radial basis function neural network-based adaptive critic control of induction motors," Applied Soft Computing, vol. 11, no. 3, pp. 3066-3074, 2011.

[149] **R. V. Kulkarni and G. K. Venayagamoorthy**, "Adaptive critics for dynamic optimization," Neural Networks, vol. 23, no. 5, pp. 587-591, 2010.

[150] **S. Mohagheghi, G. K. Venayagamoorthy, and R. G. Harley**, "Adaptive Critic Design Based Neuro-Fuzzy Controller for a Static Compensator in a Multimachine Power System," IEEE Transactions on Power Systems, vol. 21, no. 4, pp. 1744-1754, 2006.

[151] **S. Bhasin, N. Sharma, P. M. Patre, and W. E. Dixon**, "Robust asymptotic tracking of a class of nonlinear systems using an adaptive critic based controller," in *American Control Conference (ACC), 2010*, 2010, pp. 3223-3228.

[152] **S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon**, "A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems," Automatica, vol. 49, no. 1, pp. 82-92, 2013.

[153] **L. Jiaqi, R. G. Harley, and G. K. Venayagamoorthy**, "Adaptive critic design based dynamic optimal power flow controller for a smart grid," in *2011 IEEE Symposium on Computational Intelligence Applications In Smart Grid (CIASG)*, 2011, pp. 1-8.

[154] **D. Vrabie, K. Vamvoudakis, and F. Lewis**, "Adaptive optimal controllers based on Generalized Policy Iteration in a continuous-time framework," in *17th Mediterranean Conference on Control and Automation, 2009. MED '09* 2009, pp. 1402-1409.

[155] **D. Vrabie, O. Pastravanu, and F. L. Lewis**, "Policy iteration for continuous-time systems with unknown internal dynamics," in *Mediterranean Conference on Control & Automation, 2007. MED '07.* , 2007, pp. 1-6.

[156] **D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis**, "Adaptive optimal control for continuous-time linear systems based on policy iteration," Automatica, vol. 45, no. 2, pp. 477-484, 2009.

[157] **K. Vamvoudakis, D. Vrabie, and F. Lewis**, "Online policy iteration based algorithms to solve the continuous-time infinite horizon optimal control problem," in *IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning, 2009. ADPRL '09.*, 2009, pp. 36-41.

[158] **K. G. Vamvoudakis and F. L. Lewis**, "Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem," Automatica, vol. 46, no. 5, pp. 878-888, 2010.

[159] **D. Vrabie and F. L. Lewis**, "Adaptive optimal control algorithm for continuous-time nonlinear systems based on policy iteration," in *47th IEEE Conf. on Decision and Control*, Cancun, Mexico, 2008, pp. 73-79.

[160] **D. Vrabie and F. Lewis**, "Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems," Neural Networks, vol. 22, no. 3, pp. 237-246, 2009.

[161] **K. G. Vamvoudakis and F. L. Lewis**, "Online Synchronous Policy Iteration Method for Optimal Control," in *Recent Advances in Intelligent Control Systems*, W. Yu, Ed.: Springer London, 2009, pp. 357-374.

[162] **K. G. Vamvoudakis and F. L. Lewis**, "Online solution of nonlinear two-player zero-sum games using synchronous policy iteration," International Journal of Robust and Nonlinear Control, vol. 22, no. 13, pp. 1460-1483, 2012.

[163] **H. Modares, M.-B. Naghibi Sistani, and F. L. Lewis**, "A policy iteration approach to online optimal control of continuous-time constrained-input systems," ISA Transactions, vol. 52, no. 5, pp. 611-621, 2013.

[164] **S. Bhasin, M. Johnson, and W. E. Dixon**, "A model-free robust policy iteration algorithm for optimal control of nonlinear systems," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 3060-3065.

[165] **M. Johnson, S. Bhasin, and W. E. Dixon**, "Nonlinear two-player zero-sum game approximate solution using a Policy Iteration algorithm," in *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 2011, pp. 142-147.

[166] **B. Luo and H.-N. Wu**, "Online policy iteration algorithm for optimal control of linear hyperbolic PDE systems," Journal of Process Control, vol. 22, no. 7, pp. 1161-1170, 2012.

[167] **T. Dierks and S. Jagannathan**, "Online Optimal Control of Affine Nonlinear Discrete-Time Systems With Unknown Internal Dynamics by Using Time-Based Policy Update," IEEE Transactions on Neural Networks and Learning Systems, vol. 23, no. 7, pp. 1118-1129, 2012.

[168] **H. Li and D. Liu**, "Optimal control for discrete-time affine non-linear systems using general value iteration," IET Control Theory & Applications, vol. 6, no. 18, pp. 2725-2736, 2012.

[169] **J. Y. Lee, J. B. Park, and Y. H. Choi**, "Integral Q-learning and explorized policy iteration for adaptive optimal control of continuous-time linear systems," Automatica, vol. 48, no. 11, pp. 2850-2859, 2012.

[170] **R. Padhi, S. N. Balakrishnan, and T. Randolph**, "Adaptive-critic based optimal neuro control synthesis for distributed parameter systems," Automatica, vol. 37, no. 8, pp. 1223-1234, 2001.

[171] **R. Padhi and S. N. Balakrishnan**, "Development and analysis of a feedback treatment strategy for parturient paresis of cows," IEEE Transactions on Control Systems Technology, vol. 12, no. 1, pp. 52-64, 2004.

[172] **R. Padhi, N. Unnikrishnan, and S. N. Balakrishnan**, "Optimal control synthesis of a class of nonlinear systems using single network adaptive critics," in *Proceedings of the 2004 American Control Conference*, 2004, pp. 1592-1597 vol.2.

[173] **Radhakant Padhi, N. Unnikrishnan, X. Wang, and S. N. Balakrishnan**, "A single network adaptive critic (SNAC) architecture for optimal control synthesis for a class of nonlinear systems," Neural Networks, vol. 19, pp. 1648-1660, 2006.

[174] **S. Kumar, R. Padhi, and L. Behera**, "Direct Adaptive Control using Single Network Adaptive Critic," in *IEEE International Conference on System of Systems Engineering, 2007. SoSE '07*, 2007, pp. 1-6.

[175] **G. Gurrala, R. Padhi, and I. Sen**, "Power System Stability Enhancement by Single Network Adaptive Critic Stabilizers," in *International Joint Conference on Neural Networks, 2009, IJCNN*, 2009, pp. 1061-1066.

[176] **G. Gurrala, I. Sen, and R. Padhi**, "Single network adaptive critic design for power system stabilisers," IET Generation, Transmission & Distribution, vol. 3, no. 9, pp. 850-858, 2009.

[177] **S. F. Ali and R. Padhi**, "Optimal blood glucose regulation of diabetic patients using single network adaptive critics," Optimal Control Applications and Methods, vol. 32, no. 2, pp. 196-214, 2011.

[178] **K. P. Prem, L. Behera, N. H. Siddique, and G. Prasad**, "A T-S fuzzy based adaptive critic for continuous-time input affine nonlinear systems," in *IEEE International Conference on Systems, Man and Cybernetics, 2009. SMC 2009.*, 2009, pp. 4329-4334.

[179] **K. J. Åström and K. Furuta**, "Swinging up a pendulum by energy control," Automatica, vol. 36, no. 2, pp. 287-295, 2000.

[180] **D. Chatterjee, A. Patra, and H. K. Joglekar**, "Swing-up and stabilization of a cart–pendulum system under restricted cart track length," Systems & Control Letters, vol. 47, no. 4, pp. 355-364, 2002.

[181] **C. Ibañez, O. G. Frias, and M. S. Castañón**, "Lyapunov-Based Controller for the Inverted Pendulum Cart System," Nonlinear Dynamics, vol. 40, no. 4, pp. 367-374, 2005/06/01 2005.

[182] **O. I. Elgerd**, *Electric Energy Systems Theory: An Introduction*, 2 ed. New Delhi: Tata McGraw-Hill Publishing Company Ltd., 2004.

[183] **P. Kundur**, *Power System Stability and Control*, 1 ed. New Delhi: Tata McGraw-Hill Education Pvt. Ltd., 2011.

[184] **D. P. Kothari and I. J. Nagrath**, *Modern Power System Analysis*. New Delhi: Tata McGraw-Hill Publishing Company Ltd., 2003.

[185] **A. Chakrabarty and S. Halder**, *Power System Analysis: Operation and Control*, 3rd ed. New Delhi: PHI Learning Pvt. Ltd. , 2010.

[186] **Hassan Bevrani and Takashi Hiyama**, *Intelligent Automatic Generation Control*. Boca Raton: CRC Press, Taylor & Francis Group LLC, 2011.

[187] **I. Ibraheem, P. Kumar, and D. P. Kothari**, "Recent philosophies of automatic generation control strategies in power systems," Power Systems, IEEE Transactions on, vol. 20, no. 1, pp. 346-357, 2005.

[188] **T. Wen, S. Liying, and X. Zhan**, "Robust analysis and design of load frequency controller for power systems," in *Control Applications, 2008. CCA 2008. IEEE International Conference on*, 2008, pp. 329-334.

[189] **Y. Wang, R. Zhou, and C. Wen**, "Robust load-frequency controller design for power systems," IEE Proc.-C*,* vol. 140(1), pp. 11-16, January 1993.

[190] **Y. Wang, R. Zhou, and C. Wen**, "New robust adaptive load-frequency control with system parametric uncertainties," IEE Proc.-Gener. Transm. Distrib.*,* vol. 141(3), pp. 184-190, May 1994.

[191] **M. H. Kazemi, M. Karrari, and M. B. Menhaj**, "Decentralized robust adaptive-output feedback controller for power system load frequency control," Electrical Engineering*,* vol. 84, no. 2, pp. 75-83, 2002/05/01 2002.

[192] **M. Rahmani and N. Sadati**, "Hierarchical optimal robust load-frequency control for power systems," Generation, Transmission & Distribution, IET*,* vol. 6, no. 4, pp. 303-312, 2012.

[193] **G. Ray, A. Prasad, and T. Bhattacharyya**, "Design of decentralized robust load-frequency controller based on SVD method," Computers & Electrical Engineering*,* vol. 25, no. 6, pp. 477-492, 1999.

[194] **G. Ray, S. Dey, and T. K. Bhattacharyya**, "Multi-area load frequency control of power systems: A decentralized variable structure approach," Electric Power Components and Systems*,* vol. 33, no. 3, pp. 315-331, 2004.

[195] **S. Patra, S. Sen, and G. Ray**, "Design of Robust Load Frequency Controller: H ∞ Loop Shaping Approach," Electric Power Components and Systems*,* vol. 35, no. 7, pp. 769-783, 2007/04/27 2007.

[196] **R. Dey, S. Ghosh, and G. Ray**, "A robust H∞ load-frequency controller design using LMIs," in *Control Applications,(CCA) & Intelligent Control,(ISIC), 2009 IEEE*, 2009, pp. 1501-1504.

[197] **R. Dey, S. Ghosh, G. Ray, and A. Rakshit**, "H∞ load frequency control of interconnected power systems with communication delays," International Journal of Electrical Power & Energy Systems*,* vol. 42, no. 1, pp. 672-684, 2012.

[198] **S. Patra, S. Sen, and G. Ray**, "Load Frequency Control of Interconnected Power Systems via Low-Order H∞ Loop Shaping Controller," International Journal of Power and Energy Systems*,* vol. 32, no. 4, pp. 167, 2012.

[199] **H. Razmi, H. A. Shayanfar, and M. Teshnehlab**, "Steady state voltage stability with AVR voltage constraints," International Journal of Electrical Power & Energy Systems*,* vol. 43, no. 1, pp. 650-659, 2012.

[200] **S. Ayasun and A. Gelen**, "Stability analysis of a generator excitation control system with time delays," Electrical Engineering*,* vol. 91, no. 6, pp. 347-355, 2010/01/01 2010.

[201] **J. M. Kim and S. I. Moon**, "A study on a new AVR parameter tuning concept using on-line measured data with the real-time simulator," European Transactions on Electrical Power, vol. 16, no. 3, pp. 235-246, 2006.

[202] **S. S. Lee and J. K. Park**, "Design of power system stabilizer using observer/sliding mode, observer/sliding mode-model following and H∞/sliding mode controllers for small-signal stability study," International Journal of Electrical Power & Energy Systems, vol. 20, no. 8, pp. 543-553, 1998.

[203] **M. Jalili and M. Yazdanpanah**, "Transient stability enhancement of power systems via optimal nonlinear state feedback control," Electrical Engineering, vol. 89, no. 2, pp. 149-156, 2006.

[204] "DC Motor Control," in *MATLAB Control System Toolbox*: Mathworks Inc., 2011.

[205] **----**, "Band limited white noise," in *MATLAB-SIMULINK help*: Mathworks Inc.

[206] "DC Motor Control. ," in *MATLAB Control System Toolbox User' Guide*: Mathworks Inc.

[207] **K. G. Vamvoudakis and F. L. Lewis**, "Online actor critic algorithm to solve the continuous-time infinite horizon optimal control problem," in *International Joint Conference on Neural Networks, 2009, IJCNN 2009*, 2009, pp. 3180-3187.

APPENDIX – A CONTINUOUS-TIME SMOOTH FUNCTIONS

This appendix discusses in brief about the continuous-time smooth functions which are taken as assumption on nonlear dynamic systems in proofs of convergence of PI algorithms.

A.1 SMOOTH FUNCTION

A continuous function is called as smooth function if its differentives of all orders exist. According to the properties of the derivatives of functions the differentiability class of a function is defined. The differentiability class is a classification of functions. The higher order differentiability classes correspond to the existence of more derivatives (Refer Wikipedia).

Let a function is defined on open set of real number having real values. Consider *k* be a nonnegative integer. The function f belongs to differentiability class C^k if its derivatives, $f^{(1)}$, $f^{(2)}$,.... $f^{(k)}$ exist and are continuous. The continuity is automatic for all the derivatives except for $f^{(k)}$. The function f belongs to the differentiability class C^∞, or smooth, if its derivatives of all orders exist. The function f belongs to the differentiability class C^ω, or analytic, if f is smooth and if it equals its Taylor series expansion around any point in its domain. Thus,

1. Class C^0 consists of all continuous functions.
2. Class C^1 consists of all differentiable functions whose derivative is continuous. Thus, C^1 function is exactly a function whose derivative exists and is of class C^0.
3. Class C^k for any positive integer *k*, be the set of all differentiable functions whose derivative is in C^{k-1}. In particular, C^k is contained in C^{k-1}.

A.2 EXAMPLES

Consider a function defined as

$$f(x)=\begin{cases} x & if\ x\geq 0 \\ 0 & if\ x<0 \end{cases} \tag{A.1}$$

This function is continuous but not differential at $x=0$, therefore it belongs to class C^0 but does not belong to class C^1.

Consider another function defined as

$$f(x)=\begin{cases} x^2\sin\left(\dfrac{1}{x}\right) & if\ x\neq 0 \\ 0 & if\ x=0 \end{cases} \tag{A.2}$$

This function is differentiable, with derivative

$$f'(x)=\begin{cases}-\cos\left(\frac{1}{x}\right)+2x\sin\left(\frac{1}{x}\right) & if\ x\neq 0\\ 0 & if\ x=0\end{cases} \qquad \text{(A.3)}$$

Because $\cos\left(\frac{1}{x}\right)$ oscillates as $x\to 0$, $f'(x)$ is not continuous at zero. Therefore, this function is differentiable but not of class C^1.

For further detail refer smooth function in Wikipedia.

APPENDIX – B LIPSCHITZ CONTINUITY

This appendix discusses in brief about the Lipschitz continuity which is taken as an assumption on nonlinear dynamic systems in proofs of convergence of PI algorithms.

Lipschitz continuity is a strong form of uniform continuity for functions. Intuitively, a Lipschitz continuous function is limited in how fast it can change. There exists a definite real number such that, for every pair of points on the graph of this function, the absolute value of the slope of the line connecting them is not greater than this real number; this bound is called the function's "Lipschitz constant" (or "modulus of uniform continuity"). In differential equation theory, Lipschitz continuity is the central condition of the Picard–Lindelöf theorem which guarantees the existence and uniqueness of the solution to an initial value problem. A special type of Lipschitz continuity, called contraction, is used in the Banach fixed point theorem (Refer Lipschitz continuity in Wikipedia).

Lipschitz continuity is defined as: Given two metric spaces (X, d_X) and (Y, d_Y), where d_X denotes the metric on the set *X* and d_Y is the metric on set *Y* (for example, *Y* might be the set of real numbers R with the metric $dY(x, y) = |x - y|$ and X might be a subset of R), a function $f : X \to Y$ is called Lipschitz continuous if there exists a real constant $K \geq 0$ such that, for all x_1 and x_2 in *X*,

$$d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2) \tag{B.1}$$

Any such *K* is referred to as a Lipschitz constant for the function *f*. The smallest constant is sometimes called the (best) Lipschitz constant; however in most cases the latter notion is less relevant. If $K = 1$ the function is called a short map, and if $0 \leq K < 1$ the function is called a contraction.

The inequality is (trivially) satisfied if $x_1 = x_2$. Otherwise, one can equivalently define a function to be Lipschitz continuous if and only if there exists a constant $K \geq 0$ such that, for all $x_1 \neq x_2$,

$$\frac{d_Y(f(x_1), f(x_2))}{d_X(x_1, x_2)} \leq K \tag{B.2}$$

For real-valued functions of several real variables, this satisfies if and only if the absolute value of the slopes of all secant lines are bounded by *K*. The set of lines of slope *K* passing through a point on the graph of the function forms a circular cone, and a function is Lipschitz if and only if the graph of the function everywhere lies completely outside of this cone.

A function is called locally Lipschitz continuous if for every x in *X* there exists a neighborhood U of x such that f restricted to U is Lipschitz continuous. Equivalently, if X is

a locally compact metric space, then f is locally Lipschitz if and only if it is Lipschitz continuous on every compact subset of X. In spaces that are not locally compact, this is a necessary but not a sufficient condition.

For further detail refer Lipschitz continuity in Wikipedia.