Continuous Time Model Identification using Sinusoidal Response

by

©Shaikh Mohammad Fahim

A Thesis
submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree of

Master of Engineering

Faculty of Engineering and Applied Science
Memorial University of Newfoundland
October 2018

St. John’s Newfoundland
Abstract

System identification is an interface that unites the mathematical world of control theory and practical applications of control; as such its significance is omnipresent. Identification techniques involve differential equations where the coefficients are closely related to the physical parameters in the system; continuous time models have greater appeal than its discrete-time counterpart in understanding these interpretations. In this study, we have considered sinusoidal input for identification purpose as it has been discussed in the context of designing optimal input and also because it facilitates to excite processes with particular frequencies of interest. The primary objective of this work focuses on process parameter estimation. At first, integer order model is studied due to its simplicity, as order estimation is not necessary and thus the structure of the model. In addition, a comparison between different identification methods for better parameter estimates is performed on integer order model. Following on, fractional order model is taken into consideration with known and unknown order estimates. When solving for unknown model order, more emphasis is given on the logarithmic derivative term. According to literature, the unknown model order is estimated numerically whereas we provide an analytical expression of logarithmic derivative of sinusoidal inputs considering deterministic approach. For integer order model, although satisfactory results were achieved in terms of parameter estimates for different approaches varying different input constraints, it was evident that the
performances varied with data length, and more importantly with the frequency of the input signal. The developed methodology for fractional order model identification with known model order lead fairly accurate estimates of the process parameters and when extended for unknown model order, exhibited highly satisfactory results as well but with higher computational time. The main challenge of this study was optimizing process parameters based on convergence; this issue was studied in simulation and corresponding numerical results for diverse noise levels met our expectations.
I would like to take the opportunity to thank Dr. Salim Ahmed and Dr. Syed A. Imtiaz for all their generous help, support and valuable suggestions. Since the first day, they have been steady guides on this difficult voyage. There were times of frustration during research, when they encouraged me and provided me with directions, spending hours with me solving problems.

During my study period at Memorial University, I was lucky to have Professors Dr. Geoff Rideout and Dr. Leonard Lye as my course teachers. Their course have assisted me getting better insight on dynamics of mechanical systems and statistical analysis. In addition, courses offered by my supervisors, Dr. Salim Ahmed and Dr. Syed A. Imtiaz have helped me to improve my analytical skill and enhanced my knowledge in control and safety engineering.

Finally, the financial support provided by the Research and Development Corporation, School Graduate Studies, Memorial University of Newfoundland and the Faculty of Engineering and Applied Science is much appreciated.
# Table of Contents

Abstract ................................................................. ii  
Acknowledgments ........................................................ iv  
Table of Contents ......................................................... v  
List of Tables ........................................................... ix  
List of Figures ........................................................... x  
1 Introduction ............................................................ 1  
   1.1 Background .......................................................... 1  
   1.1.1 Modelling approaches ............................................ 2  
   1.1.2 Approaches for System Identification ......................... 4  
   1.1.2.1 Discrete time identification ................................. 4  
   1.1.2.2 Continuous time identification ............................... 5  
   1.1.3 Continuous time identification using sinusoidal input ....... 7  
   1.1.4 Identification of Fractional Order models .................... 8  
1.2 Motivation ............................................................ 9  
1.3 Objectives .......................................................... 10  
1.4 Structure of the thesis .............................................. 11
## 2 Literature Review

2.1 System Identification approaches ........................................... 12

2.2 Identification using sinusoidal input ........................................ 15
  2.2.1 Frequency domain identification ....................................... 18
  2.2.2 Time domain identification ............................................. 20

2.3 Continuous Time identification ............................................. 21

2.4 Fractional Order model identification .................................... 25
  2.4.1 Significance of fractional derivative ................................... 26
  2.4.2 Fractional order controller ............................................ 28
  2.4.3 Identification methods .................................................. 30

## 3 Comparative Study on Continuous Time Integer Order Model Identification

3.1 Introduction ................................................................. 36

3.2 Methodology .............................................................. 37
  3.2.1 The direct approach [Ahmed, 2014] ..................................... 38
  3.2.2 Identification by estimating step response [Ahmed et al., 2009] 40
  3.2.3 Identification by piece-wise linear approximation [Ahmed, 2010a] 44

3.3 Simulation study ............................................................ 46
  3.3.1 Effect of NSR ............................................................. 46
  3.3.2 Effect of data length .................................................... 48
  3.3.3 Effect of frequency ...................................................... 50

3.4 Concluding remarks ....................................................... 52

## 4 Fractional Order Model Identification

4.1 Introduction ................................................................. 56
4.2 Methodology- Parameter estimation with known fractional order of the model ............................................. 56
  4.2.1 Parameter estimation using sinusoidal response .............. 56
    4.2.1.1 Type-I model ................................................ 57
    4.2.1.2 Type-II model ................................................ 60
    4.2.1.3 Extended Type-II model ................................... 62
  4.2.2 Order estimation ................................................. 64
  4.2.3 Simulation and Results ......................................... 66
  4.2.4 Simulation environment ...................................... 66
    4.2.4.1 Effect of NSR .............................................. 67
    4.2.4.2 Effect of frequency ....................................... 67
    4.2.4.3 Effect of number of samples ........................... 69
    4.2.4.4 Comparison based on average error criterion .......... 72
  4.3 Methodology- Simultaneous parameter and fractional order estimation 72
  4.4 Mathematical formulations .................................... 75
    4.4.1 Identification method ..................................... 75
    4.4.2 Evaluation of the logarithmic derivative .................. 78
  4.5 Implementation issues .......................................... 79
    4.5.1 Initialization .............................................. 79
    4.5.2 Simulation and results .................................... 81
      4.5.2.1 Identification of processes and parameter convergence
                   for noise free scenario ............................... 81
      4.5.2.2 Identification of time delay models considering NSR 82
      4.5.2.3 Effect of NSR ........................................... 84
      4.5.2.4 Effect of data length and sampling interval .......... 84
  4.6 Concluding remarks ........................................... 87
List of Tables

4.1 Identification results of processes considering zero NSR . . . . . . . 82
4.2 Identification results of processes considering NSR=10% . . . . . . . 82
4.3 Effect of NSR on parameter estimates . . . . . . . . . . . . . . . . . . . 84
4.4 Effect of data length and sampling interval on parameter estimates . 86
List of Figures

3.1 Estimated process parameters with different methodology for different values of NSR .............................................. 48
3.2 Estimated process parameters with different methodology for different data length .................................................. 50
3.3 Estimated process parameters with different methodology for various frequency factors .............................................. 52
3.4 Comparison of different algorithms based on average error criterion changing one input variable at a time ....................... 54
4.1 Estimated process parameters for different values of NSR .......... 68
4.2 Estimated process parameters for different values of frequency factors ................................................................. 70
4.3 Estimated process parameters over a range of data length .......... 71
4.4 Comparison of Type-I and Type-II model ................................................................. 73
4.5 Convergence of process parameters for $\frac{3}{10s^3+1}e^{-3s}$ ................................................................. 83
4.6 Effect of NSR on parameter estimates for different models ........ 86
4.7 Effect of data length and sampling interval on parameter estimates .. 87
Chapter 1

Introduction

1.1 Background

Use of system models in different aspects of control engineering is evident with the rapid development of control technology impacting the entire control discipline. This phenomenon has led many researchers to pursue new developments concerning system identification and consequently highlighting the importance of system models in the modern era of advanced control design.

In perspective of control engineering, system models are constructed in the form of dynamical systems, which can be categorized into following different types:

- *Mental models*—intuitive notion that people have of a system behavior.
- *Software models*—descriptions that are contained in software programs.
- *Graphical models*—descriptions in the form of characteristics and graphs.
- *Mathematical models*—descriptions in the form of mathematical relations.
Mathematical models of dynamical systems mostly concern dynamical relationships between physical quantities. This implies that utilizing the basic laws of physics (first principles) a model of the system can be developed. Complementary to this, measurement data of the input and output variables of the system can contain all relevant information of the underlying system dynamics as well. Therefore, rather than building models from first principles, information from experimental data can be an effective approach to build models of the actual and emerging behavior of dynamic systems. This leads to an area which is known as system identification.

1.1.1 Modelling approaches

In general there are two ways of deriving models of physical processes:

- *First principles modeling (or physical modeling)* - Physical knowledge about the process in the form of first principles relations is employed in order to arrive at a model that will generally consist of a multitude of differential/partial differential/algebraic relations between physical quantities. The first principles relations concern e.g. the laws of conservation of energy and mass and Newton’s law of movement.

- *Data-driven modeling (more commonly system identification)* - Measurements of several variables of the process are taken and a model is constructed by identifying a model structure and estimating a set of parameters which match the dynamics captured in the data.

In many situations the first approach is followed, e.g. in chemical process, mechanical systems etc. where sufficient knowledge regarding the basic principles governing the system behavior is a prerequisite. Afterwards, numerical values of physical coefficients e.g. masses, stiffnesses, material properties, properties of chemical reactions or
other (presumed) constants have to be substituted in the model relations. Yet the developed model for the dynamical system might not practically perform as a perfect set of deterministic equations because in real-life the system variables will be subject to all kinds of disturbances influencing the input-output relation within a dynamical system. These disturbances will limit the validity of deterministic models in any type of application. Additionally, modeling large systems might often lead to high level of complexity, ultimately making the model intractable for particular applications. Considering these facts, experimental data of the process can be used in order to arrive at an appropriate model. The situation that a model is identified purely on the basis of data and without taking particular account of the physical structure is referred to as black box identification, contrary to the first principles approach, that is taken in the case of pure physical modeling [Garnier et al., 2008].

In general, system identification is the art and science of using measurements obtained from a system to characterize the system, typically in some mathematical format. The term ‘System Identification’ was coined by Zadeh [Zadeh, 1956] to categorize model estimation problem for dynamic systems in the control community. Later a more formal definition was given by Ljung- The art and science of building mathematical models of dynamic systems from observed input-output data [Ljung, 2010]; input being the external signal that can be manipulated and output being the observable signal that are of interest to an observer. Since dynamical systems are abundant in our surroundings and system identification refers as an interface between the applications of the real world and mathematical world of control theory, its undisputed importance is highly recognized with a wide range of application area including but not limited to mechanical engineering, biology, physiology, meteorology, economics, model-based control design etc.

3
1.1.2 Approaches for System Identification

System identification has been an active research area for more than four decades with considerable development. Although this area has numerous facets and there are many approaches and methods, the overall concept stemmed from two different approaches:

- Discrete time identification, and
- Continuous time identification.

1.1.2.1 Discrete time identification

A model that directly expresses the relationship between the measurements of the input and output signals at the time-instants is defined as a discrete-time model. While in continuous time models, the relations between the variables are realized in the form of continuous-time differential equations [Iwase et al., 2002]. With the advancements of computers, a lot of system identification methods benefiting from the digital processing have been developed, and identification for discrete-time systems has been studied to facilitate the analysis and data processing.

The popularity of discrete-time identification over continuous time identification despite the fact that most real world physical processes are continuous in character is because of the following conveniences [Bedoui et al., 2012]:

- Reduction of calculus into algebra,
- Ease in implementing dynamic strategies in real time,
Well established theory for deterministic and stochastic situations,

Not necessary to estimate physical parameters in case of controller designing.

Gaining from these conveniences, the attention received by discrete time methods was so enormous that the continuous time counterpart was completely overshadowed. In fact people who were curious in continuous time treatment were tagged either as ‘old timers of the analog age’ or ‘those having academic interest only’. However, the situation has gradually changed ever since as the relevance of continuous time treatment has been reinstated [Rao and Sinha, 1991]. Several indentation methods for discrete time models have been documented in the literature based on cross-correlation analysis [Zhang and Li, 2003, Zheng and Feng, 1990], error minimization between process output and process predictive model output [Gao et al., 2003, Gao et al., 2005], standard recursive least squares [Ferretti et al., 1991], variable regression estimator [Elnaggar et al., 1990] etc.

1.1.2.2 Continuous time identification

In early 1950s continuous-time based contributions to system identification started commencing. However, overshadowed by the ‘digital’ spirit instigated by parallel developments in digital computers during the following two decades. Nevertheless, again 1970s witnessed a resurgence of continuous-time concepts. Ever since the field of continuous-time system identification is maturing [Unbehauen and Rao, 1998]. Much has been elaborated on the significance of continuous time models [Rao and Unbehauen, 2006, Unbehauen and Rao, 1990, Rao and Sinha, 1991]. Some of the main arguments in favor of CT models are documented as followed.

- Models of physical systems derived from physical principles are inherently continuous in time, because physical laws on which such modeling is based are in
CT.

- CT models describe the physical phenomena of systems and processes more accurately, as because the model parameters are strongly correlated with the physical properties of the system.

- Redundant sensitivity issues with respect to model parameters do not arise, unlike in the event of discretization.

- CT models are capable of preserving partial knowledge whereas in the process of discretization of a CT model containing some known parameter, information loss takes place.

- Discretization of CT models may give rise to unnatural non-minimum phase character.

- Conventional DT methods are not in harmony with the CT spirit; in the limit of reduced sampling period, they do not converge to the results corresponding to the original CT model.

- CT models are very useful to deal with mildly non-uniformly distributed sampled data, dominant system modes, fast sampled data etc.

The various approaches reported in the literature for identification of continuous-time systems may be classified into three broad categories:

- A: Approaches using discrete time signals to identify a DT model which is then converted into native CT form.

- B: Approaches using CT signals to directly identify a native CT model.
• C: Approaches using DT signals giving rise to a unconventional discrete time (UDT) model which converges to its native CT with the help of an UDT operator which is in harmony with its CT counterpart in the sense that the DT model converges to the original CT version as the sampling interval approaches zero.

From the above mentioned approaches, indirect method which translates the identified discrete time model to a continuous time one, introduces a numerically ill conditioned problem or difficulty in transforming the zeros of the discrete time model thus leading towards direct method which directly identifies a continuous time model from the sampled input-output data has received much attention [Young, 1981, Unbehauen and Rao, 1990, Sinha and Rao, 2012].

### 1.1.3 Continuous time identification using sinusoidal input

The quality of an estimated model is closely related with the choice of the input signal for identification purpose [Ahmed, 2010b]. Although step and step like signals have been extensively studied and used for process identification for their simplicity in nature, periodic signals for continuous time model identification has caught attention of researchers of recent times and thus became an active research area.

Step signals require sudden changes in variables between two operating points which might not be feasible all the time and even though feasible might hamper the stability of an process. To avoid the circumstance, periodic signals are considered to be a great alternative as it provides much more grip towards the user.

The use of periodic signals as an input for identification of continuous time linear dynamic models has many advantages. Among various types of periodic signal, sinusoid is considered to be the most flexible one due to its optimal properties [Schoukens
et al., 1994] and plant friendliness [Rivera et al., 2009]. In addition, sinusoids help excite a process at the frequencies of interest and its smooth nature makes it desirable for making gradual changes in variables [Ahmed, 2014]. Applications of the sinusoid as input for model identification have been reported in the literature [Braun et al., 2002, Doraiswami et al., 1986, Godfrey, 1993, Kalafatis et al., 2005, Zaremba and Pavlov, 2002].

1.1.4 Identification of Fractional Order models

Contrary to the traditional integer order models, fractional order systems possess transfer functions of an arbitrary real order. Although a significant amount of literature can be found relating continuous time model identification for integer order models, in recent trend, focus has been shifted towards fractional order models as it has been observed that many real-world physical systems with long memory transients and infinite dimensional structures are better characterized by fractional order differential equations rather than classical integer order models [Narang et al., 2011]. Some examples of fractal systems include mass diffusion, heat conduction, transmission lines, electrochemical processes, dielectric polarisation, viscoelastic materials etc.

The mathematical foundation behind fractional order model identification lies with ‘fractional calculus’. Before 20th century, the term ‘fractional calculus’ was only developed in theory and its application was confined only within theoretical limitations in the field of mathematics [Oldham and Spanier, 1974, Podlubny, 1998]. Recently, especially over the last two decades it has gained much popularity and became an interesting topic of research in both scientific and industrial communities with growing computational power and thus there have been a considerable development in the use of fractional operators in various fields [Narang et al., 2011]. In general, fractional
calculus is about differentiation and integration of non-integer orders which have been well defined through Grünwald–Letnikov (GL) discrete form of the definition and the Riemann–Liouville (RL) definition respectively [Oldham and Spanier, 1974]. Later in 1990s, Lay [Le Lay, 1998], Lin [Lin, 2001] and Cois [Cois, 2002] developed system identification technique in time domain implementing the concept of fractional differentiation models following two basic approaches: equation error and output error.

System identification is treated as a standard tool for unknown systems. However, identifying a system consisting fractional orders is more complicated compared to integer order systems, as for integer-order systems, upon identifying the maximum order of the system the parameters of the model can be optimized directly; conversely for fractional-order systems, identification requires the choice of the number of fractional operators, the fractional power of the operators, and finally the coefficients of the operators. In order to better understand and to deal with the above mentioned situation, Malti et al. has reviewed and detailed some of the progresses achieved [Malti et al., 2008a]. Fractional-order Systems and Controls: Fundamentals and Applications [Monje et al., 2010] published by Springer under the series Advances in Industrial Control has tremendously motivated us to carry our work forward.

1.2 Motivation

Theoretically a system can consist of infinite number of integer orders; in reality, processes deal with first and second order systems and some in rare cases systems with higher order. But what if a system is not of an integer order, which is rare but true is some cases. These sort of systems are now known as fractional order models. This interesting topic has motivated to carry this work identifying and developing theory
behind integer order models and fractional order models with known and unknown value of the fractional order. Although the case for integer order models are sort of straight forward, a fractional order model can be very complex as because either fractional orders can be an integral multiples of a commensurate order or they can be very different from each other.

Moreover, previous studies have indicated to identify a system usually a step or step like signal was used. But, recent studies have shown selecting a sinusoid as an input signal for identification purpose gives a lot more advantages over step or step like signals due to fact that it gives a lot more control over the input parameters of a signal towards the user. Later in the study input signals with multiple sinusoids were used and different factors of the study were discussed.

1.3 Objectives

This study started with evaluation of several developed identification techniques for integer order models and later on the focus was to identify fractional order models. To outline the work, the prime objectives are as follows:

- Review and compare performance of continuous time identification methods using sinusoidal input.
- Developed fractional order model identification method using sinusoidal response.
1.4 Structure of the thesis

In the following chapter, a detail literature can be found on the related works done in this field of study. The literature briefly explains about ‘system identification’—both discrete time and continuous time along with their advantages and disadvantages. Furthermore, this section also points out the reason for using integral equation approach rather than differential approach for estimation purpose. In addition, a wide range of techniques for integer order and fractional order model identification is documented. And finally a brief literature on optimization technique is discussed.

Chapter 3 presents a comparative study done on a continuous time integer order model. In this chapter three major approaches for identification using sinusoidal signal are highlighted and a comparison is done among the three different techniques in terms of different factors of the input signal and finally comments are made on overall performance.

Chapter 4 can be sub-categorized in to two different parts. The first part is dedicated on the theory behind the identification technique of a fractional order model for a range of fractional orders with user defined interval. And afterwards, based on minimum error criterion parameters along with fractional order was selected. Whereas, the second part is more automated and follows optimization technique based on parameter convergence for optimal parameter and fractional order estimation.

Chapter 5 presents conclusions and recommendations for future works.
Chapter 2

Literature Review

2.1 System Identification approaches

Many of the system identification methods available nowadays date back to the basic principles of least-squares, as introduced by Gauss [Gauss, 1963]. System identification has received a growing interest over the last decades. The basic methods were developed in the sixties and seventies of the previous century starting from the introduction of computers for performing the often heavy calculations [Åström and Bohlin, 1966, Åström and Eykhoff, 1971]. A number of books written in the seventies have recorded these developments [Eykhoff, 1974, Goodwin and Payne, 1977]. Being based on the theories of stochastic processes and statistical analysis, system identification at that time was specifically seen as a problem of parameter estimation. It was commonly assumed that one knew the correct model structure or order and the character of the noise disturbance on the data. The main underlying assumption in this approach appeared to be the assumption that the data generating system can be modeled exactly by a linear, time-invariant, finite-order model, where ‘system’ refers not only to the input-output transfer function but also to the particular description
of how noise affects the measurement data. This assumption was reflected by the situation that an exact parameter was supposed to be present in the parameter set.

In the 1980s, this basic assumption had been relaxed, giving more attention to the more realistic situation that system identification generally comes down to ‘approximate modeling’ rather than ‘exact modeling’. Issues of approximation have become popular, a development which was pulled mainly by the Swedish school of researchers in Lund, Linköping and Uppsala [Ljung and Caines, 1980, Ljung and Söderström, 1983, Wahlberg and Ljung, 1986]. A good overview of this development, which turned ‘parameter estimation’ into ‘system identification’ is documented in the works of Ljung [Ljung, 1999] and Söderström and Stoica [Söderström and Stoica, 1989].

Interest in the issue of approximation made people move away from notions as consistency, and made them pay attention to the type of approximation that becomes involved. A related issue that comes into the picture is the issue of the intended application or goal of the model. As identifying a system no longer means finding an exact representation, but rather finding an approximation, then specific modeling goals might dictate which type of approximations are desirable; or in other words, which aspects of the system dynamics will be incorporated in the model, and which aspects will be neglected.

Especially in the area of approximate modeling, the 1990s have shown an increasing interest in identifying approximate models that are suitable for serving as a basis for model-based control design. This means that, although one realizes that models obtained are only approximate, one would like to obtain models that are accurate descriptions of the system dynamics in those aspects of the system that are specifi-

Another area of interest which is extremely relevant from an applications point of view, is the question concerning the accuracy of identified models. Experimental data provides us with information concerning the dynamical system; besides the problem of extracting an appropriate model from the measured data, it is important to be able to make statements concerning the accuracy and reliability of this result. This area, sometimes denoted as model uncertainty estimation has been a part of the classical analysis in the form of providing confidence intervals for parameter estimates, however, restricted to the situation in which consistent models were estimated. In an approximate setting of identification, this issue is still an important subject of research, being closely related to the question of model validation and to the goal-oriented design of experiments [Bombois et al., 2006].

Important challenges are faced while identifying models with nonlinear dynamics. Whereas in many applications it suffices to consider linear models of a linearized nonlinear plant, the challenge to express the nonlinear dynamical phenomena of the plant into a nonlinear model often enhances the capabilities of the model, e.g. when designing a control system that moves the plant through several operating regimes. A detail contribution in this area can be found in the literature [Suykens and Vandewalle, 2012, Nelles, 2013, Tóth, 2010].

Additionally the final question that has to be dealt with, is the question whether one is satisfied with the model obtained. This latter step in the procedure is indicated by
the term model validation. The question whether one is satisfied with the result will in
many situations be very much dependent on the question what the model is intended
for. A decisive answer to the validation question is then, that one is satisfied with
the model if in the intended model application one is satisfied with the result. If the
model is invalidated, then a redesign of the identification experiment or adjustment
of model set and identification criterion may lead to an improved model.

2.2 Identification using sinusoidal input

The process of selecting a suitable input signal for identification is termed as ‘input
design’. This amounts to determine a signal that excites the appropriate dynamics of
the system, subject to given constraints. Issues that are essential to construct a good
model in practice include the problem of input and experiment design.

For identification purpose, data should be informative meaning for open loop opera-
tion the input should be persistently exciting of a certain order consisting sufficient
distinct frequencies. For identification of linear systems, there are three basic facts
that govern the choices:

1. The asymptotic properties of the estimate depends only on the input spectrum,
not on the actual waveform of the input.

2. The input must have limited amplitude. The crest factor measures how well a
given signal utilizes such a given amplitude span.

3. Periodic inputs may have certain advantages.
For linear system identification it is desirable to achieve a desired input spectrum for a signal with a small crest factor as possible. Unfortunately these properties are somewhat in conflict with each other. Following are some commonly used input signals:

1. Filtered Gaussian White Noise: A simple choice is to let the signal be generated as white Gaussian noise, filtered through a linear filter. With this we can achieve virtually any signal spectrum by proper choice of filters. Since the signal is generated off-line, non causal filters can be applied and transient effects can be eliminated, which gives even better spectral behavior.

2. Random Binary Noise: A random binary signal is a random process which assumes only two values. It can be generated in a number of different ways.

3. Pseudo Random Binary Noise: A Pseudo Random Binary signal is a periodic deterministic signal with white noise like properties.

4. Multi-Sines: A natural choice of input is to form it as a sum of sinusoids.

5. Chirp Signals or Swept Sinusoids: A chirp signal is a sinusoid with a frequency that changes continuously over a certain band.

Three basic constituents of an identification procedure are as follows:

- **Data** - Data might be available from normal operating records, but it may as well be possible to design tailor-made experiments to perform on the process in order to obtain specific information, e.g. step responses, sinusoidal responses
etc.

- **Model set**- It has to be specified beforehand within which set of models one is going to evaluate the most accurate model for the process. In the model set several basic properties of the models have to be fixed, e.g. linearity/non-linearity, time in-variance, discrete/continuous-time, and other structural properties e.g. the order of the models.

- **Identification criterion**- Given measurement data and a model set, one has to specify in which way the optimal model from the model set is going to be determined. In applying the criterion, the models in the model set are going to be confronted with the measurement data.

In all three different aspects, a prior knowledge about the system to be identified can play a vital role. Given specific choices for the three phenomena described above, it is generally a matter of numerical optimization to construct an identified model.

Whereas most identification techniques are developed and analyzed in the time domain, the frequency domain also offers a multitude of methods and tools, and sometimes particular advantages. In the course of years, the difference between the two domains has become less, and has been characterized as a difference between the used excitation signals, being periodic or not. An account of this development can be found in Schoukens et al. [Schoukens and Pintelon, 2014] and Pintelon et al. work [Pintelon and Schoukens, 2012].

A description of a system should specify how the output signal(s) depend on the input signal(s). In the following sections two representations considered in the frequency and in the time domain are explained briefly. For a long time, frequency domain
identification and time domain identification were considered as competing methods to solve the same problem—building a model for a linear time-invariant dynamic system, but in the end, the frequency domain approach got a negative reputation because the transformation of the data from the time domain to the frequency domain is prone to leakage errors, i.e. noiseless data in the time domain resulted in noisy frequency response function (FRF) measurements [Schoukens et al., 2004].

2.2.1 Frequency domain identification

The majority of the measurements originating from real-world devices intrinsically belong to the time domain, and consequently, system identification methods and the theory developed around those deals with how to determine models from such time domain measurements in general [Ljung, 1999, Söderström and Stoica, 1989]. However, in some application areas such as vibration analysis, it is common to subject the raw data to the Fourier transform before fitting them to parametric models—a classical technique well known as ‘frequency analysis’. In frequency analysis the linear dynamical system is excited by a pure sinusoidal signal. When the output has settled to a stationary sinusoidal signal, the complex value of the transfer function at the specific excitation frequency is determined by comparing the amplitudes and phases of the input and output signals, respectively. Repeating the experiment for many frequencies yields a non-parametric estimate of the system’s frequency response. In a second step, a parameterized transfer function model can then be fitted to the transfer function data using some complex curve fitting techniques [Levy, 1959]. During the last decade the frequency domain techniques have received much attention in the system identification literature [Schoukens and Pintelon, 2014, Pintelon et al., 1994, Ljung, 1999].

The relation between the Fourier transforms of input and output signal(s) gives a
frequency domain representation of a particular system. Frequency response of the system and its models provide valuable insight. A distinctive feature of frequency domain techniques is that the modeling of continuous-time systems from sampled data can be done in a straightforward fashion if a certain class of band limited excitation signals is employed. This is a great advantage in contrast to the rather involved time domain techniques, which even in the noise-free case are only approximate if a finite set of sampled data is available. A continuous-time system with a time delay is also rather difficult to model in the time domain because it cannot be described by a finite dimensional system of ordinary differential equations. However, in the frequency domain, a nice finite-dimensional parametric description exists that lends itself to identification using parametric methods. Following are some of the features of frequency domain approach contrary time domain counterpart [McKelvey, 2002]:

- **Partial modeling**- Often it is sufficient to find a model that accurately describes the true system in a limited frequency band. A low order model could thus be sufficient rather than to fit a more complex model at all frequencies. In the frequency domain, this can be simply accomplished by fitting a model only at the desired frequencies, which corresponds to use an ideal bandpass filter on the raw time domain data.

- **Continuous time systems**- If the experimental conditions are such that a multiple sine input can be used, then modeling in the frequency domain is straightforward. In this case the Fourier transformed data is exactly described by the continuous-time frequency function. Systems with time delay are also easy to describe in the frequency domain.

- **Merging data**- If data is obtained by different experiments, all frequency data can be merged into one data set. Continuous-time models that are valid for
large frequency ranges can be estimated from data sets obtained from several experiments, each using a different sampling frequency.

- **Equivalency** - Frequency domain identification can deal equivalently with time continuous as with time discrete models.

### 2.2.2 Time domain identification

When the data considered during the identification process is taken the form of time series the method is known to be a time-domain method. Such techniques have the advantage that the signals are directly provided by current measurement devices; thus spending less time and effort on data acquisition and processing.

One of the approaches to time-domain identification, the restoring force surface (RFS) method, began with Masri et al. work [Masri and Caughey, 1979]; a parallel approach named force-state mapping was developed independently as well [Crawley and O’Donnell, 1986, Crawley and Aubert, 1986]. The RFS method initiated the analysis of nonlinear structural systems in terms of their internal RFSs. However, the initial version depended on the rather arbitrary use of Chebyshev polynomials for the expansion of the nonlinear restoring forces making numerical analysis rather complicated. The approach also suffered from bias unless the identification was iterated, and this made the process time-consuming.

A technique which was widely applied in Control Engineering at first, but was taken up by structural dynamicists, was time-series analysis. The linear variant of the approach based on ARMA (Auto-Regressive Moving Average) models has long been used for modeling and prediction purposes [Box et al., 2015]. There have been numerous
attempts to generalize the model structure to the nonlinear case, arguably the most versatile and enduring structure has been the NARMAX (Nonlinear ARMA with eX-ogeneous input) model proposed by Leontaritis and Billings [Leonaritis and Billings, 1985]. Since the inception of the method, there have been many developments, notably the introduction of an orthogonal estimation algorithm [KORENBERG et al., 1988], which allows model parameters to be estimated sequentially so that the complexity of the model can be controlled. Also noteworthy are the correlation tests designed to assess model validity [Billings et al., 1989]. The NARMAX structure is general enough to admit many forms of model including neural networks although the estimation problem becomes nonlinear and the orthogonal estimator will not work [Billings et al., 1992]. Several other time-domain techniques have been proposed in the literature.

2.3 Continuous Time identification

The field of system identification has grown in size and diversity over several decades and is now a matured field. Aström and Eykhoff in [Åström and Eykhoff, 1971] presented a survey mainly focused on system identification in discrete-time. A first significant development in the field of continuous-time system identification is a survey report by Young in [Young, 1981], which is a review on the progress of research on parameter estimation of dynamic systems in continuous-time. Subsequently, rapid developments were made in this field, which is described in surveys on continuous-time system identification by Unbehauen and Rao [Unbehauen and Rao, 1990, Unbehauen and Rao, 1998]. Furthermore, several books [Sinha and Rao, 2012, Ljung, 1999] and publications [Ding et al., 2009, Garnier et al., 2003, Rao and Unbehauen, 2006, Wang
and Zhang, 2001, Hwang and Lai, 2004] are found on the subject, which are widely discussed.

Contrast to the present day, the control world of 1950s and 1960s was dominated by CT models as most control system design was concerned with CT systems and most control system implementations employed analogue techniques. Moreover, almost all CT identification methods were largely deterministic, in the sense that they did not explicitly model the additive noise process nor attempt to quantify the statistical properties of parameter estimates. Nevertheless, it is fascinating to see that some of the early papers introduced interesting concepts that foreshadowed later. For instance Young [Young, 1964, Young, 1965] suggested the use of pre-filters to solve the derivative measurement problem and later ‘state-variable filter’ (SVF) approach [Saha and Rao, 1983] was rediscovered under the title ‘Poisson-moment functionals’ (PMF). Early state of the research used complete analogue implementation with both pre-filters and estimation algorithm which afterwards developed to be a hybrid implementation consisting analogue pre-filtering combined with digital identification algorithm. Besides, non-linear system identification using a purely deterministic ‘state-dependent parameter’ approach was also attempted in the early days of identification [Hoberock and Kohr, 1967, Lion, 1967].

Later in 1960s, it was realized that measurement noise could cause asymptotic bias on the parameter estimates when linear least-squares method was used to estimate the parameters in dynamic systems unless the additive measurement perturbation was zero-mean white noise, which in reality was impractical as the corrupting noise was correlated. As a solution to the problem, Young [Young, 1970] proposed an instrumental variable (IV) method to generate unbiased estimates of the parameters,
which was highly appreciated, adopted by the many in the identification research community followed up with new research opportunities. Although both the LS and IV methods worked perfectly for non-delay systems or systems with a known input delay, after a while issues raised regarding complex industrial plants comprising input delays. Identification of such delays along with the parameters of the continuous models was definitely a challenging issue and was later resolved through gradient search approach [Ferretti et al., 1991, Zhao and Sagara, 1991], dedicated three step procedure [Kozłowski and Kowalczuk, 2009], on-line method based on Taylor’s expansion of the delayed input [Kozłowski and Kowalczuk, 2015].

In 1970s, a dominant interest in DT identification and estimation did not let a stochastic formulation of CT estimation to appear until 1980. Following the optimal pre-filtering and recursive iterative estimation procedures for DT systems, Young and Jakeman suggested an optimal 'hybrid' refined instrumental variable solution to the CT identification (RIVC) problem [YOUNG and JAKEMAN, 1980, Young, 1976]. However, it was implemented in a simplified form (SRIVC) involving a CT model of the system and a discrete time ARMA model for the noise yielding statistically efficient parameter estimates when the additive noise was white in nature. Responding to the research on RIVC estimation, Huang et al. implemented an alternative hybrid solution that allowed for colored noise and utilized gradient optimization algorithm rather than iterative solution used in SRIVC algorithm [HUANG et al., 1987]. However the entire study was done considering a DT form and thus implementation of the pre-filters and auxiliary model was not performed explicitly in continuous time.

Two other approaches that have attracted a lot of attention in the identification community in the 1990s are sub-spaced based methods and finite difference methods-
replacing the differentiation operator with finite differences [Bastogne et al., 2001, Li et al., 2003, Pham, 2000, Soderstrom et al., 1997]. More recently, stochastic model identification with optimal CT estimation procedure has attracted researchers of this field. Initiated by Wang and Gawthrop’s work on optimal CT identification [Wang and Gawthrop, 2001], Young drew attention to the virtues of the existing SRIVC estimation algorithm and demonstrated its superiority [Young, 2002], leading towards implementing hybrid RIVC algorithm [Young et al., 2006] together with the development of associated closed-loop identification algorithm [Gilson et al., 2006]; as a result of which, optimal RIV algorithms for Box-Jenkins type stochastic transfer function models of CT and DT systems are now available, providing a unified approach to the identification and estimation of transfer function models [Young, 2008].

Lately, the problem of system identification from irregularly sampled time instants has received much attention as it is commonly used in time-series analysis, radar imaging, medical imaging, bio-medicine etc. [Adorf, 1995]. A main concern when dealing with irregularly observed data is that computational complexity of conventional methods increases substantially. Another inevitable fact is that missing data scenario will result in loss of information eventually impacting parameter estimation process. Consequently, CT modeling with direct estimation of parameters from measured data has been proven to be a way forward in a number of applications although some of the mathematics associated with CT stochastic dynamic system is more complicated than the corresponding theory for DT systems.
2.4 Fractional Order model identification

Studying various phenomenon of diverse physical processes and to achieve control over them, mathematical framework representing systems is indispensable. And as there are many processes that can be more accurately modeled using fractional differ-integrals, fractional order systems has been adopted by the contemporary fields of science in order to extend our notion of modeling the real world around us for better understanding and perception [Das and Pan, 2011, chapter 2]. Literature [Torvik and Bagley, 1984, Podlubny, 1994, Caponetto, 2010, Petras, 2011] says, with the ability of expressing in a compact manner, real dynamical systems are better characterized with fractional order differential equations. Therefore, in recent years, fractional calculus has been applied in modeling and control of various kinds of physical systems, and is well documented in control theory and application literature; whereas for past 300 years, fractional calculus did not had much clear physical and geometric interpretations in general. A typical example of a non-integer (fractional) order system is the voltage-current relation of a semi-infinite lossy transmission line [Wang, 1987] or diffusion of the heat through a semi-infinite solid, where heat flow is equal to the half-derivative of the temperature [Podlubny, 1998].

Based on the experimental data of a solid-core magnetic bearing (MB), a comparison between fractional-order system identification with its equivalent integer-order system identification reveals that the FO model is more effective describing the dynamics of the MB as well as the FO controller can significantly improve the transient and steady-state performance of the MB system [Zhong and Li, 2014]. Although a numerous number of references can be found dealing with the identifiability for both structure and parameters of a model set regarding integer order systems, it was Nazarian et al. work which first documented this phenomenon for fractional order models.
Their results revealed that the identifiability is lost for smaller commensurate order ($\alpha$) and observations suggested that determining a unique set of parameters and reaching a unique structure (order combination) for fractional order models would be difficult and often impossible; leading one to reconsider the identifiability of parameters and use identifiability of the system instead.

2.4.1 Significance of fractional derivative

According to experimental observation and or analytic solution, both the time domain and frequency domain behaviors of some linear systems and processes do not fit the standard laws, i.e., exponential evolution in time domain or integer-order slopes in their frequency responses [Monje et al., 2010, Chapter 14]. Whereas in the time domain, these complicated dynamics can be described by generalized hyperbolic functions, but in order to fit the non-integer-order slopes in the frequency responses, irrational-order transfer functions: transfer functions constructed as products of zeros and poles of fractional power or ratios of polynomials in $s^\alpha$ ($\alpha$ being non-integer), is being introduced. Such progression has led fractional-order integration and differentiation unlike integer-order derivatives and integrals to represent a rapidly growing field both in theory and in applications to real world problems [Dalir and Bashour, 2010]. Some fundamental definitions of fractional order operators in both time and Laplace domain, together with their dynamic properties, as well as analytical and numerical solutions of the fractional-order ordinary differential equations is briefly documented in Monje et al. book [Monje et al., 2010, chapter 2]. In addition, several other definitions on fractional derivatives and integrals including some elementary functions, explicit formula of fractional derivative and integral together with some applications of fractional calculus in science and engineering is elucidated in literature [Miller and
Ross, 1993, Dalir and Bashour, 2010, Tavassoli et al., 2013].

Gutiérrez et al. review article [Gutiérrez et al., 2010], concisely explains several concepts of fractional order calculus (FOC), with some of its applicability in system identification, control, mechatronics etc., and furthermore reviews several approaches on geometrical interpretation of FOC. Detailed literature on FOC can be found in [Magin et al., 2011, Ortigueira, 2000a, Ortigueira, 2000b, Ortigueira, 2003]. Utilizing Riemann-Liouville definition [Samko et al., 1993, Chapter 1] for fractional integration and differentiation, some approximation methods for fractional-order operators both in continuous and discrete time models are briefly explained and later illustrated and compared considering fractional integrator of order 0.5 by Vinagre et al. [Vinagre et al., 2000]. Integer order dynamic systems consisting of very high number of parameters can be approximated in a fractional model yet using only few parameters, due to its “long memory” characteristic; and therefore such usage of fractional derivative is named as reduced-parameters modeling or model compression and has started becoming attractive for analysis and design of large systems [Mansouri et al., 2010]. Some application of fractional order models include but are not limited to modeling of isotope separation columns [Dulf et al., 2012], bio-reactors [Ahmad and Abdel-Jabbar, 2006], pressurized heavy water reactors [Das et al., 2011], liquid/liquid interfaces [Spasic and Lazarevic, 2005], biological systems [Ionescu et al., 2011], thermal systems [Narang et al., 2011], hydro-logic processes [Benson et al., 2013]. Especially in bio-engineering, many real systems are modeled or fitted by fractional order systems [Magin, 2006].
2.4.2 Fractional order controller

In the last two decades the possibility of using fractional order controller has been considered [Oustaloup et al., 2000, Podlubny, 1999, Podlubny, 1998]. It is established that fractional order controllers are more robust in nature compared to integer ones [Chen, 2006], as they require less coefficients [Xue et al., 2006] and can capture complex behaviors [Gutiérrez et al., 2010]. A detailed study on typical fractional controllers can be found in literature [Xue and Chen, 2002, Chen, 2006, Chen et al., 2009, Gutiérrez et al., 2010]. The idea of using fractional-order controllers for the control of dynamic systems belongs to Oustaloup, who developed the “Commande Robuste d’Ordre Non Entier” (CRONE) controller [Oustaloup et al., 1995]. Indeed, Ouastaloup with the CRONE controller (Commande Robuste d’Ordre Non Entier controller) and Podlubny with the $PI^\lambda D^\mu$ controller [Podlubny, 1999], involving the fractional-order integrator and the fractional-order differentiator, have demonstrated the advantages of the fractional-controllers over the classical ones. However, before the actual design of controllers for dynamical systems it is necessary to identify these systems [Torvik and Bagley, 1984].

The stability analysis of fractional order linear systems using the technique of Root Locus (RL) through performing a transformation of a FO system into its integer order counterpart is proven to be simpler [Patil et al., 2014]. A novel auto-tuning method for fractional order PI or PD controllers that yields a robust controller despite the lack of an actual process model is presented in a work by Keyser et al. that leads towards similar results when compared with the tuning of a fractional order PI/PD controller based on classical approach (with knowledge of the process model) [De Keyser et al., 2016]. Several design methods regarding parameter tuning of fractional order controllers can be found in literature [Vu and Lee, 2013, Yeroglu and Tan,
A relatively new approach to control the linear fractional order systems of arbitrary order by designing fractional order predictive functional control (PFC) can be found in literature [Bigdeli, 2015]. Where, at first, the non-minimal input–output fractional-order state space (NMSS) model of the system has been derived. Afterwards, through defining a fractional order cost function over the fractional non-minimal state vector, the fractional-order predictive functional controller (PFC) has been designed for the NMSS model structure. Finally, a genetic algorithm has been employed to obtain the optimal PFC control coefficients. Besides, any change to an existing control loop may lead towards termination of an industrial process and thereby potentially result in production losses. In order to deal with this issue, integrating a fractional-order controller into a working loop in a non-intrusive way through tapping system’s input and output signals has been the main focus of Tepljakov et al. work [Tepljakov et al., 2016].

Implementing fractional order modeling to fractional-order controller has improved the control of the real systems through achieving a better trade-off between dynamic performance and robust stability, which in return has been beneficial to industrial processes, automotive systems, mechatronics, robotic systems, unmanned vehicles etc. In the recent literature, application of fractional order controller has been highlighted on the position control electrodes of an industrial electric arc furnace which required tuning of three parameters only and thus saving time for designing [Feliu-Batlle and Rivas-Perez, 2016], profitable execution of CRONE methodology to control a high dynamic engine test bed coupled with a spark-ignition engine from system identification based on frequency-domain approach [Lanusse et al., 2016], solving heat generation for domestic and industrial purposes via solar furnace [Beschi et al., 2016], prevention of blade icing and deicing [Sabatier et al., 2016], regarding vibration control in smart
structures [Feliu-Talegon et al., 2016] and many more. In fact, these recent works in the field of fractional calculus has made the researches realize that fractional calculus is indeed a viable mathematical tool that will accomplish far more surpassing boundaries of integral order calculus and is meant to be for the future. A survey on theoretical developments of fractional calculus and applications of fractional modeling is documented in [Sabatier et al., 2007].

2.4.3 Identification methods

Liu et al. article [Liu et al., 2013b] works both as a survey and review on several identification methods for industrial processes that have been developed over the last three decades or so. In this study, authors mainly adopted time delay systems of first and second order integer models for classification and categorized them into two separate groups: 1. Step identification and 2. Relay identification, considering both zero and nonzero initial process conditions and or load disturbance. Additionally, non-linear systems were also taken into account using multiple or modified step or relay tests. Maiti et al. have considered unit step signal as test input signal to generate output response of a fractional order process model with three unknown parameters in their work [Maiti et al., 2008]. Later, using two very basic definitions: the Riemann-Liouville and the Grunwald-Letnikov [Samko et al., 1993, Chapter 1,4], the input signal of the system was rewritten in terms of coefficients and output response of the system. As there were three parameters, to identify them, the expression relating the input and output were integrated two more times consecutively to formulate three simultaneous equations, solving which ultimately led to estimate the parameters. The first part of this study required the fractional powers of the system to be known to demonstrate the accuracy of the identification technique and the second part dealt
with identifying the fractional powers of the system using Particle Swarm optimization (PSO) [Song and Gu, 2004] algorithm. However, this article lacks to deal with systems that have time delayed response. Tavakoli-Kakhki and Tavazoei in their work [Tavakoli-Kakhki and Tavazoei, 2014], considered an unstable fractional ordered first order system with dead time (delayed time). Then a proportional controller was developed to stabilize the system in order to get the closed loop step response data which were later used to estimate the order and parameters of the process. However, the limitation of this article was that the proposed methodology could only work for systems that are fractional counterparts of first order system. In Malek et al. work [Malek et al., 2013], although authors mainly focused on designing a fractional order proportional integral controller, one of the sections dealt with identification of the system for which the controller was to be designed. Afterwards, for illustration purpose, authors identified the parameters of a heat flow equipment using unconstrained nonlinear optimization. Limitation of such identification technique involved extensive calculation and prior knowledge of the fractional order of the process. For identification purpose of fractional order models, Malti et al. extended the existing algorithm of simplified refined instrumental variable for continuous-time systems (srivc) to fractional models (srivcf) [Malti et al., 2008b]. Again this work lacks to estimate the fractional order of the system. Combining The well-known srivc algorithm with two other criterion- 1. Young information criterion (YIC) and 2. $R_T^2$ criterion, Victor and Malti proposed a new algorithm named order-optimization-srivcf (oosrivcf) to estimate coefficients and fractional order of a transfer function [Victor and Malti, 2013]. Similar work can be found in [Victor et al., 2013]. A unique yet efficient way to estimate parameters and fractional order of a model was proposed in [Narang et al., 2011]. First, with an initial assumption of system order and coefficients, applying least square and instrumental variable technique, the parameters were estimated. Afterwards, the value of system
order was updated based on iterative repetition until the difference between two consecutive iterations was less that $10^{-4}$. And hence the optimal values for parameters and order were estimated. Identification of fractional order models considering step response using integral equation approach was observed in [Ahmed, 2015]. Later simulations were carried out to demonstrate the efficacy of the proposed methodology. Nevertheless, order estimation for such system were beyond the scope of this study.

Based on the process step response, a new model identification technique has been documented for a class of delay fractional-order system [Nie et al., 2016]. Two identification schemes, first by utilizing three exact points on the step response of the process and secondly by employing optimal searching to adjust the fractional order, model parameters were calculated for higher-order, under-damped/over-damped, and minimum phase/non-minimum phase processes. Another approach on parameter identification of Fractional Order System (FOS) based on Haar wavelet operational matrix is detailed in Li et al. work [Li et al., 2015b]. Where, taking use of the Haar wavelet in order to represent input, output signals as well as operational matrix of fractional order integration, a system was converted to a sum to algebraic equations. Afterwards, based on subspace technique and non linear programming, parameter matrices and the commensurate orders were identified respectively solving the non linear optimization function for minimum error between the output of the real system and the identified system. One of the challenges while following this technique lies in expanding the Haar wavelet, where first $N$ (power of 2) terms of Haar coefficients and Haar/Heaviside function is a prerequisite. Selecting this $N$ requires expert knowledge. In summary, although this method is simple and easily implementable with expert process knowledge, it may also require higher computational time for greater value of $N$. 

32
In contrast to the conventional methods of analyzing fractional order systems by means of determining fairly accurate results based on minimum error optimization algorithm, a recent approach underpinning formalization of the Grünwald–Letnikov (GL) definition in a higher-order logic (HOL) theorem proving system can be found in [Zhao et al., 2016]. Some recent practices dealing with the issue of estimating parameters although involves State Variable Filter (SVF), Poisson moment function as filter, nevertheless avoids selection procedure of acquiring optimal filters. To overcome the scenario, Dai et al. focused on the exact study of using modulating functions for modulation in identification for fractional order systems [Dai et al., 2016]; a concept, first introduced in [Liu et al., 2013a], yet lacking real modulation itself. The outlining factor of Dai et al. work was to estimate parameters without prior initial conditions with the drawback of known fractional orders of the models.

Utilizing Levenberg-Marquardt method, another identification algorithm for fractional systems in frequency domain compares two different methods, method I- excluding steady state gain and method II- introducing steady state gain as a constant [Li et al., 2015a]. However, a proper selection of the frequency range was necessary with a future scope of work on better initializing the unknown estimates. To eliminate the estimation bias, a combination between the least-squares estimator and state variable filter, termed as bias correction method has been proposed by Yakoub et al. [Yakoub et al., 2015]. Later applying a nonlinear optimization, both coefficients and commensurate-order of a process were estimated. Nevertheless, being an indirect approach, where parameter estimation of fractional process requires parameter estimation of fractional closed-loop system, the main challenge lies with the computation time and complexity with this procedure.
Some recent development on fractional-order modeling and parameter identification can be observed for lithium-ion batteries- derived from a modified Randles model applying hybrid multi-swarm particle swarm optimization (HMPSO) [Wang et al., 2015], thermal dynamics of buildings- formulated by fractional order auto-regressive model with exogenous input using least-squares technique [Chen et al., 2016], Polymer Electrolyte Membrane Fuel Cell systems [Taleb et al., 2017], ultracapacitor- using cubic spline interpolation technique on linear parameter varying (LPV) model [Gabano et al., 2015], controlled auto regressive moving average (CARMA) systems- based on fractional least mean squares identification (FLMSI) algorithm [Raja and Chaudhary, 2015], diffusion process- modeled via lumped RC network [Sierociuk et al., 2015], thermal conductivity and diffusivity with constrained fractional order- implementing Levenberg–Marquardt algorithm, a combination of Gradient method and Gauss–Newton method to better achieve stability and convergence [Gabano and Poinot, 2009], three-dimensional random RC network- adopting flexible polyhedron algorithm [Galvão et al., 2013].
Chapter 3

Comparative Study on Continuous Time Integer Order Model Identification

The main objective of this study is to evaluate the performance of a set of recently developed approaches for continuous-time identification. Three major approaches for identification using sinusoidal response are considered (i) direct identification using the integral equation approach (ii) by estimating the step response from the sinusoidal response and (iii) by linear approximation of the sinusoidal input. The effect of frequency of the input signal and that of the data length and the noise to signal ratio are studied in simulation. A first order plus time delay model is considered in the case study and the properties of the estimates of the gain, the time constant and the delay are studied. Monte Carlo simulations are performed to estimate the bias and variance of the estimated parameters. Using an average error criterion defined based on the estimated bias and variance of the parameters, the performance of the methods are compared. The performances of the three algorithms were found to be comparable at
low to mid level frequencies of the input. However, at high frequencies of the input, the performance of the piece-wise linear approximation method deteriorated.

### 3.1 Introduction

The integral equation approach [Diamessis, 1965] is a well-known technique for parameter estimation of continuous-time transfer function models. However, use of this technique has so far been limited to step and step-like input signals [Hwang and Lai, 2004, Liu et al., 2007, Ahmed et al., 2007]. Ahmed [Ahmed, 2014] recently proposed an integral equation approach for direct identification of the parameters and the delay using the integral equation approach. The sinusoid has a distinct advantage over some other signals due to its smooth nature which helps it to change variables gradually. Besides, use of the sinusoid as input gives the opportunity to excite process at the frequencies of interest and thus helps to obtain precise frequency response of the plant.

In recent times using sinusoidal input for system identification has become popular [Kalafatis et al., 2005, Zaremba and Pavlov, 2002, Hwang et al., 2004]. At the same time identification from step response has some distinct advantages [Ahmed et al., 2007, Wang and Zhang, 2001, Wang et al., 2001]. To combine these two procedures, Ahmed et al. [Ahmed et al., 2009] proposed a novel idea of transforming the sinusoidal response to step response by passing it through a linear filter and then estimating the process model parameters using the integral equation approach.

Any signal can be expressed as an amalgamation of linear approximated signals for small sampling intervals. With this idea, a sinusoid was represented approximately with piece wise linear signals and subsequently model parameters were extracted.
[Ahmed, 2010a] using the integral equation approach. This technique turned out to be computationally less complex than estimating model parameters directly from the sinusoidal or step response.

In this article, using the above mentioned three methods, the parameters for a FOPTD model were estimated to evaluate the performance of these algorithms with varying noise to signal ratio (NSR), input frequency and data length. The estimated model parameters were compared using an average error criterion [S.Ahmed, 2006]. The following section briefly describes the above mentioned methodologies.

3.2 Methodology

Let us consider a first order plus time delay (FOPTD) system stated by the Equation 3.1

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0}{s + a_0} e^{-\delta s}$$  \hspace{2cm} (3.1)

where, $G(S)$ is the transfer function between the input, $U(S)$ and the output $Y(S)$; $[a_0, b_0, \delta]$ are the model parameters. The concept is, using sinusoidal signal as an input and introducing measurement noise we will generate noisy response, with which following the later mentioned three methods, the model parameters i.e. time constant, gain and time delay of the system will be estimated; however the methods are applicable to higher order models as well.
3.2.1 The direct approach [Ahmed, 2014]

In the time domain, a sinusoidal input with amplitude $\alpha$, frequency $\omega$, and phase $\nu$ is expressed as follows

$$u(t) = \alpha \sin(\omega t + \nu) \quad (3.2)$$

In the Laplace domain Equation 4.4 can be written as

$$U(s) = \alpha \frac{\sin(\nu)s + \omega \cos(\nu)}{s^2 + \omega^2} \quad (3.3)$$

Considering $\mu = \alpha \sin(\nu)$, $\lambda = \alpha \omega \cos(\nu)$ and $\beta = \omega^2$, the above Equation 4.5 can be rewritten as following

$$U(s) = \frac{\mu s + \lambda}{s^2 + \beta} \quad (3.4)$$

Considering the initial transient state of the process output, $y(0)$, is zero, Equation 3.1 can be expressed in the following equation error format as

$$\begin{bmatrix} s^3 + \beta s \end{bmatrix} Y(s) + a_0 \begin{bmatrix} s^2 + \beta \end{bmatrix} Y(s) = [\mu s + \lambda] b_0 e^{-\delta s} + E(s) \quad (3.5)$$

Taking inverse Laplace transform of Equation 4.26, we end up with the time domain equation as

$$\begin{bmatrix} y(t) + \beta y^{[2]}(t) \end{bmatrix} + a_0 \begin{bmatrix} y^{[1]}(t) + \beta y^{[3]}(t) \end{bmatrix} = b_0 \mu[t - \delta]$$

$$+ b_0 \lambda \frac{(t - \delta)^2}{2!} + \xi(t) \quad (3.6)$$

where $y^{[k]}(t)$ denotes the $k$-th order integral of $y(t)$. To estimate parameters, Equation 3.6 is written in the following suitable form to use the least-squares solution technique.
\[
y(t) + \beta y^{[2]}(t) = -a_0 [y^{[1]}(t) + \beta y^{[3]}(t)] + b_0 \lambda \frac{t^2}{2!} + (b_0 \mu - b_0 \lambda \delta)t + (-b_0 \mu \delta + b_0 \lambda \delta^2 \frac{t^2}{2!}) + \xi(t)
\]

(3.7)

Or equivalently,
\[
\gamma(t) = \phi^T(t)\theta + \xi(t)
\]

(3.8)

where, \(\gamma(t) = y(t) + \beta y^{[2]}(t)\)

\[
\phi(t) = \begin{bmatrix} -(y^{[1]}(t) + \beta y^{[3]}(t)) \\ \lambda \frac{t^2}{2!} \\ \lambda t \\ \lambda \end{bmatrix} \quad \text{and} \quad \theta = \begin{bmatrix} a_0 \\ b_0 \\ b_0(\eta - \delta) \\ b_0(-\eta \delta + \frac{\delta^2}{2!}) \end{bmatrix}
\]

Here, \(\eta = \mu / \lambda\). Equation 4.9 can be written for \(t = t_{d+1}, t_{d+2}...t_N\) and then combined together to give a set of estimation equations

\[
\Gamma = \Phi \theta + \Xi
\]

(3.9)

with

\[
\Gamma(t) = \begin{bmatrix} \gamma(t_{d+1}) \\ \gamma(t_{d+2}) \\ \vdots \\ \gamma(t_N) \end{bmatrix} \quad \text{and} \quad \Phi = \begin{bmatrix} \phi^T(t_{d+1}) \\ \phi^T(t_{d+2}) \\ \vdots \\ \phi^T(t_N) \end{bmatrix}
\]

Here, \(d\) is the time delay in terms of number of sampling intervals (\(\Delta t\)), i.e. \(d = \delta / \Delta t\) and \(N\) is the total no of samples available. When the time delay is not an integer multiple of sampling interval, \(d\) is chosen as the nearest integer in the positive direction. Finally, the least-squares (LS) solution of the estimation equations in.
Equation 4.10 gives us the estimated parameters of the system

$$\theta^{LS} = (\Phi^T \Phi)^{-1} \Phi^T \Gamma$$  \hspace{1cm} (3.10)

Due to integration operation, the LS solution may be biased even for a white measurement noise. To get an unbiased estimate, the instrumental variable (IV) method is implemented [Young, 1970]. To generate the instruments, the LS solution of the parameters are used to get the predicted output. The instrument vector, \( \psi(t) \) is then derived by replacing the terms related to the output, \( y(t) \), in the regressor by their predicted values, \( \hat{y}(t) \). Afterwards, \( \psi(t) \) is written for \( t = t_{d+1}, t_{d+1}...t_N \) and combined to get the instrument matrix \( \Psi \). The instrumental variable estimate of the parameters is given by

$$\theta^{IV} = (\Psi^T \Phi)^{-1} \Psi^T \Gamma$$  \hspace{1cm} (3.11)

From Equation 3.11, the parameters \( a_0 \), \( b_0 \) can be directly obtained and delta can be obtained as \( \delta = \eta - [\theta(3)/\theta(2)] \). Thus, we get the entire set of parameters \([a_0 \ b_0 \ \delta]\) for the FOPTD model.

3.2.2 Identification by estimating step response [Ahmed et al., 2009]

Equation 3.1 can be rewritten as

$$Y(s) = G(s)U(s)$$  \hspace{1cm} (3.12)
If the input is a unit step, we have $U(s) = \frac{1}{s}$ and the unit step response, $Y_{\text{step}}(s)$ can be obtained as

$$Y_{\text{step}}(s) = G(s)\frac{1}{s} \quad (3.13)$$

Comparing Equation 3.12 and 3.13, we get the relation to obtain the unit step response from output data due to other type of input signal, i.e. in this case the sinusoidal input.

$$Y_{\text{step}}(s) = \frac{Y(s)}{sU(s)} \quad (3.14)$$

Here, $Y(s)$ is the response due to the same sinusoidal input that was used in previous section and for a deterministic input, $U(s)$, which is a sinusoid in this case, can be obtained mathematically. Now, considering the term $\frac{1}{sU(s)}$ as a filter, $F(s)$, which is known, from Equation 3.14, it can be said that the step response of the unknown process can be obtained from its output due to a sinusoidal input and then by passing it through the filter.

Using Equation 3.4, for a sinusoidal input with single frequency, the filter becomes

$$F(s) = \frac{1}{sU(s)} = \frac{s^2 + \beta}{s(\mu s + \lambda)} \quad (3.15)$$

Using Equation 3.14 and Equation 3.15, we get the relation to obtain the step response from sinusoidal response as

$$Y_{\text{step}}(s) = \frac{s^2 + \beta}{s(\mu s + \lambda)}Y_{\text{sine}}(s) \quad (3.16)$$

As can be seen from Equation 3.16, for $F(s)$ to be stable, the phase $(\psi)$ of the sinusoid needs to be bounded by $0 \leq \psi \leq \pi/2$. But the issue arises when the phase $(\psi)$ of the
input sinusoid becomes zero, making the filter $F(s)$ unstable with lesser number of poles compared to zeros. Besides, with zero phase ($\nu$), the filtering involves direct differentiation of the output signal which is not desired for the noisy outputs. To solve this problem, we used the linear filtering technique through defining an extra filter $F'(s)$ in Laplace domain as $\frac{1}{\sigma s + 1}$ with a known value of $\sigma$. In this article, for simplicity, we have used the value of $\sigma$ as 1 for parameter estimation purpose. Afterwards, the filtered unit step response of the system, $Y_{step}(s)$, was obtained from Equation 3.16 as

$$Y_{step}(s) = \frac{s^2 + \beta}{s^2 \sigma + s^2 (\mu + \sigma \lambda)} Y_{sine}(s)$$

(3.17)

As in this particular case, we obtain the step response of the augmented system $G'(s) = G(s)F'(s)$, the parameters have to be estimated from the augmented system rather than the original process.

Considering the same FOPTD system used in the previous section, the relation between input and output of the augmented system in Laplace domain is as followed

$$Y_{step}(s) = \frac{b_0}{s + a_0} e^{-\delta s} \frac{1}{\sigma s + 1} U_{step}(s) + E(s)$$

(3.18)

For unit step, using $U_{step}(s) = 1/s$, Equation 3.18 becomes

$$[\sigma s^2 + s] Y_{step}(s) = -a_0[\sigma s + 1] Y_{step}(s) + b_0 e^{-\delta s} \frac{1}{s} + E(s)$$

(3.19)

Taking inverse Laplace of Equation 3.19, assuming zero initial condition, and after-
wards following integral equation approach, we get,

\[
\sigma y(t) + y^{[1]}(t) = -a_0[\sigma y^{[1]}(t) + y^{[2]}(t)] \\
+ b_0 u^{[2]}[t - \delta] + \xi(t)
\] (3.20)

For a unit step input applied at time \( t = 0 \), the following integral holds for \( t \geq \delta \)

\[
u^{[k]} = \frac{[t - \delta]^k}{k!}
\] (3.21)

where, \( k \) is the order of the integral. Using Equation 3.20 and 3.21, rearrangement leads to following least-square formulation

\[
\sigma y(t) + y^{[1]}(t) = -a_0[\sigma y^{[1]}(t) + y^{[2]}(t)] \\
+ b_0 \frac{t^2}{2!} - b_0 \delta t + \frac{b_0 \delta^2}{2!} + \xi(t)
\] (3.22)

Or equivalently,

\[
\gamma_s(t) = \phi_s^T(t)\theta_s + \xi(t)
\] (3.23)

where,

\[
\gamma_s(t) = \sigma y(t) + y^{[1]}(t)
\]

\[
\phi_s(t) = 
\begin{bmatrix}
-(\sigma y^{[1]}(t) + y^{[2]}(t)) \\
\frac{t^2}{2!} \\
-t \\
1
\end{bmatrix}
\quad \text{and} \quad
\theta_s = 
\begin{bmatrix}
a_0 \\
b_0 \\
b_0 \delta \\
\frac{b_0 \delta^2}{2!}
\end{bmatrix}
\]

Equation 3.23 can be written for \( t = t_{d+1}, t_{d+2}, \ldots t_N \) and then combined together to give a set of estimation equations to solve for the model parameters.
\[ \Gamma_s = \Phi_s \theta_s + \Xi \quad (3.24) \]

### 3.2.3 Identification by piece-wise linear approximation [Ahmed, 2010a]

Following integral equation approach, assuming zero initial condition the model equation can be written as

\[ y(t) + a_0 y^{[1]}(t) = b_0 u^{[1]}(t - \delta) + \xi(t) \quad (3.25) \]

If an input is piece wise linear, it can be mathematically expressed as

\[ u(t) = \sum_{i=0}^{N} \zeta_i [t - L_i] \Omega(t - L_i) \quad (3.26) \]

Here, \( i \) corresponds to the sampling instant, \( \zeta_i \) is the rate of change of slopes of the input signal at the \( i \)-th sample point, which can be obtained from Equation 3.27 and \( L_i = t_{i-1} \).

\[ \zeta_i = \epsilon_i - \epsilon_{i-1} \quad (3.27) \]

where, \( \epsilon_i \)s are the slopes of the signals at the sampling instants, estimated through backward approximation i.e.

\[ \epsilon_i = \frac{u_i - u_{i-1}}{t_i - t_{i-1}} \quad (3.28) \]

and \( \Omega \) is the unit step signal defined as

\[ \Omega(t - L_i) = \begin{cases} 
0 & \text{for } (t - L_i) < 0 \\
1 & \text{for } (t - L_i) \geq 0
\end{cases} \quad (3.29) \]
For any \( t = t_k \), where \( t_k \) is the \( k \)-th sampling time, in Equation 3.26, for all the terms with \( i > k \), \( \Omega(t - L_i) = 0 \). So, for \( t = t_k \), we have

\[
u(t) = \sum_{i=0}^{k} \zeta_i [t - L_i] \Omega(t - L_i)
\] (3.30)

For such an input the delayed signal can be expressed as

\[
u(t - \delta) = \sum_{i=0}^{k} \zeta_i [t - L_i - \delta] \Omega(t - L_i - \delta)
\] (3.31)

For simplicity in the presentation we will use the notation \( \Omega_i = \Omega(t - L_i - \delta) \). Using the notation, the integral of the delayed input signal can be expressed as

\[
u^{[1]}(t - \delta) = \sum_{i=0}^{k} \zeta_i [t - L_i - \delta^2] \frac{\Omega_i}{2} - \sum_{i=0}^{k} \zeta_i [t - L_i] \delta \Omega_i + \sum_{i=0}^{k} \frac{\zeta_i \delta^2}{2} \Omega_i
\] (3.32)

Using Equation 3.25 and 3.32, the estimation equation then becomes

\[
y(t) = -a_0 \nu^{[1]}(t) + b_0 \sum_{i=0}^{k} \zeta_i [t - L_i - \delta^2] \frac{\Omega_i}{2} - b_0 \delta \sum_{i=0}^{k} \zeta_i [t - L_i] \Omega_i + \sum_{i=0}^{k} \zeta_i \Omega_i + \xi(t)
\] (3.33)

Or equivalently,

\[
\gamma_p(t) = \phi_p^T(t) \theta_p + \xi(t)
\] (3.34)

with \( \sum_{i=0}^{k} \zeta_i \Omega_i = \sum_{i=0}^{k} \zeta_i \) as the \( \sum \) does not have a \( t \) component, we denote \( \gamma_p(t) = y(t) \)
\[
\phi_p(t) = \begin{bmatrix}
-y^{[1]}(t) \\
\sum_{i=0}^{k} \zeta_i \frac{(t-L_i)^2}{2} \Omega_i \\
- \sum_{i=0}^{k} \zeta_i [t - L_i] \Omega_i \\
\sum_{i=0}^{k} \zeta_i \Omega_i
\end{bmatrix}
\] and \( \theta_p = \begin{bmatrix}
a_0 \\
b_0 \\
b_0 \delta \\
a_0 \delta^2
\end{bmatrix} \)

Equation 3.34 can be written for \( t = t_{d+1}, t_{d+2}...t_N \) and then combined together to give a set of estimation equations to solve for the model parameters.

\[ \Gamma_p = \Phi_p \theta_p + \Xi \]  (3.35)

### 3.3 Simulation study

For simulation purpose, we have considered the FOPTD model to be \( \frac{1.25}{20s+1} e^{-7s} \). Simulations were carried out with a fixed value of amplitude and phase of the sinusoid which were in this case 20 and 0 rad respectively. In each case, the parameters were obtained for 50 Monte-Carlo simulations (MCS), and then based on the obtained parameters, the methods were compared based on the average error criterion defined as

\[
E_{avg} = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} \frac{(\bar{\theta}(i))^2 - \theta(i)^2 + \text{var}(\hat{\theta}(i))}{\theta(i)^2}
\]  (3.36)

where, \( \theta(i) \) represents the true values of the \( i \)-th parameter, \( \bar{\theta}(i) \) is the mean of the estimated values and \( \hat{\theta}(i) \) is the set of estimated values. \( N_\theta \) is the number of parameters.

#### 3.3.1 Effect of NSR

To check the robustness of the algorithms in the presence of noise, the noise to signal ratio was varied from 5% to 60% for a data length and frequency of 1500 and 0.05
Parameter estimation from sinusoidal input

(a) Estimated process parameters using direct methodology

Parameter estimation from step response of a sinusoidal input

(b) Estimated process parameters using sinusoid-to-step response methodology

47
Figure 3.1: Estimated process parameters with different methodology for different values of NSR.

rad/s, respectively. The results are shown in Fig. 3.1. This gives us the idea that with the increased NSR the parameter estimation deviates from the true value which is expected and in terms of performance the step response method has a slightly better estimation compared to the other two methods.

### 3.3.2 Effect of data length

Data length, which actually corresponds to sampling time plays an important role in estimation of the parameters. It is obvious that with the increase in data length the estimation results will improve. For demonstration purpose the data length was varied from 500 to 1500 points and the obtained parameters are presented in Fig. 3.2. The simulation results suggests that both the direct and step response techniques have comparable performance which is slightly better than that of the piece-wise linear
Parameter estimation from sinusoidal input

<table>
<thead>
<tr>
<th>Data Length</th>
<th>Time Constant</th>
<th>Gain</th>
<th>Time Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>600</td>
<td>5</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>800</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>1000</td>
<td>15</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>1200</td>
<td>20</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>1400</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Estimated process parameters using direct methodology

Parameter estimation from step response of a sinusoidal input

<table>
<thead>
<tr>
<th>Data Length</th>
<th>Time Constant</th>
<th>Gain</th>
<th>Time Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>600</td>
<td>5</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>800</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>1000</td>
<td>15</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>1200</td>
<td>20</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>1400</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Estimated process parameters using sinusoid-to-step response methodology
Parameter estimation from PLA of a sinusoidal input

<table>
<thead>
<tr>
<th>Data Length</th>
<th>Time Constant</th>
<th>Gain</th>
<th>Time Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.2: Estimated process parameters with different methodology for different data length approximation method.

### 3.3.3 Effect of frequency

The third input variable, frequency, plays a vital role in parameter estimation. The FOPTD system considered in this simulation has cut-off frequency of 0.05 rad/s. For simulation purpose, the parameters were obtained for a range of input frequencies starting from 0.5 times to 1.5 times of system cut-off frequency. Defining the ratio of the input signal frequency to the cut-off frequency as the frequency factor, Fig. 3.3 represents estimated parameters w.r.t different values of frequency factors. The simulation results suggest that while the input sinusoid has a frequency up to process cut-off frequency the parameters obtained with different techniques do not have that much of a significant difference but as the input frequency tends to increase more,
Parameter estimation from sinusoidal input

(a) Estimated process parameters using direct methodology

Parameter estimation from step response of a sinusoidal input

(b) Estimated process parameters using sinusoid-to-step response methodology
Figure 3.3: Estimated process parameters with different methodology for various frequency factors

the piece-wise linear approximation has a better performance. But for very high frequencies the piece-wise approximation procedure results in higher error. Fig. 3.4 represents the estimated average error for different conditions. Based on the average error criterion, it can be said that both direct and step response methods have slightly better estimation for a frequency factor less than 1.2.

3.4 Concluding remarks

There has been some new developments in the field of identification from sinusoidal response. However, the users needs to know the applicability and performance of different methods. A comparative study is carried out in this study to compare the performances of three recently developed methods for identification using sinusoidal response. The purpose of this study was to provide a guideline for the user with
(a) Estimated average error with different methodology for different values of NSR

(b) Estimated average error with different methodology for different data length
Figure 3.4: Comparison of different algorithms based on average error criterion changing one input variable at a time

regards to the choice of identification method when a sinusoid is used as input. The comparable performance of the three methods indicate that the users have a wide range of choices to use the sinusoids. However, the performance varies with data length, and more importantly with the frequency of the input signal. Performance of the algorithms in identification of higher order models and the effects of frequency as well the phase will be studied in the future.
Chapter 4

Fractional Order Model

Identification

This chapter presents methods for identification of process models with fractional orders. The main highlighting factor of this study is the development of identification techniques for fractional order models for both known and unknown fractional order using multiple sinusoid as an input. The overall work done in this chapter can be categorized mainly in three different parts. At the very first part, an identification technique for known fractional order has been developed and verified in simulations. Afterwards, the work has been extended considering unknown fractional order of a system and identification technique based on minimum error criterion is proposed. The efficacy of the proposed method is tested through Monte Carlo Simulation (MCS) and effects of different input parameters have also been studied. Lastly, a more rigorous approach based on Gauss-Newton optimization is developed for simultaneous fractional order and parameter estimates of fractional order models. In all the cases, performance of the developed methods were evaluated and effects of input parameters were studied.
4.1 Introduction

In the fields of dynamical systems and control theory, a fractional-order system is a dynamical system that can be modeled by a fractional differential equation containing derivatives of non-integer order. Identification techniques of such systems have gathered enough attention from the researchers working in the field of system identification. Although there is a vast number of work done to identify the integer order models with high precision, very few literature can be found to identify fractional order models. In fact, in identification methodology, step responses have always been prioritized. However sinusoidal input has different advantages and those has motivated us to carry on our work using sinusoidal signal as an input for identification purpose and effort is made not only in identifying parameters but also to estimate the fractional orders of the considered process model.

Recent studies show that fractional-order models can describe the system better as compared to traditional integer order models [Monje et al., 2010]. An important feature of fractional-order systems is that they exhibit hereditary properties and long memory transients. This aspect is taken into account in modeling, namely with state-space representation, parameter estimation, identification, controller design etc.

4.2 Methodology- Parameter estimation with known fractional order of the model

4.2.1 Parameter estimation using sinusoidal response

To estimate parameters of fractional order models with known model order, two types of generic model: Type-I and Type-II are considered and their corresponding identi-
fication technique is developed in the following sections.

4.2.1.1 Type-I model

A generic Type-I model with time delay is represented by (4.1)

\[ G(s) = \frac{Y(s)}{U(s)} = \frac{b}{s^\alpha + a} e^{-\delta s} \] (4.1)

Where \( s \) indicates the Laplace domain, \( G(s) \) is the transfer function between the input, \( U(s) \), and the output \( Y(s) \); \([a b]\) are the model parameters, \( \delta \) is the time delay and \( \alpha \) is the model order. Considering the process output was initially at steady state, the input-output relation can be expressed in the equation error format by (4.2)

\[ s^\alpha Y(s) + aY(s) = be^{-\delta s}U(s) + E(s) \] (4.2)

Where \( E(s) \) is the error term arising due to measurement noise. Defining a term \( \beta = \text{ceil}(\alpha) + 2 \) i.e. \( \beta \) is the nearest integer of \((\alpha+2)\) in the positive direction. Integrating (4.2) \( \beta \) times which is equivalent of multiplying both sides of the equation by \( s^{-\beta} \) and expressing the resulting equation in the form of an estimating equation in s-domain gives (4.3)

\[ s^{\alpha-\beta}Y(s) = -as^{-\beta}Y(s) + bs^{-\beta}e^{-\delta s}U(s) + s^{-\beta}E(s) \] (4.3)

If the input is a sinusoidal signal of amplitude \( V \), frequency \( \omega \) and phase \( \vartheta \), in time domain it expressed as follows

\[ u(t) = V \sin(\omega t + \vartheta) \] (4.4)
In the Laplace domain equation (4.4) can be written as

\[ U(s) = V \frac{\sin(\vartheta)s + \omega \cos(\vartheta)}{s^2 + \omega^2} \]  

(4.5)

Considering \( \lambda = V \sin(\vartheta) \), \( \mu = V \omega \cos(\vartheta) \) and \( \sigma = \omega^2 \), the above equation (4.5) can be rewritten as following

\[ U(s) = \frac{\lambda s + \mu}{s^2 + \sigma} \]  

(4.6)

For a sinusoidal input, the estimation equation (4.3) becomes

\[ s^{2+\alpha-\beta}Y(s) + \sigma s^{\alpha-\beta}Y'(s) = -as^{2-\beta}Y(s) - a\sigma s^{-\beta}Y'(s) + b\lambda s^{1-\beta}e^{-\delta s} \]
\[ + b\mu s^{-\beta}e^{-\delta s} + (s^{2-\beta} + \sigma s^{-\beta})E(s) \]  

(4.7)

Taking inverse Laplace transform the following time domain equation can be obtained which allows simultaneous estimation of model parameters along with the delay term.

\[ I^{[\beta-2-\alpha]}y(t) + \sigma I^{[\beta-\alpha]}y(t) = -aI^{[\beta-2]}y(t) - a\sigma I^{[\beta]}y(t) + b\lambda[t-\delta] + \frac{b\mu}{2}[t-\delta]^2 + \xi(t) \]  

(4.8)

Where for any signal \( x(t) \), \( I^{[\alpha]}(t) \) is the \( \alpha \)-th order integral, where \( \alpha \) can be real or integer number and \( \xi(t) = L^{-1}\{(s^{2-\beta} + \sigma s^{-\beta})E(s)\} \), \( L \) being the Laplace operator. Equation (4.8) is valid for any bounded input signal \( u(t) \). In the above equation, the time delay term remains as an implicit parameter which cannot be directly estimated. Also to estimate other parameters, the time delay should be known. In the least-squares form equation (4.8) is written as

\[ \gamma(t) = \phi^T(t)\theta + \xi(t) \]  

(4.9)
where, \( \gamma(t) = I^{[\beta-2-\alpha]}y(t) + \sigma I^{[\beta-\alpha]}y(t) \)

\[
\phi(t) = \begin{bmatrix}
-I^{[\beta-2]}y(t) - \sigma I^{[\beta]}y(t) \\
\mu t^2 \\
\mu t \\
\mu
\end{bmatrix}
\]
and \( \theta = \begin{bmatrix}
a \\
b \\
b(\eta - \delta) \\
b(-\eta \delta + \frac{\delta^2}{2\mu})
\end{bmatrix} \)

Here, \( \eta = \lambda / \mu \). Equation (4.9) can be written for \( t = t_{d+1}, t_{d+2}, ... t_N \) and then combined together to give a set of estimation equations

\[
\Gamma(t) = \Phi(t)\theta + \Xi(t) \tag{4.10}
\]

with

\[
\Gamma(t) = \begin{bmatrix}
\gamma(t_{d+1}) \\
\gamma(t_{d+2}) \\
... \\
\gamma(t_N)
\end{bmatrix}
\]

and

\[
\Phi(t) = \begin{bmatrix}
\phi^T(t_{d+1}) \\
\phi^T(t_{d+2}) \\
... \\
\phi^T(t_N)
\end{bmatrix}
\]

Here, \( d \) is the time delay in terms of number of sampling intervals (\( \Delta t \)), i.e. \( d = \delta / \Delta t \) and \( N \) is the total no of samples available. When the time delay is not an integer multiple of sampling interval, \( d \) is chosen as the nearest integer in the positive direction. Finally, the least-squares (LS) solution of the the estimation equations (4.10) gives us the estimated parameters of the system.

\[
\theta^{LS} = (\Phi^T\Phi)^{-1}\Phi^T\Gamma \tag{4.11}
\]
4.2.1.2 Type-II model

A generic Type-II model with time delay is represented by (4.12)

\[ G(s) = \frac{Y(s)}{U(s)} = \frac{b}{s^{2\alpha} + a_1 s^\alpha + a} e^{-\delta s} \] (4.12)

The above input-output relation can be expressed in the equation error format (4.13)

\[ s^{2\alpha}Y(s) + a_1 s^\alpha Y(s) + a Y(s) = be^{-\delta s}U(s) + E(s) \] (4.13)

Defining a term \( \beta = \text{ceil}(2\alpha) + 2 \) i.e. \( \beta \) is the nearest integer of \((2\alpha+2)\) in the positive direction. Considering input to be sinusoidal, integrating (4.13) \( \beta \) times which is equivalent of multiplying both sides of the equation by \( s^{-\beta} \) and expressing the resulting equation in the form of an estimating equation in the s-domain gives (4.14)

\[ s^{2+2\alpha-\beta}Y(s) + \sigma s^{2\alpha-\beta}Y(s) = -a_1 s^{2+\alpha-\beta}Y(s) - as^{2-\beta}Y(s) - a_1 \sigma s^{\alpha-\beta}Y(s) - a \sigma s^{-\beta}Y(s) \]
\[ + b\lambda s^{1-\beta}e^{-\delta s} + b\mu s^{-\beta}e^{-\delta s} + (s^{2-\beta} + \sigma s^{-\beta})E(s) \] (4.14)

Taking inverse Laplace transform the following time domain equation can be obtained which allows simultaneous estimation of model parameters along with the delay term.

\[ I[^{\beta-2-2\alpha}]y(t) + \sigma I[^{\beta-2\alpha}]y(t) = -a_1 I[^{\beta-2-\alpha}]y(t) - a_1 \sigma I[^{\beta-\alpha}]y(t) - a I[^{\beta-\alpha}]y(t) \]
\[ - a \sigma I[^{\beta}]y(t) + \frac{b\lambda}{2} [t - \delta]^2 + \frac{b\mu}{3!} [t - \delta]^3 + \xi(t) \] (4.15)

Where for any signal \( x(t) \), \( I[^{\alpha}]x(t) \) is the \( \alpha \)-th order integral, where \( \alpha \) can be real or integer number and \( \xi(t) = \mathcal{L}^{-1}\{ (s^{2-\beta} + \sigma s^{-\beta})E(s) \} \), \( \mathcal{L} \) being the Laplace operator. Equation (4.15) is valid for any bounded input signal \( u(t) \). In the above equation,
the time delay term remains as an implicit parameter which cannot be directly estimated. Also to estimate other parameters, the time delay should be known. In the least-squares form equation (4.15) is written as

\[ \gamma(t) = \phi^T(t)\theta + \xi(t) \]  

(4.16)

where, \( \gamma(t) = I^{[\beta-2-2\alpha]}y(t) + \sigma I^{[\beta-2\alpha]}y(t) \)

\[ \phi(t) = \begin{bmatrix} -I^{[\beta-2-\alpha]}y(t) - \sigma I^{[\beta-\alpha]}y(t) \\ -I^{[\beta-2]}y(t) - \sigma I^{[\beta]}y(t) \\ \mu^{\frac{3}{3!}} \\ \mu^{\frac{2}{2!}} \\ \mu \\ \mu \end{bmatrix} \quad \text{and} \quad \theta = \begin{bmatrix} a_1 \\ a \\ b \\ b(\eta - \delta) \\ b(-\eta \delta + \frac{\delta^2}{2!}) \\ b(\eta \frac{\delta^2}{2} - \frac{\delta^3}{3!}) \end{bmatrix} \]

Here, \( \eta = \lambda / \mu \). Equation (4.16) can be written for \( t = t_{d+1}, t_{d+2}, \ldots t_N \) and then combined together to give a set of estimation equations

\[ \Gamma(t) = \Phi(t)\theta + \Xi(t) \]  

(4.17)

with

\[ \Gamma(t) = \begin{bmatrix} \gamma(t_{d+1}) \\ \gamma(t_{d+2}) \\ \ldots \\ \gamma(t_N) \end{bmatrix} \quad \text{and} \quad \Phi(t) = \begin{bmatrix} \phi^T(t_{d+1}) \\ \phi^T(t_{d+2}) \\ \ldots \\ \phi^T(t_N) \end{bmatrix} \]

Here, \( d \) is the time delay in terms of number of sampling intervals (\( \Delta t \)), i.e. \( d = \delta / \Delta t \) and \( N \) is the total no of samples available. When the time delay is not an integer multiple of sampling interval, \( d \) is chosen as the nearest integer in the positive
direction. Finally, the least-squares (LS) solution of the estimation equations (4.17) gives us the estimated parameters of the system.

$$\theta^{LS} = (\Phi^T\Phi)^{-1}\Phi^T\Gamma$$ (4.18)

### 4.2.1.3 Extended Type-II model

Type-II model with time delay for the previous section is extended and represented by (4.19)

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_1s^\alpha + b}{s^{2\alpha} + a_1s^\alpha + a} e^{-\delta s}$$ (4.19)

The above input-output relation can be expressed in the equation error format (4.20)

$$s^{2\alpha}Y(s) + a_1s^\alpha Y(s) + aY(s) = [b_1s^\alpha + b]e^{-\delta s}U(s) + E(s)$$ (4.20)

Defining a term $\beta = \text{ceil}(2\alpha) + 2$ i.e. $\beta$ is the nearest integer of $(2\alpha+2)$ in positive direction. Considering input to be sinusoidal, integrating (4.20) $\beta$ times which is equivalent of multiplying both sides of the equation by $s^{-\beta}$ and expressing the resulting equation in the form of an estimating equation in s-domain gives (4.21)

$$s^{2+2\alpha-\beta}Y(s) + \sigma s^{2\alpha-\beta}Y(s) = -a_1s^{2+\alpha-\beta}Y(s) - as^{2-\beta}Y(s) - a_1s^{\alpha-\beta}Y(s)$$

$$-a\sigma s^{-\beta}Y(s) + b_1\lambda s^{1+\alpha-\beta}e^{-\delta s} + b_1\mu s^{\alpha-\beta}e^{-\delta s}$$

$$+ b\lambda s^{1-\beta}e^{-\delta s} + b\mu s^{-\beta}e^{-\delta s} + (s^{2-\beta} + \sigma s^{-\beta})E(s)$$ (4.21)
Taking inverse Laplace transform the following time domain equation can be obtained which allows simultaneous estimation of model parameters along with the delay term.

\[
I^{[\beta-2-2\alpha]}y(t) + \sigma I^{[\beta-2\alpha]}y(t) = -a_1 I^{[\beta-2-\alpha]}y(t) - a_1 \sigma I^{[\beta-\alpha]}y(t) - a_1 \sigma I^{[\beta]}y(t) + b_1 \lambda \left[ t - \delta \right]^{[\beta-\alpha-2]} + \frac{b_1 \mu}{(\beta-\alpha-1)!} \left[ t - \delta \right]^{[\beta-\alpha-1]} \\
+ \frac{b_2 \lambda}{2!} [t - \delta]^2 + \frac{b_3 \mu}{3!} [t - \delta]^3 + \xi(t)
\]

(4.22)

Where for any signal \( x(t) \), \( I^{[\alpha]}(t) \) is the \( \alpha \)-th order integral, where \( \alpha \) can be real or integer number and \( \xi(t) = L^{-1} \{(s^{\beta-\alpha} + \sigma s^{\alpha-\beta})E(s)\} \), \( L \) being the Laplace operator. Equation (4.22) is valid for any bounded input signal \( u(t) \). In the above equation, the time delay term remains as an implicit parameter which can not be directly estimated. Also to estimate other parameters, the time delay should be known. In least-square form equation (4.22) is written as

\[
\gamma(t) = \phi^T(t)\theta + \xi(t)
\]

(4.23)

where, \( \gamma(t) = I^{[\beta-2-2\alpha]}y(t) + \sigma I^{[\beta-2\alpha]}y(t) \)

\[
\phi(t) = \begin{bmatrix}
-I^{[\beta-2-\alpha]}y(t) - \sigma I^{[\beta-\alpha]}y(t) \\
-I^{[\beta-2]}y(t) - \sigma I^{[\beta]}y(t) \\
\frac{\lambda}{(\beta-\alpha-2)!} [t - \delta]^{[\beta-\alpha-2]} + \frac{\mu}{(\beta-\alpha-1)!} [t - \delta]^{[\beta-\alpha-1]} \\
\mu t \\
\mu
\end{bmatrix}
\]

and \( \theta = \begin{bmatrix} a_1 \\ a \\ b_1 \\ b \\ b(\eta - \delta) \\ b(-\eta \delta + \frac{s^2}{2!}) \\ b(\eta^2 - \frac{s^3}{3!}) \end{bmatrix} \)
Here, \( \eta = \lambda / \mu \). Equation (4.23) can be written for \( t = t_{d+1}, t_{d+2} ... t_N \) and then combined together to give a set of estimation equations

\[
\Gamma(t) = \Phi(t)\theta + \Xi(t) \quad (4.24)
\]

with

\[
\Gamma(t) = \begin{bmatrix}
\gamma(t_{d+1}) \\
\gamma(t_{d+2}) \\
... \\
\gamma(t_N)
\end{bmatrix}
\quad \text{and} \quad \Phi(t) = \begin{bmatrix}
\phi^T(t_{d+1}) \\
\phi^T(t_{d+2}) \\
... \\
\phi^T(t_N)
\end{bmatrix}
\]

Here, \( d \) is the time delay in terms of number of sampling intervals (\( \Delta t \)), i.e. \( d = \delta / \Delta t \) and \( N \) is the total no of samples available. When the time delay is not an integer multiple of sampling interval, \( d \) is chosen as the nearest integer in the positive direction. Finally, the least-squares (LS) solution of the the estimation equations (4.24) gives us the estimated parameters of the system.

\[
\theta^{LS} = (\Phi^T\Phi)^{-1}\Phi^T\Gamma \quad (4.25)
\]

### 4.2.2 Order estimation

While following the above methodology to estimate parameters of the linear time invariant systems, it is clearly evident that we must have a prior knowledge of the fractional order (\( \alpha \)) of the system, which is quite impractical. For this reason, we need to estimate the order along with the parameters of the process model. This seems rather a paradox as without identifying the parameters of the system, the fractional order can not be estimated properly. So, basically both the parameters and the order estimation depends on each other. To solve this problem, we have applied a simple
yet effective technique which will help us estimate the parameters along with the order of the system. And to implement the technique, we need to start estimating the parameters with an initial guess of the system order. Then, with that initially selected system order and estimated parameters for that initial selection, the error is calculated using following equation (4.26)

\[
Error = \Sigma (y - \hat{y})^2
\]

(4.26)

here, \(y\) represents noise-free sinusoidal response of the actual system and \(\hat{y}\) represents sinusoidal response of a system using estimated parameters for the selected fractional order.

Afterwards, this loop will continue for a particular range of system order with very small interval value. This will led to a set of matrix consisting of parameters and error corresponding to each value of fractional order within the range. From the generated matrix, based on minimum error criterion, the parameters and the corresponding fractional order is selected.

The following algorithm outlines the procedure for estimating order of fractional process models for this study.

Algorithm:

Step I: Specify a range of fractional order with very small interval for which parameters and errors to be calculated.

Step II: Estimate the parameters using the methodology stated in 4.2.1 and errors using Equation (4.26), corresponding to each value of fractional orders for the entire range of orders selected in Step I.
Step III: Select a particular set consisting of estimated parameters and corresponding order based on minimum error criterion.

### 4.2.3 Simulation and Results

For this study, the fractional order models: Type-I and Type-II were considered to be as 
\[
\frac{1.25}{100s^{1.75}+1}e^{-7s} \quad \text{and} \quad \frac{2}{100s^{1.5}+10s^{0.75}+1}e^{-7s}.
\]
Simulations were carried out with a fixed value of amplitude and phase of the sinusoid which were in this case 100 and 0.2 rad respectively. In each case, the parameters were obtained for 50 Monte-Carlo simulations (MCS), and for each MCS there was an inner loop to identify the system order based on minimum error criteria. In order to interpret the simulated parameters and the system order, they were graphically plotted and categorized into two different vertical axes. Furthermore, for Type-II model, the two different vertical axes were scaled differently, i.e., one was scaled at a logarithmic base of 10 and another one was on 2. Based on the obtained parameters, the methods were compared and analyzed for a range of NSR, frequency factor, and data length.

### 4.2.4 Simulation Environment

For the general fractional order differentials and integrals of a function \( g(t)\), the Grunwald-Letnikov (GL) definition (4.27) is commonly used see e.g. [Oldham and Spanier, 1974].

\[
t_0D^\rho tg(t) = \lim_{\eta \to 0} \frac{1}{\eta^\rho} \sum_{j=0}^{[\frac{t-t_0}{\eta}]} (-1)^j \left( \begin{array}{c} \rho \\ j \end{array} \right) g(t - j\eta)
\]  

(4.27)

Here, \( t_0 \) and \( t \) are the limits of the operator, \( \eta \) is the step size and \( \rho \) is the order with \( \rho > 0 \) means a derivative operation and \( \rho < 0 \) means integral operation. Also \( \lfloor . \rfloor \) means the integer part and
with $\Gamma(.)$ being the Euler’s Gamma function.

For numerical computation, a revised version of (4.27), presented in [Chen et al., 2009] is used where

$$
t_0 D_t^\rho g(t) = \lim_{\eta \to 0} \frac{1}{\eta^\rho} \sum_{j=0}^{\lfloor t-t_0 \rfloor} w_j(\rho) g(t-j\eta)
$$

(4.29)

where $w_j(\rho)$ can be evaluated recursively from

$$
w_0(\rho) = 1
$$

(4.30)

$$
w_j(\rho) = \left(1 - \frac{\rho + 1}{j}\right) w_{j-1}(\rho) \quad j = 1, 2, \ldots
$$

(4.31)

4.2.4.1 Effect of NSR

The noise to signal ratio was varied from 5% to 25% for a data length and frequency of 10,000 and 0.628 rad/s respectively. The results are shown in Fig. 4.1. The results clearly indicate that with the increased NSR the parameter estimation deviates from the true value which is expected.

4.2.4.2 Effect of frequency

To analyze the effect of frequency we have plotted the estimated parameters over a range of frequency factors. Frequency factor is defined as the ratio of the input signal frequency to the cut-off frequency of the system. The cut-off frequency of the Type-I and Type-II models that are being considered for this study are 0.028 rad/s and 0.0314 rad/s respectively. While performing the simulation to estimate the parameters for
Figure 4.1: Estimated process parameters for different values of NSR
both of the models, we initially considered to vary the frequency factor between 0.5 times to 2.5 times of the cut-off frequency for the individual systems. The simulations revealed that although for the Type-I model our developed methodology worked as expected but for Type-II model, the least squares solution was unable to generate parameters of the system for frequency factors less than 2. This happened as because the phi ($\phi$) matrix were being singular and that lead to very poor set of estimations. This simulation has led us to believe that the Type-II pole model has a distinctive characteristic compared to the Type-I model over a range of frequency factors. Fig. 4.2 represents the estimated parameters for both of the systems. This simulation also interprets that whereas for the Type-I model, the estimated parameters are better near the system cut-off frequency which is in this case for the frequency factor of 1.25 or 1.5; but for Type-II model, higher frequency factor led to better estimation of the parameters. For this study, we carried out the simulation for the Type-II model till frequency factor of 10 and other input variables such as the data length and NSR were set to 10,000 and 10% respectively for both of the systems.

4.2.4.3 Effect of number of samples

For this study, we have varied the number of samples from 5,000 to 15,000 for both of the systems and performed the simulation to estimate the parameters. Fig. 4.3 represents the estimated parameters from this simulation where for Type-I model, the frequency and the NSR of the input variable were 0.028 rad/s and 15% respectively and for the Type-II model the simulation was run for the frequency of 0.314 rad/s and NSR of 20%. Also, Fig. 4.3 clearly explains that with more samples the estimation is more accurate.
(a) Estimated process parameters for Type-I model for a range of frequency factors

(b) Estimated process parameters for Type-II model for a range of frequency factors

Figure 4.2: Estimated process parameters for different values of frequency factors
Figure 4.3: Estimated process parameters over a range of data length
4.2.4.4 Comparison based on average error criterion

Finally, based on the obtained parameters, both of the models- Type-I and Type-II were compared based on the average error criterion defined as

\[
E_{avg} = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} \left( \bar{\theta}(i) - \theta(i) \right)^2 + \text{var}(\hat{\theta}(i))
\]  \hspace{1cm} (4.32)

where, \( \theta(i) \) represents the true values of the \( i \)-th parameter, \( \bar{\theta}(i) \) is the mean of the estimated values and \( \hat{\theta}(i) \) is the set of estimated values. \( N_\theta \) is the number of parameters.

Fig. 4.4 represents the comparison between fractional order models w.r.t average error over a range of values for different conditions, i.e NSR, data length and frequency factor.

4.3 Methodology- Simultaneous parameter and fractional order estimation

This section of our study deals with simultaneous estimation of the fractional order of the model together with its parameters. Although, in the previous section of this chapter we have discussed how to estimate the fractional order of a model alongside with its parameters, the study was more dependent on prior process knowledge due to the fact that the accuracy of the simulation results were dependent on the range of factional order that was selected for the simulation to run. Another key factor that played an important role for better estimation was the interval between two consecutive fractional orders for a particularly selected range. All these factors played a crucial role for the better accuracy on estimating the parameters as well as the time required for running the entire simulation because the entire process was tedious.
Figure 4.4: Comparison of Type-I and Type-II model
and time consuming. Therefore, we have approached towards a more developed and automated method which have allowed us to reduce the dependency on prior process knowledge ans has also saved us ample amount of time in terms of running MCS to generate the results. This advanced technique is based on optimization and convergence theory. In other words, it is fair to say that the following approach closely relates to optimization solving for fractional ordered models which in return converges the fractional order and the parameters of a fractional model towards its true values.

A vast number of literature can be found related to optimization and convergence. However, only a few have dealt with continuous time identification especially with fractional order models. Among the very few, the work done in Narang et al. and Victor et al. article [Narang et al., 2011, Victor et al., 2013] has a close resemblance with our method and has basically served as a backbone to our proposed frame of work. In addition, the other core part of our optimization procedure is immensely inspired and adopted from Hines’s article, where a definition is made for logarithm of an operator along with the study of the logarithm of the derivative [Hines, 1955].

In order to begin with an optimization, this approach also requires some initial guess of parameters like most of the other optimization techniques which will later lead to optimal values of the parameters as well as the fractional order of the model. To reduce the complexity and dependency on prior process knowledge for making initial guess for a particular model, we have first considered a simple first order model without any time delay for Type-I model and a simple second order model without any time delay for Type-II model. Afterwards, using the least-squares technique, the parameters i.e the gain and the time constant/s have been extracted which were later used as the initial guess values for the purpose of future optimization and convergence.
4.4 Mathematical formulations

The optimization based identification method is outlined in this section. The method follows the output error approach. A technique to evaluate the logarithmic derivative required to evaluate the error gradient is also outlined.

4.4.1 Identification method

For a single input single output system, the relation between the input and the output can be expressed using the following Laplace domain equation.

\[ Y(s) = G(s, \nu)e^{-\delta s}U(s) + W(s) \]  \hspace{1cm} (4.33)

where, \( Y(s) \) and \( U(s) \) are the input and output, respectively, \( G(s, \nu) \) is the model transfer function, \( s \) being the Laplace variable with \( \nu \) as the set of coefficients and degrees of derivatives of different terms in the numerator and the denominator polynomials and \( \delta \) is the time delay. \( W(s) \) represents the noise in the output measurements. The unknown parameter vector is denoted as \( \theta = [\nu \delta]. \)

A time domain expression for the input output relation is written as

\[ y(t) = G(p, \nu)u(t - \delta) + w(t) \]  \hspace{1cm} (4.34)

The lower case letters correspond to variables in the time domain; \( p \) represents the derivative operator. The objective of an identification algorithm is to estimate the parameter vector \( \theta \) from a set of time domain measurements \( [u(t_k) y(t_k)], k = 1, 2, \cdots N \)
and $N$ is the number of data points available. The goal of the output error (OE)
approach is to estimate $\theta$, by minimizing a norm of the errors between measured and
model output.

\[ e(t, \theta) = y(t) - G(p, \nu)u(t - \delta) \]  \hspace{1cm} (4.35)

Using the notation $e_k = e(t_k, \theta)$, the following objective function can be defined for
the OE algorithm.

\[ f(\theta) = \sum_{k=1}^{N} \frac{1}{2} e_k^2 = \frac{1}{2} \| e \|^2 \] \hspace{1cm} (4.36)

A number of different approaches can be taken for solution of the optimization prob-
lem. We follow the Gauss-Newton approach to simultaneously estimate all the pa-
rameters. In this algorithm, estimate of the parameters at an iteration step $i$ is given
by

\[ \theta^i = \theta^{i-1} - \left[ H(\theta^{i-1}) \right]^{-1} \nabla f(\theta^{i-1}) \] \hspace{1cm} (4.37)

where $\nabla f(\theta)$ is the error gradient given by

\[ \nabla f(\theta) = J^T e \] \hspace{1cm} (4.38)

with $J$ being the Jacobian.

\[ J = \begin{pmatrix} \frac{\partial e_1}{\partial \theta_1} & \cdots & \frac{\partial e_1}{\partial \theta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial e_N}{\partial \theta_1} & \cdots & \frac{\partial e_N}{\partial \theta_n} \end{pmatrix} \] \hspace{1cm} (4.39)

In the Gauss-Newton approach for optimization, the Hessian, $H$ is approximated by

\[ H \approx J^T J \] \hspace{1cm} (4.40)
The advantages and disadvantages of Newton’s method and the Gauss-Newton method have been widely addressed in the literature [Chong and Zak, 2013, Fletcher, 2013]. The advantage of the Gauss-Newton method is that the second derivative matrix of the error function is not required. The optimization step follows a standard procedure. Using an initial guess of the parameters, the Jacobian and Hessian are evaluated and the parameters are iteratively updated until convergence.

To evaluate the Jacobian, we need $\frac{\partial e_i}{\partial \theta}$. A model of the following form is considered for illustration.

$$G(s) = \frac{be^{-\delta s}}{s^\alpha + a}$$  \hspace{1cm} (4.41)

Where, the parameter vector is $\theta^T = [a \ b \ \delta \ \alpha]$. Each row of the Jacobian matrix represents a sampling instant; the $k$-th row $J_k$ can be expressed in the Laplace domain as

$$J_k = \begin{bmatrix} \frac{be^{-\delta s}}{(s^\alpha + a)^2} U(s) & -\frac{1}{s^\alpha + a} U(s) & \frac{se^{-\delta s}}{s^\alpha + a} U(s) & \frac{b e^{-\delta s}}{(s^\alpha + a)^2} \ln(s) U(s) \end{bmatrix}$$  \hspace{1cm} (4.42)

An equivalent time domain expression is given by

$$J_k = \begin{bmatrix} b \frac{1}{(p^\alpha + a)^2} u(t_k - \delta) & \frac{-1}{p^\alpha + a} u(t_k - \delta) & \frac{p}{p^\alpha + a} u(t_k - \delta) & \frac{bp^\alpha}{(p^\alpha + a)^2} \ln(p) u(t_k - \delta) \end{bmatrix}$$  \hspace{1cm} (4.43)

For a model without time delay $\frac{\partial e_i}{\partial \theta}$ is expressed as

$$J_k = \begin{bmatrix} \frac{b}{(p^\alpha + a)^2} u(t_k) & \frac{-1}{p^\alpha + a} u(t_k) & \frac{bp^\alpha}{(p^\alpha + a)^2} \ln(p) u(t_k) \end{bmatrix}$$  \hspace{1cm} (4.44)

For a model, $G(s) = \frac{b}{s^{\alpha_2} + a_1 s^{\alpha_1} + a_0}$ with the parameter vector $\theta^T = [a_1 \ a_0 \ b \ \alpha_2 \ \alpha_1]$, the corresponding expression becomes
\[ J_k = \left[ \frac{b p^{\alpha_1}}{(p^{\alpha_2} + a_1 p^{\alpha_1} + a_0)^2} u(t_k) - \frac{b}{(p^{\alpha_2} + a_1 p^{\alpha_1} + a_0)^2} u(t_k) - \frac{1}{p^{\alpha_2} + a_1 p^{\alpha_1} + a_0} u(t_k) \right] \] (4.45)

Here \( \ln(s) \) and \( \ln(p) \) are the logarithm of the derivative operator, expressed in the Laplace and time domain, respectively. Evaluation of the Jacobian needs estimation of \( \ln(p) u(t_k) \). [Victor et al., 2013] suggested numerical estimation of the Jacobian as logarithm of the derivative operator is not trivial to simulate. We propose a method to evaluate the logarithmic derivative of the input signal.

### 4.4.2 Evaluation of the logarithmic derivative

The input is assumed to be deterministic. For many cases, for example, for sinusoids, the analytical expression of the input is known. A sinusoidal input can be expressed as

\[ u(t) = \sin(\omega t) \] (4.46)

[Hines, 1955] derived the logarithmic derivative \( L_{\theta} = \ln \partial_t \) of sinusoid as

\[ L_{\theta} \sin t = \frac{n \pi}{2} \cos t \quad n = 1, 2, 3, \cdots \] (4.47)

Following [Hines, 1955] we derive the logarithmic derivative of a single frequency sinusoid as

\[ L_{\theta} \sin(\omega t) = \frac{n \pi}{2} \cos(\omega t) + \ln(\omega) \sin(\omega t) \] (4.48)
For a sinusoid with a phase angle $\mu$, the corresponding expression becomes

\[
L_\theta \sin(\omega t + \mu) = \cos \mu \left[ \frac{n\pi}{2} \cos(\omega t) + \ln(\omega) \sin(\omega t) \right]
+ \sin \mu \left[ -\frac{n\pi}{2} \sin(\omega t) + \ln(\omega) \cos(\omega t) \right]
\]

(4.49)

A multi-frequency sinusoid can be expressed as

\[
u(t) = \sum_{i=1}^{m} \sin(\omega_i t)
\]

(4.50)

where, $m$ is the number of frequencies in the signal. Considering that the operator and the summation commute, we find the $L_\theta$ of a multi-frequency sinusoid.

\[
L_\theta \sum_{i=1}^{m} \sin(\omega_i t) = \frac{n\pi}{2} \sum_{i=1}^{m} \cos(\omega_i t) + \sum_{i=1}^{m} \ln(\omega_i) \sin(\omega_i t)
\]

(4.51)

### 4.5 Implementation issues

#### 4.5.1 Initialization

A major issue with an optimization algorithm is initialization of parameters. In the proposed methodology, initialization regarding model orders, coefficients as well as that of the time delay is required. We propose to initiate the optimization algorithm by estimating an integer order model. For models of the form (4.41), a first (integer) order process is assumed to estimate the coefficients. The estimated coefficients along with the integer order is used as the parameter set to initialize the estimation procedure. If a fractional order model with a time delay is to be estimated, a small time delay is assumed to initiate the algorithm.
For estimation of the initial model coefficients assuming an integer order, the integral equation approach is used. In this procedure a differential equation representing the input-output relation of the form (4.52) is considered.

\[
y(t) = \frac{\beta}{p + \mu} u(t) + \epsilon_1(t) \tag{4.52}
\]

The relation can be presented in the equation error form as

\[
\frac{dy(t)}{dt} + \mu y(t) = \beta u(t) + \epsilon_2(t) \tag{4.53}
\]

The equation is then integrated to get

\[
y(t) + \mu y[1](t) = \beta u[1](t) + \epsilon(t) \tag{4.54}
\]

where, for any variable, \(y(t)\),

\[
y[1](t) = \int_0^t y(t)dt \tag{4.55}
\]

The estimation equation (4.55) can be reformulated to get in a least-squares form

\[
y(t) = \begin{bmatrix} -y[1](t) & u[1](t) \end{bmatrix} \begin{bmatrix} \mu \\ \beta \end{bmatrix} + \epsilon(t) \tag{4.56}
\]

Or equivalently

\[
\psi(t) = \phi^T(t) \vartheta + \epsilon(t) \tag{4.57}
\]

where,

\[
\psi(t) = y(t), \quad \phi^T(t) = \begin{bmatrix} -y[1](t) & u[1](t) \end{bmatrix}, \quad \vartheta = \begin{bmatrix} \mu \\ \beta \end{bmatrix}
\]

Equation (4.57) can be written for \(t = t_1, t_2 \cdots t_N\) and combined to give the estimation
equation

\[ \Psi = \Phi \vartheta + \epsilon \]  \hspace{1cm} (4.58)

with

\[ \Psi(t) = \begin{bmatrix} \psi(t_1) \\ \psi(t_2) \\ \vdots \\ \psi(t_N) \end{bmatrix}, \quad \Phi(t) = \begin{bmatrix} \phi^T(t_1) \\ \phi^T(t_2) \\ \vdots \\ \phi^T(t_N) \end{bmatrix} \]  \hspace{1cm} (4.59)

The parameter vector \( \vartheta \) is then obtained as the solution of the least-square equation as

\[ \vartheta = (\Phi^T \Phi)^{-1} \Phi^T \Psi \]  \hspace{1cm} (4.60)

The main focus of this section is to update all the parameters and the fractional order with each iteration, whereas in previous literature it was more of a two step procedure: first step was to update the value of the fractional order based on minimization of a particular objective function or gradient based algorithm leading to complex computation of sensitivity function; and secondly update the parameters of a fractional order model with respect to the updated value of the fractional order. This study merges both the steps into one and with the help of operator mathematics, computing sensitivity function was made easier as well.

4.5.2 Simulation and results

4.5.2.1 Identification of processes and parameter convergence for noise free scenario

Different process models with unknown fractional order has been considered to verify simulation results for noise free scenario. Table 4.1 represents estimated models compared to true models and total number of iterations needed to reach the parameter
convergence. The simulation was carried out using multiple sinusoidal input with \( \omega = [0.1 \ 0.2] \) rad/s and 1000 data points with an interval of 0.15 seconds for all the models. Figure 4.5 represents convergence of all the unknown parameters for process model \( \frac{3}{10s^{4}+1}e^{-3s} \).

Table 4.1: Identification results of processes considering zero NSR

<table>
<thead>
<tr>
<th>True models</th>
<th>Estimated models</th>
<th>Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{3}{10s^{4}+1}e^{-3s} )</td>
<td>( 3.0007 ) ( 1.2090 ) ( e^{-2.9634s} )</td>
<td>12</td>
</tr>
<tr>
<td>( \frac{1}{20s^{4}+1}e^{-4s} )</td>
<td>( 19.9864 ) ( 1.9999 ) ( e^{-3.9376s} )</td>
<td>9</td>
</tr>
<tr>
<td>( \frac{1}{20s^{4}+1}e^{-7s} )</td>
<td>( 20.0073 ) ( 1.9999 ) ( e^{-6.9622s} )</td>
<td>13</td>
</tr>
<tr>
<td>( \frac{5}{10s^{4}+1}e^{-1s} )</td>
<td>( 10.0180 ) ( 1.9999 ) ( e^{-6.9261s} )</td>
<td>18</td>
</tr>
<tr>
<td>( \frac{5}{16s^{4}+1}e^{-5s} )</td>
<td>( 16.0119 ) ( 1.9999 ) ( e^{-4.9916} )</td>
<td>13</td>
</tr>
</tbody>
</table>

4.5.2.2 Identification of time delay models considering NSR

A list of Type-I models with time delay is considered in this section. Table 4.2 shows the results. The parameters presented are the means of 100 MCS with the corresponding standard deviation in the parentheses. NSR for each cases is 10\% and 1600 data points with an interval of 0.05 second is used. The convergence rate is 100\% for each of the cases. Two frequency sinusoids of 0.1 and 0.2 rad/s are used in all cases.

Table 4.2: Identification results of processes considering NSR=10\%

<table>
<thead>
<tr>
<th>True models</th>
<th>Estimated models</th>
<th>Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1.25}{20s^{4}+1}e^{-4s} )</td>
<td>( 20.3536(\pm0.2189) ) ( 1.6166(\pm0.0618) ) ( 1.9031(\pm0.0539) ) ( e^{-3.919(\pm0.3412)s} )</td>
<td>22.82(\pm3.2547)</td>
</tr>
<tr>
<td>( \frac{1.25}{10s^{4}+1}e^{-1s} )</td>
<td>( 9.9327(\pm0.7078) ) ( 1.2023(\pm0.0201) ) ( e^{-0.9904(\pm0.2056)s} )</td>
<td>13.47(\pm1.7721)</td>
</tr>
<tr>
<td>( \frac{1}{20s^{4}+1}e^{-7s} )</td>
<td>( 20.0597(\pm1.3608) ) ( 1.9999(\pm0.0329) ) ( 1.0047(\pm0.0703) ) ( e^{-6.9716(\pm0.2874)s} )</td>
<td>15.63(\pm2.5767)</td>
</tr>
</tbody>
</table>
(a) Convergence of process parameter: Gain
(b) Convergence of process parameter: Time-constant
(c) Convergence of process parameter: Time delay
(d) Convergence of process parameter: Fractional Order

Figure 4.5: Convergence of process parameters for $\frac{3}{10s^{1.4}+1}e^{-3s}$
4.5.2.3 Effect of NSR

Figure 4.6 clearly helps us understand the effect of NSR on parameter estimates of fractional order models. It is seen that the effect of NSR is consistent with theories i.e parameter estimates are better with lesser value of NSR. For simulation purpose a two frequency sinusoid of 0.1 and 0.2 rad/s were used in all the cases for noise to signal ratio of 2, 5, 10, 15, 20 and 50. For each of the cases and for every NSR value, 100 MCS were run to generate the results. Besides for process models $\frac{1.25}{10s^{0.75}+1}e^{-7s}$ and $\frac{3}{10s^{1.4}+1}e^{-3s}$ total number of data points were considered as 1000 with sampling interval of 0.15 second, 1500 with sampling interval of 0.1 second and 400 with sampling interval of 0.2 second consecutively. Table 4.3 is the tabular representation of Figure 4.6b.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimated values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSR=2</td>
</tr>
<tr>
<td>$\alpha \equiv 0.75$</td>
<td>0.7491 (0.0209)</td>
</tr>
<tr>
<td>$K \equiv 1.25$</td>
<td>1.2502 (0.0378)</td>
</tr>
<tr>
<td>$\tau \equiv 10$</td>
<td>9.9728 (0.2309)</td>
</tr>
<tr>
<td>$\delta \equiv 7$</td>
<td>6.9511 (0.1278)</td>
</tr>
</tbody>
</table>

4.5.2.4 Effect of data length and sampling interval

To study the effect of data length and sampling interval, data are generated with the sinusoidal input with $\omega = [0.1 0.2]$ for process model $\frac{1.25}{20s^{1.4}+1}e^{-4s}$. Table 4.4 shows the identification results. It is seen that the data length and sampling interval affect the
(a) Effect of NSR on parameter estimates for process model: 
\[ \frac{1.25}{20s+1}e^{-4s} \]

(b) Effect of NSR on parameter estimates for process model: 
\[ \frac{1.25}{10s+1}e^{-7s} \]
Figure 4.6: Effect of NSR on parameter estimates for different models

results which is consistent with theories. All the parameter estimates are better with higher number of data points and lower sampling interval. Figure 4.7 is the graphical representation of Table 4.4.

Table 4.4: Effect of data length and sampling interval on parameter estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimated values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=400</td>
</tr>
<tr>
<td></td>
<td>$\Delta t = 0.375$</td>
</tr>
<tr>
<td>$\alpha \equiv 1.4$</td>
<td>1.3992 (0.0504)</td>
</tr>
<tr>
<td>$K \equiv 1.25$</td>
<td>1.2613 (0.0726)</td>
</tr>
<tr>
<td>$\tau \equiv 20$</td>
<td>20.3349 (2.3969)</td>
</tr>
<tr>
<td>$\delta \equiv 4$</td>
<td>3.8303 (0.4331)</td>
</tr>
</tbody>
</table>
Figure 4.7: Effect of data length and sampling interval on parameter estimates

4.6 Concluding remarks

In the genre of system identification, significance of fractional modeling is undeniable and consequently, in recent era, fractional order modeling has been the center of attention as it serves as the core whether it comes to implement fractional order controller or to design a physical system itself. This study primarily focuses on identifying the parameter estimates of a fractional order model based on integral equation approach followed by instrumental variable method. The purpose of this study was to develop an improved identification algorithm for fractional order model with a guideline on the variables of the input sinusoid for better parameter estimation. During this study, it was obvious that while the fractional order of a system was known, implementing the identification technique was trouble-free but got complicated for unknown model order. To deal with the issue, parameter estimates were generated for a pre-selected range of fractional orders and based on minimum error criteria w.r.t actual output
of the model, a set of parameters and corresponding model order was selected. This seemed to be a tedious solution in terms of computational time. As a better approach to the problem, we incorporated Gauss-Newton optimization which lead to solve logarithmic derivative of input signal analytically as contrary to the numerical solutions found in the recent literature. This iterative approach significantly improved computational time and when tested, convergence of parameters including fractional order of the model were found 100% for noise free scenario and highly accurate when noise was introduced in the system. Although in this study, we have used two types of generic models: Type-I and Type-II, the methodology can be extended for higher order models as well. Performance of the identification algorithm for non-commensurate fractional order models and the role of input sinusoid’s phase will be the future scope of this study.
Chapter 5

Conclusion

In this study we highlighted one of the important segments of control engineering—‘system identification’ along with its applicability in the field of control systems. The main outcomes of this study are as follows:

1. Comparative study- A comparison on three different identification techniques, (i) direct approach, (ii) converting sinusoid to step response and (iii) linear approximation is carried out on integer order continuous time models. All three methods proved to be robust for different variations of the input parameters such as signal to noise ratio, data length and input frequency. Later on changing only one variable at a time the efficacy of the mentioned algorithms are checked based on an average error criterion. This study provides user with general guidelines while selecting an identification method considering sinusoidal input together with the choice of input parameters outlining the more accurate set of model parameter estimates.

2. Fractional order model identification- the direct identification from sinusoidal response method was adopted and expanded for fractional order continuous time models
with different structures. Simulation results were generated to validate the identification algorithm which proved to be robust for a wide range of input variables such as signal to noise ration, number of samples and frequency factor. Besides, the simulation results also followed the general trend that the estimated set of model parameters is suppose to follow with the variation in input changes leading towards validation of the applied technique.

3. Simultaneous parameter and order estimation- While developing the identification algorithm for fractional order models, a parallel problem to solve fractional calculus especially fractional integral in our case was dealt. At first without making the problem complex, a simple fractional order model with known fractional order was considered for validation purpose of the algorithm. Afterwards, the fractional order of the model was considered to be unknown and identification algorithm was run several times for a predefined range of fractional orders of the model which required some prior knowledge. These multiple simulations lead towards the best possible set of parameter estimates based on the minimum error criterion. But as can be seen this was a semi automated process requiring a vast amount of computation and time, finally an optimization technique was introduced to make the overall identification techniques automatic. It is also worth mentioning that the developed identification method can only deal with fractional order models of commensurate model orders.

4. Optimization- The optimization technique that has been utilized in this study is based on Gauss-Newton method. The reason behind selecting this method is because this method has proven to perform better compared to other conventional methods regarding convergence of the parameter estimates for a wider range of initialized guess values of time delay. Using this optimization technique to update the fractional order
and the parameters of the fractional model with every iteration lead to an unusual dilemma i.e. calculation of logarithmic of a differential operator. At present, methods regarding optimization involving logarithm of a derivative operator is handled numerically such as by introducing central difference method for example. In our study this difficulty is resolved applying operator mathematics which uses an analytical expression rather than numerical estimation ultimately leading towards convergence of parameter estimates together with fractional order of the model.

5.1 Future recommendations

This study deals with some of the recently developed identification methods taking sinusoid as an input and making a comparison based on input variables. However, following are some of the issues that can be considered in future work:

1. Consideration of higher order integer models- For the simplicity this study was based on first order plus time delay models without any phase difference. The work can be generalized for higher order time delay models considering phase difference as well as for multiple sinusoidal inputs.

2. Simulation time limitation for fractional order model identification- Simulation time has been an issue for our study, as because it was simulated for a range of fractional orders first and then again replicated over several Monte-Carlo simulations which resulted in a very long processing time. In future, there might be some scope to improve on the processing time and thus saving time before estimation.

3. Consideration of different types of periodic input- In future, the work can be ex-
tended considering other periodic signals of similar characteristics of sinusoid.

4. Optimization efficiency- The operator mathematics that we have applied in this study to deal with logarithm of a derivative operator is based on sinusoidal inputs only which can be extended for different types of periodic and or aperiodic signals. Besides, the optimization algorithm also requires initial values to start the procedure. In future, work can be done to improve the initial guess or to remove any dependency on initial guess so that irrespective of the initial guess value of the parameters global optimization can be achieved.

5. Input optimization- So far our work indicates robustness of the used and developed identification procedures together with providing some insight knowledge of the input parameters and their effects on identification. In future, a significant amount of work can be done on selecting the best set of input parameters for a particular type of system and identifications technique for more accurate parameter estimation.
Bibliography


94


