DEPARTMENT OF INFORMATICS,
BIOENGINEERING, ROBOTICS, AND SYSTEMS
ENGINEERING

UNIVERSITÀ DEGLI STUDI
DI GENOVA

Optimal Input Design for Active Parameter Identification of Dynamic Nonlinear Systems

by
Muhammad Zeeshan Babar

Supervised by
Prof. Marco Baglietto

A Thesis submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy
in Bioengineering and Robotics, Curriculum of
Robotics and Autonomous Systems.

December 2017

I want to dedicate this work to my dearest father, mother and my beloved wife for their continuous support during this period.

ACKNOWLEDGMENT

First and foremost, I would like to thank ALMIGHTY ALLAH, who gave me the courage, guidance and atmosphere to complete my postgraduate studies at University of Genoa, Italy. The perseverance and determination granted by ALLAH helped me to bear the hard times to produce this thesis. All respects to Prophet MUHAMMAD (P.B.U.H), who is the last messenger, whose life is a perfect model for the whole humanity.

I would like to thank my supervisor, **Prof. Marco Baglietto**, with whom it has been an honor to work. I am grateful for giving me this opportunity to pursue a PhD degree and all the guidance along the way. I am thankful for all the patience with my questions and good advice over the years. His wealth of ideas, clarity of thoughts, enthusiasm and energy have made my working with him, an exceptional experience.

I want to thank all the past and present colleagues at the department for the fun times, great discussions, early lunches and late hangouts all over the Genoa.

I want to thank the administration and coordinators of my program for taking care of all the things that make the department run so smoothly.

I am really thankful on the efforts of my loving parents and family who kept me motivated, guided and focused throughout my PhD work. Their help in various regards contributed in keeping my moral high.

I would like to acknowledge the efforts of Dr. Yasir Awais Butt for his help and guidance at several steps of my research period.

In the end, I am thankful to University of Genoa, Italy for providing me the scholarship and great atmosphere to complete my PhD studies.

Muhammad Zeeshan Babar
February 2018

ABSTRACT

There are many important aspects to be considered while designing optimal excitation signal for system identification experiment in control applications. Active parameter identification is an important issue in system and control theory. In this dissertation, the problem of optimal input design for active parameter identification of dynamic nonlinear system is addressed.

Real life physical systems are identified by excitation with a suitable input signal and observing the resulting output behavior of the system. It is important to choose the input signal intelligently in the sense that it is responsible to determine the accuracy and nature of the unknown system characteristics. This leads to a spurred interest in designing such an optimal excitation signals that can yield maximal information from the identification experiment. The information obtained from parameter identification is usually not accurate due to incomplete knowledge of the system, disturbance as exogenous inputs and noisy measurements. Hence, the input spectrum is designed in such a way that it can improve the system performance and shape the quality of obtained information. A well-designed input signal can maximize the amount of information and reduce the experimental cost and time. The input signal is usually given some a-priori characteristics (knowledge on the pdf) so that "excitation" of the system is guaranteed. In this thesis, a closed-loop method is investigated which is able to improve the parameter identification on the basis of the actual system's behavior. The effectiveness of the proposed algorithm is presented by the experimental results which corresponds to the perfect identification of the unknown parameter vector.

The major technical contribution of this work is to propose an optimal feedback input design method for active parameter identification of dynamic nonlinear systems. The proposed framework can design such optimal excitation signals, considering the information from the identified parameters, that can maximize the amount of information from the identified parameters, guarantee to meet the specified control performance and minimize some cost function of the error covariance matrix of the identified parameters. The problem is formulated in a receding horizon framework where extended Kalman filter is used for system identification and the optimal input is designed in a nonlinear model predictive control framework. In order to carry out a comparison study, also Unscented Kalman Filter and Gaussian Sum Filter are used for the active parameter identification of dynamic nonlinear system. Towards this end, a suitable optimality criterion related to the unknown parameters is proposed and motivated as an information measure. The aim of the optimal input design is to yield maximal information from the unknown system by minimizing the cost related to the unknown parameters while maintaining some process performance and satisfying the possible constraints. Simulations are performed to show the effectiveness of the proposed algorithm.

TABLE OF CONTENTS

Acknowledgment . . . . . **ii**
Abstract . . . . . **iii**
Table of Contents . . . . . **iv**
List of Figures . . . . . **viii**
List of Tables . . . . . **ix**
List of Acronyms . . . . . **x**

Chapter 1
Introduction **1**
1.1 Motivating Examples 3
1.2 Problem Formulation 5
1.3 Thesis Structure 6

Chapter 2
State of the Art **8**
2.1 Nonlinear System Identification 9
2.1.1 History 12
2.1.1.1 Identification Strategies: Overview 13
2.1.1.2 Estimation Strategies: Overview 15
2.1.1.3 Kalman Filter and it's Variants 16
2.1.1.4 Particle Filter 17
2.1.1.5 Gaussian Sum Filter 18
2.2 Optimal Input Design in Nonlinear System Identification 18
2.2.1 Background 19
2.2.1.1 Open loop Optimal Input Design 20
2.2.1.2 Closed loop Optimal Input Design 21
2.2.1.3 Sequential Input Design 22
2.2.2 Applications of Optimal Input Design 23
2.2.2.1 Input Design for Industrial Process 24
2.2.2.2 Input Design for System Biology 25
2.2.3 Model Predictive Control for Optimal Input Design 25
2.2.3.1 Model Predictive Control 26
2.3 Chapter Summary 27

Chapter 3
Fundamentals of Optimal Input Design **29**
3.1 Optimality Criteria for System Identification 30
3.1.1 Nonlinear Discrete-time System 30
3.1.2 Optimality Criteria 30
3.1.3 Information-based Criteria 31
3.1.3.1 G-Optimality Criterion 31
3.1.3.2 D-Optimality Criterion 32

3.1.3.3 A-Optimality Criterion 33
3.1.3.4 E-Optimality Criterion 33
3.1.3.5 I-Optimality Criterion 34
3.1.4 Distance-based Criteria 35
3.1.4.1 U-Optimality Criterion 35
3.1.4.2 S-Optimality Criterion 35
3.1.5 Compound-design Criteria 36
3.1.5.1 DT-Optimality Criterion 36
3.1.5.2 CD-Optimality Criterion 36
3.2 System Identification Strategies 37
3.2.1 Extended Kalman Filter 37
3.2.2 Unscented Kalman Filter 40
3.2.3 Gaussian Sum Filter 42
3.2.3.1 Gaussian Sum Approximations 43
3.3 Nonlinear Model Predictive Control for Optimal Input Design . . . 48
3.3.1 Model Predictive Control Framework 48
3.4 Chapter Summary 51

Chapter 4
Optimal Input Design for Active Parameter Identification **52**
4.1 System Description 52
4.1.1 Preliminaries 52
4.1.2 Active Parameter Identification Problem 53
4.2 Identification Step 55
4.2.1 Extended Kalman Filter 56
4.2.2 Unscented Kalman Filter 56
4.2.3 Gaussian Sum Filter 57
4.3 Optimality Criterion 57
4.3.1 A-Optimality Criterion 58
4.4 Optimal Input Design Framework 59
4.4.1 Nonlinear Model Predictive Control 59
4.4.2 Proposed Framework 59
4.5 Chapter Summary 62

Chapter 5
Numerical Examples **63**
5.1 Numerical Examples 63
5.1.1 A Toy Example 64
5.1.1.1 Active Parameter Identification 65
5.1.1.1.1 Case 1 66
5.1.1.1.2 Case 2 68
5.1.1.1.3 Case 3 69
5.1.1.2 Classical Optimal Control Problem 71
5.1.1.2.1 Case 1 71
5.1.1.2.2 Case 2 73
5.1.1.2.3 Case 3 74

5.1.1.2.4 Case 4 . . . 75
5.1.1.3 Example Summary . . . 77
5.1.2 3-DOF Mobile Robot Model . . . 77
5.1.2.1 Overview . . . 77
5.1.2.2 Equations of Motion . . . 78
5.1.2.3 Proposed Framework . . . 80
5.1.2.4 Control Strategy . . . 80
5.1.2.5 Simulation Results . . . 81
5.1.2.5.1 Active Parameter Identification . . . 82
5.1.2.5.2 Classical Optimal Control Problem . . . 83
5.1.2.5.3 Trade-off between Identification and Control . . . 84
5.1.3 Example Summary . . . 86
5.2 Comparison Study for Proposed Framework . . . 86
5.2.1 2-DOF Mobile Robot Model . . . 86
5.2.1.1 Overview . . . 87
5.2.1.2 Equations of Motion . . . 87
5.2.1.3 OID Strategy for Two-Wheeled Mobile Robot . . . 89
5.2.2 Simulation Results . . . 91
5.2.2.1 Extended Kalman Filter . . . 92
5.2.2.1.1 EKF Case 1: Identification Experiment . . . 92
5.2.2.1.2 EKF Case 2: Classical Optimal Control Problem . . . 93
5.2.2.1.3 EKF Case 3: Trade-off between Identification and Control Performance . . . 95
5.2.2.2 Unscented Kalman Filter . . . 97
5.2.2.2.1 UKF Case 1: Identification Experiment . . . 97
5.2.2.2.2 UKF Case 2: Classical Optimal Control Problem . . . 98
5.2.2.2.3 UKF Case 3: Trade-off between Identification and Control Performance . . . 100
5.2.2.3 Gaussian Sum Filter . . . 101
5.2.2.3.1 GSF Case 1: Identification Experiment . . . 101
5.2.2.3.2 GSF Case 2: Classical Optimal Control Problem . . . 103
5.2.2.3.3 GSF Case 3: Trade-off between Identification and Control Performance . . . 103
5.2.3 Mean Square Error Comparison: EKF, UKF and GSF . . . 106
5.3 Chapter Summary . . . 109

Chapter 6
Conclusions, Recommendations and Future Work . . . **110**
6.1 Conclusions . . . 110
6.2 Recommendations . . . 112
6.3 Open Problems . . . 112

Chapter A
Appendix A **114**

Chapter B
Appendix B **116**

Chapter C
Appendix C **119**

References **120**

LIST OF FIGURES

2.1 Elements of System Identification 10
2.2 Sequential Input Design 23

3.1 EKF Flow Chart (One Cycle) 40
3.2 Receding horizon principle 50

4.1 Block Diagram: Active Parameter Identification 54

5.1 Measurement Regions 65
5.2 $x_0 = 0, \hat{x}_0 = 0.5, \hat{\theta}_0 = -0.5, \alpha = 0$ 67
5.3 White Gaussian Noise as Input Signal 68
5.4 $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5, \alpha = 0$ 69
5.5 $x_0 = -4, \hat{x}_0 = -3.5, \hat{\theta}_0 = -0.5, \alpha = 0$ 70
5.6 $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5, \alpha = 1, x_k^\star = 0$ 72
5.7 $x_0 = -4, \hat{x}_0 = -3.5, \hat{\theta}_0 = -0.5, \alpha = 1, x_k^\star = 0$ 73
5.8 $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta} = -0.5, \alpha = 1, x_k^\star = -4$ 75
5.9 $x_0 = 0, \hat{x}_0 = 0.5, \hat{\theta}_0 = -0.5, x_k^\star = -4$ 76
5.10 Side-view and Plane-view of two wheeled Robot 78
5.11 Active Parameter Identification ($\alpha = 0$) 83
5.12 Classical Optimal Control Problem ($\alpha = 1$) 84
5.13 Trade-off between identification and control ($\alpha = 0.1$) 85
5.14 Side-view of two wheeled Robot 88
5.15 EKF: Parameter Identification $\alpha = 0$ 93
5.16 EKF: Classical Optimal Control Problem $\alpha = 1$ 94
5.17 EKF: Trade-off between Identification and Control Performance $\alpha = 0.1$ 96
5.18 UKF: Parameter Identification $\alpha = 0$ 98
5.19 UKF: Classical Optimal Control Problem $\alpha = 1$ 99
5.20 UKF: Trade-off between Identification and Control Performance $\alpha = 0.1$ 101
5.21 GSF: Parameter Identification $\alpha = 0$ 102
5.22 GSF: Classical Optimal Control Problem $\alpha = 1$ 104
5.23 GSF: Trade-off between Identification and Control Performance $\alpha = 0.1$ 105
5.24 Case $\alpha = 0$: Error in Estimation of the State 106
5.25 Case $\alpha = 1$: Error in Estimation of the State 107
5.26 Case $\alpha = 1$: Error in Reference Tracking 108

A.1 Kalman Filter for Linear Systems (one cycle) 115

LIST OF TABLES

5.1 List of Parameters based on [1] 78
5.2 List of Parameters based on [2] 87
5.3 Mean Square Error: EKF 108
5.4 Mean Square Error: UKF 109
5.5 Mean Square Error: GSF 109

LIST OF ACRONYMS

API	Active Parameter Identification
OID	Optimal Input Design
KF	Kalman Filter
EKF	Extended Kalman Filter
UKF	Unscented Kalman Filter
MPC	Model Predictive Control
NMPC	Nonlinear Model Predictive Control
SISO	Single-Input Single-Output
MIMO	Multi-Input Multi-Output
MLE	Maximum Likelihood Estimation
PEM	Prediction Error Method
LSE	Least Square Estimator
PF	Particle Filter
GSF	Gaussian Sum Filter
PDF/pdf	Probability Density Function
LS	Least Square
MMSE	Minimum Mean Square Error
ML	Maximum Likelihood
RLS	Recursive Least Square
MSE	Mean Square Error
LMS	Least Mean Square
UT	Unscented Transformation
LMP	Least Mean p-Power

Chapter 1

INTRODUCTION

"To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of."

– Ronald Fisher

Most of the real life physical problems are characterized by their models in order to have better understanding and visualization. Models are divided in different types and classes, depending on the aim and field of study. The models are meant to represent some physical process or phenomena objectively. Nevertheless, they are just an approximation (or simplification) of some real physical system which they are intended to represent. A good model has all the useful information about the system and abstracts away the amount of information which is of little or no importance to the studied phenomena. The focus of this work is to develop such an algorithm that can maximize the amount of information from the unknown or uncertain system and achieve some desired control performance while respecting the input and state constraints.

The problem of optimal input design (OID) is to generate the external excitation signals in such a way that it can yield maximum information from the unknown system through the identification strategy. The nature of the exogenous input signal determines the quality and type of the model obtained from the system identification experiment. These excitations are usually manipulated by means of model builder which helps to acquire as much information as possible. In this thesis, a combined framework of OID for active parameter identification of nonlinear dynamic system is addressed. The objective is to extract maximal information about the unknown or uncertain parameters of the system. The proposed formulation allows the treatment of several important problems such as active identification of the parameter, classical optimal control problem and a trade-off between the two cases by acting on the parameter of interest. The formulation presented in this dissertation is based on the theoretical concepts of *System Identification and Estimation Theory, Optimal Control Theory, Information Theory and Optimal Experiment Design*. Hence, in this first chapter, we have discussed a brief introduction of these topics.

The topic of OID for experiment design is a widely researched topic among the control and estimation community. However, the topic of OID for active parameter identification, addressed in this work, has not been yet faced directly. Some similar concepts are discussed in literature like optimal experiment design, where the objective is to obtain the maximum information from the system. These problems arise in Chemistry, Biology, Physics, etc., where complex and expensive experiments are required. The related theory serves as a source of inspiration but is not directly related to our proposed field of study. In fact, the proposed theory is for the unknown/uncertain dynamic systems while the optimal experiment design deals with static systems.

After the work of Shannon [3], the field of Information Theory has grown considerably. The work done by Shannon was motivated by the problem he faced in Communication Theory but the field of Information Theory cannot be treated as a subset of Communication Theory. The field has contributed to Thermodynamics (Statistical Physics), Kolmogorov complexity and algorithmic complexity (Computer Science) and Occam's Razor (Statistical Inference). It is interesting to link the concepts of Information Theory with the ideas of Control Theory. Some interesting results are presented in [4, 5] but yet, the link between the two fields is not well established.

In this dissertation, we have proposed a new formulation for the design of optimal excitation signals for active parameter identification of nonlinear dynamic systems. By active parameter identification, we mean that the problem is finding such an optimal feedback control law that can maximize the amount of information on the unknown system parameters while the system evolves using available information online. The problem is formulated as a stochastic optimal control problem in Model Predictive Control (MPC) framework, where a suitable measure of uncertainty is added as the information cost on the unknown parameters. The introduction of the proposed work and problem formulation is given in Chapter 1 and a through discussion is carried in subsequent chapters.

1.1 Motivating Examples

In order to introduce the reader with the key concepts of the subject and to discuss the wide class of problems that can be faced with proposed framework, we have discussed some of those examples here.

- **Map Building of Unknown Environment:** Consider an autonomous vehicle that must explore an unknown or partially known environment. In order to build a map of the unknown building, it is obvious for the autonomous vehicle to explore the unknown environment. To build an effective map of the environment, it is necessary to have a feedback control law that will help the vehicle to maneuver and perform the necessary tasks of exploration.

- **Telecommunication Network Exploration:** The problem of exploring a telecommunication network where some nodes or links are broken, could be an interesting problem to investigate. The proposed algorithm can build a map of the network by means of "intelligent" tokens and communicate about the broken links.

- **Estimation of Target Position:** The problem of estimating a target position from noisy sensor measurements which is mounted on a board (moving observer) could be an interesting application for proposed framework. This example is a typical case where the control technique affects the observability of the plant. Due to the presence of nonlinearity, the estimation problem makes the observer maneuvers fundamental in order to have a perfect estimate of the target position.

- **Parameter Identification in Robotic Applications:** In order to identify the parameters of interest in some robotic applications like wheeled robot or robotic arm, it is not desirable to generate random excitation signals (motor torques). This may lead to poor controllability and stability of the system. Definitely, the choice of interest is to generate such control or excitation signals that can maximize the information from the system and respect the physical constraints on the system.

- **Bioengineering and Systems Biology:** In last few years, the use of control strategies for the identification of different diseases and health issues has seen an exponential increase. The use of mechanistic models in systems biology is a well researched topic where the identification and control of different system models consisting of biochemical pathways of interest in oncology are studied. However, due to unavailability of data and state of the art methodologies, the understanding of the intrinsic characteristics of complex pathologies like cancer is limited. Information retrieval in system biology related experiments is still a major area of interest for the researchers.

The fundamental objective of all the examples discussed above is to maximize the information about the unknown parameters from the system. Although every single example has its own particularity, yet a need of a general framework to deal with these broad class of problems is required. In this work, we have proposed such a framework which will use a suitable measure of uncertainty in the cost to be minimized and maximize the amount of information while respecting the physical constraints and bounds on the system.

1.2 Problem Formulation

It is difficult for a defined model to exactly capture all the aspects of a physical system. Also, it is not desirable to capture all the information/aspects of the system as it may lead to a highly complex models. The desire is to capture only the relevant information which depends on the intended use of the model application.

In this thesis, a novel formulation is presented for OID for nonlinear dynamic systems. In particular, the problem of active parameter identification is addressed in a receding horizon framework using some uncertainty measure related to the identified parameters as an information cost. The quality of the obtained information highly depends on the type of the identification experiment and the excitation properties of the applied input signal. For the system identification of nonlinear dynamic systems, Extended Kalman Filter (EKF) has been used which provides the information about the unknown states, parameters and the covariance matrix related to the unknown parameters. A-optimality criterion is defined as the information measure on the unknown parameter vector, which tend to minimize the *trace* of the covariance matrix. The excitation signals are generated using the MPC strategy, where a cost function is defined consisting of a process cost and information cost on the parameters. The proposed framework has the following objectives:

- The cost related to the identification of the parameter should be as small as possible which results in better identification of the parameter.
- The performance of the identified model should be very good.
- The constraints and bounds on the physical plant should be respected.

The general cost function for the optimization problem is given as:

$$\begin{aligned} \underset{input}{\text{minimize}} \quad & cost\ of\ experiment \\ subject\ to \quad & Performance\ specifications \\ & System\ constraints \end{aligned} \tag{1.1}$$

The cost function defined above is for general optimization problem solved in OID framework.

1.3 Thesis Structure

This section gives an overview and outline of the different chapters of the thesis. The chapters are given as follows:

- **Chapter 2:** In Chapter 2, a detailed description on the theoretical background of nonlinear system identification is presented. Different system identification and system estimation strategies are discussed. The motivation for the use of EKF as an identification strategy is presented. Some related estimation strategies like Unscented Kalman Filter (UKF), Particle Filters (PF) and Gaussian Sum Filter (GSF) are also discussed. A detailed background for the OID of nonlinear dynamic system is studied with the most relevant literature. The applications of OID in different fields like industrial process and system biology are presented. In the end, a detailed description of MPC framework for the OID is given.

- **Chapter 3:** Chapter 3 has focused on the fundamental of the OID for active parameter identification. It begins with a formal system description and then a detailed discussion is carried on the choice of the optimality criteria for the information cost. Different optimality criteria are discussed and a motivation for the use of A-optimality criterion (information-based criterion) is presented. To identify the information, a detailed description is given on EKF, UKF and GSF which are used to carry a comparison study on the identification of unknown parameter vector. In the end, the use of MPC for the active parameter identification and its problem formulation in MPC framework is presented.

- **Chapter 4:** In Chapter 4, the complete framework of OID for active parameter identification is presented. First, a complete description of the nonlinear discrete-time dynamic system is presented where the initial preliminaries and problem formulation is discussed. A detailed description on the use of system identification strategies (EKF, UKF or GSF), discussed in previous chapters is provided. In order to get the best estimates, A-optimality criterion is used and its detailed description is presented. In the end, the combined EKF/NMPC strategy is presented in the receding horizon framework.

- **Chapter 5:** In order to validate the proposed framework, a thorough study is carried with some abstract and realistic numerical examples. To get a deeper insight into the proposed formulation, it is important to implement it on some examples which can give clear idea on the effectiveness of the proposed work. A simple "toy model" is used as an abstract example which is considered as a stable system. Simulations are performed for different scenarios with different initial conditions and the results show the superiority of the proposed framework. In order to see the effectiveness of the proposed framework on a more realistic example, it is implemented on a 2-DOF and 3-DOF model of two-wheeled mobile robot model. Simulations are performed for different scenarios and a comparison study is carried out on the basis of the identified information on the unknown parameters. The simulation results shows that the proposed strategy performs exceptionally well in all cases and gives perfect identification of the parameter while respecting the physical constraints on the system.

- **Chapter 6:** In the last chapter, the thesis is concluded with the summary of major proposed contributions and discussion of the obtained results. Some recommendations and suggestions for the further future research are also presented.

Chapter 2

STATE OF THE ART

"Science may be described as the art of systematic over-simplification — the art of discerning what we may with advantage omit."

– Karl Popper

For the system identification of real physical systems, it is excited by a suitable excitation signal and observing the input and output behavior of the system. The choice of input signal highly influences the quality and accuracy of the information obtained on the unknown system. In recent decades, extensive work has been done to design such optimal input signals that can yield maximal information on the system. The accuracy of the information obtained from the system identification experiment is usually poor (due to incomplete knowledge of the system, external disturbances and noisy measurements, etc,). Hence, a suitable input spectrum has to be designed that can shape the quality of the information and improve the system performance. A well-designed input signal has the ability to maximize the amount of information on the unknown parameters of the system and also reduced the cost and time of the experiment. The input signals are designed on the a-priori knowledge of the system so it can "excite" the system well. In this chapter, we have presented a state of the art literature related to OID framework for system identification of nonlinear dynamic systems. Some useful literature is presented for active parameter identification and how to formulate the problem in a receding horizon framework.

This chapter briefly summarizes the basic concepts, preliminaries of nonlinear system identification using the OID and some useful reading suggestions on the topic from the literature which will be helpful for the material developed in the succeeding chapters of the thesis. Among good references for OID framework for system identification are [6–8].

2.1 Nonlinear System Identification

To understand the behavior of systems (either natural or man-made), modern science and technology is highly dependent on the mathematical models. A mathematical model can be roughly defined as a mathematical law that links the causes (system inputs) to the effects (system outputs). It is necessary to perfectly model a system mathematically and its application range from simulation and prediction to control and diagnosis of heterogeneous fields. To build a mathematical model of uncertain or unknown phenomena, system identification is a widely used approach. It is used to identify the model based on the observed uncertain or noisy measurements from the unknown system.

The idea of system identification is defined explicitly by many researchers. In [9], it is defined as: "system identification is the determination on the basis of observations of input and output of a system within a specified class of systems to which the system under test is equivalent". Due to presence of noise and uncertainty, it is highly impossible to identify a model that matches the actual physical plant. Hence, only an approximation of the practical plant can be obtained from the identification. The description given in [10] explained that the system identification tries to built a model that can describe the essential characteristics of a unknown system and the resulting model can be expressed in a useful form. The definition given by Ljung [11] is rather interesting one: "The identification procedure is based on three entities: the data, the set of models, and the criterion.

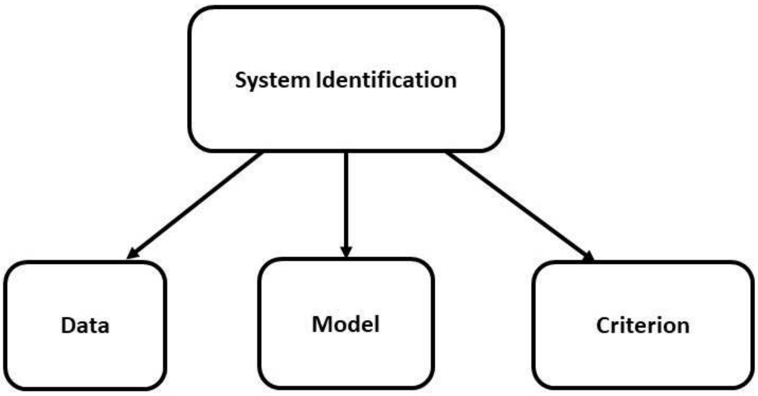


FIGURE 2.1: Elements of System Identification

Identification, then, is to select the model in the model set that describes the data best, according to the criterion."

The definition of Ljung has divided the system identification procedure into three main parts: data, model and criterion, see Fig. 2.1. The performance of the identification algorithm, its accuracy to identify the unknown parameters, robustness, convergence rate and computational complexity is directly dependent on these three elements [12]. The choice and design of these three elements is of critical importance in the system identification.

The general definition of system identification is given as:

Definition 2.1. System identification describes a way to construct the mathematical model of a uncertain dynamic system from the observed data obtained from inputs and outputs of the system. The input signal is designed in a such a way that it can maximzie the amount of the information on the unknown system parameters.

It is a well-studied field which provides the necessary tools to construct the mathematical models of the unknown or uncertain systems which are helpful to describe the behavior of the dynamic system. The different applications of system identification can be found in numerous engineering fields. In general, the primary task of any system identification method is to choose a model class which is parametrized by unknown or uncertain parameter vector. The objective is then to identify the unknown parameter vector as efficiently as possible so that the selected model can describe the true system dynamics.

The choice of a suitable model is a crucial step in system identification which gives a trade-off between complexity and the quality of the model. During the past few decades, numerous model structures were discussed for both linear and nonlinear systems. For linear systems, finite impulse response (FIR), auto-regressive (AR) and auto-regressive moving average (ARMA) models are very commonly used. While for nonlinear systems, nonlinear auto-regressive, radial basis function (RBF) and multilayer perceptron (MLP) are widely used model types. It is possible to have some prior knowledge of the model structure or may be the the model structure is confined to a particular tractable structure to have a good approximation of the system. Different model selection criteria are discussed in literature such as minimum description length (MDL) [13, 14], cross-validation criterion (CVC) [15], Bayesian information criterion (BIC) [16] and Akaike's information criterion (AIC) [17].

The choice of data selection in terms of measured variables and the design of optimal input for the system identification is of critical interest. For an experiment design, the key objective is to adjust the experimental conditions in such a way that maximal information can be obtained from the unknown system. In literature, different information matrices are used as an optimality criterion. The choice of input signal can significantly improve the quality of the identified information [18].

The third element to discuss is the equivalent criterion which has also major importance in system identification. The criterion helps to improve the quality of the information by measuring the similarity (or difference) between the identified model and the actual system. There are variety of optimality criteria which are used for system identification and a detailed discussion is carried out in Chapter 3. The choice of different criteria may lead to different type of estimates. The aim of OID is to find such a control law that can minimize (or maximize) the approximation criterion in order to maximize (or minimize) the information on the system. The optimality criterion serves as the measure of the accuracy and has a significant influence on the convergence behavior and optimal solution of the system. The next section will elaborate a detailed history of system identification.

2.1.1 History

The earliest work on system identification is reported back in the end of 18^{th} and beginning of 19^{th} century by statistics and time series communities. In [19], a break through work was proposed in which the method of least square is described. In the start of 20^{th} century, the development of statistical theory of regression and correlation analysis was also a major milestone achieved [20]. Some very good references for earliest history of system identification and time series analysis can be found in [21–23].

During the early 20^{th} century, two major developments took place: the theory of stationary process was proposed which serves as the main model class for time series and the first systematic approach was presented for modern system identification as Cowles Commission Econometrics. During almost the same time period, the ergodic theory for strict stationary process was proposed in [24]. The asymptotic characteristics of ordinary least square estimators were derived in [25] while in [26], stochastic behavior of macro models were presented. In [27], for the linear ARX model, a complete theory of system identification is presented. The identifiably analysis along with the asymptotic properties and Gaussian maximum likelihood estimation was also discussed. During the mid of 19^{th} century, different strategies were proposed for non-parametric estimation of spectral densities. The first smoothed spectral estimator was proposed in [28]. After the development of non-parametric spectral estimation strategies, the estimation strategies for AR and ARMA models gained lot of attention.

The first signs of system identification in the field of engineering was found in early 1960's. The work of R. E. Kalman on realization and parametrization [29, 30] has given birth to model based control era which later proved to be the foundation of pole placement and LQG control. These techniques were only applicable to the systems where the information on the model is perfectly known which is not true in all cases. The problem is addressed by the two papers which have given birth to subspace identification method [31] and prediction error identification [32]. The first correct proof of consistency of maximum likelihood estimation (MLE) for single-input single-output (SISO) system was presented in [33] which leads to asymptotic normality of the MLE's in different other papers [34, 35]. The concepts of system identification for SISO as well as MIMO systems reached a level of maturity by the end of 1980's. Several books were written on the mainstream identification of linear systems [7, 36–38] and presented the idea of system identification as a design problem. A well described work on the topic of frequency domain techniques is presented in [39] while the subspace identification strategies were covered in [40].

2.1.1.1 Identification Strategies: Overview

Traditionally, for system identification or state estimation, least square (LS) methods [41], minimum mean square error (MMSE) method [42] and the maximum likelihood (ML) methods [43, 44] are commonly used. It was first introduced by Carl Friedrich Gauss in 1795 and defined as a criterion which tends to minimize the sum of square errors. The error is defined as the difference between the actual value and the observed value of the model which corresponds to maximum likelihood criterion if the error has an Gaussian distribution. Due to its straight forward implementation steps, mathematical tractability and efficiency in terms of identification, the LS method is widely used strategy to solve the problems of estimation, identification and regression. The LS methods gives a closed form solution to a linear problem. In [45], a regularized version of LS method has been introduced. Different identification criteria, i.e. recursive least square (RLS) and its types [7], are developed on the basis of LS method. It is also common in signal processing and statistical community to use MMSE method as a measure of quality of estimation which tends to minimize the mean square error (MSE) of the observed value. It is used as a stochastic approximation method in system identification. These methods are used to find the extrema of a those functions, of which it is not possible to compute it directly, by noisy observations. A variant of MMSE method is famous least mean square (LMS) method which is based on the gradient decent algorithm [46–48]. In case of ML methods which provides a unified criterion for estimation, the model parameters are selected in such way a that can maximize the likelihood function. The method possess the characteristics as asymptotic normality, consistency and efficiency in terms of parameter identification.

The LS, MMSE and ML criteria work really well in most of the real life problems and still playing a fundamental part in system identification. However, the applicability of these criteria have some limitations. For example, only the second-order modes of the data are captured by LS and MMSE, which might result in poor approximation in case of nonlinear and non-Gaussian distribution. Similarly, a-prior knowledge of conditional distribution is required by ML criterion, which is not always available in practical problems. Also, in some complicated nonlinear problems, ML methods are not suitable to use. Hence, a need of criterion, which can identify beyond the second-order modes is a attractive problem in system identification community.

To address the above mentioned problems and discuss the optimal criterion for identification, non-MMSE criteria were introduced in [49]. The article showed that the non-MMSE criterion produces the same prediction results as produced by linear MMSE. The work was extended and several related articles were published [50–52]. For recursive parameter identification, the idea of general error criterion was proposed in [53]. An optimal criterion was proposed which tends to minimize the error covariance matrix of the parameter estimate. Another approach to select the optimality criteria from least-mean fourth family was proposed in [54, 55], where a cost related to the moments of the interfering noise was minimized. The use of calculus of variations to determine the optimal criterion among a large number of general optimality criteria was proposed in [56]. A method to optimize the derivative of the error criterion by optimizing the performance in steady state was proposed in [57]. In [58], least mean p-power (LMP) method was proposed. Several other non-MSE criteria discussed in literature are: M-estimation method [59], risk-sensitive method [60, 61], mixed norm method [62–64] and high-order cumulant method [65–67].

2.1.1.2 Estimation Strategies: Overview

Most of the real life physical systems have been represented with some mathematical models. These mathematical models, categorized in the two groups: deterministic and stochastic. The models are useful to effectively understand the past behavior of the system and predict the future behavior up to some extent. It is easy to represent and work with deterministic models. The shortcoming of this kind of system is that it does not provide enough information which gives rise to the use of stochastic models. This can be stated as:

- Mathematical representation of a physical system is never perfect. The models shows only the dominant modes of the physical systems.
- Due to the presence of uncertainty and approximation of the plant parameters, the accuracy of the model is highly effected.
- It is not possible to deterministically model the effects of exogenous disturbances.
- The measurement noise is always present in the information provided by the sensors.

The earliest work of R. A. Fisher on the ML estimation method is the major building block of classical estimation theory. Different parameter estimation strategies have been discussed in literature which differ due to several assumptions made regarding the prior probability and the optimality criterion. For example, linear regression method and the least square method assume that the optimality criterion is a scalar quantity while the ML method assume the maximization of the probability density function. Another interesting and widely used estimation strategy is Kalman filter and its variants. It accommodates both discrete and continuous time systems, especially for linear systems, where KF gives the optimal solution. As this work focuses only on discrete time systems, we will only discuss the discrete form of KF.

2.1.1.3 Kalman Filter and it's Variants

For the parameter identification and state estimation of linear systems, KF is the optimal recursive algorithm. The ease of implementation due to the recursive nature of the algorithm and the applicability in high dimensional state spaces are the key benefits of implementing the KF. It is most widely used method for state estimation in Control Theory as it produces the optimal estimate of the unknown or uncertain system in a sense that the sum of the estimation error is minimized. The application of KF to different physical systems is addressed in [68–72].

The nonlinear version of KF is known as extended Kalman filter in Estimation Theory. The filter tries to linearize about the current value of the estimated mean and covariance. In control theory, EKF has been used as a sub-optimal state estimator for uncertain nonlinear systems [73–76]. Its application for the parameter identification was first proposed in [77], where the unknown parameter vector θ is treated in the state vector with other states.

For the systems with highly nonlinear dynamics, EKF does not perform up to the mark because a linearized model is used to propagate the covariance matrix of the system. In the case of complex nonlinear systems, it involves costly computation of the Jacobin matrices which leads to slow convergence and implementation difficulties. Also, if the sampling time is not sufficiently small, this linearization leads to filter instability. In order to address the limitations possessed by EKF in terms of linearization, unscented transformation (UT) is used to estimate the mean and covariance matrix instead of linearization by the Jacobian matrices. The UKF addresses the assumption that it is easy to estimate a Gaussian distribution rather than to approximate an arbitrary complex nonlinear system. In Control community, UKF emerges as a strong and powerful nonlinear estimation method and proved its superiority over the EKF in many applications [78–80].

Lot of work has been carried out on the application of Kalman filters [81, 82]. In [81], a predictive likelihood approach has been used for estimating the noise filters for linear, EKF and UKF. In [82], an UKF is demonstrated. The UKF is capable of reconstructing the dynamics and estimating the unknown parameters of a neural mass model. A closed loop strategy has been demonstrated for modeling the dynamics of the model. In [83], a dual extended Kalman filter (DEKF) technique has been used for model based vehicle estimation. Two EKF are used in parallel to estimate the state and the parameters separately. Some other problems using EKF are presented in [84–87].

2.1.1.4 Particle Filter

For the case of linear systems with Gaussian noise, the KF is the optimal solution for estimation. For nonlinear systems with Gaussian noise, KF can gives you good results but the PF may give you better results. The main advantage of using particle filter instead of KF is that for a higher dimensional system, these PF are tractable while the KF is not. The other reason is that KF try to make the problem tractable by solving a simpler model instead of a complex model and find an exact solution by solving the simpler system. The problem with this kind of solution is that it might be still computationally expensive to solve it or the simplified model is not good enough to find an exact estimate. This problem is overcome by the use of PF which uses the full complex model to find an approximate solution of the system. The comparison of these two methods are carried out in various papers [88–94] and the references therein. The principal benefit of particle filtering is that they do not rely on any local linearization techniques or any crude functional approximation. The price that must be paid for this behavior is computational: these methods are computationally expensive.

Particle filter has been widely used in many real life applications. In [95], a comparison study has been carried out between EKF and PF for the state estimation of an industrial robot. Online state estimation is also a key element in process engineering. In [96], a PF based on sequential Monte-Carlo method has been proposed for the process engineering. Some other research papers has also been published on the state estimation of a dynamical system like [97, 98].

Particle filters have been effectively used for the estimation of static parameters [99]. An interesting application of particle filter for the parameter estimation in geophysics has been proposed in [100]. The paper describes the parameter identification of a pressure regulator with a nonlinear structure by sequential Bayes estimation in the framework of data assimilation. A damping coefficient of feedback system in the pressure regulator that cannot be observed directly is estimated using a particle filter and a nonlinear state space model. In process engineering, online state and parameter estimation is a key component in the modeling of batch processes. A kernel smoothing approach using the particle filter algorithm has been introduced for the robust estimation of unknown and time-varying model parameters [101]. Some interesting references on parameter estimation using particle filters are [102, 103].

2.1.1.5 Gaussian Sum Filter

In order to address various factors which effects the performance of EKF and PF, such as, conditional pdf of the system is non-vanishing, stationarity of the problem, the decay rate of the conditional pdf in state space, analytical structure of the problem and effective dimensionality of the problem, an efficient and simple approach for general nonlinear filtering is GSF. The approach is a Gaussian mixture approximation of state space pdf based on the Bayesian estimation [104–106].

The GSF tries to propagate the first two moments of the Gaussian component using the linearized model of a nonlinear dynamic system and the new weights are chosen as the prior weights which are updated accordingly using the Bayes rule. Variety of literature is discussed on GSF in theoretical aspect [107, 108]. The application of GSF is presented, e.g. for target tracking [109–111], geosciences [112], computer vision [113, 114].

2.2 Optimal Input Design in Nonlinear System Identification

Optimal input design for linear and nonlinear system identification has a long tradition in Statistical and Control community. In this section, a detail discussion on the design of optimal signals for nonlinear system identification is presented.

2.2.1 Background

The results on the input design in statistical literature trace back to the beginning of 20^{th} century. The input signal is designed in such a way that the estimated error in the identified parameters is minimized in the presence of some constraints. In [115], a good literature review is presented on the input design in statistical framework. The results presented about input design in statistics in used for System Theory [116, 117].

The Control community has recognized very early that accuracy of a model highly depends on the input signal [118–120]. In order to judge the performance of the control input, it seems logical to see the accuracy of the identified information [121]. In 1970, OID for dynamic system has started to attract the attention of the researchers [117, 122]. In 1986, the model purpose is explicitly incorporated in the OID framework by introducing a measure of performance degradation in the estimate of transfer function [123]. In [124, 125], the power of the excitation signal is used as the cost of the identification experiment which is more commonly used in application oriented input design. This idea was first originated in late 90's from the concept of plant-friendly input design for chemical process [126] which is somewhat related to the idea of designing the control signal from the identified information and achieving the desired control performance. Usually, the power of the probing signal or the magnitude of the perturbation signal is minimized as a cost for the identification experiment [124].

A common technique discussed in literature is to simultaneously identify and control the system which aims at increasing the control performance as the the knowledge on the unknown system is gradually improved. Such a method is used in [127] where the identification experiment and dual control strategy are used sequentially. In [128], a separate identification experiment is proposed in open loop in order to collect the data in the presence of already working control strategy. In [129, 130], the OID problem is presented as a receding horizon control problem in an MPC framework. In [131], a similar problem is discussed for the active identification of the unknown parameter in an information theoretic setting using the Shannon entropy as the information measure. An information-based multi-agent exploration framework is addressed in [132].

2.2.1.1 Open loop Optimal Input Design

The system identification for nonlinear systems highly depends on the relation of achieved accuracy of the identified model and the exogenous input signal used to excite it. This idea motivated many researchers in mid 70's to study the OID in open loop [117,133]. The research conducted in 1970's on OID for the identification of the systems, is mostly focused on the open loop systems [134–136]. The design variables were selected as either the sampling interval or the choice of the input signal. The optimization is performed in a statistical framework for both time-domain and frequency domain by considering the scalar functions of parameter covariance matrix as criteria for OID [137]. It is often assumed that the error is only due to the variance of the system and an open loop design are introduced based on some scalar criteria of information matrix X. The commonly used criteria are:

- **A-optimality criterion:** $\min \; trace(X^{-1})$.
- **D-optimality criterion:** $\max \; det(X)$.
- **E-optimality criteria:** $\max \; \lambda_{min}(X^{-1})$.

where, λ_{min} is the minimum eigenvalue of the information matrix. The time-domain solutions are defined as finite sequence of input data while the frequency domain solutions are based on the spectrum of designed optimal input [123]. There are two important results associated with it, i.e,

1. For the power constrained inputs, the set of average information matrices are convex in the spectrum of the input signal.

2. The information matrix X, can be obtained from input signals containing no more than $^{n(n+1)}/_{2} + 1$ sinusoidal components, where n is the number of parameter estimated.

For the model-based control design, usually some distances between the identified information and the nominal value of the system is minimized. It is noted that the closed loop performance is better and the sequence of control used during the identification process should match the desired model based controller [138]. Some recent developments in the field of OID have led to the *least costly identification* [139] which tries to minimize the cost subjected to some constraints. In [140], a more balanced approach is proposed between the least-costly identification and classical experiment design problems. Some other examples are presented in the H_∞ control framework [125, 141, 142]. In [143], an novel idea is presented which proposes that a good OID can yield only important properties of the process while the less important are neglected. The idea is extended to *application oriented input design* in [124, 144].

2.2.1.2 Closed loop Optimal Input Design

The major issue faced in open loop optimal control design is non-convex problems which lead to increasing computational complexities. Also, safety regulations or the cost of interrupting the normal operation, makes it difficult to perform it. It is not advisable to change the existing closed loop controller with an open loop controller in order to yield the maximum information. The problem is addressed by using the high-order expressions to achieve the desired model accuracy. Results obtained from high-order terms are quite accurate but there are cases where they failed to perform, see for example [145]. Also, the frequency-wise constraints are not handled perfectly which is an important aspect of robust control design.

The limitations associated with open loop OID are addressed by many researchers [125, 146–148] by introducing a new approach to solve the input design problem. These methods have shown that most of the open loop problems can be solved in a convex program. For the control applications, various arguments are presented in favor of identification in closed loop. For example, in [123], it is shown that for a high order variance situation, the experiment design under closed loop with minimum variance control are optimal. A similar problem is addressed in [135] where it is claimed that the closed loop input signal can outperform any fixed input design signal provided the experiment horizon is long enough. In [136], the use of high-order variance terms showed that the closed loop input design are optimal provided the variance constrained are fulfill. In [149, 150], it is also shown that closed loop performance is better than open loop for bias error.

The idea of closed loop OID needs to incorporate the controller information in the experiment design. The usefulness of the closed loop optimal design has been shown in [123, 135, 151, 152]. In [123], it is considered that the model is used to design the controller which is responsible to minimize the variance. Hence, the optimal controller is the minimum variance controller in closed loop setting. The problem associated with this design framework is that the controller depends on the complete knowledge of the system to be identified which is not true in all practical scenarios. In [153], it is shown that the optimal control parameters depends on the model structure by solving the input design problem in a minimum variance framework.

2.2.1.3 Sequential Input Design

The major problem arises during an OID is that the solution typically depends on the true knowledge of the model itself which is unknown for most of the physical problems. In [154], an optimal robust input design procedure has been proposed to handle the initial parametric uncertainties for a time-continuous system with one unknown parameter. The dependence on the system's unknown parameters for optimal control design is addressed by sequential procedure which is shown in Fig. 2.2. The excitation signal is improved online as the information on the system becomes available.

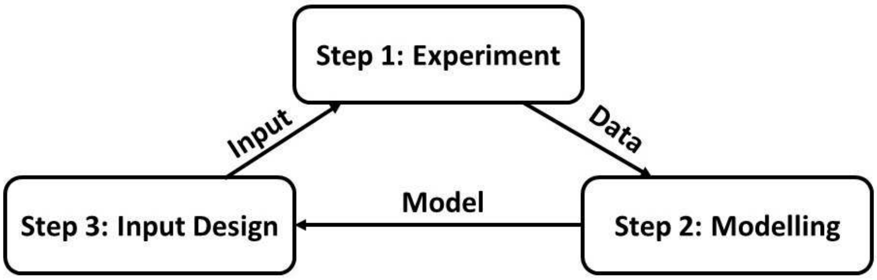


Figure 2.2: Sequential Input Design

The procedure is performed in three steps which are given as:

- **Step 1:** An identification experiment is performed on the unknown system using some excitation signals.
- **Step 2:** The identified information is used to construct a model of the system.
- **Step 3:** An input signal is generated using the model obtained in the previous step. The procedure repeats itself as this information is used in Step 1 again.

Many researchers have shown the usefulness of sequential approaches to OID [135, 147, 155] for engineering applications. The use of sequential methods in Statistics [156] and system biology has been presented in [157, 158]. It is also used for dual control problems where the control has to achieve both identification and control performance [159]. The key objective is to control the system and then generate such excitation signal that can excite the system sufficiently.

2.2.2 Applications of Optimal Input Design

The idea of OID is widely researched topic in Statistics where the systems are often assumed with uncontrollable inputs and static characteristics. In this section, we will consider the application of OID for industrial process and system biology only.

2.2.2.1 Input Design for Industrial Process

It is always difficult to design optimal input signal for industrial process as the practical constraints are often conflicted with the theoretical requirements. It is discussed in Section 2.2.1.2 that a powerful input signal for a longer time period can yield more accurate information about the system. This corresponds to industrial process where the system dynamics are slow and disturbances are high. The cost of the experiment procedure is low due to short time and small signals. The objective of the identification step in industry is to identify the data from the system that must lead to the nominal system in finite time period and deviate as little as possible from the normal system behavior [160]. It is stated in [161] that modeling is the most time consuming and expensive part of the model based control design framework. Most of the industrial process are nonlinear in nature and it is a topic of concern to design optimal input signal in open loop or closed loop frameworks [162].

It is always desirable to make the OID experiment user friendly as most of the control engineers are not familiar with identification. In [162], it is stated that "*the optimal input design framework for identification must be simple and clear; guidelines must be given. As the information is available, the calculations should be carried out automatically without the intervention of user*". This idea corresponds to trade-off between the computational complexity associated with the OID and the accuracy of the obtained model.

In most of the user-friendly input design frameworks, it is always a trade-off between the demands of actual plant and theoretical concepts [163, 164]. The constraints are imposed on the input signal in-line with the industrial system which are usually time-domain constraints. The use of identification strategies has reduced the experimental time and cost and hence it is used widely in industrial applications [162, 165]. These methods can reduced 70% of the computation complexity and time of MPC framework.

2.2.2.2 Input Design for System Biology

The field of system biology addresses the concepts of modeling and analysis of intracellular process like cell networks, proteins and genes, which are helpful to understand a disease and to develop some medicine treatments for it. A common experimental technique used to see the gene activity, which is changed due to concentration of alcohol or some toxic substance, is microarrays. The measurement of this input (alcohol concentration) and output (gene activity) gives information about the connection of different genes.

The concept of microarrays is similar to the idea of OID in system biology but the process includes high number of input and outputs. For example, humans have approximately 27000 genes which are inter-connected. Assume that we are taking one input (alcohol concentration), there will be 27000 outputs which will be excited by the input signal and if we are interested in the excitation of only one gene, let suppose X, one can analyze it using microarrays but it is difficult to tell that the output change in gene is directly influenced by the input or due to interconnection of other genes.

The idea of input design in system biology is catching the eyes of researchers, see for example, [166–168]. It is difficult to quantify the modeling errors in system biology as the gene activity is nonlinear and the uncertainty is large in the output information. This leads to a poor quality of measurement due to errors and low signal to noise ratio.

2.2.3 Model Predictive Control for Optimal Input Design

The model-based controllers significantly depends on quality of the model. The model is usually identified by some system identification strategies which are discussed in Section 2.1. In this section, we will present some useful material related to the problem formulation of OID in MPC framework.

2.2.3.1 Model Predictive Control

Model-based designs are widely used in today's industrial applications in order to meet the industry cutting-edge methods for identifying the necessary nonlinear models. Model predictive control is an optimization based feedback control strategy for nonlinear dynamic systems. Due to its flexibility to handle multivariate processes, straight forward implementation steps, ability to incorporate the input and output constrains explicitly and generally applicable control design method, MPC is a widely used strategy in process industry [169–171]. The use of MPC in petrochemical industry is very common and most of the process is controlled by it [168]. Due to increasing speed of processors and explicit MPC (where control is calculated offline), it is used in faster processes as well.

The property to incorporate the constraints makes MPC a popular choice in different applications but this is considered as a limitation in OID. This is due to the fact that when input and output constraints are considered, there is no explicit solution to the MPC optimization problem [172]. There are different problems associated with the design of optimal input signal in MPC framework:

1. The problem of OID usually relies on the assumption that the true system parameters are perfectly known. However, this is not true for all physical problems. This problem is addressed in two ways in literature.
   (a) The designed input signal should be robust to parameter changes. This leads to a robust input design framework, see for example [18].
   (b) The problem can also be solved by using the adaptive strategies by initializing the problem with best available estimate of the parameters and then the output predictions are updated as more information is obtained [173].
2. The other problem associated with OID in MPC framework is the choice of the suitable application cost function. The choice of a good cost function is still an open problem to address. The cost should be selected in way that

it can relate the control performance with the identified information on the parameter. The optimal control signal should point in the direction of the parameter space for which the performance of the control is sensitive to parameter changes. These directions are excited more during the identification procedure in order to get maximum information. Thus, it is important to chose the cost function intelligently as the performance of the MPC depends on it.

For the reference tracking problem in MPC framework, the appropriate choice of cost function could be the difference of the obtained output trajectory resulted by estimated value of the parameter and the desired reference signal. This could be given as:

$$J_{cost}(\theta) = \frac{1}{N}\sum_{t=1}^{N} \left\| \hat{y}_k(\hat{\theta}) - r_{k_d} \right\|_2^2 + \|u_k\|^2 \tag{2.1}$$

where, $\hat{y}_k(\hat{\theta})$ is the output with estimated parameters, r_{k_d} is the desired reference signal and u_k is the control signal. This is the most commonly used cost function in literature [129, 174].

3. Another problem is the approximation of the application set. The difficulty arises due to the unavailability of explicit solution for the MPC to solve the input-output constraint problem. This leads to large number of time-consuming simulations and numerical approximations.

2.3 Chapter Summary

This chapter presented a detailed literature survey on OID for nonlinear system identification. Some system identification strategies were discussed for nonlinear systems and then use of estimation strategy is motivated for the parameter identification. The use of Kalman filter and its variants for the system identification was discussed with some other estimation strategies like particle filters and GSF. The

OID in both open loop and closed loop framework was presented with some applications to industrial process and system biology. The use of MPC was motivated for the OID and relevant literature was presented.

Chapter 3

FUNDAMENTALS OF OPTIMAL INPUT DESIGN

"Prediction is very difficult, especially about the future."

– Niels Bohr

One of the main contributions of this thesis is to introduce a quite general framework for OID for system identification of nonlinear dynamic systems. The aim of this chapter is to introduce and discuss the fundamentals of this framework. Furthermore, the system identification strategies, the optimality criteria and the MPC used in this work will be more thoroughly studied. For a historical background on these topics, we referred the reader's to see Chapter 2 and the references therein.

First we will give an overview of the main strategies used for identification and the optimality criteria used in the framework. This introduction will also give a flavor of what kind of optimality criteria can be used for OID problems in order to improve the quality of the identified information. Later, a complete description of MPC framework for OID will be presented.

3.1 Optimality Criteria for System Identification

3.1.1 Nonlinear Discrete-time System

Consider a general discrete-time stochastic nonlinear system of the form

$$x_{k+1} = f(x_k, u_k, \theta) + \xi_k \tag{3.1a}$$

$$y_k = h(x_k, \theta) + \eta_k \tag{3.1b}$$

where, $k = 0, 1, \ldots$ is the current sampling instance, $x_k \in \Re^{n_x}$, $y_k \in \Re^{n_y}$, $\theta \in \Re^{n_\theta}$ and $u_k \in \Re^{n_u}$ are the state vector, the measurement vector, the unknown parameter vector and the input vector, respectively. The measurement noise is given as $\eta_k \in \Re^{n_y}$ while the process noise is represented by $\xi_k \in \Re^{n_x}$. The initial condition x_0 of the system is unknown but it is assumed that some a-priori information is available about its mean and covariance matrix.

3.1.2 Optimality Criteria

In system identification, information theoretic methods are used by many researchers for the solution of related identification problems. The earliest work presented in [175] has shown a way to use information theory for general system identification problems. For the parameter identification, the inverse of the Fisher information matrix provides a lower bound (also known as the Cramer-Rao lower bound) on the variance of the estimator [176–178]. For parameter identification, the use of rate distortion function to find the performance limitations is a common method [179–181]. The use of information theoretic measures (divergence, entropy, mutual information) with classical identification methodologies (MSE) has been widely addressed by many researchers [182–184]. It is suggested by many researchers that entropy and divergence can be used as an information theoretic identification criterion and it can be helpful to improve the performance of the identification in more realistic problems.

In control theory, optimal design is referred to experiment design which are generated by the use of some optimality criterion. The optimality criterion is responsible for the quality of the design and shows how good a design is. There are several optimality criteria discussed in literature which are:

- Information-based criteria.
- Distance-based criteria.
- Compound design criteria.

3.1.3 Information-based Criteria

The information based criteria are referred to the one which depends on the information matrix for the design. The information matrix is defined as a matrix which is proportional to the inverse of the variance-covariance matrix for the least-square estimate of the linear parameter of the unknown model. The information-based criteria are further subdivided in the following categories depending upon the number of parameters used.

3.1.3.1 G-Optimality Criterion

In [185], the G-optimality criterion was first proposed for the optimal design of regression problems. This criterion is also stated as global criterion in [119] which aims as a response estimation criterion. It is defined as:

$$\min_{x_i, i=1,\ldots,n} \max_{x \in \chi} var(\hat{y}_x) \tag{3.2}$$

which corresponds to minimizing the maximum variance of any predicted state over the complete experiment horizon. Where, $\hat{y}_x$ is the predicted state and its variance is given as:

$$var\{\hat{y}_x\} = \sigma^2 f^\top (X^\top X)^{-1} f \tag{3.3}$$

where, $X^\top X$ is the information matrix for the design. Let define a probability measure ξ on χ, then the normalized generalization form corresponding to $var(\hat{y}_x)$ is given as:

$$
\begin{aligned}
d(x,\xi) &= f_x^\top M_\xi^{-1} f_x \\
&= n\frac{var(\hat{y}_x)}{\sigma^2}
\end{aligned}
$$

Hence, $\xi^\star$ will be G-optimal, if and only if

$$\min_{\xi}\max_{x\in\chi} d(x,\xi) = \max_{x\in\chi} d(x,\xi^\star)$$

The sufficient condition for $\xi^\star$ to be a G-optimal is

$$\max_{x\in\chi} d(x,\xi^\star) = p$$

where p represents the number of unknown parameters in the model. The design efficiency for G-optimal design is defined as:

$$G_\xi = \frac{p}{\max\limits_{x\in\chi}\ d(x,\xi)}$$

3.1.3.2 D-Optimality Criterion

The aim of D-optimality criterion is to emphasis on the quality of the identified information. It is the most important and popular design criterion among the optimal control community and was first proposed as a determinant criterion in [186]. The problem of D-optimality is addressed by many researchers [118, 137, 187, 188] in variety of examples. It is defined as:

$$\max_{x_i, i=1,\ldots,n} \left|X^{\intercal}X\right| = \min_{x_i, i=1,\ldots,n} \left|(X^{\intercal}X)^{-1}\right| \tag{3.4}$$

which is stated as maximizing the determinant of the information matrix or it is equivalent to minimizing the determinant of the inverse of information matrix. The efficiency of the D-optimal design ξ is given as:

$$D_{\xi} = \left\{ {}^{|M_{\xi}|} \Big/ {}_{\left|M_{\xi^{\star}_{D}}\right|} \right\}^{1/p}$$

where the term $\xi^{\star}_{D}$ is assumed to be D-optimal.

3.1.3.3 A-Optimality Criterion

The A-optimality criterion is most widely used criterion for OID. It was first introduced by [189] which uses the knowledge of Fisher information matrix. In [190], an algebraic approach for generalized linear systems was proposed. The criterion is defined as:

$$\min_{x_i, i=1,\ldots,n} \quad trace(X^{\intercal}X)^{-1} \tag{3.5}$$

which is minimizing the *trace* of the information matrix or equivalent to minimizing the average variance of the estimated value of the parameter. The efficiency of the design ξ is defined as:

$$A_{\xi} = {}^{trace\left[M^{-1}_{\xi^{\star}_{A}}\right]} \Big/ {}_{trace\left[M^{-1}_{\xi}\right]}$$

where, $\xi^{\star}_{A}$ is A-optimal.

3.1.3.4 E-Optimality Criterion

The E-optimality criterion was first proposed in [191]. Later, the computations of E-optimal polynomial regression design was introduced in [192] and to compute the

E-optimal criterion for a broad class of systems, a method was proposed in [193]. It is defined as:

$$\max\ \lambda_{\min}(X^{\top}X) = \min\ \lambda_{\max}(X^{\top}X)^{-1} \tag{3.6}$$

which aims at minimizing the maximum eigenvalue of the $(X^{\top}X)^{-1}$ or equivalently maximizing the minimum eigenvalue of the information matrix. The efficiency of the design ξ is given as:

$$E_{\xi} = \frac{\lambda_{min}(M_{\xi})}{\lambda_{min}(M_{\xi_E^{\star}})}$$

where, the term $\xi_E^{\star}$ is assumed as E-optimal.

3.1.3.5 I-Optimality Criterion

The I-optimality criterion or I_v-optimality criterion is the "integrated variance" criterion, which was first introduced in [137, 194]. It tries to minimize the normalized average variance and defined as:

$$I = \frac{n}{\sigma^2} \int_R var(\hat{y}_x) dx \tag{3.7}$$

where the term R represents the region of interest. The efficiency of the design ξ is given as:

$$I_{\xi} = \frac{trace\left[MM_{\xi_I^{\star}}^{-1}\right]}{trace\left[MM_{\xi}^{-1}\right]}$$

where, the term $\xi_I^{\star}$ is I-optimal. There are several other information-based criteria are discussed in literature such as: D_A-optimality criterion [195, 196], D_s-optimality criterion [120, 197, 198], E_A-optimality criterion [188], L-optimality criterion [137], C-optimality criterion [198–200].

3.1.4 Distance-based Criteria

The distance-based criteria are defined by the distance $d(x, A)$ from a point x in the Euclidean space $\Re^p$ of dimension p to a set $A \subset \Re^p$. The distance-based criterion are further subdivided in these categories.

- U-optimality criterion.
- S-optimality criterion

3.1.4.1 U-Optimality Criterion

The U-optimality criterion is defined as a combination of A-,D-, and E-optimality criterion which was first proposed in [201]. It is defined as the minimization of the sum of the distance from each candidate point to the design. It is defined as

$$\min \sum_{x \in C} d(x, D) \tag{3.8}$$

where, the term C and D are referred to the set of candidate points and set of design points respectively. It is also named as "uniform coverage" design due to its ability to uniformly covers the candidate points.

3.1.4.2 S-Optimality Criterion

The S-optimality criterion seeks to maximize the harmonic distance from each design point to all the other points in the desired design region. It was first introduced in [202] and defined as:

$$\frac{N_D}{\sum_{y \in D} {}^{1}/_{d(y, D-y)}} \tag{3.9}$$

where the term N_D is the number of points in the region D and D is the set of the design points. The distance $d(y, D - y)$ represents the spread of the points in the maximal area and it is also named as "maximum spread" design.

3.1.5 Compound-design Criteria

The compound-design criteria is defined as the maximization of the weighted product of efficiencies [203]. It is defined as:

Definition 3.1. Let $\psi_i(M_i(\xi)), i = 1, \ldots, n$ be defined as a set of n convex design criteria in a experimental region χ and $\alpha_i, i = 1, \ldots, n$ be the positive weights, then the compound design criterion is defined as:

$$\psi_\xi = \sum_{i=1}^{n} \alpha_i \psi_i \{M_i(\xi)\} \tag{3.10}$$

which is minimized by selecting the value of ξ.

There are two different types of compound design criteria discussed in literature:

- DT-optimality criterion
- CD-optimality criterion

3.1.5.1 DT-Optimality Criterion

The idea of DT-optimality criterion was introduced in [203] which is a combination of D- and T-optimality criterion. It gives the advantage of a balance between parameter identification and model discrimination. It is defined as:

$$\phi_\xi^{DT} = (1-k)\log\Delta_2(\xi) + ({}^{k}/_{p}\log|M_\xi|) \tag{3.11}$$

where the term ϕ_ξ^{DT} is the combination of T-optimality criterion ($\log\Delta_2(\xi)$) and the D-optimality criterion.

3.1.5.2 CD-Optimality Criterion

It was first introduced in [203] as a combination of parameter estimation and C-optimality. It seeks to determine the estimate of the parameter and minimize the area under the curve. It is given as:

$$\phi_{\xi}^{CD} = ({}^{k}/_{p}) \log |M_{\xi}| - (1-k) \log c^{\mathsf{T}} M_{x}^{-1} ic \tag{3.12}$$

The term $(1-k) \log c^{\mathsf{T}} M_{x}^{-1} ic$ is the C-optimality criterion. There are several other criteria are also discussed in literature, for example, T-optimality which was proposed in [204].

3.2 System Identification Strategies

The choice of a suitable model is a crucial step in system identification which gives a trade-off between complexity and the quality of the model. In this work, a system of the form (3.1) is used with different KF variants and GSF for the estimation and identification of states and parameters respectively. A detailed methodology of simple KF is given in Appendix A while EKF, UKF and GSF are presented in this section which will be used as system identification strategy in the proposed framework given in Fig. 4.1.

3.2.1 Extended Kalman Filter

The nonlinear version of KF is known as EKF in estimation theory. The filter tries to linearize about the current value of the estimated mean and covariance. In control theory, EKF has been used as a sub-optimal state estimator for uncertain nonlinear systems. In this work, we have treated the unknown parameter vector θ as an augmented state with the state vector. Consider a general discrete-time nonlinear system given as:

$$x_{k+1} = f_k(x_k, u_k, \theta) + \xi_k \tag{3.13}$$

$$y_k = h_k(x_k, \theta) + \eta_k \tag{3.14}$$

where, $k = 0, 1, \ldots$ is the sampling index, $x_k \in \Re^{n_x}, u_k \in \Re^{n_u}, \theta \in \Re^{n_\theta}$ and $y_k \in \Re^{n_y}$ are state vector, input vector, unknown parameter and the output vector respectively. The term ξ_k is the white-noise sequence with zero-mean and covariance Q_k^ξ and η_k is the white-noise sequence with zero-mean and covariance R_k. It is assumed that the initial condition x_o is obtained from a known density function $p(x_o)$. The two noise quantities ξ_k and η_k do not depend on the initial condition x_o and are assumed to be mutually independent.

Consider the nonlinear discrete time system given in (3.13), the prediction step of EKF strategy is given as

$$\begin{aligned}
\hat{x}_{k+1|k} &= f(\hat{x}_{k|k}, u_k, \hat{\theta}_{k|k}) \\
\hat{\theta}_{k+1|k} &= \hat{\theta}_{k|k} \\
P_{k+1|k} &= F_k P_{k|k} F_k^\mathsf{T} + Q_k
\end{aligned}$$

where, $\hat{x}_{k+1|k}$ is the predicted state vector, $\hat{\theta}_{k+1|k}$ is the predicted parameter vector and the prediction of covariance matrix is given as $P_{k+1|k}$. The term Q_k is a covariance matrix given as

$$Q_k = \begin{bmatrix} Q_k^\xi & 0 \\ 0 & Q_k^\theta \end{bmatrix}$$

where, the term Q_k^ξ is the part of the covariance matrix related to the system states while the term Q_k^θ accounts for the possible uncertainties in the evolution of identified parameter vector. Where the latter are considered constant over time, we put $Q_k^\theta = 0$. A term different from zero can be used to model slow changes in their values. The term F_k represents the Jacobian of the system equation given in (3.13) and represented as

$$F_k = \begin{bmatrix} \dfrac{\partial f(x, u_k, \theta)}{\partial x} & \dfrac{\partial f(x, u_k, \theta)}{\partial \theta} \\ 0 & I_{n_\theta} \end{bmatrix}_{\substack{x = \hat{x}_{k|k} \\ \theta = \hat{\theta}_{k|k}}}$$

The innovation step is given as

$$
\begin{aligned}
\hat{y}_{k+1|k} &= h(\hat{x}_{k+1|k}, \hat{\theta}_{k|k}) \\
v_{k+1} &= y_{k+1} - \hat{y}_{k+1|k} \\
S_{k+1} &= H_{k+1} P_{k+1|k} H_{k+1}^{\mathsf{T}} + R_{k+1} \\
W_{k+1} &= P_{k+1|k} H_{k+1}^{\mathsf{T}} S_{k+1}^{-1}
\end{aligned}
$$

where, H_{k+1} is the Jacobian of the measurement equation and given as

$$
H_{k+1} = \left[\begin{array}{cc} \frac{\partial h(x,\theta)}{\partial x} & \frac{\partial h(x,\theta)}{\partial \theta} \end{array} \right]_{\substack{x = \hat{x}_{k+1|k} \\ \theta = \hat{\theta}_{k+1|k}}}
$$

The updated state estimate and updated covariance matrix estimate is given as

$$
\left[\hat{x}_{k+1|k+1}^{\mathsf{T}}, \hat{\theta}_{k+1|k+1}^{\mathsf{T}}\right]^{\mathsf{T}} = \left[\hat{x}_{k+1|k}^{\mathsf{T}}, \hat{\theta}_{k|k}^{\mathsf{T}}\right]^{\mathsf{T}} + W_{k+1} v_{k+1} \tag{3.15}
$$

$$
P_{k+1|k+1} = P_{k+1|k} - W_{k+1} S_{k+1} W_{k+1}^{\mathsf{T}} \tag{3.16}
$$

where, $\hat{x}_{k+1|k+1}$ is the updated state estimate, $\hat{\theta}_{k+1|k+1}$ is the updated parameter estimate and $P_{k+1|k+1}$ is the updated covariance matrix of the system. The flow chart for one cycle of EKF estimation is shown in Fig. 3.1.

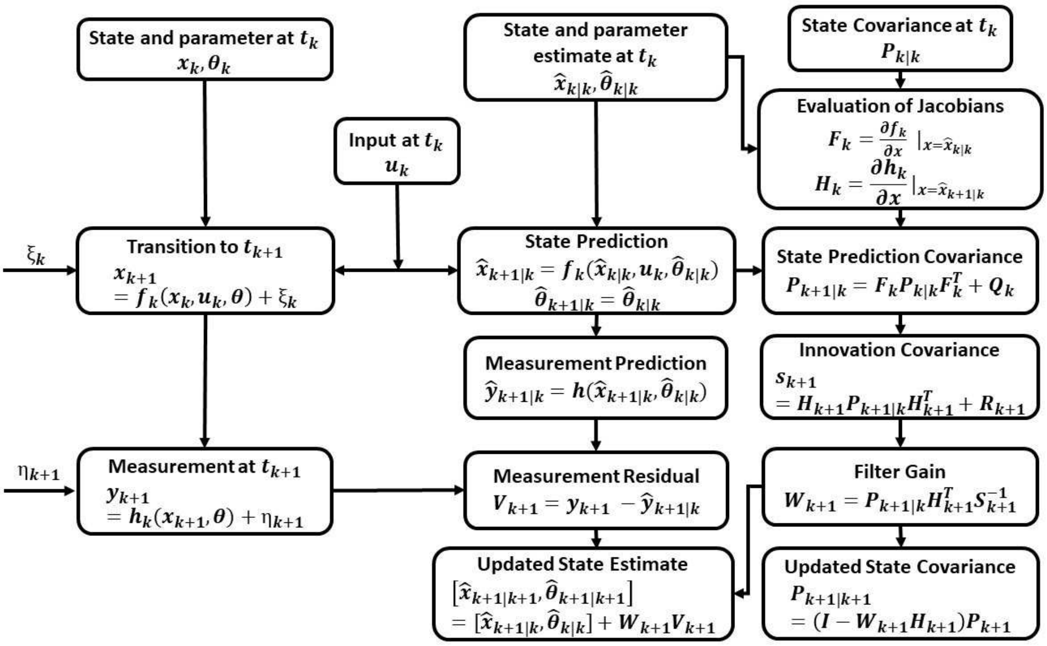


FIGURE 3.1: EKF Flow Chart (One Cycle)

3.2.2 Unscented Kalman Filter

For the systems with highly nonlinear dynamics, EKF does not perform up to the mark because a linearized model is used to propagate the covariance matrix of the system. In the case of complex nonlinear systems, it involves costly computation of the Jacobian matrices which leads to slow convergence and implementation difficulties. Also, if the sampling time is not sufficiently small, this linearization leads to filter instability. In order to address the limitations possessed by EKF in terms of linearization, unscented transformation (UT) is used to estimate the mean and covariance matrix instead of linearization by the Jacobian matrices. The UKF addresses the assumption that it is easy to estimate a Gaussian distribution rather than to approximate an arbitrary complex nonlinear system.

Consider the nonlinear dynamic system given in (3.1). To explain the UKF algorithm in detail, we assumed that unknown parameter vector θ is augmented with other states in the vector $z_k \triangleq \left[x_k^\mathsf{T} \;\; \theta^\mathsf{T}\right]^\mathsf{T}$. For the sake of simplicity, in this section we will use z_k to represent the state vector. With a little abuse of notation, we will use z_k as argument of the state and measurement function given in eq. (3.1) instead of the couple (x_k, θ). Let $L \triangleq n_x + n_\theta$ be the dimension of the vector z_k

and $\mathcal{Z}$ be the matrix of $2L+1$ sigma vectors $\mathcal{Z}_i$ (with corresponding weights ω_i) using the unscented transformation as

$$
\begin{aligned}
\mathcal{Z}_0 &= \hat{z}_{0|0} \\
\mathcal{Z}_j &= \hat{z}_{0|0} + (\sqrt{(L+\lambda)P})_j \qquad j = 1, ..., L \\
\mathcal{Z}_j &= \hat{z}_{0|0} - (\sqrt{(L+\lambda)P})_{j-L} \quad j = L+1, ..., 2L
\end{aligned}
$$

where, the weight of each sigma point $\mathcal{Z}_j$ can be calculated as:

$$
\begin{aligned}
\omega_0^m &= \frac{\lambda}{L+\lambda} \\
\omega_0^c &= \frac{\lambda}{L+\lambda} + (1 - \rho^2 + \upsilon) \\
\omega_j^m = \omega_j^c &= \frac{1}{2(L+\lambda)} \qquad j = 1, ..., 2L
\end{aligned}
$$

where, $\lambda = \rho^2(L+\kappa) - L$ is a scaling parameter having any arbitrary value except $\lambda \neq -n$. The value of ρ determines the spread of the sigma points around $\hat{x}_k$ while κ is secondary scaling parameter which is usually set to zero. The value of υ is set to incorporate prior knowledge of the distribution of z_k. At time $k = 0$, take $\mathcal{Z}_0^{k|k} = \hat{z}_{k|k}$. The time update equations of UKF algorithms for $k = 1, ..., N$ is given as:

$$
\begin{aligned}
\mathcal{Z}_j^{k|k} &= \left[\hat{z}_{k|k} \quad \hat{z}_{k|k} \pm \sqrt{(L+\lambda)P_{k|k}}\right]_j \\
\mathcal{Z}_j^{k+1|k} &= f\left[\mathcal{Z}_j^{k|k}, u_k\right] \\
\hat{z}_{k+1|k} &= \sum_{j=0}^{2L} \omega_j^m \mathcal{Z}_j^{k+1|k} \\
P_{k+1|k} &= \sum_{j=0}^{2L} \omega_j^c \left[\mathcal{Z}_j^{k+1|k} - \hat{z}_{k^-}\right] \left[\mathcal{Z}_j^{k+1|k} - \hat{z}_{k^-}\right]^{\mathsf{T}} \\
\Upsilon_j^{k+1|k} &= h\left[\mathcal{Z}_j^{k+1|k}\right] \\
\hat{y}_{k+1|k} &= \sum_{j=0}^{2L} \omega_j^m \Upsilon_j^{k+1|k}
\end{aligned}
$$

The measurement update equations are given as:

$$
\begin{aligned}
P_{y_{k+1}y_{k+1}} &= \sum_{j=0}^{2L} \omega_j^c \left[\Upsilon_j^{k+1|k} - \hat{y}_{k+1|k}\right] \left[\Upsilon_j^{k+1|k} - \hat{y}_{k+1|k}\right]^{\mathsf{T}} \\
P_{x_{k+1}y_{k+1}} &= \sum_{j=0}^{2L} \omega_j^c \left[\mathcal{Z}_j^{k+1|k} - \hat{z}_{k+1|k}\right] \left[\Upsilon_j^{k+1|k} - \hat{y}_{k+1|k}\right]^{\mathsf{T}} \\
K_{k+1} &= P_{x_{k+1}y_{k+1}} P_{y_{k+1}y_{k+1}}^{-1} \\
\hat{z}_{k+1|k+1} &= \hat{z}_{k+1|k} + K_{k+1}(y_{k+1} - \hat{y}_{k+1|k}) \\
P_{k+1|k+1} &= P_{k+1|k} - K_{k+1} P_{y_{k+1}y_{k+1}} K_{k+1}^{\mathsf{T}}
\end{aligned}
$$

where $\hat{z}_{k+1|k+1} \triangleq \left[\hat{x}_{k+1|k+1}^{\mathsf{T}}, \theta_{k+1|k+1}^{\mathsf{T}}\right]^{\mathsf{T}}$ is the updated state estimate and $P_{k+1|k+1}$ is the updated covariance matrix of the system. There are many other variants of KF which are studied in literature but those are out of scope of this thesis.

3.2.3 Gaussian Sum Filter

In more sophisticated nonlinear estimation schemes, an attempts is usually made to calculate the desired a posteriori probability density function (pdf) or at least the sufficient statistics of these functions. Though the approach used for this purpose is conceptually appealing, but a high level of complexity is associated with it in case when no approximation is used. In general, the computation of pdf requires a large storage of bits as for every value of the state, the corresponding value of the pdf must be stored. Also, the computation cost for such strategies is high, as in each iteration, an integration is required. Approximations are made to compensate for the storage and computation problems.

In previous section, a detailed discussion is carried on EKF algorithm which involves approximation of first and second order terms of the densities. The idea can be improved by using the higher order moments which focused on the approximation of the density near to mean value. In GSF, a Bayesian estimation algorithm is introduced by using the Gaussian sum approximation for the probability densities. The fundamental idea behind the Gaussian sum approximation is to approximate a density function as a weighted sum of several Gaussian densities. The covariance associated with these densities is usually assumed very small and its mean and variance is calculated by the EKF algorithm. The resulting filter is a bank of weighted EKF, where every EKF provides the evaluation of the assigned density function and the weights are computed from the EKF residuals. If the noise is considered small, the resulting estimator can give near optimal solutions.

3.2.3.1 Gaussian Sum Approximations

Gaussian sum filter is most frequently used state estimation strategy for the nonlinear systems. The filter evolves on the linearization of the system using the current value of the estimates and utilizing the EKF algorithm equations. This requires two types of approximation: first, the linear model is used instead of nonlinear model and second, the a posteriori density function is approximated by a Gaussian density function. The GSF has increased the validity of the approximation of the physical system to a great extant and eliminated the assumption of density function to be Gaussian.

Consider a general discrete-time nonlinear system given as:

$$x_{k+1} = f_k(x_k, u_k, \theta) + \xi_k \tag{3.17}$$

$$y_k = h_k(x_k, \theta) + \eta_k \tag{3.18}$$

where, $k = 0, 1, \ldots$ is the sampling index, $x_k \in \Re^{n_x}, u_k \in \Re^{n_u}, \theta \in \Re^{n_\theta}$ and $y_k \in \Re^{n_y}$ are state vector, input vector, unknown parameter and the output vector respectively. The term ξ_k is the white-noise sequence with zero-mean and covariance Q_k^ξ and η_k is the white-noise sequence with zero-mean and covariance R_k. In the probabilistic context of GSF, the a posteriori density function $p(x_k|Z^k)$[1] provides the most complete information on x_k, where $Z^k = [y_1, \ldots, y_k, u_1, \ldots, u_k]$.

Let $\mathcal{N}(x_k - \hat{x}_{i,k|k}, P_{i,k|k})$ is a normal Gaussian density function which is expressed as:

$$\mathcal{N}(x_k - \hat{x}_{i,k|k}, P_{i,k|k}) = (2\pi)^{-n/2} \left|P_{i,k|k}\right|^{-1/2} e^{-(1/2)(x_k - \hat{x}_{i,k|k})^\top P_{i,k|k}^{-1}(x_k - \hat{x}_{i,k|k})} \tag{3.19}$$

where, $\hat{x}_{i,k|k}$ is the mean vector and $P_{i,k|k}$ is the covariance matrix. The following Lemma states the approximation properties of GSF.

Lemma 3.2. *Any Probability density function $p(x_k)$ can be approximated as closely as desired in the space* [2] *$L_1(\Re^n)$ by a Gaussian sum representation of the form*

$$p_{\mathcal{A}}(x) = \sum_{i=1}^{m} \alpha_{i,k} \mathcal{N}(x_k - \hat{x}_{i,k|k}, P_{i,k|k}) \tag{3.20}$$

for some integer m, positive scalars $\alpha_{i,k|k}$ with $\sum_{i=1}^{m} \alpha_{i,k} = 1$, mean vector $\hat{x}_{i,k|k}$ and positive definite covariance matrices $P_{i,k|k}$ [106, 205].

The $p_{\mathcal{A}}(\cdot)$ is non-negative and integrates to 1 over the $\Re^n$. As the number of Gaussian m increases, the approximation error is minimized accordingly. For

[1] In GSF context, the superscript $()^k$ will denote the sequence up to and including k.

[2] The approximation is such that $\int_{\Re^n} |p(x) - p_{\mathcal{A}}(x)|\, dx$ can be made arbitrarily small.

the sake of simplicity, we are assuming that $X_k = [x_k, \theta]^\intercal$ represents the new augmented state vector. To help understand the Gaussian sum approximation algorithm, let assume that the $p(X_k|Z^k)$ is expressed as:

$$p(X_k|Z^k) \approx= \sum_{i=1}^{m} \alpha_{i,k} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) \tag{3.21}$$

Then, $X_{k|k} = \mathbb{E}(X_k|Z^k)$ and $P_{k|k} = Var(X_k|Z^k)$ are readily calculated by using these equations:

$$\hat{X}_{k+k} = \sum_{i=1}^{m} \alpha_{i,k} \hat{X}_{i,k|k} \tag{3.22}$$

$$P_{k|k} = \sum_{i=1}^{m} \alpha_{i,k} \Big\{ P_{i,k|k} + (\hat{X}_{k|k} - \hat{X}_{i,k|k})(\hat{X}_{k|k} - \hat{X}_{i,k|k})^\intercal \Big\} \tag{3.23}$$

For the proof, the readers are referred to see Appendix B.

Time-update Equation

The pdf $p(X_k|Z^{k-1})$ is given as a sum of Gaussian probability densities

$$\begin{aligned} p(X_k|Z^{k-1}) &= \int p(X_k|X_{k-1}) p(X_{k-1}|Z^{k-1}) dX_{k-1} \\ &\equiv \sum_{i=1}^{m} \alpha_{i,k-1} \mathcal{N}(X_k - \hat{X}_{i,k|k-1}, P_{i,k|k-1}) \end{aligned}$$

For each Gaussian density (value of i), the prediction equations are obtained by using the EKF equations as

$$\hat{X}_{i,k|k-1} = f(\hat{X}_{i,k-1|k-1}, u_k) \tag{3.24}$$

$$P_{i,k|k-1} = F_{i,k|k-1} P_{i,k-1|k-1} F_{i,k|k-1}^\intercal + Q_k \tag{3.25}$$

where, $\hat{X}_{i,k|k-1}$ is the predicted state vector for i^{th} density function and the prediction of the covariance matrix is given as $P_{i,k|k-1}$. The term Q_k is the noise covariance matrix. The matrix $F_{i,k|k-1}$ is computed as:

$$F_{i,k|k-1} = \left.\frac{\partial f(X_{k-1}, u_k)}{\partial X_{k-1}}\right|_{X_{k-1} = \hat{X}_{i,k-1|k-1}} \tag{3.26}$$

In the same way, the value of $\hat{X}_{k|k-1}$ and $P_{k|k-1}$ are computed as

$$\hat{X}_{k|k-1} = \sum_{i=1}^{m} \alpha_{i,k-1}\hat{X}_{i,k|k-1} \tag{3.27}$$

$$P_{k|k-1} = \sum_{i=1}^{m} \alpha_{i,k-1}\Big\{P_{i,k|k-1} + (\hat{X}_{k|k-1} - \hat{X}_{i,k|k-1})(\hat{X}_{k|k-1} - \hat{X}_{i,k|k-1})^{\mathsf{T}}\Big\} \tag{3.28}$$

Measurement-update Equation

The probability density function $p(X_k|Z^k)$ can be found from $p(X_k|Z^{k-1})$, when the new measurement y_k becomes available. The probability density function $p(X_k|Z^{k-1})$ is the weighted sum of Gaussian densities and similarly, $p(X_k|Z^k)$ can be calculated and expressed as

$$\begin{aligned} p(X_k|Z^k) &= \frac{p(y_k|X_k)p(X_k|Z^{k-1})}{\int p(y_k|X_k)p(x_k|Z^{k-1})dX_k} \\ &= \sum_{i=1}^{m} \alpha_{i,k}\mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) \end{aligned}$$

The updating equations are computed by using the extended Kalman filter equations as

$$\hat{X}_{i,k|k} = \hat{X}_{i,k|k-1} + W_{i,k}\left[y_k - h(\hat{X}_{i,k|k})\right] \tag{3.29}$$

$$P_{i,k|k} = P_{i,k|k-1} - W_{i,k}S_{i,k|k-1}W_{i,k}^{\mathsf{T}} \tag{3.30}$$

where,

$$W_{i,k} = P_{i,k|k-1} H_{i,k|k-1}^{\mathsf{T}} S_{i,k|k-1}^{-1} \tag{3.31}$$

$$S_{i,k|k-1} = H_{i,k|k-1} P_{i,k|k-1} H_{i,k|k-1}^{\mathsf{T}} + R_{k-1} \tag{3.32}$$

$$H_{i,k|k-1} = \frac{\partial h(X_k)}{\partial X_k}\bigg|_{X_k = \hat{X}_{i,k|k-1}} \tag{3.33}$$

As, $p(X_k|Z^k)$ is approximated by the linear combination of the densities with weight $\alpha_{i,k}$, which is calculated as

$$\alpha_{i,k} = \frac{\alpha_{i,k-1}\mathcal{N}(y_k - h(\hat{X}_{i,k|k-1}), S_{i,k|k-1})}{\sum_{j=1}^{m} \alpha_{j,k-1}\mathcal{N}(y_k - h(\hat{X}_{j,k|k-1}), S_{j,k|k-1})} \tag{3.34}$$

which motivates to the following result,

Theorem 3.3. *With the measurement available according to (3.1) and probability function $p(X_k|Z^{k-1})$, the updated probability density function $p(X_k|Z^k)$ is given as*

$$p(X_k|Z^k) = \sum_{i=1}^{m} \alpha_{i,k}\mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) \tag{3.35}$$

and converges to X_k and y_k (which refers to minimization of the approximation error) as $P_{i,k|k} \to 0$ for $i = 1, 2, \ldots, m$ [206].

Key Points for Filter Implementation

In the above discussion, a GSF algorithm is presented. For the updated mean and covariance, both time-update and measurement-update equations are taken from EKF algorithm. The overall algorithm requires the implementation of m separate EKF running in parallel and tuning the weight $\alpha_{i,k}$ as the new measurement is available. The updated mean-square-error estimate $\hat{X}_{k|k}$ is the weighted sum of all the estimates obtained from EKF algorithm in (3.29).

The above discussed GSF algorithm can be summarized as a recursive estimation strategy for updating a posteriori density which is initialed at time $k = 0$.

- Initialize the system with the following initial values $\hat{X}_{i,0|0}$ and $P_{i,0|0}$. The recursive algorithm presented in equations (3.24), (3.25) and (3.29)-(3.33) is used to compute $\hat{X}_{i,k|k}$ and $P_{i,k|k}$ for all the Gaussian densities $i = 1, \dots, m$ and $k = 1, \dots, .$
- For $i = 1, \dots, m$ and $k = 1, \dots, ,$ using the initial weight $\alpha_{i,0}$ and the measurement $y_{i,k|k-1}$, the updated weight $\alpha_{i,k}$ can be obtained.
- Apply the measurement-update equation as discussed in Theorem 3.3 to obtained the updated estimate of augmented state vector and covariance matrix.
- Update the time instance as $k + 1 = k$.

3.3 Nonlinear Model Predictive Control for Optimal Input Design

NMPC is an optimization based feedback control strategy for nonlinear dynamic systems. Due to its flexibility to handle multivariate processes, straight forward implementation steps, ability to incorporate the input and output constrains explicitly and generally applicable control design method, NMPC is a widely used strategy in process industry. At every sampling instant k, a numerical optimization problem is solved using the identified information. The first control of the optimized sequence is applied and the process is repeated at next time instant $k + 1$. Due to the shift of control horizon window on every sampling instant, the optimization problem is also named as *receding horizon* optimization problem.

3.3.1 Model Predictive Control Framework

At the core of any MPC implementation is a model of the process that is to be controlled. Typically, MPC uses a deterministic model for predicting the results of the control action and disturbances. Consider the discrete-time nonlinear dynamic system given in (3.1) which is used to estimate the unknown states and parameters by the EKF or UKF as discussed in Section 3.2 and the control is designed in MPC framework using this information. The *prediction horizon* is used to define the number of sample or iterations for which the identification is performed while *control horizon* represent the number of samples in the optimization horizon. The general quadratic cost function J_k used to define the controller in MPC framework is given as:

$$J_k = \sum_{k=0}^{N-1} g_k(x_k, u_k, \theta) + g_N(x_N) \tag{3.36}$$

where the term $g_k(x_k, u_k, \theta)$ represents the transition cost computed at time instance $k = 0, 1, \ldots, N-1$ and $g_N(x_N)$ is the final cost computed at time instance $k = N$. The term T in Fig. 3.2 represents the prediction horizon while the term N represents the control horizon. The optimization problem is solved at every sampling instance k as:

$$\bar{U}_k^\star = \arg\min_{\bar{U}_k} J_k(\bar{U}_k)$$

where the column vector $\bar{U}_k \in \Re^{mN}$ containing the full sequence of candidate control vectors considered at time k for $\ell = 0, 1, \ldots, N-1$ which is subjected to

$$\begin{aligned}
\hat{x}_{k+1} &= f(x_k, u_k, \theta) \;\; k = 0, 1, \ldots, N_y - 1 \\
\hat{y}_k &= h(x_k, \theta) \;\; k = 1, \ldots, N_y, \\
\hat{x}_0 &= x_k, \\
\hat{u}_0 &= u_{k-1}, \\
\hat{x}_k &\in \mathcal{X}, \hat{u}_k \in \mathcal{U}, \hat{y}_k \in \mathcal{Y}
\end{aligned} \tag{3.37}$$

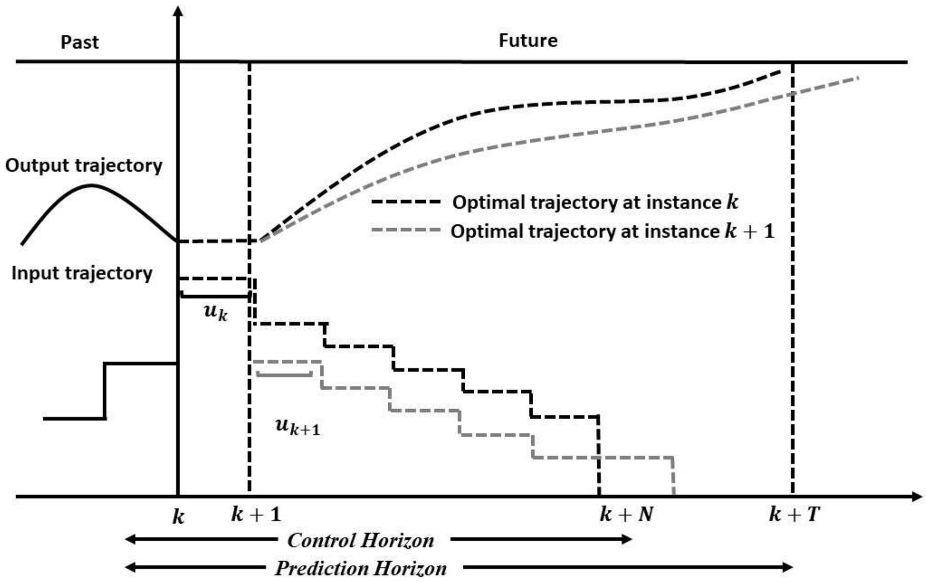


FIGURE 3.2: Receding horizon principle

where, the term u_{k-1} is the applied input at sampling instance $k-1$, $\mathcal{X}$,$\mathcal{U}$ and $\mathcal{Y}$ are set of constraints on states, outputs and inputs respectively.

The solution to the optimization problem presented in (3.37) produced sequence of controls over the complete control horizon. However, at every sampling instance k, only the first control input of the sequence is applied to system, i.e. $u_k = \bar{u}_k^\star$. The procedure is repeated at next sampling instance $k+1$ and the whole principle is illustrated in Fig. 3.2.

As both prediction and control horizons are shifted at next sampling instance, a measurement of the state trajectory is taken and used in the optimization framework. But in real life applications, it is difficult to measure all the system states and parameters due to possible physical constraint on the system states or due to the presence of noise or uncertainty in the measurement. The problem is addressed by introducing the concepts of nonlinear filters, observers or estimators which results in least-costly identification of the unknown parameters of the system.

The model obtained from the system identification steps is responsible for the quality of the MPC control sequence and system performance. Models with uncertainties or imprecise information leads to tracking errors due to wrong gain estimates. The problem is address to some extent by introducing the concept of integral action to MPC framework. The modeling errors due to measurement noise or uncertainties may also lead to violation of the constraints or in worst case scenario may cause instability of the system. Hence, the quality of the model is of central importance for OID in MPC framework.

3.4 Chapter Summary

In this chapter, a detailed description was presented on fundamentals of proposed OID framework for active parameter identification. For system identification of nonlinear dynamic systems, extended Kalman filter, unscented Kalman filter and Gaussian sum filter were used. Detailed discussion was presented on the three strategies and the complete algorithm was discussed. The details of linear Kalman filter is presented in Appendix A. To address the optimality criterion for the identified information, a detailed study was presented for different optimality criteria. From the information-based criteria, A-optimality criteria was used in this thesis. The general model predictive control formulation is presented later to understand the framework.

Chapter 4

OPTIMAL INPUT DESIGN FOR ACTIVE PARAMETER IDENTIFICATION

"We definitely use nonlinear systems and nonlinear indicators. Linear indicators, such as filters with moving averages, have been mined dry."

– William Eckhardt

In this chapter, the proposed OID framework for active parameter identification (API) is presented. The details discussed in Chapter 3 will be used to develop a feedback control law for parameter identification and control performance.

4.1 System Description

In this section, a detailed description of general nonlinear dynamic system is presented which will be used in this chapter and rest of the thesis.

4.1.1 Preliminaries

Consider the state vector $x_k \in \Re^{n_x}$ of a given nonlinear discrete time system which evolves according to a given equation

$$x_{k+1} = f_k(x_k, u_k, \theta) + \xi_k \quad k = 0, 1, \dots \tag{4.1}$$

where, k is the current sampling index, $u_k \in \Re^{n_u}$ is the control vector, $\theta \in \Re^{n_\theta}$ is the unknown parameter vector and ξ_k is the white-noise sequence with zero-mean and covariance Q_k^ξ. It is assumed that the initial condition x_o is obtained from a known density function $p(x_o)$. The system behavior is observed through the measurement equation for the quantities $y_k \in \Re^{n_y}$ as

$$y_k = h_k(x_k, \theta) + \eta_k \quad k = 0, 1, \ldots \tag{4.2}$$

where, η_k is the white-noise sequence with zero-mean and covariance R_k. The two noise quantities ξ_k and η_k do not depend on the initial condition x_o and are assumed to be mutually independent.

4.1.2 Active Parameter Identification Problem

The problem of OID for active parameter identification is to find such an optimal control law that can maximize the amount of information on the unknown system parameters while satisfying some desired system performance measures. In literature, two methods are often proposed for the solution of OID problem. In the first case, it is assumed that the system states and parameters are all measurable and the optimal control signal is designed on the basis of the measured information [207]. But in real life applications, it is difficult to measure all the system states and parameters due to possible physical constraint on the system states or due to the presence of noise or uncertainty in the measurement. The problem is addressed by introducing the concepts of nonlinear filters, observers or estimators which results in least-costly identification experiment [208–210].

In order to formally state the problem of active parameter identification in OID framework, let us first introduce the idea of active parameter identification.

Definition 4.1. The active parameter identification (API) problem is to find such an optimal feedback control law which when subjected to the system given in (4.1), yields the maximal amount of information on the uncertain parameter vector θ and achieves some desired system performances.

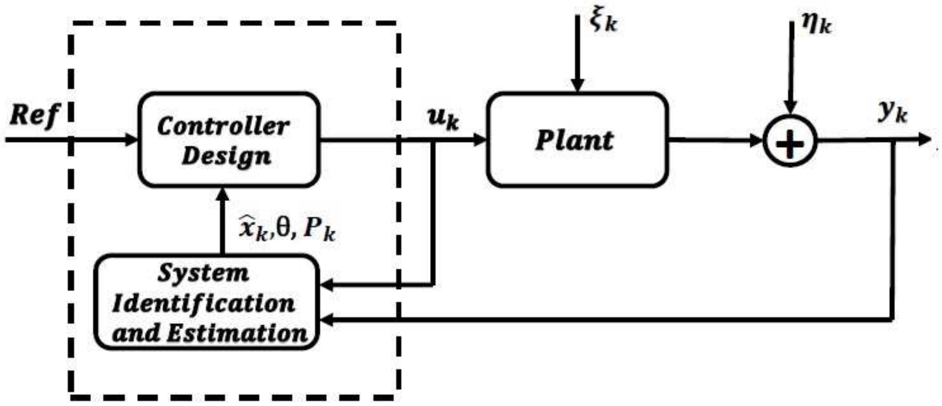


FIGURE 4.1: Block Diagram: Active Parameter Identification

Here, the idea of API is presented as an OID problem. The information obtained from the identification experiment is used to solve a receding horizon optimal control problem in NMPC framework. The complete idea of the proposed strategy is given in Fig. 4.1, where *plant* is the nonlinear discrete time dynamic system and *system identification and estimation* block represents the identification strategy (EKF, UKF or GSF) providing the necessary information about the unknown state vector x_k, unknown parameter vector θ_k and the covariance matrix P_k. The *controller design* block represents the MPC framework to design the excitation signal. The terms ξ_k and η_k are the process and measurement noise respectively.

The fundamental objective of this strategy is to design an optimal control law for the system given in (4.1) which can maximize the amount of information on the unknown system parameters by minimizing some cost related to the identified parameters while considering also some process performance and constraints on the system. The problem of API in the OID framework can be stated as

Problem 4.2. *Let $\alpha \in [0, 1]$ be a trade-off factor and N be the maximum sampling index in control horizon for which the API problem is solved in the OID framework for $k = 0, 1, \ldots,$ as*

$$J_k = \alpha J_k^{proc} + (1 - \alpha) J_k^{info} + \beta J_k^{cons} \tag{4.3}$$

where, J_k is the total cost computed at every sampling instance k, J_k^{proc} represents the process cost related to the desired performance of the system and J_k^{info} represents the information cost related to the identified parameter. The term J_k^{cons} represents a soft constraint function (exterior penalty function) on estimated system states $\hat{x}_k$ and the sequence of controls u_k to keep it in possibly specified bounds while β is the penalty parameter used to weight different constraint functions. The choice of β is made with monotonically increasing values by the use of sequential unconstrained optimization technique ([211]) which is presented in Appendix C. By acting on the trade-off parameter α, one can solve between a conventional receding horizon optimal control problem $\alpha = 1$ and an optimal identification experiment $\alpha = 0$. It is also possible to chose a value of $\alpha \in (0, 1)$ where on can achieve both identification of the parameter and achieve the desired system performance.

Since the aim of the proposed OID framework is to collect the maximum system information from the identification experiment, it is necessary to minimize some cost (identification criterion) related to the identified parameters. As a measure of accuracy of the identified information, it has substantial influence on the design of optimal input signal. In order to excite all the process modes, the designed input signal must be sufficiently rich to persistently excite the unknown system which depends on the identified information but also should account for the (estimated) states of the system in order to avoid configurations where the gain in information on the parameter is poor. Thus it is important to choose a suitable optimality criterion for the identification experiment. A detailed description on the selection of optimality criterion is presented in Section 3.1.

4.2 Identification Step

The aim of OID framework is to obtain the maximum amount of information from the unknown nonlinear system by minimizing a cost related to the identification of the parameter and generate a system model which when subjected to some control application, guarantees the acceptable desired control performance. As the designed control input directly influences the estimate of the system states and its parameters accordingly, the desired control performance is affected by its design. In this work, we have tried to make a comparison study for OID based on different nonlinear system identification techniques i.e. EKF, UKF and GSF which are discussed in detail in Section 3.2. A comparison is made on the basis of identification performance for parameters and states and also on the system performance to track the reference signal.

4.2.1 Extended Kalman Filter

In systems and control community, the use of EKF for nonlinear system identification is very common. Due to statistical nature of the algorithm, it provides information about the (approximate) probability distribution of identified data which is helpful in improving the quality of obtained information. It is beneficial to use of EKF over other classical identification method as it provides information about variance of the estimation error which is partially or completely ignored in other identification techniques. This identified information is used to shape the input spectrum which results in achieving the desired control specifications. In this work, the idea of using EKF is to accurately identify the unknown/uncertain system parameters and states and design the optimal input signal using the identified information. The detail EKF methodology is presented in Section 3.2.1.

4.2.2 Unscented Kalman Filter

For the systems with highly nonlinear dynamics, EKF does not perform up to the mark because a linearized model is used to propagate the covariance matrix of the system. In the case of complex nonlinear systems, it involves costly computation of the Jacobian matrices which leads to slow convergence and implementation difficulties. Also, if the sampling time is not sufficiently small, this linearization may lead to filter instability. In order to address the limitations possessed by EKF in terms of linearization, UT is used to estimate the mean and covariance matrix instead of linearization by the Jacobian matrices. The UKF addresses the assumption that it is easy to estimate a Gaussian distribution rather than to approximate an arbitrary complex nonlinear system. In control community, UKF emerges as a strong and powerful nonlinear estimation method and proved its superiority over the EKF in many applications. In this work, UKF is used with the MPC framework to identify the unknown parameters and states of the system. The information obtained from UKF is used for the OID and the results are compared with the one obtained by EKF. The detailed formulation of UKF algorithm is presented in Section 3.2.2.

4.2.3 Gaussian Sum Filter

For the systems with high estimation error or high level of measurement noise or non-Gaussian noise case, the performance of EKF (or UKF) degrades. Also, the EKF algorithm only involves the approximation of first and second order moments of the probability density function. This idea is refined by considering the high order moments of the densities in GSF. The Gaussian sum approximation for the nonlinear state estimation is catching the interest of the researchers in context of Bayesian estimation. In GSF algorithm, a more sophisticated Gaussian sum approximation is used for nonlinear system estimation which calculate the a priori density by means of Gaussian sum densities. The algorithm involves a bank of extended Kalman filters, which are solved, in parallel, for each term in the Gaussian sum. In this dissertation, we have used the GSF algorithm as a estimation strategy for both unknown states and parameters of nonlinear dynamic system in the proposed combined framework. The information obtained from the GSF is used for the OID and a comparison study is carried out with the results obtained from EKF or UKF. A detailed discussion of GSF is presented in Section 3.2.3.

4.3 Optimality Criterion

The key idea behind the active parameter identification is to shape the optimal input signal in such a way that the quality of the identified information become as good as possible. In practice, some information measure related to the covariance matrix of the identified parameter vector $P_k^{\theta\theta}$ is optimized in order to maximize the amount of information from the system. A detailed discussion on the optimality criteria was presented in Section 3.1 where three different types of optimality criteria are discussed i.e, information-based criteria, distance-based criteria and compound design criteria. Due to high level of complexity and extensive computational cost associated with *compound design* criteria and *distance-based* criteria, the use of both criteria is not very frequent.

In this work, we are using A-optimality criterion from the *information-based* criteria as the possible cost related to the identified parameter. The choice is motivated by the fact that the information is directly available and it depends on the information matrix for the design problem. It minimizes the variance of the unknown parameter vector, which is similar to maximizing the information on the unknown parameters.

4.3.1 A-Optimality Criterion

Among different optimality criteria discussed in Section 3.1, the A-optimality criterion is a popular choice to use as an information cost in OID framework. In this work, A-optimality criterion is used as the information cost which tends to improve the quality of the identified information by minimizing the *trace* of the covariance matrix related to the unknown parameter. The A-optimality criterion is described as

$$A\text{−}optimality = trace\,(P_k^{\theta\theta}) \tag{4.4}$$

where $P_k^{\theta\theta}$ is the part of the covariance matrix related to the identified parameter vector θ_k. The choice will help to improve the quality of the excitation signal which correspondingly will improve the expected knowledge of the unknown system. The full covariance matrix is given as

$$P_{k|k} = \begin{bmatrix} P^{xx}_{k|k} & P^{x\theta}_{k|k} \\ P^{\theta x}_{k|k} & P^{\theta\theta}_{k|k} \end{bmatrix}$$

where $P^{xx}_{k|k}$ is the part of the covariance matrix related to the estimated states while $P^{\theta\theta}_{k|k}$ is the part of the covariance matrix related to the unknown parameter. The cross-diagonal terms $P^{x\theta}_{k|k}$ and $P^{\theta x}_{k|k}$ represents the part of the covariance matrix associated with both estimated states and parameters.

4.4 Optimal Input Design Framework

In this section, all the fundamentals of OID framework (presented in previous sections and chapters) for active parameter identification are put together to propose a combined algorithm.

4.4.1 Nonlinear Model Predictive Control

Due to straightforward implementation algorithm, bulk of research literature, flexibility to deal with several model types and ability to explicitly incorporate the constraints, NMPC has been widely used in process industry. At every time instant, the system is updated using the information obtained from the identification experiment and the optimization process is repeated simultaneously. As the control horizon is shifted on every time instant, this problem is also called receding horizon optimization problem. The process is repeated on every time instant using only the first control sequence. In this work, a receding horizon optimization problem is formulated as discussed in Section 3.3 and solved by using MPC framework which tends to minimize a cost related to the identified information and design an optimal control signal that can help to achieve the desired system performance.

4.4.2 Proposed Framework

In this work, the problem of OID for API is proposed in NMPC framework. At every sampling instance k, the proposed framework is used to solve a receding horizon optimization problem over a control horizon of N future steps based on the identified information. The general cost function given in (4.3) can be written in the form

$$
\begin{aligned}
J_k(\bar{U}_k) &= \alpha J_k^{proc}(\bar{U}_k) + (1-\alpha) J_k^{info}(\bar{U}_k) + \beta J_k^{cons}(\bar{U}_k) \\
&= \alpha \left[\sum_{\ell=1}^{N-1} \Gamma_{\ell}(\bar{x}_{k+\ell}, \bar{u}_{k+\ell}, \hat{\theta}_{k+\ell|k+\ell}) + \Gamma_{N}(\bar{x}_{k+N}, \hat{\theta}_{k+N|k+N}) \right] \\
&+ (1-\alpha)\, \mathbb{E} \left[\sum_{\ell=1}^{N-1} \Pi_{\ell}(\bar{P}_{k+\ell|k+\ell}) + \Pi_{N}(\bar{P}_{k+N|k+N}) \right] \\
&+ \beta \left[\sum_{\ell=0}^{N-1} \Xi_{l}(\bar{x}_{k+\ell}, \bar{u}_{k+\ell}) + \Xi_{N}(\bar{x}_{k+N}) \right]
\end{aligned} \tag{4.5}
$$

where, k is the current discrete time index and ℓ represents the time index running in the control horizon of length N considered at time k. The terms Γ, Π and Ξ represents the process cost function, the information cost function and the soft constraint function. The function Γ depends on the states $\bar{x}_k$ and the estimate of the unknown parameter vector $\hat{\theta}$ (states depends on the value of the parameter). The function Ξ is a soft constraint function on the state and the sequence of control in order to keep them in the desired constrained region. It is desirable to keep the sequence of control in certain bound (physical constraints). The candidate control vectors required to be optimized at every time instance in control horizon $\ell = 0, 1, \ldots, N-1$ are denoted by $\bar{u}_{k+\ell}$. The column vector $\bar{U}_k \in \Re^{mN}$ containing the full sequence of candidate control vectors considered at time k for $\ell = 0, 1, \ldots, N-1$ is defined as

$$
\bar{U}_k \triangleq col\,[\bar{u}_k, \bar{u}_{k+1}, \ldots, \bar{u}_{k+N-1}]
$$

subjected to the following set of state and input constraints

$$\hat{x}_{k_{min}} \leq \hat{x}_k \leq \hat{x}_{k_{max}}$$

$$u_{k_{min}} \leq u_k \leq u_{k_{max}}$$

In the control horizon, the system evolve according to the given equation

$$\bar{x}_{k+\ell+1} = f_k(\bar{x}_{k+\ell}, \bar{u}_{k+\ell}, \bar{\theta}) \;\; \ell = 0, 1, \ldots, N-1$$

where for every sampling instance k, the *process cost* $J_k^{proc}(\cdot)$ is computed by the certainty equivalence principle using the online state estimates $\bar{x}_k = \hat{x}_{k|k}$ and $\bar{\theta} = \hat{\theta}_{k|k}$. The process cost includes any possible functions of desired control performance which need to be achieved along with active parameter identification. For the information cost $J_k^{info}(\cdot)$, the expectation is approximated by Monte-Carlo simulations, averaging among V realization of initial states $\bar{x}_k$ and parameter vector $\bar{\theta}$, generated with mean value $\hat{x}_{k|k}$ and $\hat{\theta}_{k|k}$ respectively and covariance matrix $P_{k|k}$. Inside the control horizon $\ell = 0, 1, \ldots, N$, the predicted covariance matrices $\bar{P}_{k+\ell|k+\ell}$ given in (3.16) which are associated with the state vector $x_{k+\ell}$ and parameter vector θ are computed by propagating the identification algorithm (in simulation) inside the optimization horizon. The system identification method is initialized with $\bar{\hat{x}}_{k|k} = \hat{x}_{k|k-1}, \bar{\hat{\theta}}_{k|k} = \hat{\theta}_{k|k-1}$ and $\bar{P}_{k|k} = P_{k|k-1}$ and carried on by using the sequence of measurements generated (in simulation) as $\bar{y}_{k+\ell} = h(\bar{x}_{k+\ell}, \bar{\theta}_{k|k})$. For the process cost J_k^{proc} given in (4.5), at any time instance k, the transition cost for $\ell = 0, 1, \ldots, N-1$ and the terminal cost are represented as $\Gamma_\ell(\cdot)$ and $\Gamma_{_N}(\cdot)$ respectively. The transition cost $\Gamma_\ell(\cdot)$ is the part of the cost computed during the propagation of the process at every sampling instance $\ell = k+1, \ldots, k+N-1$ while the terminal cost $\Gamma_{_N}(\cdot)$ is the part of the cost computed at $\ell = k+N$. For the information cost, at every time index k, the transition cost is represented as $\Pi_{_\ell}(\cdot)$ for $\ell = k+1, \ldots, k+N-1$ and the terminal cost is denoted as $\Pi_{_N}(\cdot)$ for $\ell = k+N$. The terms $\Xi_l(\cdot)$ represent the soft constraint function (penalty function) for $\ell = k+1, \ldots, k+N-1$. The term $\Xi_N(\cdot)$ represents the value of the constraint function at $\ell = k + N$. The soft constraint function serves as a penalizing function for the system states and the sequence of controls to remain in the desired bound. The constant $\beta > 0$ is used to weight the penalty function.

The proposed OID framework uses the NMPC approach at every sampling instance $k = 0, 1, ...,$ to generate the full control sequences $\bar{U}(k)$ composed of all the control vectors $\bar{u}(k + l)$ for $l = 0, 1, .., N - 1$ which is optimized by minimizing the cost given in (4.5)

$$\bar{U}_k^{\star} = \arg\min_{\bar{U}_k} J_k(\bar{U}_k)$$

but only the first control vector of the sequence is applied

$$u(k) = \bar{u}^{*}(k)$$

The procedure is repeated at next time $k+1$ by getting a new value of the estimated parameter and repeat the optimization with a one step forward shift of the moving horizon.

4.5 Chapter Summary

In this chapter, the problem of OID for the active parameter identification of nonlinear dynamic system was addressed. The problem was formulated in a combined framework of system identification strategy and NMPC method, where EKF, UKF or GSF was used for system identification of nonlinear dynamic system and on the basis of the identified information, an OID problem was solved in NMPC framework. The A-optimality criterion was proposed as a measure of information on the unknown parameter which tends to minimize the *trace* of the covariance matrix related to the identified parameter as the information cost. The problem was formulated in a receding horizon context, where a trade-off parameter α has been introduced to weight between the *process cost* and *information cost*.

Chapter 5

NUMERICAL EXAMPLES

"A good simulation, be it a religious myth or scientific theory, gives us a sense of mastery over experience. To represent something symbolically, as we do when we speak or write, is somehow to capture it, thus making it one's own. But with this appropriation comes the realization that we have denied the immediacy of reality and that in creating a substitute we have but spun another thread in the web of our grand illusion."

– Heinz R. Pagels

The proposed algorithm of OID for active parameter identification of nonlinear dynamic system is evaluated on relatively abstract example like A-toy model and some complex systems like 2-DOF or 3-DOF nonlinear model of two-wheeled mobile robot. These examples are well discussed and understood in literature and make it suitable for experimental validation of the proposed algorithm. Evaluation of the proposed method on the discussed examples are made in simulations, which provides a situation close to theory and are helpful in understanding the effectiveness of the proposed scheme.

5.1 Numerical Examples

In this section, we have illustrated (in simulations) the effectiveness of the proposed algorithm with the help of numerical examples. The first example, a simple "toy" model, provides insight on the objectivity of the OID framework while the second example, a 3-DOF two-wheeled mobile robot model, is used to see the implementation of the proposed algorithm on a more realistic and complex non-linear system. For the comparison study in terms of performance in identifying the parameter and achieving the desired control performance, UKF and GSF are used as the identification strategy to compare the results with EKF on a 2-DOF mobile robot model.

5.1.1 A Toy Example

Consider a simple first order system

$$\dot{x} = \theta x + u \tag{5.1}$$

where $\theta = -1$ is the unknown parameter. The corresponding discrete time non-linear Euler approximation is given as

$$x_{k+1} = x_k + \delta_t[\theta x_k + u_k] \tag{5.2}$$

where, x_k is the system state which is assumed to be measurable, θ is the unknown parameter and u_k is the control input. For simplicity we have assumed only one unknown parameter. The measurement equation is given as

$$\begin{cases} y_k = (x_k + 1) + \eta_k, & \text{if } x_k \geq 0 \\ y_k = e^{x_k} + \eta_k, & \text{if } x_k < 0 \end{cases} \tag{5.3}$$

The measurement equations given in (5.3) is shown in Fig. 5.1 in which it can be seen that for $x < 0$ (region 1), the measurement is available according to

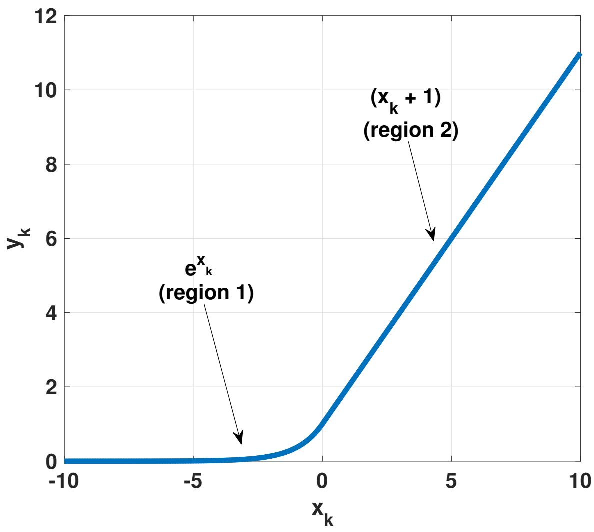


FIGURE 5.1: Measurement Regions

equation $y_k = e^{x_k} + \eta_k$ which makes it difficult to obtain any information on the unknown parameter while for $x \geq 0$ (region 2), the measurement is available as $y_k = (x_k + 1) + \eta_k$ and hence it is easy to identify the unknown parameter. Multiple simulations are performed for a control horizon $N = 30$ and sampling time $\delta_t = 0.01$ with different initial conditions in both regions.

The value of the tunable parameter α is critical to define the control objective as discussed in Section 4.4. The simulations are performed for two different scenarios corresponding to the value of α:

- Active parameter identification ($\alpha = 0$).
- Classical optimal control problem ($\alpha = 1$).

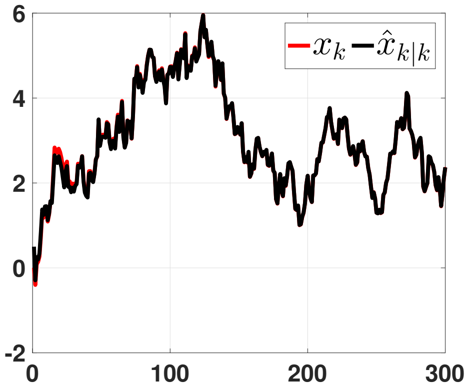
5.1.1.1 Case 1: Active Parameter Identification ($\alpha = 0$)

The control objective is to generate such a sequence of controls that can actively identify the unknown parameters by minimizing the cost related to the covariance matrix of the unknown parameter vector. Thus, by setting the trade-off factor in general cost function given in (4.5) to $\alpha = 0$, the problem of active parameter identification is solved. The information cost is given as

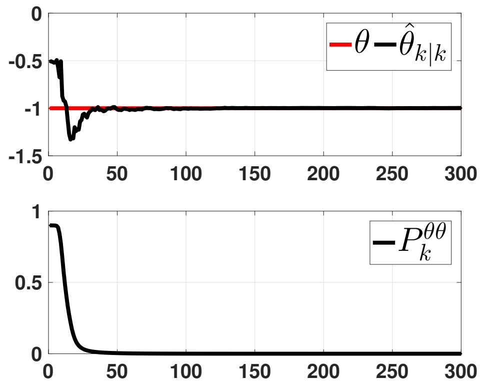
$$J_k^{info} = \mathbb{E}\left[\sum_{\ell=1}^{N-1} \kappa_l \, trace(\bar{P}^{\theta\theta}_{k+\ell|k+\ell}) + \kappa_N \, trace(\bar{P}^{\theta\theta}_{k+N|k+N})\right]$$

where, κ_l and κ_N are positive scalars used to weight the transition and terminal cost respectively. The simulations are performed with different initial conditions in order to verify the effectiveness of the proposed algorithm.

5.1.1.1.1 $x_0 = 0, \hat{x}_0 = 0.5, \hat{\theta}_0 = -0.5$ In this scenario, the state trajectory x_k and the estimated value of the state $\hat{x}_k$ is at the origin initially $x_0 = 0$ and $\hat{x}_0 = 0.5$ respectively. The initial estimated value of the parameter is given as $\hat{\theta} = -0.5$. At the initial point, it is not possible to identify the parameter as the information on the system trajectory is unavailable. The control is expected to bring the system to region 2 where it is possible to identify the unknown parameter.

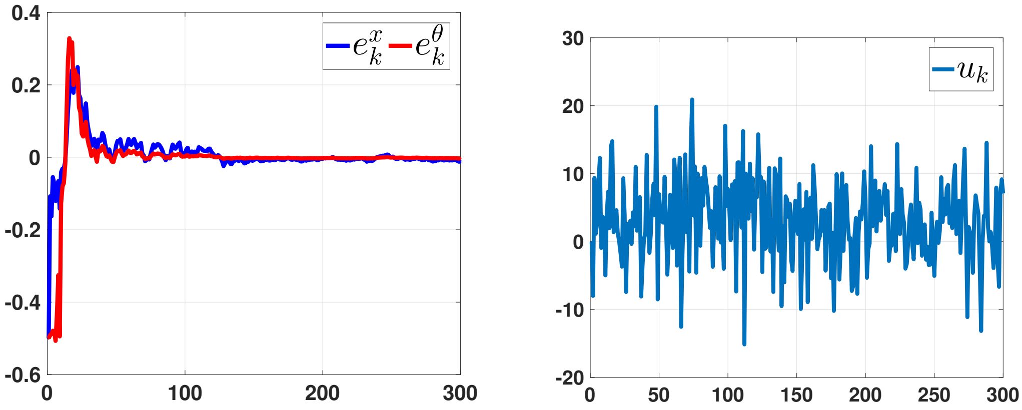


(a) System Trajectory



(b) Identified Parameter & Covariance Matrix

The state trajectory, x_k as shown in Fig. 5.2(a), moves in fact to region 2 where reliable information on state x_k can be collected and a quasi-random behavior can be seen which allows the identification of the parameter. The covariance matrix related to the identified parameter is minimized which results in a perfect identification of the unknown parameter as shown in Fig. 5.2(b). The error in the estimated state (e_k^x) and the identified value of the parameter (e_k^θ) are brought to zero after some iterations as shown in Fig. 5.2(c) which corresponds to a very



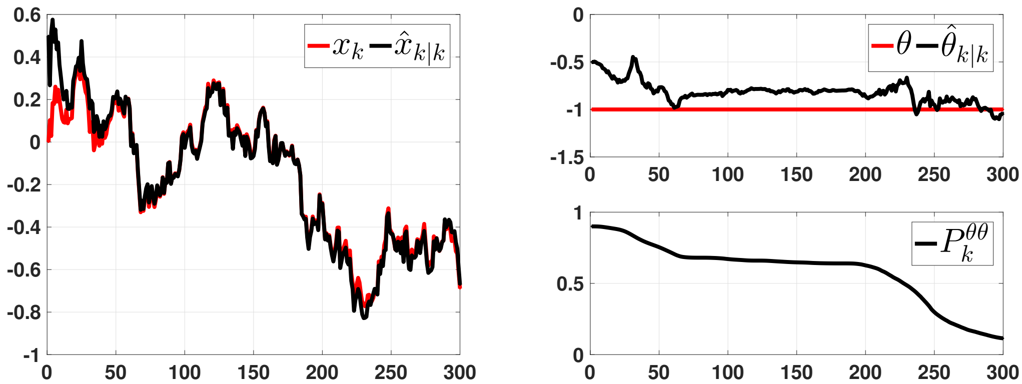
(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}$

(d) Control Signal

FIGURE 5.2: $x_0 = 0, \hat{x}_0 = 0.5, \hat{\theta}_0 = -0.5, \alpha = 0$

good state and parameter estimation. The sequence of control to achieve this active parameter identification is shown in Fig. 5.2(d).

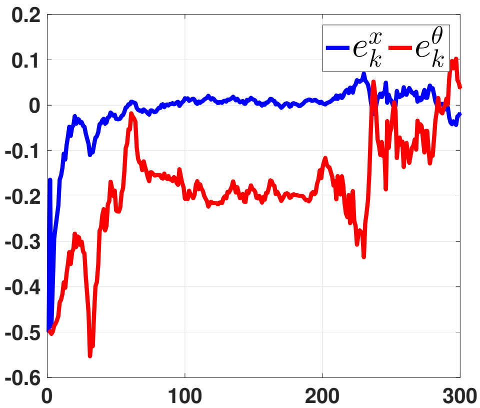
To show the superiority of the proposed method in terms of parameter identification, a random white Gaussian noise of same power as generated by our proposed algorithm is used to excite the system given in (5.2) with the same initial conditions.



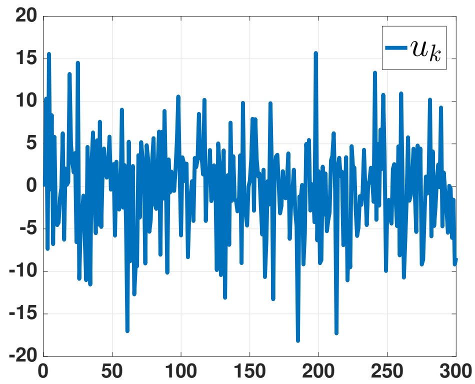
(a) System Trajectory

(b) Identified Parameter & Covariance Matrix

The state trajectory x_k moves to region 1 as shown in Fig. 5.3(a) which is not the desired behavior as the measurement does not give reliable information there and it is not possible to identify the parameter. The cost related to the covariance of the identified parameter is not minimized which results in a poor identification of



(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}$



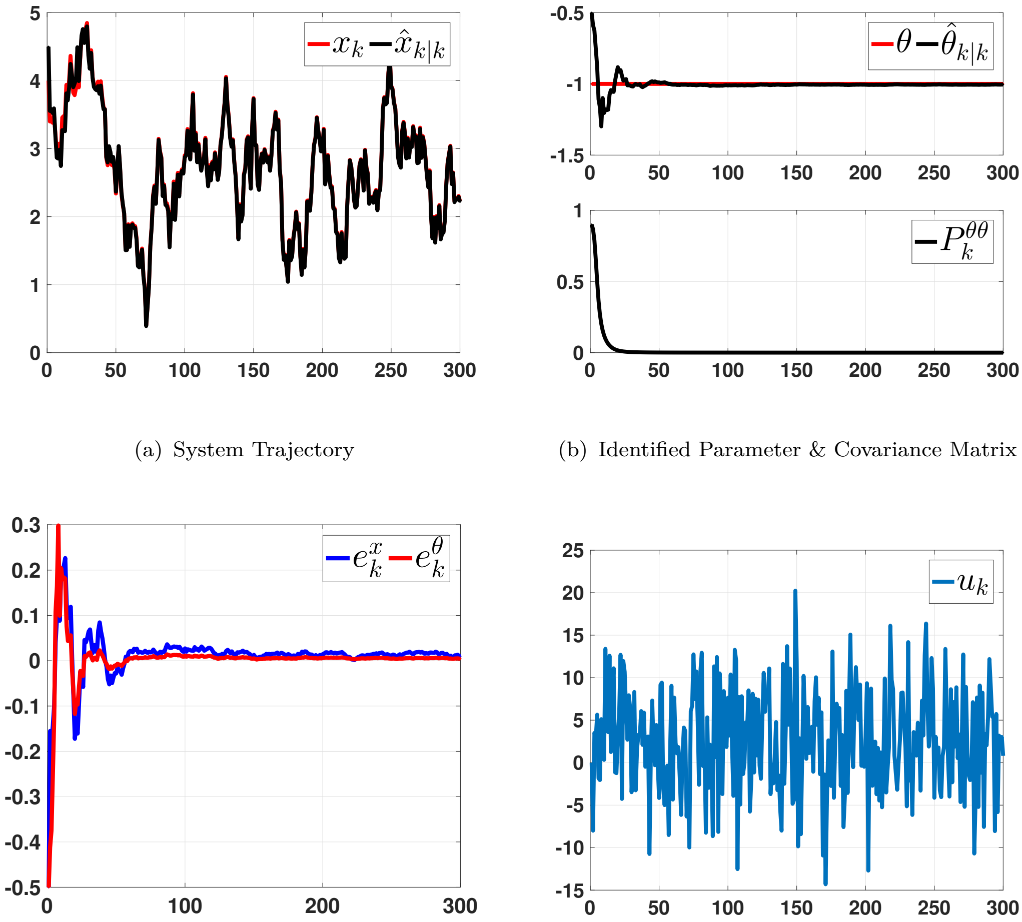
(d) Control Signal

FIGURE 5.3: White Gaussian Noise as Input Signal

the unknown parameter as shown in Fig. 5.3(b). The estimation error, for both state and parameter, is not brought to zero as shown in Fig. 5.3(c) which refers to a poor performance of the control input. The white Gaussian input signal is shown in Fig. 5.3(d).

These simulations indicates that the proposed strategy is effective enough to generate such excitation signals that provides the maximum information on the unknown system parameters as compared to a random white Gaussian signal.

5.1.1.1.2 $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5$ In this case, the system is initialized with a value $x_0 = 4$ and the estimated value of state and parameter is given as $\hat{x}_0 = 4.5$ and $\hat{\theta}_0 = -0.5$ in region 2, where the information on the system trajectory is reliable. The control sequence is expected to keep the system trajectory in the same region in order to identify the unknown parameter.



(a) System Trajectory

(b) Identified Parameter & Covariance Matrix

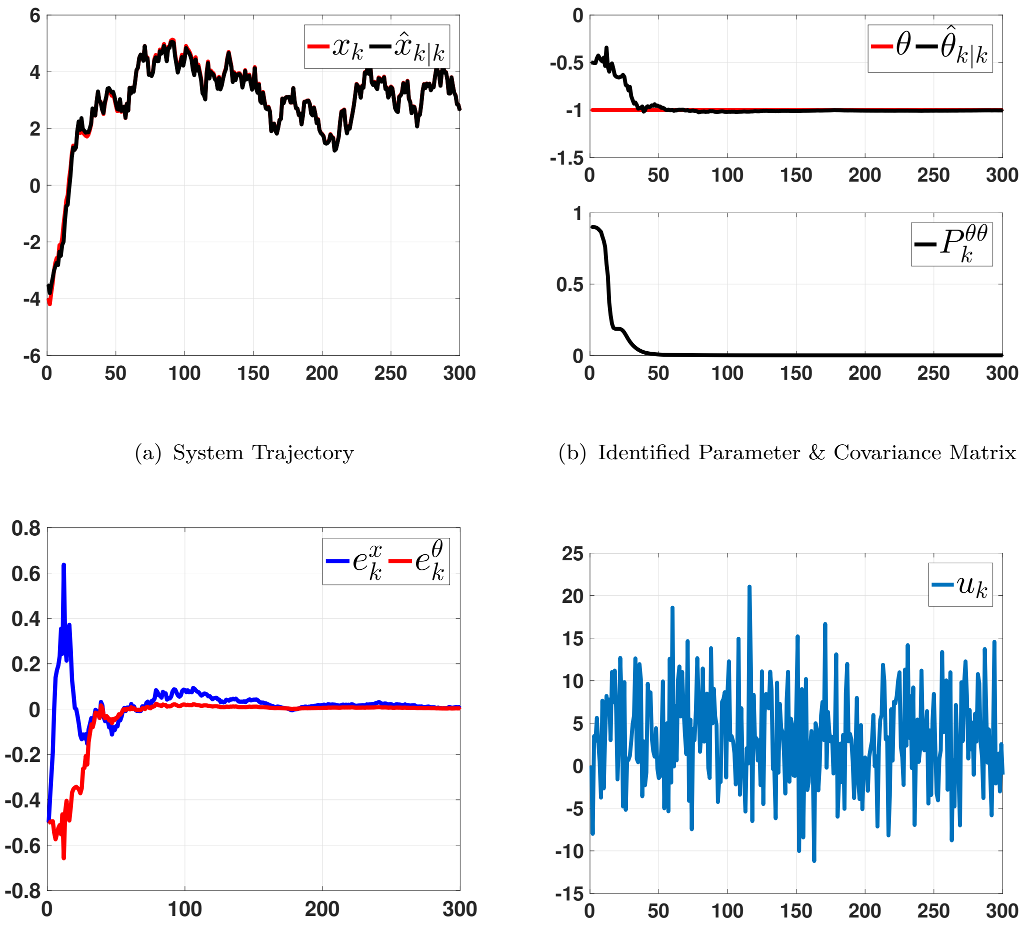
(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}$

(d) Control Signal

FIGURE 5.4: $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5, \alpha = 0$

Indeed, the state x_k shows a random behavior but remains in the region 2 as shown in Fig. 5.4(a) which results in the identification of the unknown parameter, while the part of the covariance matrix related to the unknown parameter is minimized as shown in Fig. 5.4(b). The error in actual and estimated value of state and parameter is brought to zero as shown in Fig. 5.4(c) which corresponds to a very good identification of the parameter. The almost random sequence of control is shown in Fig. 5.4(d).

5.1.1.1.3 $x_0 = -4, \hat{x}_0 = -3.5, \hat{\theta}_0 = -0.5$ Another possibility to see the effectiveness of the proposed algorithm is to consider the system with an initial value in region 1, where (due to the lack of information on the system state) the measurement is not reliable. In order to gain information on the uncertain parameter, we expect the state x_k to be brought to region 2.



(a) System Trajectory

(b) Identified Parameter & Covariance Matrix

(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}$

(d) Control Signal

FIGURE 5.5: $x_0 = -4, \hat{x}_0 = -3.5, \hat{\theta}_0 = -0.5, \alpha = 0$

In Fig. 5.5(a), it is shown that the system trajectory moves in fact to the region 2 (which is the desired system behavior). In region 2, a random behavior of the state x_k is shown which results in the identification of the parameter. The covariance matrix related to the unknown parameter is not minimized until the state x_k is in region 1. As the state x_k enters region 2, the part of the covariance matrix related to the identified parameter is minimized to zero which results in perfect identification of the parameter as shown in Fig. 5.5(b). The estimation error for both state and unknown parameter is minimized to zero as represented in Fig. 5.5(c) which also refers to state and parameter estimation. The sequence of control to achieve the desired identification performance is shown in Fig. 5.5(d).

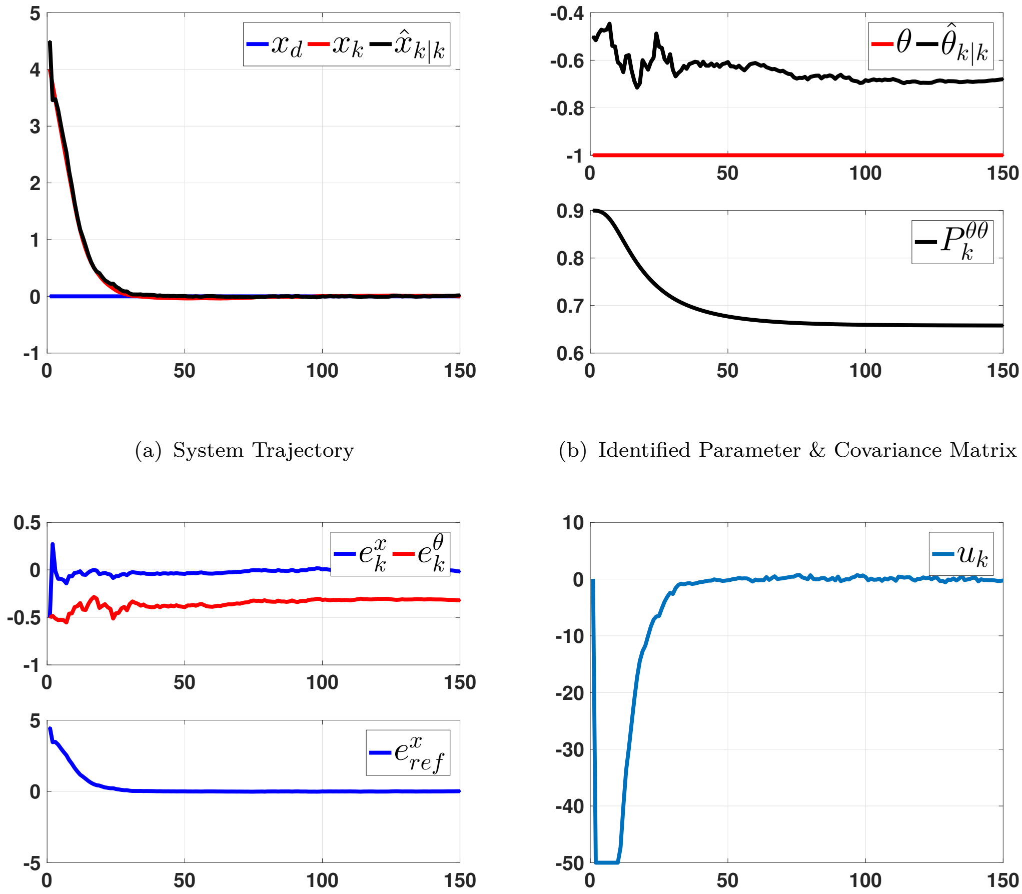
5.1.1.2 Case 2: Classical Optimal Control Problem ($\alpha = 1$)

In order to achieve some desired control performances, the trade-off factor $\alpha = 1$ is selected in the general cost function given in (4.5) to solve a classical optimal control design problem. The process cost is given as

$$J_k^{proc} = \left[\sum_{\ell=1}^{N-1} \rho_\ell (\bar{x}_{k+\ell} - x_{k+\ell}^\star)^2 + \gamma_\ell (\bar{u}_{k+\ell})^2\right] + \left[\rho_N (\bar{x}_{k+N} - x_{k+N}^\star)^2 + \gamma_N (\bar{u}_{k+N})^2\right] \tag{5.4}$$

where $x_k^\star$ is the desired reference trajectory, $\rho_\ell, \gamma_\ell, \rho_N$ and γ_N are positive scalars used to weight the transition and terminal cost related to state vector and control input respectively. The control objective is to track some desired reference trajectories starting from any initial conditions regardless of the information on the unknown parameter. The process cost given in (5.4) involves a cost related to the system trajectory $(x_k - x_k^\star)^2$ which is minimized over a control horizon of $N = 30$ future steps. Simulations are performed with different initial conditions to show the effectiveness of the proposed algorithm.

5.1.1.2.1 $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5, x_k^\star = 0, k = 1, 2, \ldots$ The system is initialized with an initial value in region 2 where it is possible to have useful measurements of the state x_k. The control objective is to regulate the system to the origin.



(a) System Trajectory

(b) Identified Parameter & Covariance Matrix

(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}, e_{ref}^x = \hat{x}_{k|k} - x_k^\star$

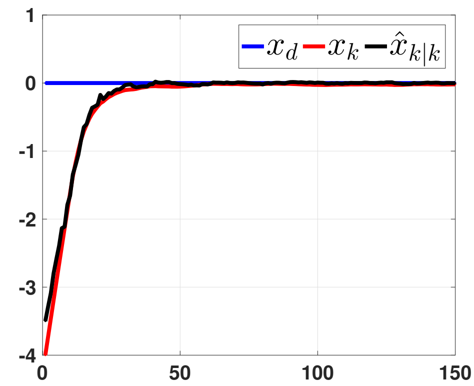
(d) Control Signal

FIGURE 5.6: $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5, \alpha = 1, x_k^\star = 0$

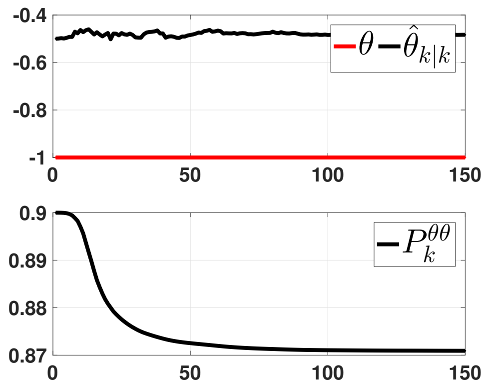
As the measurement on the state x_k is available on the initial value, it is clear from Fig. 5.6(a) that the system is regulated to the desired reference $x_k^\star = 0$ with a perfect estimate of the state. At the initial condition, it is possible to identify the parameter; thus, the identified value of the parameter moves to the true value for few iterations but as the state x_k reaches the origin (where information is not reliable), the value of the identified parameter remains constant for rest of the iterations. The cost related to the covariance matrix of the unknown parameter is not minimized which results in poor identification of the parameter as shown in Fig. 5.6(b). As the state is perfectly estimated, the error in state estimation e_k^x goes to zero but the error in the actual and estimated value of the parameter e_k^θ is not brought to zero as shown in Fig. 5.6(c). The tracking error e_{ref}^x is shown in the same figure, which represents a perfect tracking of the desired reference signal $x_k^\star = 0$. The control input is shown in Fig. 5.6(d), which shows that a maximum control is generated to bring the system to desired reference trajectory and as the desired objective is achieved, the control effort is brought to zero.

5.1.1.2.2 $x_0 = -4, \hat{x}_0 = -3.5, \hat{\theta}_0 = -0.5, x_k^\star = 0, k = 1, 2, \ldots$

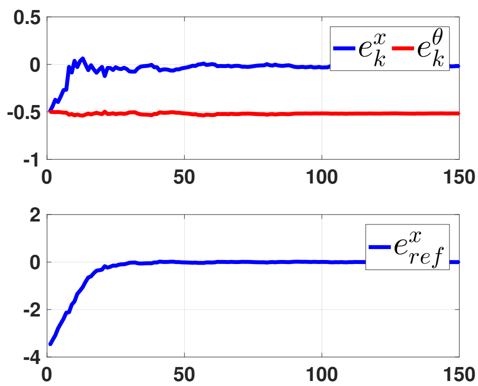
In this scenario, the system starts from an initial condition in region 1 where it is difficult to estimate the state and parameter. As $\alpha = 1$, the control objective is to regulate the state x_k to the origin ($x_k^\star = 0$) irrespective of the information gain on the unknown parameter.



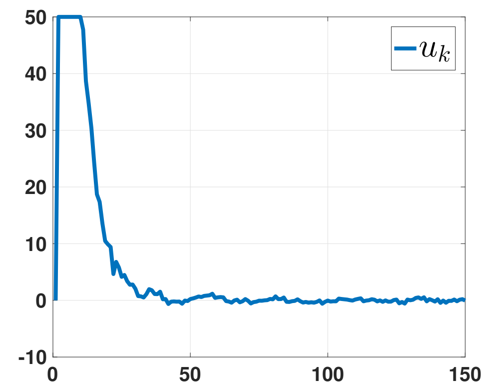
(a) System Trajectory



(b) Identified Parameter & Covariance Matrix



(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}, e_{ref}^x = \hat{x}_{k|k} - x_k^\star$

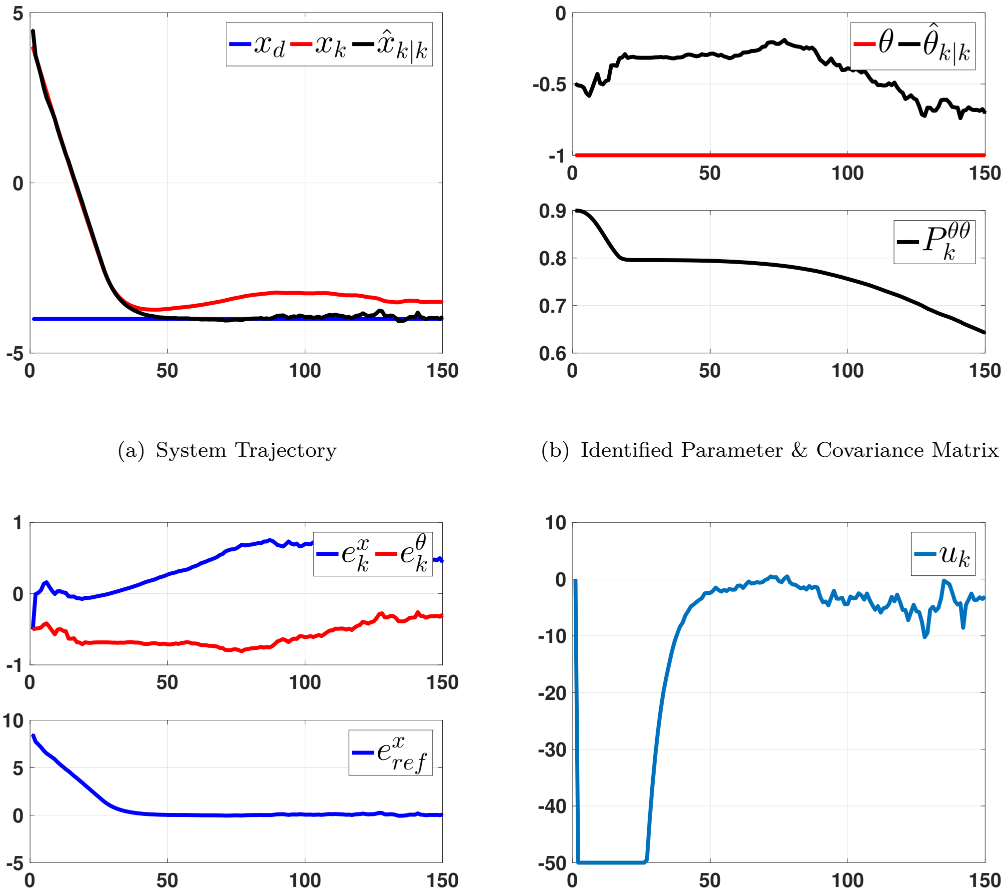


(d) Control Signal

FIGURE 5.7: $x_0 = -4, \hat{x}_0 = -3.5, \hat{\theta}_0 = -0.5, \alpha = 1, x_k^\star = 0$

Indeed, the state x_k regulates to the origin as shown in Fig. 5.7(a) with a poor estimate of the state for first few iterations but the estimate improves as the trajectory reaches the origin for rest of the iterations. At the initial value, due to lack of available information on the measurement, it is not possible to identify the parameter. Since, the state x_k remains in region 1 for all the iterations, the parameter is poorly identified and a constant value of the parameter is shown in Fig. 5.7(b). The part of the covariance matrix related to the unknown parameter is shown in the same figure which is consistent with the fact that the cost J_k^{info} is not minimized. The error in the state estimation e_k^x is brought to zero after some iterations. The error in parameter identification e_k^θ persists which indicates a poor parameter identification as shown in Fig. 5.7(c). The tracking error e_{ref}^x goes to zero as shown in the same figure which corresponds to the regulation of the $\hat{x}_k$ to the origin. The control effort is shown in Fig. 5.7(d). The identification of the parameter could have been better if the desired trajectory was in the region 2. Simulations are performed also for a case in which system trajectory starts from $x_0 = -4$ and track the desired reference trajectory $x_k^\star = 4, k = 0, 1, \ldots$. The identification of the parameter is much better than the above discussed case (due to shortage of space, the results are not presented).

5.1.1.2.3 $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta}_0 = -0.5, x_k^\star = -4, k = 1, 2, \ldots$ It is interesting to see the performance of the proposed scheme in terms of tracking a reference signal $x_k^\star = -4$ in region 1, where it is not possible to measure the state reliably. The system is initialized in region 2.



(a) System Trajectory

(b) Identified Parameter & Covariance Matrix

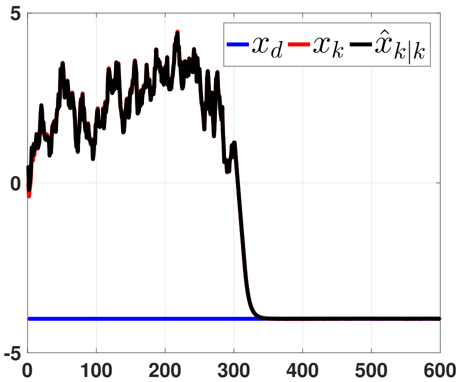
(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}, e_{ref}^x = \hat{x}_{k|k} - x_k^\star$

(d) Control Signal

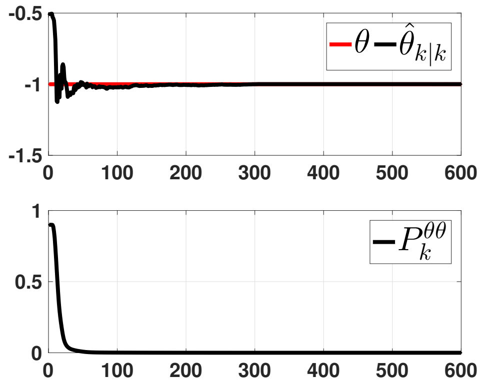
FIGURE 5.8: $x_0 = 4, \hat{x}_0 = 4.5, \hat{\theta} = -0.5, \alpha = 1, x_k^\star = -4$

Fig. 5.8(a) shows that the state x_k tracks the desired reference signal $x_k^\star$ in some iterations. After tracking the reference signal, due to poor information on the unknown parameter (as shown in Fig. 5.8(b)), also the state estimate becomes poor as entering region 1 EKF relies almost only on the prediction. The cost related to the covariance matrix of the identified parameter is not minimized as shown in Fig. 5.8(b) which results in poor parameter identification. The control effort is generated as shown in Fig. 5.8(d).

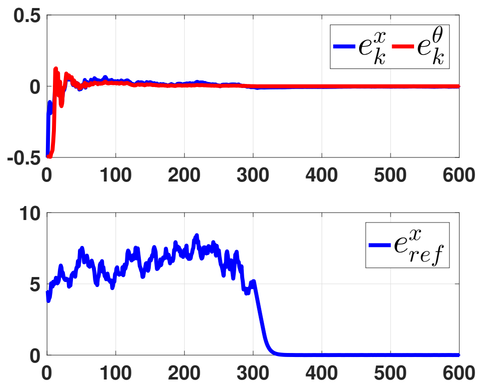
5.1.1.2.4 $x_0 = 0, \hat{x}_0 = 0.5, \hat{\theta}_0 = -0.5, x_k^\star = -4, k = 1, 2, \ldots$ The problem of poor state and parameter estimates in the previous case is addressed here by identifying the parameter first and once the information on the unknown parameter is available, the desired reference trajectory is tracked perfectly. The simulations are performed for $T = 600$ iterations. For the first 300 iterations, the active parameter identification is performed setting $\alpha = 0$, then the desired reference trajectory is tracked for the next 300 iterations by setting $\alpha = 1$ in the cost that is minimized.



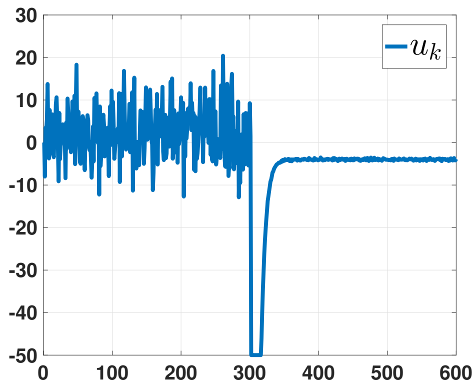
(a) System Trajectory



(b) Identified Parameter & Covariance Matrix



(c) $e_k^x = x_k - \hat{x}_{k|k}, e_k^\theta = \theta - \hat{\theta}_{k|k}, e_{ref}^x = \hat{x}_{k|k} - x_k^\star$



(d) Control Signal

FIGURE 5.9: $x_0 = 0, \hat{x}_0 = 0.5, \hat{\theta}_0 = -0.5, x_k^\star = -4$

In Fig. 5.9(a), it can be seen that the state trajectory x_k shows a quasi-random behavior for first 300 iterations consistent with the control objective of parameter identification ($\alpha = 0$). In this first phase, the trajectory moves to region 2 where the measurement on x_k is reliable and it is possible to identify the parameter correctly. For $k > 300$, as ($\alpha = 1$), the tracking performance improves. In Fig. 5.9(b) it is shown that the parameter identification is achieved and the cost related to covariance of the identified parameter is minimized. The reference tracking of desired signal $x_k^\star = -4$ is shown in Fig. 5.9(a). The error in state and parameter estimation is shown in Fig. 5.9(c) which is minimized indicating a very good estimate of state and parameter. The reference tracking error is also brought to zero as shown in Fig. 5.9(c). The sequence of control for this case is shown in Fig. 5.9(d).

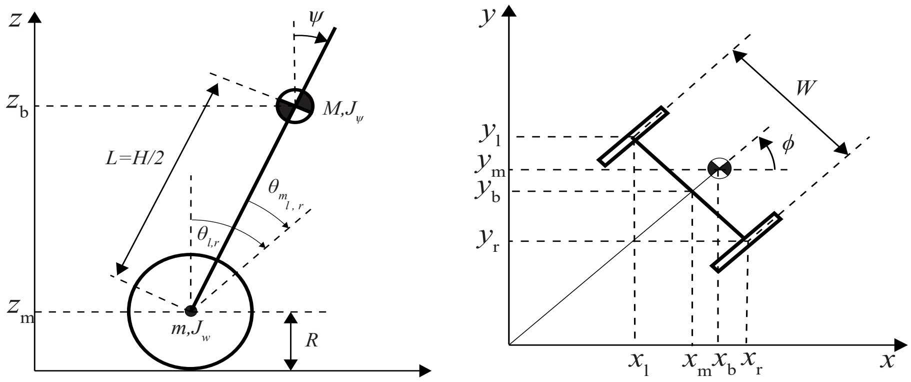
5.1.1.3 Example Summary

The problem of OID for the active parameter identification of a toy example was addressed. The problem was formulated in a combined EKF/NMPC framework, where EKF was used for system identification of nonlinear dynamic system and on the basis of the identified information, an OID problem was solved in NMPC framework. The problem was formulated in a receding horizon context, where a trade-off parameter α was introduced to weight between the *process cost* and *information cost*. The proposed strategy was implemented (in simulations) on a simple toy example and the simulation results were presented to show the effectiveness of the proposed methodology.

5.1.2 3-DOF Mobile Robot Model

5.1.2.1 Overview

The dynamics of two-wheeled mobile robot (2-DOF) is similar to that of an inverted pendulum on a wheeled cart. The 3-DOF model of two-wheeled mobile robot has an extra degree of freedom which incorporates the extra dynamics of the heading angle. It is a widely used control design platform for researchers, both from academia and industry [212–214]. To understand the dynamic equations of the system, the complete nomenclature, list of symbols and their values are given is Table 5.1. The nonlinear equations are taken from [1] which are derived from Fig. 5.10 in which the side and plane view of the mechanical model of two-wheeled robot is shown.



(a) Side View (b) Plane View

FIGURE 5.10: Side-view and Plane-view of two wheeled Robot

Symbol	Parameter	Value	Unit
m_b	Mass of the main body of robot	0.6	kg
m_w	Mass of each wheel	0.13	kg
L	Distance of the center of mass		
	from the wheel axle	0.20	m
R	Radius of each wheel	0.106	m
J_ψ	Body pitch inertia moment	0.63	kgm^2
J_ϕ	Body yaw inertia moment	1.12	kgm^2
G	Gravitational acceleration	9.81	$\frac{m}{s^2}$
D	Distance between the center of the		
	wheels axis C and the center of gravity G	0.212	m
ϕ	Pendulum tilt angle		rad
ψ	Yaw angle (orientation of the robot)		rad
x_l & x_r	Left and right wheel position		m
τ_1 & τ_2	Torque on the two wheels		Nm

TABLE 5.1: List of Parameters based on [1]

5.1.2.2 Equations of Motion

The 3-DOF dynamic model of two-wheeled mobile robot has six states: linear position x, linear velocity $\dot{x}$, yaw/heading angle ψ, yaw rate $\dot{\psi}$, pitch angle ϕ and pitch rate $\dot{\phi}$. It is assumed that the robot can move in any possible direction on plane by changing the heading angle. The input to the system is torque τ_1 and τ_2 which is applied to both wheels. It is assumed that the wheels of the robot will always remain in contact with the ground and there is no slip or cornering force on the wheel. To express the system in general form of (4.1), it is assumed that $x_1 = x, x_2 = \dot{x}, x_3 = \psi, x_4 = \dot{\psi}, x_5 = \phi, x_6 = \dot{\phi}$ and used the Euler approximation to rewrite the nonlinear equations discussed in [1] in discrete time nonlinear system equations as

$$x_{1_{k+1}} = x_{1_k} + \delta_t x_{2_k} \tag{5.5}$$

$$\begin{aligned} x_{2_{k+1}} = x_{2_k} + \delta_t \Bigg[& \frac{m_b d \sin x_{5_k}(m_b d^2 cos^2 x_{5_k} - (m_b d^2 + J_\phi))}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} \dot{x}_{3_k}^2 + \\ & \frac{m_b^2 g d^2}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} \sin x_{5_k} \cos x_{5_k} + \\ & \frac{(m_b d^2 + J_\phi)(m_b d \sin x_{5_k})}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} \dot{x}_{5_k}^2 + \\ & \frac{m_b d R \cos x_{5_k} + (m_b d^2 + J_\phi)}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} (\tau_1 + \tau_2) \Bigg] \end{aligned} \tag{5.6}$$

$$x_{3_{k+1}} = x_{3_k} + \delta_t x_{4_k} \tag{5.7}$$

$$\begin{aligned} x_{4_{k+1}} = x_{4_k} + \delta_t \Bigg[& - \frac{(m_b d^2 \sin x_{5_k} \cos x_{5_k})}{m_w(3L^2 + \frac{1}{2}R^2) + m_b d^2 \sin^2 x_{5_k} + J_\psi} \dot{x}_{3_k} \dot{x}_{5_k} - \\ & \frac{L}{R(m_w(3L^2 + \frac{1}{2}R^2) + m_b d^2 \sin^2 x_{5_k} + J_\psi)} (\tau_1 - \tau_2) \Bigg] \end{aligned} \tag{5.8}$$

$$x_{5_{k+1}} = x_{5_k} + \delta_t x_{6_k} \tag{5.9}$$

$$\begin{aligned} x_{6_{k+1}} = x_{6_k} + \delta_t \Bigg[& \frac{(3m_w m_b d^2 \sin x_{5_k} \cos x_{5_k})}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} \dot{x}_{3_k}^2 - \\ & \frac{(3m_w + m_b)}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} m_b g d \sin x_{5_k} + \\ & \frac{3m_w m_b d^2 \sin x_{5_k} \cos x_{5_k}}{m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi)} \dot{x}_{5_k}^2 - \\ & \frac{R(3m_w + m_b) + m_b d \cos x_{5_k}}{R(m_b^2 d^2 \cos^2 x_{5_k} - (3m_w + m_b)(m_b d^2 + J_\phi))} (\tau_1 + \tau_2) \Bigg] \end{aligned} \tag{5.10}$$

For simplicity, we have assumed that only the mass of the robot body $\theta = m_b$ is an unknown parameter whose value is estimated together with the state vector components. For a detail discussion on the above model, readers are encouraged to see [1] and the references therein. It is assumed that the linear velocity x_{2_k}, yaw angle x_{3_k} and pitch angle x_{5_k} are measurable and the information about the other states are not directly available.

5.1.2.3 Proposed Framework

For the case of 3-DOF two-wheeled mobile robot model, the simulations are performed for three cases by selecting different values of trade-off parameter α.

- Active parameter identification ($\alpha = 0$).
- Classical optimal control problem ($\alpha = 1$).
- Both active parameter identification and "classical" optimal control problem ($\alpha = 0.1$)

5.1.2.4 Control Strategy

The classical process cost J_k^{proc} given in (4.5) is defined for this particular case as

$$J_k^{proc} = J_{tran}^{proc} + J_{term}^{proc} \tag{5.11}$$

where J_k^{proc} is the total process cost which consists of the *transition cost* J_{tran}^{proc} and the *terminal cost* J_{term}^{proc}. The transition cost is computed at every sampling instance for $\ell = k+1, \ldots, k+N-1$ while the terminal cost is computed at $\ell = k+N$. The transition cost is given as

$$\begin{aligned} J_{tran}^{proc} = \sum_{\ell=0}^{N-1} \Bigg[& \rho_{\ell_{x_2}} \Big((\bar{x}_{2_{k+\ell}} - x_{2_d})^2 + \rho_{\ell_{x_4}} (\bar{x}_{4_{k+\ell}} - x_{4_d})^2 + \rho_{\ell_{x_5}} (\bar{x}_{5_{k+\ell}} - x_{5_d})^2 \Big) \\ & + \gamma_\ell \Big((\bar{\tau}_{1_{k+\ell}})^2 + (\bar{\tau}_{2_{k+\ell}})^2 \Big) \Bigg] \end{aligned} \tag{5.12}$$

where $\rho_{\ell_{x_2}}, \rho_{\ell_{x_4}}$ and $\rho_{\ell_{x_5}}$ are positive scalars and used to weight the cost related to linear velocity, yaw rate and pitch angle respectively. The positive scalars γ_ℓ are used to weight the cost on the sequence of control inputs. The terminal cost is given as

$$\begin{aligned} J^{proc}_{term} = \Bigg[\bigg(&\rho_{N_{x_2}}(\bar{x}2_{k+N} - x2_d)^2 + \rho_{N_{x_4}}(\bar{x}4_{k+N} - x4_d)^2 + \rho_{N_{x_5}}(\bar{x}5_{k+N} - x5_d)^2\bigg) \\ &+ \gamma_N\bigg((\bar{\tau}1_{k+N})^2 + (\bar{\tau}2_{k+N})^2\bigg)\Bigg] \end{aligned} \tag{5.13}$$

where $\rho_{N_{x_2}}, \rho_{N_{x_4}}, \rho_{N_{x_5}}$ and γ_N are positive scalars used to weight the cost related to linear velocity, yaw rate, pitch angle and control effort respectively at every sampling instance $l = k + N$. The information cost J^{info}_k defined in (4.5) is given as:

$$J^{info}_k = \mathbb{E}\left[\sum_{\ell=1}^{N-1} \kappa_l \; trace(\bar{P}^{\theta\theta}_{k+\ell|k+\ell}) + \kappa_N \; trace(\bar{P}^{\theta\theta}_{k+N|k+N})\right]$$

where, κ_l and κ_N are the positive scalars used to weight the transition cost and terminal cost respectively. As it is important to keep some states of the system in bound, penalty functions (soft constraints) on the desired states are introduced as

$$J^{cons}_k = \left[\sum_{\ell=0}^{N-1} \beta_l \; \Xi_l(\bar{x}1_{k+\ell}, \bar{x}3_{k+\ell}, \bar{x}5_{k+\ell}) + \beta_N \; \Xi_N(\bar{x}1_{k+N}, \bar{x}3_{k+N}, \bar{x}5_{k+N})\right]$$

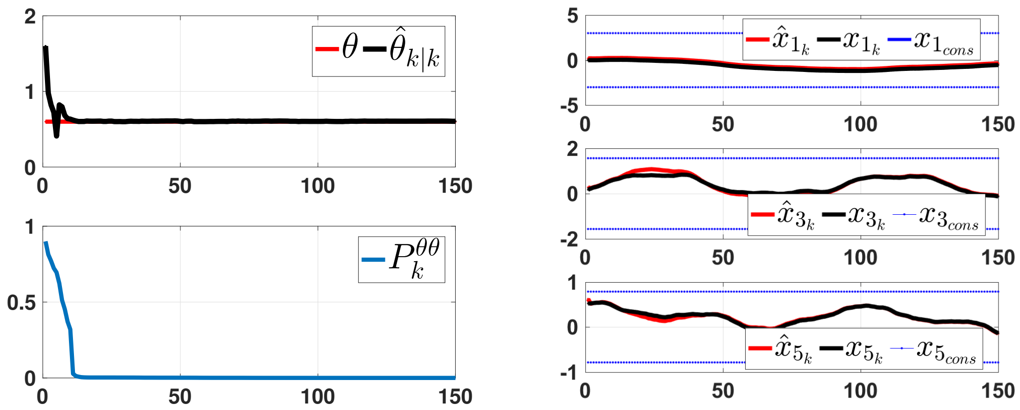
where β_l and β_N are positive tunable parameters used to weight different penalty functions on the states.

5.1.2.5 Simulation Results

It is intended to keep the robot in closed region (physical constraint) which imposes a soft constraint on the robot position as $x1_k = [max\{(\bar{x}1 - x1_{max}), (x1_{min} - \bar{x}1), 0\}]^2$, where $x1_{max} = +3m$ and $x1_{min} = -3m$. To avoid a big change in the heading angle of the robot, a penalty function is imposed as $x_{3_k} = [max\{(\bar{x}_3 - x_{3_{max}}), (x_{3_{min}} - \bar{x}_3), 0\}]^2$, where $x_{3_{max}} = \pi/2\ rad$ and $x_{3_{min}} = -\pi/2\ rad$. Also, it is important to keep the pitch angle x_{5_k} in a bound which is important to keep the robot vertically stable. The penalty function is imposed on the pitch angle $x_{5_k} = [max\{(\bar{x}_5 - x_{5_{max}}), (x_{5_{min}} - \bar{x}_5), 0\}]^2$, where $x_{5_{max}} = \pi/4\ rad$ and $x_{5_{min}} = -\pi/4\ rad$. It is also intended to keep the control sequences within a desired bound; hence a maximum ($\tau_{max} = +50Nm$) and minimum ($\tau_{min} = -50Nm$) bound is imposed on torques applied to both the wheels. The initial condition for the true system is chosen as $x_0 = [0\ 0\ \pi/18\ 0\ \pi/6\ 0]^\mathsf{T}$ and the mass of the robot body (which is assumed unknown parameter) is $m_b = 0.6kg$. The a-priori estimates are $\hat{x}_0 = x_0 + [0.2\ 0.2\ 0.1\ 0.1\ 0.1\ 0.1]^\mathsf{T}$ and $\hat{\theta}_0 = m_b + 1$. The simulations are performed for a control horizon of $N = 30$.

5.1.2.5.1 Active Parameter Identification ($\alpha = 0$)

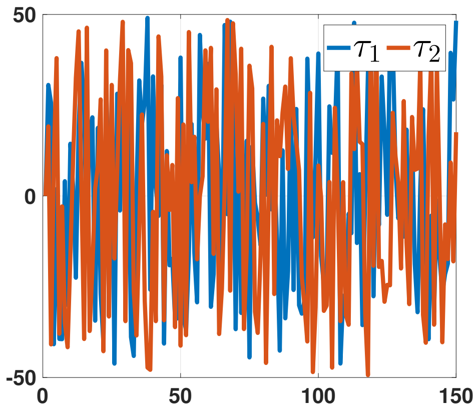
As the control objective is to obtain maximum information from the system, by acting on the trade-off factor $\alpha = 0$, the problem is addressed as an optimal experiment design problem.



(a) Identified Parameter & Covariance Matrix

(b) System Response with Constraints

The identified parameter and the part of the covariance matrix related to the identified parameter is shown in Fig. 5.11(a). It is shown that covariance matrix is minimized which results in a perfect identification of the parameter. The problem of avoiding the robot to go beyond the penalty margins are achieved by the introduction of soft constraints on the position x_{1_k}, yaw angle x_{3_k} and the pitch angle x_{5_k}. It is shown in Fig. 5.11(b) that all the states are within the desired bounds. As the control objective



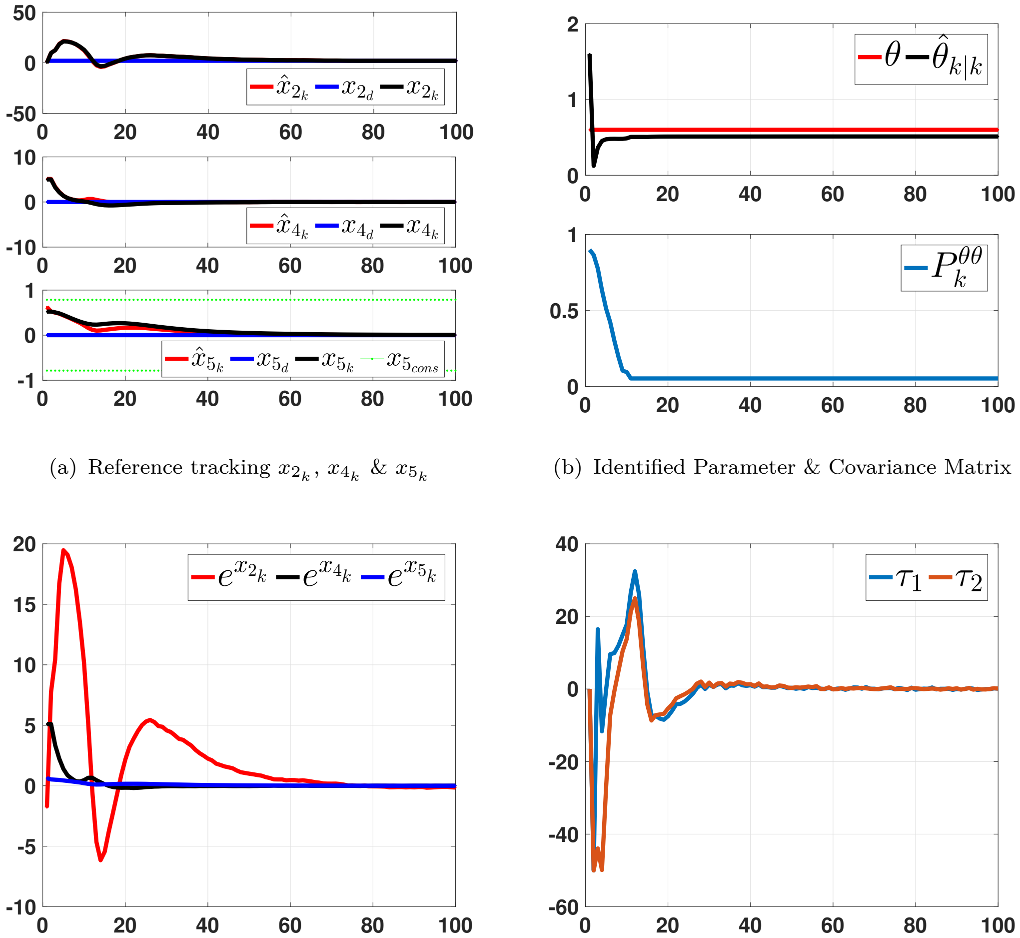
(c) Torque τ_1 and τ_2

FIGURE 5.11: Active Parameter Identification ($\alpha = 0$)

is to identify the parameter, almost random control sequences are generated as shown in Fig. 5.11(c) which are also within the desired bound.

5.1.2.5.2 Classical Optimal Control Problem ($\alpha = 1$) In order to achieve some desired system performance, the trade-off parameter $\alpha = 1$ is selected in the general cost function given in (4.5). The control objective is to stabilize the robot in vertical upward position which corresponds to bring the initial pitch angle $x_{5_k} = \pi/6\ rad$ to the desired reference pitch angle $x_{5_d} = 0\ rad$. It is also desirable to bring the initial yaw rate $x_{4_k} = 5\ rad/sec$ to the desired value $x_{4_d} = 0\ rad/sec$ and track the reference linear velocity $x_{2_d} = 2\ m/sec$ starting from an initial value of $x_{2_k} = 0\ m/sec$.

It is shown in Fig. 5.12(a) that as the pitch angle x_{5_k} is brought to its desired value x_{5_d}, the linear velocity x_{2_k} is changed from its initial value and an increase is shown which corresponds to stabilizing the robot. The yaw rate x_{4_k} is also brought to zero from its initial value to achieve the desired control performance. It is shown that the robot has tracked the desired linear velocity and also brought the yaw rate and pitch angle to zero to stabilize the robot in a vertically upward position. The EKF has tried to estimate the unknown parameter θ as shown in Fig. 5.12(b) but it can be seen that a constant biasing is always there. The covariance matrix related to the identified parameter is not minimized to zero which results in poor identification of the parameter. The tracking error between the estimated states $\hat{x}_{2_k}, \hat{x}_{4_k}, \hat{x}_{5_k}$ and the desired reference trajectories $x_{2_d}, x_{4_d}, x_{5_d}$ are shown in Fig. 5.12(c). All the three estimated states tracks the desired



(a) Reference tracking x_{2_k}, x_{4_k} & x_{5_k}

(b) Identified Parameter & Covariance Matrix

(c) $e^{x_{2_k}} = \hat{x}_{2_k} - x_{2_d}, e^{x_{4_k}} = \hat{x}_{4_k} - x_{4_d}, e^{x_{5_k}} = \hat{x}_{5_k} - x_{5_d}$

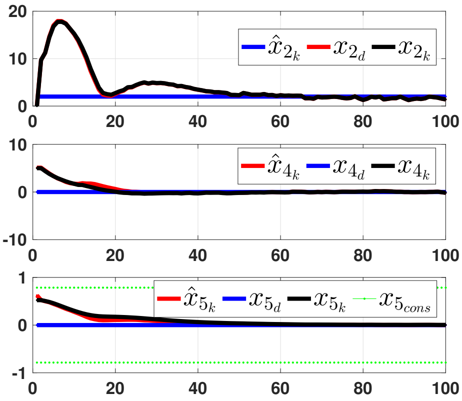
(d) Torques applied to both wheels

FIGURE 5.12: Classical Optimal Control Problem ($\alpha = 1$)

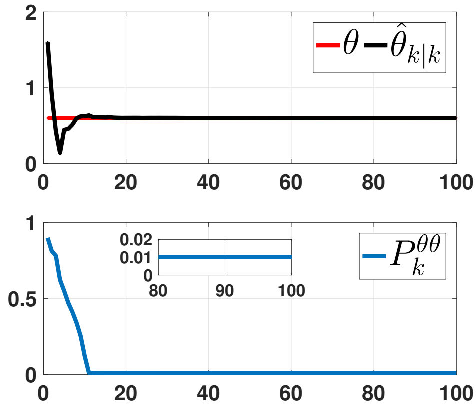
trajectories perfectly and the steady state error is brought to zero. The sequence of control inputs are shown in Fig. 5.12(d), which shows that the control effort is brought to zero as the desired reference trajectories are tracked.

5.1.2.5.3 Trade-off between Identification and Control ($\alpha = 0.1$) In this case, the choice of the trade-off weight α should guarantee to achieve both the parameter identification and control performance. The general cost function given in (4.5) is used to minimize both information cost J_k^{info} and the process cost J_k^{proc}. This choice of α will weight more to the information cost as compared to the process cost.

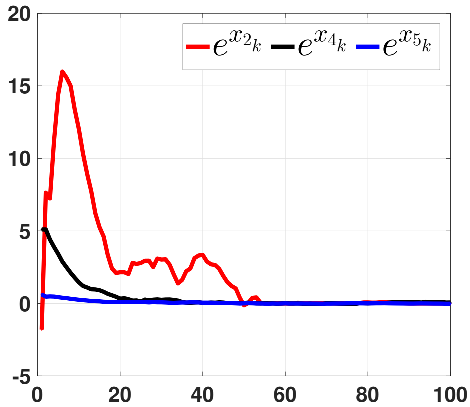
It is shown in Fig. 5.13(a) that the starting from a rest position, the robot achieve the desired reference velocity x_{2_d}. As the control objective is to identify the parameter as



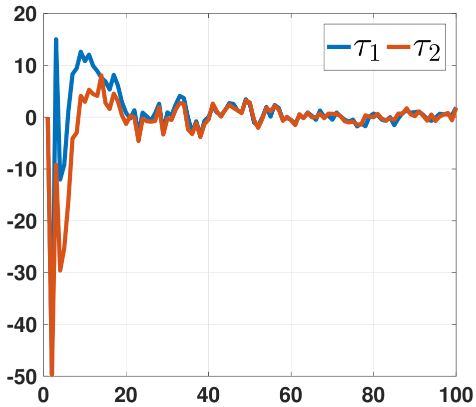
(a) Reference tracking x_{2_k}, x_{4_k} & x_{5_k}



(b) Identified Parameter & Covariance Matrix



(c) $e^{x_{2_k}} = \hat{x}_{2_k} - x_{2_d}, e^{x_{4_k}} = \hat{x}_{4_k} - x_{4_d}, e^{x_{5_k}} = \hat{x}_{5_k} - x_{5_d}$



(d) Torque τ_1 and τ_2

FIGURE 5.13: Trade-off between identification and control ($\alpha = 0.1$)

well, the linear velocity changes rapidly before reaching the desired value. The yaw rate x_{4_k} is also brought to zero from an initial value. The robot is stabilized in a vertically upward position and the initial pitch angle is brought to zero. The identified parameter and the covariance matrix related to the identified parameter are shown in Fig. 5.13(b) which indicates that the parameter is tracked better than the previous case and the cost related to covariance matrix is minimized. The tracking error in the different state estimates and the respective desired reference trajectory are shown in Fig. 5.13(c) which indicates that the error are minimized to zero. The sequence of control applied to the two wheels is shown in Fig. 5.13(d).

5.1.3 Example Summary

The problem of OID for the active parameter identification of nonlinear dynamic system was addressed. The problem was formulated in a combined EKF/NMPC framework, where EKF was used for system identification of nonlinear dynamic system and on the basis of the identified information, an OID problem was solved in NMPC framework. The A-optimality criterion was proposed as a measure of information on the unknown parameter which tends to minimize the *trace* of the covariance matrix related to the identified parameter as the information cost. The problem was formulated in a receding horizon context, where a trade-off parameter α was introduced to weight between the *process cost* and *information cost*. The proposed strategy was implemented (in simulations) on a 3-DOF two-wheeled mobile robot model and the simulation results were presented to show the effectiveness of the proposed methodology.

5.2 Comparison Study for Proposed Framework

In order to show the comparison of the system identification results obtained by the extended Kalman filter, we have tried to use the UKF or GSF as the identification strategy instead of EKF. The comparison among the three identification strategies is made by simulating a 2-DOF two-wheeled mobile robot model for different scenarios with different initial conditions and the obtained results are compared on the basis of the identification performance and achieved desired system characteristics.

5.2.1 2-DOF Mobile Robot Model

Two-wheeled mobile robot has the same dynamics as that of a inverted pendulum on a wheeled cart. It serves as a great platform for control design experiment and is widely studied by many researchers [212–214]. In this work, we have tried to use the 2-DOF model of two-wheeled robot which has tendency to move on a straight line and the control objective is to stabilize the robot in vertical upward position and the identification objective is to estimate the unknown mass of the robot body (unknown parameter).

Parameter	Symbol	Value
Rotation Angle of the Chassis	θ_p	rad
Mass of Wheel connected to both sides of the robot	M_w	0.03 kg
Mass of the Robot's chassis	M_p	0.6 kg
Radius of the Wheel	r	0.04 m
Nominal terminal Resistance	R	6.69 Ω
Height of the Robot	H	0.144 m
Dist. between z-axis and COG of the Robot	L	$\frac{H}{2}$
Moment of Inertia of the wheels	I_w	$\frac{1}{2}M_w r^2 kgm^2$
Moment of Inertia of the Robot's chassis	I_p	$0.0041kgm^2$
Back emf constant	k_e	0.268 $\frac{V.s}{rad}$
Torque constant	k_m	0.317 $\frac{Nm}{A}$
Applied Terminal Voltage	V_a	V

TABLE 5.2: List of Parameters based on [2]

5.2.1.1 Overview

To understand the dynamic equations of the system, the complete nomenclature, list of symbols and their values are given is Table 5.2. The nonlinear equations are derived from Fig. 5.10 in which the side view of the mechanical model of two wheeled robot is shown.

5.2.1.2 Equations of Motion

The dynamic equations of two-wheeled mobile robot are presented for linear motion case in which it is assumed that the robot can move only in a straight line. The dynamics related to the linear position x, linear velocity $\dot{x}$, angular position ψ and angular velocity $\dot{\psi}$ are considered in the model. The input to the system is voltage V_a which is applied to both wheels. It is assumed that the wheels of the robot will always remain in contact with the ground and there is no slip or cornering force on the wheel. The nonlinear equations of motion for a two-wheeled mobile robot have the following form

$$-M_p L\ddot{x}\cos\psi = (I_p + M_p L^2)\ddot{\psi} - \frac{2k_m k_e}{Rr}\dot{x} + \frac{2k_m}{R}V_a + M_p g L\sin\psi \tag{5.14}$$

$$\frac{2k_m}{Rr}V_a = (2M_w + \frac{2I_w}{r^2} + M_p)\ddot{x} + \frac{2k_m k_e}{Rr^2}\dot{x} + M_p L\ddot{\psi}\cos\psi - M_p L\dot{\psi}^2\sin\psi \tag{5.15}$$

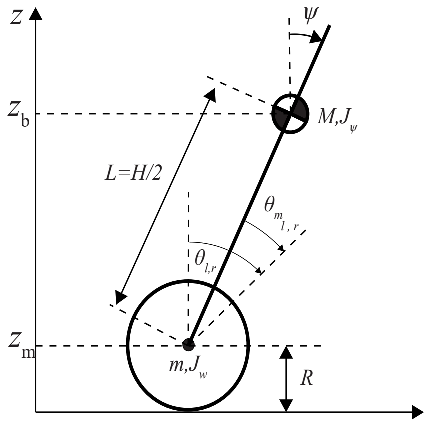


Figure 5.14: Side-view of two wheeled Robot

where x and $\dot{x}$ are the linear position and linear velocity of the robot while ψ and $\dot{\psi}$ are the angular position and angular velocity of the robot chassis. The mass of the chassis is assumed as the unknown parameter $\theta = M_p$ and used as the augmented state. To express the system in general form of (4.1), it is assumed that $x_1 = x$, $x_2 = \dot{x}$, $x_3 = \psi$ and $x_4 = \dot{\psi}$ and used the Euler method to rewrite the above equations in discrete time nonlinear system equations as

$$
\begin{aligned}
x_{1_{k+1}} &= x_{1_k} + \delta_t x_{2_k} \\
x_{2_{k+1}} &= x_{2_k} + \delta_t \Bigg[\frac{2k_m k_e}{\Lambda r R^2} \left[(I_p + M_p L^2) + r M_p L \cos x_{3_k} \right] x_{2_k} - \frac{M_p L (I_p + M_p L^2)}{\Lambda} x_{4_k}^2 \sin x_{3_k} \\
&\quad - \frac{(M_p L)^2 g}{\Lambda} \sin x_{3_k} \cos x_{3_k} - \frac{2k_m}{\Lambda r R} \left[(I_p + M_p L^2) + r M_p L \cos x_{3_k} \right] V_a \Bigg] \\
x_{3_{k+1}} &= x_{3_k} + \delta_t x_{4_k} \\
x_{4_{k+1}} &= x_{4_k} + \delta_t \Bigg[\frac{2k_m k_e}{\Lambda r R^2 (I_p + M_p L^2)} \Big[\Lambda R - M_p L \cos x_{3_k} \left[(I_p + M_p L^2) + r M_p L \cos x_{3_k} \right] \Big] x_{2_k} \\
&\quad - \frac{M_p g L}{(I_p + M_p L^2)} \sin x_{3_k} + \frac{(M_p L)^2}{\Lambda} x_{4_k}^2 \cos x_{3_k} \sin x_{3_k} + \frac{(M_p L)^3}{\Lambda (I_p + M_p L^2)} \cos^2 x_{3_k} \sin x_{3_k} \\
&\quad - \frac{2k_m}{\Lambda r R (I_p + M_p L^2)} \Big[\Lambda r - M_p L \cos x_{3_k} \left[(I_p + M_p L^2) + r M_p L \cos x_{3_k} \right] \Big] V_a \Bigg]
\end{aligned}
$$

Where, $\Lambda = (M_p l \cos x_{3_k})^2 - (I_p + M_p l^2)(2M_w + \frac{2I_w}{r^2} + M_p)$ and $\theta = M_p$. The model has also incorporated the dynamics of the DC motor attached to the wheels. For a more detailed explanation, readers are encouraged to see [2] and the references therein. We are assuming the case in which only the angular velocity $x_4 = \dot{\psi}$ is measurable state and the information about the other states are not directly available.

5.2.1.3 OID Strategy for Two-Wheeled Mobile Robot

This section presents the OID problem for a two-wheeled mobile robot in a combined framework of nonlinear system identification technique and model predictive control. As discussed in Section 5.2, the dynamic model has four states; two associated with the linear motion and two with the angular position of the robot. The fifth component of the augmented state is the unknown parameter which in this case is mass of the chassis M_p. It is assumed that only the angular velocity x_4 is measurable. In order to get the information of the unmeasured states and parameter, we have used the EKF, UKF or GSF as the identification strategy. This identified information is used in MPC framework to generate an optimal control signal which guaranteed the desired system performance subjected to control input. It is desirable for the two wheeled robot to achieve the vertical self-balance position starting from some initial angle and after achieving the stability, the linear velocity of the robot must be zero. For N future steps, cost related to the control performance and the identified parameter is minimized over a finite horizon. The general cost function given in (4.3) is formulated for the proposed strategy as

$$J_k(\bar{U}_k) = \alpha \left[\sum_{\ell=0}^{N-1} \Gamma_\ell(\bar{x}_{k+\ell}, \bar{u}_{k+\ell}, \hat{\theta}_{k|k}) + \Gamma_N(\bar{x}_{k+N}) \right] + (1-\alpha)\,\mathbb{E}\left[\sum_{\ell=1}^{N-1} \Pi_\ell(\bar{P}_{k+\ell|k+\ell}) + \Pi_N(\bar{P}_{k+N|k+N}) \right] + \beta \sum_{\ell=0}^{N} \Xi(\bar{x}_{k+\ell}) \tag{5.16}$$

Where, k is the current sampling instance and ℓ is the control horizon time index. For the control horizon $\ell = 0, 1, \ldots, N-1$, the desired control vectors needed to be optimized are given by $\bar{u}_{k+\ell}$. The column vector $\bar{U}_k \in \Re^{mN}$ is defined as

$$\bar{U}_k \triangleq \mathrm{col}[\bar{u}_k, \bar{u}_{k+1}, \ldots, \bar{u}_{k+N-1}]$$

The states of the system evolve in the control horizon fulfilling the given equation

$$\bar{x}_{k+\ell+1} = f(\bar{x}_{k+\ell}, \bar{u}_{k+\ell}, \bar{\theta}) \;\; \ell = 0, 1, \ldots, N-1$$

where the parameter vector $\bar{\theta}$ inside the control horizon is assumed as $\bar{\theta} = \hat{\theta}_{k|k}$. The certainty equivalence principle is used to compute the J_k^{proc} by setting $\bar{x}_k = \hat{x}_{k|k}$ and $\bar{\theta} = \hat{\theta}_{k|k}$, i.e. as their online estimates at time k . For the J_k^{info}, Monte-Carlo simulations are used to approximate the expectation by taking average of $\mathbf{V}$ realizations obtained by assuming the initial state vector $\bar{x}_k$, initial parameter vector $\bar{\theta}$ and the initial covariance matrix $\bar{P}_{k|k}$ equal to the $\hat{x}_{k|k}$, $\hat{\theta}_{k|k}$ and $P_{k|k}$ respectively. The covariance matrices $\bar{P}_{k+\ell|k+\ell}$ associated with the states $x_{k+\ell}$ and the parameter vector θ are calculated by propagating the nonlinear identification algorithms inside the control horizon $\ell = 0, 1, \ldots, N$. The values of $\hat{\bar{x}}_{k|k} = \hat{x}_{k|k-1}$, $\hat{\bar{\theta}}_{k|k} = \hat{\theta}_{k|k-1}$ and $\bar{P}_{k|k} = P_{k|k-1}$ are used as the initial conditions for these identification techniques and the measurement sequence are yielded (in simulations) as $\bar{y}_{k+\ell} = h(\bar{x}_{k+\ell})$. In (5.16), the transition cost at any time instant $k+\ell$ is represented by $\Gamma_\ell(\cdot)$ for $\ell = 0, 1, \ldots, N-1$ while the terminal cost is given as $\Gamma_N(\cdot)$. The transition and terminal cost related to the information of the system is given as $\Pi_\ell(\cdot)$ and $\Pi_N(\cdot)$ respectively.

For the particular two-wheeled mobile robot case, the problem can be formulated similarly to the general problem formulation discussed in (5.16). As the aim of the control performance is to balance the robot vertically upward, a cost related to the angular position $x_3 = \psi$ is included in the process cost. Also, a cost related to the control signal is added to the cost function which tends to keep the input signal in desired control region. The process cost for the proposed algorithm is given as

$$J_k^{proc} = \left[\sum_{\ell=0}^{N-1} \rho_\ell (\bar{x}_{3_{k+\ell}} - x_{r_{k+\ell}})^2 + \gamma_\ell (\bar{u}_{k+\ell})^2\right] + \left[\rho_N (\bar{x}_{3_{k+N}} - x_{r_{k+N}})^2 + \gamma_N (\bar{u}_{k+N})^2\right]$$

where ρ_ℓ, γ_ℓ, ρ_N and γ_N are positive scalars for transition and terminal cost and helpful in achieving the desired control specifications and x_r is the reference signal. The information cost for the proposed algorithm is given by

$$J_k^{info} = \mathbb{E}\left[\sum_{\ell=1}^{N-1} trace(\bar{P}^{\theta\theta}_{k+\ell|k+\ell}) + trace(\bar{P}^{\theta\theta}_{k+N|k+N})\right]$$

In order to keep the system states in bounds specially in the case of information retrieval $\alpha = 0$, a penalty function is imposed. The penalty function on $\bar{x}_k$ is given as

$$J_k^{cons} = \beta \sum_{\ell=0}^{N} \Xi(\bar{x}_{3_{k+\ell}})$$

where, β is the tunable weight on the penalty function which is given in this equation $\Xi(x_3) = [max\{(\bar{x}_3 - x_{max}), (x_{min} - \bar{x}_3), 0\}]^2$. It is intended to keep the angular position x_{3_k} in a bound $[x_{min}, x_{max}]$ chosen as $[+1rad, -1rad]$.

5.2.2 Simulation Results

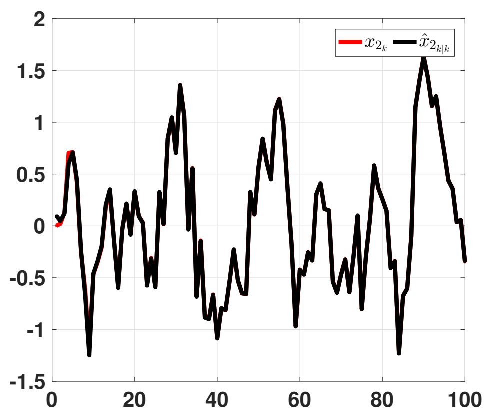
Simulations are performed with three different identification strategies (EKF, UKF or GSF) and results are presented in the subsequent sections. From (5.16), the proposed OID problem can be simulated in three cases by acting on the value of trade-off parameter α. If the purpose of OID design is just to acquire information about the unknown nonlinear system, the problem is solved for $\alpha = 0$. For $\alpha = 1$, the problem can be seen as a classical stochastic optimal control problem to achieve the desired control performance. For the third case, the trade-off parameter α can be tuned to a desired value where one can achieve both identification and control performance. The simulations are performed with a control horizon of $N = 35$. We have performed the observability analysis using the Lie Algebra and checked the full rank condition for the observability matrix. From the analysis, it is evident that the robot position is unobservable.

The detailed discussion on the three identification strategies (EKF, UKF and GSF) was given in Chapter 3. We have used EKF as the principle strategy for our proposed framework and then make a comparison with the other two strategies (UKF or GSF) and to analyze the performance of our proposed scheme in terms of system identification and achieving the desired control performance.

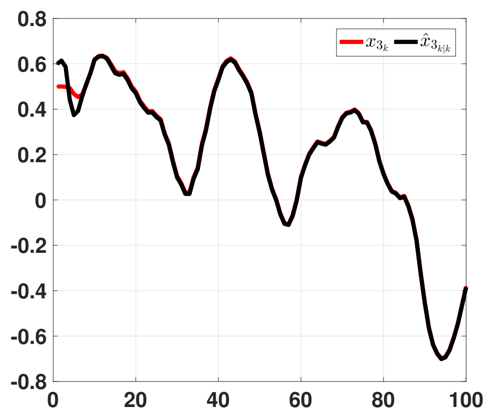
5.2.2.1 Extended Kalman Filter

Simulations are performed with EKF as the identification strategy in the proposed framework and results are presented in this section. Three cases are simulated i.e, $\alpha = 0$ corresponds to the identification of the parameter, $\alpha = 1$ corresponds to solving a classical optimal control problem in order to achieve some desired control performances and $\alpha = 0.1$ corresponding to the case, where one want to have both good identification and control performance. Here, $\alpha = 0.1$ means that more weight is given to the identification cost as compared to the process cost.

5.2.2.1.1 Case 1: Identification Experiment ($\alpha = 0$) This case corresponds to the information retrieval only. It is important to keep the robot body in a specified bound $[-1rad, +1rad]$ when the information is retrieved. The constants $\rho_l = 10^{-2}$, $\gamma_l = 10^{-3}$, $\rho_N = 10^{-1}$, $\gamma_N = 10^{-2}$, and $\beta = 10^{-4}$ are chosen to improve the performance of the designed controller. The initial condition for the true system is $x_0 = [1\ 0\ 0.5\ 0]^{\mathsf{T}}$. The a-priori estimates are $\hat{x}_{0|-1} = x_0 + [0.1\ 0.1\ 0.1\ 0.1]^{\mathsf{T}}$ and $\hat{\theta}_{0|-1} = M_p + 0.5$.

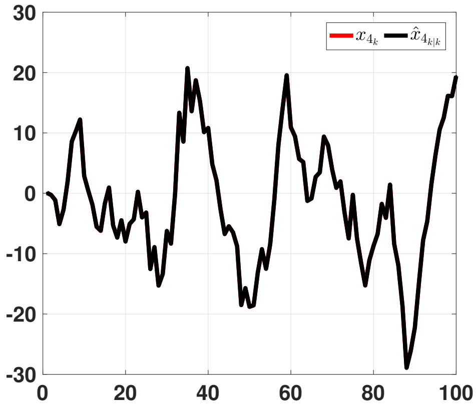


(a) Linear Velocity

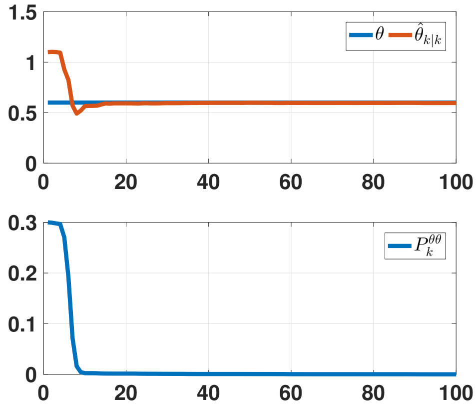


(b) Angular Position

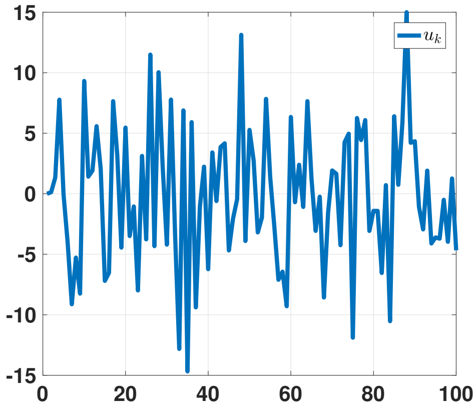
It is evident from the Fig. 5.15 the random behavior of the system states as the purpose of the OID design is just to identify the parameter. As expected, a random input signal is generated just to identify the information from the system as shown in Fig. 5.15(e). In Fig. 5.15(a), the linear velocity is shown for both true and estimated value of the system trajectories. x_{2_k} corresponds to the true value of the state while the $\hat{x}_{2_{k|k}}$ correspond to the estimated value of the state. It can be seen that the states x_{2_k} and x_{3_k} are



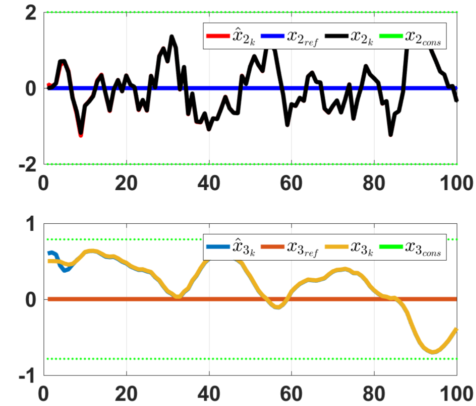
(c) Angular Velocity



(d) Identified Parameter & Covariance Matrix



(e) Optimal Control Signal

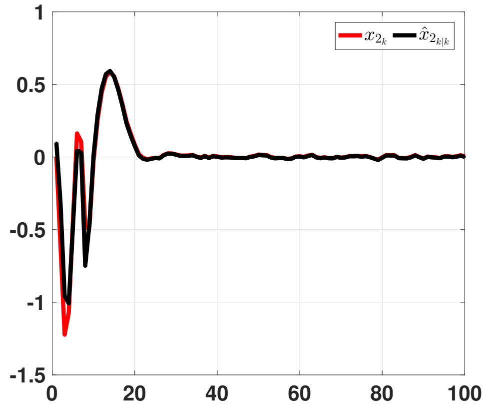


(f) Constraints on x_{2_k} and x_{3_k}

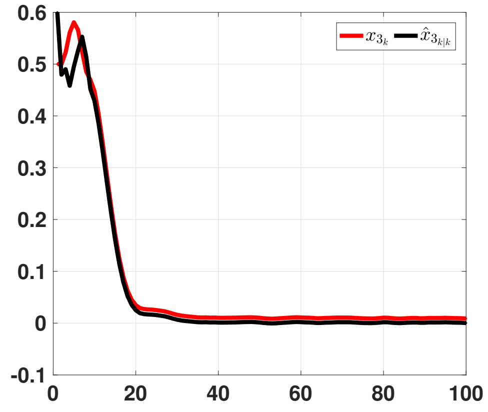
FIGURE 5.15: EKF: Parameter Identification $\alpha = 0$

estimated perfectly as shown in Fig. 5.15(b) while the state x_{4_k} is measured hence a perfect estimate is shown in Fig. 5.15(c). The constraints on state $x_{2_k} \leq 2^m/_s$ and $x_{3_k} \leq \pm^{\pi}/_4$ are not violated which shows that the designed control has respected the constraints as shown in Fig. 5.15(f). The identified parameter and the covariance matrix $P_k^{\theta\theta}$ is shown in Fig. 5.15(d). The parameter is estimated perfectly and the covariance matrix related to the identified parameter is minimized which is used as the information cost. The optimal input is visible in Fig. 5.15(e) which is as expected a random signal.

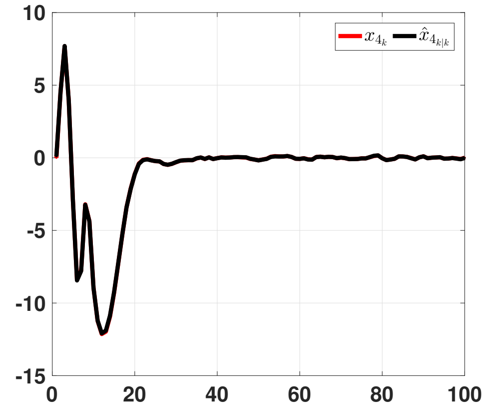
5.2.2.1.2 Case 2: Classical Optimal Control Problem $(\alpha = 1)$ In this case, the aim of the OID design is to stabilize the robot vertically upward and bring the robot velocity to zero. The simulations are performed with the same initial conditions for both true and estimated value of the system and similar value of the parameters are used except the value of the $\gamma_l = 10^{-5}$ and $\gamma_N = 10^{-6}$ which corresponds to a strict penalty function in the previous case.



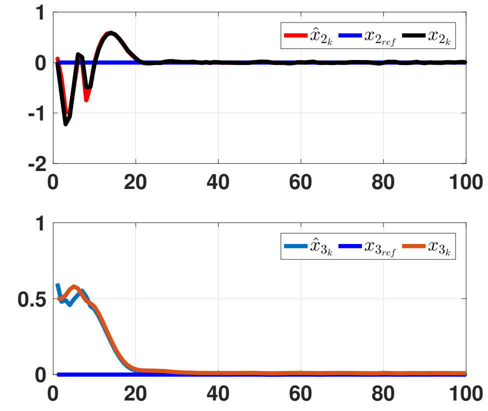
(a) Linear Velocity



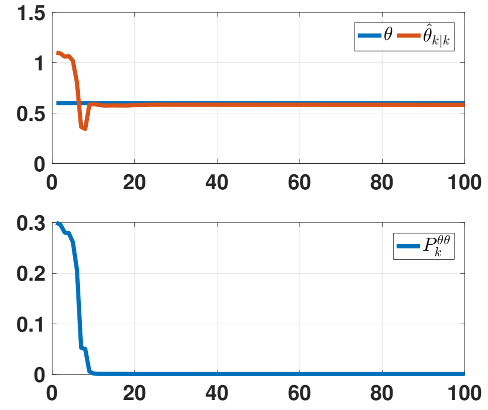
(b) Angular Position



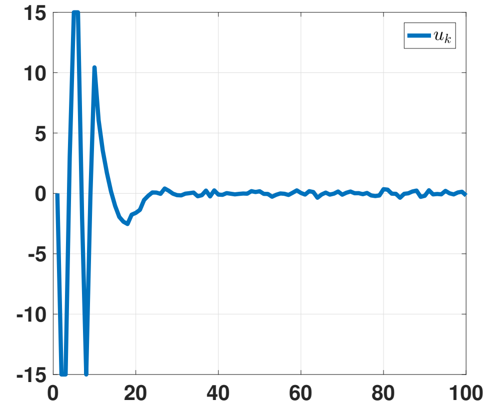
(c) Angular Velocity



(d) Reference Tracking



(e) Identified Parameter & Covariance Matrix



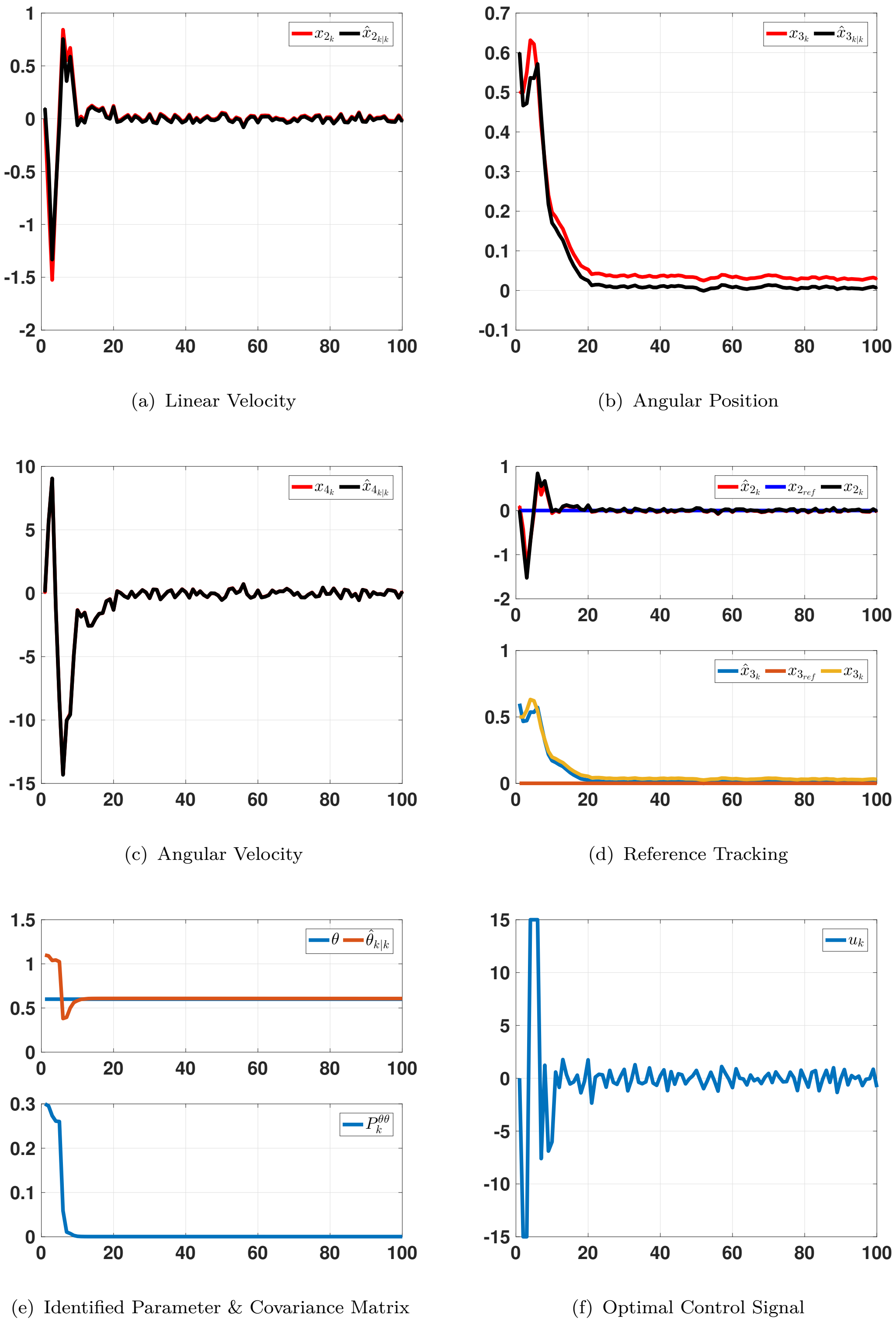
(f) Optimal Control Signal

FIGURE 5.16: EKF: Classical Optimal Control Problem $\alpha = 1$

In Fig. 5.16, all the system states are shown for the case of $\alpha = 1$. The robot is stabilized vertically upward starting from an initial value of $0.5rad$ as shown in Fig. 5.16(b) and as the robot stabilize itself, the linear velocity is also brought to zero as shown in Fig. 5.16(a). The angular velocity also become zero as the robot stabilizes itself as shown in Fig. 5.16(c). A particular case of regulation is shown in Fig. 5.16(d), it is visible that the linear velocity of the robot is brought to zero as the robot stabilizes itself vertically upward. As the control objective is to stabilize the robot vertically upward, the parameter identification is subjected to a constant biasing. Starting from the estimated value, the estimate converges towards the true parameter rapidly but a constant biasing remains near the actual value of the parameter and the steady state error persist. Here, convergence is somewhat slow but the estimation is perfect without biasing. The optimal control signal is shown in Fig. 5.16(f) which indicates that in the start in order to achieve the desired control performance, the control has a high value but as the desired control specifications are achieved, the control efforts comes to zero.

5.2.2.1.3 Case 3: Trade-off between Identification and Control Performance $(\alpha = 0.1)$ In this case, the aim of the OID design is to have a trade-off between the parameter identification and the desired control performance. It is desirable to identify the parameter as well as stabilize the robot vertically upward. Simulations are performed with the same initial conditions for both true and estimated system and similar values of the parameters are used.

It is shown in Fig. 5.17 that both identification and control performance are achieved. The trade-off parameter $(\alpha = 0.1)$ is selected to weight the identification cost more as compared to the process cost. It is shown in Fig. 5.17(a) that as the robot is stabilized vertically upward, the linear velocity is brought to zero and the estimate of the state x_{2_k} is also perfect. The estimate of the pitch angle x_{3_k} is poor as shown in Fig. 5.17(b) which is due to the trade-off between the identification of the parameter and state. The robot is brought to its vertical position starting from an initial value as shown in Fig. 5.17(d). It is also shown that the linear velocity is also zero when the robot is stabilized vertically upward. The identification of the parameter and the covariance matrix related to the identified parameter is shown in Fig. 5.17(e), which indicates a perfect identification of the parameter. The control sequence is shown in Fig. 5.17(f).



(a) Linear Velocity

(b) Angular Position

(c) Angular Velocity

(d) Reference Tracking

(e) Identified Parameter & Covariance Matrix

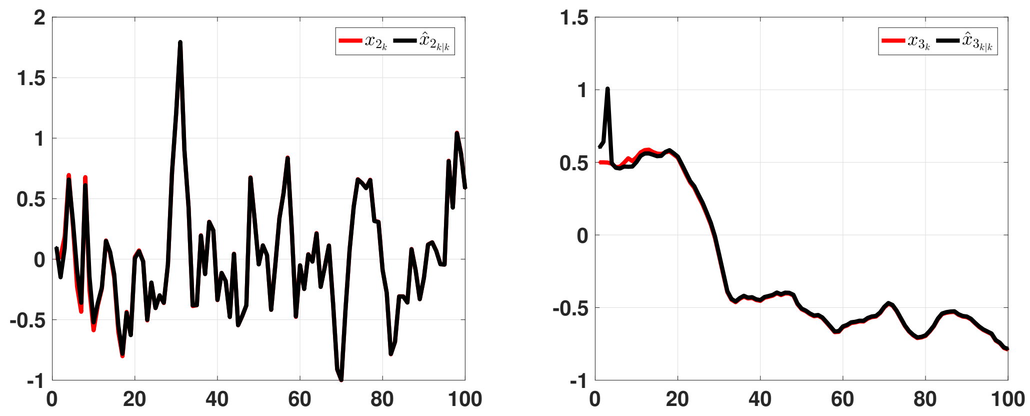
(f) Optimal Control Signal

FIGURE 5.17: EKF: Trade-off between Identification and Control Performance $\alpha = 0.1$

5.2.2.2 Unscented Kalman Filter

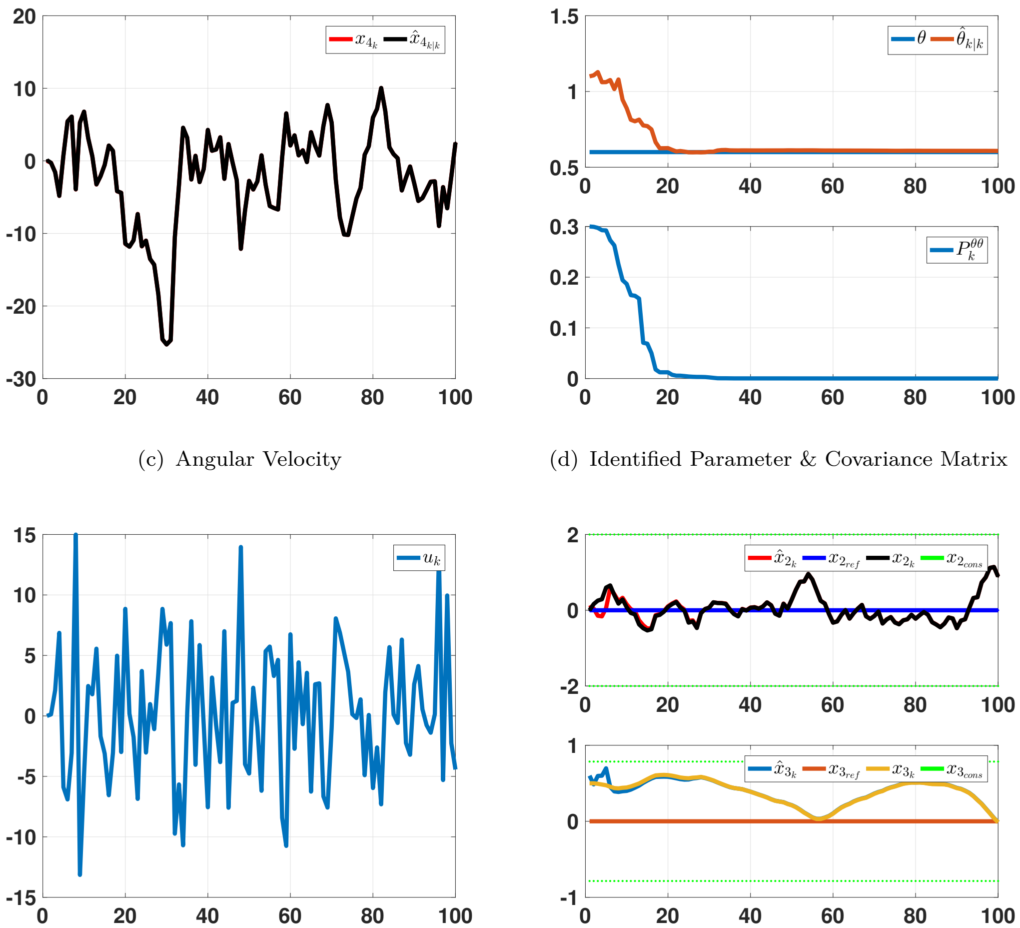
In order to have a comparison of the results in terms of identification of the parameter and control performance, we have performed simulations with unscented Kalman Filter. UKF is used as the identification strategy in the proposed framework and results are presented in this section. Simulations are performed for same three cases i.e, $\alpha = 0$ corresponds to the identification of the parameter, $\alpha = 1$ corresponds to solve a classical optimal control problem in order to achieve some desired control performance and $\alpha = 0.1$ corresponds to the case, where we want to have both good identification and control performance.

5.2.2.2.1 Case 1: Identification Experiment $(\alpha = 0)$ We have performed the simulations with the same initial conditions and the control objective is to identify the parameter. In order to make a comparison, we have analyzed the errors in identification of the parameter and achieved control performance which is presented in later sections.



(a) Linear Velocity (b) Angular Position

In Fig. 5.18(a), it is shown that in order to identify the parameter, the robot is moved randomly which results in random linear velocity. It can be seen that the linear velocity is estimated perfectly. Similarly the angle of the robot body x_{3_k} is estimated perfectly as shown in Fig. 5.18(b). The identified parameter and its covariance is shown in Fig. 5.18(d) which shows a perfect estimate of the parameter but as compared to the results obtained in EKF, it is little slow in terms of number of iterations. The constraints are satisfied here as well as shown in Fig. 5.18(f) and the random sequence of control is



(c) Angular Velocity

(d) Identified Parameter & Covariance Matrix

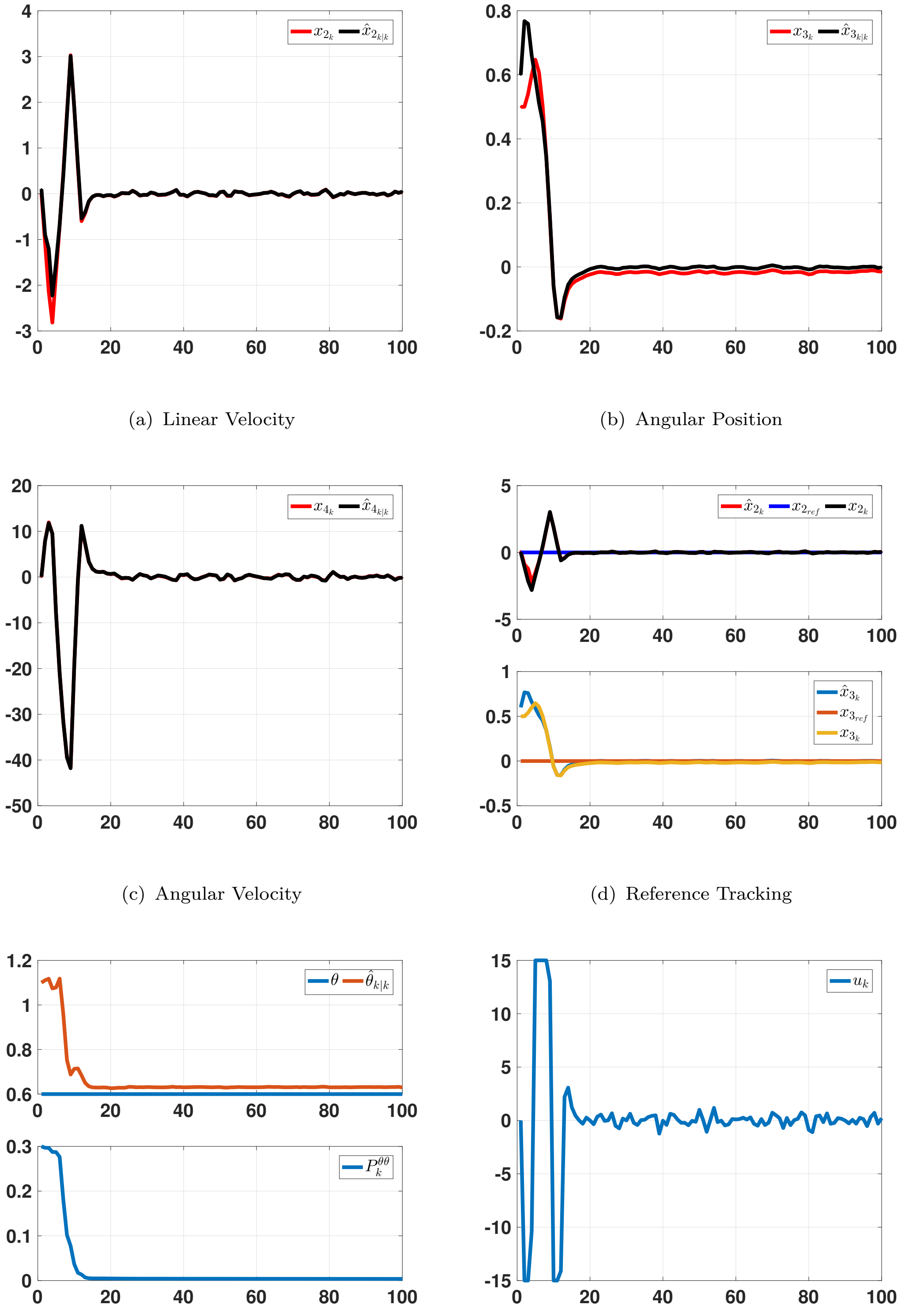
(e) Optimal Control Signal

(f) Constraints on x_{2_k} and x_{3_k}

FIGURE 5.18: UKF: Parameter Identification $\alpha = 0$

shown in Fig. 5.18(e). The mean square error comparison between the EKF and UKF results is presented in later section.

5.2.2.2.2 Case 2: Classical Optimal Control Problem $(\alpha = 1)$ In this case, the control performance is evaluated on the basis of reference tracking by the proposed framework. It is desirable to bring the linear velocity to zero and stabilize the robot body to vertically upward position. Simulations are performed with same initial conditions.



(a) Linear Velocity

(b) Angular Position

(c) Angular Velocity

(d) Reference Tracking

(e) Identified Parameter & Covariance Matrix

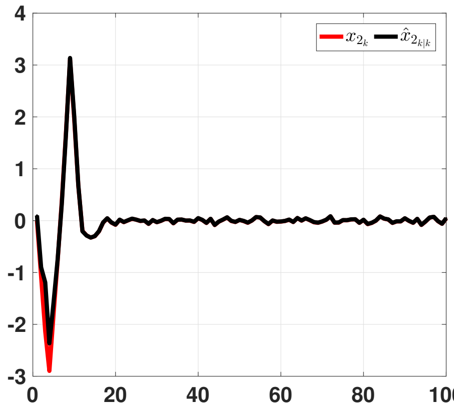
(f) Optimal Control Signal

Figure 5.19: UKF: Classical Optimal Control Problem $\alpha = 1$

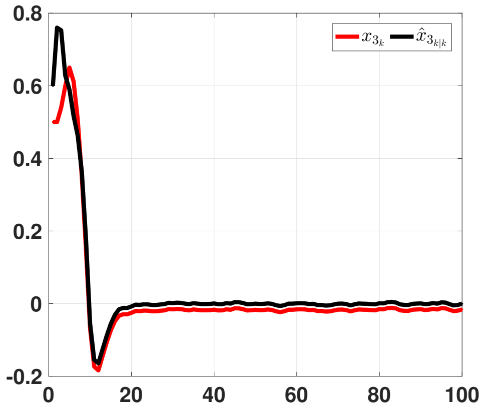
The linear velocity x_{2_k} is shown in Fig. 5.19(a) which indicates perfect estimation. It is brought to zero again as the robot stabilize itself as shown in Fig. 5.19(d). The parameter is identified with a constant biasing (poor identification) and covariance matrix related to identified parameter is not minimized. The sequence of control is given in Fig. 5.19(f) which shows that the control effort is brought to zero as the robot is stabilized vertically upward. The error comparison is given in later section.

5.2.2.2.3 Case 3: Trade-off between Identification and Control Performance ($\alpha = 0.1$)

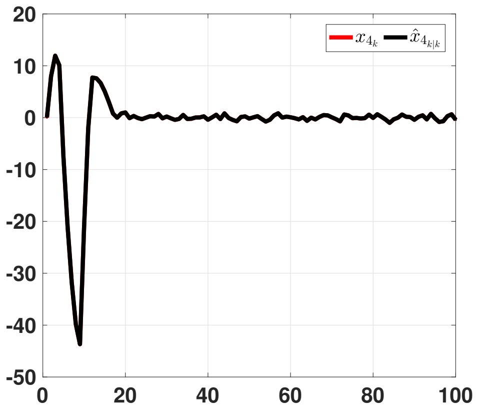
In this case, the aim of the OID design is to have a trade-off between the parameter identification and the desired control performance. Simulations are performed with similar initial conditions.



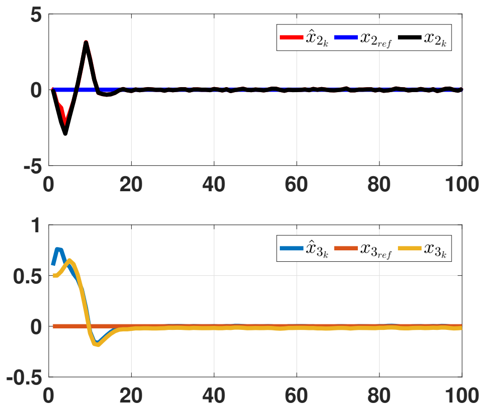
(a) Linear Velocity



(b) Angular Position

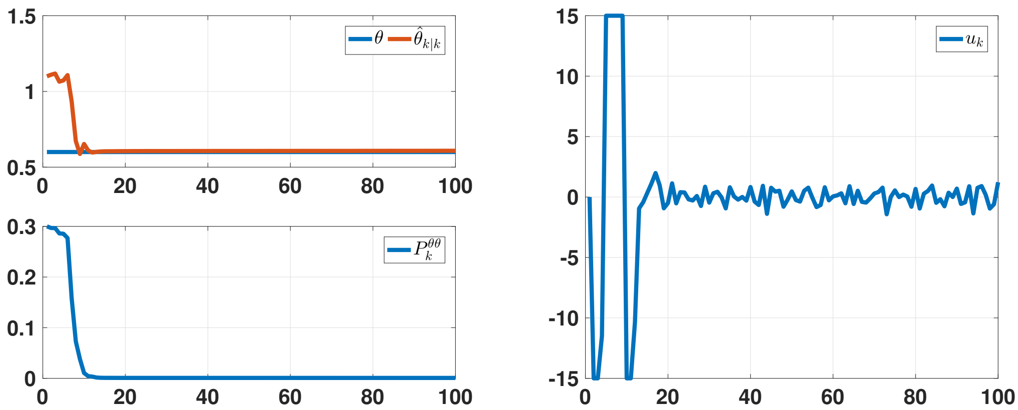


(c) Angular Velocity



(d) Reference Tracking

The results obtained in this case is similar to the one in EKF case. The linear velocity x_{2_k}, pitch angle of the robot body x_{3_k} and the angular velocity x_{4_k} are shown in



(e) Identified Parameter & Covariance Matrix

(f) Optimal Control Signal

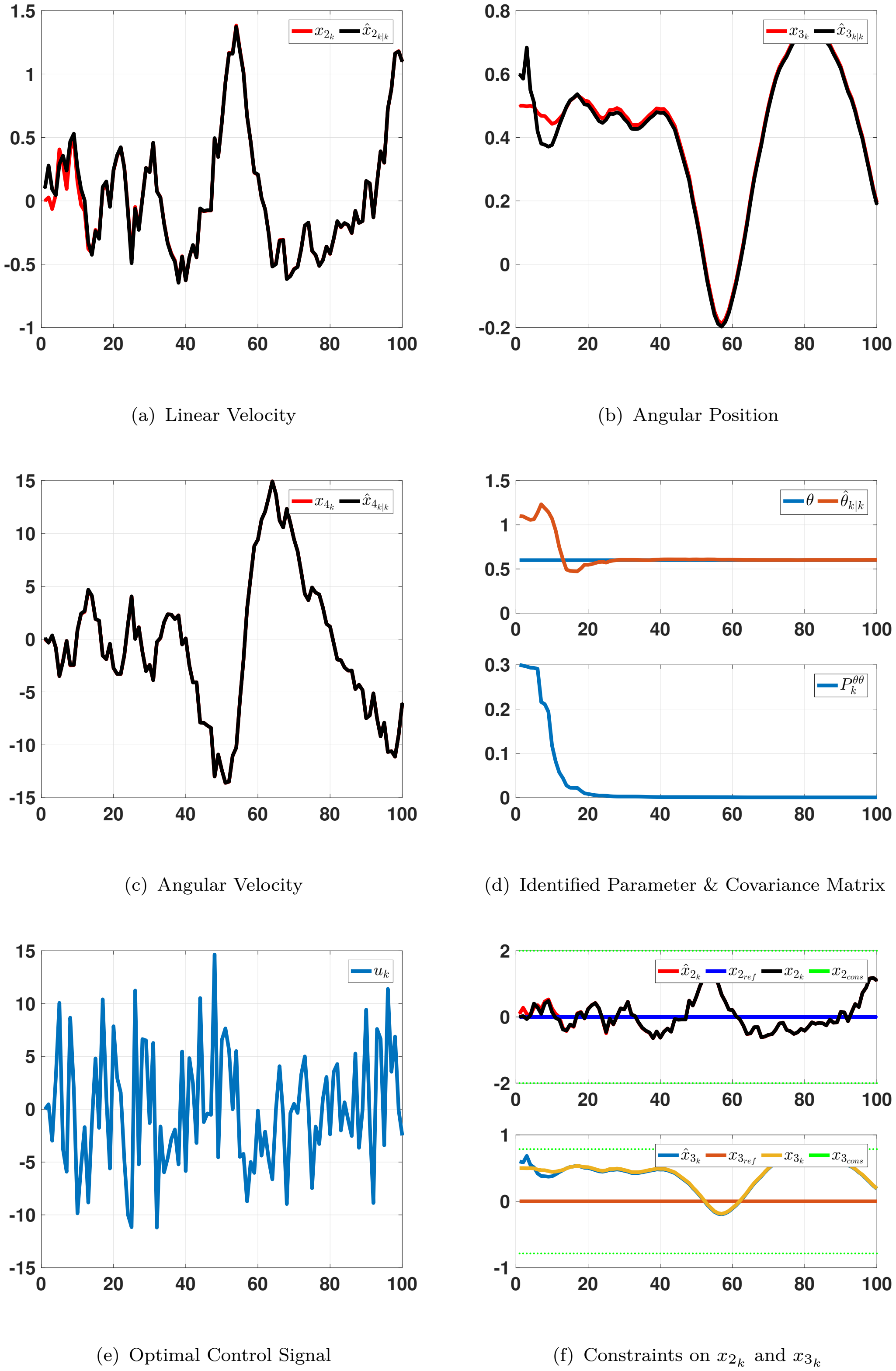
FIGURE 5.20: UKF: Trade-off between Identification and Control Performance $\alpha = 0.1$

Fig. 5.20(a), Fig. 5.20(b) and Fig. 5.20(c) respectively which represents that these states are estimated perfectly. The pitch angle is brought to zero (vertical upward position) which was the desired control objective as shown in Fig. 5.20(d). The perfectly identified parameter and the covariance matrix related to the identified parameter is shown in Fig. 5.20(e). The optimal control signal is given in Fig. 5.20(f).

5.2.2.3 Gaussian Sum Filter

In order to have a comparison of the results in terms of identification of the parameter and control performance, we have performed simulations with Gaussian sum filter as well. GSF is used as the identification strategy in the proposed framework and results are presented in this section. For the simulation purpose, $m = 10$ Gaussian densities are chosen to compute the estimate of the state and the covariance matrix. The detail of the GSF algorithm is given in Section 3.2.3. The results are compared with the one obtained in EKF and UKF case.

5.2.2.3.1 Case 1: Identification Experiment $(\alpha = 0)$ In this case, we have performed the simulations with the same initial conditions as done in case of EKF and UKF and the control objective is to identify the parameter. In order to make a comparison, we have analyzed the error in identification of the parameter and achieved control performance which is presented in later section.



(a) Linear Velocity (b) Angular Position

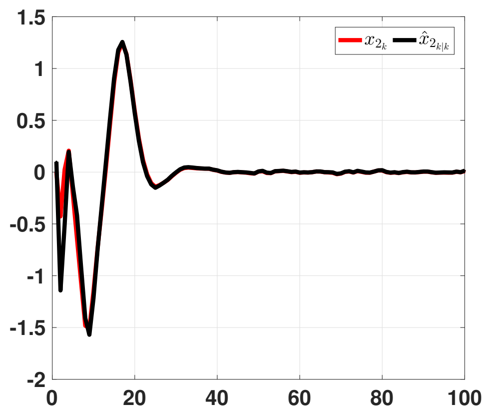
(c) Angular Velocity (d) Identified Parameter & Covariance Matrix

(e) Optimal Control Signal (f) Constraints on x_{2_k} and x_{3_k}

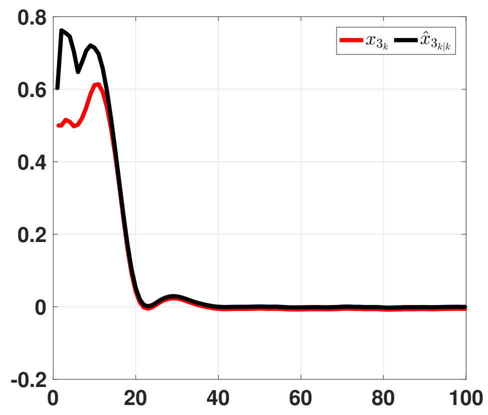
FIGURE 5.21: GSF: Parameter Identification $\alpha = 0$

In order to identify the parameter, a random sequence of control is generated as shown in Fig. 5.21(e) which results in random linear velocity as shown in Fig. 5.21(a). The state estimate in the case of linear velocity x_{2_k} is good but for the case of x_{3_k}, it is shown that the estimate is not good in the start. The identified parameter and its covariance is shown in Fig. 5.21(d) which shows a perfect estimate of the parameter. The comparison in terms of mean square error between the three identification strategies is presented in next section. The constraints are satisfied as shown in Fig. 5.21(f).

5.2.2.3.2 Case 2: Classical Optimal Control Problem $(\alpha = 1)$ In this case, we want to bring the robot to vertically upward position and also the linear velocity should be zero. It is desirable to achieve the reference tracking of the linear velocity and bring the robot body to vertically upward position from given initial condition and stabilize it there. Simulations are performed with same initial conditions.



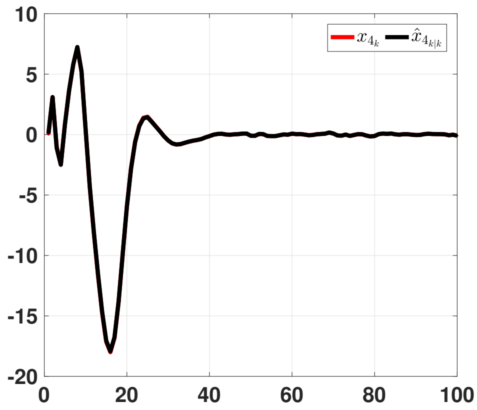
(a) Linear Velocity



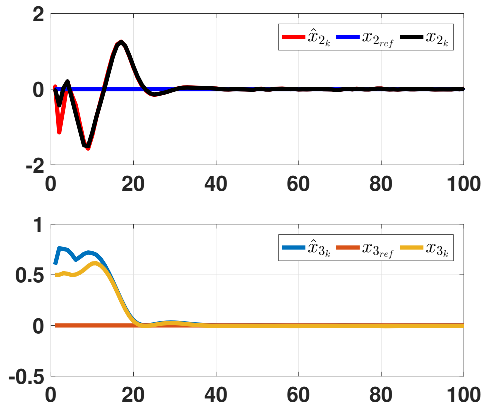
(b) Angular Position

The linear velocity x_{2_k} is shown in Fig. 5.19(a) which indicates perfect estimation. It is brought to zero as the robot stabilize itself as shown in Fig. 5.19(d). The parameter is identified with a constant biasing (poor identification) and covariance matrix related to identified parameter is not minimized. The sequence of control is given in Fig. 5.19(f) which shows that the control effort is brought to zero as the robot is stabilized vertically upward. The reference tracking of linear velocity $x_{2_k} = 0$ and the angle $x_{3_k} = 0$ is shown in Fig. 5.22(d). The error comparison is given in later section.

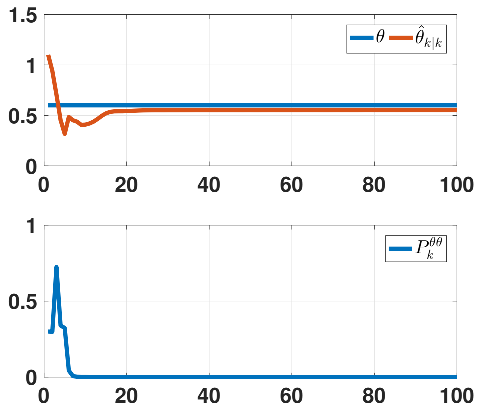
5.2.2.3.3 Case 3: Trade-off between Identification and Control Performance $(\alpha = 0.1)$ In this case, we have performed the simulations to achieve both



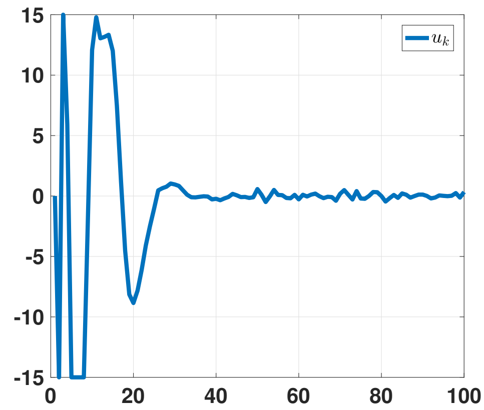
(c) Angular Velocity



(d) Reference Tracking



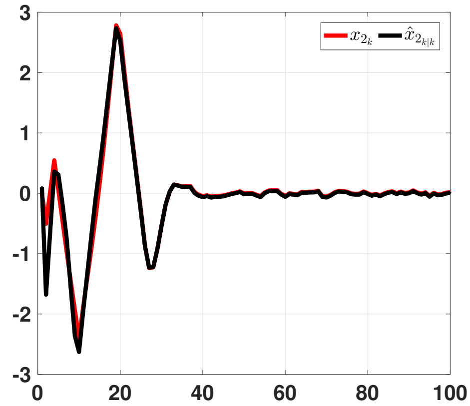
(e) Identified Parameter & Covariance Matrix



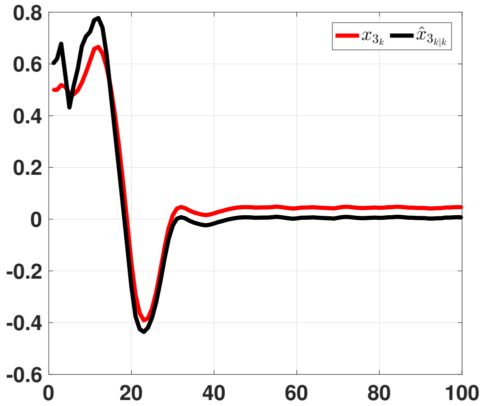
(f) Optimal Control Signal

FIGURE 5.22: GSF: Classical Optimal Control Problem $\alpha = 1$

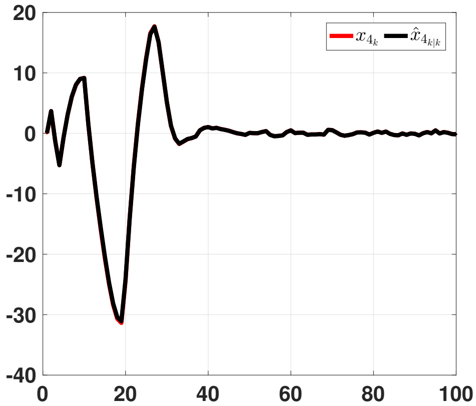
system identification as well as the control performance in terms of tracking the desired reference signals. The results are compared with the one obtained in case of EKF and UKF. Simulations are performed with similar initial conditions.



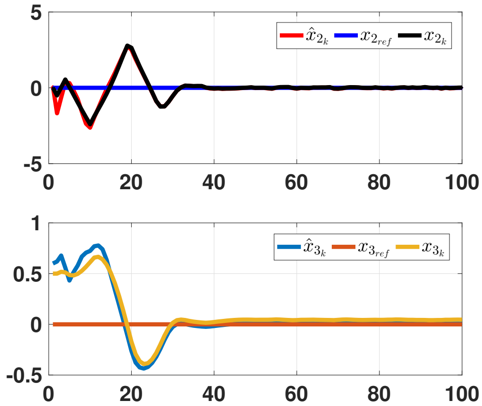
(a) Linear Velocity



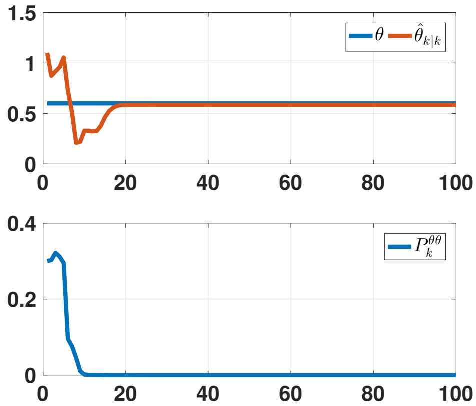
(b) Angular Position



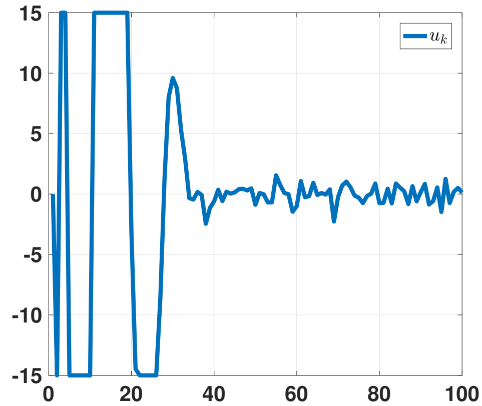
(c) Angular Velocity



(d) Reference Tracking



(e) Identified Parameter & Covariance Matrix



(f) Optimal Control Signal

Figure 5.23: GSF: Trade-off between Identification and Control Performance $\alpha = 0.1$

The linear velocity x_{2_k} is estimated perfectly as shown in Fig. 5.23(a) while the pitch angle of the robot body x_{3_k} is estimated poorly in the start as shown in Fig. 5.23(b). The pitch angle is brought to zero (vertical upward position) which was the desired control objective as shown in Fig. 5.23(d) and also the linear velocity x_{2_k} is brought to zero. The identified parameter and the covariance matrix related to the identified parameter is shown in Fig. 5.23(e) which indicates that parameter is identified perfectly without biasing. The optimal control signal is given in Fig. 5.23(f).

5.2.3 Mean Square Error Comparison: EKF, UKF and GSF

In order to evaluate the performance of the three cases, the error in parameter identification and error in reference tracking of the desired signal is presented here. The mean square error values computed for all three estimation strategies and then compared to show that EKF performs better than the UKF and GSF.

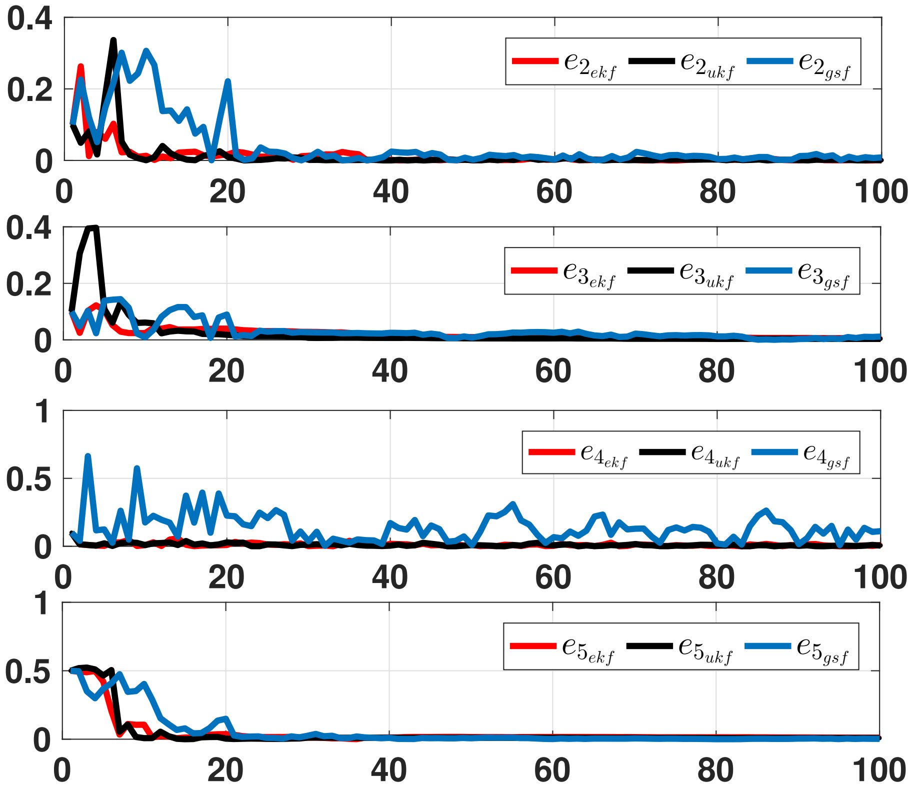


FIGURE 5.24: Case $\alpha = 0$: Error in Estimation of the State

In Fig. 5.24, the error in the estimation of state and parameter is shown. The error $e_{2_{ekf}}$ represents the error computed in state x_{2_k} which is given as $x_{2_k} - \hat{x}_{2_{k|k}}$ using the EKF strategy, while $e_{2_{ukf}}$ represents the error in state $x_{2_k} - \hat{x}_{2_{k|k}}$ using the UKF algorithm. Similarly, $e_{2_{gsf}}$ represents the error in state $x_{2_k} - \hat{x}_{2_{k|k}}$ using the GSF method. For all the other states, x_{3_k}, x_{4_k} and x_{5_k}, the error representation remain the same. It is obvious in the estimation of the state that error due to EKF is reduced to zero quicker than the other two strategies and also the parameter estimated using the EKF is better as the error is minimized faster. The performance of GSF is poor as compared to other two strategies.

For the case of $\alpha = 1$, error is calculated for both estimation of the state and parameter as well as error in reference tracking of the desired signals.

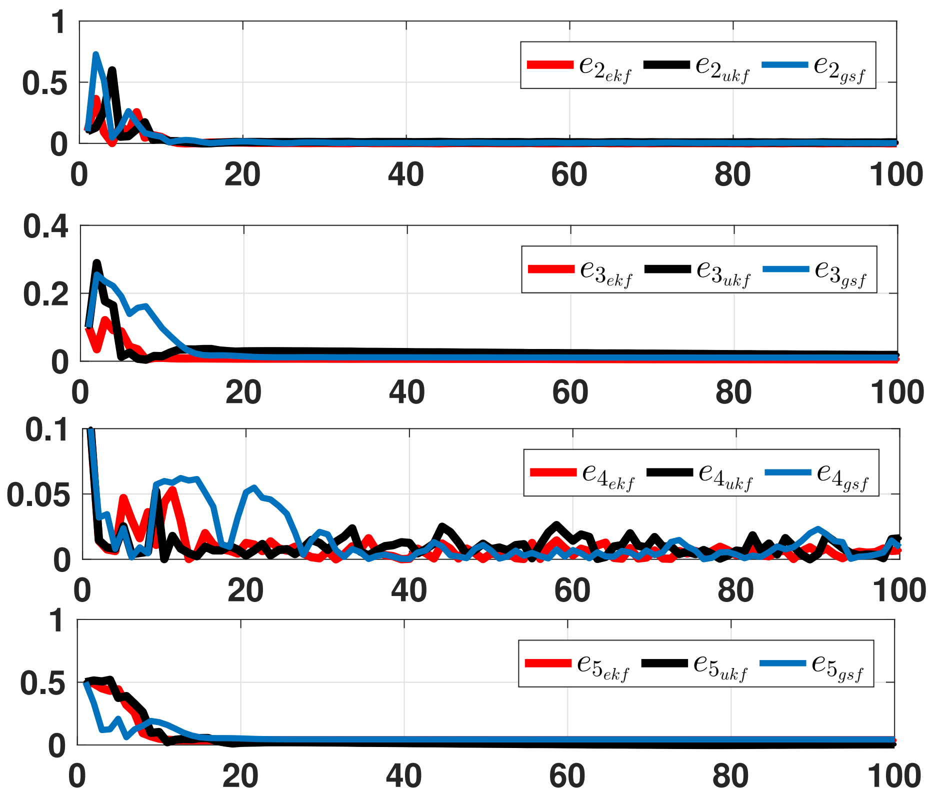


FIGURE 5.25: Case $\alpha = 1$: Error in Estimation of the State

It is shown in Fig. 5.25 that the error is minimized faster in case of EKF for all the three states x_{2_k}, x_{3_k} and x_{4_k}. The error in parameter identification is also better in case of EKF, while the GSF has the worst result.

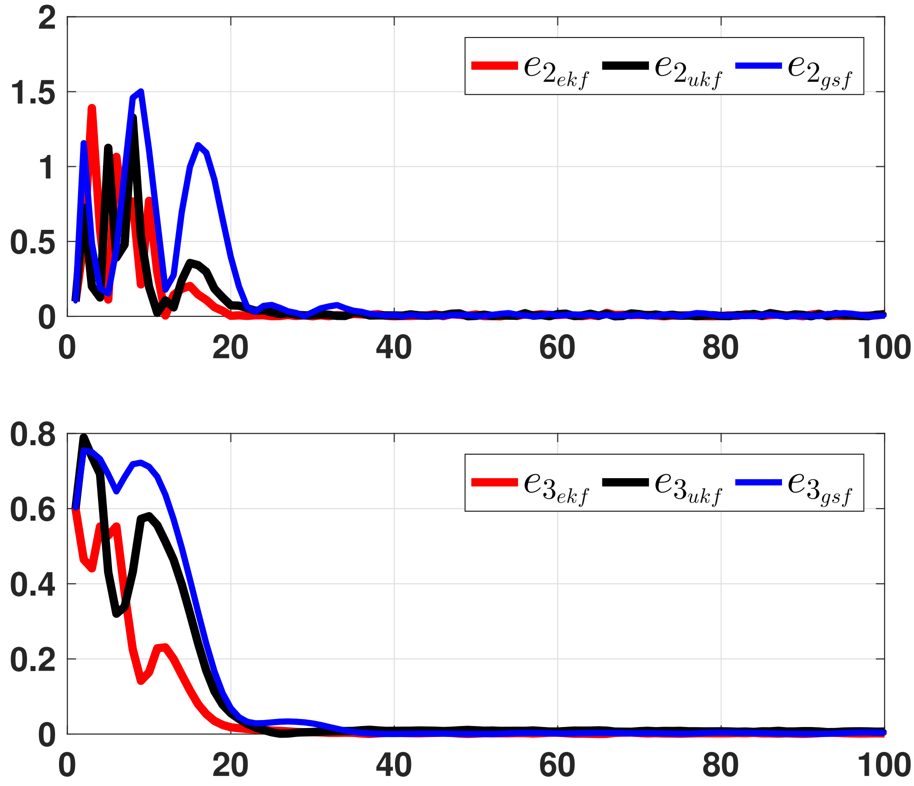


FIGURE 5.26: Case $\alpha = 1$: Error in Reference Tracking

It is also evident from the Fig. 5.26 that the error in reference tracking of the estimated value of the state and the desired reference signal is minimized better in case of EKF while GSF performs poorly. The comparison of mean square error for all the three cases is given in tables below.

	EKF		
	Error: Parameter Identification	**Error: Reference Tracking**	
	e_{5_k}	e_{2_k}	e_{3_k}
$\alpha = 0$	$1.53e^{-04}$	$1.69e^{-02}$	$1.89e^{-02}$
$\alpha = 0.5$	$1.31e^{-03}$	$1.87e^{-02}$	$2.43e^{-02}$
$\alpha = 1$	$1.59e^{-02}$	$4.82e^{-03}$	$4.43e^{-03}$

TABLE 5.3: Mean Square Error: EKF

	UKF		
	Error: Parameter Identification	Error: Reference Tracking	
	e_{5_k}	e_{2_k}	e_{3_k}
$\alpha = 0$	$4.32e^{-03}$	$2.93e^{-02}$	$3.53e^{-02}$
$\alpha = 0.5$	$1.60e^{-03}$	$1.16e^{-02}$	$1.02e^{-02}$
$\alpha = 1$	$2.34e^{-02}$	$4.21e^{-03}$	$4.01e^{-03}$

Table 5.4: Mean Square Error: UKF

	GSF		
	Error: Parameter Identification	Error: Reference Tracking	
	e_{5_k}	e_{2_k}	e_{3_k}
$\alpha = 0$	$5.74e^{-02}$	$1.64e^{-01}$	$3.50e^{-01}$
$\alpha = 0.5$	$5.89e^{-02}$	$1.86e^{-01}$	$8.06e^{-02}$
$\alpha = 1$	$6.66e^{-02}$	$1.63e^{-01}$	$1.13e^{-01}$

Table 5.5: Mean Square Error: GSF

5.3 Chapter Summary

In order to see objectivity and applicability of the proposed OID framework for active parameter identification (presented in Chapter 4), we performed some simulations with different numerical examples. This chapter presented all the numerical results and helped the reader to understand the concepts presented in previous chapters of the dissertation. A simple abstract example of a *toy model* was used to understand the objective of the proposed framework. Later, a complex model of 3-DOF two-wheeled mobile robot was chosen to show the performance of the proposed scheme on more complex systems. For the identification purpose, EKF was used as the identification strategy in the proposed framework and results obtained in simulations shows the effectiveness of the framework. For the purpose of comparison, we simulated the 2-DOF robot model with two different identification strategies, UKF and GSF. Simulation results were presented and a mean square error comparison was made. The results shows that EKF performs better than the other two strategies but both UKF or GSF can be used for more complex systems where EKF performance degrade.

Chapter 6

CONCLUSIONS, RECOMMENDATIONS AND FUTURE WORK

> *"We have the duty of formulating, of summarizing, and of communicating our conclusions, in intelligible form, in recognition of the right of other free minds to utilize them in making their own decisions."*
>
> – Ronald Fisher

In this work, the problem of OID for active parameter identification of nonlinear dynamic system was addressed. The major focus of the work was on the design of such an optimal excitation signals, that can yield maximal information from the unknown or uncertain system and achieve some desired control performance. This chapter summarizes the main contribution of the proposed work, some useful recommendations to address those issues which were not done effectively and some suggestions for future work.

6.1 Conclusions

In this dissertation, the problem of OID for the active parameter identification of nonlinear dynamic system was addressed as a novel strategy which combines the identification strategy with model predictive control in a receding horizon framework. The problem has been formulated in a combined EKF/NMPC framework (for comparison: UKF and GSF also used), where EKF was used for system identification of nonlinear dynamic system and on the basis of the identified information, an OID problem was solved in

NMPC framework. A-optimality criterion was proposed as a measure of information on the unknown parameter which tends to minimize the *trace* of the covariance matrix related to the identified parameter as the information cost. The problem was formulated in a receding horizon context, where a trade-off parameter α was introduced to weight between the *process cost* and *information cost*. The proposed strategy was implemented on different numerical examples and the results shown the effectiveness of the proposed methodology.

In order to make a comparison of proposed strategy in terms of parameter identification, we have used UKF and GSF as the identification strategy. The choice is motivated due to the limitations of EKF for highly nonlinear and complex systems. The results obtained with EKF provided the best identification of the parameter while UKF performance was better than the GSF. As the complexity of these examples were not very high, EKF performs exceptionally well than the other two. The results were presented and mean square error comparison is also given to enlighten the performance in terms of identification and achieving the desired control performance.

The numerical examples presented in this dissertation were chosen to show the effectiveness and objectivity of the proposed work. The problem is addressed in two different ways, i.e, active parameter identification ($\alpha = 0$) and classical optimal control problem ($\alpha = 1$) by acting on the trade-off parameter. In order to have a greater understanding of the proposed scheme, a simple abstract example of a toy model was chosen. Simulations were performed for different initial conditions under different scenarios and the results showed the effectiveness of the proposed framework. It was also desirable to see the performance of the system on a more complex example like 3-DOF two-wheeled mobile robot model. The performance of the proposed framework was considered in three cases by altering the value of the trade-off parameter and results were presented to show the effectiveness of the proposed framework. For the comparison of three strategies (EKF, UKF and GSF), we used 2-DOF model of two-wheeled robot which was subjected to only linear motion. The performance of the three cases were presented and the superiority of the EKF over the other two strategies was shown in the results.

6.2 Recommendations

After the successful completion of this thesis, some recommendations are made which must be considered in future work.

- The proposed algorithm has applications in *map building* of unknown environment, *fault detection* in different real life systems, identification and control of uncertain disease in living beings along with robotic applications.
- The problem should be formulated carefully as it can effect the system performance.
- The cost related to the system performance should be defined precisely.
- The performance of the dynamic system is improved by improving the accuracy of the dynamic model.
- The constraints and bounds on the input and states of the system are of critical importance. Hence, it should be defined with lot of care and knowledge on the system.
- The computational cost must be minimized by using some parallel processing architectures.

6.3 Open Problems

After going through some hard efforts to produce this research work, there are some areas which are still open and need future attention.

- The first and obvious future work is to perform the implementation of the proposed work on a real application like two-wheeled robot.
- It will be interesting to verify the performance of the proposed work while identifying more than one parameter in the examples.
- For the implementation on real problems, the computational costs must be reduced. Some parallel processing strategies must be explored or by using some high speed processors.

- It will be interesting to see the performance of algorithm with some other identification strategy like particle filters. But for this case, the optimality criterion must be selected accordingly.

Appendix A

Kalman Filter

Kalman filter (KF) is an optimal recursive algorithm for the state estimation of linear systems. It is most widely used method for state estimation in control theory as it produces the optimal estimate of the unknown or uncertain system in a sense that the sum of the estimation error is minimized. The application of KF to different physical systems is addressed in [68–72]. Consider a general discrete-time linear system of the form

$$x_{k+1} = f_k x_k + g_k u_k + w_k \tag{A.1}$$

$$y_k = h_k x_k + v_k \tag{A.2}$$

where k is the current sampling index, $x_k \in \Re^{n_x}$ is the state vector, $u_k \in \Re^{n_u}$ is the control vector, and $y_k \in \Re^{n_y}$ is the output vector. The matrix $f_k \in \Re^{n_x \times n_x}$ is the feedback matrix, $g_k \in \Re^{n_x \times n_u}$ is the input matrix and $h_k \in \Re^{n_y \times n_x}$ is the output matrix. The terms w_k and v_k are process noise and observation noise with zero-mean and covariances Q_k and R_k respectively. The Kalman filter algorithm is divided in two steps: prediction and update. The prediction step can be written as follow:

$$\hat{x}_{k+1|k} = f_k \hat{x}_{k|k} + g_k u_k$$

$$P_{k+1|k} = f_k P_{k|k} f_k^{\mathsf{T}} + Q_k$$

The update equations are written as:

$$\hat{y}_{k+1|k} = h_{k+1}\hat{x}_{k+1|k}$$

$$V_{k+1} = y_{k+1} - \hat{y}_{k+1|k}$$

$$S_{k+1} = h_{k+1}P_{k+1|k}h_{k+1}^{\mathsf{T}} + R_{k+1}$$

$$W_{k+1} = P_{k+1|k}h_{k+1}^{\mathsf{T}}S_{k+1}^{-1}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + W_{k+1}V_{k+1} \tag{A.3}$$

$$P_{k+1|k+1} = (I - W_{k+1}h_{k+1})P_{k+1|k} \tag{A.4}$$

The eq. (A.3) and eq. (A.4) represents the updated state estimate and updated covariance matrix respectively. The one complete cycle of Kalman filter for linear systems is shown in Fig. A.1.

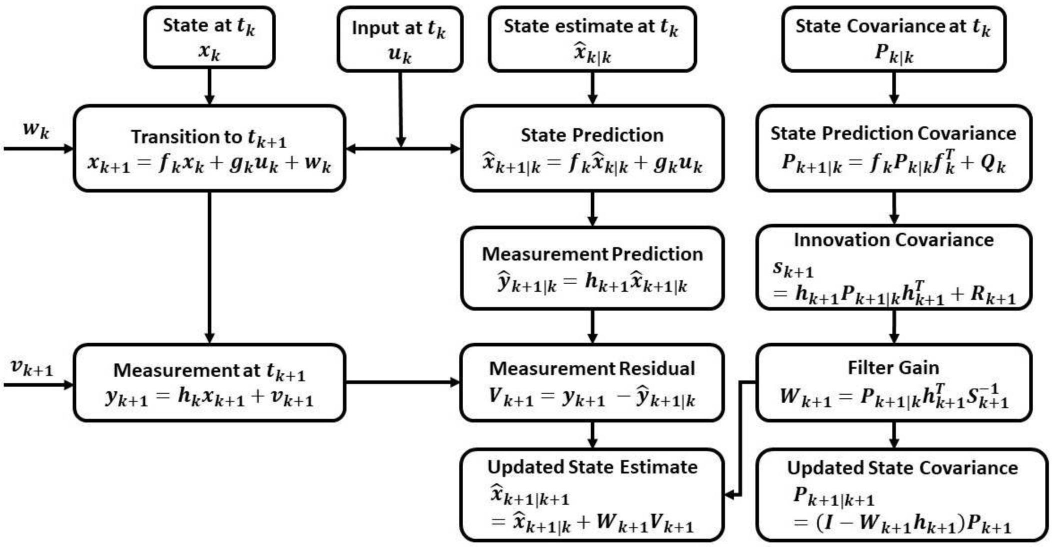


FIGURE A.1: Kalman Filter for Linear Systems (one cycle)

As the paper is focused on the nonlinear discrete-time systems, different variants of Kalman filter are used for the estimation of nonlinear systems.

Appendix B

Gaussian Sum Filter: Filtering Estimates

For the sake of simplicity, we are assuming that $X_k = [x_k, \theta]^\mathsf{T}$ represents the new augmented state vector. The two filtering estimate equations are given as

$$\hat{X}_{k|k} = \sum_{i=1}^{m} \alpha_{i,k} \hat{X}_{i,k|K} \tag{B.1}$$

$$P_{k|k} = \sum_{i=1}^{m} \alpha_{i,k} \{ P_{i,k|k} + (\hat{X}_{k|k} - \hat{X}_{i,k|k})(\hat{X}_{k|k} - \hat{X}_{i,k|k})^\mathsf{T} \} \tag{B.2}$$

The conditional density function of X_k computed at time k is given as sum of weighted Gaussian distribution densities as:

$$p(X_k|Z^k) = \sum_{i=1}^{m} \alpha_{i,k} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) \tag{B.3}$$

$$= \sum_{i=1}^{m} \alpha_{i,k} (2\pi)^{-n/2} \left| P_{i,k|k} \right|^{-1/2} exp\{-(1/2)(X_k - \hat{X}_{i,k|k})^\mathsf{T} P_{i,k|k}^{-1} (X_k - \hat{X}_{i,k|k})\} \tag{B.4}$$

Hence, the mean and covariance of x_k can be computed as:

$$
\begin{aligned}
\hat{X}_{k|k} &= \mathbb{E}(X_k|Z^k) \\
&= \int X_k p(X_k|Z^k) dX_k \\
&= \int X_k \sum_{i=1}^{m} \alpha_{i,k} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \\
&= \sum_{i=1}^{m} \alpha_{i,k} \int X_k \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \\
&= \sum_{i=1}^{m} \alpha_{i,k} \hat{X}_{i,k|k}
\end{aligned}
$$

The equation for the covariance matrix is given as

$$
\begin{aligned}
P_{k|k} &= \mathbb{E}\left[(X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^{\mathsf{T}}|Z^k\right] \\
&= \int (X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^{\mathsf{T}} p(X_k|Z^k) dX_k \\
&= \int (X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^{\mathsf{T}} \sum_{i=1}^{m} \alpha_{i,k} \mathcal{N}(X_k - \hat{X}_{k|k}, P_{i,k|k}) dX_k \\
&= \sum_{i=1}^{m} \alpha_{i,k} \int (X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^{\mathsf{T}} \mathcal{N}(X_k - \hat{X}_{k|k}, P_{i,k|k}) dX_k \\
&= \sum_{i=1}^{m} \alpha_{i,k} \int \left[(X_k - \hat{X}_{i,k|k}) + (\hat{X}_{i,k|k} - \hat{X}_{k|k})\right]\left[(X_k - \hat{X}_{i,k|k}) + (\hat{X}_{i,k|k} - \hat{X}_{k|k})\right]^{\mathsf{T}} \\
&\times \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \\
&= \sum_{i=1}^{m} \alpha_{i,k} \Big(\int (X_k - \hat{X}_{i,k|k})(X_k - \hat{X}_{i,k|k})^{\mathsf{T}} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \\
&+ \int (X_k - \hat{X}_{i,k|k})(\hat{X}_{i,k|k} - \hat{X}_{k|k})^{\mathsf{T}} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \\
&+ \int (\hat{X}_{i,k|k} - \hat{X}_{k|k})(X_k - \hat{X}_{i,k|k})^{\mathsf{T}} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \\
&+ \int (\hat{X}_{i,k|k} - \hat{X}_{k|k})(\hat{X}_{i,k|k} - \hat{X}_{k|k})^{\mathsf{T}} \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k \Big)
\end{aligned}
$$

Recall that,

$$
\int X_k \mathcal{N}(X_k - \hat{X}_{i,k|k}, P_{i,k|k}) dX_k = (\hat{X}_{i,k|k} - \hat{X}_{k|k})(\hat{X}_{i,k|k} - \hat{X}_{k|k})^{\mathsf{T}}
$$

Thus, one can write the final equation as

$$P_{k|k} = \sum_{i=1}^{m} \alpha_{i,k} \Big[P_{i,k|k} + (\hat{X}_{i,k|k} - \hat{X}_{k|k})(\hat{X}_{i,k|k} - \hat{X}_{k|k})^{\mathsf{T}} \Big] \tag{B.5}$$

Appendix C

Sequential Unconstrained Minimization Technique (SUMT)

Let the penalty parameter β_k, for $k = 1, 2, \ldots$ be the increasing sequence such that $\beta_k > 0$ and $\beta_{k+1} > \beta_k$. For every value of k, the problem is solved as

$$minimize \quad \{J_k^{cons}(\beta_k, x_k) : x_k \in \Re^{n_x}\} \qquad (C.1)$$

to obtain the optimum x_k which satisfies the penalty function. It is assumed that for all positive value of β_k, the problem given in *(C.1)* has a solution. The steps taken to solve the problem is given as:

- **Initialize the parameters:** Let at $k = 0$, the algorithm is initialized with the penalty parameter β_0, stopping parameter $\sigma > 0$ and growth parameter $\rho > 1$.

- **Iterate the process:** Let at $k = 1$, the cost $J_k^{cons}(\beta_{k-1}, x_{k-1})$ is minimized. Call the solution x_k and check the violation of the constraints.

- **Stopping Criteria:** If the constraints are violated and the distance between x_k and x_{k-1} is smaller than the stopping parameter σ, i.e, $\|x_{k-1} - x_k\| < \sigma$, stop the iteration with x_k be the optimal solution. Otherwise, update the value of penalty parameter $\beta_k = \rho\beta_{k-1}$ and repeat the iteration with x_k and $k = k + 1$.

REFERENCES

[1] M. Muhammad, S. Buyamin, and A. A. Bature, "Multiple operating points model-based control of a two-wheeled inverted pendulum mobile robot," in *International Journal of Mechanical & Mechatronics Engineering IJMME-IJENS Vol:13 No:05*, 2013.

[2] R. C. Ooi, *Balancing a Two-Wheeled Autonomous Robot.* Undergraduate honors thesis, The University of Western Australia, `http://robotics.ee.uwa.edu.au/theses/2003-Balance-Ooi.pdf/` 2003.

[3] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, jul 1948.

[4] G. N. Saridis, "Entropy formulation of optimal and adaptive control," *IEEE Transactions on Automatic Control*, vol. 33, pp. 713–721, Aug 1988.

[5] X. Feng, K. A. Loparo, and Y. Fang, "Optimal state estimation for stochastic systems: an information theoretic approach," *IEEE Transactions on Automatic Control*, vol. 42, pp. 771–785, Jun 1997.

[6] G. Goodwin and R. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, vol. 136. Academic Press, NY, 1 ed., 1977.

[7] L. Ljung, ed., *System Identification: Theory for the User.* Canada: Pearson Education, 1987.

[8] R. R. Pintelon, J. J. Schoukens, and Wiley InterScience (Online service), *System identification : a frequency domain approach.* Wiley, 2012.

[9] L. A. Zadeh, "From circuit theory to system theory," *Proceedings of the IRE*, vol. 50, pp. 856–865, May 1962.

[10] P. Eykhoff, *System Identification : Parameter and State Estimation.* Wiley, London, 1974.

[11] L. Ljung, "Convergence analysis of parametric identification methods," *IEEE Transactions on Automatic Control*, vol. 23, pp. 770–783, Oct 1978.

[12] L. Ljung, ed., *System Identification (2^{nd} Ed.): Theory for the User.* Upper Saddle River, NJ, USA: Prentice Hall PTR, 1999.

[13] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465 – 471, 1978.

[14] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Transactions on Information Theory*, vol. 44, pp. 2743–2760, Oct 1998.

[15] P. Zhang, "Model selection via multifold cross validation.," *The Annals of Statistics*, vol. 21, no. 1, pp. 299–313, 1993.

[16] G. Schwarz, "Information theory and an extension of the maximum likelihood principle," *Estimating the Dimension of a Model*, pp. 461–464, 1978.

[17] H. Akaike, *Information Theory and an Extension of the Maximum Likelihood Principle*, pp. 199–213. New York, NY: Springer New York, 1998.

[18] C. R. Rojas, J. S. Welsh, G. C. Goodwin, and A. Feuer, "Robust optimal experiment design for system identification," *Automatica*, vol. 43, no. 6, pp. 993 – 1008, 2007.

[19] A. M. Legendre, *New methods for the determination of comet orbits.* Firmin Didot Publishers, 1805.

[20] F. Galton, "Co-relations and their measurement, chiefly from anthropometric data.," *Proceedings of the Royal Society of London*, vol. 45, no. 1-4, pp. 135–145, 1888.

[21] H. T. Davis, *The Analysis of Economic Time Series.* The Principia Press, Inc., Bloomington, Indiana, 1st ed., 1941.

[22] E. J. Hannan, *Multiple Time Series.* John Wiley & Sons, Inc, New York, 1st ed., 1970.

[23] M. S. Morgan, *The History of Econometric Ideas.* Cambridge University Press, 1st ed., 1992.

[24] M. Deistler, *System Identification and Time Series Analysis: Past, Present, and Future*, pp. 97–109. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[25] H. B. Mann and A. Wald, "On the statistical treatment of linear stochastic difference equations.," *Econometrica*, vol. 11, no. 3/4, pp. 173–220, 1943.

[26] H. Trygve, "The probability approach in econometrics," *Econometrica*, vol. 12, pp. iii–115, 1944.

[27] T. Koopmans, ed., *Statistical Inference in Dynamic Economic Models.* Upper Saddle River, NJ, USA: John Wiley, New York., 1950.

[28] P. J. Daniell, "Discussion o symposium on auto correlation in time series.," *Journal of the Royal Statistical Society*, vol. 8, no. 1, pp. iii–115, 1946.

[29] R. Kalman, "Contributions to the theory of optimal control," 1960.

[30] R. Kalman, "A new approach to linear filtering and prediction problems," *ASME Journal of Basic Engineering*, vol. 82(1), pp. 35–45, 1960.

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

[32] K. J. Astrom and T. Bohlin, *Numerical Identification of Linear Dynamic Systems from Normal Operating Records*, pp. 96–111. Boston, MA: Springer US, 1966.

[33] E. J. Hannan, "The asymptotic theory of linear time-series models.," *Journal of Applied Probability*, vol. 10, no. 1, pp. 130 – 145, 1973.

[34] W. Dunsmuir and E. J. Hannan, "Vector linear time series models.," *Advances in Applied Probability*, vol. 8, no. 2, pp. 339 – 364, 1976.

[35] L. Lennart and P. E. Caines, "Asymptotic normality of prediction error estimators for approximate system models," *Stochastics*, vol. 3, no. 1-4, pp. 29–46, 1980.

[36] P. E. Caines, S. P. Meynt, and E. Moscas, "On the steady state behaviour of stochastic systems," *International Journal of Adaptive Control and Signal Processing*, vol. 2, no. 2, pp. 121–131, 1988.

[37] E. Hannan and M. Deistler, *The Statistical Theory of Linear Systems.* Society for Industrial and Applied Mathematics, 1988.

[38] G. C. Reinsel, *Elements of Multivariate Time Series Analysis.* Springer Series in Statistics, 1993.

[39] P. V. Overschee and B. de Moor, *Subspace Identification for Linear Systems: Theory - Implementation - Applications.* Springer, 1996.

[40] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach.* IEEE Press, New York, 2nd ed., 2001.

[41] J. Wolberg, *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments.* Springer, 2nd ed., 2006.

[42] T. Kailath, H. S. Ali, and B. Hassibi, *Linear Estimation.* Prentice-Hall, 2000, 1 ed., 2000.

[43] J. Aldrich, "R. a. fisher and the making of maximum likelihood 1912-1922," *Statistical Science*, vol. 12, no. 3, pp. 162–176, 1997.

[44] A. Hald, "On the history of maximum likelihood in relation to inverse probability and least squares," *Statist. Sci.*, vol. 14, pp. 214–222, 05 1999.

[45] A. N. Tikhonov and Y. A. Vasiliy, *Solutions of Ill-posed Problems.* V. H. Winston & Sons, Washington, 1 ed., 1977.

[46] B. Widrow and S. Sterns, *Adaptive Signal Processing.* Prentice Hall, Englewood Cliffs, NJ, 1 ed., 1985.

[47] S. Haykin, *Adaptive Filter Theory.* Prentice Hall, Englewood Cliffs, NJ, 2002.

[48] S. Haykin and B. W. (Eds.), *Least-Mean-Square Adaptive Filters.* Wiley, New York, 2003.

[49] S. Sherman, "Non-mean-square error criteria," *IRE Transactions on Information Theory*, vol. 4, pp. 125–126, September 1958.

[50] J. Brown, "Asymmetric non-mean-square error criteria," *IRE Transactions on Automatic Control*, vol. 7, pp. 64–66, January 1962.

[51] M. Zakai, "General error criteria (corresp.)," *IEEE Transactions on Information Theory*, vol. 10, pp. 94–95, Jan 1964.

[52] E. B. Hall and G. L. Wise, "On optimal estimation with respect to a large family of cost functions," *IEEE Transactions on Information Theory*, vol. 37, pp. 691–693, May 1991.

[53] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA, 1983.

[54] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Transactions on Information Theory*, vol. 30, pp. 275–283, March 1984.

[55] E. Walach, "On high-order error criteria for system identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, pp. 1634–1635, Dec 1985.

[56] S. C. Douglas and T. H. Y. Meng, "Stochastic gradient adaptation under general error criteria," *IEEE Transactions on Signal Processing*, vol. 42, pp. 1335–1351, Jun 1994.

[57] T. Y. Al-Naffouri and A. H. Sayed, "Adaptive filters with error nonlinearities: Mean-square analysis and optimum design," *EURASIP J. Appl. Signal Process.*, vol. 2001, pp. 192–205, Jan. 2001.

[58] S. C. Pei and C. C. Tseng, "Least Mean P-Power Error Criterion for Adaptive FIR Filter," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 9, pp. 1540–1547, 1994.

[59] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection.* John Wiley & Sons, Inc., New York, 1987.

[60] R. K. Boel, M. R. James, and I. R. Petersen, "Robustness and risk-sensitive filtering," *IEEE Transactions on Automatic Control*, vol. 47, pp. 451–461, Mar 2002.

[61] J. T. Lo and T. Wanner, "Existence and uniqueness of risk-sensitive estimates," *IEEE Transactions on Automatic Control*, vol. 47, pp. 1945–1948, Nov 2002.

[62] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides, "Least mean mixed-norm adaptive filtering," *Electronics Letters*, vol. 30, pp. 1574–1575, Sep 1994.

[63] O. Tanrikulu and J. A. Chambers, "Convergence and steady-state properties of the least-mean mixed-norm (LMMN) adaptive algorithm," *IEE Proceedings - Vision, Image and Signal Processing*, vol. 143, pp. 137–142, Jun 1996.

[64] J. Chambers and A. Avlonitis, "A robust mixed-norm adaptive filter algorithm," *IEEE Signal Processing Letters*, vol. 4, pp. 46–48, Feb 1997.

[65] A. N. Delopoulos and G. B. Giannakis, "Strongly consistent identification algorithms and noise insensitive mse criteria," *IEEE Transactions on Signal Processing*, vol. 40, pp. 1955–1970, Aug 1992.

[66] C.-Y. Chi, W.-J. Chang, and C.-C. Feng, "A new algorithm for the design of linear prediction error filters using cumulant-based mse criteria," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2876–2880, 1994.

[67] C.-C. Feng and C.-Y. Chi, "Design of wiener filters using a cumulant based mse criterion," *Signal Processing*, vol. 54, no. 1, pp. 23 – 48, 1996.

[68] K. Lin, Z. Guo, F. Yang, J. Huang, and Y. Zhang, "Kalman filter-based multi-object tracking algorithm by collaborative multi-feature," in *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 1239–1244, March 2017.

[69] Y. Jiang and Q. Hui, "Kalman filter with diffusion strategies for detecting power grid false data injection attacks," in *2017 IEEE International Conference on Electro Information Technology (EIT)*, pp. 254–259, May 2017.

[70] N. H. Nguyen and K. Dogançay, "Improved pseudolinear Kalman filter algorithms for bearings-only target tracking," *IEEE Transactions on Signal Processing*, vol. 65, pp. 6119–6134, Dec 2017.

[71] M. V. Kulikova and J. V. Tsyganova, "Improved discrete-time Kalman filtering within singular value decomposition," *IET Control Theory Applications*, vol. 11, no. 15, pp. 2412–2418, 2017.

[72] J. V. Tsyganova and M. V. Kulikova, "Svd-based Kalman filter derivative computation," *IEEE Transactions on Automatic Control*, vol. 62, pp. 4869–4875, Sept 2017.

[73] R. Ambrosino, F. Baselice, G. Ferraioli, and G. Schirinzi, "Extended Kalman filter for multichannel InSAR height reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 5854–5863, Oct 2017.

[74] Y. Huang, Y. Zhang, B. Xu, Z. Wu, and J. Chambers, "A new adaptive extended Kalman filter for cooperative localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. PP, no. 99, pp. 1–1, 2017.

[75] Y. Wang, H. Fang, L. Zhou, and T. Wada, "Revisiting the state-of-charge estimation for lithium-ion batteries: A methodical investigation of the extended Kalman filter approach," *IEEE Control Systems*, vol. 37, pp. 73–96, Aug 2017.

[76] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Transactions on Power Systems*, vol. 32, pp. 3205–3216, July 2017.

[77] S. Singhal and L. Wu, *Advances in Neural Information Processing Systems 1*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1 ed., 1989.

[78] M. St-Pierre and D. Gingras, "Comparison between the unscented Kalman filter and the extended Kalman filter for the position estimation module of an integrated navigation information system," in *IEEE Intelligent Vehicles Symposium, 2004*, pp. 831–835, June 2004.

[79] A. Giannitrapani, N. Ceccarelli, F. Scortecci, and A. Garulli, "Comparison of ekf and ukf for spacecraft localization via angle measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, pp. 75–84, January 2011.

[80] L. D'Alfonso, W. Lucia, P. Muraca, and P. Pugliese, "Mobile robot localization via EKF and UKF: A comparison based on real data," *Robotics and Autonomous Systems*, vol. 74, Part A, pp. 122 – 127, 2015.

[81] F. Kwasniok, "Estimation of noise parameters in dynamical system identification with Kalman filters," *PHYSICAL REVIEW E*, vol. 86, p. 036214, September 2012.

[82] X. Liu and Q. Gao, "Parameter estimation and control for a neural mass model based on the unscented Kalman filter," *PHYSICAL REVIEW E*, vol. 88, p. 042905, October 2013.

[83] T. A. Wenzel, K. J. Burnham, M. V. Blundell, and R. A. Williams, "Dual extended Kalman filter for vehicle state and parameter estimation," *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, vol. 44, pp. 153–171, Feb 2007.

[84] X. Sun, L. Jin, and M. Xiong, "Extended Kalman filter for estimation of parameters in nonlinear state-space models of biochemical networks," *PLoS ONE*, vol. 3, no. 11, 2008.

[85] E. Blanchard, A. Sandu, and C. Sandu, "Parameter estimation method using an extended Kalman filter," *Proceedings of the Joint North America, Asia-Pacific ISTVS Conference and Annual Meeting of Japanese Society for Terramechanics, Fairbanks, Alaska, USA*, vol. 30, No. 6, June 2007.

[86] F. C. Ma and S. H. Tong, "Real time parameters identification of ship dynamic using the extended Kalman filter and the second order filter," in *Control Applications, 2003. CCA 2003. Proceedings of 2003 IEEE Conference on*, vol. 2, pp. 1245–1250 vol.2, June 2003.

[87] S. Stubberud, "System identification using the neural-extended Kalman filter for control modification," in *Neural Networks, 2006. IJCNN '06. International Joint Conference on*, pp. 4449–4455, 2006.

[88] M. Myersa, A. Jorgeb, M. Mutton, and D. Walker, "A comparison of extended Kalman filter, Particle filter, and Least Squares localization methods for a high heat flux point source," *International journal of heat and mass transfer*, vol. 55(9-10), pp. 2219–2228, April 2012.

[89] F. Ababsa, M. Mallem, and D. Roussel, "Comparison between Particle filter approach and Kalman filter-based technique for head tracking in augmented reality systems," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 1, pp. 1021–1026 Vol.1, April 2004.

[90] M. Marron, J. Garcia, M. Sotelo, M. Cabello, D. Pizarro, F. Huerta, and J. Cerro, "Comparing a Kalman filter and a Particle filter in a multiple objects tracking application," in *Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on*, pp. 1–6, Oct 2007.

[91] N. Bellotto and H. Hu, "People tracking with a mobile robot: A comparison of Kalman and Particle filters," in *Proceedings of the 13th IASTED International Conference on Robotics and Applications*, RA '07, (Anaheim, CA, USA), pp. 388–393, ACTA Press, 2007.

[92] L. Bazzani, D. Bloisi, and V. Murino, "A comparison of multi hypothesis Kalman filter and Particle filter for multi-target tracking," in *Performance Evaluation of Tracking and Surveillance workshop at CVPR*, (Miami, Florida), pp. 47–54, 2009.

[93] D. C. K. Yuen and B. A. Macdonald, "A comparison between extended Kalman filtering and sequential monte carlo techniques for simultaneous localisation and map-building," in *In Proceedings of the 2002 Australasian Conference on Robotics and Automation*, 2002.

[94] M. F. Bugallo, S. Xu, and P. M. Djuri, "Performance comparison of EKF and Particle filtering methods for maneuvering targets," *Digital Signal Processing*, vol. 17, no. 4, pp. 774 – 786, 2007.

[95] G. Rigatos, "Particle filtering for state estimation in nonlinear industrial systems," *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, pp. 3885–3900, Nov 2009.

[96] T. Chen, J. Morris, and E. Martin, "Particle filters for the estimation of a state space model," *Proceedings of European Symposium on Computer Aided Process Engineering*, April 2004.

[97] X. Shao, B. Huang, and J. M. Lee, "Constrained Bayesian state estimation, a comparative study and a new Particle filter based approach," *Journal of Process Control*, vol. 20, no. 2, pp. 143 – 157, 2010.

[98] Z. Zhao, B. Huang, and F. Liu, "Constrained Particle filtering methods for state estimation of nonlinear process," *Proceedings of European Symposium on Computer Aided Process Engineering*, vol. 60, 2014.

[99] X. Yang, K. Xing, K. Shi, and Q. Pan, "Joint state and parameter estimation in Particle filtering and stochastic optimization," *Journal of Control Theory and Applications*, vol. 6, no. 2, pp. 215–220, 2008.

[100] T. Ishigaki and T. Higuchi, "Parameter identification of a pressure regulator with a nonlinear structure using a Particle filter based on the nonlinear state space model," in *Information Fusion, 2008 11th International Conference on*, pp. 1–6, June 2008.

[101] T. Chen, J. Morris, and E. Martin, "Particle filters for state and parameter estimation in batch processes," *Journal of Process Control*, vol. 15, no. 6, pp. 665 – 673, 2005.

[102] F. Gustafsson and T. Schön, "Particle filters for system identification of state-space models linear in either parameters or states," in *In proceedings of the 13th IFAC Symposium on System Identification*, pp. 1287–1292, 2003.

[103] F. Gustafsson and P. Hriljac, "Particle filters for system identification with application to CHAOS prediction."

[104] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45, pp. 910–927, May 2000.

[105] H. Sorenson and D. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465 – 479, 1971.

[106] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, pp. 439–448, Aug 1972.

[107] R. Chen and J. S. Liu, "Mixture Kalman filters," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 62, no. 3, pp. 493–508, 2000.

[108] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2602–2612, Oct 2003.

[109] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, pp. 780–783, Aug 1988.

[110] W. I. Tam, K. Plataniotis, and D. Hatzinakos, "An adaptive Gaussian sum algorithm for radar tracking," *Signal Processing*, vol. 77, no. 1, pp. 85 – 104, 1999.

[111] D. Musicki, B. F. L. Scala, and R. J. Evans, "Integrated track splitting filter - efficient multi-scan single target tracking in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, pp. 1409–1425, October 2007.

[112] I. Santamaria, C. J. Pantaleon, J. Ibanez, and A. Artes, "Deconvolution of seismic data using adaptive Gaussian mixtures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 855–859, Mar 1999.

[113] P. A. Brasnett, L. Mihaylova, N. Canagarajah, and D. Bull, "Particle filtering with multiple cues for object tracking in video sequences," *Proc.SPIE*, vol. 5685, pp. 5685 – 5685 – 12, 2005.

[114] B. Han, Y. Zhu, D. Comaniciu, and L. Davis, "Kernel-based Bayesian filtering for object tracking," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 227–234 vol. 1, June 2005.

[115] A. C. Atkinson and A. N. Donev, *Optimum experimental designs*. Clarendon Press, 1992.

[116] R. Mehra, "Optimal input signals for parameter estimation in dynamic systems–survey and new results," *IEEE Transactions on Automatic Control*, vol. 19, pp. 753–768, Dec 1974.

[117] G. C. Goodwin and R. L. Payne, *Dynamic system identification : experiment design and data analysis*. Academic Press, 1977.

[118] J. Kiefer, "Optimum experimental designs," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 21, no. 2, pp. 272–319, 1959.

[119] J. Kiefer and J. Wolfowitz, "Optimum designs in regression problems," *The Annals of Mathematical Statistics*, vol. 30, no. 2, pp. 271–294, 1959.

[120] S. Karlin and W. J. Studden, "Optimal experimental designs," *Ann. Math. Statist.*, vol. 37, pp. 783–815, 08 1966.

[121] K. Åström and P. Eykhoff, "System identification - a survey," *Automatica*, vol. 7, no. 2, pp. 123 – 162, 1971.

[122] R. Mehra, "Optimal inputs for linear system identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 192–200, Jun 1974.

[123] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application," *Automatica*, vol. 22, no. 5, pp. 543 – 554, 1986.

[124] H. Hjalmarsson, "System identification of complex and structured systems," *European Journal of Control*, vol. 15, no. 3, pp. 275 – 310, 2009.

[125] X. Bombois, G. Scorletti, M. Gevers, P. V. den Hof, and R. Hildebrand, "Least costly identification experiment for control," *Automatica*, vol. 42, no. 10, pp. 1651 – 1662, 2006.

[126] B. L. Cooley, J. H. Lee, and S. P. Boyd, "Control-relevant experiment design: a plant-friendly, LMI-based approach," in *Proceedings of the 1998 American Control Conference. ACC (IEEE Cat. No.98CH36207)*, vol. 2, pp. 1240–1244 vol.2, Jun 1998.

[127] M. Forgione, X. Bombois, and P. M. V. den Hof, "Data-driven model improvement for model-based control," *Automatica*, vol. 52, pp. 118 – 124, 2015.

[128] M. G. Potters, X. Bombois, M. Forgione, P. E. Modén, M. Lundh, H. Hjalmarsson, and P. M. J. V. den Hof, "Optimal experiment design in closed loop with unknown, nonlinear and implicit controllers using stealth identification," in *2014 European Control Conference (ECC)*, pp. 726–731, June 2014.

[129] C. A. Larsson, M. Annergren, H. Hjalmarsson, C. R. Rojas, X. Bombois, A. Mesbah, and P. E. Modén, "Model predictive control with integrated experiment design for output error systems," in *2013 European Control Conference (ECC)*, pp. 3790–3795, July 2013.

[130] M. Z. Babar and M. Baglietto, "MPC based optimal input design for nonlinear system identification," in *2016 20th International Conference on System Theory, Control and Computing (ICSTCC)*, pp. 619–625, Oct 2016.

[131] M. Baglietto, L. Scardovi, and R. Zoppoli, "Active identification of unknown systems: an information theoretic approach," in *Proceedings of the 2004 American Control Conference*, vol. 4, pp. 3826–3830 vol.4, June 2004.

[132] M. Baglietto, M. Paolucci, L. Scardovi, and R. Zoppoli, "Information-based multi-agent exploration," in *Robot Motion and Control, 2002. RoMoCo '02. Proceedings of the Third International Workshop on*, pp. 173–179, Nov 2002.

[133] M. B. Zarrop, "Preliminaries," in *Optimal Experiment Design for Dynamic System Identification*, pp. 1–13, Berlin/Heidelberg: Springer-Verlag, 1979.

[134] L. Ljung, "Asymptotic variance expressions for identified black-box transfer function models," *IEEE Transactions on Automatic Control*, vol. 30, pp. 834–844, Sep 1985.

[135] H. Hjalmarsson, M. Gevers, and F. de Bruyne, "For model-based control design, closed-loop identification gives better performance," *Automatica*, vol. 32, no. 12, pp. 1659 – 1673, 1996.

[136] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, no. 7, pp. 1215 – 1241, 1999.

[137] V. Fedorov, *Theory Of Optimal Experiments.* Academic Press, 1^{st} ed., 1st January 1972.

[138] M. Gevers, "Connecting identification and robust control: A new challenge," *IFAC Proceedings Volumes*, vol. 25, no. 15, pp. 1 – 10, 1992.

[139] X. Bombois, G. Scorletti, P. V. den Hof, and M. Gevers, "Least costly identification experiment for control. A solution based on a high-order model approximation," in *Proceedings of the 2004 American Control Conference*, vol. 3, pp. 2818–2823 vol.3, June 2004.

[140] C. R. Rojas, J. C. Agüero, J. S. Welsh, and G. C. Goodwin, "On the equivalence of least costly and traditional experiment design for control," *Automatica*, vol. 44, no. 11, pp. 2706 – 2715, 2008.

[141] M. Barenthin, H. Jansson, and H. Hjalmarsson, "Applications of mixed H_2 and H_∞ input design in identification," in *16th Triennial World Congress, Prague, Czech Republic*, vol. 16, pp. 458–463, 2005.

[142] X. Bombois and G. Scorletti, "Design of least costly identification experiments : The main philosophy accompanied by illustrative examples," *Journal Européen des Systèmes Automatisés (JESA)*, vol. 46, no. 6-7, pp. 587–610, 2012.

[143] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, no. 3, pp. 393 – 438, 2005.

[144] M. Annergren, C. A. Larsson, H. Hjalmarsson, X. Bombois, and B. Wahlberg, "Application-oriented input design in system identification: Optimal input design for control [applications of control]," *IEEE Control Systems*, vol. 37, pp. 31–56, April 2017.

[145] B. Ninness and H. Hjalmarsson, "Exact quantification of variance error," *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 277 – 282, 2002.

[146] K. Lindqvist and H. Hjalmarsson, "Optimal input design using linear matrix inequalities," in *Proceedings of the 12th IFAC Symposium on System Identification*, 2000.

[147] K. Lindqvist and H. Hjalmarsson, "Identification for control: adaptive input design using convex optimization," in *Proceedings of the 40th IEEE Conference on Decision and Control*, vol. 5, pp. 4326–4331 vol.5, 2001.

[148] X. Bombois, G. Scorletti, M. Gevers, R. Hildebrand, and P. V. den Hof, "Cheapest open-loop identification for control," in *2004 43rd IEEE Conference on Decision and Control (CDC)*, vol. 1, pp. 382–387 Vol.1, Dec 2004.

[149] Z. Zang, R. R. Bitmead, and M. Gevers, "Iterative weighted least-squares identification and weighted LQG control design," *Automatica*, vol. 31, no. 11, pp. 1577 – 1594, 1995.

[150] W. S. Lee, B. D. O. Anderson, R. L. Kosut, and I. M. Y. Mareels, "A new approach to adaptive robust control," *International Journal of Adaptive Control and Signal Processing*, vol. 7, pp. 183–211, May 1993.

[151] L. Ljung and Z.-D. Yuan, "Asymptotic properties of black-box identification of transfer functions," *IEEE Transactions on Automatic Control*, vol. 30, pp. 514–530, Jun 1985.

[152] Y. Zhu and P. P. van den Bosch, "Optimal closed-loop identification test design for internal model control," *Automatica*, vol. 36, no. 8, pp. 1237 – 1241, 2000.

[153] R. Hildebrand and G. Solari, "Identification for control: Optimal input intended to identify a minimum variance controller," *Automatica*, vol. 43, no. 5, pp. 758 – 767, 2007.

[154] G. C. Goodwin, J. S. Welsh, A. Feuer, and M. Derpich, “Utilizing prior knowledge in robust optimal experiment design,” *IFAC Proceedings Volumes*, vol. 39, no. 1, pp. 1358 – 1363, 2006. 14th IFAC Symposium on Identification and System Parameter Estimation.

[155] L. Gerencser and H. Hjalmarsson, “Adaptive input design in system identification,” in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 4988–4993, Dec 2005.

[156] E. E. Walter and L. Pronzato, *Identification of parametric models from experimental data*. Springer, 1997.

[157] D. B. Kell and S. G. Oliver, “Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era,” *BioEssays*, vol. 26, pp. 99–105, Jan 2004.

[158] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, “Active Learning with Statistical Models,” *Journal of Artificial Intelligence Research*, pp. 129 – 145, February 1996.

[159] A. A. Feldbaum, “Theory of dual control,” *Avtomat. i Telemekh.*, vol. 22, no. 1, 1961.

[160] D. Rivera, M. Braun, and H. Mittelmann, “Constrained multisine inputs for plant-friendly identification of chemical processes,” *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 425 – 430, 2002. 15th IFAC World Congress.

[161] B. A. Ogunnaike, “A contemporary industrial perspective on process control theory and practice,” *Annual Reviews in Control*, vol. 20, no. Supplement C, pp. 1 – 8, 1996.

[162] Y. Zhu, “System identification for process control: recent experience and outlook,” *International Journal of Modelling, Identification and Control*, vol. 6, no. 2, p. 89, 2009.

[163] D. E. Rivera, H. Lee, M. W. Braun, and H. D. Mittelmann, ““Plant-Friendly” system identification: a challenge for the process industries,” *IFAC Proceedings Volumes*, vol. 36, no. 16, pp. 891 – 896, 2003. 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands, 27-29 August, 2003.

[164] J. H. Lee, “Control-relevant design of periodic test input signals for iterative open-loop identification of multi-variable FIR systems,” *IFAC Proceedings Volumes*, vol. 36, no. 16, pp. 903 – 908, 2003. 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands, 27-29 August, 2003.

[165] H. Zhao, M. Harmse, J. Guiver, and W. M. Canney, “Subspace identification in industrial APC applications - a review of recent progress and industrial experience,” *IFAC Proceedings Volumes*, vol. 39, no. 1, pp. 1074 – 1079, 2006. 14th IFAC Symposium on Identification and System Parameter Estimation.

[166] D. Densmore and S. Hassoun, "Design automation for synthetic biological systems," *IEEE Design Test of Computers*, vol. 29, pp. 7–20, June 2012.

[167] F. Menolascina, D. Bellomo, T. Maiwald, V. Bevilacqua, C. Ciminelli, A. Paradiso, and S. Tommasi, "Developing optimal input design strategies in cancer systems biology with applications to microfluidic device engineering.," *BMC bioinformatics*, vol. 10 Suppl 12, p. S4, oct 2009.

[168] C. Kreutz and J. Timmer, "Systems biology: experimental design," *FEBS Journal*, vol. 276, pp. 923–942, feb 2009.

[169] G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel, "Nominally robust model predictive control with state constraints," in *42nd IEEE International Conference on Decision and Control*, vol. 2, pp. 1413–1418 Vol.2, Dec 2003.

[170] D. Limon, T. Alamo, F. Salas, and E. Camacho, "Input to state stability of min-max MPC controllers for nonlinear systems with bounded uncertainties," *Automatica*, vol. 42, no. 5, pp. 797 – 803, 2006.

[171] D. Limon, T. Alamo, F. Salas, and E. F. Camacho, "On the stability of constrained MPC without terminal constraint," *IEEE Transactions on Automatic Control*, vol. 51, pp. 832–836, May 2006.

[172] J. M. J. M. Maciejowski, *Predictive control : with constraints*. Prentice Hall, 2002.

[173] L. Gerencsér, H. Hjalmarsson, and J. Mårtensson, "Identification of ARX systems with non-stationary inputs: asymptotic analysis with application to adaptive input design," *Automatica*, vol. 45, no. 3, pp. 623 – 633, 2009.

[174] B. Wahlberg, H. Hjalmarsson, and M. Annergren, "On optimal input design in system identification for control," in *49th IEEE Conference on Decision and Control (CDC)*, pp. 5548–5553, Dec 2010.

[175] J. Zaborszky, "An information theory viewpoint for the general identification problem," *IEEE Transactions on Automatic Control*, vol. 11, pp. 130–131, Jan 1966.

[176] D. L. Snyder and I. B. Rhodes, "Filtering and control performance bounds with implications on asymptotic separation," *Automatica*, vol. 8, no. 6, pp. 747 – 753, 1972.

[177] J. Galdos, "A Cramer-Rao bound for multidimensional discrete-time dynamical systems," *IEEE Transactions on Automatic Control*, vol. 25, pp. 117–119, Feb 1980.

[178] P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Transactions on Signal Processing*, vol. 49, pp. 87–90, Jan 2001.

[179] M. Zakai and J. Ziv, “Lower and upper bounds on the optimal filtering error of certain diffusion processes,” *IEEE Transactions on Information Theory*, vol. 18, pp. 325–331, May 1972.

[180] J. Galdos, “A rate distortion theory lower bound on desired function filtering error (corresp.),” *IEEE Transactions on Information Theory*, vol. 27, pp. 366–368, May 1981.

[181] R. Washburn and D. Teneketzis, “Rate distortion lower bound for a special class of nonlinear estimation problems,” *Systems & Control Letters*, vol. 12, no. 3, pp. 281 – 286, 1989.

[182] M. Zakai, “On mutual information, likelihood ratios, and estimation error for the additive Gaussian channel,” *IEEE Transactions on Information Theory*, vol. 51, pp. 3017–3024, Sept 2005.

[183] T. E. Duncan, “On the calculation of mutual information,” *SIAM Journal on Applied Mathematics*, vol. 19, no. 1, pp. 215–220, 1970.

[184] S. Verdu, “Mismatched estimation and relative entropy,” *IEEE Transactions on Information Theory*, vol. 56, pp. 3712–3720, Aug 2010.

[185] K. Smith, “On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations,” *Biometrika*, vol. 12, no. 1/2, pp. 1–85, 1918.

[186] A. Wald, “On the efficient design of statistical investigations,” *Ann. Math. Statist.*, vol. 14, pp. 134–140, 06 1943.

[187] S. D. Silvey, *Optimal Design: An Introduction to the Theory for Parameter Estimation.* Ettore Majorana International Science Series, 1^{st} ed., 1980.

[188] A. Pazman, *Foundations of Optimum Experimental Design.* Theoretical Computer Science, 1986.

[189] H. Chernoff, “Locally optimal designs for estimating parameters,” *The Annals of Mathematical Statistics*, vol. 24, no. 4, pp. 586–602, 1953.

[190] M. Yang, “A-optimal designs for generalized linear models with two parameters,” *Journal of Statistical Planning and Inference*, vol. 138, no. 3, pp. 624 – 641, 2008.

[191] S. Ehrenfeld, “On the efficiency of experimental designs,” *Ann. Math. Statist.*, vol. 26, pp. 247–255, 06 1955.

[192] B. Heiligers, “Computing E-optimal polynomial regression designs,” *Journal of Statistical Planning and Inference*, vol. 55, no. 2, pp. 219 – 233, 1996.

[193] H. Dette and L. M. Haines, “E-optimal designs for linear and nonlinear models with two parameters,” *Biometrika*, vol. 81, no. 4, pp. 739–754, 1994.

[194] N. R. Draper and R. C. S. John, “Designs in three and four components for mixtures models with inverse terms,” *Technometrics*, vol. 19, no. 2, pp. 117–130, 1977.

[195] R. Sibson and A. Kenny, “Coefficients in D-optimal experimental design,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 37, no. 2, pp. 288–292, 1975.

[196] A. C. Atkinson, “The comparison of designs for sequential clinical trials with covariate information,” *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 165, no. 2, pp. 349–373, 2002.

[197] C. L. Atwood, “Optimal and efficient designs of experiments,” *Ann. Math. Statist.*, vol. 40, pp. 1570–1602, 10 1969.

[198] S. D. Silvey and D. M. Titterington, “A geometric approach to optimal design theory,” *Biometrika*, vol. 60, no. 1, pp. 21–32, 1973.

[199] G. Elfving, “Optimum allocation in linear regression theory,” *The Annals of Mathematical Statistics*, vol. 23, no. 2, pp. 255–262, 1952.

[200] D. M. Titterington, “Optimal design: Some geometrical aspects of d-optimality,” *Biometrika*, vol. 62, no. 2, pp. 313–320, 1975.

[201] R. M. Heiberger, D. K. Bhaumik, and B. Holland, “Optimal data augmentation strategies for additive models,” *Journal of the American Statistical Association*, vol. 88, no. 423, pp. 926–938, 1993.

[202] E. Läuter, “Experimental design in a class of models,” *Mathematische Operationsforschung und Statistik*, vol. 5, no. 4-5, pp. 379–398, 1974.

[203] A. Atkinson, “DT-optimum designs for model discrimination and parameter estimation,” *Journal of Statistical Planning and Inference*, vol. 138, no. 1, pp. 56 – 64, 2008.

[204] A. C. Atkinson and V. V. Fedorov, “The design of experiments for discriminating between two rival models,” *Biometrika*, vol. 62, no. 1, pp. 57–70, 1975.

[205] F. Girosi, *Regularization Theory, Radial Basis Functions and Networks*, pp. 166–187. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994.

[206] H. Tanizaki, “Density-Based Nonlinear Filters,” in *Nonlinear Filters*, pp. 71–111, Berlin, Heidelberg: Springer Berlin Heidelberg, 1996.

[207] S. Rastegarpour, A. Shahmansoorian, and A. H. Mazinan, “Designing an optimal control law for nonlinear system by using intelligent algorithm,” in *Measurement, Information and Control (ICMIC), 2013 International Conference on*, vol. 02, pp. 830–836, Aug 2013.

[208] S. Kwak, U. C. Moon, and J. C. Park, "Predictive-control-based direct power control with an adaptive parameter identification technique for improved AFE performance," *IEEE Transactions on Power Electronics*, vol. 29, pp. 6178–6187, Nov 2014.

[209] L. Scardovi, M. Baglietto, and T. Parisini, "Active state estimation for nonlinear systems: A neural approximation approach," *IEEE Transactions on Neural Networks*, vol. 18, pp. 1172–1184, July 2007.

[210] G. Gatto, I. Marongiu, and A. Serpi, "Discrete-time parameter identification of a surface-mounted permanent magnet synchronous machine," *IEEE Transactions on Industrial Electronics*, vol. 60, pp. 4869–4880, Nov 2013.

[211] P. Venkataraman, *Applied optimization with MATLAB programming.* John Wiley & Sons, 2009.

[212] S. Kwon, S. Kim, and J. Yu, "Tilting-type balancing mobile robot platform for enhancing lateral stability," *IEEE/ASME Transactions on Mechatronics*, vol. 20, pp. 1470–1481, June 2015.

[213] C. H. Chiu, "The design and implementation of a wheeled inverted pendulum using an adaptive output recurrent cerebellar model articulation controller," *IEEE Transactions on Industrial Electronics*, vol. 57, pp. 1814–1822, May 2010.

[214] S. Yuan, G. Lei, and X. Bin, "Dynamic modeling and sliding mode controller design of a two-wheeled self-balancing robot," in *2016 IEEE International Conference on Mechatronics and Automation*, pp. 2437–2442, Aug 2016.