ASPECTS OF INTUITIVE CONTROL FRAMEWORK: STABILIZE, OPTIMIZE, AND IDENTIFY

by

PAVAN KUMAR NUTHI

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2015

To my loving parents Neeraja and Vigneswara Rao for their support.

To Mr. Bertrand Russel for his 'teapot'.

ACKNOWLEDGEMENTS

I would like to thank my supervising professor Dr. Kamesh Subbarao for his invaluable advice and guidance. I am grateful for his role as my mentor throughout my doctoral research.

I have had the good fortune of honing my skills and knowledge under the tutelage of great teachers . I thank them all for inspiring my learning process. In particular, I wish to thank Dr. Frank Lewis, Dr. Atilla Dogan, Dr. Kent Lawrence, Dr. Gaik Ambartsoumian for their interest in my research and for taking time to serve in my dissertation committee.

My interest in the scientific method and engineering disciplines is spurred by hours spent helping my Grandfather in his machine shop. As an optical scientist, he showed me that application of theoretical knowledge to engineering practices can be a rewarding experience. I am grateful to him for getting me started on this path and my identity as an engineer.

I am thankful to my family for being supportive of my decisions throughout my life. Mom and Dad, thank you for being patient and for your faith in my choices. To my little sister, thank you for being my partner in crime when I needed one.

A man is a sum total of his experiences. I am thankful to all my friends and acquaintances for shaping these experiences. I am grateful for all the laughs we shared, and discussions that broadened my horizons. To my labmates, thank you for all the ideas we shared, and all the good times. Especially Alok and Gus, thank you for entertaining your delightfully quirky housemate.

Most of all, I am grateful to my heroes from all walks of life for shaping my identity as a philosophical skeptic.

April 13, 2015

ABSTRACT

ASPECTS OF INTUITIVE CONTROL FRAMEWORK: STABILIZE, OPTIMIZE, AND IDENTIFY

PAVAN KUMAR NUTHI, Ph.D.

The University of Texas at Arlington, 2015

Supervising Professor: Kamesh Subbarao

The duality of estimation and control problems is a well known fact in control theory literature. Parameter convergence and closed loop stability are usually competing interests for a given control scheme. This motivates identification routines to be performed only in offline experiments. On the other hand stable controllers do not guarantee parameter convergence to true parameters. Thus there is a need for a higher level abstraction for a control scheme which acts in stages and prioritizes various aspects at different stages.

The stage abstraction for controller is inspired by human intuition towards dealing with control and identification simultaneously and hence named Intuitive control framework. The first stage prioritizes stabilization of states only. The controller moves onto the next stage after the unknown system is stabilized. The subsequent stages involves optimization with different performance metrics through adaptive learning. After enough information for identification is acquired, the control schemes developed for various optimal metrics are used to estimate the unknown parameters in the final stage. This narrative for selective prioritization of objectives and a higher

level abstraction for control schemes is illustrated for a continuous linear time invariant state space realization with state feedback. Numerous real-world applications can benefit from this online system identification routine inspired by the human cognitive process. This offers a seamless integration of control and identification with a higher level of priorities. Such framework is presented with explicit formulations for certain classes of dynamic systems, and evaluated with computer simulations as well as experimental results. Further computation of forward reachable sets after identification also offers the only way to perform such computation for an unknown system without the need for experimentation. Identified reachable sets are also presented with a discussion on their accuracy.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Motivation

Performance of man made control systems have often been compared with control mechanisms found in nature. Conversely numerous instances of bio-inspired control have been documented in [1, 2]. In humanoid walking algorithms, we can often see control schemes which try to mimic humans to emulate biped walking as evident in [3]. Although humans are a highly intelligent species on earth capable of articulated form of communication, almost all living organisms with or without perceivable consciousness exhibit intelligence at some level.

An interpretation of such cognitive process for control is outlined in Fig. [1.1]. An example can be the cognitive process behind riding a bicycle. First block shows that stabilization is the first response to an unfamiliar control situation. During this stage, stability is the only priority for the rider, irrespective of performance. After a certain stable strategy is developed, performance is prioritized in optimization stage. The rider progresses to ride the bicycle in a straight line, then in a circle *etc.* The experiences from optimization stage are distilled into a model during the identification stage. Upon learning the optimal control strategies, the rider develops a model for riding a bicycle. After this point if the rider needs to ride even a different bicycle, he already possesses a reasonable model, hence the expression goes 'like riding a bike'. Eventually a higher understanding of the control process can be developed as outlined in the final stage of awareness. This stage is an assessment of the limits and

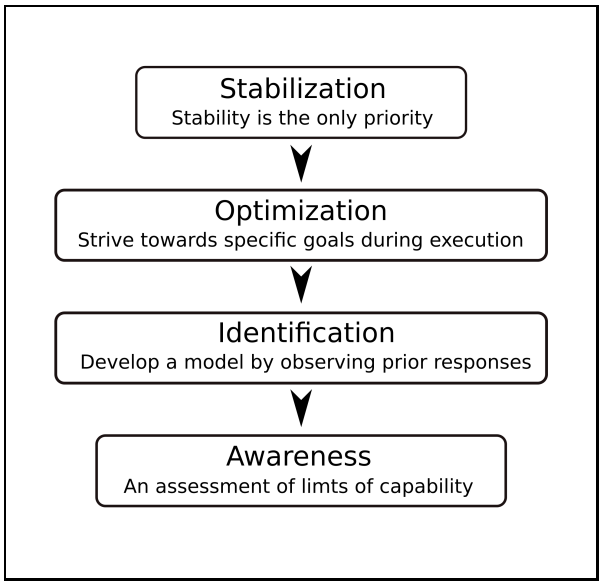

Figure 1.1. An interpretation of cognitive control process.

capabilities of the controller. Such interpretation of a cognitive control process can be realized using tools and methods from control theory.

One of the key aspects seen in nature is the concept of optimization. For instance, A flock of birds fly in formation to conserve collective energy expended by the flock. Most birds have a certain clap and flying pattern to their flapping motion in order to conserve energy. Any swimmer can understand the rationale behind this clap and fling rather than a constant pitch flapping. Optimal control theory postulated in [4, 5] abstracts this concept with a mathematical rigor, and provides control solutions which can minimize control energy among other criterion.

Another important aspect of natural reflexes is the adaptability or learning. This aspect allows for online assessment of the situation and reconfiguration of the control law accordingly. Adaptive control theory postulated in [6, 7] captures this aspect using parametric uncertainties and unmodeled dynamics. Most common flavor of the adaptive framework parameterizes the control law as a linear state feedback to deal with uncertainties. There have been adaptive approaches shown in [8] which

reconfigure the controller complexity using newly learnt information from Artificial Neural Networks.

There have been numerous contributions in the field of adaptive optimal control which combine the optimality aspect with the robustness to uncertainties from adaptive control. This branch is increasingly referred to as intelligent control. [9] discusses adaptive optimal controller implemented in a reinforcement learning framework. Whereas, a policy iteration based learning algorithm is proposed in [10] which approximates the optimal control policy with unknown internal dynamics. This work would be a natural extension to the policy iteration in [10] by the expulsion of initial stabilizing policy. System identification as a byproduct of adaptive optimal control is akin to the natural process of learning.

1.2 Background

1.2.1 Adaptive Dynamic Programming

A major group of adaptive optimal control tools are provided by methods collectively known as Adaptive Dynamic Programming. ADP has its origins in Heuristic Dynamic Programming where Reinforcement learning tools have been used to learn the optimal control strategy for stochastic models. The term dynamic programming refers to the Bellman equation which is at the root of these methods. [9] discusses adaptive optimal controller implemented in a reinforcement learning framework for a deterministic model. The backwards in time fixed point Bellman equation is solved so that the optimal control policy can be approximated using *forward-in-time* methods. This framework involves an actor and a critic which can interact with each other. The critic evaluates the performance of control policy, whereas the actor changes the control policy according to the critic's evaluation. Adaptive Dynamic Programming

methods are applied to both continuous and discrete time applications. For the sake of following discussion only continuous time systems are considered.

Policy iteration as described in [11, 12] is a change in control law which is evaluated over a period of time. This framework learns the optimal feedback controller without the knowledge of model parameters. A policy iteration based learning algorithm is proposed in [10] which approximates the optimal control policy with unknown internal dynamics. The policy iteration scheme implemented is equivalent to the Kleinman iteration from [13] which requires an initial stabilizing gain.

Another ADP method is online value iteration where the parameterized value function is iterated to yield optimal controller. Although [14] eliminates the need for an initial stabilizing gain by an online value iteration scheme, there is no known proof of convergence to the best of author's knowledge. [15,16] further analyze the proposed value iteration scheme from [14]. It has been proved that if the Value Iteration scheme converges, it converges to the optimal control policy.

Note that policy iteration hinges on solving a Lyapunov equation whenever the policy is iterated. Whereas value iteration is only a recursion equation. Both techniques rely on an assumption that there is enough information in the measurements. [17] combines both techniques to describe the developments in Adaptive Dynamic Programming and gives a comprehensive account of the field. [18] describes an iterative algorithm similar to Value Iteration known as Computational Adaptive control to estimate the optimal feedback controller without the knowledge of system parameters (for a stable system). All these methods are collectively called Adaptive Dynamic Programming methods which provide a measurement based algorithm to converge towards an optimal controller in spite of unknown system parameters.

1.2.2 Online System Identification

Determination of a model which fits the input and output measurements of a process is essentially system identification. A majority of early work in system identification revolved around discrete time models, since the measurements are collected discretely in real world applications. The discussion in this document is restricted to only deterministic models. Earlier impetus for System Identification came from statistics and econometrics community. Majority of work in this area is focused on linear models and linear approximations to nonlinear models. When the discussion is restricted to linear models, one of the major divides in identification methods is between frequency domain and time domain techniques. It has been shown that these methods are complementary rather than rivaling. [19] settled the debate at the time. [20], [21] outline identification design and discuss the importance of experiment design and choice of model structure.

[22] is a seminal work in system identification which outlines the determination of minimal state space representation from impulse response data for a deterministic problem. This paper is the basis of developments in subspace identification. This paper solved the state space realization problem for the first time using tools such as the Hankel Matrix.

[23] is another seminal paper from the same era which paved the way for estimating parameters by minimizing prediction error. This paper introduced maximum likelihood methods to approximate parameters from ARMA (Auto regressive moving average) models.

[24] is a more recent comprehensive survey of developments in System Identification. It also discusses various branches of system identification and has chapters dedicated to topics like closed loop identification and frequency domain approaches.

A major contribution to the identification community is summarized in [25] with a survey on identification methods and discussion on identifiability of system parameters in closed loop. This paper established that closed loop identification of processes is not just theoretically possible but not necessarily inferior to open loop experiments. This marked the beginning of system identification using closed loop data.

Control is the motivation for most of the applications to identification. System identification gives the description of a model within certain variance which makes design of a controller more tractable. Robust control framework is effective when uncertainty in the knowledge of the model is characterized. Both these theories were developed by two different research communities, but invariably depend on each other. Adaptive control framework either direct or indirect eliminates the necessity for a-priori knowledge of system model. This relation between control and identification led to combined methods in identification which guarantee robustness margins.

[26–28] were some of the first works to propose an iterative control and identification scheme although all the results were derived for discrete time input output models only.

Assuming the measurement of states is available in real-time, the linear model for the continuous time process can be approximated as outlined in [7, 29]. This involves a *series-parallel model* which ensures closed loop stability the system and guarantees boundedness of errors in system parameter estimates $\mathbf{A}, \mathbf{B}$. It is however imperative that the plant be stable in order to guarantee the identification goal. This is true for both MRAC (Model Reference Adaptive Control) as well as APPC (Adaptive Pole Placement Control). [30] presents a new class of adaptive control schemes with stronger convergence properties compared to the traditional adaptive controllers in the presence of over parametrization. In all the adaptive control schemes

it is observed that a trade off between closed loop performance and system parameter convergence is inevitable.

1.2.3 Reachable sets

Computation of reachable and safe sets for dynamical systems has very far-reaching applications in fields such as robotics and air-traffic management. Active control decisions can be made at a higher level using the reachable and safe set data. The question of feasibility is answered through these solutions which makes a number of path planning decisions easier, and a majority of unfeasible control algorithms trivially useless. But the calculation of such solutions is far from trivial even if the dynamics are completely known. [31] gives a very good understanding of the field along with a lucid introduction to the concepts of reachable sets, viable sets, and safe sets.

The connection of the above mentioned sets to the viscosity solution of the special forms of Hamilton Jacobi (HJ) equations using level set methods is also explained at length. [32] provides a very accessible means of solving the above problem using a MATLAB based solver. Two most important concepts in this area of research are reachable sets and safe sets. Reachable sets are computed by solving a HJ partial differential equation (PDE) forward in time, while safe sets are computed by solving the same PDE backwards in time. Reachable sets are computed such that the temporal derivative of an implicit function is always positive whereas safe sets are computed such that the temporal derivative of the implicit function is always negative. Thus it can be observed that reachable sets only grow (in the sense of inclusion), whereas safe sets only shrink with increasing horizon. Several publications [31, 33–35] provide safe set computation results for various models of aircraft longitudinal dynamics. The computation of safe set or the largest controlled invariant subset of a given flight

envelope is an interesting problem to the aerospace community. Whereas the computation of reachable set can be used in any area of automation to answer the feasibility question for a controller.

1.3 Summary

This work presents a compelling case for the proposed intuitive control framework. This is done by applying the framework to continuous time realizations of linear time invariant systems. Computer simulation results are presented to validate the proposed control methods, and identification routine.

The presented work includes numerical results to show the closed loop identification of unknown linear system parameters for a continuous time MIMO realization using online control techniques. Robust adaptive schemes have complemented identification schemes for partially unknown stable models under certain assumptions. Our contribution will be extending the idea of iterative control schemes which converges to the optimal controller for a MIMO system which is not necessarily stable. In this present work the stability assumption in the ADP framework is eliminated by complementing with classical adaptive control techniques. Model matching conditions from adaptive control literature are imposed instead of stricter stability conditions.

In addition to extending ADP for a wider class of systems, the presented work can be looked upon as a natural extension to closed loop system identification. The proposed control framework is divided into phases as an interpretation of intuitive control for methodically identifying unknown parameters.

Simulation results with aircraft models ensure that the proposed control framework is relevant to real-world aerospace applications. A 2-DOF helicopter control experiment from [36] were used to evaluate the control scheme. This evaluates the proposed control scheme for a real world example which is weakly nonlinear.

The proposed method has been extended to linear systems with lipschitz nonlinearities and simulation results are presented with identified parameters. The scope of intuitive control framework is further extended to a special class of nonlinearities (Hamiltonian systems).

Computation of forward reachable sets after identification is a valuable addition for aerospace applications. Although such computations for unknown Lipschitz nonlinear systems have been performed in [37], their applicability is limited due to the offline nature of implementation (presence of offline experiments for learning parameters). Reachable sets for the identified models will be computed by propagation of level sets using Hamilton-Jacobi-Bellman equation. This work provides the only possible way to identify the unknown parameters and thus calculate reachable sets for an unknown linear dynamic system.

1.4 Objectives and Contributions

Objectives of the proposed work are classified as a list of primary and secondary objectives.

1.4.1 List of Primary Objectives

Below mentioned primary objectives amount to a major part of the original contribution to the research community and hence are not expendable.

I. Develop an online controller framework for arriving at the optimal controller for unknown continuous time models (Linear Time Invariant) of known order.

II. Validate and evaluate the proposed framework using computer simulation of numerical models.

III. Identify the unknown model parameters in closed loop using the proposed online method.

1.4.2 List of Secondary Objectives

Although the primary objectives themselves constitute a complete contribution. Secondary objectives enhance the proposed work by making it more relevant for a larger number of problems. These objectives are mentioned below.

IV. Investigate the extension of the framework to weakly nonlinear systems.

V. Evaluate the framework for a relevant aerospace application.

VI. Evaluate the framework for a real-world MIMO system.

VII. Compute the forward reachable sets using the identified parameters for an otherwise unknown model.

1.4.3 List of Contributions

(a) Objectives I., II., III., V. :

Nuthi, Pavan, and Kamesh Subbarao. "Aspects of Intuitive Control: Stabilize, Optimize, and Identify " Proceedings of AIAA Guidance Navigation and Control Conference, AIAA Scitech 2015. [38]

(b) Objectives V., VI. :

Nuthi, Pavan, and Kamesh Subbarao. "Experimental Verification of Linear and Adaptive Control Techniques for a 2-DOF Helicopter" Journal of Dynamic Systems, Measurement and Control. [36]

(c) Objective V., VI. :

Nuthi, Pavan, and Kamesh Subbarao. "Implementation and Testing of Adaptive Augmentation Techniques on a 2-DOF Helicopter." ASME 2013 International Mechanical Engineering Congress and Exposition. American Society of Mechanical Engineers, 2013. [39]

(d) Objective V. :

Nuthi, Pavan, and Kamesh Subbarao. "Autonomous vertical landing on a marine

vessel" Proceedings of AIAA Atmospheric Flight Mechanics Conference, AIAA Scitech 2014. [40]

(e) Objectives V., VII. :

Nuthi, Pavan, and Kamesh Subbarao. "Computation of Safe and Reachable Sets for Model-Free Dynamical Systems: Aircraft Longitudinal Dynamics" Proceedings of AIAA Atmospheric Flight Mechanics Conference, AIAA Scitech 2014. [37]

CHAPTER 2

Intuitive Control Framework

2.1 Problem Formulation

Consider a linear system realization as shown in Eq. (2.1) with n states and m control inputs. It is assumed that measurements of state space variable $\mathbf{x} \in \mathbb{R}^n$, and control input $\mathbf{u} \in \mathbb{R}^m$ are available in real time. This assumption of full state measurement allows the design of full state feedback controllers.

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{f}(\mathbf{x}) \\ \mathbf{y} &= \mathbf{x} \end{aligned} \tag{2.1}$$

Assuming varying levels of knowledge for linear system description $(\mathbf{A}, \mathbf{B}, \mathbf{f}(\mathbf{x}))$ in Eq. (2.1), the problem is to develop an online controller implementation which regulates the states of the unknown system and eventually estimates the unknowns in a closed loop fashion. The following material discusses four different cases with varying levels of uncertainty in system parameters. First two cases assume linear system description with $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

It is assumed that the pair $(\mathbf{A}, \mathbf{B})$ although unknown is stabilizable. In the absence of this assumption, existence of a static linear feedback controller is not guaranteed which makes the whole exercise of designing a linear state feedback controller moot. Note that unlike typical problems handled by Robust control methods, no explicit assumptions on bounding sets for unknown parameters$(\mathbf{A}, \mathbf{B})$ are made.

The proposed control framework is demonstrated using a continuous time simulation model of linearized lateral flight dynamics of a Harrier AV-8B for linear cases. The underlying nonlinear model has been used in prior publications [37, 40–42]. Note that Case 2 is a generalization of Case 1. But they are presented separately in order to show the chronological progress made on the solution framework. Both the methods also differ in their implementation of controllers.

The assumptions made on the unknown parameters $\mathbf{A}, \mathbf{B}$ are similar to the ones from classical adaptive control literature and note that no assumption on the stability of unknown matrix $\mathbf{A}$ is made in contrast to the online Policy iteration based controllers. This results in less strict conditions on the unknowns and applicability of the solution to a wider variety of systems.

2.1.1 Case 1: Unknown linear internal dynamics ($\mathbf{A}$)

In this case it is assumed that only the parameter $\mathbf{A} \in \mathbb{R}^{n \times n}$ is unknown, and $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. The controller implementation is free to use explicit knowledge of known parameter $\mathbf{B} \in \mathbb{R}^{n \times m}$.

It is also assumed that $\exists\ \mathbf{K}^* \in \mathbb{R}^{m \times n}$ such that $\mathbf{A} - \mathbf{B}\mathbf{K}^* = \mathbf{A_m}$ where $\mathbf{A_m} \in \mathbb{R}^{n \times n}$ represents the chosen stable reference model.

2.1.2 Case 2: Unknown linear dynamics ($\mathbf{A}, \mathbf{B}$)

In this case it is assumed that both parameters $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{B} \in \mathbb{R}^{n \times m}$ are unknown, and $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. The controller implementation must be done without the explicit knowledge of $(\mathbf{A}, \mathbf{B})$.

It is also assumed that $\exists\ \mathbf{K}^* \in \mathbb{R}^{m \times n}, \mathbf{L}^* \in \mathbb{R}^{m \times m}$ such that $\mathbf{A} - \mathbf{B}\mathbf{K}^* = \mathbf{A_m}$ and $\mathbf{B}\mathbf{L}^* = \mathbf{B_m}$ where $\mathbf{A_m} \in \mathbb{R}^{n \times n}$ is Hurwitz, and $\mathbf{L}^*$ is either positive definite or negative definite.

2.1.3 Case 3: Unknown $(\mathbf{A}, \mathbf{B})$ with Lipschitz nonlinearity

In this case it is assumed that both parameters $\mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{B} \in \mathbb{R}^{n\times m}$ are unknown, and $\mathbf{f}(\mathbf{x}) \leq \alpha\|\mathbf{x}\|\ \forall \mathbf{x}$ for a known $\alpha > 0$. The controller implementation must be done without the explicit knowledge of $(\mathbf{A}, \mathbf{B})$.

It is also assumed that $\exists\ \mathbf{K}^* \in \mathbb{R}^{m\times n}, \mathbf{L}^* \in \mathbb{R}^{m\times m}$ such that $\mathbf{A} - \mathbf{B}\mathbf{K}^* = \mathbf{A_m}$ and $\mathbf{B}\mathbf{L}^* = \mathbf{B_m}$ where $\mathbf{A_m} + \alpha\mathbf{I} \in \mathbb{R}^{n\times n}$ is Hurwitz, and $\mathbf{L}^*$ is either positive definite or negative definite.

2.1.4 Case 4: Rigid body attitude dynamics with unknown inertia

In this case it is assumed that parameter $\mathbf{B} \in \mathbb{R}^{n\times m}$ is unknown,$\mathbf{A} = \mathbf{0}$, and $\mathbf{f}(\mathbf{x})$ is such that $\frac{\partial V}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) = \mathbf{0}$ for some positive definite Lyapunov-like function $V(\mathbf{x})$. The nonlinear term is passive, and does not result in change of energy for the closed loop system. Such systems are called Hamiltonian systems, and do not have a internal mechanism for energy dissipation or gain. The controller implementation must be done without the explicit knowledge of $(\mathbf{A}, \mathbf{B})$.

It is also assumed that $\exists\ \mathbf{K}^* \in \mathbb{R}^{m\times n}, \mathbf{L}^* \in \mathbb{R}^{m\times m}$ such that $\mathbf{A} - \mathbf{B}\mathbf{K}^* = \mathbf{A_m}$ and $\mathbf{B}\mathbf{L}^* = \mathbf{B_m}$ where $\mathbf{A_m} \in \mathbb{R}^{n\times n}$ is Hurwitz, and $\mathbf{L}^*$ is either positive definite or negative definite.

2.2 Solution Methodology

The Intuitive control framework identifies an unknown linear system using an online control technique while ensuring stable regulation in closed loop. This section gives a bird's eye view of the framework without going into the specific details of implementation for all the cases.

In this framework, it is interpreted that a conscious act of control can be crudely divided into two phases, namely stabilize and then optimize. To illustrate this inter-

pretation of intuitive control, concepts from both adaptive and optimal control theory are employed. The first phase of control entails stabilization of the unknown internal dynamics of a linear system through adaptive control methods. The second phase entails further learning in which an optimal control policy is learnt using a flavor of dynamic programming from optimal control theory. A novel use of resulting information gained in the optimization phase for the identification of unknown internal dynamics is also presented.

Idea for the framework originates from an over simplified notion of complementing Online Policy Iteration controller with Model Reference Adaptive control methods found in literature. There have been several online implementations of generalized policy iteration on continuous time systems which guarantee identification of optimal linear feedback controller. All these methods are restricted to use with stable systems only. The same restriction is sometimes mentioned as the knowledge of initial stabilizing gain. Conventionally model reference adaptive techniques have been used to stabilize an unknown system with nominal assumptions on the unknown parameters. Hence the framework starts with matching condition assumptions from classical adaptive control.

2.2.1 Structure

The framework consists of a hybrid controller which progresses through namely three stages: Stabilization, and Optimization followed by an Identification method. The schematic in Fig. [2.2.1] shows a state machine for the controller. Specific nature of these controllers in each stage will be discussed in subsequent chapters.

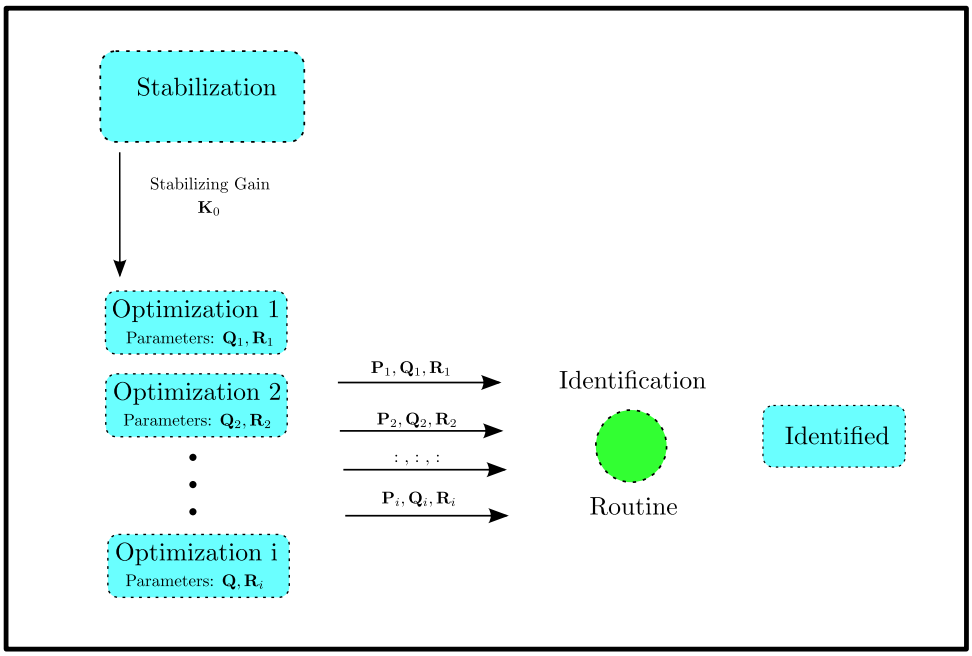

Figure 2.1. Intuitive Control framework for unknown dynamics.

2.2.2 Stabilization

The controller with no knowledge of the system parameters starts in the Stabilization phase. The goal of this phase is to stabilize the unknown dynamic system and yield a stabilizing gain $\mathbf{K}_0$. An exit condition based on the norm of error from the reference model is developed to ensure stability after the adaptation is ceased.

2.2.3 Optimization

The stabilization phase results in a statically stabilizing feedback gain for the optimization phase which consists of iteratively evaluating optimal feedback controllers for a variety of parameters. The optimization phase continues until enough information is gathered for full parameter identification. The exit condition for the optimization phase ensures unique solution to the unknown linear system parameters.

2.2.4 Identification

The identification routine is performed using a linear least squares solution of vectorized Algebraic Ricatti Equations from Optimization phase. The Identified phase simply implements an optimal controller by solving the Algebraic Ricatti Equation using identified parameters from the Identification routine.

CHAPTER 3

Solution for unknown linear internal dynamics

This chapter develops the ideas mentioned in Ch. 2 and explicitly formulates the controller for the unknown internal dynamics case (unknown $\mathbf{A}$, known $\mathbf{B}$, $\mathbf{f}(\mathbf{x}) = \mathbf{0}$) for a linear time invariant system shown below. In addition to the implementation details further analysis is presented for the stabilizing controller which also identifies the unknown matrix $\mathbf{A}$.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{3.1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$ represent the state and control vectors respectively.

3.1 Solution Methodology - Stabilize

The first part of the controller involves a Model Reference Adaptive Control (MRAC) approach which stabilizes the partially known dynamic system . $\mathbf{A_m}$ is chosen to be a Hurwitz matrix. It is further assumed that there is sufficient structural flexibility to ensure the existence of $\mathbf{K}^*$ ($\exists\ \mathbf{K}^*$ s.t. $\mathbf{A_m} = \mathbf{A} - \mathbf{B}\mathbf{K}^*$).

The state signal $\mathbf{x}(t)$ can be characterized as shown in Eq. (3.2) with state feedback control law $\mathbf{u}(t) = -\hat{\mathbf{K}}(t)\mathbf{x}(t)$. The estimation error for the ideal parameter $\mathbf{K}^*$ is defined as $\tilde{\mathbf{K}}(t) = \mathbf{K}^* - \hat{\mathbf{K}}(t)$.

$$\begin{aligned} \dot{\mathbf{x}}(t) &= (\mathbf{A_m} + \mathbf{B}\mathbf{K}^*)\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ &= \mathbf{A_m}\mathbf{x}(t) + \mathbf{B}\tilde{\mathbf{K}}(t)\mathbf{x}(t) \end{aligned} \tag{3.2}$$

The adaptive law for $\hat{\mathbf{K}}$ follows from a straightforward Lyapunov analysis by choosing the candidate Lyapunov function to be $V = \mathbf{x}^\top(t)\boldsymbol{\mathcal{P}}\mathbf{x}(t) + \mathrm{Tr}(\tilde{\mathbf{K}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{K}}(t))$. Note that $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{n\times n}, \boldsymbol{\Gamma} \in \mathbb{R}^{m\times m}$ are chosen symmetric positive definite matrices. Following the developments in [**?**,43], asymptotic stability of the closed loop dynamics in Eq. (3.2) can be shown by Lyapunov-Like Lemma(motivated by Barbalat's Lemma) for non-autonomous systems if the adaptive law for $\hat{\mathbf{K}}$ is chosen as

$$\dot{\hat{\mathbf{K}}}(t) = -\boldsymbol{\Gamma}\mathbf{B}^\top\boldsymbol{\mathcal{P}}\mathbf{x}\mathbf{x}^\top \tag{3.3}$$

where $\boldsymbol{\mathcal{P}}$ is the solution to the Lyapunov equation ${\mathbf{A_m}}^\top\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\mathbf{A_m} = -\boldsymbol{\mathcal{N}}$, for a chosen $\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{N}}^\top > 0$. The prescribed feedback adaptation ensures $\dot{V} = -\mathbf{x}^\top(t)\boldsymbol{\mathcal{N}}\mathbf{x}(t)$. Note that the convergence of $\hat{\mathbf{K}}(t)$ to $\mathbf{K}^*$ is not guaranteed. $\hat{\mathbf{K}}(t)$ does however converge to a stabilizing gain $\hat{\mathbf{K}}_\infty$ eventually. Thus the resulting closed loop system upon convergence to $\hat{\mathbf{K}}_\infty$ is asymptotically stable i.e. $\mathbf{A} - \mathbf{B}\hat{\mathbf{K}}_\infty$ is Hurwitz. Barbalat's Lemma can be applied since $\dot{V}$ is clearly uniformly continuous in time. Note that the Lyapunov function V converges to a constant but not necessarily zero, whereas the derivative along the trajectory $\dot{V}$ vanishes as $t \to \infty$.

3.2 Exit condition for Stabilization phase

Although Barbalat's Lemma guarantees asymptotic stability as $t \to \infty$, ideally the adaptation should be continued for only a finite time. Assume that the MRAC style adaptation of feedback gain $\hat{\mathbf{K}}$ is ceased for time $t \geq T_s$ for some chosen T_s .

$$\begin{aligned}
\dot{\mathbf{x}} &= (\mathbf{A} - \mathbf{B}\hat{\mathbf{K}})\mathbf{x} \\
&= (\mathbf{A} - \mathbf{B}\mathbf{K}^* + \mathbf{B}\mathbf{K}^* - \mathbf{B}\hat{\mathbf{K}})\mathbf{x} \\
&= \mathbf{A_m}\mathbf{x} + \mathbf{B}\tilde{\mathbf{K}}\mathbf{x}
\end{aligned}$$

Upon integration for feedback gain error $\tilde{\mathbf{K}}(t)$, and the state $\mathbf{x}(t)$ before the adaptation is ceased *i.e* $\forall\ 0 \leq t \leq T_s$.

$$\begin{aligned}
\tilde{\mathbf{K}}(t) &= \tilde{\mathbf{K}}(0) + \int_0^t \mathbf{\Gamma B}^\top \boldsymbol{\mathcal{P}} \mathbf{x}(s)\mathbf{x}^\top(s) ds \quad \text{where} \quad \dot{\tilde{\mathbf{K}}} = -\dot{\hat{\mathbf{K}}} = \mathbf{\Gamma B}^\top \boldsymbol{\mathcal{P}} \mathbf{x}\mathbf{x}^\top \quad (3.4)\\
\mathbf{x}(t) &= \mathrm{e}^{\mathbf{A_m}t}\mathbf{x}(0) + \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B}\tilde{\mathbf{K}}(\mathbf{x}(\tau))\mathbf{x}(\tau) d\tau \quad (3.5)\\
&= \mathrm{e}^{\mathbf{A_m}t}\mathbf{x}(0) + \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B}\left(\tilde{\mathbf{K}}(0) + \int_0^\tau \mathbf{\Gamma B}^\top \boldsymbol{\mathcal{P}} \mathbf{x}(s)\mathbf{x}^\top(s) ds\right)\mathbf{x}(\tau) d\tau
\end{aligned}$$

The state relation obtained is implicit but it is not a kind of implicitness which can be dealt with Bellman-Gronwall Lemma. However $\mathbf{x}(T_s), \tilde{\mathbf{K}}(T_s)$ can be evaluated using Eq. (3.5,3.4). After the adaptation is ceased, feedback gain will remain constant $\hat{\mathbf{K}}(T_s)$. Thus the states are governed by linear time invariant dynamics after T_s. Explicit form for the state can be given $\forall\ t \geq T_s$ using Eq. (3.6).

$$\begin{aligned}
\mathbf{x}(t) &= \mathrm{e}^{(\mathbf{A_m}+\mathbf{B}\tilde{\mathbf{K}}(T_s))(t-T_s)}\mathbf{x}(T_s) \quad (3.6)\\
\|\mathbf{x}(t)\| &\leq \|\mathrm{e}^{(\mathbf{A_m}+\mathbf{B}\tilde{\mathbf{K}}(T_s))(t-T_s)}\|\ \|\mathbf{x}(T_s)\|\\
\|\mathbf{x}(t)\| &\leq \mathrm{e}^{\mu(\mathbf{A_m}+\mathbf{B}\tilde{\mathbf{K}})(t-T_s)}\ \|\mathbf{x}(T_s)\| \quad (3.7)
\end{aligned}$$

Note that $\mathbf{x}(t) \in \mathbb{R}^n$ and $\|\mathbf{x}(t)\| \in \mathbb{R}$. The expression $\|\cdot\|$ for a square matrix should be interpreted as the induced norm from vector 2-norm $\|\cdot\|$. Above steps use triangle inequality for vector norms and the definition of induced norm for square matrices. Let $\mu(\mathbf{A})$ represent logarithmic norm of a matrix $\mathbf{A}$, and signifies the maximal growth rate of $\log\|\mathbf{x}\|$ if $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. The logarithmic norm properties include $\|\mathrm{e}^{\mathbf{P}t}\| \leq \mathrm{e}^{\mu(\mathbf{P})t}$, and $\mu(\mathbf{P}+\mathbf{Q}) \leq \mu(\mathbf{P}) + \|\mathbf{Q}\|$.

The upper limit on $\|\mathbf{x}(t)\|$ can also be obtained by using Bellman-Gronwall Lemma. If $\mu(\mathbf{A_m} + \mathbf{B}\tilde{\mathbf{K}}) \leq -\delta$ is satisfied for some $\delta > 0$ and $\|\mathbf{x}(T_s)\|$ is finite, the state trajectories will be bounded by a decaying exponential.

Note that $\mu(\mathbf{A_m} + \mathbf{B}\tilde{\mathbf{K}}(T_s)) \leq -\delta \implies |\mu(\mathbf{A_m} + \mathbf{B}\tilde{\mathbf{K}}(T_s))| \geq \delta$.

$$\mathbf{B}\tilde{\mathbf{K}}(T_s) = \mathbf{B}\tilde{\mathbf{K}}(0) + \mathbf{B}\boldsymbol{\Gamma}\mathbf{B}^\top\boldsymbol{\mathcal{P}}\int_0^{T_s}\mathbf{x}(s)\mathbf{x}^\top(s)ds \tag{3.8}$$

If the condition $\int_0^{T_s}\|\mathbf{x}(s)\mathbf{x}^\top(s)\|ds \geq \frac{|\mu(\mathbf{A_m}+\mathbf{B}\tilde{\mathbf{K}}(0))|+\delta}{\|\mathbf{B}\boldsymbol{\Gamma}\mathbf{B}^\top\boldsymbol{\mathcal{P}}\|}$ is satisfied, then the gain $(\hat{\mathbf{K}}(T_s))$ is stabilizing after the adaptation is ceased. Further it can be established that $\|\mathbf{x}\mathbf{x}^\top\| = \|\mathbf{x}\|^2$(Consider an arbitrary vector $\mathbf{p} \in \mathbb{R}^n$, then $\|\mathbf{x}\mathbf{x}^\top\mathbf{p}\| = |\mathbf{x}^\top\mathbf{p}|\|\mathbf{x}\| \leq \|\mathbf{x}\|^2\|\mathbf{p}\|$).

$$\boxed{\int_0^{T_s}\|\mathbf{x}(s)\|^2ds \geq \frac{|\mu(\mathbf{A_m} + \mathbf{B}\tilde{\mathbf{K}}(0))| + \delta}{\|\mathbf{B}\boldsymbol{\Gamma}\mathbf{B}^\top\boldsymbol{\mathcal{P}}\|}} \tag{3.9}$$

Note that the lower limit from stabilizing condition cannot be explicitly obtained using known parameters. However such a positive limit can be calculated if the parameters were known. Thus a positive limit $\sigma > 0$ is chosen which yields an implementable stabilizing condition.

$$\int_0^{T_s}\|\mathbf{x}(s)\|^2ds \geq \sigma > 0 \tag{3.10}$$

The stabilizing condition can also be interpreted as a lower limit on the decay of Lyapunov function V.

$$\begin{aligned}
\lambda_{\min}(\boldsymbol{\mathcal{N}})\|\mathbf{x}\|^2 \quad &\leq \mathbf{x}^\top\boldsymbol{\mathcal{N}}\mathbf{x} \quad &&\leq \lambda_{\max}(\boldsymbol{\mathcal{N}})\|\mathbf{x}\|^2 \quad \forall\mathbf{x} \\
\frac{\mathbf{x}^\top\boldsymbol{\mathcal{N}}\mathbf{x}}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} \quad &\leq \|\mathbf{x}\|^2 \quad &&\leq \frac{\mathbf{x}^\top\boldsymbol{\mathcal{N}}\mathbf{x}}{\lambda_{\min}(\boldsymbol{\mathcal{N}})} \quad \forall\mathbf{x} \\
\int_0^{T_s}\frac{-\dot{V}}{\lambda_{\max}(\boldsymbol{\mathcal{N}})}ds \quad &\leq \int_0^{T_s}\|\mathbf{x}\|^2ds \quad &&\leq \int_0^{T_s}\frac{-\dot{V}}{\lambda_{\min}(\boldsymbol{\mathcal{N}})}ds \\
\frac{V(0)-V(T_s)}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} \quad &\leq \int_0^{T_s}\|\mathbf{x}\|^2ds \quad &&\leq \frac{V(0)-V(T_s)}{\lambda_{\min}(\boldsymbol{\mathcal{N}})}
\end{aligned} \tag{3.11}$$

3.3 Solution Methodology - Optimize

The second phase of the controller uses the adaptive scheme developed in [10] using the Bellman equation to recursively approach the optimal feedback gain. This construction leads to a control gain which converges to the optimal feedback gain for the Linear Quadratic regulator with unknown internal dynamics.

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{3.12}$$

Assuming that the pair $(\mathbf{A}, \mathbf{B})$ is stabilizable, the infinite horizon linear quadratic regulator problem would be to find $\mathbf{u}^*(t)$.

$$\mathbf{u}^*(t) = \underset{\mathbf{u}(t),\ t\in[t_0,\infty]}{\operatorname{argmin}} V(t_0, \mathbf{x}(t_0), \mathbf{u}(t)) \tag{3.13}$$

The infinite horizon cost for the optimal control problem is posed as

$$V(\mathbf{x}(t_0), t_0) = \int_{t_0}^{\infty} (\mathbf{x}^\top(\tau)\mathbf{Q}\mathbf{x}(\tau) + \mathbf{u}^\top(\tau)\mathbf{R}\mathbf{u}(\tau))d\tau \tag{3.14}$$

where $\mathbf{Q} > \mathbf{0},\ \mathbf{R} > \mathbf{0}$, and the pair $((\mathbf{A}, \sqrt{\mathbf{Q}})$ is detectable. The solution to this particular optimal control problem is known to be a state feedback controller $\mathbf{u}(t) = -\mathbf{K}\mathbf{x}(t)$ and the gain $\mathbf{K} = \mathbf{R}^{-\mathbf{1}}\mathbf{B}^\top\mathbf{P}$ where $\mathbf{P}$ is the positive definite solution to the following Algebraic Ricatti Equation.

$$\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-\mathbf{1}}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} = \mathbf{0} \tag{3.15}$$

Of course the control law mentioned above can be synthesized if $\mathbf{A}$ is known.

The policy iteration proposed by [10] is used with $\mathbf{K}_0 = \hat{\mathbf{K}}(T_s)$ as the initial stabilizing gain. The result is an adaptive controller which converges to the optimal feedback controller obtained from ARE in Eq. (3.15) without the knowledge of internal dynamics.

The cost-to-go with a stabilizing controller gain $\mathbf{K}$ can be written as

$$V(\mathbf{x}(t)) = \int_t^\infty \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}^\top \mathbf{R}\mathbf{K})\mathbf{x}(\tau)d\tau = \mathbf{x}^\top(t)\mathbf{P}\mathbf{x}(t) \tag{3.16}$$

where $\mathbf{P}$ is the solution of the following Lyapunov equation

$$(\mathbf{A} - \mathbf{B}\mathbf{K})^\top \mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) = -(\mathbf{K}^\top \mathbf{R}\mathbf{K} + \mathbf{Q}) \tag{3.17}$$

The cost function can be incrementally written as

$$V(\mathbf{x}(t)) = \int_t^{t+T} \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}^\top \mathbf{R}\mathbf{K})\mathbf{x}(\tau)d\tau + V(\mathbf{x}(t+T)) \tag{3.18}$$

A policy iteration scheme proposed in [10] is used. Considering an initial stabilizing gain $\mathbf{K_0}$ the following policy iteration scheme is implemented online

$$\mathbf{x}^\top(t)\mathbf{P}_k\mathbf{x}(t) = \int_t^{t+T} \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k)\mathbf{x}(\tau)d\tau + \mathbf{x}^\top(t+T)\mathbf{P}_k\mathbf{x}(t+T) \tag{3.19}$$

$$\mathbf{K}_{k+1} = \mathbf{R}^{-1}\mathbf{B}^\top \mathbf{P}_k \tag{3.20}$$

Repeat the iteration until subsequent estimates for $\mathbf{P}$ are close enough. Let $N \in \mathbb{N}$ such that $\|\mathbf{P}_k - \mathbf{P}_{k-1}\|_F < \epsilon$ is true for all $k > N$, where $\|\cdot\|_F$ represents the Frobenius norm.

3.4 Optimize - Proof of Convergence

The above mentioned policy iteration scheme can be shown to be convergent with the assumption of an initial stabilizing $\mathbf{K}_0$. A few supporting lemmas are mentioned prior to the proof of convergence.

Lemma 1: Assuming $\mathbf{A} - \mathbf{B}\mathbf{K}_k$ is Hurwitz, the solution $\mathbf{P}_k$ in Eq. (3.19) is equivalent to finding the solution of following Lyapunov equation.

$$(\mathbf{A} - \mathbf{B}\mathbf{K}_k)^\top \mathbf{P}_k + \mathbf{P}_k(\mathbf{A} - \mathbf{B}\mathbf{K}_k) = -(\mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k + \mathbf{Q}) \tag{3.21}$$

Proof: There exists a positive definite solution $\mathbf{P}_k$ to the above equation due to the assumption that $\mathbf{A} - \mathbf{B}\mathbf{K}_k$ is Hurwitz and $\mathbf{Q} + \mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k$ is positive definite from previous assumptions ($\mathbf{Q} > \mathbf{0},\ \mathbf{R} > \mathbf{0}$). Note that $V_k(\mathbf{x}(t)) = \mathbf{x}^\top(t)\mathbf{P}_k\mathbf{x}(t),\ \forall \mathbf{x}(t)$ is a Lyapunov function for the system $\dot{\mathbf{x}} = \mathbf{A}_k\mathbf{x}$ where $\mathbf{A}_k = \mathbf{A} - \mathbf{B}\mathbf{K}_k$.

$$
\begin{aligned}
\dot{V}_k &= \mathbf{x}^\top(t)(\mathbf{A}_k^\top \mathbf{P}_k + \mathbf{P}_k\mathbf{A}_k)\mathbf{x}(t) \\
&= -\mathbf{x}^\top(t)(\mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k + \mathbf{Q})\mathbf{x}(t) \\
-\frac{d(\mathbf{x}^\top(t)\mathbf{P}_k\mathbf{x}(t))}{dt} &= \mathbf{x}^\top(t)(\mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k + \mathbf{Q})\mathbf{x}(t) + \mathbf{x}^\top(t)\mathbf{P}_k\mathbf{x}(t) - \mathbf{x}^\top(t+T)\mathbf{P}_k\mathbf{x}(t+T) \\
&= \int_t^{t+T} \mathbf{x}^\top(\tau)(\mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k + \mathbf{Q})\mathbf{x}(\tau)d\tau \qquad (3.22)
\end{aligned}
$$

Eq. (3.22) is true $\forall\ T > 0$, thereby Eq. (3.21) $\Rightarrow$ Eq. (3.19). It is easily shown that Eq. (3.19) $\Rightarrow$ Eq. (3.21). Thus Eq. (3.19) and (3.21) are equivalent. Hence, the solution for $\mathbf{P}_k$ can be obtained from Eq. (3.19) without the knowledge of $\mathbf{A}$. Since it is proved that Eq. (3.19) $\Leftrightarrow$Eq. (3.21), the iteration scheme described in Eq. (3.19) and (3.20) is equivalent to iterating in between Eq. (3.21) and (3.20).

Lemma 2: Assuming that $\mathbf{K}_k$ is a stabilizing gain for the system $\dot{\mathbf{x}} = \mathbf{A}_k\mathbf{x}$ with the cost $V_k(\mathbf{x}(t)) = \mathbf{x}^\top(t)\mathbf{P}_k\mathbf{x}(t)$, if Eq. (3.20) was used for updating $\mathbf{K}_k$ then the resulting new control policy $\mathbf{K}_{k+1}$ will be stabilizing.

Proof: Let $V_k(\mathbf{x}(t))$ be a Lyapunov function candidate for the system with new control policy $\mathbf{K}_{k+1}$.

$$
\begin{aligned}
\dot{V}_k(\mathbf{x}(t)) &= \mathbf{x}^\top(t)[\mathbf{P}_k(\mathbf{A} - \mathbf{B}\mathbf{K}_{k+1}) + (\mathbf{A} - \mathbf{B}\mathbf{K}_{k+1})^\top \mathbf{P}_k]\mathbf{x}(t) \\
&= -\mathbf{x}^\top(t)[(\mathbf{K}_k - \mathbf{K}_{k+1})^\top \mathbf{R}(\mathbf{K}_k - \mathbf{K}_{k+1}) + \mathbf{Q} + \mathbf{K}_{k+1}^\top \mathbf{R}\mathbf{K}_{k+1}]\mathbf{x}(t)
\end{aligned}
$$

A negative definite $\dot{V}_k(\mathbf{x}(t))$ proves that the new control policy $\mathbf{u} = -\mathbf{K}_{k+1}\mathbf{x}$ is stabilizing if the previous gain $\mathbf{K}_k$ is stabilizing. The lemma also implies that all the iterated control policies $\mathbf{K}_k$ starting from $\mathbf{K}_0$ will be stabilizing.

Let $\mathbf{Ric}(\mathbf{P}_k)$ be defined as

$$\mathbf{Ric}(\mathbf{P}_k) \equiv \mathbf{A}^\top \mathbf{P}_k + \mathbf{P}_k \mathbf{A} + \mathbf{Q} - \mathbf{P}_k \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^\top \mathbf{P}_k \tag{3.23}$$

with $\mathbf{Ric'}_{\mathbf{P}_k}$ being its Frechet derivative with respect to $\mathbf{P}_k$.

$$\mathbf{Ric'}_{\mathbf{P}_k}(\mathbf{M}) = (\mathbf{A} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{P}_k)^\top \mathbf{M} + \mathbf{M}(\mathbf{A} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{P}_k) \tag{3.24}$$

Note, the above equation evaluates the derivative at any given matrix $\mathbf{M}$.

Lemma 3: Newton's iteration method using the Frechet derivative is equivalent to iterating between Eq. (3.19) and Eq. (3.20).

$$\mathbf{P}_{k+1} = \mathbf{P}_k - (\mathbf{Ric'}_{\mathbf{P}_k})^{-1} \mathbf{Ric}_{\mathbf{P}_k} \tag{3.25}$$

Proof: Eq. (3.19,3.20) $\Rightarrow$ Eq. (3.20,3.21), and substituting Eq. (3.20) in Eq. (3.21) yields

$$\mathbf{A}_k^\top \mathbf{P}_k + \mathbf{P}_k \mathbf{A}_k = -(\mathbf{Q} + \mathbf{P}_{k-1} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^\top \mathbf{P}_{k-1}) \tag{3.26}$$

Subtracting $\mathbf{A}_k^\top \mathbf{P}_{k-1} + \mathbf{P}_{k-1}\mathbf{A}_k$ from both sides yields

$$\mathbf{A}_k^\top (\mathbf{P}_k - \mathbf{P}_{k-1}) + (\mathbf{P}_k - \mathbf{P}_{k-1}) \mathbf{A}_k = -(\mathbf{Q} - \mathbf{P}_{k-1} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^\top \mathbf{P}_{k-1}) \tag{3.27}$$

which is the Newton's iteration method using the introduced notation.

Theorem 4 (*Convergence of* $\mathbf{P}$ *in the Optimize phase*): Assuming the pair $(\mathbf{A}, \mathbf{B})$ is stabilizable, and the pair $(\mathbf{A}, \sqrt{\mathbf{Q}})$ is detectable, and $\mathbf{R} > \mathbf{0},\ \mathbf{Q} > \mathbf{0}$, the iteration of Eq. (3.19) and (3.20) will converge to the optimal controller given by the ARE solution corresponding to the cost function in Eq. (3.14).

Proof of Convergence: It has been shown in [13] that the Newton's iteration using Frechet derivative will converge to the solution of the ARE. Using the equivalence

results established from Lemma 1 and 3, we can say that iteration of Eq. (3.19) and (3.20) will converge to the optimal controller. Lemma 2 establishes that all the iterations of feedback gain are stabilizing, thus rendering the iteration convergent.

Alternate Proof of Convergence: The stability of the dynamic system with online Policy iteration can also be shown by using the framework of Lyapunov functions for switched systems [44]. Under the assumption that the gain at the end of stabilization phase is stabilizing, [16] show that the corresponding $\mathbf{P}_k$ form a monotonically decreasing sequence ($\mathbf{P}_k > \mathbf{P}_{k-1}$). This fact is also evident from Eq. (3.27). The sequence of Lyapunov functions $\mathbf{x}^\top \mathbf{P}_k \mathbf{x}$ form the required set of non-increasing positive definite functions which in turn show the stability of the hybrid switched control scheme implemented.

3.5 Solution Methodology - Identify

The first two phases stabilize and then optimize a quadratic performance metric on a partially unknown linear system. The third and final identification phase uses the information gathered from the optimization phase to identify the unknown matrix $\mathbf{A}$. Since the identification step involves optimization solutions for different LQR parameters, it is convenient to introduce a more accommodating notation for approximations of $\mathbf{P}$. The parameters $N, \mathbf{Q}, \mathbf{R}, \mathbf{P}_k$ used in previous section for an optimization phase will be generalized for multiple optimization phases.

Let $\mathbf{P}_{i,k}$ be the solution of k^{th} iteration for LQR problem posed with parameters $\mathbf{Q}_i, \mathbf{R}_i$. Corresponding state feedback gain approximation for next iteration is computed as $\mathbf{K}_{i,k+1} = \mathbf{R}_i^{-1}\mathbf{B}^\top \mathbf{P}_{i,k}$. Also note that the convergence for each optimization phase is indicated by $\|\mathbf{P}_{i,k} - \mathbf{P}_{i,k-1}\|_F < \epsilon$ which is true for all $k > N_i$. Thus $\mathbf{P}_{i,N_i}$ represents the approximation of LQR solution for parameters $\mathbf{Q}_i, \mathbf{R}_i$

Let $\hat{\mathbf{A}}$ represent the estimated value of unknown parameter $\mathbf{A}$. Since the optimization phase ends when $\mathbf{P}_k$ converges to the optimal solution $\mathbf{P}$, the investigation for a closed form solution of $\mathbf{A}$ starts with the inspection of Eq. (3.15). The governing Algebraic Ricatti Equation(ARE) reduces to the following Eq. (3.28), where $\mathbf{A}$ is the only unknown.

$$\hat{\mathbf{A}}\mathbf{P}_N + \mathbf{P}_N\hat{\mathbf{A}}^\top = \mathbf{X} \tag{3.28}$$

Note that Eq. (3.28) is a symmetric linear matrix equation and thus yields $\frac{n^2+n}{2}$ linear scalar equations for a dynamic system of order n. This poses an underdetermined system of linear equations in terms of n^2 unknown elements of $\hat{\mathbf{A}}$. The problem of insufficient information is solved by using another optimal solution for a new set of performance metrics $(\mathbf{Q}, \mathbf{R})$. Consider two cases yielding linear matrix equations with unknown matrix $\hat{\mathbf{A}}$.

$$\hat{\mathbf{A}}\mathbf{P}_{1,N_1}+\mathbf{P}_{1,N_1}\hat{\mathbf{A}}^\top = \mathbf{X}_1 \tag{3.29}$$

$$\hat{\mathbf{A}}\mathbf{P}_{2,N_2}+\mathbf{P}_{2,N_2}\hat{\mathbf{A}}^\top = \mathbf{X}_2 \tag{3.30}$$

where $\mathbf{P}_{1,N_1} > \mathbf{0}, \mathbf{P}_{2,N_2} > 0, \mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times n}$. Note that the matrices $\mathbf{X}_i$ are evaluated as $\mathbf{P}_{i,N_i}\mathbf{B}\mathbf{R}_i^{-1}\mathbf{B}^\top\mathbf{P}_{i,N_i}-\mathbf{Q}_i$ for $i = 1, 2$. Eq. (3.29) and (3.30) represent the optimal state feedback control problem for the same system with different set of $(\mathbf{Q}, \mathbf{R})$. Kronecker algebra [45] can be employed in formulating an analytic solution for $\hat{\mathbf{A}}$. Refer to Appendix A for a Primer in Kronecker Algebra.

Upon vectorization, the linear matrix equation in Eq. (3.28) is reduced to a system of linear equations of order n^2 with the introduction of a known permutation matrix $\mathbf{\Pi}$(Refer to Ch. A).

$$\begin{aligned}(\mathbf{P}_N{}^\top \otimes \mathbf{I}_n)vec(\hat{\mathbf{A}}^\top) + (\mathbf{I}_n \otimes \mathbf{P}_N)vec(\hat{\mathbf{A}}) &= vec(\mathbf{X}) \\ \left[(\mathbf{P}_N \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_N)\right] vec(\hat{\mathbf{A}}) &= vec(\mathbf{X}) \end{aligned} \tag{3.31}$$

where $\mathbf{I}_n$ represents an identity matrix of the order n. Upon vectorizing the Eq. (3.29) and (3.30), they can be combined to solve for $\hat{\mathbf{A}}$.

$$\begin{bmatrix} (\mathbf{P}_{1,N_1} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_{1,N_1}) \\ (\mathbf{P}_{2,N_2} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_{2,N_2}) \end{bmatrix} vec(\hat{\mathbf{A}}) = \begin{bmatrix} vec(\mathbf{X}_1) \\ vec(\mathbf{X}_2) \end{bmatrix} \tag{3.32}$$

This set of overdetermined linear equations can be solved in a least squares sense using a psuedo-inverse to obtain a closed form solution for $\mathbf{A}$.

$$vec(\hat{\mathbf{A}}) = \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \begin{bmatrix} vec(\mathbf{P}_{1,N_1}\mathbf{B}\mathbf{R}_1{}^{-1}\mathbf{B}^\top \mathbf{P}_{1,N_1} - \mathbf{Q}_1) \\ vec(\mathbf{P}_{2,N_2}\mathbf{B}\mathbf{R}_2{}^{-1}\mathbf{B}^\top \mathbf{P}_{2,N_2} - \mathbf{Q}_2) \end{bmatrix} \tag{3.33}$$

$$\text{where } \mathbf{H} = \begin{bmatrix} (\mathbf{P}_{1,N_1} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}^n \otimes \mathbf{P}_{1,N_1}) \\ (\mathbf{P}_{2,N_2} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}^n \otimes \mathbf{P}_{2,N_2}) \end{bmatrix}$$

Above solution can be generalized in a case where employing only two sets of $(\mathbf{Q}, \mathbf{R})$ does not provide sufficient information to solve for a unique solution $\hat{\mathbf{A}}$. Let p be the number of pairs of $(\mathbf{Q}, \mathbf{R})$ employed in the solution of $\hat{\mathbf{A}}$

$$\boxed{\begin{aligned} vec(\hat{\mathbf{A}}) &= \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \mathbf{Y} \\ \text{where } \mathbf{H} &= \begin{bmatrix} \mathbf{H_1} \\ \mathbf{H_2} \\ \vdots \\ \mathbf{H_p} \end{bmatrix}, \ \mathbf{Y} = \begin{bmatrix} \mathbf{Y_1} \\ \mathbf{Y_2} \\ \vdots \\ \mathbf{Y_p} \end{bmatrix}, \\ \mathbf{H_i} &= (\mathbf{P}_{i,N_i} \otimes \mathbf{I}^n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_{i,N_i}), \\ \mathbf{Y_i} &= vec(\mathbf{P}_{i,N_i}\mathbf{B}{\mathbf{R}_i}^{-1}\mathbf{B}^\top \mathbf{P}_{i,N_i} - \mathbf{Q}_i) \end{aligned}} \tag{3.34}$$

The existence of a solution is guaranteed if and only if the pairs $(\mathbf{Q}_i, \mathbf{R}_i)$ are such that $rank(\mathbf{H}) = rank([\mathbf{H}\ \mathbf{Y}]) = n^2$. Thus if two instances of optimal control solution are not enough to solve for a unique $\hat{\mathbf{A}}$, more instances can be incorporated into the solution.

3.6 Algorithm for online implementation

Pseudocode for the three phase identification algorithm is listed as follows.

1. Employ the control law $\mathbf{u} = -\hat{\mathbf{K}}\mathbf{x}$ with adaptation law $\dot{\hat{\mathbf{K}}}(t) = -\mathbf{\Gamma}\mathbf{B}^\top \boldsymbol{\mathcal{P}}\mathbf{x}\mathbf{x}^\top$ for sampling time interval T seconds.
2. Go back to Step 1 to stabilize for another T seconds if the stabilization condition in Eq. (3.10) is not satisfied. Continue to Step 3 if satisfied.
3. Initialize the optimization phase for the first time by setting $k = 0$,$i = 1$ and $\mathbf{K}_{1,0} = \hat{\mathbf{K}}(T_s)$. Set the LQR parameters $\mathbf{Q}_1, \mathbf{R}_1$
4. Employ the control law $\mathbf{u} = -\mathbf{K}_{i,k}\mathbf{x}$ for the next sampling time interval T seconds.The state information is used to solve for $\mathbf{P}_{i,k}, \mathbf{K}_{i,k+1}$ from Eq. (3.19,3.20) respectively.

5. Check if $\|\mathbf{P}_{i,k} - \mathbf{P}_{i,k+1}\|_F \leq \epsilon$ for some $k \in \mathbb{N}$ where $\epsilon > 0$ is a predefined threshold for convergence. If the condition is not satisfied continue to Step 6, and if satisfied then continue to Step 7.
6. Continue the policy iteration for converging solution by setting $k = k + 1$ and go to Step 4.
7. Note down converging solution $\mathbf{P}_{i,N_i}$ corresponding to parameters $\mathbf{Q}_i, \mathbf{R}_i$. Check for the rank condition $rank(\mathbf{H}) = rank([\mathbf{H}\ \mathbf{Y}]) = n^2$ from Eq. (3.33). If the condition is not satisfied continue to Step 8, and if satisfied then continue to Step 9.
8. Reinitialize the optimization phase with a different set of $\mathbf{Q}_i, \mathbf{R}_i$, by resetting the iteration counter $k = 0,\ \mathbf{K}_{i+1,0} = \mathbf{K}_{i,N_i},\ i = i + 1$ and continue to Step 4.
9. Use Eq.(3.33) to calculate the identified internal dynamics $\hat{\mathbf{A}}$.

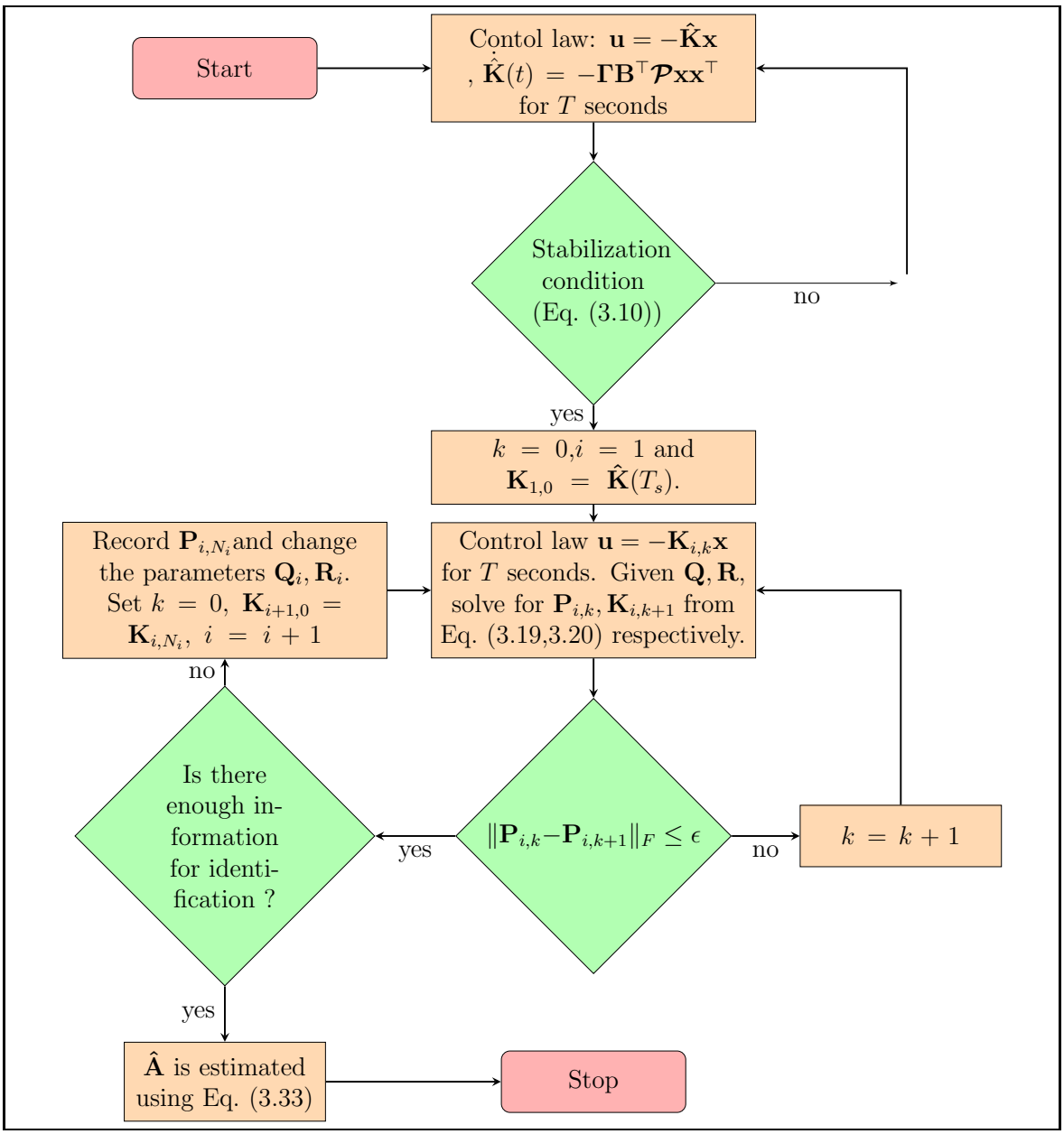

Figure 3.1. Flowchart for Online Implementation - linear, unknown $\mathbf{A}$.

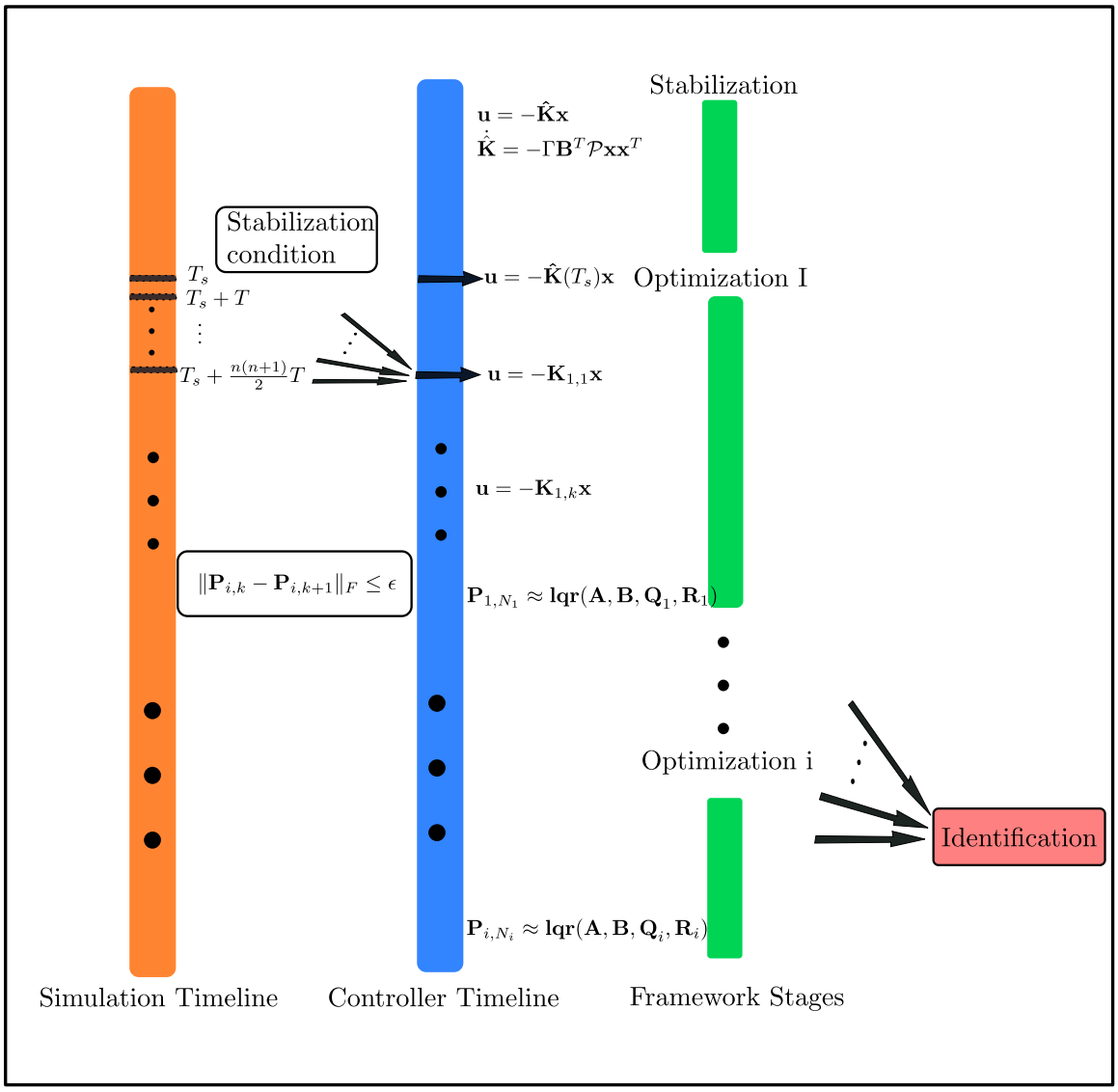

Figure 3.2. Timelines for framework operation - linear, unknown $\mathbf{A}$.

3.7 Simulation results

A continuous time simulation is setup to implement the three stages of learning. The dynamic model is chosen to be a linearized lateral dynamics of Harrier AV-8B from [40]. This linearization is valid at an airspeed of $50ft/s$ at an altitude $50ft$ above sea level. Four simulated states are lateral velocity in body frame v in ft/s, body axis roll angular velocity p in deg/s, body axis yaw angular velocity r in deg/s, roll angle ϕ in deg. The LQR weights are chosen to be identity matrices. The linearized model is unstable in the absence of control. The initial estimates for feedback gains are set to $\mathbf{0}$. The simulation represents a scenario in which the controller regulates the lateral oscillations of a Harrier AV-8B in near-hover conditions close to sea level.

$$\mathbf{A} = \begin{bmatrix} -0.0283 & 0.1823 & -0.8588 & 0.5493 \\ 0.0414 & -0.6662 & 0.2962 & 0 \\ -0.1926 & -0.0447 & -0.0891 & 0 \\ 0 & 1 & 0.2125 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3070 & 0.9224 \\ 1.0060 & -1.4070 \\ 0 & 0 \end{bmatrix}$$

The simulation demonstrates the identification and closed loop stabilization of a linear time invariant system without the knowledge of internal dynamics. Initial conditions of states are set to [0 ft/s 0 deg/s 0 deg/s 4 deg].

Fig. [3.3] shows the state history for a case when $\gamma = 10$. The Stabilization phase lasts till about $t = 1$ seconds. The plots clearly show that the states are regulated as desired. A stable reference model characterized by $\mathbf{A_m}$ is chosen to accommodate the structural flexibility requirements.

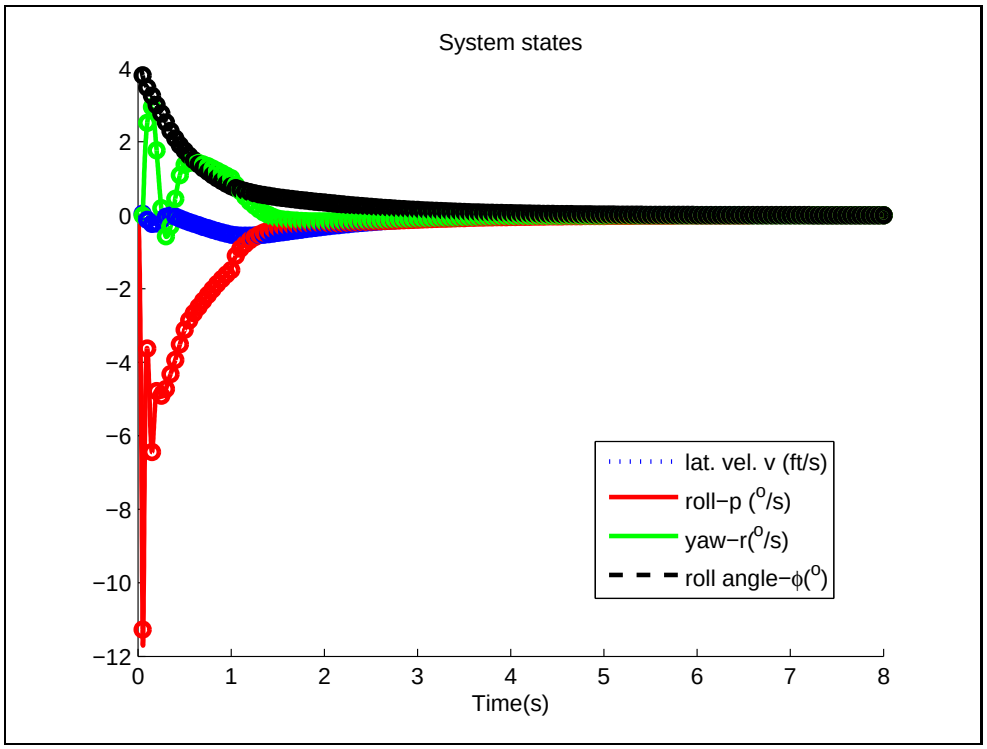

Figure 3.3. Closed Loop system response - linear, unknown $\mathbf{A}$.

$$\mathbf{A_m} = \begin{bmatrix} -0.10 & 0.1 & -1 & 0.5 \\ -7 & -15 & 1.5 & -10 \\ 0.5 & -0.8 & -2 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Fig. [3.4] shows the control history which is continuous in the stabilization phase, but has discrete updates during both optimization phases.

Note that the open loop eigen values of $\mathbf{A}$ are at $(-0.6391 \pm 0.2167i, 0.2473 \pm 0.1913i)$. By setting $T_s = 1s$, the stabilizing gain $\hat{\mathbf{K}}(T_s)$ moves the closed loop poles to $(-13.01, -0.21, -1.93 \pm 0.74i)$.

$$\hat{\mathbf{K}}(T_s) = \begin{bmatrix} 0.0018 & -0.9143 & -0.1214 & -1.5473 \\ -0.4591 & -1.0369 & 1.4956 & -0.0078 \end{bmatrix}$$

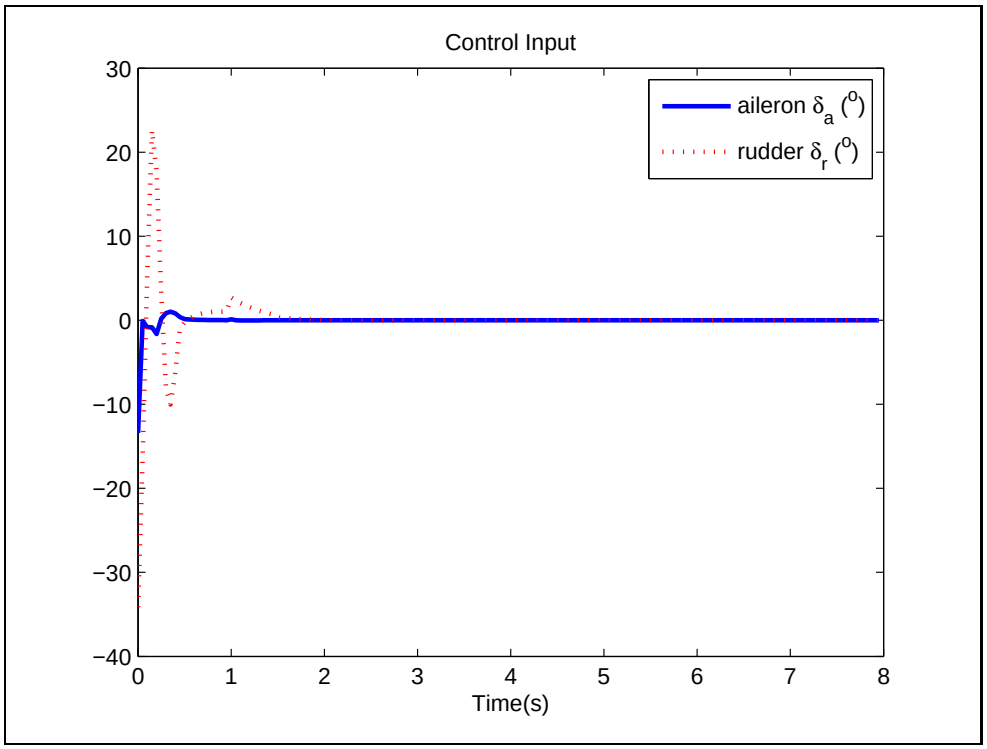

Figure 3.4. Control input history - linear, unknown $\mathbf{A}$.

The LQR weights are chosen to be identity matrices for first phase of optimization $(\mathbf{Q}_1, \mathbf{R}_1)$. Fig. [3.5] shows the Frobenius norm of error $\mathbf{P}_{1,k} - \mathbf{P}_1^*$ for the first phase of optimization. Note that an update to the policy is made only after data collection over 10 samples. The plot shows convergence of $\mathbf{P}_{1,k}$ to the optimal solution given by LQR weights $\mathbf{Q}_1, \mathbf{R}_1$. The limiting solution is recorded as $\mathbf{P}_{1,N_1}$. This phase lasts till $t = 4$ seconds where the tolerance condition on $\mathbf{P}_{1,k}$ update is satisfied.

A change in $\mathbf{Q}, \mathbf{R}$ parameters is introduced by setting $\mathbf{Q}_2 = 4\mathbf{Q}_1,\ \mathbf{R}_2 = 0.5\mathbf{R}_1$. Fig. [3.6] shows the Frobenius norm of error $\mathbf{P}_{1,k} - \mathbf{P}_1^*$ for the second phase of optimization. The plot shows convergence of $\mathbf{P}_{2,k}$ to the optimal solution given by LQR

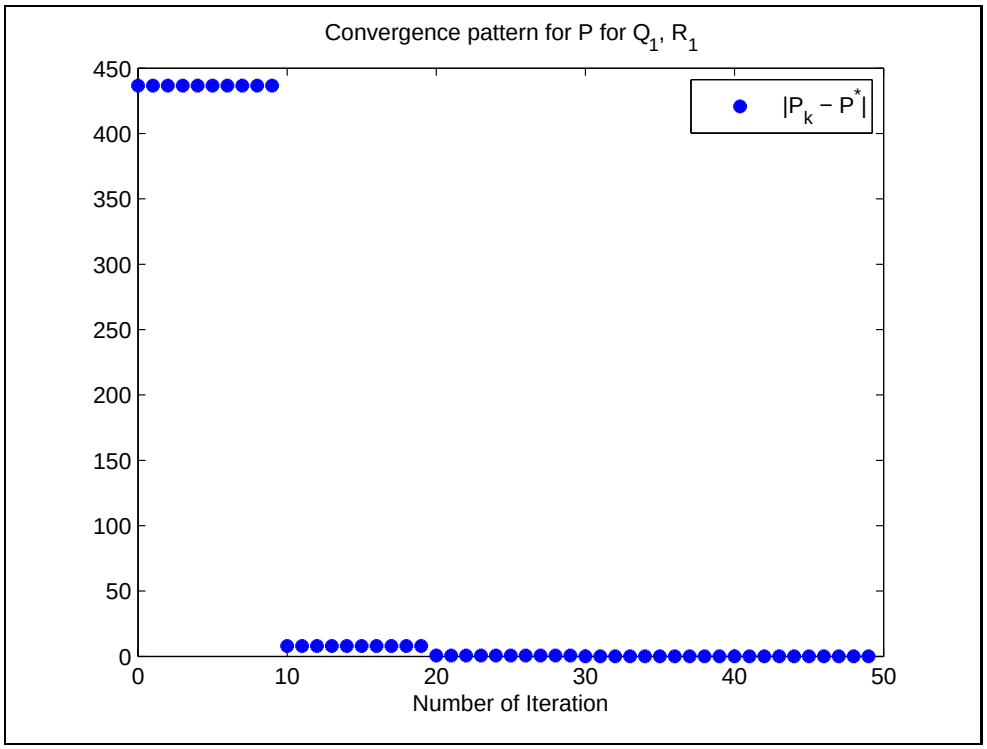

Figure 3.5. Iteration history of $\mathbf{P}_k$ for $\mathbf{Q}_1, \mathbf{R}_1$ - linear, unknown $\mathbf{A}$.

weights $\mathbf{Q}_2, \mathbf{R}_2$. The limiting solution is recorded as $\mathbf{P}_{2,N_2}$. This phase lasts till the end of simulation at $t = 8$ seconds.

$$\mathbf{P}_1^* = \begin{bmatrix} 1.4416 & 0.0726 & -0.5330 & 0.3965 \\ 0.0726 & 0.0770 & -0.0604 & 0.0894 \\ -0.5330 & -0.0604 & 0.8477 & -0.1021 \\ 0.3965 & 0.0894 & -0.1021 & 1.2013 \end{bmatrix}$$

$$\mathbf{P}_{1,N_1} = \begin{bmatrix} 1.4422 & 0.0726 & -0.5334 & 0.3967 \\ 0.0726 & 0.0770 & -0.0604 & 0.0895 \\ -0.5334 & -0.0604 & 0.8478 & -0.1024 \\ 0.3967 & 0.0895 & -0.1024 & 1.2013 \end{bmatrix}$$

It is observed that the difference between the approximation $\mathbf{P}_{1,N_1}$ and the actual solution to the ARE ($\mathbf{P}_1^*$) has a Frobenius norm of 1×10^{-3}. The second policy

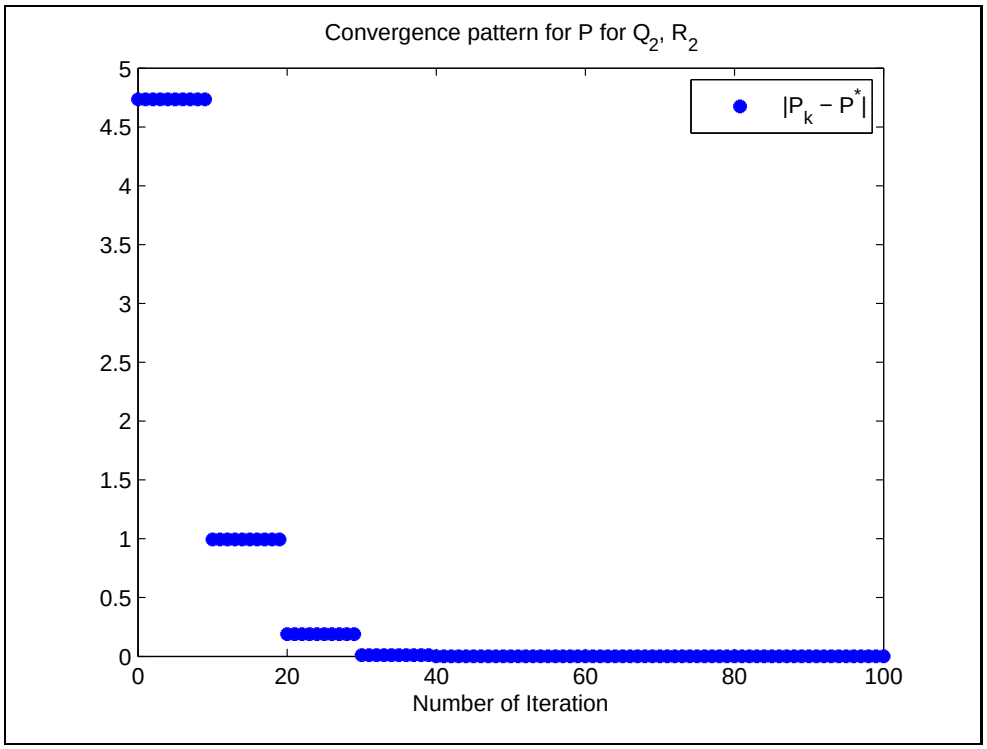

Figure 3.6. Iteration history of $\mathbf{P}_k$ for $\mathbf{Q}_2, \mathbf{R}_2$ - linear, unknown $\mathbf{A}$.

iteration converges to $\mathbf{P}_{2,N_2}$, whereas the actual solution to the ARE ($\mathbf{P}_2^*$) differs by a Frobenius norm of about 7×10^{-5}.

$$\mathbf{P}_2^* = \begin{bmatrix} 4.6058 & 0.0843 & -0.5592 & 1.2426 \\ 0.0843 & 0.1040 & -0.0614 & 0.1173 \\ -0.5592 & -0.0614 & 1.0102 & -0.0401 \\ 1.2426 & 0.1173 & -0.0401 & 4.5084 \end{bmatrix}$$

$$\mathbf{P}_{2,N_2} = \begin{bmatrix} 4.6059 & 0.0843 & -0.5592 & 1.2426 \\ 0.0843 & 0.1040 & -0.0614 & 0.1173 \\ -0.5592 & -0.0614 & 1.0102 & -0.0401 \\ 1.2426 & 0.1173 & -0.0401 & 4.5084 \end{bmatrix}$$

The closed form solution of $\mathbf{A}$ is solved from the approximated solutions $\mathbf{P}_{1,N_1}, \mathbf{P}_{2,N_2}$ using Eq. (3.33).

$$\hat{\mathbf{A}} = \begin{bmatrix} -0.0247 & 0.1823 & -0.8583 & 0.5640 \\ 0.0757 & -0.6620 & 0.2608 & -0.0016 \\ -0.1897 & -0.0416 & -0.0910 & 0.0114 \\ -0.0142 & 1.0001 & 0.2123 & -0.0038 \end{bmatrix}$$

The approximation $\hat{\mathbf{A}}$ has eigen values at $(-0.6371 \pm 0.2436i, 0.2464 \pm 0.1585i)$ which are close to their counterparts of $\mathbf{A}$ at $(-0.6391 \pm 0.2167i, 0.2473 \pm 0.1913i)$.

CHAPTER 4

Solution for unknown linear dynamics

This chapter develops ideas mentioned in Ch. 2 and explicitly formulates the controller for unknown dynamics case (unknown $\mathbf{A}$, unknown $\mathbf{B}$, $\mathbf{f}(\mathbf{x}) = \mathbf{0}$) as shown below. In addition to the implementation details further analysis is presented for the stabilizing controller which also identifies the unknown matrices $\mathbf{A}, \mathbf{B}$.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{4.1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$ represent the state and control vectors respectively.

4.1 Solution Methodology - Stabilize

First phase of the controller involves a Model Reference Adaptive Control (MRAC) approach which solves a tracking problem for the unknown dynamic system . $\mathbf{A_m} \in \mathbb{R}^{n\times n}$ is chosen to be a Hurwitz matrix, $\mathbf{B_m} \in \mathbb{R}^{n\times m}$ is chosen such that $\exists \mathbf{K}^* \in \mathbb{R}^{m\times n}, \mathbf{L}^* \in \mathbb{R}^{m\times m}$ satisfying $\mathbf{A_m} = \mathbf{A} - \mathbf{B}\mathbf{K}^*, \mathbf{B}\mathbf{L}^* = \mathbf{B_m}$.

The reference model is characterized by $\mathbf{A_m}, \mathbf{B_m}$, where $\mathbf{r} \in \mathbb{R}^m$ represents the given reference input signal for tracking.

$$\dot{\mathbf{x}}_\mathbf{m}(t) = \mathbf{A_m}\mathbf{x_m}(t) + \mathbf{B_m}\mathbf{r}(t) \tag{4.2}$$

The error ($\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{x_m}(t)$) between states of the unknown dynamic system and the chosen reference model is minimized. The closed loop error dynamics can be characterized as shown in Eq. (4.3) with state feedback control law $\mathbf{u}(t) =$

$-\hat{\mathbf{K}}(t)\mathbf{x}(t)+\hat{\mathbf{L}}(t)\mathbf{r}(t)$ where $\tilde{\mathbf{K}}(t)=\mathbf{K}^*-\hat{\mathbf{K}}(t),\ \tilde{\mathbf{L}}(t)=\mathbf{L}^*-\hat{\mathbf{L}}(t)$ represent the errors in the estimation of unknown ideal gains $\mathbf{K}^*,\mathbf{L}^*$.

$$\begin{aligned}
\dot{\mathbf{e}}(t) &= \dot{\mathbf{x}}-\dot{\mathbf{x}}_{\mathbf{m}} \\
&= (\mathbf{A_m}+\mathbf{BK}^*)\mathbf{x}(t)+\mathbf{Bu}(t)-\mathbf{A_m x_m}-\mathbf{B_m r} \\
&= \mathbf{A_m e}(t)+\mathbf{B}(\tilde{\mathbf{K}}(t)\mathbf{x}(t)-\tilde{\mathbf{L}}(t)\mathbf{r}(t)) \\
&= \mathbf{A_m e}(t)+\mathbf{B_m L}^{*-\mathbf{1}}(\tilde{\mathbf{K}}(t)\mathbf{x}(t)-\tilde{\mathbf{L}}(t)\mathbf{r}(t)) \qquad (4.3)
\end{aligned}$$

The adaptive laws for $\hat{\mathbf{K}}(t),\hat{\mathbf{L}}(t)$ follow from a straightforward Lyapunov analysis by choosing the candidate Lyapunov function to be as follows.

$$V=\mathbf{e}^\top(t)\boldsymbol{\mathcal{P}}\mathbf{e}(t)+\mathrm{Tr}(\tilde{\mathbf{K}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{K}}(t)+\tilde{\mathbf{L}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{L}}(t))$$

Note that $\boldsymbol{\mathcal{P}}\in\mathbb{R}^{n\times n},\boldsymbol{\Gamma}\in\mathbb{R}^{m\times m}$ are chosen symmetric positive definite matrices. Following the developments in [**?**], asymptotic stability of the closed loop dynamics in Eq. (4.3) can be shown by Lyapunov-Like Lemma(motivated by Barbalat's Lemma) for non-autonomous systems if the adaptive laws for $\hat{\mathbf{K}},\hat{\mathbf{L}}$ is chosen as

$$\dot{\hat{\mathbf{K}}}(t) = -\gamma sgn(\mathbf{L}^*)\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top \qquad (4.4)$$

$$\dot{\hat{\mathbf{L}}}(t) = \gamma sgn(\mathbf{L}^*)\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{r}^\top \qquad (4.5)$$

where $\boldsymbol{\mathcal{P}}$ is the solution to the Lyapunov equation $\mathbf{A_m}^\top\boldsymbol{\mathcal{P}}+\boldsymbol{\mathcal{P}}\mathbf{A_m}=-\boldsymbol{\mathcal{N}}$, for a chosen $\boldsymbol{\mathcal{N}}=\boldsymbol{\mathcal{N}}^\top>0$. It can be shown that the adaptive law along with $\boldsymbol{\Gamma}=\gamma sgn(\mathbf{L}^*)\mathbf{L}^{*-\mathbf{1}}$ results in $\dot{V}=-\mathbf{e}^\top(t)\boldsymbol{\mathcal{N}}\mathbf{e}(t)$

Note that the convergence of $\hat{\mathbf{K}}(t),\hat{\mathbf{L}}(t)$ to $\mathbf{K}^*,\mathbf{L}^*$ is not guaranteed. $\hat{\mathbf{K}}(t)$ does however converge to a stabilizing gain $\hat{\mathbf{K}}_\infty$ eventually. Thus the resulting closed loop system upon convergence to $\hat{\mathbf{K}}_\infty$ is asymptotically stable i.e. $\mathbf{A}-\mathbf{B}\hat{\mathbf{K}}_\infty$ is Hurwitz. Barbalat's Lemma can be applied since $\dot{V}$ is clearly uniformly continuous in time.

Note that the Lyapunov function V converges to a constant but not necessarily zero, whereas the derivative along the trajectory $\dot{V}$ vanishes as $t \to \infty$.

4.2 Exit condition for Stabilization phase

Although Barbalat's Lemma guarantees asymptotic stability as $t \to \infty$, ideally the adaptation should be continued for only a finite time. Assume that the MRAC style adaptation of feedback gain $\hat{\mathbf{K}}$ is ceased for time $t \geq T_s$ for some chosen T_s .

$$\dot{\mathbf{e}} \quad = \mathbf{A_m}\mathbf{e}(t) + \mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t))$$

Upon integration for feedback gain error $\tilde{\mathbf{K}}(t)$, and the state error $\mathbf{e}(t)$ before the adaptation is ceased $i.e\ \forall\ 0 \leq t \leq T_s$.

$$\tilde{\mathbf{K}}(t) \quad = \tilde{\mathbf{K}}(0) + \int_0^t \gamma sgn(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top ds \tag{4.6}$$

$$\tilde{\mathbf{L}}(t) \quad = \tilde{\mathbf{L}}(0) - \int_0^t \gamma sgn(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{r}^\top ds \tag{4.7}$$

$$\begin{aligned}\mathbf{e}(t) \quad &= \mathrm{e}^{\mathbf{A_m}t}\mathbf{e}(0) + \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{K}}(\mathbf{x}(\tau),\mathbf{e}(\tau))\mathbf{x}(\tau)d\tau \\ &\quad - \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{L}}(\mathbf{r}(\tau),\mathbf{e}(\tau))\mathbf{r}(\tau)d\tau\end{aligned} \tag{4.8}$$

The state relation obtained is implicit but it is not a kind of implicitness which can be dealt with Bellman-Gronwall Lemma. However $\mathbf{e}(T_s), \tilde{\mathbf{K}}(T_s)$ can be evaluated using Eq. (4.8,4.6). After the adaptation is ceased, feedback gain will remain constant $\hat{\mathbf{K}}(T_s)$. Thus the states are governed by linear time invariant dynamics after T_s. Explicit form for the state can be given $\forall\ t \geq T_s$ using Eq. (4.9).

$$
\begin{aligned}
\mathbf{e}(t) \quad &= \mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-T_s)}\mathbf{e}(T_s) \\
&+ \int_0^{\top} \mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-\tau)}\mathbf{B_m}\mathbf{L^{*-1}}(\tilde{\mathbf{K}}(T_s)\mathbf{x_m}(\tau) - \tilde{\mathbf{L}}(T_s)\mathbf{r}(\tau))d\tau \quad (4.9) \\
\|\mathbf{e}(t)\| \quad &\leq \|\mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-T_s)}\| \ (\|\mathbf{e}(T_s)\| + \\
&\|\mathbf{B_m}\mathbf{L^{*-1}}\|(\|\tilde{\mathbf{K}}(T_s)\|\|\mathbf{x_m}\| + \|\tilde{\mathbf{L}}(T_s)\|\|\mathbf{r}\|)) \\
&\leq \mathrm{e}^{\mu(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-T_s)} \ (\|\mathbf{e}(T_s)\| + \\
&\|\mathbf{B_m}\mathbf{L^{*-1}}\|(\|\tilde{\mathbf{K}}(T_s)\|\|\mathbf{x_m}\| + \|\tilde{\mathbf{L}}(T_s)\|\|\mathbf{r}\|))
\end{aligned}
\tag{4.10}
$$

Note that $\mathbf{e}(t) \in \mathbb{R}^n$ and $\|\mathbf{e}(t)\| \in \mathbb{R}$. The expression $\|\cdot\|$ for a square matrix should be interpreted as the induced norm from vector 2-norm $\|\cdot\|$. Above steps use triangle inequality for vector norms and the definition of induced norm for square matrices. Let $\mu(\mathbf{A})$ represent logarithmic norm of a matrix $\mathbf{A}$, and signifies the maximal growth rate of $\log\|\mathbf{x}\|$ if $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. The logarithmic norm properties include $\|\mathrm{e}^{\mathbf{P}t}\| \leq \mathrm{e}^{\mu(\mathbf{P})t}$, and $\mu(\mathbf{P}+\mathbf{Q}) \leq \mu(\mathbf{P}) + \|\mathbf{Q}\|$.

The upper limit on $\|\mathbf{e}(t)\|$ can also be obtained by using Bellman-Gronwall Lemma. If $\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s)) \leq -\delta$ is satisfied for some $\delta > 0$ and the norms $\|\mathbf{e}(T_s)\|$,$\|\mathbf{x_m}\|$,$\|\mathbf{r}\|$ are bounded, the state trajectories will be bounded by a decaying exponential.

Note that $\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s)) \leq -\delta \implies |\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))| \geq \delta$.

$$
\begin{aligned}
\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s) &= \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(0) + \gamma sgn(\mathbf{L}^*)\mathbf{B_m}\mathbf{L^{*-1}}\mathbf{B_m}^{\top}\boldsymbol{\mathcal{P}}\int_0^{T_s}\mathbf{x}(s)\mathbf{x}^{\top}(s)ds \\
\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s) &= \frac{1}{\gamma}sgn(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0) + \mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^{\top}\boldsymbol{\mathcal{P}}\int_0^{T_s}\mathbf{x}(s)\mathbf{x}^{\top}(s)ds \quad (4.11)
\end{aligned}
$$

If the condition $\int_0^{T_s}\|\mathbf{e}(s)\mathbf{e}^{\top}(s)\|ds \geq \frac{|\mu(\mathbf{A_m}+\frac{1}{\gamma}sgn(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0))|+\delta}{\|\mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^{\top}\boldsymbol{\mathcal{P}}\|}$ is satisfied, then the gain $(\hat{\mathbf{K}}(T_s))$ is stabilizing after the adaptation is ceased. Further it can be

established that $\|\mathbf{x}\mathbf{x}^\top\| = \|\mathbf{x}\|^2$(Consider an arbitrary vector $\mathbf{p} \in \mathbb{R}^n$, then $\|\mathbf{x}\mathbf{x}^\top\mathbf{p}\| = |\mathbf{x}^\top\mathbf{p}|\|\mathbf{x}\| \leq \|\mathbf{x}\|^2\|\mathbf{p}\|$).

$$\boxed{\int_0^{T_s} \|\mathbf{e}(s)\|^2 ds \geq \frac{|\mu(\mathbf{A_m} + \frac{1}{\gamma} sgn(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0))| + \delta}{\|\mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\|}} \tag{4.12}$$

Note that the lower limit from stabilizing condition cannot be explicitly obtained using known parameters. However such a positive limit can be calculated if the parameters were known. Thus a positive limit $\sigma > 0$ is chosen which yields a stabilizing condition which can be verified.

$$\int_0^{T_s} \|\mathbf{e}(s)\|^2 ds \geq \sigma > 0 \tag{4.13}$$

The stabilizing condition can also be interpreted as a lower limit on the decay of Lyapunov function V.

$$\begin{array}{rcll}
\lambda_{\min}(\boldsymbol{\mathcal{N}})\|\mathbf{e}\|^2 & \leq \mathbf{e}^\top\boldsymbol{\mathcal{N}}\mathbf{e} & \leq \lambda_{\max}(\boldsymbol{\mathcal{N}})\|\mathbf{e}\|^2 & \forall \mathbf{e} \\
\dfrac{\mathbf{e}^\top\boldsymbol{\mathcal{N}}\mathbf{e}}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} & \leq \|\mathbf{e}\|^2 & \leq \dfrac{\mathbf{e}^\top\boldsymbol{\mathcal{N}}\mathbf{e}}{\lambda_{\min}(\boldsymbol{\mathcal{N}})} & \forall \mathbf{e} \\
\int_0^{T_s} \dfrac{-\dot{V}}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} ds & \leq \int_0^{T_s} \|\mathbf{e}\|^2 ds & \leq \int_0^{T_s} \dfrac{-\dot{V}}{\lambda_{\min}(\boldsymbol{\mathcal{N}})} ds & \\
\dfrac{V(0) - V(T_s)}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} & \leq \int_0^{T_s} \|\mathbf{e}\|^2 ds & \leq \dfrac{V(0) - V(T_s)}{\lambda_{\min}(\boldsymbol{\mathcal{N}})} &
\end{array} \tag{4.14}$$

4.3 Solution Methodology - Optimize

Second phase of the controller uses the adaptive scheme developed in [18] to recursively approach the optimal feedback gain. This construction leads to a control gain which converges to the optimal feedback gain for the Linear Quadratic regulator with unknown internal dynamics. The iterative algorithm proposed by [18] is used with $\mathbf{K}_0 = \hat{\mathbf{K}}(T_s)$ as the initial stabilizing gain. The result is an adaptive controller

which converges to the optimal feedback controller obtained from ARE in Eq. (4.18) without the knowledge of $\mathbf{A}, \mathbf{B}$.

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{4.15}$$

Assuming that the pair $(\mathbf{A}, \mathbf{B})$ is stabilizable, the infinite horizon linear quadratic regulator problem would be to find $\mathbf{u}^*(t)$.

$$\mathbf{u}^*(t) = \underset{\mathbf{u}(t),\ t\in[t_0,\infty]}{\operatorname{argmin}} V(t_0, \mathbf{x}(t_0), \mathbf{u}(t)) \tag{4.16}$$

The infinite horizon cost for the optimal control problem is posed as

$$V(\mathbf{x}(t_0), t_0) = \int_{t_0}^{\infty} (\mathbf{x}^\top(\tau)\mathbf{Q}\mathbf{x}(\tau) + \mathbf{u}^\top(\tau)\mathbf{R}\mathbf{u}(\tau))d\tau \tag{4.17}$$

where $\mathbf{Q} > \mathbf{0},\ \mathbf{R} > \mathbf{0}$, and the pair $((\mathbf{A}, \sqrt{\mathbf{Q}})$ is detectable. The solution to this particular optimal control problem is known to be a state feedback controller $\mathbf{u}(t) = -\mathbf{K}\mathbf{x}(t)$ and the gain $\mathbf{K} = \mathbf{R}^{-\mathbf{1}}\mathbf{B}^\top\mathbf{P}$ where $\mathbf{P}$ is the positive definite solution to the following Algebraic Ricatti Equation.

$$\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-\mathbf{1}}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} = \mathbf{0} \tag{4.18}$$

Of course the control law mentioned above can be synthesized if $\mathbf{A}, \mathbf{B}$ are known.

The cost-to-go with a stabilizing controller gain $\mathbf{K}$ can be written as

$$V(\mathbf{x}(t)) = \int_t^{\infty} \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}^\top\mathbf{R}\mathbf{K})\mathbf{x}(\tau)d\tau = \mathbf{x}^\top(t)\mathbf{P}\mathbf{x}(t) \tag{4.19}$$

where $\mathbf{P}$ is the solution of the following Lyapunov equation

$$(\mathbf{A} - \mathbf{B}\mathbf{K})^\top\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) = -(\mathbf{K}^\top\mathbf{R}\mathbf{K} + \mathbf{Q}) \tag{4.20}$$

The cost function can be incrementally written as

$$V(\mathbf{x}(t)) = \int_t^{t+T} \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}^\top\mathbf{R}\mathbf{K})\mathbf{x}(\tau)d\tau + V(\mathbf{x}(t+T)) \tag{4.21}$$

Some intermediate notation is defined which will help in postulating the algorithm. Consider the following index representations of $\mathbf{P} \equiv p_{ij}$, and $\mathbf{x} \equiv x_i$.

$\bar{\mathbf{P}} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$ is a vectorized minimal representation of symmetric $\mathbf{P} \in \mathbb{R}^{n \times n}$.

$$\bar{\mathbf{P}} = [p_{11}\ 2p_{12}\ ..\ 2p_{1n}\ p_{22}\ 2p_{23}\ ..\ p_{nn}]^\top \tag{4.22}$$

$\bar{\mathbf{x}} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$ is a minimal representation of the outer product $\mathbf{x} \otimes \mathbf{x}$.

$$\bar{\mathbf{x}} = [x_1^2\ x_1x_2\ ..x_1x_n\ x_2^2\ x_2x_3\ ..\ x_n^2]^\top \tag{4.23}$$

Above notation is used to propagate the quadratic forms, and note that $\mathbf{x}^\top \mathbf{P} \mathbf{x} = \bar{\mathbf{P}}^\top \bar{\mathbf{x}}$. Matrices $\delta_{\mathbf{xx}} \in \mathbb{R}^{l \times \frac{1}{2}n(n+1)}, \mathbf{I}_{\mathbf{xx}} \in \mathbb{R}^{l \times n^2}, \mathbf{I}_{\mathbf{xu}} \in \mathbb{R}^{l \times mn}$ are defined for l time intervals as below.

$$\delta_{\mathbf{xx}} = [\bar{\mathbf{x}}(t_1) - \bar{\mathbf{x}}(t_0)\ \bar{\mathbf{x}}(t_2) - \bar{\mathbf{x}}(t_1)\ ..\ \bar{\mathbf{x}}(t_{l-1}) - \bar{\mathbf{x}}(t_l)]^\top \tag{4.24}$$

$$\mathbf{I}_{\mathbf{xx}} = \left[\int_{t_0}^{t_1} \mathbf{x} \otimes \mathbf{x} d\tau\ \int_{t_1}^{t_2} \mathbf{x} \otimes \mathbf{x} d\tau\ ..\ \int_{t_{l-1}}^{t_l} \mathbf{x} \otimes \mathbf{x} d\tau \right]^\top \tag{4.25}$$

$$\mathbf{I}_{\mathbf{xu}} = \left[\int_{t_0}^{t_1} \mathbf{x} \otimes \mathbf{u} d\tau\ \int_{t_1}^{t_2} \mathbf{x} \otimes \mathbf{u} d\tau\ ..\ \int_{t_{l-1}}^{t_l} \mathbf{x} \otimes \mathbf{u} d\tau \right]^\top \tag{4.26}$$

For a chosen length of time interval T, sampling times $t_j = t_0 + jT\ \forall\ j \in \{1, 2, .., l\}$. Simulation data in the form of matrices $\delta_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}$ is collected in the presence of exploration noise and initial stabilizing gain. Data collection is continued until $rank([\mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}]) = \frac{n(n+1)}{2} + mn$ for all subsequent samples. This condition is ensured by the persistent excitation from the exploration noise $\mathbf{d}$.

Corollary 5 (*A special condition with known* $\mathbf{B}$): The rank condition can be relaxed to $rank([\mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}]) = \frac{n(n+1)}{2}$ for all subsequent samples if matrix $\mathbf{B}$ were known. This would be a value iteration case with unknown internal dynamics as opposed to the policy iteration presented in previous chapter.

The iterative scheme is based on the following vectorized equation.

$$\mathbf{\Theta}\begin{bmatrix} \bar{\mathbf{P}}_k \\ vec(\mathbf{K}_{k+1}) \end{bmatrix} = \mathbf{\Xi}_k \tag{4.27}$$

where $\mathbf{\Theta}_k$, $\mathbf{\Xi}_k$ are defined as follows.

$$\mathbf{\Theta}_k = [\delta_{\mathbf{xx}} \ -\mathbf{2I_{xx}}(\mathbf{I}_n \otimes \mathbf{K}_k^\top \mathbf{R}) - \mathbf{2I_{xu}}(\mathbf{I}_n \otimes \mathbf{R})] \tag{4.28}$$

$$\mathbf{\Xi}_k = -\mathbf{I_{xx}} vec(\mathbf{Q} + \mathbf{K}_k^\top \mathbf{R}\mathbf{K}_k) \tag{4.29}$$

The recursive relations from Eq. (4.28,4.29) are iterated starting from an initial stabilizing gain $\mathbf{K}_0$. Eq. (4.27) is solved for $\mathbf{P}_k, \mathbf{K}_{k+1}$ using pseudo inverse. The following iteration scheme is implemented online due to the presence of rank condition which also ensures convergence to optimal feedback without knowledge of $\mathbf{A}, \mathbf{B}$.

$$\begin{bmatrix} \bar{\mathbf{P}}_k \\ vec(\mathbf{K}_{k+1}) \end{bmatrix} = (\mathbf{\Theta}_k^\top \mathbf{\Theta}_k)^{-1} \mathbf{\Theta}_k^\top \mathbf{\Xi}_k \tag{4.30}$$

4.4 Solution Methodology - Identify

The first two phases stabilize and optimize a quadratic performance metric on a partially unknown linear system. The third and final identification phase uses the information gathered from the optimization phase to identify the unknown parameters $\mathbf{A}, \mathbf{B}$. Since the identification step involves optimization solutions for different LQR parameters, it is convenient to introduce a more accommodating notation for approximations of $\mathbf{P}$. The parameters $N, \mathbf{Q}, \mathbf{R}, \mathbf{P}_k, \mathbf{K}_k$ used in previous section for an optimization phase will be generalized for multiple optimization phases.

Let $\mathbf{P}_{i,k}$ be the solution of k^{th} iteration for LQR problem posed with parameters $\mathbf{Q}_i, \mathbf{R}_i$. Corresponding state feedback gain approximation for next iteration is represented as $\mathbf{K}_{i,k+1}$. Also note that the convergence for each optimization phase is indicated by $\|\mathbf{P}_{i,k} - \mathbf{P}_{i,k-1}\|_F < \epsilon$ which is true for all $k > N_i$. Thus $\mathbf{P}_{i,N_i}, \mathbf{K}_{i,N_i}$ represent the approximation of LQR solution for parameters $\mathbf{Q}_i, \mathbf{R}_i$

Let $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ represent the estimated value of unknown parameter $\mathbf{A}, \mathbf{B}$. Since the optimization phase ends when $\mathbf{P}_k$ converges to the optimal solution $\mathbf{P}$, the investigation for a closed form solution of $\mathbf{A}, \mathbf{B}$ starts with the inspection of Eq. (4.18). The governing Algebraic Ricatti Equation(ARE) reduces to the following Eq. (4.31), where $\hat{\mathbf{A}}$ is the only unknown.

$$\hat{\mathbf{A}}\mathbf{P}_N + \mathbf{P}_N\hat{\mathbf{A}}^\top = \mathbf{X} \tag{4.31}$$

$$\mathbf{K}_N = \mathbf{R}^{-1}\hat{\mathbf{B}}^\top\mathbf{P}_N \tag{4.32}$$

Note that Eq. (4.31) is a symmetric linear matrix equation and thus yields $\frac{n^2+n}{2}$ linear scalar equations for a dynamic system of order n. This poses an underdetermined system of linear equations in terms of n^2 unknown elements of $\hat{\mathbf{A}}$. The problem of insufficient information is solved by using another optimal solution for a new set of performance metrics $(\mathbf{Q}, \mathbf{R})$. Consider two cases yielding linear matrix equations with unknown matrix $\hat{\mathbf{A}}$.

$$\hat{\mathbf{A}}\mathbf{P}_{1,N_1} + \mathbf{P}_{1,N_1}\hat{\mathbf{A}}^\top = \mathbf{X}_1 \tag{4.33}$$

$$\hat{\mathbf{A}}\mathbf{P}_{2,N_2} + \mathbf{P}_{2,N_2}\hat{\mathbf{A}}^\top = \mathbf{X}_2 \tag{4.34}$$

where $\mathbf{P}_{1,N_1} > \mathbf{0}, \mathbf{P}_{2,N_2} > 0, \mathbf{X}_1, \mathbf{X}_2 \in \mathbb{R}^{n\times n}$. Note that the matrices $\mathbf{X}_i$ are evaluated as $\mathbf{P}_{i,N_i}\mathbf{B}\mathbf{R}_i^{-1}\mathbf{B}^\top\mathbf{P}_{i,N_i} - \mathbf{Q}_i$ for $i = 1, 2$. Equations (4.33) and (4.34) represent the optimal state feedback control problem for the same system with different set of $(\mathbf{Q}, \mathbf{R})$. Kronecker algebra [45] can be employed in formulating an analytic solution for $\hat{\mathbf{A}}$. Refer to Ch. A for a Primer in Kronecker Algebra.

The estimate of $\mathbf{B}$ is computed as an average of estimates from both solutions in case of two optimization phases. This result can be extended to p optimization phases accordingly.

$$\hat{\mathbf{B}} = \frac{1}{2}(\mathbf{P}_{1,N_1}^{-1}\mathbf{K}_{1,N_1}^{\top}\mathbf{R}_1 + \mathbf{P}_{2,N_2}^{-1}\mathbf{K}_{2,N_2}^{\top}\mathbf{R}_2) \tag{4.35}$$

The estimate for $\hat{\mathbf{B}}$ is used in the computation of estimate $\hat{\mathbf{A}}$. Upon vectorization, the linear matrix equation in Eq. (4.31) is reduced to a system of linear equations of order n^2 with the introduction of a known permutation matrix $\mathbf{\Pi}$(Refer to Ch. A).

$$\begin{aligned} (\mathbf{P}_N{}^{\top} \otimes \mathbf{I}_n)vec(\hat{\mathbf{A}}^{\top}) + (\mathbf{I}_n \otimes \mathbf{P}_N)vec(\hat{\mathbf{A}}) &= vec(\mathbf{X}) \\ \left[(\mathbf{P}_N \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_N)\right] vec(\hat{\mathbf{A}}) &= vec(\mathbf{X}) \end{aligned} \tag{4.36}$$

where $\mathbf{I}_n$ represents an identity matrix of the order n. Upon vectorizing the Eq. (4.33) and (4.34), they can be combined to solve for $\hat{\mathbf{A}}$.

$$\begin{bmatrix} (\mathbf{P}_{1,N_1} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_{1,N_1}) \\ (\mathbf{P}_{2,N_2} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_{2,N_2}) \end{bmatrix} vec(\hat{\mathbf{A}}) = \begin{bmatrix} vec(\mathbf{X}_1) \\ vec(\mathbf{X}_2) \end{bmatrix} \tag{4.37}$$

This set of overdetermined linear equations can be solved in a least squares sense using a psuedo-inverse to obtain a closed form solution for $\hat{\mathbf{A}}$.

$$vec(\hat{\mathbf{A}}) = \left(\mathbf{H}^{\top}\mathbf{H}\right)^{-\mathbf{1}} \mathbf{H}^{\top} \begin{bmatrix} vec(\mathbf{P}_{1,N_1}\hat{\mathbf{B}}\mathbf{R}_1{}^{-\mathbf{1}}\hat{\mathbf{B}}^{\top}\mathbf{P}_{1,N_1} - \mathbf{Q}_1) \\ vec(\mathbf{P}_{2,N_2}\hat{\mathbf{B}}\mathbf{R}_2{}^{-\mathbf{1}}\hat{\mathbf{B}}^{\top}\mathbf{P}_{2,N_2} - \mathbf{Q}_2) \end{bmatrix} \tag{4.38}$$

$$\text{where } \mathbf{H} = \begin{bmatrix} (\mathbf{P}_{1,N_1} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}^n \otimes \mathbf{P}_{1,N_1}) \\ (\mathbf{P}_{2,N_2} \otimes \mathbf{I}_n)\mathbf{\Pi} + (\mathbf{I}^n \otimes \mathbf{P}_{2,N_2}) \end{bmatrix}$$

This solution can be generalized in a case where employing only two sets of $(\mathbf{Q}, \mathbf{R})$ does not provide sufficient information to solve for a unique solution $\hat{\mathbf{A}}$. Let p be the number of pairs of $(\mathbf{Q}, \mathbf{R})$ employed in the solution of $\hat{\mathbf{A}}$

$$\boxed{\begin{aligned} vec(\hat{\mathbf{A}}) &= \left(\mathbf{H}^\top \mathbf{H}\right)^{-\mathbf{1}} \mathbf{H}^\top \mathbf{Y} \\ \text{where } \mathbf{H} &= \begin{bmatrix} \mathbf{H_1} \\ \mathbf{H_2} \\ \vdots \\ \mathbf{H_p} \end{bmatrix}, \ \mathbf{Y} = \begin{bmatrix} \mathbf{Y_1} \\ \mathbf{Y_2} \\ \vdots \\ \mathbf{Y_p} \end{bmatrix}, \\ \mathbf{H_i} &= (\mathbf{P}_{i,N_i} \otimes \mathbf{I}^n)\mathbf{\Pi} + (\mathbf{I}_n \otimes \mathbf{P}_{i,N_i}), \\ \mathbf{Y_i} &= vec(\mathbf{P}_{i,N_i} \hat{\mathbf{B}} {\mathbf{R}_i}^{-\mathbf{1}} \mathbf{B}^\top \mathbf{P}_{i,N_i} - \mathbf{Q}_i) \end{aligned}} \tag{4.39}$$

The existence of a solution is guaranteed if and only if the pairs $(\mathbf{Q}_i, \mathbf{R}_i)$ are such that $rank(\mathbf{H}) = rank([\mathbf{H}\ \mathbf{Y}]) = n^2$. Thus if two instances of optimal control solution are not enough to solve for a unique $\hat{\mathbf{A}}$, more instances can be incorporated into the solution.

4.5 Algorithm for online implementation

The algorithm for the three phase identification algorithm is postulated.

1. Employ the control law $\mathbf{u} = -\hat{\mathbf{K}}\mathbf{x}$ with adaptation law $\dot{\hat{\mathbf{K}}}(t) = -\gamma \text{sgn}(\mathbf{L}^*) {\mathbf{B_m}}^\top \boldsymbol{\mathcal{P}} \mathbf{e} \mathbf{x}^\top$ for sampling time interval T seconds.
2. Go back to Step 1 to stabilize for another T seconds if the stabilization condition in Eq. (4.13) is not satisfied. Continue to Step 3 if satisfied.
3. Initialize the optimization phase for the first time by setting $k = 0$,$i = 1$ and $\mathbf{K}_{1,0} = \hat{\mathbf{K}}(T_s)$. Set the LQR parameters $\mathbf{Q}_i, \mathbf{R}_i$
4. Employ the control law $\mathbf{u} = -\mathbf{K}_{i,k}\mathbf{x} + \mathbf{d}$ for the next sampling time interval T seconds. The state information is used to $\delta_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}$ from Eq. (4.24,4.25,4.26) respectively.

5. Continue to Step 6 if the rank condition rank($[\mathbf{I_{xx}}, \mathbf{I_{xu}}]$) = $\frac{n(n+1)}{2} + mn$ is met. Go back to Step 4 if not satisfied.
6. Iterate with k from equations 4.28,4.29 until $\|\mathbf{P}_{i,k} - \mathbf{P}_{i,k+1}\|_F \leq \epsilon$ where $\epsilon > 0$ is a predefined threshold for convergence.
7. Note down converging solution $\mathbf{P}_{i,N_i}$ corresponding to parameters $\mathbf{Q}_i, \mathbf{R}_i$. Check for the rank condition $rank(\mathbf{H}) = rank([\mathbf{H}\ \mathbf{Y}]) = n^2$ from Eq.(4.38). If the condition is not satisfied continue to Step 8, and if satisfied then continue to Step 9.
8. Set the parameters $\mathbf{Q}_{i+1}, \mathbf{R}_{i+1}$ and reinitialize the optimization phase by setting $k = 0, \mathbf{K}_{i+1,0} = \mathbf{K}_{i,N_i}$, $i = i + 1$ and continue to Step 4.
9. Use Eq.(4.38,4.35) to calculate the identified parameters $\hat{\mathbf{A}}, \hat{\mathbf{B}}$.

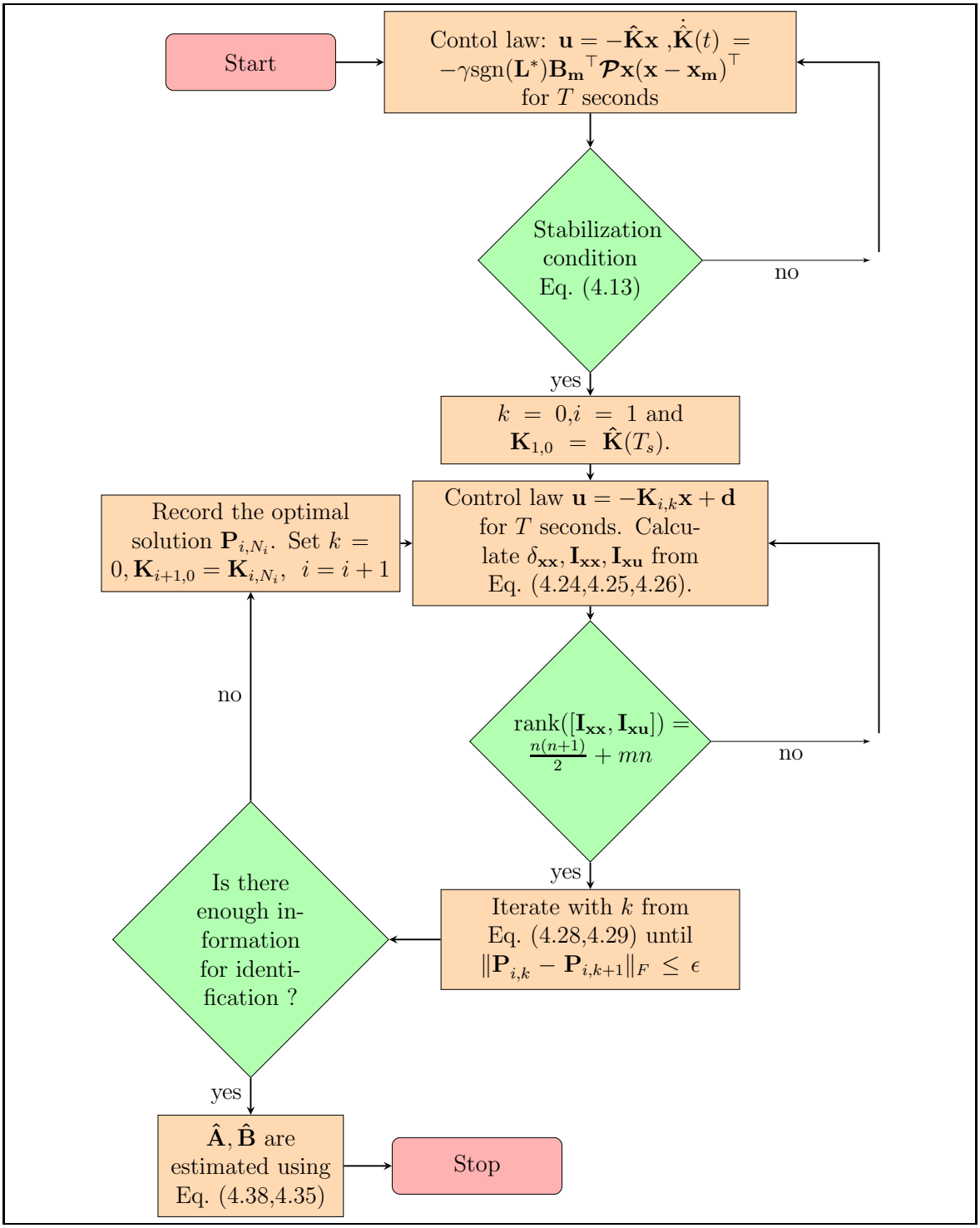

Figure 4.1. Flowchart for Online Implementation - linear, unknown $\mathbf{A}, \mathbf{B}$.

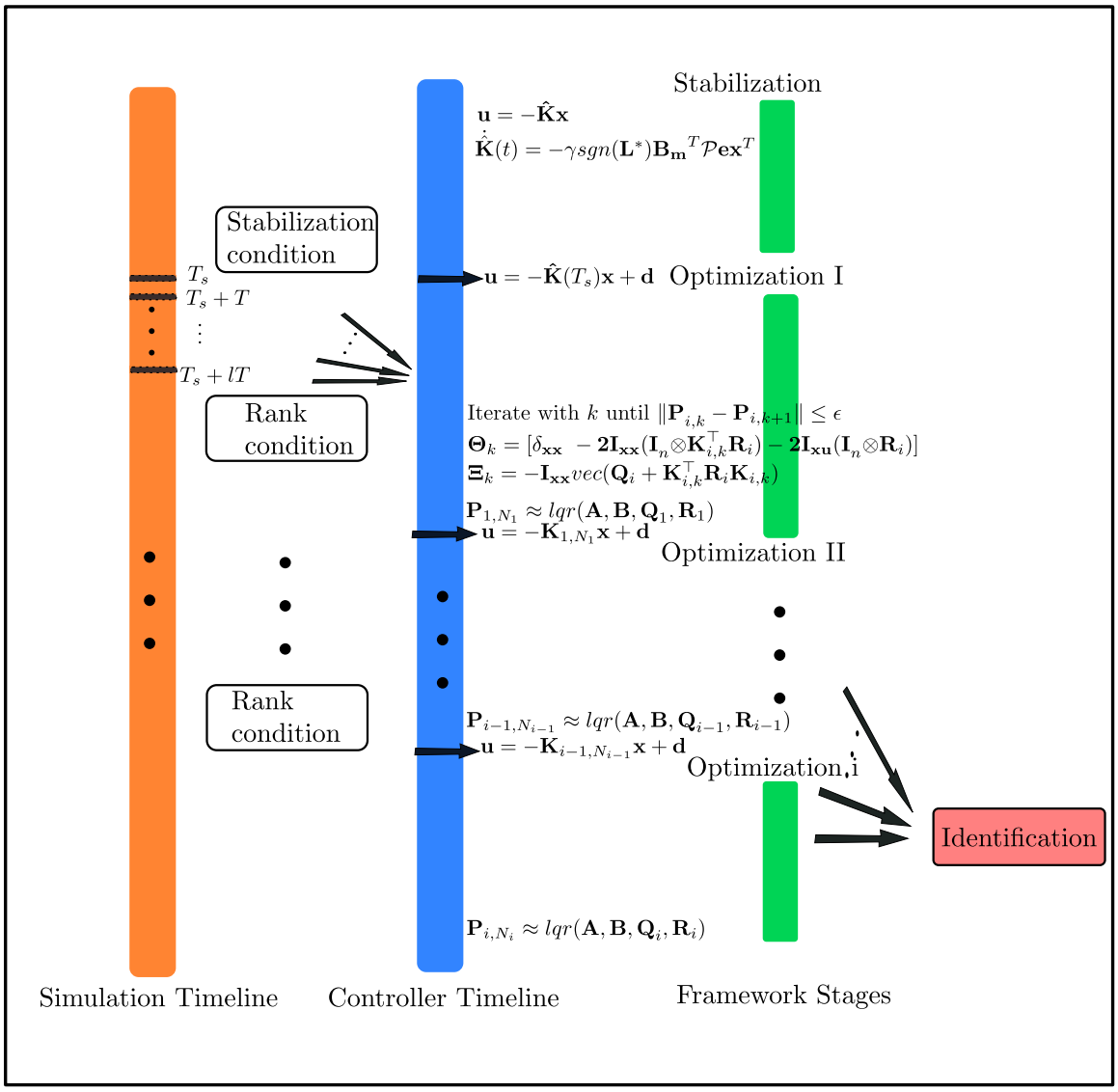

Figure 4.2. Timelines for framework operation - linear, unknown $\mathbf{A}, \mathbf{B}$.

4.6 Results

A continuous time simulation is setup to implement the three stages of learning. The dynamic model is chosen to be a linearized lateral dynamics of Harrier AV-8B from [40]. This linearization is valid at an airspeed of $50ft/s$ at an altitude $50ft$ above sea level. Four simulated states are lateral velocity in body frame v in ft/s, body axis roll angular velocity p in deg/s, body axis yaw angular velocity r in deg/s, roll angle ϕ in deg. The LQR weights are chosen to be identity matrices. The linearized model is unstable in the absence of control. The initial estimates for feedback gains are set to $\mathbf{0}$. The simulation represents a scenario in which the controller regulates the lateral oscillations of a Harrier AV-8B in near-hover conditions close to sea level.

$$\mathbf{A} = \begin{bmatrix} -0.0283 & 0.1823 & -0.8588 & 0.5493 \\ 0.0414 & -0.6662 & 0.2962 & 0 \\ -0.1926 & -0.0447 & -0.0891 & 0 \\ 0 & 1 & 0.2125 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3070 & 0.9224 \\ 1.0060 & -1.4070 \\ 0 & 0 \end{bmatrix}$$

The simulation demonstrates the identification and closed loop stabilization of a linear time invariant system without the knowledge of matrices $\mathbf{A}, \mathbf{B}$. The reference input $\mathbf{r}(t)$ is set to $\mathbf{0}$ to simulate a regulation case. Initial conditions of states are set to [0 ft/s 0 deg/s 0 deg/s 2 deg].

Fig. [4.3] shows the state history for a case when $\gamma = 10$. The Stabilization phase lasts till about $t = 2$ seconds. The plots clearly show that the states are regulated as desired. A stable reference model characterized by $(\mathbf{A_m}, \mathbf{B_m})$ is chosen to accommodate the structural flexibility requirements. The only information used in the controller formulation is that $\text{sgn}(\mathbf{L}^*) = +1$.

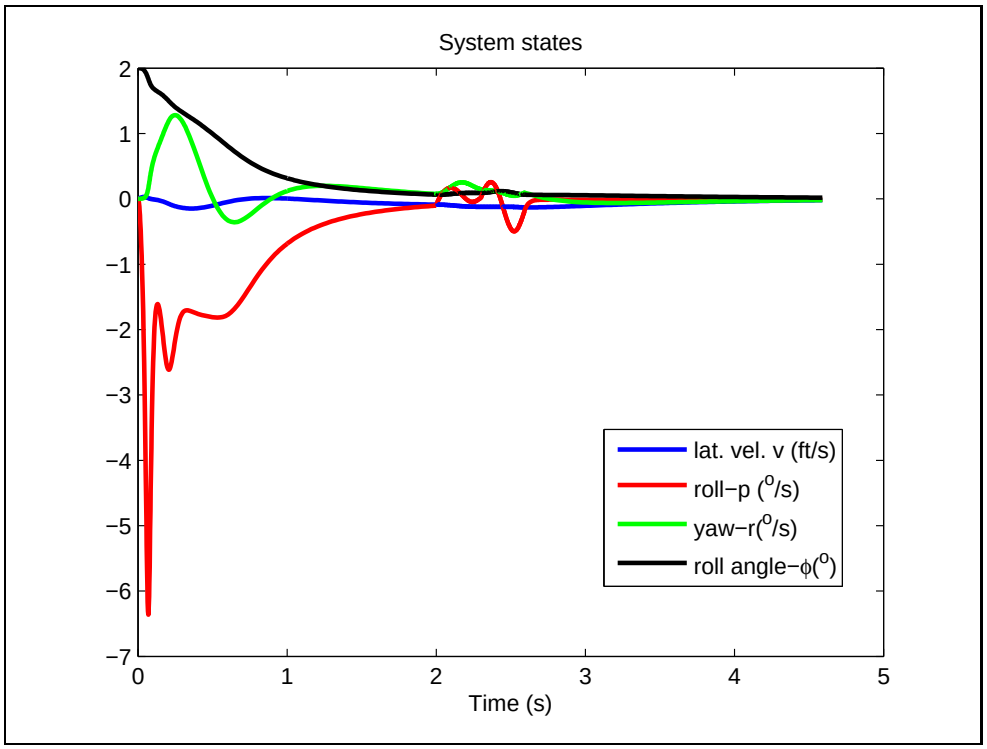

Figure 4.3. Closed Loop system response - linear, unknown $\mathbf{A}, \mathbf{B}$.

$$\mathbf{A_m} = \begin{bmatrix} -0.10 & 0.1 & -1 & 0.5 \\ -7 & -15 & 1.5 & -10 \\ 0.5 & -0.8 & -2 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{B_m} = \begin{bmatrix} 0.0536 & 0.1669 \\ 29.5364 & 16.1518 \\ 0.6050 & -1.8080 \\ 0 & 0 \end{bmatrix} \tag{4.40}$$

Fig. [4.4] shows the control history which is continuous for the stabilization phase, but exhibits discrete updates due to the optimization phases. Note that the open loop eigen values of $\mathbf{A}$ are at $(-0.6391 \pm 0.2167i, 0.2473 \pm 0.1913i)$. By setting $T_s = 2s$, the stabilizing gain $\hat{\mathbf{K}}(T_s)$ moves the closed loop poles to $(-30.87, -0.07, -3.2, -1.39)$.

$$\hat{\mathbf{K}}(T_s) = \begin{bmatrix} 0.0049 & 2.1728 & 0.3314 & 3.1665 \\ 0.2112 & 0.5760 & -1.9938 & 1.9219 \end{bmatrix} \tag{4.41}$$

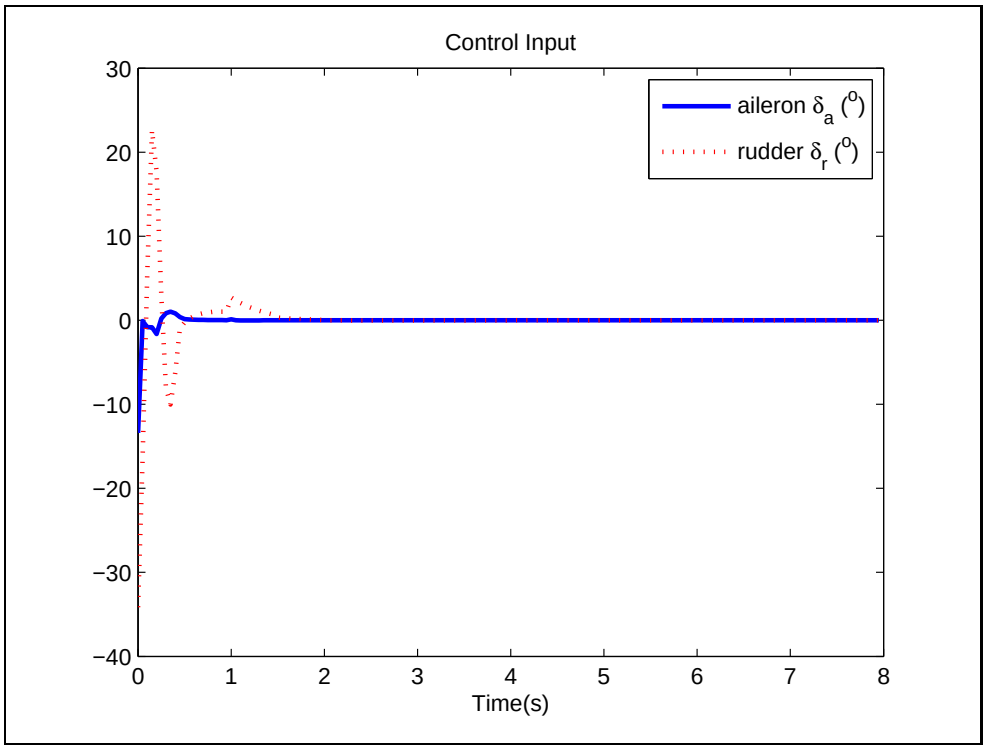

Figure 4.4. Control input history - linear, unknown $\mathbf{A}, \mathbf{B}$.

The LQR weights are chosen to be identity matrices for first phase of optimization $(\mathbf{Q}_1, \mathbf{R}_1)$. Note that a band limited random sinusoid is added to the control input during the optimization phases. This exploration noise is necessary for the convergence to optimal solution. The noise $\mathbf{d}$ is introduced into the control input as shown in Eq. (4.42) where ω_i are uniformly distributed random frequencies in the range $[-25, 25]Hz$.

$$\mathbf{u}(t) = -\mathbf{K}(T_s)\mathbf{x}(t) + 0.1\sum_{i=1}^{200}\sin(\omega_i t) \tag{4.42}$$

Data is collected until the rank condition is satisfied. Recursive relations from Eq. (4.28,4.29) are used to calculated the optimal feedback gain $\mathbf{K}_{1,N_1}$ and corresponding $\mathbf{P}_{1,N_1}$. Fig. [4.5] shows the Frobenius norm of error $\mathbf{P}_{1,k} - \mathbf{P}^*$ during the recursion for first phase of optimization. The plot shows convergence of $\mathbf{P}_{1,k}$ to the

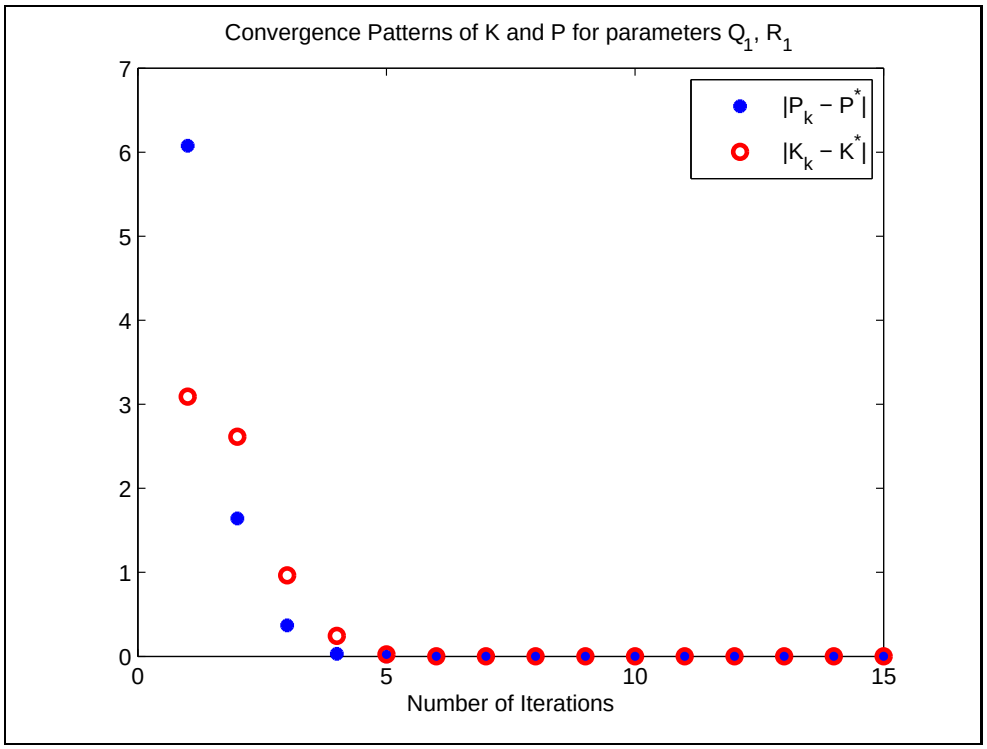

Figure 4.5. Iteration history of $\mathbf{P}_k, \mathbf{K}_k$ for $\mathbf{Q}_1, \mathbf{R}_1$ - linear, unknown $\mathbf{A}, \mathbf{B}$.

optimal solution given by LQR weights $\mathbf{Q}_1, \mathbf{R}_1$ which is recorded as $\mathbf{P}_{1,N_1}$. This phase lasts till about $t = 2.4$ seconds.

A change in $\mathbf{Q}, \mathbf{R}$ parameters is introduced by setting $\mathbf{Q}_2 = 4\mathbf{Q}_1,\ \mathbf{R}_2 = 0.5\mathbf{R}_1$. Fig. [4.6] shows the Frobenius norm of error $\mathbf{P}_{2,k} - \mathbf{P}_2^*$ during the recursion for second phase of optimization. The plot shows convergence of $\mathbf{P}_{2,k}$ to the optimal solution

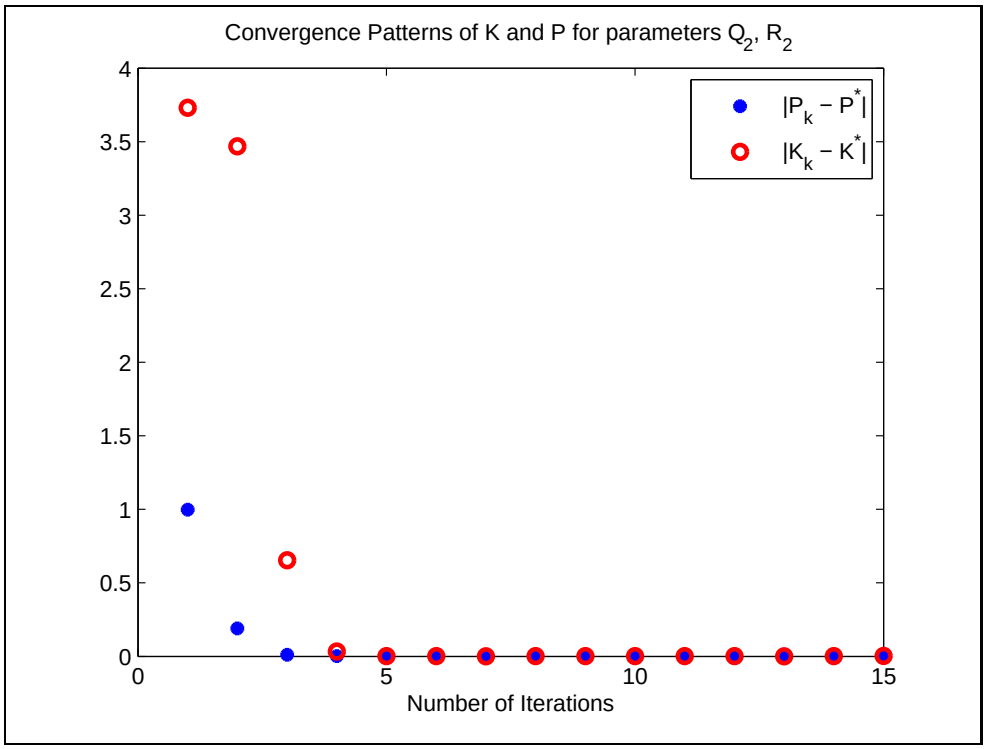

Figure 4.6. Iteration history of $\mathbf{P}_k, \mathbf{K}_k$ for $\mathbf{Q}_2, \mathbf{R}_2$ - linear, unknown $\mathbf{A}, \mathbf{B}$.

given by LQR weights $\mathbf{Q}_2, \mathbf{R}_2$ which is recorded as $\mathbf{P}_{2,N_2}$. This phase lasts till about $t = 2.8$ seconds.

$$\mathbf{P}_1^* = \begin{bmatrix} 1.4416 & 0.0726 & -0.5330 & 0.3965 \\ 0.0726 & 0.0770 & -0.0604 & 0.0894 \\ -0.5330 & -0.0604 & 0.8477 & -0.1021 \\ 0.3965 & 0.0894 & -0.1021 & 1.2013 \end{bmatrix}$$

$$\mathbf{P}_{1,N_1} = \begin{bmatrix} 1.4416 & 0.0726 & -0.5330 & 0.3963 \\ 0.0726 & 0.0770 & -0.0604 & 0.0894 \\ -0.5330 & -0.0604 & 0.8477 & -0.1021 \\ 0.3963 & 0.0894 & -0.1021 & 1.2011 \end{bmatrix}$$

It is observed that the difference between the approximation $\mathbf{P}_{1,N_1}$ and the actual solution to the ARE ($\mathbf{P}_1^*$) has a Frobenius norm of 2×10^{-4}. The second

phase of optimization converges to $\mathbf{P}_{2,N_2}$, whereas the actual solution to the ARE ($\mathbf{P}_2^*$) differs by a Frobenius norm of about 6×10^{-4}.

$$\mathbf{P}_2^* = \begin{bmatrix} 4.6058 & 0.0843 & -0.5592 & 1.2426 \\ 0.0843 & 0.1040 & -0.0614 & 0.1173 \\ -0.5592 & -0.0614 & 1.0102 & -0.0401 \\ 1.2426 & 0.1173 & -0.0401 & 4.5084 \end{bmatrix}$$

$$\mathbf{P}_{2,N_2} = \begin{bmatrix} 4.6059 & 0.0843 & -0.5593 & 1.2426 \\ 0.0843 & 0.1040 & -0.0614 & 0.1173 \\ -0.5593 & -0.0614 & 1.0101 & -0.0401 \\ 1.2426 & 0.1173 & -0.0401 & 4.5078 \end{bmatrix}$$

The closed form solution of $\mathbf{A}$ can be solved from the approximated solutions $\mathbf{P}_{1,N_1}, \mathbf{P}_{2,N_2}$ using Eq. (4.38).

$$\hat{\mathbf{A}} = \begin{bmatrix} -0.0303 & 0.1825 & -0.8591 & 0.5410 \\ 0.0329 & -0.6656 & 0.2932 & 0.0044 \\ -0.1932 & -0.0444 & -0.0894 & -0.0048 \\ 0.0080 & 0.9997 & 0.2128 & 0.0018 \end{bmatrix}$$

The approximation $\hat{\mathbf{A}}$ has eigen values at ($-0.6398 \pm 0.2022i$,$0.2481 \pm 0.1933i$) which are close to their counterparts of $\mathbf{A}$ at ($-0.6391 \pm 0.2167i, 0.2473 \pm 0.1913i$).

$$\mathbf{B} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3070 & 0.9224 \\ 1.0060 & -1.4070 \\ 0 & 0 \end{bmatrix}, \quad \hat{\mathbf{B}} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3071 & 0.9224 \\ 1.0060 & -1.4070 \\ -0.0001 & -0.0000 \end{bmatrix}$$

Tuning for convergence only requires the choice of a few parameters as mentioned. It is observed that a shorter stabilization phase results in diverging approximations for policy-iteration. The accuracy of identification phase is observed to be a strong function of the accuracy of solutions in the optimization phase.

CHAPTER 5

Solution for a Lipschitz nonlinear system

This section explicitly formulates the controller for systems with Lipschitz nonlinearities. In case of the nonlinear dynamic system shown below, it is assumed that $\mathbf{A}, \mathbf{B}$ are unknown along with nonlinear function $\mathbf{f}(\mathbf{x})$ for which the Lipschitz constant α is known (unknown $\mathbf{A}$, unknown $\mathbf{B}$, known α).

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{f}(\mathbf{x}) \tag{5.1}$$

where $\mathbf{x} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^p$ are state and control input respectively.

Assumption 1: The pair $(\mathbf{A}, \mathbf{B})$ is controllable. **Assumption 2:** The function $\mathbf{f}(\mathbf{x})$ is Lipschitz continuous ($\|\mathbf{f}(\mathbf{x})\| \leq \alpha\|\mathbf{x}\| \quad \forall \mathbf{x}$) such that there exists a Lipschitz constant $\alpha > 0$.

Before deriving a controller for the proposed case, the given system is studied for state bounds and controller design if $\mathbf{A}, \mathbf{B}$ were known.

5.1 State bounds

An explicit solution for state $\mathbf{x}(t)$ can be written and an upper bound for the norm of state vector can derived.

$$\begin{aligned}
\mathbf{x}(t) &= e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}(\mathbf{f}(\mathbf{x}) + \mathbf{B}\mathbf{u})d\tau \\
\|\mathbf{x}(t)\| &\leq \|e^{\mathbf{A}t}\|\|\mathbf{x}(0)\| + \int_0^t \|e^{\mathbf{A}(t-\tau)}\|(\|\mathbf{f}(\mathbf{x})\| + \|\mathbf{B}\mathbf{u}\|)d\tau \\
&\leq e^{\mu(\mathbf{A})t}\|\mathbf{x}(0)\| + \int_0^t e^{\mu(\mathbf{A})(t-\tau)}(\alpha\|\mathbf{x}\| + \|\mathbf{B}\|\|\mathbf{u}\|)d\tau
\end{aligned}$$

Examining the BIBO stability of the system at $\mathbf{0}$.

$$\|\mathbf{x}(t)\| \leq e^{\mu(\mathbf{A})t}\|\mathbf{x}(0)\| + \alpha e^{\mu(\mathbf{A})t}\int_0^t e^{-\mu(\mathbf{A})\tau}\|\mathbf{x}(\tau)\|d\tau \tag{5.2}$$

At this point Bellman Gronwall Lemma can be applied to get an explicit bound where $\|\mathbf{x}(t)\|$ is the considered non-negative function.

$$\begin{aligned}
\|\mathbf{x}(t)\| &\leq \|\mathbf{x}(0)\|e^{\mu(\mathbf{A})t} + \alpha e^{\mu(\mathbf{A})t}\int_0^t \|\mathbf{x}(0)\|e^{\int_s^t \alpha d\tau}ds \\
&\leq \|\mathbf{x}(0)\|e^{\mu(\mathbf{A})t} + \alpha e^{\mu(\mathbf{A})t}\int_0^t \|\mathbf{x}(0)\|e^{\alpha(t-s)}ds \\
&\leq \|\mathbf{x}(0)\|e^{\mu(\mathbf{A})t} + \alpha\|\mathbf{x}(0)\|e^{\mu(\mathbf{A}+\alpha)t} - \|\mathbf{x}(0)\|e^{\mu(\mathbf{A})t} \\
\|\mathbf{x}(t)\| &\leq \alpha\|\mathbf{x}(0)\|e^{\mu(\mathbf{A}+\alpha)t}
\end{aligned} \tag{5.3}$$

Norm of the state in absence of control is upper bounded by a decaying exponential if $\mathbf{A}+\alpha < 0$. The positive constant α for a stable system would have an upper bound $\alpha \leq -\mu(\mathbf{A})$.

5.2 Controller design with known parameters

If $\mathbf{A}, \mathbf{B}$ were known, an asymptotically stable controller can be derived using only the Lipschitz constant for unknown $\mathbf{f}(\mathbf{x})$. Consider a candidate Lyapunov function $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$ where $\mathbf{P} = \mathbf{P}^T > 0 \in \mathbb{R}^{n\times n}$. Augment the control signal with a stabilizing gain for the linear part $\mathbf{u} = \mathbf{v} - \mathbf{K}\mathbf{x}$ where $\mathbf{A}_c = \mathbf{A} - \mathbf{B}\mathbf{K}$ is Hurwitz. Note that the existence of a gain $\mathbf{K}$ can be guaranteed from the assumption of controllability of $(\mathbf{A}, \mathbf{B})$ pair.

$$\begin{aligned}
\dot{V} &= \mathbf{x}^T(\mathbf{A}_c^T\mathbf{P} + \mathbf{P}\mathbf{A}_c)\mathbf{x} + 2\mathbf{x}^T\mathbf{P}\mathbf{B}\mathbf{u} + 2\mathbf{x}^T\mathbf{P}\mathbf{f}(\mathbf{x}) \\
&\leq \mathbf{x}^T(\mathbf{A}_c^T\mathbf{P} + \mathbf{P}\mathbf{A}_c)\mathbf{x} + 2\mathbf{x}^T\mathbf{P}\mathbf{B}\mathbf{u} + 2\alpha\|\mathbf{P}\mathbf{x}\|\ \|\mathbf{x}\| \\
&\leq \mathbf{x}^T(\mathbf{A}_c^T\mathbf{P} + \mathbf{P}\mathbf{A}_c + \mathbf{P}^2 + \alpha^2\mathbf{I})\mathbf{x} + 2\mathbf{x}^T\mathbf{P}\mathbf{B}\mathbf{v}
\end{aligned}$$

Consider a control law $\mathbf{v} = -\mathbf{B}^T\mathbf{P}\mathbf{x}$.

$$\begin{aligned} \dot{V} &\leq \mathbf{x}^T(\mathbf{A}_c^T\mathbf{P} + \mathbf{P}\mathbf{A}_c + \mathbf{P}^2 + \alpha^2\mathbf{I})\mathbf{x} - \mathbf{x}^T\mathbf{P}(\mathbf{B}\mathbf{B}^T)\mathbf{P}\mathbf{x} \\ \dot{V} &\leq \mathbf{x}^T(\mathbf{A}_c^T\mathbf{P} + \mathbf{P}\mathbf{A}_c + \mathbf{P}(\mathbf{I} - \mathbf{B}\mathbf{B}^T)\mathbf{P} + \alpha^2\mathbf{I})\mathbf{x} \end{aligned} \tag{5.4}$$

The resulting closed loop system would be asymptotically stable if $\dot{V} \leq -\mathbf{x}^T\mathbf{Q}\mathbf{x}$ for a chosen $\mathbf{Q} > 0$ which in turn is ensured by the following condition.

$$\mathbf{A}_c^T\mathbf{P} + \mathbf{P}\mathbf{A}_c + \mathbf{P}(\mathbf{I} - \mathbf{B}\mathbf{B}^T)\mathbf{P} + \alpha^2\mathbf{I} = -\mathbf{Q} \tag{5.5}$$

The resulting Algebraic Riccati Equation has been studied extensively and necessary conditions for the existence of a solution $\mathbf{P} > 0$ have been established. Following equation represents the necessary condition for such existence.

$$\lambda_{\min}(\mathbf{I} - \mathbf{B}\mathbf{B}^T)tr(\alpha^2\mathbf{I} + \mathbf{Q}) - n\lambda_{\min}^2\left(\frac{\mathbf{A} + \mathbf{A}^T}{2}\right) < 0 \tag{5.6}$$

Choice of $\mathbf{Q}$ can greatly effect the existence of a positive definite solution. This would be a stabilizing controller design if $\mathbf{A}, \mathbf{B}, \alpha$ were known.

5.3 Solution Methodology - Stabilize

First phase of the controller involves a Model Reference Adaptive Control (MRAC) approach which solves a tracking problem for the unknown dynamic system . $\mathbf{A_m} + \alpha\mathbf{I} \in \mathbb{R}^{n\times n}$ is chosen to be a Hurwitz matrix, $\mathbf{B_m} \in \mathbb{R}^{n\times m}$ is chosen such that $\exists \mathbf{K}^* \in \mathbb{R}^{m\times n}, \mathbf{L}^* \in \mathbb{R}^{m\times m}$ satisfying $\mathbf{A_m} = \mathbf{A} - \mathbf{B}\mathbf{K}^*, \mathbf{B}\mathbf{L}^* = \mathbf{B_m}$.

The reference model is characterized by $\mathbf{A_m}, \mathbf{B_m}$, where $\mathbf{r} \in \mathbb{R}^m$ represents the given reference input signal for tracking.

$$\dot{\mathbf{x}}_\mathbf{m}(t) = (\mathbf{A_m} + \alpha\mathbf{I})\mathbf{x_m}(t) + \mathbf{B_m}\mathbf{r}(t) \tag{5.7}$$

The error ($\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{x_m}(t)$) between states of the unknown dynamic system and the given reference model are minimized. The closed loop error dynamics can be characterized as shown in Eq. (5.8) with state feedback control law $\mathbf{u}(t) = -\hat{\mathbf{K}}(t)\mathbf{x}(t) + \hat{\mathbf{L}}(t)\mathbf{r}(t)$ where $\tilde{\mathbf{K}}(t) = \mathbf{K}^* - \hat{\mathbf{K}}(t),\ \tilde{\mathbf{L}}(t) = \mathbf{L}^* - \hat{\mathbf{L}}(t)$ represent the errors in the estimation of unknown ideal gains $\mathbf{K}^*, \mathbf{L}^*$.

$$\begin{aligned}
\dot{\mathbf{e}}(t) &= \dot{\mathbf{x}} - \dot{\mathbf{x}}_\mathbf{m} \\
&= (\mathbf{A_m} + \mathbf{B}\mathbf{K}^*)\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) - (\mathbf{A_m} + \gamma\mathbf{I})\mathbf{x_m} - \mathbf{B_m}\mathbf{r} \\
&= \mathbf{A_m}\mathbf{e}(t) + \mathbf{f}(\mathbf{x}) - \alpha\mathbf{x_m} + \mathbf{B}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t)) \\
&= \mathbf{A_m}\mathbf{e}(t) + \mathbf{f}(\mathbf{x}) - \alpha\mathbf{x_m} + \mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t)) \qquad (5.8)
\end{aligned}$$

The adaptive laws for $\hat{\mathbf{K}}(t), \hat{\mathbf{L}}(t)$ follow from a straightforward Lyapunov analysis by choosing the candidate Lyapunov function as follows

$$V = \mathbf{e}^\top(t)\boldsymbol{\mathcal{P}}\mathbf{e}(t) + \mathrm{Tr}(\tilde{\mathbf{K}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{K}}(t) + \tilde{\mathbf{L}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{L}}(t))$$

Note that $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{n\times n}, \boldsymbol{\Gamma} \in \mathbb{R}^{m\times m}$ are chosen symmetric positive definite matrices. Following the developments in [**?**], asymptotic stability of the closed loop dynamics in Eq. (5.8) can be shown by Lyapunov-Like Lemma(motivated by Barbalat's Lemma) for non-autonomous systems if the adaptive laws for $\hat{\mathbf{K}}, \hat{\mathbf{L}}$ is chosen as

$$\dot{\hat{\mathbf{K}}}(t) = -\gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top \qquad (5.9)$$

$$\dot{\hat{\mathbf{L}}}(t) = \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{r}^\top \qquad (5.10)$$

Further $\boldsymbol{\mathcal{P}}$ can be solved which ensures asymptotic stability.

$$\begin{aligned}
\dot{V} &= \mathbf{e}^\top(\mathbf{A_m}^\top\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\mathbf{A_m})\mathbf{e} + 2\mathbf{e}^\top\boldsymbol{\mathcal{P}}(\mathbf{f}(\mathbf{x}) - \alpha\mathbf{x_m}) \\
&\leq \mathbf{e}^\top(\mathbf{A_m}^\top\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\mathbf{A_m})\mathbf{e} + 2\|\boldsymbol{\mathcal{P}}\mathbf{e}\|\|\alpha\mathbf{e}\| \\
&\leq \mathbf{e}^\top(\mathbf{A_m}^\top\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\mathbf{A_m} + \boldsymbol{\mathcal{P}}^2 + \alpha^2\mathbf{I})\mathbf{e} \qquad (5.11)
\end{aligned}$$

If $\boldsymbol{\mathcal{P}}$ is the solution to the Riccati equation $\mathbf{A_m}^\top\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\mathbf{A_m} + \boldsymbol{\mathcal{P}}^2 = -\boldsymbol{\mathcal{N}} - \alpha^2\mathbf{I}$, for a chosen $\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{N}}^\top > 0$. It can be shown that the adaptive law along with $\boldsymbol{\Gamma} = \gamma\text{sgn}(\mathbf{L}^*)\mathbf{L}^{*-\mathbf{1}}$ results in $\dot{V} \leq -\mathbf{e}^\top(t)\boldsymbol{\mathcal{N}}\mathbf{e}(t)$

Note that the convergence of $\hat{\mathbf{K}}(t), \hat{\mathbf{L}}(t)$ to $\mathbf{K}^*, \mathbf{L}^*$ is not guaranteed. $\hat{\mathbf{K}}(t)$ does however converge to a stabilizing gain $\hat{\mathbf{K}}_\infty$ eventually. Thus the resulting closed loop system upon convergence to $\hat{\mathbf{K}}_\infty$ is asymptotically stable i.e. $\mathbf{A} - \mathbf{B}\hat{\mathbf{K}}_\infty$ is Hurwitz. Barbalat's Lemma can be applied since $\dot{V}$ is clearly uniformly continuous in time. Note that the Lyapunov function V converges to a constant but not necessarily zero, whereas the derivative along the trajectory $\dot{V}$ vanishes as $t \to \infty$.

The Lipschitz nonlinearity can be compensated by a robust adaptive controller if the Lipschitz constant is given. Note that the Lyapunov function is parametrized by $\boldsymbol{\mathcal{P}}$ which in turn is solved from an Algebraic Riccati Equation.

5.4 Exit condition for Stabilization phase

Although Barbalat's Lemma guarantees asymptotic stability as $t \to \infty$, ideally the adaptation should be continued for only a finite time. Assume that the MRAC style adaptation of feedback gain $\hat{\mathbf{K}}$ is ceased for time $t \geq T_s$ for some chosen T_s .

$$\dot{\mathbf{e}} \quad = \mathbf{A_m}\mathbf{e}(t) + \mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t)) + \mathbf{f}(\mathbf{x}) - \alpha\mathbf{x_m}$$

Upon integration for feedback gain error $\tilde{\mathbf{K}}(t)$, and the state error $\mathbf{e}(t)$ before the adaptation is ceased *i.e* $\forall\ 0 \leq t \leq T_s$.

$$\begin{aligned}
\tilde{\mathbf{K}}(t) &= \tilde{\mathbf{K}}(0) + \int_0^t \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top ds && (5.12)\\
\tilde{\mathbf{L}}(t) &= \tilde{\mathbf{L}}(0) - \int_0^t \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{r}^\top ds && (5.13)\\
\mathbf{e}(t) &= \mathrm{e}^{\mathbf{A_m}t}\mathbf{e}(0) + \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{K}}(\mathbf{x}(\tau),\mathbf{e}(\tau))\mathbf{x}(\tau)d\tau && (5.14)\\
&\quad - \int_0^t \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{L}}(\mathbf{r}(\tau),\mathbf{e}(\tau))\mathbf{r}(\tau)d\tau + \int_0^t \mathrm{e}^{\mathbf{A_m}(t-\tau)}(\mathbf{f}(\mathbf{x}) - \alpha\mathbf{x_m})d\tau
\end{aligned}$$

The state relation obtained is implicit but it is not a kind of implicitness which can be dealt with Bellman-Gronwall Lemma. However $\mathbf{e}(T_s), \tilde{\mathbf{K}}(T_s)$ can be evaluated using Eq. (5.14,5.12). After the adaptation is ceased, feedback gain will remain constant $\hat{\mathbf{K}}(T_s)$. Thus the states are governed by linear time invariant dynamics after T_s. Explicit form for the state can be given $\forall\ t \geq T_s$ using Eq. (5.15).

$$\begin{aligned}
\mathbf{e}(t) &= \mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{K}}(T_s))(t-T_s)}\mathbf{e}(T_s) + \int_0^t \mathrm{e}^{\mathbf{A_m}(t-\tau)}(\mathbf{f}(\mathbf{x}) - \alpha\mathbf{x_m})d\tau\\
&\quad + \int_0^t \mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{K}}(T_s))(t-\tau)}\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}(\tilde{\mathbf{K}}(T_s)\mathbf{x_m}(\tau) - \tilde{\mathbf{L}}(T_s)\mathbf{r}(\tau))d\tau && (5.15)\\
\|\mathbf{e}(t)\| &\leq \|\mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{K}}(T_s))(t-T_s)+\alpha}\|\ (\alpha\|\mathbf{e}(T_s)\| +\\
&\quad \|\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\|(\|\tilde{\mathbf{K}}(T_s)\|\|\mathbf{x_m}\| + \|\tilde{\mathbf{L}}(T_s)\|\|\mathbf{r}\|)\\
\|\mathbf{e}(t)\| &\leq \mathrm{e}^{\mu(\mathbf{A_m}+\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\tilde{\mathbf{K}}(T_s))(t-T_s)+\alpha}\ (\alpha\|\mathbf{e}(T_s)\| +\\
&\quad \|\mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}\|(\|\tilde{\mathbf{K}}(T_s)\|\|\mathbf{x_m}\| + \|\tilde{\mathbf{L}}(T_s)\|\|\mathbf{r}\|))
\end{aligned} \qquad (5.16)$$

Note that $\mathbf{e}(t) \in \mathbb{R}^n$ and $\|\mathbf{e}(t)\| \in \mathbb{R}$. The expression $\|\cdot\|$ for a square matrix should be interpreted as the induced norm from vector 2-norm $\|\cdot\|$. Above steps use triangle inequality for vector norms and the definition of induced norm for square

matrices. Let $\mu(\mathbf{A})$ represent logarithmic norm of a matrix $\mathbf{A}$, and signifies the maximal growth rate of $\log \|\mathbf{x}\|$ if $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. The logarithmic norm properties include $\|e^{\mathbf{P}t}\| \leq e^{\mu(\mathbf{P})t}$, and $\mu(\mathbf{P}+\mathbf{Q}) \leq \mu(\mathbf{P}) + \|\mathbf{Q}\|$.

The upper limit on $\|\mathbf{e}(t)\|$ can also be obtained by using Bellman-Gronwall Lemma. If $\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s)) \leq -\delta$ is satisfied for some $\delta > 0$ and the norms $\|\mathbf{e}(T_s)\|, \|\mathbf{x_m}\|, \|\mathbf{r}\|$ are bounded, the state trajectories will be bounded by a decaying exponential.

$$\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s)) \leq -\delta - \alpha \implies |\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))| \geq \delta + \alpha.$$

$$\begin{aligned}
\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s) &= \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(0) + \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\mathbf{L^{*-1}}\mathbf{B_m}^\top \boldsymbol{\mathcal{P}} \int_0^{T_s} \mathbf{x}(s)\mathbf{x}^\top(s)ds \\
\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s) &= \frac{1}{\gamma}\mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0) + \mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top \boldsymbol{\mathcal{P}} \int_0^{T_s} \mathbf{x}(s)\mathbf{x}^\top(s)ds \qquad (5.17)
\end{aligned}$$

If the condition $\int_0^{T_s} \|\mathbf{e}(s)\mathbf{e}^\top(s)\|ds \geq \frac{|\mu(\mathbf{A_m}+\frac{1}{\gamma}\mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0))|+\delta+\alpha}{\|\mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\|}$ is satisfied, then the gain $(\hat{\mathbf{K}}(T_s))$ is stabilizing after the adaptation is ceased. Further it can be established that $\|\mathbf{x}\mathbf{x}^\top\| = \|\mathbf{x}\|^2$(Consider an arbitrary vector $\mathbf{p} \in \mathbb{R}^n$, then $\|\mathbf{x}\mathbf{x}^\top\mathbf{p}\| = |\mathbf{x}^\top\mathbf{p}|\|\mathbf{x}\| \leq \|\mathbf{x}\|^2\|\mathbf{p}\|$).

$$\boxed{\int_0^{T_s} \|\mathbf{e}(s)\|^2 ds \geq \frac{|\mu(\mathbf{A_m} + \frac{1}{\gamma}\mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0))| + \delta + \alpha}{\|\mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\|}} \qquad (5.18)$$

Note that the lower limit from stabilizing condition cannot be explicitly obtained using known parameters. However such a positive limit can be calculated if the parameters were known. Thus a positive limit $\sigma > 0$ is chosen which yields a stabilizing condition which can be verified.

$$\int_0^{T_s} \|\mathbf{e}(s)\|^2 ds \geq \sigma > 0 \qquad (5.19)$$

The stabilizing condition can also be interpreted as a lower limit on the decay of Lyapunov function V.

5.5 Solution Methodology - Optimize

If the unknowns $\mathbf{A}, \mathbf{B}, \mathbf{f}(\mathbf{x})$ were all known, a worst case optimal controller can be formulated using the Hamilton-Jacobi approach. Consider an infinite horizon quadratic cost function $J = \frac{1}{2}\int_0^\infty (\mathbf{x}(s)^\top \mathbf{Q}\mathbf{x}(s) + \mathbf{u}(s)^\top \mathbf{R}\mathbf{u}(s))ds$. Hamiltonian H can be formed by augmenting the integrand of cost function J with the dynamic constraint.

$$H = \frac{1}{2}(\mathbf{x}^\top \mathbf{Q}\mathbf{x} + \mathbf{u}^\top \mathbf{R}\mathbf{u}) + \boldsymbol{\lambda}^\top (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{f}(\mathbf{x})) \tag{5.20}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^n$ is the costate vector. The control law and costate dynamics are obtained by partial derivatives of Hamiltonian H.

$$\frac{\partial H}{\partial \mathbf{u}} \equiv \mathbf{0} \;\Rightarrow \mathbf{u} = \mathbf{R}^{-1}\mathbf{B}^\top \boldsymbol{\lambda} \tag{5.21}$$

$$\frac{\partial H}{\partial \mathbf{x}} + \dot{\boldsymbol{\lambda}} \equiv \mathbf{0} \;\Rightarrow -\dot{\boldsymbol{\lambda}} = \mathbf{Q}\mathbf{x} + \mathbf{A}^\top \boldsymbol{\lambda} + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)^\top \boldsymbol{\lambda}, \; \lim_{t\to\infty} \boldsymbol{\lambda}(t) = \mathbf{0} \tag{5.22}$$

Assuming linear feedback parametrized by $\mathbf{P} = \mathbf{P}^\top > \mathbf{0}$, by substituting $\boldsymbol{\lambda} = \mathbf{P}\mathbf{x}$. Note that the control law is governed by a linear state feedback $\mathbf{u} = -\mathbf{R}^{-1}\mathbf{B}^\top \mathbf{P}\mathbf{x}$, costate dynamics are governed by the following equation.

$$
\begin{aligned}
\dot{\boldsymbol{\lambda}} + \mathbf{Q}\mathbf{x} + \mathbf{A}^\top \boldsymbol{\lambda} + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)^\top \boldsymbol{\lambda} &= \mathbf{0} \\
\dot{\mathbf{P}}\mathbf{x} + \mathbf{P}\dot{\mathbf{x}} + \mathbf{Q}\mathbf{x} + \mathbf{A}^\top(\mathbf{P}\mathbf{x}) + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)^\top (\mathbf{P}\mathbf{x}) &= \mathbf{0} \\
\dot{\mathbf{P}}\mathbf{x} + \mathbf{P}(\mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P}\mathbf{x} + \mathbf{f}(\mathbf{x})) + \mathbf{Q}\mathbf{x} + \mathbf{A}^\top(\mathbf{P}\mathbf{x}) + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)^\top (\mathbf{P}\mathbf{x}) &= \mathbf{0} \\
\left(\dot{\mathbf{P}} + \mathbf{P}\mathbf{A} + \mathbf{A}^\top\mathbf{P} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)^\top \mathbf{P}\right)\mathbf{x} + \mathbf{P}\mathbf{f}(\mathbf{x}) &= \mathbf{0}
\end{aligned}
$$

If the above equation is solved for known $\mathbf{A}, \mathbf{B}, \mathbf{f}(\mathbf{x})$, a linear optimal controller can be synthesized. A static feedback solution (*i.e* $\dot{\mathbf{P}} = \mathbf{0}$) can be solved for the worst case (*i.e* $\mathbf{f}(\mathbf{x}) = \alpha\mathbf{x}$) if $\mathbf{f}(\mathbf{x})$ is unknown nonlinear function with Lipschitz constant α.

$$
\begin{aligned}
\left(\mathbf{P}\mathbf{A} + \mathbf{A}^\top\mathbf{P} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} + (\alpha\mathbf{I})^\top \mathbf{P}\right)\mathbf{x} + \mathbf{P}(\alpha\mathbf{I}\mathbf{x}) &= \mathbf{0} \\
\left(\mathbf{P}(\mathbf{A} + \alpha\mathbf{I}) + (\mathbf{A} + \alpha\mathbf{I})^\top\mathbf{P} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P} + \mathbf{Q}\right)\mathbf{x} &= \mathbf{0}
\end{aligned}
$$

Above relation holds $\forall \mathbf{x}$ if the governing Riccati Equation is satisfied. Thus an optimal controller for worst case unknown lipschitz term $\mathbf{f}(\mathbf{x})$ is obtained by the solution of an Algebraic Riccati Equation.

$$
\mathbf{P}(\mathbf{A} + \alpha\mathbf{I}) + (\mathbf{A} + \alpha\mathbf{I})^\top\mathbf{P} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} \tag{5.23}
$$

The formulation from Ch. 4 is valid for approximating the optimal controller without the knowledge of $(\mathbf{A} + \alpha\mathbf{I}), \mathbf{B}$. Refer to Ch. 4 for approximating the optimal controller without the knowledge of $\mathbf{A}, \mathbf{B}$

5.6 Solution Methodology - Identify

The first two phases stabilize and optimize a quadratic performance metric for worst case scenario on a partially unknown Lipschitz system. The third and final iden-

tification phase uses the information gathered from the optimization phase to identify the unknown parameters $\mathbf{A}, \mathbf{B}$. Since the identification step involves worst case optimization solutions for different LQR parameters, it is convenient to introduce a more accommodating notation for approximations of $\mathbf{P}$. The parameters $N, \mathbf{Q}, \mathbf{R}, \mathbf{P}_k, \mathbf{K}_k$ used in previous section for an optimization phase will be generalized for multiple optimization phases.

Let $\mathbf{P}_{i,k}$ be the solution of k^{th} iteration for LQR problem posed with parameters $\mathbf{Q}_i, \mathbf{R}_i$. Corresponding state feedback gain approximation for next iteration is represented as $\mathbf{K}_{i,k+1}$. Also note that the convergence for each optimization phase is indicated by $\|\mathbf{P}_{i,k} - \mathbf{P}_{i,k-1}\|_F < \epsilon$ which is true for all $k > N_i$. Thus $\mathbf{P}_{i,N_i}, \mathbf{K}_{i,N_i}$ represent the approximation of LQR solution for parameters $\mathbf{Q}_i, \mathbf{R}_i$.

The parameters $\mathbf{A}, \mathbf{B}$ are identified using the expressions from Ch. 4 by replacing $\hat{\mathbf{A}}$ with $\hat{\mathbf{A}} + \alpha\mathbf{I}$. Since the controller in optimization phase is assuming worst case for the nonlinearity $\mathbf{f}(\mathbf{x})$, the accuracy of $\hat{\mathbf{A}}$ is directly dependent on α. Smaller the Lipschitz constant yields a better accuracy of $\hat{\mathbf{A}}$.

5.7 Algorithm for online implementation

The algorithm for the three phase identification algorithm is postulated.

1. Employ the control law $\mathbf{u} = -\hat{\mathbf{K}}\mathbf{x}$ for sampling time interval T seconds, with adaptation law $\dot{\hat{\mathbf{K}}}(t) = -\gamma \text{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top$.
2. Go back to Step 1 to stabilize for another T seconds if the stabilization condition in Eq. (5.19) is not satisfied. Continue to Step 3 if satisfied.
3. Initialize the optimization phase for the first time by setting $k = 0$,$i = 1$ and $\mathbf{K}_{1,0} = \hat{\mathbf{K}}(T_s)$. Set the LQR parameters $\mathbf{Q}_i, \mathbf{R}_i$

4. Employ the control law $\mathbf{u} = -\mathbf{K}_{i,k}\mathbf{x} + \mathbf{d}$ for the next sampling time interval T seconds. The state information is used to $\delta_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}$ from Eq. (4.24,4.25,4.26) respectively.
5. Continue to Step 6 if the rank condition $\text{rank}([\mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}]) = \frac{n(n+1)}{2} + mn$ is met. Go back to Step 4 if not satisfied.
6. Iterate with k from equations 4.28,4.29 until $\|\mathbf{P}_{i,k} - \mathbf{P}_{i,k+1}\|_F \leq \epsilon$ where $\epsilon > 0$ is a predefined threshold for convergence.
7. Note down converging solution $\mathbf{P}_{i,N_i}$ corresponding to parameters $\mathbf{Q}_i, \mathbf{R}_i$. Check for the rank condition $rank(\mathbf{H}) = rank([\mathbf{H}\ \mathbf{Y}]) = n^2$ from Eq.(4.38). If the condition is not satisfied continue to Step 8, and if satisfied then continue to Step 9.
8. Set the parameters $\mathbf{Q}_{i+1}, \mathbf{R}_{i+1}$ and reinitialize the optimization phase by setting $k = 0, \mathbf{K}_{i+1,0} = \mathbf{K}_{i,N_i}$, $i = i + 1$ and continue to Step 4.
9. Use Eq.(4.38,4.35) to calculate the identified parameters $\hat{\mathbf{A}}, \hat{\mathbf{B}}$.

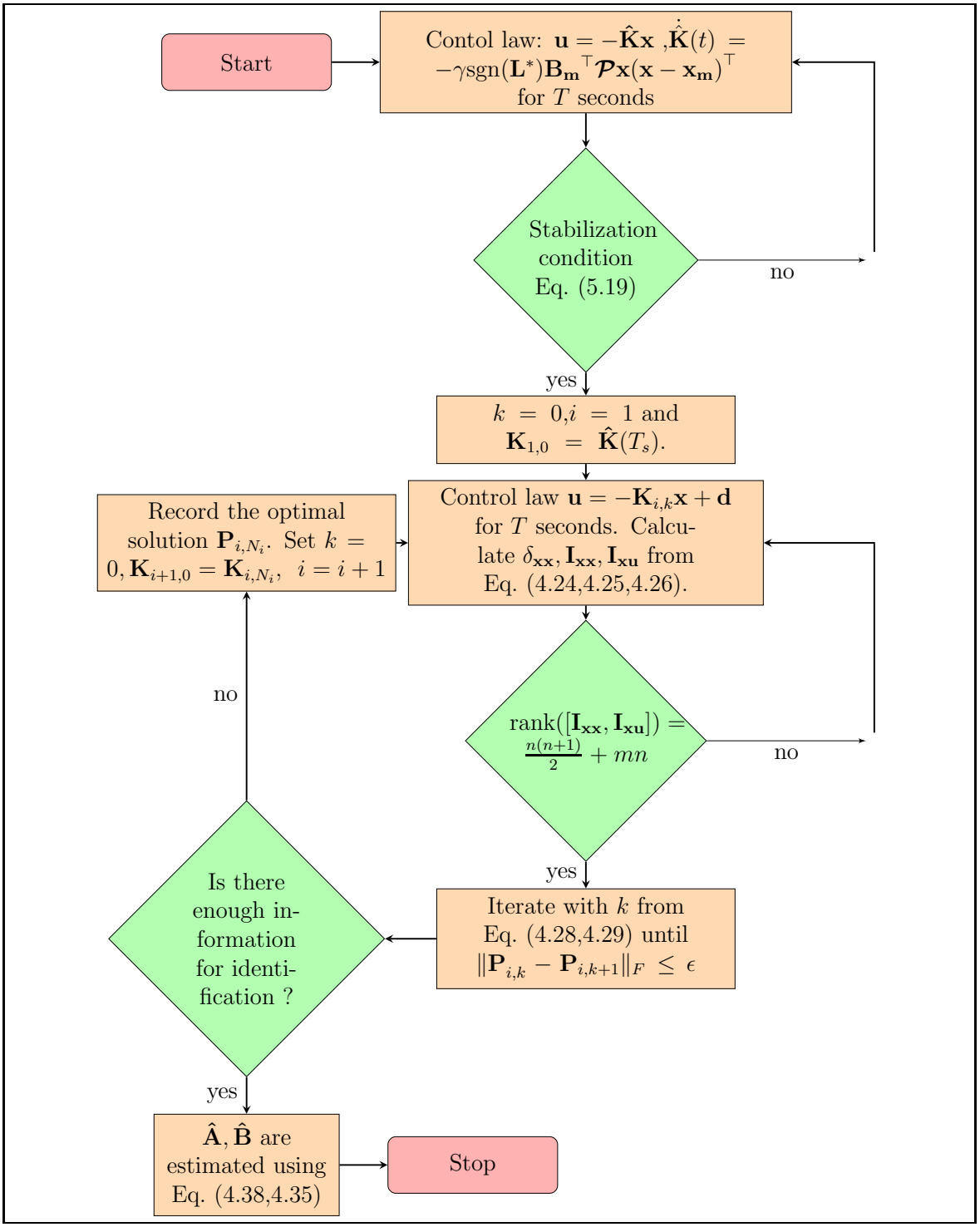

Figure 5.1. Flowchart for Online Implementation - Lipschitz.

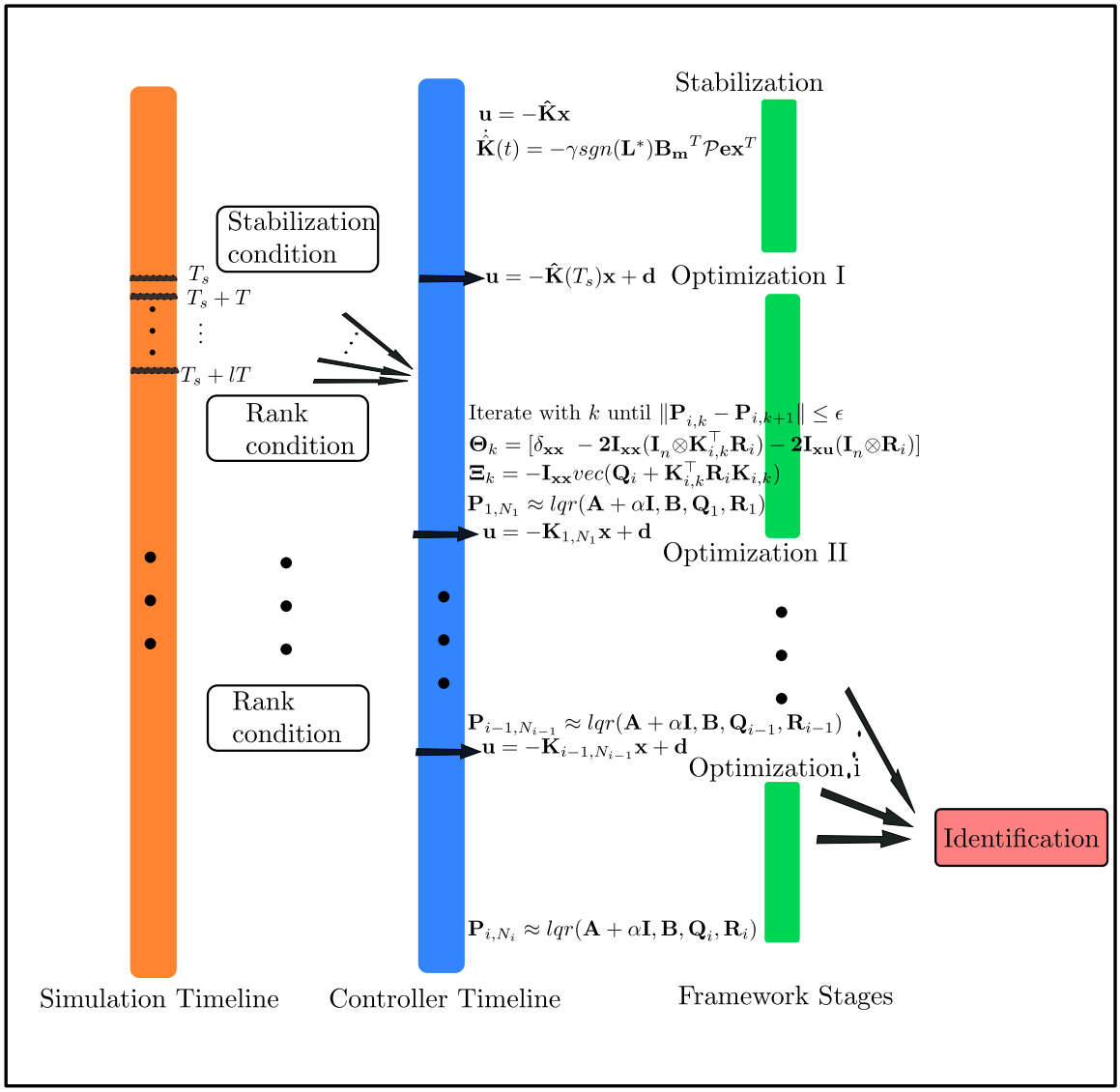

Figure 5.2. Timelines for framework operation - Lipschitz.

5.8 Simulation Results

A continuous time simulation is setup to implement the three stages of learning. The dynamic model is chosen to be a linearized lateral dynamics of Harrier AV-8B from [40] perturbed by a sinusoidal disturbance in the first state. Note that the perturbation is destabilizing and Lipschitz with $\alpha = 0.01$. This linearization is valid at an airspeed of $50ft/s$ at an altitude $50ft$ above sea level. Four simulated states are lateral velocity in body frame v in ft/s, body axis roll angular velocity p in deg/s, body axis yaw angular velocity r in deg/s, roll angle ϕ in deg.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + 0.01 \begin{bmatrix} \sin(x_1) \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.24}$$

The LQR weights are chosen to be identity matrices. Note that the given dynamics are unstable for zero control input. The initial estimates for feedback gains are set to $\mathbf{0}$. The simulation represents a scenario in which the controller regulates the lateral states, and identifies the matrices $\mathbf{A}, \mathbf{B}$ when the Lipschitz constant α is given.

$$\mathbf{A} = \begin{bmatrix} -0.0283 & 0.1823 & -0.8588 & 0.5493 \\ 0.0414 & -0.6662 & 0.2962 & 0 \\ -0.1926 & -0.0447 & -0.0891 & 0 \\ 0 & 1 & 0.2125 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3070 & 0.9224 \\ 1.0060 & -1.4070 \\ 0 & 0 \end{bmatrix}$$

The simulation demonstrates the identification and closed loop stabilization of a linear time invariant system without the knowledge of matrices $\mathbf{A}, \mathbf{B}$. The reference input $\mathbf{r}(t)$ is set to $\mathbf{0}$ to simulate a regulation case. Initial conditions of states are set to $[0 \text{ ft/s } 0\ ^o\text{/s } 0\ ^o\text{/s } 0\ ^o]^\top$.

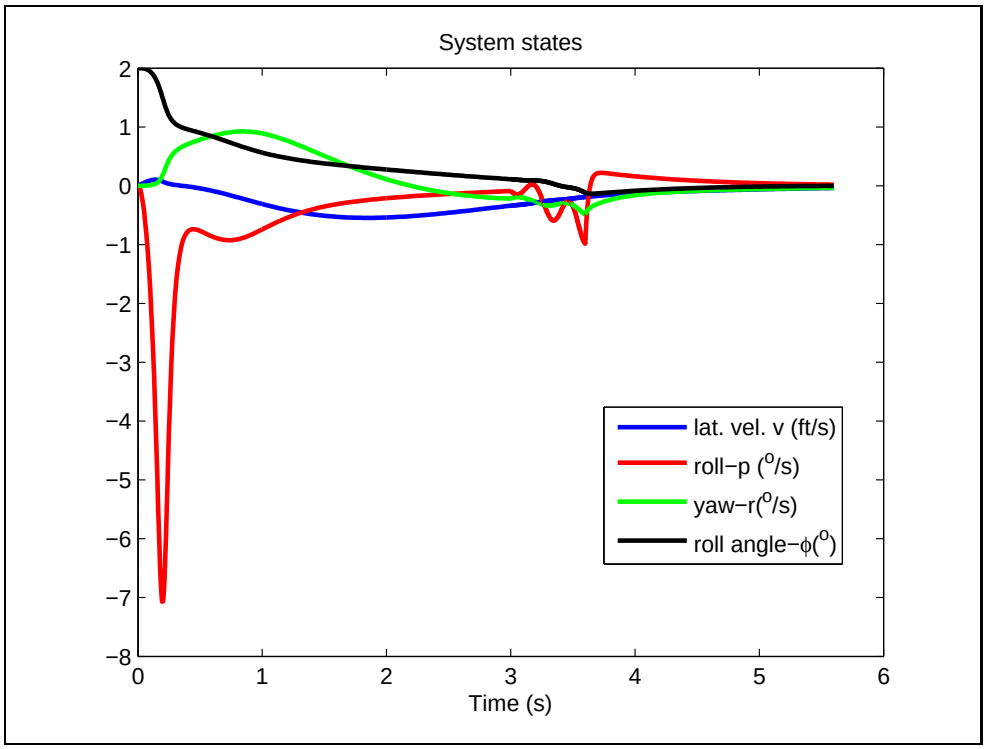

Figure 5.3. Closed Loop system response - Lipschitz.

Fig. [5.3] shows the state history for a case when $\gamma = 10$. The Stabilization phase lasts till about $t = 3$ seconds. The plots clearly show that the states are regulated as desired. A stable reference model characterized by $(\mathbf{A_m}, \mathbf{B_m})$ is chosen to accommodate the structural flexibility requirements. The only information used in the controller formulation is that $\text{sgn}(\mathbf{L}^*) = +1$ and $\alpha = 0.01$.

$$\mathbf{A_m} = \begin{bmatrix} -0.1 & 0.1 & -1 & 0.5 \\ -7 & -15 & 1.5 & -10 \\ 0.5 & -0.8 & -2 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{B_m} = \begin{bmatrix} 0.0536 & 0.1669 \\ 29.5364 & 16.1518 \\ 0.6050 & -1.8080 \\ 0 & 0 \end{bmatrix} \tag{5.25}$$

Fig. [5.4] shows the control history which is continuous for the stabilization phase, but exhibits discrete updates due to the optimization phases. Note that the open loop

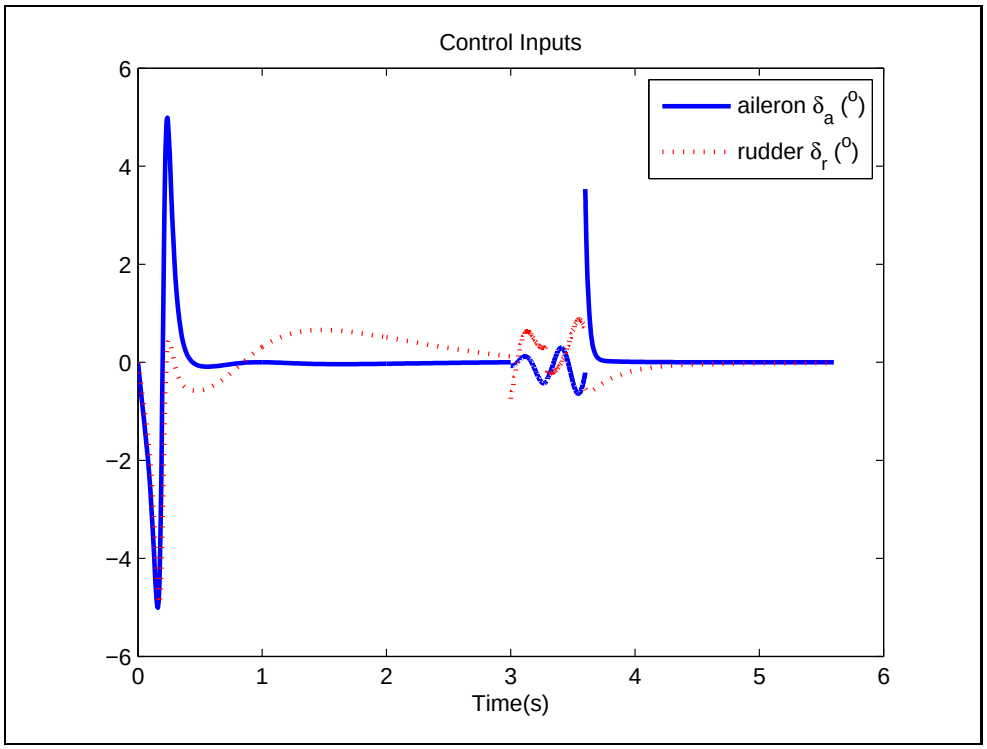

Figure 5.4. Control input history - Lipschitz.

eigen values of $\mathbf{A}$ are at $(-0.6391 \pm 0.2167i, 0.2473 \pm 0.1913i)$. The exit condition is satisfied at around 3 seconds, where the stabilizing gain $\hat{\mathbf{K}}(T_s)$ moves the closed loop poles to $(-14.4159, -1.2622, -1.0654 \pm 0.3757i)$.

$$\hat{\mathbf{K}}(T_s) = \begin{bmatrix} 0.4836 & 1.0454 & -0.0026 & 1.1785 \\ 0.9686 & 0.1641 & -1.3069 & 0.2687 \end{bmatrix} \tag{5.26}$$

The LQR weights are chosen to be identity matrices for first phase of optimization $(\mathbf{Q}_1, \mathbf{R}_1)$. Note that a band limited random sinusoid is added to the control input during the optimization phases. This exploration noise is necessary for the convergence to optimal solution. The noise $\mathbf{d}$ is introduced into the control input as shown in Eq. (5.27) where ω_i are uniformly distributed random frequencies in the range $[-25, 25]Hz$.

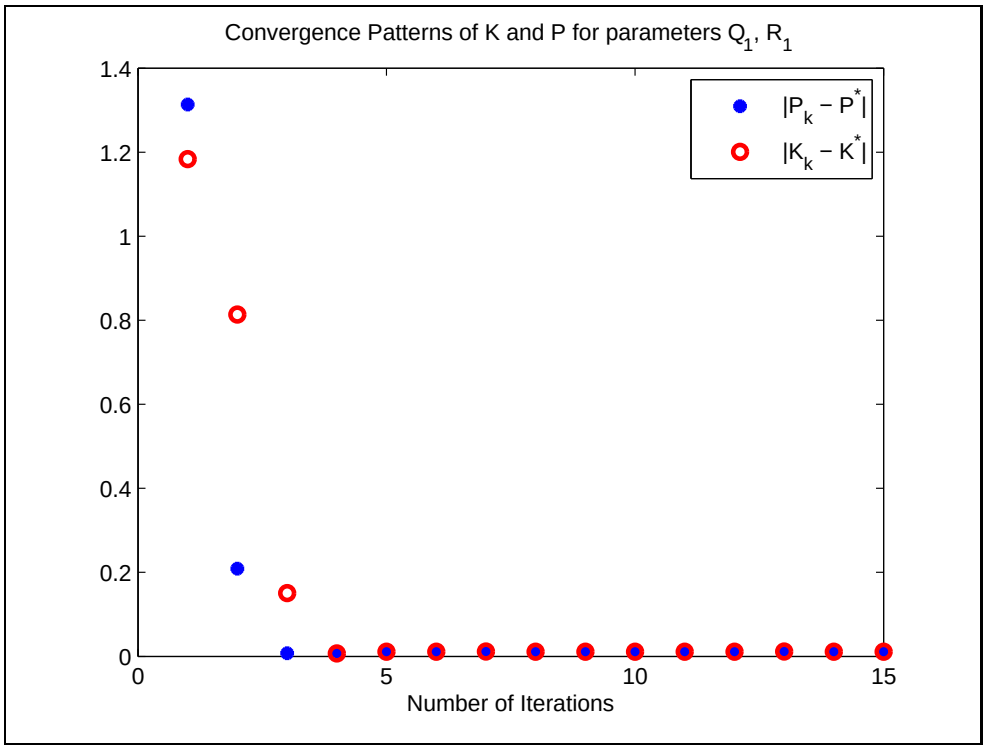

Figure 5.5. Iteration history of $\mathbf{P}_k, \mathbf{K}_k$ for $\mathbf{Q}_1, \mathbf{R}_1$ - Lipschitz.

$$\mathbf{u}(t) = -\mathbf{K}(T_s)\mathbf{x}(t) + 0.1\sum_{i=1}^{200}\sin(\omega_i t) \tag{5.27}$$

Data is collected until the rank condition is satisfied. Recursive relations from Eq. (4.28,4.29) are used to calculated the optimal feedback gain $\mathbf{K}_{1,N_1}$ and corresponding $\mathbf{P}_{1,N_1}$. Fig. [5.5] shows the Frobenius norm of error $\mathbf{P}_{1,k} - \mathbf{P}^*$ during the recursion for first phase of optimization. The plot shows convergence of $\mathbf{P}_{1,k}$ to the optimal solution given by LQR weights $\mathbf{Q}_1, \mathbf{R}_1$ which is recorded as $\mathbf{P}_{1,N_1}$. This phase lasts till about $t = 3.3$ seconds.

A change in $\mathbf{Q}, \mathbf{R}$ parameters is introduced by setting $\mathbf{Q}_2 = 4\mathbf{Q}_1,\ \mathbf{R}_2 = 0.5\mathbf{R}_1$. Fig. [5.6] shows the Frobenius norm of error $\mathbf{P}_{2,k} - \mathbf{P}_2^*$ during the recursion for second phase of optimization. The plot shows convergence of $\mathbf{P}_{2,k}$ to the optimal solution

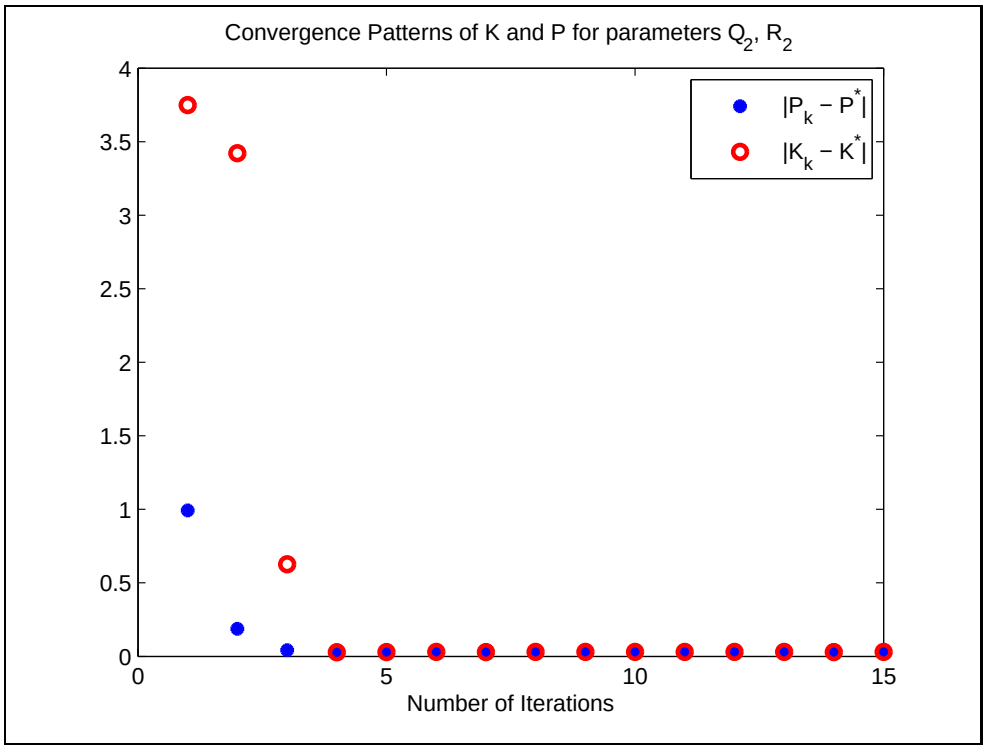

Figure 5.6. Iteration history of $\mathbf{P}_k, \mathbf{K}_k$ for $\mathbf{Q}_2, \mathbf{R}_2$ - Lipschitz.

given by LQR weights $\mathbf{Q}_2, \mathbf{R}_2$ which is recorded as $\mathbf{P}_{2,N_2}$. This phase lasts till about $t = 3.6$ seconds.

$$\mathbf{P}_1^* = \begin{bmatrix} 1.4747 & 0.0739 & -0.5424 & 0.4052 \\ 0.0739 & 0.0775 & -0.0611 & 0.0908 \\ -0.5424 & -0.0611 & 0.8576 & -0.1033 \\ 0.4052 & 0.0908 & -0.1033 & 1.2254 \end{bmatrix}$$

$$\mathbf{P}_{1,N_1} = \begin{bmatrix} 1.4750 & 0.0740 & -0.5420 & 0.4070 \\ 0.0740 & 0.0775 & -0.0609 & 0.0903 \\ -0.5420 & -0.0609 & 0.8530 & -0.1048 \\ 0.4070 & 0.0903 & -0.1048 & 1.2167 \end{bmatrix}$$

It is observed that the difference between the approximation $\mathbf{P}_{1,N_1}$ and the actual solution to the ARE ($\mathbf{P}_1^*$) has a Frobenius norm of 0.0096. The second phase

of optimization converges to $\mathbf{P}_{2,N_2}$, whereas the actual solution to the ARE ($\mathbf{P}_2^*$) differs by a Frobenius norm of about 0.0420.

$$\mathbf{P}_2^* = \begin{bmatrix} 4.6627 & 0.0853 & -0.5654 & 1.2579 \\ 0.0853 & 0.1041 & -0.0616 & 0.1185 \\ -0.5654 & -0.0616 & 1.0131 & -0.0404 \\ 1.2579 & 0.1185 & -0.0404 & 4.5532 \end{bmatrix}$$

$$\mathbf{P}_{2,N_2} = \begin{bmatrix} 4.6609 & 0.0853 & -0.5655 & 1.2575 \\ 0.0853 & 0.1040 & -0.0615 & 0.1175 \\ -0.5655 & -0.0615 & 1.0109 & -0.0414 \\ 1.2575 & 0.1175 & -0.0414 & 4.5112 \end{bmatrix}$$

The closed form solution of $\mathbf{A}$ can be solved from the approximated solutions $\mathbf{P}_{1,N_1}, \mathbf{P}_{2,N_2}$ using Eq. (4.38).

$$\hat{\mathbf{A}} = \begin{bmatrix} -0.0007 & 0.1824 & -0.8567 & 0.6558 \\ 0.1566 & -0.6764 & 0.2818 & 0.0138 \\ -0.1817 & -0.0434 & -0.0990 & 0.0650 \\ -0.1049 & 0.9991 & 0.2115 & -0.0399 \end{bmatrix} \tag{5.28}$$

The approximation $\hat{\mathbf{A}}$ has eigen values at $(-0.6645 \pm 0.3434i, 0.2566 \pm 0.1562i)$ which are close to their counterparts of $\mathbf{A}$ at $(-0.6391 \pm 0.2167i, 0.2473 \pm 0.1913i)$.

$$\mathbf{B} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3070 & 0.9224 \\ 1.0060 & -1.4070 \\ 0 & 0 \end{bmatrix}, \quad \hat{\mathbf{B}} = \begin{bmatrix} -0.0199 & 0.0934 \\ 14.3070 & 0.9225 \\ 1.0061 & -1.4070 \\ -0.0000 & -0.0000 \end{bmatrix}$$

CHAPTER 6

Solution for rigid body attitude dynamics

This section develops ideas mentioned in Ch. 2 and explicitly formulates the controller for a spacecraft with an unknown inertia tensor. Consider a spacecraft equipped with appropriate thrusters for attitude correction and gyroscopes for angular velocity measurement. A sterile and noise free space environment is assumed with rotational dynamics considered about the center of gravity of the spacecraft.

6.1 Introduction

Modern day spacecraft are heavily equipped with sensors. The sensor data is utilized by various onboard control systems to perform complex maneuvers. A majority of such controllers rely on rigid body models for most attitude controllers. One key parameter which influences the rigid body dynamics is the moment of inertia tensor. Precise measurement of the inertia matrix can be used in modeling the attitude dynamics if it is available. Such predetermined measurements will be invalid in the event of partial damage to the structure, or temporary reconfiguration of shape of the spacecraft.

Consider the case of a tumbling spacecraft with a reconfigured inertia matrix. In order to arrest the angular velocity, the knowledge of inertia matrix cannot be utilized. Such scenario would benefit from a control method which regulates the angular velocities despite the unknown inertia matrix. Estimation of the unknown inertia matrix would be beneficial for damage assessment and control of future maneuvers.

Attempts for such control framework with online inertia tensor identification have been made with considerable success. One such identification methods uses an onboard robotic arm to perturb the inertia distribution. [46] There have also been other solutions with least-squares scheme in presence of robotic arm for inertia identification. [47] An experimental implementation of another such least squares based identification can be seen in [48]. But the applicability of such schemes is restricted to the availability of a robotic arm.

The online regulation and inertia identification problem has been solved in [49, 50]. More recently globally convergent attitude tracking problem has been solved in [51]. Adaptive control methods have been used for tracking in majority of these works which ensure asymptotic identification of inertia matrix. Experimental results for such control methods are published in [52] showing promising results for inertia identification and attitude tracking.

The control framework from [38] shows the online regulation as well as identification of unknown parameters for a MIMO linear time invariant system. This paper extends the scope of control methods in [38] to spacecraft attitude dynamics problem.

6.2 Dynamics

Consider a spacecraft equipped with appropriate thrusters for attitude correction and gyroscopes for angular velocity measurement. A sterile and noise free space environment is assumed with rotational dynamics considered about the center of gravity of the spacecraft.

$$\dot{\boldsymbol{\omega}} = -\mathbf{J}^{-1}[\boldsymbol{\omega}_{\times}]\mathbf{J}\boldsymbol{\omega} + \mathbf{J}^{-1}\boldsymbol{\tau} \tag{6.1}$$

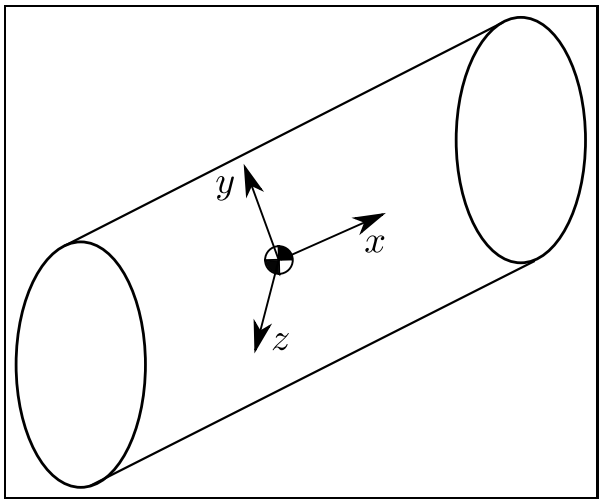

Figure 6.1. A representative rigid spacecraft.

where $\boldsymbol{\omega} \in \mathbb{R}^3$ is the angular velocity about the center of gravity, $\boldsymbol{\tau} \in \mathbb{R}^3$ is the torque applied by the thrusters about the center of gravity, and $\mathbf{J} \in \mathbb{R}^{3\times3}$ represents the unknown moment of inertia of the spacecraft. All the above quantities are measured about a set of mutually perpendicular axes which form a right handed triad.

6.2.1 Hamiltonian dynamics

The nonlinear term can be shown to have zero contribution towards $\dot{V}$ if $V = \frac{1}{2}\boldsymbol{\omega}^\top \mathbf{J}\boldsymbol{\omega}$ is considered to be an energy measure of the system. Such fact is true for dynamics governed by a Hamiltonian system in the absence of control input. In this case V represents the Hamiltonian and $\dot{V} = \boldsymbol{\omega}^\top \boldsymbol{\tau}$. The contribution of nonlinear

term towards $\dot{V}$ can be proved to be zero for any symmetric quadratic function $V = \boldsymbol{\omega}^\top \mathbf{P}\boldsymbol{\omega} \quad \forall\ \mathbf{P} = \mathbf{P}^\top$.

$$\begin{aligned} \dot{V} &= \dot{\boldsymbol{\omega}}^\top \mathbf{P}\boldsymbol{\omega} + \boldsymbol{\omega}^\top \mathbf{P}\dot{\boldsymbol{\omega}} \\ &= \left(\mathbf{J}^{-1}[\boldsymbol{\omega}_\times]\mathbf{J}\boldsymbol{\omega}\right)^\top \mathbf{P}\boldsymbol{\omega} + \boldsymbol{\omega}^\top \mathbf{P}\left(\mathbf{J}^{-1}[\boldsymbol{\omega}_\times]\mathbf{J}\boldsymbol{\omega}\right) \\ &= -\boldsymbol{\omega}^\top \mathbf{J}[\boldsymbol{\omega}_\times]\mathbf{J}^{-1}\mathbf{P}\boldsymbol{\omega} + \boldsymbol{\omega}^\top \mathbf{P}\mathbf{J}^{-1}[\boldsymbol{\omega}_\times]\mathbf{J}\boldsymbol{\omega} \\ &= 0 \end{aligned} \tag{6.2}$$

This property of the nonlinear term allows the treatment of nonlinear attitude dynamics as a linear time invariant system in the discussion of adaptive control as well as optimal control solutions, if the respective Lyapunov function and Cost function are quadratic and symmetric.

6.2.2 Optimal feedback

Consider an optimal control problem with an infinite horizon quadratic cost function $H = \frac{1}{2}\int_0^\infty \boldsymbol{\omega}^\top(s)\mathbf{Q}\boldsymbol{\omega}(s) + \boldsymbol{\tau}^\top(s)\mathbf{R}\boldsymbol{\tau}(s)ds$. The optimal controller for such cost function will be approximated by computational adaptive control in real time.

Hamilton Jacobi Formulation for the optimal control problem requires definition of $L = \frac{1}{2}(\boldsymbol{\omega}^\top\mathbf{Q}\boldsymbol{\omega} + \boldsymbol{\tau}^\top\mathbf{R}\boldsymbol{\tau}) + \boldsymbol{\lambda}^\top(-\mathbf{J}^{-1}[\boldsymbol{\omega}_\times]\mathbf{J}\boldsymbol{\omega} + \mathbf{J}^{-1}\boldsymbol{\tau})$, where $\boldsymbol{\lambda}$ represents the costate vector.

Control law is given by $\frac{\partial L}{\partial \boldsymbol{\tau}} = \mathbf{0}$. An assumption of linear feedback can be made by setting $\boldsymbol{\lambda} = \mathbf{P}\boldsymbol{\omega}$ resulting in $\boldsymbol{\tau} = -\mathbf{R}^{-1}\mathbf{J}^{-1}\mathbf{P}\boldsymbol{\omega}$.

The parameter $\mathbf{P}$ can be determined from costate dynamics given by $\frac{\partial L}{\partial \boldsymbol{\omega}} = -\dot{\boldsymbol{\lambda}}$.

$$\dot{\mathbf{P}} = -\mathbf{Q} + \mathbf{P}\mathbf{J}^{-1}\mathbf{R}^{-1}\mathbf{J}^{-1}\mathbf{P} - \mathbf{P}\mathbf{J}^{-1}[\boldsymbol{\omega}_\times]\mathbf{J} + [\boldsymbol{\omega}_\times]\mathbf{R}^{-1}\mathbf{J}^{-1}\mathbf{P} + (2\mathbf{J} - tr(\mathbf{J})\mathbf{I})[\boldsymbol{\omega}_\times]\mathbf{J}^{-1}\mathbf{R}^{-1}\mathbf{J}^{-1}\mathbf{P} \tag{6.3}$$

Note that the dynamic Riccati equation is coupled with state dynamics through state feedback $\boldsymbol{\omega}$. Hence there is seemingly no *static* linear feedback which solves the LQR problem for given rigid body dynamics. If we consider only the state invariant parts of above equation, it reduces to the LTI case. Although this approximation is stricly not optimal for given dynamics, it yields a stabilizing controller and ensures parameter identification as well. Further the residual for this approximation diminishes as $\boldsymbol{\omega} \to \mathbf{0}$

6.3 Solution Methodology

The proposed control method serves a dual purpose simultaneously. It stabilizes the rotation, and estimates the moment of inertia tensor for the unknown spacecraft without the need for prior experimentation. The algorithm uses readily available readings from the onboard gyroscopes. As a contrast to classical adaptive control, the algorithm identifies the unknown system parameters without sacrificing the closed loop stability or vice-versa. The effectiveness of such algorithm for an unknown linear system realization is shown for a MIMO aircraft model. [38] The algorithm is extended to a very specific class of nonlinear dynamics *i.e* spacecraft attitude dynamics. The control and estimation architecture operates in three stages namely stabilize, optimize and identify. The algorithm derived for a linear time invariant system is valid due to the nature of nonlinearity present in rigid body dynamics. The controller and estimation algorithm will be derived based on the standard notation for a linear time invariant system.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{6.4}$$

where $\mathbf{x} \equiv \boldsymbol{\omega} \in \mathbb{R}^3$ represents the state, $\mathbf{u} \equiv \boldsymbol{\tau} \in \mathbb{R}^3$ is the control. The matrix $\mathbf{A} \equiv \mathbf{0}$, and the unknown matrix $\mathbf{B} \equiv \mathbf{J}^{-1}$ are set to appropriate values for the attitude dynamics problem.

The stabilization phase is marked by model reference adaptive control to actively stabilize the attitude dynamics. The end of stabilization phase is triggered by a stabilization condition based on the measure of considered Lyapunov function candidate in the stabilization phase. The optimization phase applies computational optimal control to estimate the optimal feedback gain for a quadratic cost function. The end of optimization phase is triggered by the convergence of such feedback gain beyond an acceptable tolerance. The unknown matrix $\mathbf{J}$ is derived from the relation for optimal solution which constitutes the last and final phase.

6.3.1 Stabilization

First phase of the controller involves a Model Reference Adaptive Control (MRAC) approach which solves a tracking problem for the unknown dynamic system . $\mathbf{A_m} \in \mathbb{R}^{n\times n}$ is chosen to be a Hurwitz matrix, $\mathbf{B_m} \in \mathbb{R}^{n\times m}$ is chosen such that $\exists \mathbf{K}^* \in \mathbb{R}^{m\times n}, \mathbf{L}^* \in \mathbb{R}^{m\times m}$ satisfying $\mathbf{A_m} = \mathbf{A} - \mathbf{B}\mathbf{K}^*, \mathbf{B}\mathbf{L}^* = \mathbf{B_m}$.

The reference model is characterized by $\mathbf{A_m}, \mathbf{B_m}$, where $\mathbf{r} \in \mathbb{R}^m$ represents the given reference input signal for tracking.

$$\dot{\mathbf{x}}_\mathbf{m}(t) = \mathbf{A_m}\mathbf{x_m}(t) + \mathbf{B_m}\mathbf{r}(t) \tag{6.5}$$

The error ($\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{x_m}(t)$) between states of the unknown dynamic system and the given reference model are minimized. The closed loop error dynamics can be characterized as shown in Eq. (6.6) with state feedback control law $\mathbf{u}(t) = -\hat{\mathbf{K}}(t)\mathbf{x}(t) + \hat{\mathbf{L}}(t)\mathbf{r}(t)$ where $\tilde{\mathbf{K}}(t) = \mathbf{K}^* - \hat{\mathbf{K}}(t),\ \tilde{\mathbf{L}}(t) = \mathbf{L}^* - \hat{\mathbf{L}}(t)$ represent the errors in the estimation of unknown ideal gains $\mathbf{K}^*, \mathbf{L}^*$.

$$
\begin{aligned}
\dot{\mathbf{e}}(t) &= \dot{\mathbf{x}} - \dot{\mathbf{x}}_{\mathbf{m}} \\
&= (\mathbf{A_m} + \mathbf{B}\mathbf{K}^*)\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) - \mathbf{A_m}\mathbf{x_m} - \mathbf{B_m}\mathbf{r} \\
&= \mathbf{A_m}\mathbf{e}(t) + \mathbf{B}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t)) \\
&= \mathbf{A_m}\mathbf{e}(t) + \mathbf{B_m}\mathbf{L}^{*-\mathbf{1}}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t)) \qquad (6.6)
\end{aligned}
$$

The adaptive laws for $\hat{\mathbf{K}}(t), \hat{\mathbf{L}}(t)$ follow from a straightforward Lyapunov analysis by choosing the candidate Lyapunov function as shown below.

$$
V = \mathbf{e}^\top(t)\boldsymbol{\mathcal{P}}\mathbf{e}(t) + \mathrm{Tr}(\tilde{\mathbf{K}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{K}}(t) + \tilde{\mathbf{L}}^\top(t)\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{L}}(t))
$$

Note that $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{n\times n}, \boldsymbol{\Gamma} \in \mathbb{R}^{m\times m}$ are chosen symmetric positive definite matrices. Following the developments in [**?**], asymptotic stability of the closed loop dynamics in Eq. (6.6) can be shown by Lyapunov-Like Lemma(motivated by Barbalat's Lemma) for non-autonomous systems if the adaptive laws for $\hat{\mathbf{K}}, \hat{\mathbf{L}}$ is chosen as

$$
\begin{aligned}
\dot{\hat{\mathbf{K}}}(t) &= -\gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top \qquad (6.7) \\
\dot{\hat{\mathbf{L}}}(t) &= \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top\boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{r}^\top \qquad (6.8)
\end{aligned}
$$

where $\boldsymbol{\mathcal{P}}$ is the solution to the Lyapunov equation $\mathbf{A_m}^\top\boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{P}}\mathbf{A_m} = -\boldsymbol{\mathcal{N}}$, for a chosen $\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{N}}^\top > 0$. It can be shown that the adaptive law along with $\boldsymbol{\Gamma} = \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{L}^{*-\mathbf{1}}$ results in $\dot{V} = -\mathbf{e}^\top(t)\boldsymbol{\mathcal{N}}\mathbf{e}(t)$

Note that the convergence of $\hat{\mathbf{K}}(t), \hat{\mathbf{L}}(t)$ to $\mathbf{K}^*, \mathbf{L}^*$ is not guaranteed. $\hat{\mathbf{K}}(t)$ does however converge to a stabilizing gain $\hat{\mathbf{K}}_\infty$ eventually. Thus the resulting closed loop system upon convergence to $\hat{\mathbf{K}}_\infty$ is asymptotically stable i.e. $\mathbf{A} - \mathbf{B}\hat{\mathbf{K}}_\infty$ is Hurwitz. Barbalat's Lemma can be applied since $\dot{V}$ is clearly uniformly continuous in time. Note that the Lyapunov function V converges to a constant but not necessarily zero, whereas the derivative along the trajectory $\dot{V}$ vanishes as $t \to \infty$.

6.3.2 Stabilization condition

Although Barbalat's Lemma guarantees asymptotic stability as $t \to \infty$, ideally the adaptation should be continued for only a finite time. Assume that the MRAC style adaptation of feedback gain $\hat{\mathbf{K}}$ is ceased for time $t \geq T_s$ for some chosen T_s .

$$\dot{\mathbf{e}} \quad = \mathbf{A_m}\mathbf{e}(t) + \mathbf{B_m}\mathbf{L^{*-1}}(\tilde{\mathbf{K}}(t)\mathbf{x}(t) - \tilde{\mathbf{L}}(t)\mathbf{r}(t))$$

Upon integration for feedback gain error $\tilde{\mathbf{K}}(t)$, and the state error $\mathbf{e}(t)$ before the adaptation is ceased *i.e* $\forall\ 0 \leq t \leq T_s$.

$$\tilde{\mathbf{K}}(t) \quad = \tilde{\mathbf{K}}(0) + \int_0^t \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top ds \tag{6.9}$$

$$\tilde{\mathbf{L}}(t) \quad = \tilde{\mathbf{L}}(0) - \int_0^t \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{r}^\top ds \tag{6.10}$$

$$\begin{aligned}\mathbf{e}(t) \quad &= \mathrm{e}^{\mathbf{A_m}t}\mathbf{e}(0) + \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(\mathbf{x}(\tau),\mathbf{e}(\tau))\mathbf{x}(\tau)d\tau \\ &\quad - \int_0^\top \mathrm{e}^{\mathbf{A_m}(t-\tau)}\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{L}}(\mathbf{r}(\tau),\mathbf{e}(\tau))\mathbf{r}(\tau)d\tau \end{aligned} \tag{6.11}$$

The state relation obtained is implicit but it is not a kind of implicitness which can be dealt with Bellman-Gronwall Lemma. However $\mathbf{x}(T_s), \tilde{\mathbf{K}}(T_s)$ can be evaluated using Eq. (6.11,6.10). After the adaptation is ceased, feedback gain will remain constant $\hat{\mathbf{K}}(T_s)$. Thus the states are governed by linear time invariant dynamics after T_s. Explicit form for the state can be given $\forall\ t \geq T_s$ using Eq. (6.12).

$$\begin{aligned}\mathbf{e}(t) \quad &= \mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-T_s)}\mathbf{e}(T_s) \\ &\quad + \int_0^\top \mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-\tau)}\mathbf{B_m}\mathbf{L^{*-1}}(\tilde{\mathbf{K}}(T_s)\mathbf{x_m}(\tau) - \tilde{\mathbf{L}}(T_s)\mathbf{r}(\tau))d\tau \end{aligned} \tag{6.12}$$

$$\begin{aligned}\|\mathbf{e}(t)\| \quad &\leq \|\mathrm{e}^{(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-T_s)}\|\ (\|\mathbf{e}(T_s)\| + \\ &\quad \|\mathbf{B_m}\mathbf{L^{*-1}}\|(\|\tilde{\mathbf{K}}(T_s)\|\|\mathbf{x_m}\| + \|\tilde{\mathbf{L}}(T_s)\|\|\mathbf{r}\|)) \\ \|\mathbf{e}(t)\| \quad &\leq \mathrm{e}^{\mu(\mathbf{A_m}+\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))(t-T_s)}\ (\|\mathbf{e}(T_s)\| + \\ &\quad \|\mathbf{B_m}\mathbf{L^{*-1}}\|(\|\tilde{\mathbf{K}}(T_s)\|\|\mathbf{x_m}\| + \|\tilde{\mathbf{L}}(T_s)\|\|\mathbf{r}\|)) \end{aligned} \tag{6.13}$$

Note that $\mathbf{e}(t) \in \mathbb{R}^n$ and $\|\mathbf{e}(t)\| \in \mathbb{R}$. The expression $\|\cdot\|$ for a square matrix should be interpreted as the induced norm from vector 2-norm $\|\cdot\|$. Above steps use triangle inequality for vector norms and the definition of induced norm for square matrices. Let $\mu(\mathbf{A})$ represent logarithmic norm of a matrix $\mathbf{A}$, and signifies the maximal growth rate of $log\|\mathbf{x}\|$ if $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. The logarithmic norm properties include $\|\mathrm{e}^{\mathbf{P}t}\| \leq \mathrm{e}^{\mu(\mathbf{P})t}$, and $\mu(\mathbf{P}+\mathbf{Q}) \leq \mu(\mathbf{P}) + \|\mathbf{Q}\|$.

The upper limit on $\|\mathbf{e}(t)\|$ can also be obtained by using Bellman-Gronwall Lemma. If $\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s)) \leq -\delta$ is satisfied for some $\delta > 0$ and the norms $\|\mathbf{e}(T_s)\|, \|\mathbf{x_m}\|, \|\mathbf{r}\|$ are bounded, the state trajectories will be bounded by a decaying exponential.

Note that $\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s)) \leq -\delta \implies |\mu(\mathbf{A_m} + \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s))| \geq \delta$.

$$\begin{aligned}
\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s) &= \mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(0) + \gamma \mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\mathbf{L^{*-1}}\mathbf{B_m}^\top \boldsymbol{\mathcal{P}} \int_0^{T_s} \mathbf{x}(s)\mathbf{x}^\top(s)ds \\
\mathbf{B_m}\mathbf{L^{*-1}}\tilde{\mathbf{K}}(T_s) &= \frac{1}{\gamma}\mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0) + \mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top \boldsymbol{\mathcal{P}} \int_0^{T_s} \mathbf{x}(s)\mathbf{x}^\top(s)ds \qquad (6.14)
\end{aligned}$$

If the condition $\int_0^{T_s} \|\mathbf{e}(s)\mathbf{e}^\top(s)\|ds \geq \frac{|\mu(\mathbf{A_m} + \frac{1}{\gamma}\mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0))| + \delta}{\|\mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\|}$ is satisfied, then the gain $(\hat{\mathbf{K}}(T_s))$ is stabilizing after the adaptation is ceased. Further it can be established that $\|\mathbf{x}\mathbf{x}^\top\| = \|\mathbf{x}\|^2$(Consider an arbitrary vector $\mathbf{p} \in \mathbb{R}^n$, then $\|\mathbf{x}\mathbf{x}^\top\mathbf{p}\| = |\mathbf{x}^\top\mathbf{p}|\|\mathbf{x}\| \leq \|\mathbf{x}\|^2\|\mathbf{p}\|$).

$$\boxed{\int_0^{T_s} \|\mathbf{e}(s)\|^2 ds \geq \frac{|\mu(\mathbf{A_m} + \frac{1}{\gamma}\mathrm{sgn}(\mathbf{L}^*)\mathbf{B_m}\boldsymbol{\Gamma}\tilde{\mathbf{K}}(0))| + \delta}{\|\mathbf{B_m}\boldsymbol{\Gamma}\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\|}} \qquad (6.15)$$

Note that the lower limit from stabilizing condition cannot be explicitly obtained using known parameters. However such a positive limit can be calculated if the parameters were known. Thus a positive limit $\sigma > 0$ is chosen which yields a stabilizing condition which can be verified.

$$\int_0^{T_s} \|\mathbf{e}(s)\|^2 ds \geq \sigma > 0 \tag{6.16}$$

The stabilizing condition can also be interpreted as a lower limit on the decay of Lyapunov function V.

$$\begin{array}{rll}
\lambda_{\min}(\boldsymbol{\mathcal{N}})\|\mathbf{e}\|^2 & \leq \mathbf{e}^\top \boldsymbol{\mathcal{N}} \mathbf{e} & \leq \lambda_{\max}(\boldsymbol{\mathcal{N}})\|\mathbf{e}\|^2 \quad \forall \mathbf{e} \\
\dfrac{\mathbf{e}^\top \boldsymbol{\mathcal{N}} \mathbf{e}}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} & \leq \|\mathbf{e}\|^2 & \leq \dfrac{\mathbf{e}^\top \boldsymbol{\mathcal{N}} \mathbf{e}}{\lambda_{\min}(\boldsymbol{\mathcal{N}})} \quad \forall \mathbf{e} \\
\displaystyle\int_0^{T_s} \frac{-\dot{V}}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} ds & \leq \int_0^{T_s} \|\mathbf{e}\|^2 ds & \leq \displaystyle\int_0^{T_s} \frac{-\dot{V}}{\lambda_{\min}(\boldsymbol{\mathcal{N}})} ds \\
\dfrac{V(0) - V(T_s)}{\lambda_{\max}(\boldsymbol{\mathcal{N}})} & \leq \int_0^{T_s} \|\mathbf{e}\|^2 ds & \leq \dfrac{V(0) - V(T_s)}{\lambda_{\min}(\boldsymbol{\mathcal{N}})}
\end{array} \tag{6.17}$$

6.3.3 Optimization

Second phase of the controller uses the adaptive scheme developed in [18] to recursively approach the optimal feedback gain. This construction leads to a control gain which converges to the optimal feedback gain for the Linear Quadratic regulator with unknown internal dynamics. The iterative algorithm proposed by [18] is used with $\mathbf{K}_0 = \hat{\mathbf{K}}(T_s)$ as the initial stabilizing gain. The result is an adaptive controller which converges to the optimal feedback controller obtained from ARE in Eq. (6.21) without the knowledge of $\mathbf{A}, \mathbf{B}$.

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{6.18}$$

Assuming that the pair $(\mathbf{A}, \mathbf{B})$ is stabilizable, the infinite horizon linear quadratic regulator problem would be to find $\mathbf{u}^*(t)$.

$$\mathbf{u}^*(t) = \underset{\mathbf{u}(t),\ t \in [t_0, \infty]}{\operatorname{argmin}} V(t_0, \mathbf{x}(t_0), \mathbf{u}(t)) \tag{6.19}$$

The infinite horizon cost for the optimal control problem is posed as

$$V(\mathbf{x}(t_0), t_0) = \int_{t_0}^{\infty} (\mathbf{x}^\top(\tau)\mathbf{Q}\mathbf{x}(\tau) + \mathbf{u}^\top(\tau)\mathbf{R}\mathbf{u}(\tau))d\tau \tag{6.20}$$

where $\mathbf{Q} > \mathbf{0},\ \mathbf{R} > \mathbf{0}$, and the pair $((\mathbf{A}, \sqrt{\mathbf{Q}})$ is detectable. The solution to this particular optimal control problem is known to be a state feedback controller $\mathbf{u}(t) = -\mathbf{K}\mathbf{x}(t)$ and the gain $\mathbf{K} = \mathbf{R}^{-\mathbf{1}}\mathbf{B}^\top\mathbf{P}$ where $\mathbf{P}$ is the positive definite solution to the following Algebraic Ricatti Equation.

$$\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-\mathbf{1}}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} = \mathbf{0} \tag{6.21}$$

Of course the control law mentioned above can be synthesized if $\mathbf{A}, \mathbf{B}$ are known.

The cost-to-go with a stabilizing controller gain $\mathbf{K}$ can be written as

$$V(\mathbf{x}(t)) = \int_{t}^{\infty} \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}^\top\mathbf{R}\mathbf{K})\mathbf{x}(\tau)d\tau = \mathbf{x}^\top(t)\mathbf{P}\mathbf{x}(t) \tag{6.22}$$

where $\mathbf{P}$ is the solution of the following Lyapunov equation

$$(\mathbf{A} - \mathbf{B}\mathbf{K})^\top\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) = -(\mathbf{K}^\top\mathbf{R}\mathbf{K} + \mathbf{Q}) \tag{6.23}$$

The cost function can be incrementally written as

$$V(\mathbf{x}(t)) = \int_{t}^{t+T} \mathbf{x}^\top(\tau)(\mathbf{Q} + \mathbf{K}^\top\mathbf{R}\mathbf{K})\mathbf{x}(\tau)d\tau + V(\mathbf{x}(t+T)) \tag{6.24}$$

Some intermediate notation is defined which will help in postulating the algorithm. Consider the following representations for $\mathbf{P} \equiv p_{ij}$, and $\mathbf{x} \equiv x_i$.

$\bar{\mathbf{P}} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$ is a minimal vectorized representation of symmetric $\mathbf{P} \in \mathbb{R}^{n\times n}$.

$$\bar{\mathbf{P}} = [p_{11}\ 2p_{12}\ ..\ 2p_{1n}\ p_{22}\ 2p_{23}\ ..\ p_{nn}]^\top \tag{6.25}$$

$\bar{\mathbf{x}} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$ is a minimal representation of the outer product $\mathbf{x} \otimes \mathbf{x}$.

$$\bar{\mathbf{x}} = [x_1^2\ x_1x_2\ ..x_1x_n\ x_2^2\ x_2x_3\ ..\ x_n^2]^\top \tag{6.26}$$

Above notation is used to propagate the quadratic forms, and note that $\mathbf{x}^\top \mathbf{P} \mathbf{x} = \bar{\mathbf{P}}^\top \bar{\mathbf{x}}$. Matrices $\delta_{\mathbf{xx}} \in \mathbb{R}^{l \times \frac{1}{2}n(n+1)}, \mathbf{I}_{\mathbf{xx}} \in \mathbb{R}^{l \times n^2}, \mathbf{I}_{\mathbf{xu}} \in \mathbb{R}^{l \times mn}$ are defined for l time intervals as below.

$$\delta_{\mathbf{xx}} = [\bar{\mathbf{x}}(t_1) - \bar{\mathbf{x}}(t_0)\ \bar{\mathbf{x}}(t_2) - \bar{\mathbf{x}}(t_1)\ ..\ \bar{\mathbf{x}}(t_{l-1}) - \bar{\mathbf{x}}(t_l)]^\top \tag{6.27}$$

$$\mathbf{I}_{\mathbf{xx}} = \left[\int_{t_0}^{t_1} \mathbf{x} \otimes \mathbf{x} d\tau \ \int_{t_1}^{t_2} \mathbf{x} \otimes \mathbf{x} d\tau \ ..\ \int_{t_{l-1}}^{t_l} \mathbf{x} \otimes \mathbf{x} d\tau \right]^\top \tag{6.28}$$

$$\mathbf{I}_{\mathbf{xu}} = \left[\int_{t_0}^{t_1} \mathbf{x} \otimes \mathbf{u} d\tau \ \int_{t_1}^{t_2} \mathbf{x} \otimes \mathbf{u} d\tau \ ..\ \int_{t_{l-1}}^{t_l} \mathbf{x} \otimes \mathbf{u} d\tau \right]^\top \tag{6.29}$$

For a chosen length of time interval T, sampling times $t_j = t_0 + jT\ \forall\ j \in \{1, 2, .., l\}$. Simulation data in the form of matrices $\delta_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}$ is collected in the presence of exploration noise and initial stabilizing gain. Data collection is continued until $rank([\mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}]) = \frac{n(n+1)}{2} + mn$ for all subsequent samples. This condition is ensured by the persistent excitation from the exploration noise $\mathbf{d}$.

The iterative scheme is based on the following vectorized equation.

$$\mathbf{\Theta} \begin{bmatrix} \bar{\mathbf{P}}_k \\ vec(\mathbf{K}_{k+1}) \end{bmatrix} = \mathbf{\Xi}_k \tag{6.30}$$

where $\mathbf{\Theta}_k$, $\mathbf{\Xi}_k$ are defined as follows.

$$\mathbf{\Theta}_k = [\delta_{\mathbf{xx}}\ -2\mathbf{I}_{\mathbf{xx}}(\mathbf{I}_n \otimes \mathbf{K}_k^\top \mathbf{R}) - 2\mathbf{I}_{\mathbf{xu}}(\mathbf{I}_n \otimes \mathbf{R})] \tag{6.31}$$

$$\mathbf{\Xi}_k = -\mathbf{I}_{\mathbf{xx}} vec(\mathbf{Q} + \mathbf{K}_k^\top \mathbf{R} \mathbf{K}_k) \tag{6.32}$$

The recursive relations from Eq. (6.31,6.32) are iterated starting from an initial stabilizing gain $\mathbf{K}_0$. Eq. (6.30) is solved for $\mathbf{P}_k, \mathbf{K}_{k+1}$ using pseudo inverse. The

following iteration scheme is implemented online due to the presence of rank condition which also ensures convergence to optimal feedback without knowledge of $\mathbf{A}, \mathbf{B}$.

$$\begin{bmatrix} \bar{\mathbf{P}}_k \\ vec(\mathbf{K}_{k+1}) \end{bmatrix} = (\mathbf{\Theta}_k^\top \mathbf{\Theta}_k)^{-1} \mathbf{\Theta}_k^\top \mathbf{\Xi}_k \tag{6.33}$$

6.3.3.1 Identification

The optimal controller in Eq. (6.21) is solved for the unknown matrix $\mathbf{B}$ which is an approximation of $\mathbf{J}^{-1}$. $\mathbf{P}_k$ is the solution of k^{th} iteration for LQR problem posed with parameters $\mathbf{Q}, \mathbf{R}$. Corresponding state feedback gain approximation for next iteration is represented as $\mathbf{K}_{k+1}$. Also note that the convergence is indicated by $\|\mathbf{P}_k - \mathbf{P}_{k-1}\|_F < \epsilon$ which is true for all $k > N$. Thus $\mathbf{P}_N, \mathbf{K}_N$ represent the approximation of LQR solution for parameters $\mathbf{Q}, \mathbf{R}$

$$\mathbf{K}_N = \mathbf{R}^{-1} \hat{\mathbf{B}}^\top \mathbf{P}_N \tag{6.34}$$

Let $\hat{\mathbf{B}}$ represent the estimated value of unknown parameter $\mathbf{B}$. Since the optimization phase ends when $\mathbf{P}_k$ converges to the optimal solution $\mathbf{P}$, the closed form solution of $\mathbf{B}$ is approximated from Eq. (6.35). Note that the solution for $\hat{\mathbf{B}}$ is well defined even for non-square matrices.[38]

$$\hat{\mathbf{B}} = \mathbf{P}_N^{-1} \mathbf{K}_N^\top \mathbf{R} \tag{6.35}$$

6.3.3.2 Algorithm for online implementation

The algorithm for the three phase identification algorithm is postulated along with a flowchart for implementation. The online nature of the controller is evident from Figure [6.3] which shows the various timelines during the controller operation.

1. Employ the control law $\mathbf{u} = -\hat{\mathbf{K}}\mathbf{x}$ for sampling time interval T seconds, with adaptation law $\dot{\hat{\mathbf{K}}}(t) = -\gamma \text{sgn}(\mathbf{L}^*)\mathbf{B_m}^\top \boldsymbol{\mathcal{P}}\mathbf{e}\mathbf{x}^\top$.
2. Go back to Step 1 to stabilize for another T seconds if the stabilization condition in Eq. (6.16) is not satisfied. Continue to Step 3 if satisfied.
3. Initialize the optimization phase by setting $k = 0$ and $\mathbf{K}_0 = \hat{\mathbf{K}}(T_s)$. Set the LQR parameters $\mathbf{Q}, \mathbf{R}$
4. Employ the control law $\mathbf{u} = -\mathbf{K}_k\mathbf{x} + \mathbf{d}$ for the next sampling time interval T seconds. The state information is used to $\delta_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}$ from Eq. (6.27,6.28,6.29) respectively.
5. Continue to Step 6 if the rank condition $\text{rank}([\mathbf{I}_{\mathbf{xx}}, \mathbf{I}_{\mathbf{xu}}]) = \frac{n(n+1)}{2} + mn$ is met. Go back to Step 4 if not satisfied.
6. Iterate with k from Equations (6.31,6.32) until $\|\mathbf{P}_k - \mathbf{P}_{k+1}\|_F \leq \epsilon$ where $\epsilon > 0$ is a predefined threshold for convergence.
7. Note down converging solution $\mathbf{P}_N$ corresponding to parameters $\mathbf{Q}, \mathbf{R}$.
8. Use Eq.(6.35) to calculate the identified parameter $\hat{\mathbf{B}}$.

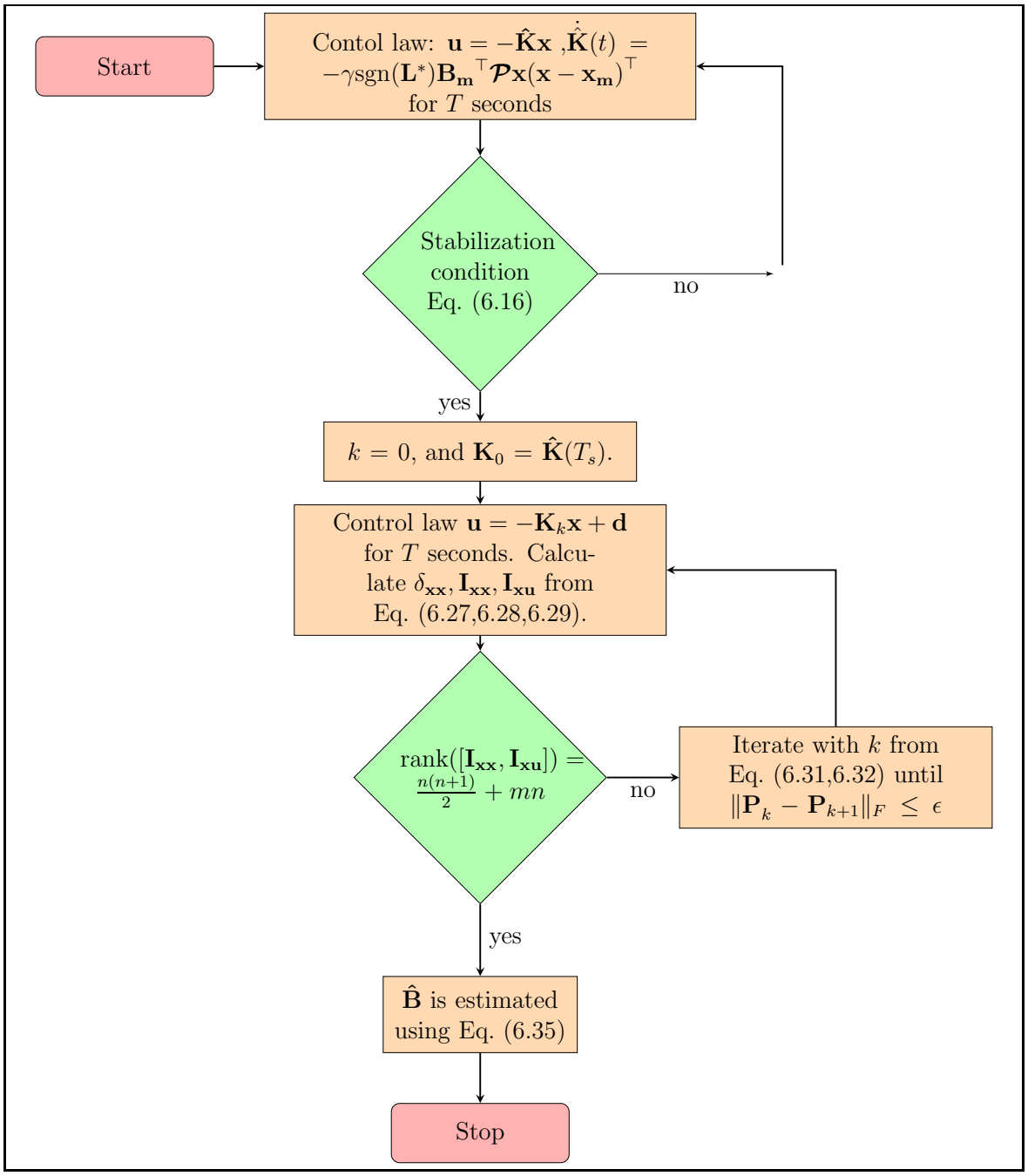

Figure 6.2. Flowchart for Online Implementation - rigid body dynamics.

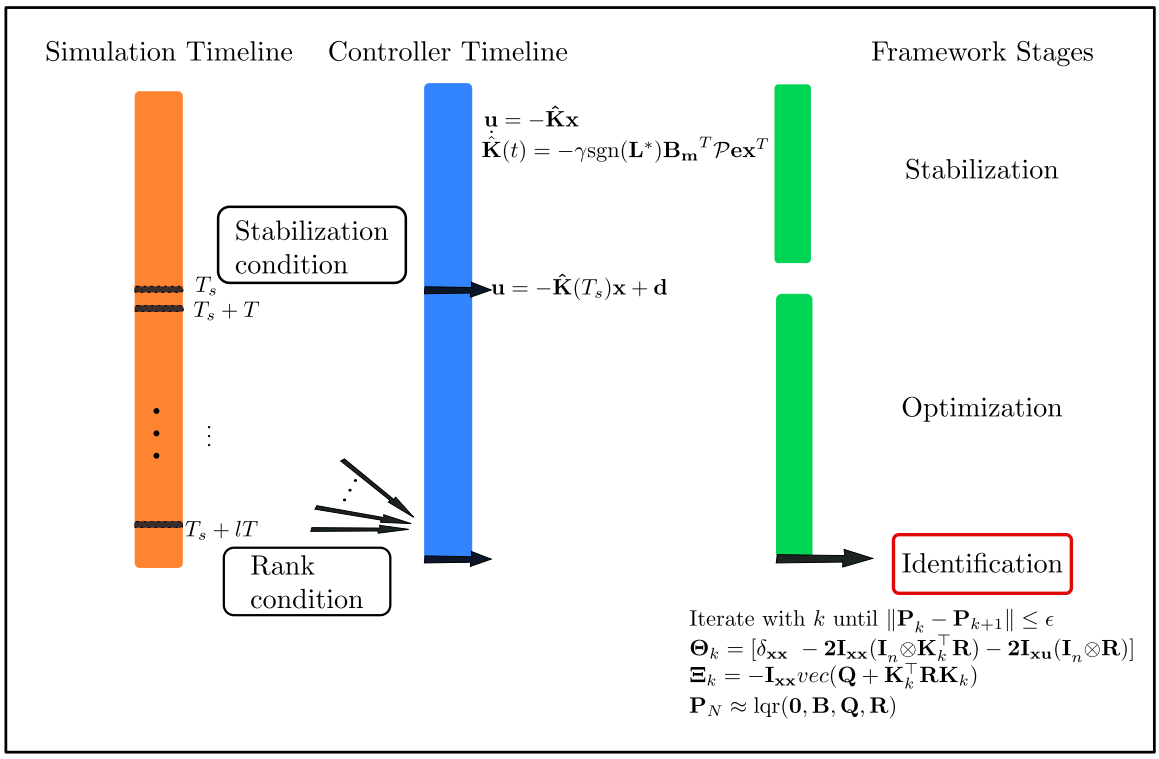

Figure 6.3. Timelines for framework operation - rigid body dynamics.

6.3.4 Simulation Results

A continuous time simulation of the nonlinear attitude dynamics from Eq. is used to implement the proposed algorithm. The dynamic model is characterized by only the moment of inertia matrix about the axes of angular velocity measurement $\mathbf{J}$. Note that the dynamics in absence of a control input are purely oscillatory due to the nature of the nonlinear term.

$$\mathbf{J} = \begin{bmatrix} 5 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{bmatrix}$$

The chosen value of $\mathbf{J}$ is only representative of a valid inertia matrix without any physical significance. The simulation represents a scenario in which the controller

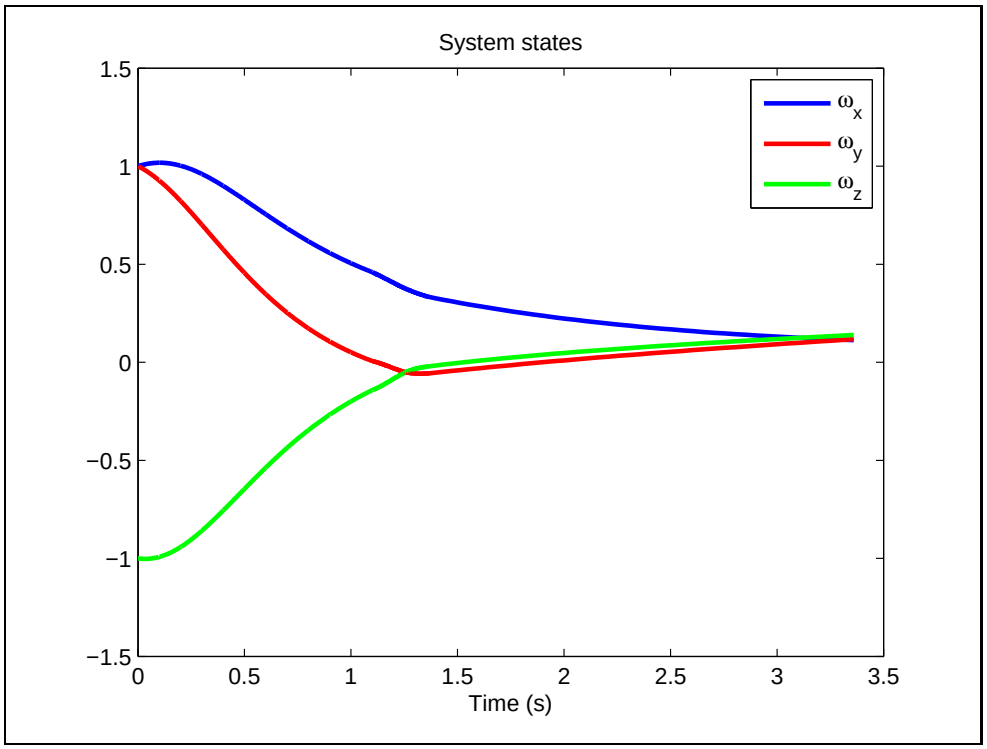

Figure 6.4. Closed Loop system response - rigid body dynamics.

regulates the initial oscillations of a rigid body of unknown moment of inertia tensor and eventually approximates the unknown inertia tensor $\mathbf{J}$.

The reference input $\mathbf{r}(t)$ is set to $\mathbf{0}$ to simulate a regulation case. Initial conditions of states are set to $[1\ 1\ -1]^\top$. The initial estimates for feedback gains are set to $\mathbf{0}$ for the stabilization phase.

Fig. [6.4] shows the state history for a case when $\gamma = 10$. The Stabilization phase lasts till about $t = 1.11$ seconds. The plots clearly show that the states are regulated as desired. A stable reference model characterized by $(\mathbf{A_m}, \mathbf{B_m})$ is chosen to accommodate the structural flexibility requirements. The only information used in the controller formulation is that $\text{sgn}(\mathbf{L}^*) = +1$ which is evident from inherent positive definite nature of $\mathbf{J}$.

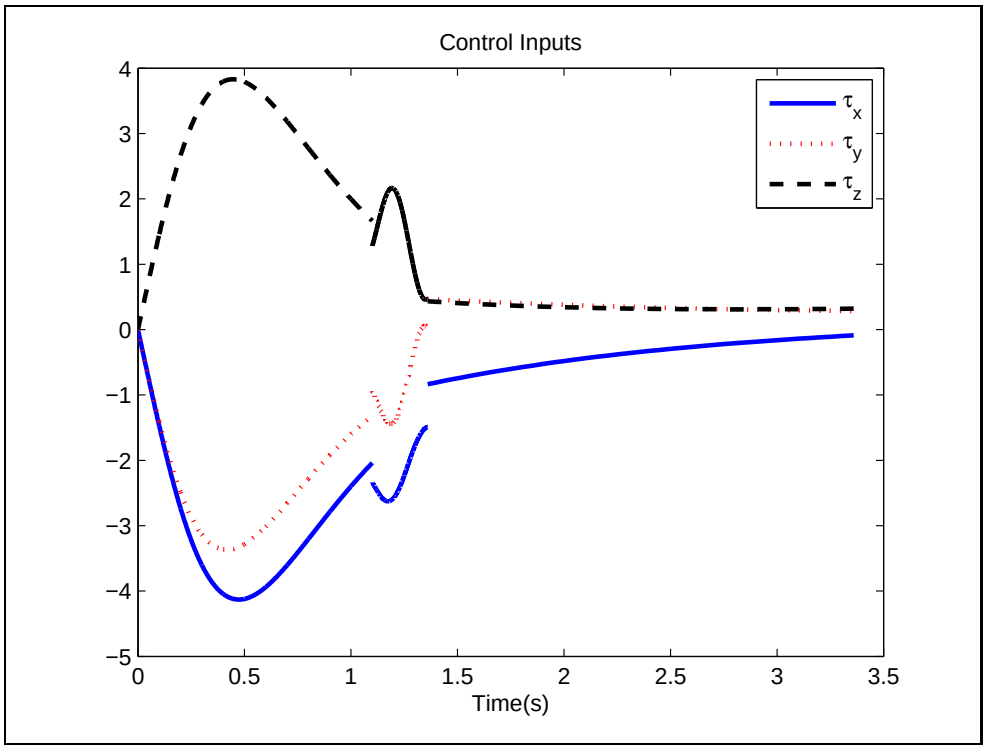

Figure 6.5. Control input history - rigid body dynamics.

$$\mathbf{A_m} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \mathbf{B_m} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.36}$$

Fig. [6.5] shows the control history which is continuous for the stabilization phase, but exhibits discrete updates due to the optimization phases. By setting $T_s = 1.11s$, the stabilizing gain $\hat{\mathbf{K}}(T_s)$.

$$\hat{\mathbf{K}}(T_s) = \begin{bmatrix} 3.4113 & 2.2248 & -2.8093 \\ 2.2248 & 1.6419 & -1.9572 \\ -2.8093 & -1.9572 & 2.3974 \end{bmatrix} \tag{6.37}$$

The LQR weights are chosen to be identity matrices for the optimization phase $(\mathbf{Q}, \mathbf{R})$. Note that a band limited random sinusoid is added to the control input

during the optimization phases. This exploration noise is necessary for the convergence to optimal solution. The noise $\mathbf{d}$ is introduced into the control input as shown in Eq. (6.38) where η_i are uniformly distributed random frequencies in the range $[-25, 25]Hz$.

$$\mathbf{u}(t) = -\mathbf{K}(T_s)\mathbf{x}(t) + 0.1 \left[\sum_{i=1}^{100} \sin(\eta_i t) \quad \sum_{i=101}^{200} \sin(\eta_i t) \quad \sum_{i=201}^{300} \sin(\eta_i t) \right]^\top \tag{6.38}$$

Data is collected until the rank condition is satisfied. Recursive relations from Eq. (6.31,6.32) are used to calculated the optimal feedback gain $\mathbf{K}_N$ and corresponding $\mathbf{P}_N$. Fig. [6.6] shows the Frobenius norm of error $\mathbf{P}_k - \mathbf{P}^*$ during the recursion in optimization phase. The plot shows convergence of $\mathbf{P}_k$ to the optimal solution given by LQR weights $\mathbf{Q}, \mathbf{R}$ which is recorded as $\mathbf{P}_N$. This phase lasts till about $t = 1.33$ seconds.

It is observed that the difference between the successive approximations of $\mathbf{P}_N$ have a Frobenius norm of 1×10^{-4} as a tolerance.

$$\mathbf{P}_N = \begin{bmatrix} 15.8114 & 3.1623 & 3.1623 \\ 3.1623 & 12.6491 & 3.1623 \\ 3.1623 & 3.1623 & 15.8114 \end{bmatrix}$$

$$\mathbf{K}_N = \begin{bmatrix} 3.1623 & 0.0000 & 0.0000 \\ 0.0000 & 3.1623 & -0.0000 \\ 0.0000 & -0.0000 & 3.1623 \end{bmatrix}$$

The closed form solution of $\mathbf{B}$ can be solved from the approximated solutions $\mathbf{P}_N, \mathbf{K}_N$ using Eq. (6.35).

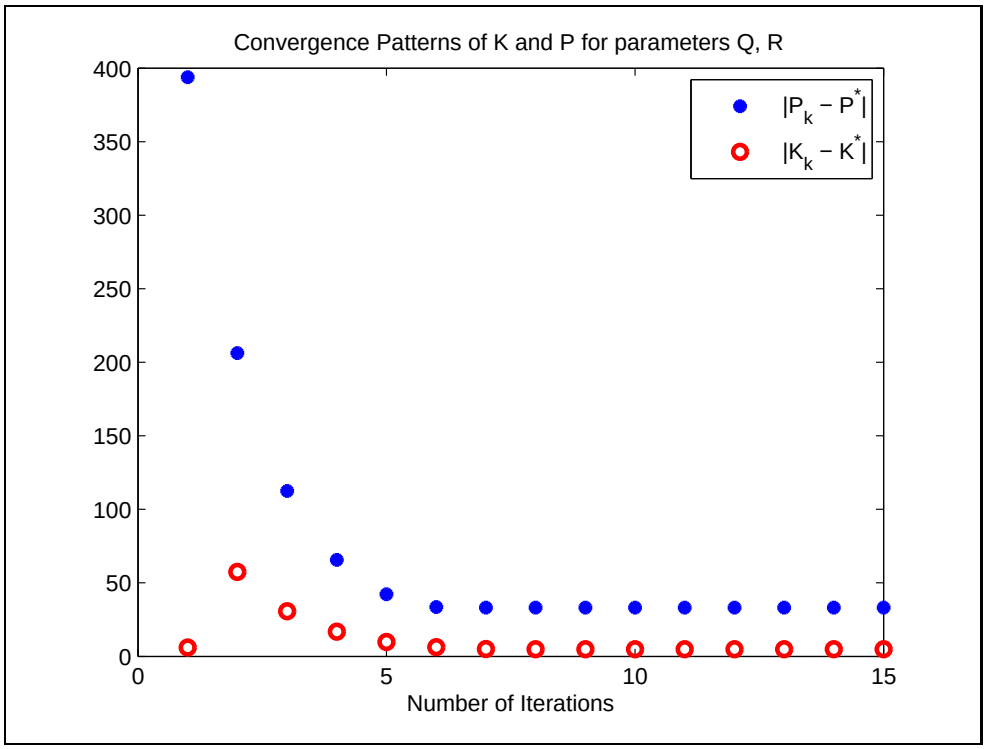

Figure 6.6. Iteration history of $\mathbf{P}_k, \mathbf{K}_k$ for $\mathbf{Q}, \mathbf{R}$ - rigid body dynamics.

$$\mathbf{J} = \begin{bmatrix} 5 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{bmatrix}, \quad \hat{\mathbf{J}} = \begin{bmatrix} 5.0000 & 1.0000 & 1.0000 \\ 1.0000 & 4.0000 & 1.0001 \\ 1.0000 & 1.0001 & 5.0000 \end{bmatrix}$$

CHAPTER 7

Estimation of Reachable set for an unknown linear system

This chapter presents the computed reachable sets for unknown linear dynamic system from Ch. 4. A brief introduction to level set methods for computation of reachable sets is presented along with numerical results for linearized lateral dynamics of Harrier AV-8B.

7.1 Reachable and Safe set formulation

This section sets up the mathematical problem which drives the reachable or safe set computation for a control affine system. Consider a nonlinear system dynamics given by Eq. (7.1)

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{7.1}$$

where $\mathbf{x} \equiv [x_1 \ ... \ x_i \ ... \ x_n]^T \in \mathbb{R}^n$ represents the n-dimensional state vector, and $\mathbf{u} \equiv [u_1 \ ... \ u_i \ ... \ u_m]^T$ represents the vector of m control inputs, and $\mathbf{A} \equiv A_{ij}$, $\mathbf{B} \equiv B_{ij}$ represent the identified linear parameters from intuitive control framework. For the sake of simplifying the optimum solution, it is also assumed that the control signal is restricted to a hyper-rectangular set $\mathcal{Q}$ defined by $\mathcal{Q} \equiv \{[u_i] : \mathbb{U}_i^{\min} \leq u_i \leq \mathbb{U}_i^{\max}, \forall \ i\}$

Using Level set methods to calculate the forward reachable set for the system involves posing a Hamiltonian H in terms of the implicit scalar field V whose zero level set represents the boundary of reachable set. It is standard practice to start with a set where the answer to the reachability question is in affirmative. It is assumed that the initial set from which the system starts is represented by a given $\mathcal{P}$ such that $\mathbf{x}(0) \in \mathcal{P} \equiv \{\mathbf{x} : l(\mathbf{x}) < 0\}$.

$$H(\mathbf{x},\mathbf{p}) = \min(\max_{u\in\mathcal{Q}} \mathbf{p}^T\dot{\mathbf{x}}, 0), \text{ where } \mathbf{p} = \frac{\partial V}{\partial \mathbf{x}} \tag{7.2}$$

It is documented in Ref. [31]. that the evolution of forwards reachable set is identical to the evolution of zero level set of scalar field V . The governing equation is of Hamilton-Jacobi form, and is given by Eq. (7.3)

$$\frac{\partial V(\mathbf{x},t)}{\partial t} + H(\mathbf{x}, \frac{\partial V(\mathbf{x},t)}{\partial \mathbf{x}})] = 0, \text{ where } V(\mathbf{x},0) = l(\mathbf{x}) \tag{7.3}$$

The original formulation uses sup() instead of max(), but both the operators are identical due to the compactness of the set $\mathcal{Q}$, such that $\mathbf{u} \in \mathcal{Q}$. Above equation is solved with an initial reachable set $\mathcal{P}$ using the subroutines from Ian Mitchell's Level Set Toolbox [32, 53]. H is the Hamiltonian for a given pair of $(\mathbf{X}, \mathbf{p})$, which has to be evaluated at any point in the state space $\mathbf{x}$, and for any value of co-state $\mathbf{p}$. For a control affine system under consideration, it is evident from the following formulation that the optimal input $\mathbf{u}^*$ is one of the corners of the hyper-rectangular set $\mathcal{Q}$.

$$\mathbf{u}^*(\mathbf{x},\mathbf{p}) = \operatorname*{argmax}_{\mathbf{u}\in\mathcal{Q}} \mathbf{p}^T(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}) \tag{7.4}$$

The quantity to be optimized is

$$\mathbf{p}^T\dot{\mathbf{x}} \;=\; \sum_{i=1}^{n} p_i \left(\sum_{j=1}^{n} A_{ij}x_j + \sum_{j=1}^{m} B_{ij}u_j \right) \tag{7.5}$$

Maximum of the above expression which is in the form of a partial sum occurs when each term attains the maximum individually. For a given $\mathbf{x}$ it is clear that $\mathbf{u}$ has to be one of the corners of the hyper-rectangle $\mathcal{Q}$ for the extremum to occur.

$$\mathbf{u}^* = \{[u_i^*]\} \tag{7.6}$$

where for any given i

$$u_i^* = \begin{array}{l} \mathbb{U}_i^{\max} \text{ if } \sum_{i=1}^n p_i B_{ij} > 0 \\ \mathbb{U}_i^{\min} \text{ if } \sum_{i=1}^n p_i B_{ij} < 0 \end{array} \tag{7.7}$$

An explicit solution for $\mathbf{u}^*$ is required since the solution of HJ PDE involves discretization of problem over a computational grid spanning the state space. If the computational grid is assumed to be rectangular with q number of points in each dimension, then the total number of points at which the optimization needs to be performed is q^n. Thus explicit solutions are the only viable option, as opposed to online optimization techniques. Explicit solutions like Eq. (7.7) give an additional computational advantage if it can be coded using matrix operations .

Above solution can be concisely coded as a matrix operation below in Eq. (7.8)

$$u_i^* = \mathbb{U}_i^{\max} \left(\sum_{i=1}^n p_i B_{ij} > 0 \right) + \mathbb{U}_i^{\min} \left(\sum_{i=1}^n p_i B_{ij} < 0 \right) \tag{7.8}$$

The use of logical operators may seem trivial and unnecessary for individual calculations, but it yields a sizable computational advantage when applied for the grid in its entirety. It is to be noted that asymmetric form of $\mathcal{Q}$ ($|\mathbb{U}_i^{\min}| \neq |\mathbb{U}_i^{\max}|$) about origin can also be handled using this operator.

7.2 Results

The algorithm described in the previous section computes the reachable sets for a linearized lateral dynamics of Harrier AV-8B with identified matrices $\hat{\mathbf{A}}, \hat{\mathbf{B}}$. The states are lateral velocity v, roll velocity p, yaw velocity r, and yaw angle ϕ respectively, and the control inputs are aileron deflection δ_a, and rudder deflection δ_r. The initial conditions are assumed to be in a ball of unit radius. The reachable set is propagated for about 0.3 seconds forward in time, for control input bounded by $-0.5^o \leq \delta_a \leq 0.5^o,\ 5^o \leq \delta_r \leq 5^o$. The set at every time instant represents the

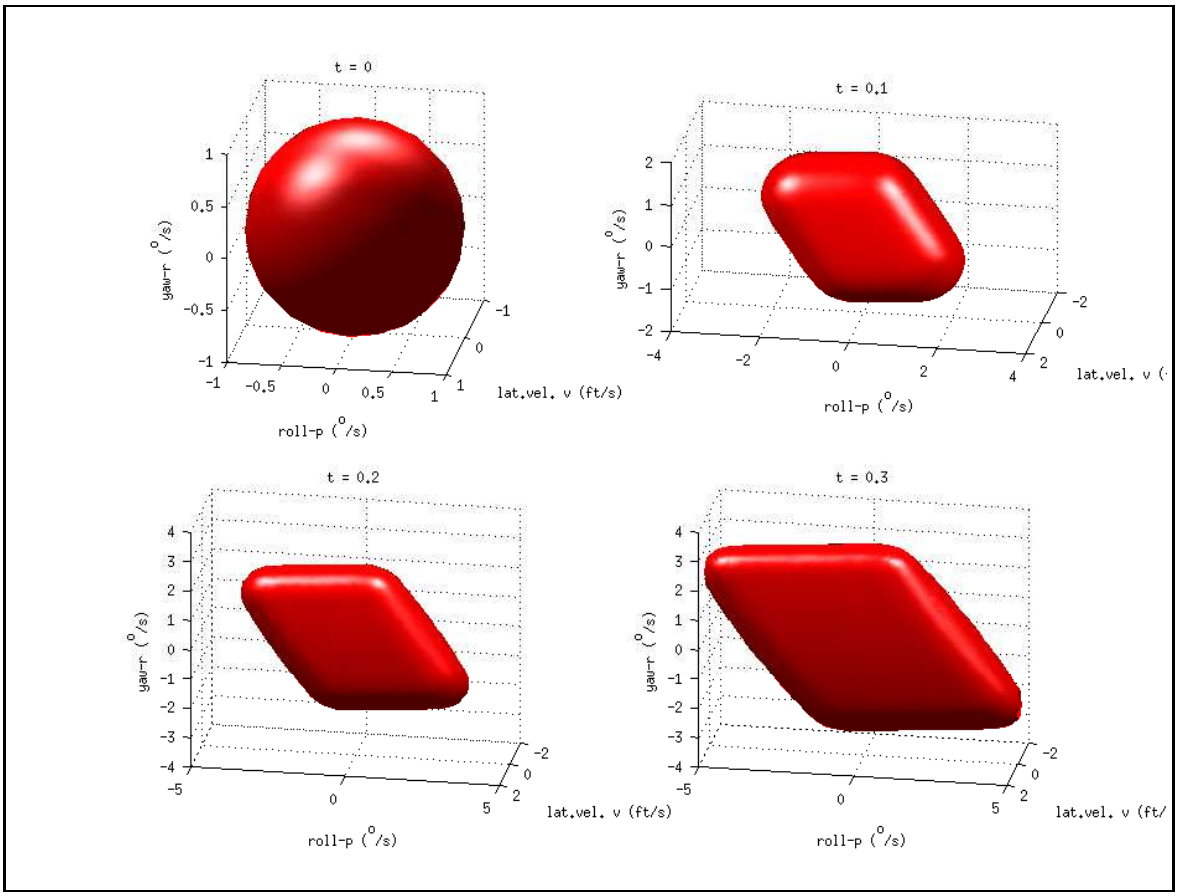

Figure 7.1. Reachable set computed for known dynamics .

control invariant set. Note that the bounds on δ_a are stricter than δ_r. This is due to the higher sensitivity of dynamics to δ_a. The scaling in control bounds is only done for ease of representation of reachable sets. Equal bounds on both controls would result in an ill proportioned reachable set which would be harder to observe. Since the model represents a linear time invariant system, the optimal control problem has the solution given by Eq. (7.7). It is to be noted that the optimal solution is not continuous in $\mathbf{X}$ and $\mathbf{p}$. Level sets are propagated using Ian Mitchell's Toolbox.

Fig. [7.1] shows the 3D projection along states (v, p, r) of computed reachable sets for known parameters $\mathbf{A}, \mathbf{B}$. Whereas Fig. [7.2] shows the same for identified parameters from $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ from Ch. 4. Both the results are computed with identical

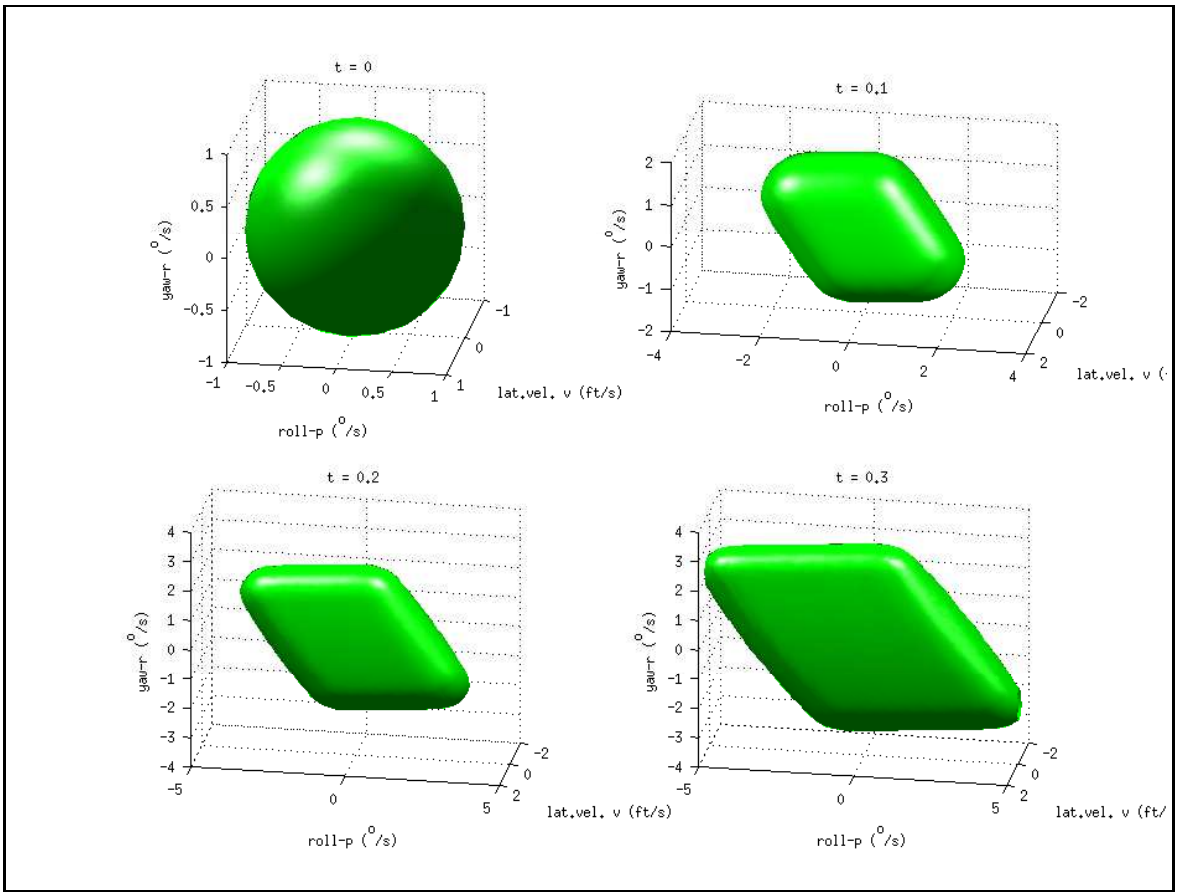

Figure 7.2. Reachable set computed for identified unknown dynamics.

solver settings. Though the differences in these plots is hard to notice, the change in percentage generalized volume of these two sets is calculated to be 0.2%. This difference in volumes suggests that the reachable sets can be calculated without the need for parameters $\mathbf{A}, \mathbf{B}$.

CHAPTER 8

Experimental results with 2-DOF Helicopter

8.1 Introduction

This chapter presents the testing of the proposed intuitive control framework and the corresponding real time experimental results for the Quanser 2-DOF helicopter. The Quanser 2-DOF Helicopter is an educational control experiment made by Quanser Consulting Inc. [54] for testing control strategies on a nonlinear MIMO system with coupled dynamics. Similar experimental testbeds have been successfully used to demonstrate and validate control techniques [55–57] on real world systems. Augmented techniques are of special interest for experimental control testbeds which are explored in [57] in the form of robust compensator.

The experimental setup shown in Fig. [8.1] consists of a helicopter model mounted on a fixed base and free to rotate in both pitch and yaw directions. Pitch angle is controlled using a propeller driven by a DC motor, and is restricted to a range of -40° to 40°. Yaw angle is controlled using a propeller driven by a smaller DC motor, and is free to rotate without any restrictions. DC voltages of pitch and yaw motors serve as control inputs. Due to disparity in sizes of motors, the helicopter rests at a pitch angle of -40° in its unactuated initial state. Both pitch and yaw angles are recorded using encoders which enable state feedback to be used in control design. The pitch and yaw rates are synthesized using a second order filter. For e.g.,

$$\frac{\theta(s)}{\theta_m(s)} = \frac{\omega_f^2}{s^2 + 2\zeta_f\omega_f s + \omega_f^2} \tag{8.1}$$

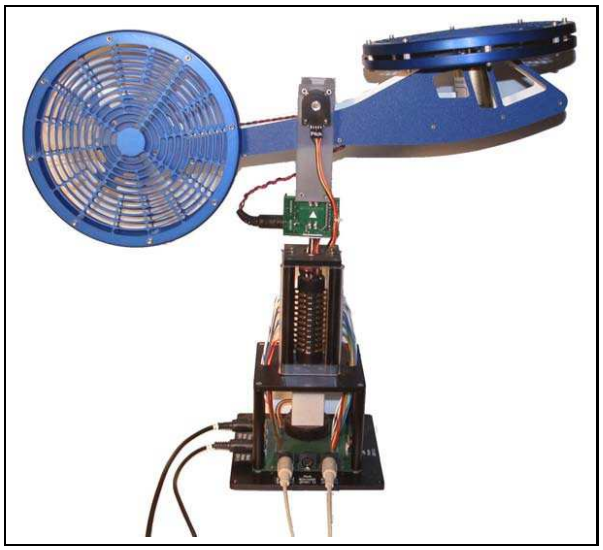

Figure 8.1. Quanser 2-DOF Helicopter.

Where $\theta_m(s)$ is the measured angle, ω_f is the natural frequency of the filter dynamics and ζ_f is the damping ratio. Since the transfer function is of second order, the two states are $\theta(t)$ and $\dot{\theta}(t)$ in the time domain. Thus the angular rate is obtained.

A baseline controller with LQR+I framework has been proved to be effective in stabilization as well as achieving non-zero set point tracking with the experimental setup at its nominal conditions. However, the above mentioned controller is not designed to handle some kinds of uncertainties in its formulation, namely mass/inertia uncertainties and unmodeled dynamics.

This necessitates the design of a controller which is robust to parametric uncertainties as well as unmodeled dynamics. Adaptive control has been used successfully in the past for compensation in the presence such uncertainties [**?**,6,58]. Throughout its development adaptive control has been formulated in several different architectures, which leaves a multitude of choices for controller design [59]. Many of the adaptive frameworks suffer from various shortcomings such as steady state parameter oscillations, robustness at the cost of performance and vice versa as illustrated in [60].

However, recent developments such as the L_1 adaptive control framework [59, 61–63] address several of the issues faced by its predecessors. It decouples the estimation problem from control, thereby allowing arbitrarily fast adaptation for certain robustness. Given its merits,implementation of an L_1 adaptive framework as stand-alone as well as an augmentation to the existing baseline controller was performed using the experimental setup. The design, implementation and real time results for such an adaptive augmentation to the baseline controller is presented in [36,39] with comparison for output tracking performance and robustness. The experimental results with proposed intuitive control framework are discussed along with tuning considerations.

The controller implementation is carried out in the MATLAB®/Simulink® environment. An interface which takes in control inputs (voltages) from the Simulink® environment to actuate the DC motors, and writes the system states (pitch and yaw angles and rates) back to Simulink® is provided by Quanser. The experimental setup requires Simulink® to transfer the code to an on-board processor to run the helicopter in real time. An Euler integration scheme is used to integrate the state space models at 1000 Hz. Although the actual system is propagated as a discrete time model, control techniques assume a continuous time formulation due to its high update rate.

8.2 Modeling

The governing equations of motion for the 2-DOF helicopter are summarized in this section . Given nonlinear model is utilized for the controller development and numerical experiments to tune the parameters for the real-time implementation. It is to be noted that the numerical model given below differs from actual system due to presence of parameter uncertainties and unmodeled dynamics. But it is established a priori that the system parameters used for the formulation of baseline controller are accurate to acceptable precision.

$$\begin{bmatrix} J_p + ml^2 & 0 \\ 0 & J_y + ml^2\cos^2\Theta \end{bmatrix} \begin{bmatrix} \ddot{\Theta} \\ \ddot{\Psi} \end{bmatrix} + \begin{bmatrix} B_p & 0 \\ 0 & B_y \end{bmatrix} \begin{bmatrix} \dot{\Theta} \\ \dot{\Psi} \end{bmatrix} \tag{8.2}$$
$$+ml^2\dot{\Psi}\sin\Theta\cos\Theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \dot{\Theta} \\ \dot{\Psi} \end{bmatrix} + \begin{bmatrix} mgl\cos\Theta \\ 0 \end{bmatrix} = \begin{bmatrix} K_{pp} & K_{py} \\ K_{yp} & K_{yy} \end{bmatrix} \begin{bmatrix} V_p \\ V_y \end{bmatrix}$$

Eq. (8.2) describes the dynamics of 2-DOF helicopter model as a rigid body where Θ and Ψ are pitch and yaw angles, V_p and V_y are voltage inputs to pitch and yaw motors and g represents acceleration due to gravity. System parameters m, l, J_p, J_y, B_p, B_y, K_{pp}, K_{py}, K_{yp}, K_{yy} are known quantities provided by Quanser for the experimental setup. The parametrization assumes the absence of actuator dynamics for the DC motors. A nonlinear simulation model with given parameters is used as a preliminary test platform for control strategies. The state space representation of above system can be written as:

$$\begin{aligned}
\dot{x}_1 &= x_3 \\
\dot{x}_2 &= x_4 \\
\dot{x}_3 &= \frac{K_{pp}u_1 + K_{py}u_2}{J_p + ml^2} - \frac{B_p x_3 + ml^2 x_4^2 \sin x_1 \cos x_1 + mgl\cos x_1}{J_p + ml^2} \\
\dot{x}_4 &= \frac{K_{yp}u_1 + K_{yy}u_2}{J_y + ml^2\cos^2 x_1} - \frac{B_y x_4 - ml^2 x_3 x_4 \sin x_1 \cos x_1}{J_y + ml^2\cos^2 x_1}
\end{aligned}$$

where $\mathbf{x} \equiv [x_1\ x_2\ x_3\ x_4]^T \equiv [\Theta\ \Psi\ \dot{\Theta}\ \dot{\Psi}]^T$ and $\mathbf{u} \equiv [u_1\ u_2]^T \equiv [V_p\ V_y]^T$. It can be observed that all the uncertainties are matched and parametric in nature.

The proposed controller is augmented with a feedforward term which linearizes the closed loop dynamics. The control law is $\mathbf{u} = \mathbf{u}_{ff} + \mathbf{v}$, where the feedforward term is given by Eq. (8.3)

$$\mathbf{u}_{ff} = \begin{bmatrix} K_{pp} & K_{py} \\ K_{yp} & K_{yy} \end{bmatrix}^{-1} \begin{bmatrix} ml^2 x_4^2 \sin x_1 \cos x_1 + mgl\cos x_1 \\ -ml^2 x_3 x_4 \sin x_1 \cos x_1 \end{bmatrix} \tag{8.3}$$

8.3 Reduced order model

The proposed control law makes the closed loop dynamics as shown below.

$$\begin{aligned}
\dot{x}_1 &= x_3 \\
\dot{x}_2 &= x_4 \\
\dot{x}_3 &= \frac{K_{pp}v_1 + K_{py}v_2}{J_p + ml^2} - \frac{B_p x_3}{J_p + ml^2} \\
\dot{x}_4 &= \frac{K_{yp}v_1 + K_{yy}v_2}{J_y + ml^2\cos^2 x_1} - \frac{B_y x_4}{J_y + ml^2\cos^2 x_1}
\end{aligned}$$

Although the helicopter has 2 degrees of freedom, the motor powering yaw actuation is slower when responding to changes in yaw direction. This observation has been reinforced by all previous experiences with the testbed. On the other hand the pitch motor is much faster to respond to any given changes in voltage. If the motors were to be modeled with first order actuator dynamics, the pitch motor would have a much lower time constant compared to the yaw motor. Actuator dynamics could potentially be used to augment the state space and is a subject of future work.

Given the vast differences in the pitch and yaw motor dynamics, it was decided that the verification be conducted on a single degree of freedom reduced dynamics, namely pitch dynamics. The reduced order model for only pitch dynamics is considered with states x_1, x_3, and control v_1. The yaw motor voltage was set to zero, hence $v_2 = 0$. The reduced order dynamics can be expressed in the standard linear state space form, i.e. $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$ with $\mathbf{A}, \mathbf{B}$ as shown below.

$$\begin{aligned}
\dot{x}_1 &= x_3 \\
\dot{x}_3 &= \frac{K_{pp}v_1}{J_p + ml^2} - \frac{B_p}{J_p + ml^2}
\end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{B_p}{J_p+ml^2} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ \frac{K_{pp}}{J_p+ml^2} \end{bmatrix} \tag{8.4}$$

This reduced order model for pitch dynamics of the helicopter was used for demonstrating the intuitive control framework and identification of parameters $\mathbf{A}$, $\mathbf{B}$.

8.4 Experimental Results

The first two stages out of the three stage controller from Ch. 4 are implemented on the experimental platform. The unknown dynamics are stabilized by the adaptive controller for the first 10 seconds, then the stabilizing control policy from stabilization phase is used along with introduction of randomized band-limited noise in the control channel. The time required for the optimization phase to converge to the optimal controller is observed to be of the order of 10 seconds. It is to be noted that the approximated optimal controller varies with repeated experiments with exactly same parameters. But all the estimates are consistently close to the optimal controller when the initial estimates are close to the optimal control policy. This strong dependence on initial estimates on the approximated optimal control parameters has been documented in literature [64] while implementing Policy Iteration methods for the same experimental setup. The mildly uncertain nature of experimental results prompt that every set of controller parameters is tested on the experiment for a chosen number of times. The results are thus post processed for a mean estimate and a standard deviation from a set of 10 experiments for every set of controller parameters. All the presented results show the stabilization phase till 10 seconds, and the length of optimization phase is characterized by two parameters (number of samples N, and sampling time T_s). The baseline controller parameters are considered 2000 samples at a rate of $100Hz$ for the optimization phase. For this case the data is collected

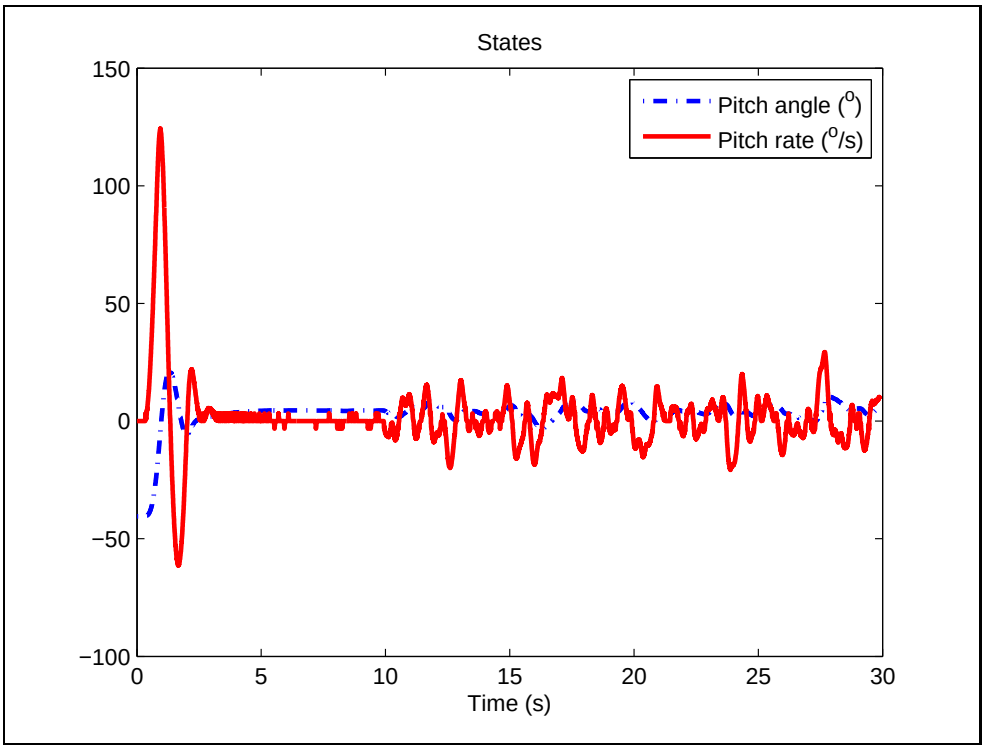

Figure 8.2. State history - baseline experimental.

for 20 seconds and the optimal controller is approximated for LQR parameters given below.

$$\mathbf{Q}_1 = \begin{bmatrix} 500 & 0 \\ 0 & 100 \end{bmatrix} \qquad \mathbf{R}_1 = 0.1 \tag{8.5}$$

One instance of the experiment carried out with the baseline controller yields states as shown in Fig. [8.2], and control signal as shown in Fig. [8.3] both of which indicate regulating behavior till 10 seconds and bounded behavior for the next 20 seconds. The optimal controller calculated using knowledge of $\mathbf{A}, \mathbf{B}$ is characterized by parameters $\mathbf{K}_{lqr}, \mathbf{P}_{lqr}$.

$$\mathbf{K}_{lqr1} = \begin{bmatrix} 70.7107 \\ 28.8699 \end{bmatrix} \qquad \mathbf{P}_{lqr1} = \begin{bmatrix} 231.8527 & 2.9877 \\ 2.9877 & 1.2198 \end{bmatrix} \tag{8.6}$$

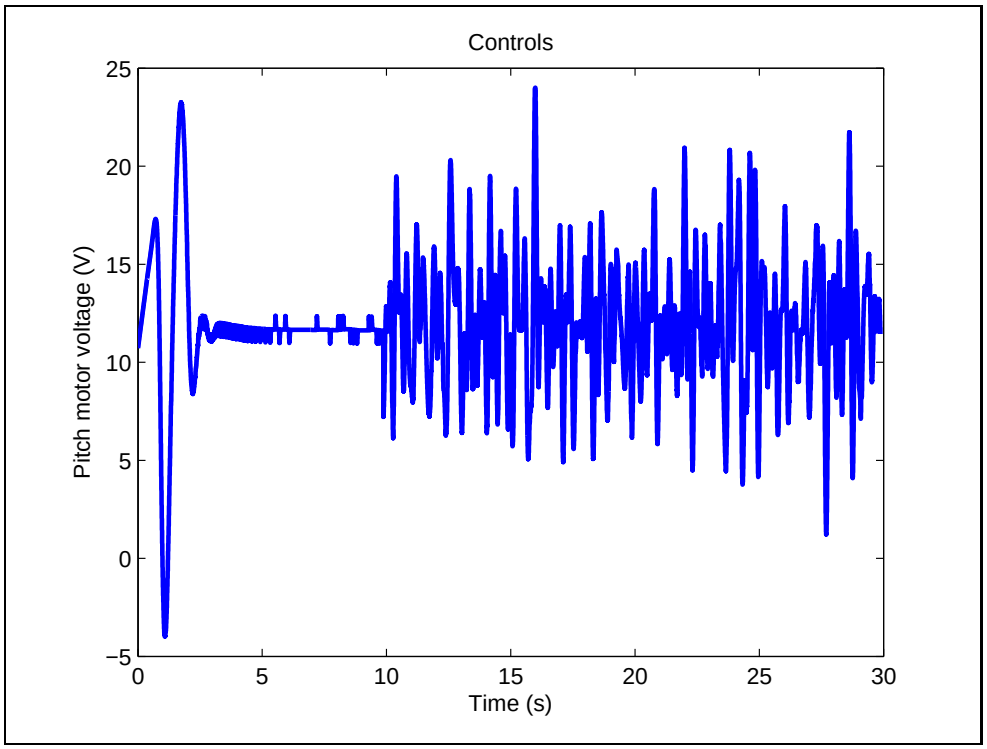

Figure 8.3. Control input history - baseline experimental.

This experiment is carried out with randomized noise in the optimization phase for 10 instances. The mean and standard deviation of the approximated controller parameters $\hat{\mathbf{K}}, \hat{\mathbf{P}}$ are tabulated for various sampling parameters ($N,\ T_s$) during optimization. This study shows the effect of sampling parameters on the results of value iteration scheme on the experimental platform. Table [8.1] shows mean and standard deviation of approximated solutions calculated over a sample of 10 experiments each. Fig. [8.4] shows the same comparative study for approximated gain on a 2D plot. The ellipses are centered around the mean value and length of major and minor axes represent the 3σ bounds for each parameter.

The most important observation is the loss of accuracy in approximating the term $\hat{\mathbf{K}}_{12}$ for all cases. This is attributed to the synthesized measurements of pitch angular rate for the experimental setup. These measurements are synthesized using

Table 8.1. Effect of sampling parameters on $\hat{\mathbf{K}}, \hat{\mathbf{P}}$

	$E(\hat{\mathbf{K}})$	$\sigma(\hat{\mathbf{K}})$	$E(\hat{\mathbf{P}})$	$\sigma(\hat{\mathbf{P}})$
Baseline ($N = 2000$, $T_s = 0.01$ seconds)	$\begin{bmatrix} 77.9597 \\ 44.0952 \end{bmatrix}^\top$	$\begin{bmatrix} 4.0755 \\ 2.4644 \end{bmatrix}^\top$	$\begin{bmatrix} 196.7443 & 1.7176 \\ 1.7176 & 1.5155 \end{bmatrix}$	$\begin{bmatrix} 25.8762 & 1.4530 \\ 1.4530 & 1.7047 \end{bmatrix}$
Lower sampling frequency ($N = 1000$, $T_s = 0.02$ seconds)	$\begin{bmatrix} 76.1145 \\ 45.5832 \end{bmatrix}^\top$	$\begin{bmatrix} 7.5025 \\ 3.9894 \end{bmatrix}^\top$	$\begin{bmatrix} 223.3859 & 0.2327 \\ 0.2327 & 2.8286 \end{bmatrix}$	$\begin{bmatrix} 35.3177 & 1.2823 \\ 1.2823 & 0.6805 \end{bmatrix}$
Higher sampling frequency ($N = 4000$, $T_s = 0.005$ seconds)	$\begin{bmatrix} 73.3145 \\ 47.2528 \end{bmatrix}^\top$	$\begin{bmatrix} 10.0982 \\ 4.6066 \end{bmatrix}^\top$	$\begin{bmatrix} 125.3759 & 4.4263 \\ 4.4263 & 0.9212 \end{bmatrix}$	$\begin{bmatrix} 23.2791 & 1.6293 \\ 1.6293 & 0.8093 \end{bmatrix}$
Lower sampling period ($N = 1000$, $T_s = 0.01$ seconds)	$\begin{bmatrix} 66.5645 \\ 43.6001 \end{bmatrix}^\top$	$\begin{bmatrix} 8.7595 \\ 2.8566 \end{bmatrix}^\top$	$\begin{bmatrix} 140.6242 & 2.8006 \\ 2.8006 & 1.5795 \end{bmatrix}$	$\begin{bmatrix} 36.2553 & 2.0020 \\ 2.0020 & 1.0795 \end{bmatrix}$
Higher sampling period ($N = 3000$, $T_s = 0.01$ seconds)	$\begin{bmatrix} 76.4449 \\ 45.6191 \end{bmatrix}^\top$	$\begin{bmatrix} 6.4604 \\ 4.0597 \end{bmatrix}^\top$	$\begin{bmatrix} 193.6810 & 2.3759 \\ 2.3759 & 2.4218 \end{bmatrix}$	$\begin{bmatrix} 22.4563 & 1.5735 \\ 1.5735 & 0.5874 \end{bmatrix}$

a high pass filter (Eq. (8.1)) which leads to a noisy angular rate measurement. The sampling parameter study suggests that both faster and slower sampling frequency yield worse results compared to the baseline case. Lower sampling frequency results in omission of higher frequency data which explains the deterioration of accuracy. On the other hand higher sampling frequency results in corruption of collected data by high frequency noise. Thus the sampling frequency needs to be high enough to capture all the frequencies in the introduced noise in input channel, but low enough to ignore the extraneous sources of noise. The baseline case seems to strike that balance for sampling frequency.

The sampling parameter study also reveals the effect of sampling period on the accuracy of results. The baseline case approximates the optimal controller by collecting data for 20 seconds. Lower sampling period results in the omission of low frequency data which can directly effect the accuracy of solution. Higher sampling period effects the accuracy due to the presence of extraneous noise. It is observed that the accuracy of results deteriorate when the sampling period is either decreased (10 seconds), or increased (30 seconds). In a simulation example both higher sampling rate and higher sampling period result in better results due to the absence of extrane-

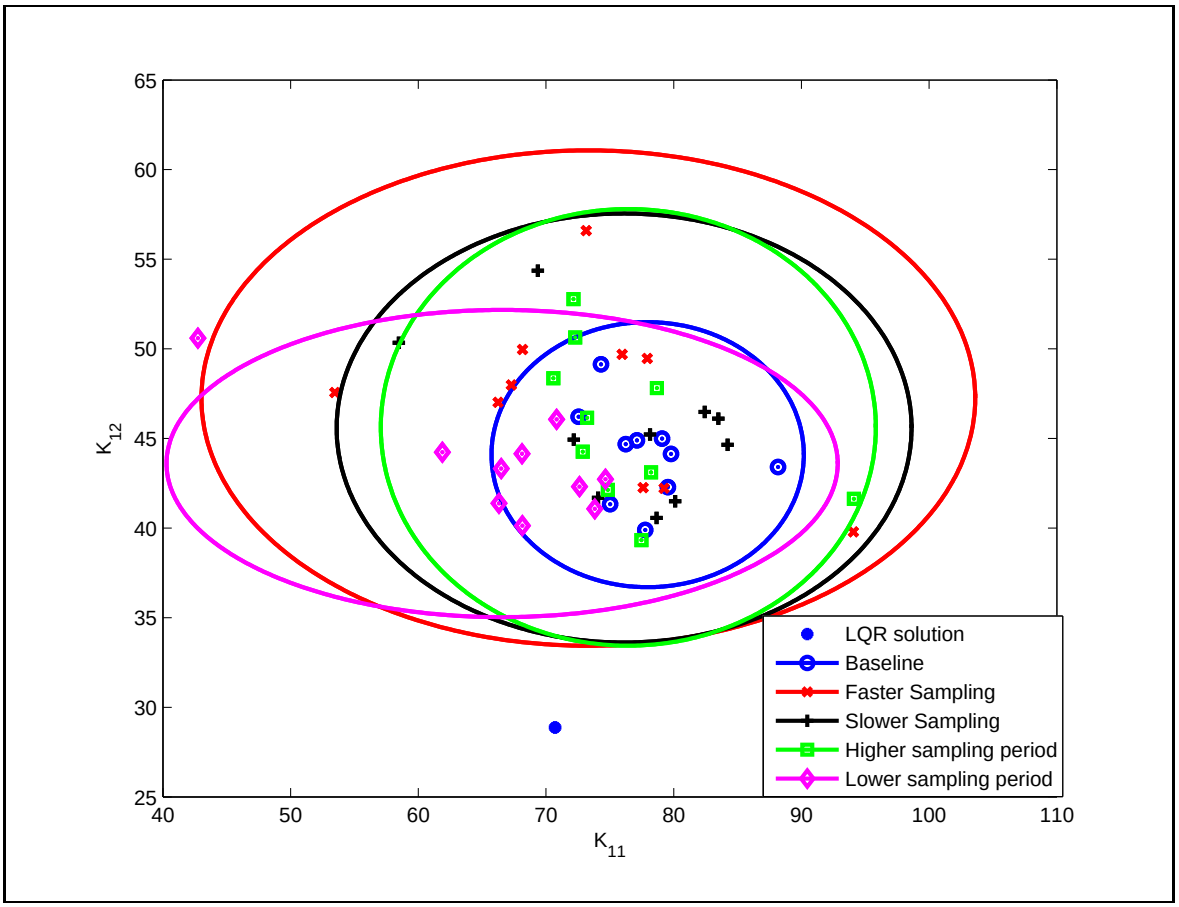

Figure 8.4. Effect of sampling parameters on $\hat{\mathbf{K}}$.

ous sources of noise. The experimental results suggest that sampling frequency and sampling time should be chosen close to the baseline case for best results.

Since the optimal controller for a given set of parameters $\mathbf{Q}, \mathbf{R}$ can be estimated with acceptable accuracy, the experiments are repeated with a second set of parameters $\mathbf{Q}_2 = 0.5\mathbf{Q}_1,\ \mathbf{R}_2 = \mathbf{R}_1$ with sampling parameters from the baseline case.

$$\mathbf{K}_{lqr2} = \begin{bmatrix} 50.0000 \\ 19.6948 \end{bmatrix} \qquad \mathbf{P}_{lqr2} = \begin{bmatrix} 118.0690 & 2.1126 \\ 2.1126 & 0.8322 \end{bmatrix} \tag{8.7}$$

A set of 10 experiments yield an approximation for the optimal control policy which is again more accurate with respect to the pitch angle.

$$\hat{\mathbf{K}}_2 = \begin{bmatrix} 57.9194 \\ 37.1559 \end{bmatrix} \qquad \hat{\mathbf{P}}_2 = \begin{bmatrix} 95.5341 & 2.9201 \\ 2.9201 & 1.4066 \end{bmatrix} \tag{8.8}$$

The identification routine developed in the previous chapters yields an estimate for parameters $\mathbf{A}$, $\mathbf{B}$.

$$\hat{\mathbf{A}} = \begin{bmatrix} -1.0965 & -0.2214 \\ 99.7724 & 50.4806 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 0 & 1.0000 \\ 0 & -9.2751 \end{bmatrix} \tag{8.9}$$

$$\hat{\mathbf{B}} = \begin{bmatrix} -0.0036 \\ 2.7897 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 2.3667 \end{bmatrix} \tag{8.10}$$

It is to be noted that although $\mathbf{B}$ and thus the constant $\frac{K_{pp}}{J_p+ml^2}$ is estimated with acceptable accuracy, the estimate $\hat{\mathbf{A}}$ does not compare well with the actual $\mathbf{A}$. This is observed to be a direct result to consistently poor accuracy of $\hat{\mathbf{K}}_{12}$, which in turn is attributed to noisy measurement of pitch rate in the experimental setup.

CHAPTER 9

Closing Remarks

The presented work outlines a hybrid control framework for unknown linear time invariant systems which identifies the unknown parameters. Although there are methods in literature to estimate the unknown parameters, such methods cannot be used in online applications due to the trade-off between closed loop stability and parameter convergence. The proposed controller methodically identifies parameters after collecting closed loop data for a certain period of time.

The proposed control framework is an additional layer of abstraction which can be represented by a state machine(stages) for the controller itself. An unknown system always starts in the stabilization stage where it is rendered stable with a linear state feedback gain. Then the stabilized system is subject to either Policy Iteration (see Ch. 3) or Value Iteration (see Ch. 4) to approximate the optimal feedback controller without the knowledge of linear system parameters. This process is repeated with various quadratic cost functions until enough data is acquired to identify the unknown parameters. A unique solution to the unknown linear system parameters makes the identification process complete. Further extension of the framework to Lipschitz and Hamiltonian nonlinearities further increases the scope of its applicability.

The simulation results from Ch. 3,4 show unique identification of linear parameters with acceptable accuracy with generalized policy iteration methods used in optimization stage. It can be observed that the accuracy of approximated parameters decreases with increasing uncertainties. It is also observed that when model reference adaptive control is used along with generalized policy iteration, the stability

requirements can be relaxed. The reachable sets calculated in Ch. 7 from identified parameters are observed to be accurate to an acceptable degree.

The simulation results from Ch. 5 show successful identification of linear parameters with accuracy comparable to the Lipschitz constant. This behavior is justified since the results are approximated from a worst case optimal control solution. It is to be noted that such estimates for linear parameters would be exceedingly erroneous for a higher Lipschitz constant.

The simulation results from Ch. 6 show successful regulation of angular velocities for rigid body dynamics without the knowledge of inertia properties. The algorithm also yields an estimate for the unknown parameter without the need for offline experiments. The resulting controller although not optimal for the nonlinear dynamics, it approaches the optimal controller close to equilibrium. This algorithm could be implemented for most spacecraft which can be considered rigid bodies operating in a relatively disturbance free environment. Such online technique for inertia estimation can serve as a spacecraft health monitoring aid which would report any changes in spacecraft integrity. Although the present work shows a regulation case for angular velocities, the algorithm can be augmented to control attitude in the outer loop.

The proposed controller is implemented on a table-top 2-DOF Helicopter which is a nonlinear MIMO system with real world issues like unwanted noise, limited controller bandwidth, and physically restricted state space. The proposed method is implemented on a reduced order model of the helicopter's dynamics due to physical restrictions of the platform. A study of sampling parameters for the optimization phase reveals that both sampling frequency and sampling period need to be chosen subject to the frequency of introduced input noise and the frequency of extraneous noise. It is also observed that some parts of the optimal control policy were approx-

imated with acceptable accuracy. The less accurate results for pitch rate gains is attributed to the use of synthesized angular rate in lieu of true angular rate measurements. As a result, the identification routine also shows mixed results with a good estimate for some parameters ($\mathbf{B}$) and mediocre estimates for others($\mathbf{A}$). The proposed framework does not model extraneous noises inherent to the experimental platform. This observation leads to a possible avenue for future work on extending the proposed framework to models with extraneous noise. The adaptive control methods may be generalized by robust adaptive control methods, and the LQR problem may be generalized by a differential game with H_∞ formulation. Such robustification of the proposed framework could make it more conducive to implementation on an experimental platform.

APPENDIX A

A Primer on Kronecker Algebra

This appendix discusses the mathematical concepts needed to pose the system identification problem in an efficient manner. A brief introduction to the vectorization operation and Kronecker product is shown below using matrices $\mathbf{M}, \mathbf{N} \in \mathbb{R}^{n \times n}$.

$$vec(\mathbf{M}) = \begin{bmatrix} m_{11} \\ m_{21} \\ \vdots \\ m_{n1} \\ m_{12} \\ \vdots \\ m_{nn} \end{bmatrix} \quad \text{where} \quad \mathbf{M} = \begin{bmatrix} m_{11} m_{12} \cdots m_{1n} \\ m_{21} m_{22} \cdots m_{2n} \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ m_{n1} m_{n2} \cdots m_{nn} \end{bmatrix} \tag{A.1}$$

Kronecker product is defined as shown in Eq. (A.2). It is a generalization of outer product for matrices and is represented by $\otimes$.

$$\mathbf{M} \otimes \mathbf{N} = \begin{bmatrix} m_{11}\mathbf{N} & m_{12}\mathbf{N} & \cdots & m_{1n}\mathbf{N} \\ m_{21}\mathbf{N} & m_{22}\mathbf{N} & \cdots & m_{2n}\mathbf{N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1}\mathbf{N} & m_{n2}\mathbf{N} & \cdots & m_{nn}\mathbf{N} \end{bmatrix} \tag{A.2}$$

Vectorized forms of $\mathbf{A}, \mathbf{A^T}$ are related by a unique permutation matrix $\mathbf{\Pi}$ for any matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. There could be numerous ways of creating such matrix. $\mathbf{\Pi}$ will be constructed using the block matrix form shown below.

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{\Pi}^{11} & \mathbf{\Pi}^{12} & \cdots & \mathbf{\Pi}^{1n} \\ \mathbf{\Pi}^{21} & \mathbf{\Pi}^{22} & \cdots & \mathbf{\Pi}^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Pi}^{n1} & \mathbf{\Pi}^{n2} & \cdots & \mathbf{\Pi}^{nn} \end{bmatrix} \tag{A.3}$$

$\mathbf{\Pi}^{ij}$ can be constructed using column and row push and pop operations which will be defined on the matrix $\mathbf{\Pi}^{11}$.

$$\mathbf{\Pi}_{ij} = \mathbf{\Phi}_D^{j-1}\mathbf{\Pi}_{11}\mathbf{\Phi}_R^{i-1} \tag{A.4}$$

$\mathbf{\Pi}^{11}$ is defined as a matrix of appropriate order with the only non zero element corresponding to first row and first column equal to 1. In index notation $\mathbf{\Pi}^{11} \in \mathbb{R}^{n\times n}, s.t. \mathbf{\Pi}^{11}_{pq} = 1 \quad \text{iff.} \quad p = q = 1.$

The matrices $\mathbf{\Phi}_D$,$\mathbf{\Phi}_R$ are special kind of permutation matrices of order n. Pre-multiplication with $\mathbf{\Phi}_D$ performs a row operation as shown below. This operation is similar to push and pop operations seen in a memory stack. Thus the operation is called push down pop up.

$$\mathbf{\Phi}_D \begin{bmatrix} \mathbf{Row_1 A} \\ \mathbf{Row_2 A} \\ \vdots \\ \mathbf{Row_n A} \end{bmatrix} = \begin{bmatrix} \mathbf{Row_n A} \\ \mathbf{Row_1 A} \\ \vdots \\ \mathbf{Row_{n-1} A} \end{bmatrix} \quad \forall \quad \mathbf{A} = \begin{bmatrix} \mathbf{Row_1 A} \\ \mathbf{Row_2 A} \\ \vdots \\ \mathbf{Row_n A} \end{bmatrix} \in \mathbb{R}^{n\times n} \tag{A.5}$$

Post-multiplication with $\mathbf{\Phi}_R$ performs a column operation as shown below. This operation is called push right pop left.

$$\begin{bmatrix} \mathbf{Col_1 A Col_2 A \cdots Col_n A} \end{bmatrix} \mathbf{\Phi}_R = \begin{bmatrix} \mathbf{Col_n A Col_1 A \cdots Col_{n-1} A} \end{bmatrix} \tag{A.6}$$
$$\forall\ \mathbf{A} = \begin{bmatrix} \mathbf{Col_1 A Col_2 A \cdots Col_n A} \end{bmatrix} \in \mathbb{R}^{n\times n}$$

These transformation matrices can be found by performing the same push and pop operations on the identity matrix $\mathbf{I^n}$.

$$\mathbf{\Phi}_D = \begin{bmatrix} \mathbf{Row_n I^n} \\ \mathbf{Row_1 I^n} \\ \vdots \\ \mathbf{Row_{n-1} I^n} \end{bmatrix} \tag{A.7}$$

$$\mathbf{\Phi}_R = \begin{bmatrix} \mathbf{Col_2 I^n} \cdots \mathbf{Col_n I^n Col_1 I^n} \end{bmatrix} \tag{A.8}$$

Eq. (A.3) and (A.4) can be used to evaluate a permutation matrix $\mathbf{\Pi}$ which relates a vectorization of a square matrix with the vectorization of its transpose for a given order n.

REFERENCES

[1] P. Antoniou, A. Pitsillides, A. Engelbrecht, and T. Blackwell, "Mimicking the bird flocking behavior for controlling congestion in sensor networks," in *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, Nov 2010, pp. 1–7.

[2] J. Dyhr, N. Cowan, D. Colmenares, K. Morgansen, and T. Daniel, "Autostabilizing airframe articulation: Animal inspired air vehicle control," in *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, Dec 2012, pp. 3715–3720.

[3] B. Jakimovski, M. Kotke, M. Harenz, and E. Maehle, "A Biologically Inspired Approach for Self-Stabilizing Humanoid Robot Walking," in *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*. Springer Berlin Heidelberg, 2010, vol. 329, pp. 302–313.

[4] A. E. Bryson and Y. C. Ho, *Applied Optimal Control: Optimization, Estimation and Control*. CRC Press, 1975.

[5] J. L. Speyer and D. H. Jacobson, *Primer on Optimal Control Theory*. Society for Industrial and Applied Mathematics, 2010, vol. 20.

[6] K. J. Astrom and B. Wittenmark, *Adaptive Control*, 2nd ed. Addison-Wesley Longman Publishing Co., Inc., 1994.

[7] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Dover Publications, 2012.

[8] W. T. Miller, P. J. Werbos, and R. S. Sutton, *Neural Networks for Control*. MIT press, 1995.

[9] R. Sutton, A. Barto, and R. J. Williams, "Reinforcement learning is direct adaptive optimal control," *Control Systems, IEEE*, vol. 12, no. 2, pp. 19–22, April 1992.

[10] 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.

[11] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[12] D. P. Bertsekas and S. Ioffe, "Temporal differences-based policy iteration and applications in neuro-dynamic programming," *Lab. for Info. and Decision Systems Report LIDS-P-2349, MIT, Cambridge, MA*, 1996.

[13] D. Kleinman, "On an iterative technique for Riccati equation computations," *IEEE Transactions on Automatic Control*, vol. 12, pp. 114–115, 1968.

[14] D. Vrabie, M. Abu-Khalaf, F. Lewis, and Y. Wang, "Continuous-Time ADP for Linear Systems with Partially Unknown Dynamics," in *Approximate Dynamic Programming and Reinforcement Learning, 2007. ADPRL 2007. IEEE International Symposium on*, April 2007, pp. 247–253.

[15] J. Y. Lee, J. B. Park, and Y. H. Choi, "A novel generalized value iteration scheme for uncertain continuous-time linear systems," in *Decision and Control (CDC), 2010. 49th IEEE Conference on*, Dec 2010, pp. 4637–4642.

[16] J. Y. Lee, J. B. Park, and Y. H. Choi, "On integral value iteration for continuous-time linear systems," in *American Control Conference (ACC), 2013*, June 2013, pp. 4215–4220.

[17] F. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *Circuits and Systems Magazine, IEEE*, vol. 9, no. 3, pp. 32–50, Third Quarter 2009.

[18] Y. Jiang and Z. P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics ," *Automatica*, vol. 48, no. 10, pp. 2699 – 2704, 2012.

[19] L. Ljung and K. Glover, "Frequency domain versus time domain methods in system identification ," *Automatica*, vol. 17, no. 1, pp. 71 – 86, 1981.

[20] L. Ljung, *System Identification: Theory for the User*, 1999, vol. 7632.

[21] M. B. Tischler and R. K. Remple, *Aircraft and rotorcraft system identification: Engineering methods with flight test examples*. American Institute of Aeronautics and Astronautics, 2012.

[22] B. L. Ho and R. E. Kalman, "Editorial: Effective construction of linear state-variable models from input/output functions," *Automatisierungstechnik*, vol. 14, no. 1-12, pp. 545–548, 1966.

[23] K. J. Astrom and T. Bohlin, "Numerical Identification of Linear Dynamic Systems from Normal Operating Records," in *Theory of self-adaptive control systems*. Springer, 1966, pp. 96–111.

[24] M. Gevers, "A personal view of the development of system identification: A 30-year journey through an exciting field," *Control Systems, IEEE*, vol. 26, no. 6, pp. 93–105, Dec 2006.

[25] I. Gustavsson, L. Ljung, and T. Soderstrom, "Identification of processes in closed loop - identifiability and accuracy aspects ," *Automatica*, vol. 13, no. 1, pp. 59 – 75, 1977.

[26] K. Liu and R. Skelton, "Closed-loop identification and iterative controller design," in *Decision and Control, 1990., Proceedings of the 29th IEEE Conference on*. IEEE, 1990, pp. 482–487.

[27] R. J. Schrama, "Accurate identification for control: The necessity of an iterative scheme," *Automatic Control, IEEE Transactions on*, vol. 37, no. 7, pp. 991–994, 1992.

[28] D. Bayard, Y. Yam, and E. Mettler, "A criterion for joint optimization of identification and robust control," *Automatic Control, IEEE Transactions on*, vol. 37, no. 7, pp. 986–991, Jul 1992.

[29] P. Ioannou and C. Johnson, "Reduced-order performance of parallel and series-parallel identifiers with weakly observable parasitics ," *Automatica*, vol. 19, no. 1, pp. 75 – 80, 1983.

[30] Y. Zhang, P. Ioannou, and C. C. Chien, "Parameter convergence of a new class of adaptive controllers," *Automatic Control, IEEE Transactions on*, vol. 41, no. 10, pp. 1489–1493, Oct 1996.

[31] J. Lygeros, "On reachability and minimum cost optimal control," *Automatica*, vol. 40, no. 6, pp. 917–927, 2004.

[32] I. M. Mitchell and J. A. Templeton, "A Toolbox of Hamilton-Jacobi Solvers for Analysis of Nondeterministic Continuous and Hybrid Systems," in *Hybrid systems: computation and control*. Springer, 2005, pp. 480–494.

[33] H. G. Kwatny, J. E. T. Dongmo, B. C. Chang, G. Bajpai, M. Yasar, and C. Belcastro, "Aircraft Accident Prevention: Loss-of-Control Analysis," in *AIAA Guidance, Navigation and Control Conference*, 2009.

[34] C. Tomlin, J. Lygeros, and S. Sastry, "Aerodynamic envelope protection using hybrid control," in *American Control Conference, 1998. Proceedings of the 1998*, vol. 3. IEEE, 1998, pp. 1793–1796.

[35] M. Oishi, I. Mitchell, C. Tomlin, and P. Saint-Pierre, "Computing viable sets and reachable sets to design feedback linearizing control laws under saturation,"

in *Decision and Control, 2006 45th IEEE Conference on*. IEEE, 2006, pp. 3801–3807.

[36] P. Nuthi and K. Subbarao, "Experimental verification of linear and adaptive control techniques for a 2-dof helicopter," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 137, 2015.

[37] P. Nuthi and K. Subbarao, "Computation of safe and reachable sets for model-free dynamical systems: Aircraft longitudinal dynamics," in *AIAA Atmospheric Flight Mechanics Conference, 2014. Proceedings of AIAA Scitech 2014*, 2014.

[38] P. Nuthi and K. Subbarao, "Aspects of intuitive control: Stabilize, optimize, and identify," in *AIAA Guidance Navigation and Control Conference, 2015.Proceedings of AIAA Scitech 2015*, 2015.

[39] P. Nuthi and K. Subbarao, "Implementation and testing of adaptive augmentation techniques on a 2-dof helicopter," in *Dynamics, Vibration and Control, 2013. Proceedings of ASME 2013 International Mechanical Engineering Congress and Exposition*. American Society of Mechanical Engineers, 2013.

[40] P. Nuthi and K. Subbarao, "Autonomous vertical landing on a marine vessel," in *AIAA Atmospheric Flight Mechanics Conference, 2014. Proceedings of AIAA Scitech 2014*, 2014.

[41] J. L. Junkins, K. Subbarao, and A. Verma, "Structured Adaptive Control for Poorly Modeled Nonlinear Dynamical Systems," *CMES- Computer Modeling in Engineering and Sciences*, vol. 1, no. 4, pp. 99–117, 2000.

[42] A. Verma, K. Subbarao, and J. L. Junkins, "A novel trajectory tracking methodology using structured adaptive model inversion for uninhabited aerial vehicles," in *American Control Conference, 2000. Proceedings of the 2000*, vol. 2. IEEE, 2000, pp. 859–863.

[43] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.

[44] M. Branicky, "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems," *Automatic Control, IEEE Transactions on*, vol. 43, no. 4, pp. 475–482, Apr 1998.

[45] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed. Princeton University Press, 2009.

[46] O. Ma, H. Dang, and K. Pham, "On-Orbit Identification of Inertia Properties of Spacecraft Using a Robotic Arm," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 31, no. 6, pp. 1761–1771, 2008.

[47] M. C. Norman, M. A. Peck, and D. J. O'Shaughnessy, "In-Orbit Estimation of Inertia and Momentum-Actuator Alignment Parameters," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 34, no. 6, pp. 1798–1814, 2011.

[48] E. Wilson, C. Lages, and R. Mah, "On-line gyro-based, mass-property identification for thruster-controlled spacecraft using recursive least squares," in *Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on*, vol. 2. IEEE, 2002, pp. II–334.

[49] J. Ahmed, V. T. Coppola, and D. S. Bernstein, "Adaptive Asymptotic Tracking of Spacecraft Attitude Motion with Inertia Matrix Identification," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 21, no. 5, pp. 684–691, 1998.

[50] J. Ahmed and D. S. Bernstein, "Globally convergent adaptive control of spacecraft angular velocity without inertia modeling," in *American Control Conference, 1999. Proceedings of the 1999*, vol. 3. IEEE, 1999, pp. 1540–1544.

[51] A. K. Sanyal, M. Chellappa, J. Valk, J. Ahmed, J. Shen, and D. S. Bernstein, "Globally convergent adaptive tracking of spacecraft angular velocity with inertia identification and adaptive linearization," in *Decision and Control, 2003. Proceedings of 42nd IEEE Conference on*, vol. 3. IEEE, 2003, pp. 2704–2709.

[52] N. A. Chaturvedi, D. S. Bernstein, J. Ahmed, F. Bacconi, and N. H. McClamroch, "Globally convergent adaptive tracking of angular velocity and inertia identification for a 3-DOF rigid body," *Control Systems Technology, IEEE Transactions on*, vol. 14, no. 5, pp. 841–853, 2006.

[53] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *Automatic Control, IEEE Transactions on*, vol. 50, no. 7, pp. 947–957, 2005.

[54] "Quanser Consulting Inc., 2-dof Helicopter," `www.quanser.com`(last visited - April 2013).

[55] P. Castillo, P. Albertos, P. Garcia, and R. Lozano, "Simple Real-time Attitude Stabilization of a Quad-rotor Aircraft With Bounded Signals," in *Decision and Control, 2006. 45th IEEE Conference on*, December 2006, pp. 1533–1538.

[56] B. Zheng and Y. Zhong, "Robust Attitude Regulation of a 3-DOF Helicopter Benchmark: Theory and Experiments," *Industrial Electronics, IEEE Transactions on*, vol. 58, no. 2, pp. 660–670, Feb 2011.

[57] H. Liu, G. Lu, and Y. Zhong, "Robust LQR attitude control of a 3-DOF laboratory helicopter for aggressive maneuvers. ," *Industrial Electronics, IEEE Transactions on*, vol. 60, no. 10, pp. 4627–4636, 2013.

[58] K. Narendra and A. Annaswamy, *Stable Adaptive Systems*, ser. Information and System Sciences. Prentice Hall, 1989.

[59] E. Kharisov, N. Hovakimyan, and K. J. Astrom, "Comparison of Several Adaptive Controllers According to Their Robustness Metrics ," in *Proceedings of AIAA Guidance, Navigation and Control Conference*, 2010.

[60] C. Cao and N. Hovakimyan, *L1 Adaptive Control Theory*. Society for Industrial and Applied Mathematics, 2010.

[61] C. Cao and N. Hovakimyan, "Design and Analysis of a Novel L1 Adaptive Control Architecture with Guaranteed Transient Performance ," *IEEE Transactions on Automatic Control*, vol. 53, no. 2, pp. 586–591, 2008.

[62] C. Cao and N. Hovakimyan, "L1 Adaptive Controller for a Class of Systems with Unknown Nonlinearities: Part I ," in *Proceedings of American Control Conference*, 2008, pp. 4093–4098.

[63] C. Cao and N. Hovakimyan, "L1 Adaptive Controller for Nonlinear Systems in the Presence of Unmodelled Dynamics: Part II ," in *Proceedings of American Control Conference*, 2008, pp. 4099–4104.

[64] G. Atmeh and K. Subbarao, "Experimental Results for Adaptive, Optimal Control of a 2-DOF Helicopter," in *AIAA InfoTech, 2015. Proceedings of AIAA Scitech 2015*, 2015.

BIOGRAPHICAL STATEMENT

Pavan Nuthi was born in Machilipatnam, India, in 1988. He received his B.Tech. degree in Aerospace Engineering from Indian Institute of Technology Madras, India, in 2010, and began a bachelor's to doctoral degree program at University of Texas at Arlington. He has been awarded the Enhanced Graduate Teaching Assistant grant to support his doctoral research. Outside of his doctoral pursuit, he worked on modeling human ear canal for developing better noise canceling headphones for Lightspeed Aviation in the summer of 2013. His current research interest is in the area of nonlinear control systems and robotics. He is a student member of IEEE, AIAA and, ASME, with aptitude for programming and building mechatronic systems.