Identification of Dynamic System in the Presence of Noise with Wavelets as Basis Functions

by
Siddhartha Mukhopadhyay

A thesis submitted to the
Board of Studies in Engineering Sciences
In partial fulfilment of requirements
For the Degree of
DOCTOR OF PHILOSOPHY
of
HOMI BHABHA NATIONAL INSTITUTE

2012

Homi Bhabha National Institute

Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by Siddhartha Mukhopadhyay entitled " Identification of Dynamic System in the Presence of Noise with Wavelets as Basis Functions" and recommend that it may be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Lalit Mohan Gantayet — **Date:** 21.9.2012

Chairman - Dr. L. M. Gantayet

BK Dutta — **Date:** 21.9.12

Member - Dr. B. K. Dutta

P Satyamurthy — **Date:** 21-9-12

Member - Dr. P Satyamurthy

APTiwari — **Date:** 21.9.12

Guide/ Member - Dr. A. P. Tiwari

U Mahapatra — **Date:** 21.09.12

Technology Advisor - Mr. U. Mahapatra

B Bandyopadhyay — **Date:** 21.09.12

Examiner - Prof. B. Bandyopadhyay

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it may be accepted as fulfilling the dissertation requirement.

Date: 21st September, 2012

Place: Mumbai

APTiwari

Dr. A. P. Tiwari

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowed without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author.

S Mukhopadhyay

Siddhartha Mukhopadhyay

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree/ diploma at this or any other institute/ university.

S. Mukhopadhyay

Siddhartha Mukhopadhyay

ACKNOWLEDGEMENTS

To begin with I dedicate this work to my parents. I take this opportunity to express my heartfelt gratitude to one and all who inspired this work. The man who inspired me most over last two decades and encouraged me to see the positive side of every failure is Mr. G. P. Srivastava, Director E and I Group, BARC. My technical advisor, Head Control and Instrumentation Division, Mr. U. Mahapatra unconditioally supported all my efforts to pursue this work. Continuous arguments with him strengthened my fundamentals and sharpened the new theory we are proposing. I am extremely thankful to Mr. T. U. Bhatt of Reactor Control Division and my colleagues in Adavanced Magnetics and Instrumentation Section, Control Instrumentation Division, especially Mr. S. K. Lahiri, Ms. Shilpi Saha, Mr. S. K. bahuguna and Mr. Debmalya Mukherjee who were to taste the new theory in its raw form to suggest what would make the perfect blend. I am grateful to Dr. Arun K. Tangirala of IIT Madras who was as excited as me with the new findings and endless discussion with him enriched the work enormously. Internal reviews by the Doctoral committee chaired by Dr. L. M. Gantayet were extremely helpful for directing this research to a successful and logical end. I thank eminent members of the doctoral committee Dr. B. K. Dutta, Dr. P. Satyamurthy, Dr. A. P. Tiwari for their kind advice and suggestions to improve the quality of this work. My guide, friend and philosopher Dr. A. P. Tiwari had always been available to give me a helping hand to control the dynamic modes of this work particularly when they became sluggish or unstable. My son Sagnik, my daughter Arunita and my wife Keya are witness of pain and excitement I experienced while pursuing this work in last four years and I could not achieve this without their love and support.

Table of Contents

Abstract

The work in the thesis addresses the problem of identification of complex systems in presence of noise. Modeling complexities with wavelet basis improves prediction capability and accuracy of the identified model. A novel technique for estimating parameters of a strictly time varying model in wavelet domain is proposed based on consistent output prediction as an alternative solution to the classical least squares minimization problem. The work introduces and justifies the use of spline biorthogonal wavelets as a modeling tool for system identification. It suggests that weighted scalar summation of projections in approximation space could be used for deriving consistent output prediction in case model structure is built with spline biorthogonal wavelets. The method of identification could be viewed as modeling with pre-filtered input and output which renders the identified model minimum-memory (ideal case) and insensitive to noise. Resulting parameter estimates are unbiased and bounded. An iterative algorithm, alternately projecting the solution in time and wavelet domain for minimization of local error in wavelet coefficients, is proposed for output reconstruction. The algorithm is computationally efficient and exhibits excellent performance in cross validation. Stability and uniqueness issues of reconstruction by alternate projection is studied. As an extension of existing methods of reconstruction from sparse wavelet representation, a new representation called wavelet maximum curvature point representation is proposed. The algorithm ensures that the reconstructed signal contains complete information for characterization. The technique is validated by characterizing NDT signals. As a case study, the paper addresses the problem of modeling a complex process called the Liquid Zone Control System (LZCS) in a large Pressurized Heavy Water Reactor based on the evolution of input and output. In this work, an identification scheme of a linear time invariant model of the LZCS is studied. Orthogonal as well as biorthogonal wavelets are used for consistent output estimate of the LZCS process. The technique is verified on the real experimental data obtained from a full scale test setup. The concept of designing an admissible control, constrained to be memory-less, with output wavelet states is introduced and controllability of the open loop and closed loop system with output feedback is studied. The theory of controller design with wavelet states is developed in the work. The technique is demonstrated with simulation examples of multiscale systems. Point kinetic model of a nuclear reactor is studied for pole assignment by designing a wavelet state controller.

List of Figures

List of Tables

Nomenclature

SYMBOLS

FIR	Finite impulse response
IIR	Infinite impulse response
L_2	Hilbert space, span of all functions that are square integrable in Lebesgue's sense
l_2	Span of all discrete functions that are square summable
$\mathcal{L}$	Span of wavelet functions at projections with significant values
$\Im$	Span of all dyadic wavelet transforms
$\Re^n$	N dimensional space of real numbers
iid	Identically and independently distributed
β_i	Neutron fraction from i^{th}group of delayed neutron precursor
m_d	Number of delayed neutron precursor group
λ_i	Decay constant of i^{th}group of delayed neutron precursor
C_i	Concentration of i^{th}group of delayed neutron precursor
l^*	Prompt neutron life time
I	Iodine Concentration
X	Xenon Concentration
$\overline{\sigma}_I$	Microscopic thermal neutron absorption cross-section of Iodine
$\overline{\sigma}_X$	Microscopic thermal neutron absorption cross-section of Xonon
ρ	Reactivity
Σ_a	Macroscopic thermal neutron absorption cross-section
Σ_f	Macroscopic thermal neutron fission cross-section

NOTATIONS

$y\left(t\right)$	Continuous function in time
y	Uniformly sampled discrete sequence
y^s	deterministic signal component of measurement y
y^n	stchastic noise component of measurement y
$\hat{y}$	estimate of y
φ_i	Analysing basis finction φ placed at i^{th}time index, $\varphi_i = \varphi\left(t-i\right)$
φ_i^*	Reconstructing basis finction φ^* placed at i^{th}time index, $\varphi_i^* = \varphi^*\left(t-i\right)$
φ^n	Basis finction φ of degree n
Wy	Wavelet transform of y, w^y
$W_k y$	Linear multiscale causal analysis filterbank, $W_k y = w^y\left[k\right]$
$\langle x, y\rangle$	Inner product of x with y
$x \star y$	Convolution of x with y
$O(N)$	Order of N
$\mathbb{Z}$	Set of integers
$\mathbb{R}$	Set of real numbers

Chapter I

Introduction

System identification, by perturbing a plant to excite its dynamic modes is carried out for numerous reasons such as plant simulation, design of process control, plant diagnostics etc. Moreover, identification from the experimental data proves superior to analytical methods of modeling while retrofitting a control design in an existing plant which has aged over time. The accuracy of the control design depends largely on the model accuracy. Hence, it is important to arrive at a plant model which is reasonably accurate and at the same time, simple enough to be relevant to the design objective.

The classical regression model for estimating output $y(t)$ of a Linear Time varying (LTV) system is expressed in terms of parameter vector $\mathbb{H}(t)$ and regression vector $v(t)$

$$\widehat{y}(t) = v^T(t)\mathbb{H}(t) + \varepsilon(t) \tag{1.1}$$

The model is linear in $\mathbb{H}(t)$ and in case the regression vector $v(t)$ is made up of past values of input and output, the model structure is called Auto-Regressive with eXogenous inputs (ARX) model. Here, $\varepsilon(t)$ is considered to be *iid* Gaussian *i.e.* $N(0,\,\sigma^2)$ distributed additive noise at the output. An important underlying assumption, for modeling a discrete time system is that prior to sampling, continuous time input and output are pre-filtered with an ideal, band limited filter and they can be perfectly reconstructed by interpolating with orthogonal sinc functions. For some systems (for example, multiple time scale systems with clusters of poles far apart) modeling with complex sine or sinc basis necessitates very high frequency resolution to discriminate closely spaced poles as well as very high time resolution for modeling multi-scale pole clusters. As both frequency and time resolution cannot be made arbitrarily high, we need to look for alternative basis. The problem can be solved if the restriction of band limit of the pre-filter and reconstruction filter is removed and identification could use many other basis functions [12, 24, 35, 69]. In recent time, identification of linear time varying, partially linear and non-linear systems has attracted very active research interest [12, 24, 25, 28, 35, 69, 89]. In most of these works, uses of basis functions, other than classical sine or sinc function have been suggested as underlying structures of the models. Wavelet basis functions [18, 84] among them are of particular interest in this work. The structure of the model, is governed to a large extent, by the wavelets selected as basis. The identification problem in the wavelet domain is usually re-formulated using orthogonal wavelets as basis functions, while the model parameters are estimated by minimizing errors in a Least Squares (LS) sense [12]. With orthogonal basis functions, error minimization (in LS sense) in time and wavelet domain gives the same solution. A constrained minimization problem is solved however, in this work, projecting the solution onto biorthogonal wavelet basis functions.

Wavelet basis functions have excellent approximation property which is very useful for signal/system modeling. In addition, efficient shrinkage associated with wavelets can render an identified model insensitive to noise. The use of wavelets as basis functions brings with it several other advantages to the modeling arena. Primary among them are their abilities to handle non-stationary and non-linear systems. Wavelets are known to provide near optimal non-linear estimates of signals. Moreover, higher-order systems in the measurement space reduce to lower-order models in the projection space because the correlation functions in wavelet domain decay faster than the correlation functions of the original signal in time [75]. In addition, the multiresolution approximation abilities of wavelet basis functions naturally accommodate pre-filtering of data resulting an accurate model.

The specific motivation to this work originates from efforts to model a complex system (possibly) consisting of nonlinearities and integrating effects. Traditionally, a large class of non-linear systems is modeled as linear systems with time-varying parameters. However, if non-linearities are accompanied with high-order and integrating type effects, modeling in the classical input-output description (regardless of time varying nature) require significantly large number of parameters. Modeling of such effects in wavelet domain is likely to greatly reduce number of model parameters.

Thus, the objective translates to building a consistent and parsimonious LTV predictor of a non-linear system using wavelets as basis.

That a wavelet basis LTV model can capture complexities such as component nonlinearity, integrating and multiple time scale behaviours very effectively in fewer parameters as shown in this thesis. This provides a strong motivation for modeling and control design of complex systems, for example a nuclear reactor, with wavelet basis LTV or a derived wavelet basis Linear Time Invariant (LTI) model. This work has been extended to the development of a method to efficiently isolate multiple time-scales in a linear time-invariant system, a problem that has been of interest and challenge in identification and control of nuclear reactors [77]. Multi-scale Principal Component Analysis (MSPCA) combining PCA with wavelet analysis has been suggested for modeling and monitoring multivariable statistical processes[2]. The property of wavelet basis to approximately decorrelate the autocorrelated measurements makes it a good choice for modeling multi-scale systems.

For estimating LTI model parameters in LS sense, the basic assumption is that the system is overdetermined *i.e.* there are more samples available than number of parameters. A strict LTV model in general is underdetermined because dynamic solution requires derivation of more than one system parameters at each sample time instant. Hence, the parameter estimation will always fall short of samples. Several approaches have been taken to obtain determinable solution of the LTV parameter estimation problem. For instance, modeling of LTV systems using wavelet basis has been addressed in [24] where Least Mean Square (LMS) adaptive filtering algorithm is used for identification. In [23] it is shown that a wavelet model is particularly suitable for adaptive identification of Linear Periodically Time Varying (LPTV) systems. Although in these two papers it is claimed that the LMS algorithm with wavelets converges faster than the LMS Finite Impulse Response (FIR) algorithm applied in measurement space, it is well known that the adaptive LMS algorithms work only for relatively slowly time varying processes. LS solution of parameter estimation problem of a wavelet basis model, proposed in the twin paper by Zhao and Bentsman [90, 91], assumes time invariance over the length of the filter. Although the values of model parameters are updated at every time instant, the approach fails to capture abrupt change in the system such as regime switching in a process. Moreover, linear approximation as suggested by the work, may give rise to ill conditioning of estimated impulse response while modeling some of the complexities mentioned above.

A different approach has been taken in this thesis, whereby it does not necessitate the assumption of local time invariance. The method of parameter estimation, proposed here, naturally accommodates non-linear approximation [46] and primarily checks the local consistency of the estimate with output signal for a determinable minimum memory solution in wavelet domain. The method proposed in this work, exploits the fact that the wavelet coefficients have less memory compared to time samples and attempts to minimize the local error in de-correlated wavelet coefficients for estimation of system parameters. Wavelet packet identification of LTI models in frequency subbands have been suggested in order to achieve a compromise between accuracy and parsimony [59]. It has also been observed that since the model only needs to represent the system in limited frequency band, only two taps are required for fine-tuning FIR filter. This supports our assumption that it is possible to have reduced memory models in subbands. It is important to note, at the outset, that consistency in the literature of system identification generally refers to an asymptotically unbiased estimate of parameters. Strictly speaking, parameter estimate is said to be consistent when the estimate tends to the true value of system parameter[71, 88]. The method given in the thesis, predicts the output signal one step ahead and to avoid ambiguity, defines the method as consistent output prediction. The definition of consistent output prediction used in this work refers to the signal which has the same representation in wavelet domain as the original output, bearing strong similarity to the idea of consistent estimate in signal processing literature [17, 57]. The technique gives a simple and elegant algorithmic solution of LS minimization problem.

In the proposed scheme, the system is modeled in an approximation space in which both input and output are projected. Spline biorthogonal wavelets span the approximation space and can be very effectively used for system identification because of their short support and ability to get excellent approximation. Wavelets constructed from splines are called spline biorthogonal wavelets. The underlying spline functions, however, are not strictly orthogonal except for those of order zero and hence energy is not preserved in transform domain. As the system dynamics is captured by a few non orthogonal basis functions in transform domain the issue of stable reconstruction (needed for cross validation of the model) in measurement space, needs to be assured. Providentially, in case splines are used as generalized bases, weighted scalar addition of projections in approximation space could be used for consistent output prediction and it can be shown that the proposed solution seeking local fit in approximation space does not necessarily require the assumption of strict orthogonality. Hence higher order spline wavelet bases are admissible for modeling. In general, the method of estimation of model parameters, with wavelet basis, could be cast as a penalized least squares problem. For a given strictly positive threshold, the solution is arrived at by retaining wavelet coefficients whose modulus values cross the threshold [9, 20, 21].

The technique of consistent output prediction in wavelet domain is proposed as an alternative solution to the classical LS minimization problem. Although derived through a different route, this parametric identification result has a striking similarity to non parametric Time Frequency Representation (TFR) of an LTV model expressed as ratio of Continuous Wavelet Transform (CWT) of output and input [68]. TFR, however, is computationally expensive and gives a frequency domain description of the system. In contrast, the technique given in the thesis identifies a system truly in multiresolution. Moreover, an elegant algorithmic solution to the problem has been proposed, for discrete implementation.

Quality of a model is tested by reconstructing system output. Proposed technique of consistent output prediction is a non-linear approximation technique and is based on a subset of projections (decided by the choice of threshold) in the subband. The system is in fact, modeled in even lower dimension subspace spanned by the wavelet functions located at the indices where projections have significant values i.e. they cross a pre-defined threshold. Inverse wavelet transform can be seen as a mapping from a space of lower dimension spanned by a few wavelet functions to the measurement space of higher dimension. Hence, the reconstruction in time and its dyadic wavelet transform are not unique. All possible solutions associated with wavelet representation in the space of all dyadic wavelet transforms $\Im$ together make the reconstruction set residing in space $\mathcal{L}$. The minimum error norm solution is reached iteratively by projecting the solution back to time and then again projecting the crude prediction (in time) forth in transform domain. The intermediate solution in wavelet domain is forced to match significant projection values in every iteration. The alternate projection on $\mathcal{L}$ and $\Im$ is known to converge strongly to the orthogonal projection on $\Im \cap \mathcal{L}$, eventually to the member of the reconstruction set that minimizes the objective function.

In classical approach of system identification, noise is modeled in addition to system parameters. However, working in equation error framework, it has been attempted here, to identify a model from pre-filtered input and output such that identified model is less sensitive to noise. From this viewpoint, the proposed method is reminiscent of the familiar technique of identification using pre-filtered data. It is therefore worthwhile pointing out certain salient differences between the pre-filtering approach and the approach under study. Firstly, in the approach using pre-filters, the data is projected back in time followed by the usual identification exercise. In contrast, the models developed here work with projections. In this respect, the proposed method takes a different stance from that of the several existing methods using wavelets for identification. Secondly, the noise is handled in the proposed approach by means of elimination using the well-known thresholding strategy, which is equivalent to a non-linear estimation of the signal - while, the approach of pre-filtering traditionally carries out noise elimination using linear filters. In this study, it is assumed that the (white) noise directly enters the output. Consequently, the focus is on achieving the best predictions of the deterministic component of the output, but in the wavelet space. Although, in general, the proposed wavelet based model is time varying, if it is known from physics that the process is mildly non-linear and time invariant, an approximate LTI model can be derived in each subspace spanned by wavelet functions. The proposed LTI model is similar to the special case discussed in [15] where parameters are considered to be constant at each scale and may vary from scale to scale. Assuming that the unmodeled noise is a zero mean Gaussian process, it is proved that parameter error is also zero mean for a derived LTI model. Hence, it is established that parameter estimation using the method of consistent output prediction is unbiased. It can be further proved that the parameter uncertainty is bounded and the bound can be reduced by increasing level of thresholds. Hence correctness of the identified model depends on the level of threshold chosen by the designer. At the same time, higher threshold could possibly remove significant signal component thereby compromising usefulness of the identified model. A tradeoff in this regard is necessary to meet the design objective as well as quality of identification. For the LTI model, how parameter estimates can be computed from the noisy measurements is stated and proved in the thesis.

The derived LTI model can be used for control design of a nuclear reactor. A nuclear reactor is a very complex dynamical system. Modeling and identification problems for nuclear power reactors have been considered extensively in [30, 39, 41, 58, 67]. In analytical method, either complicated equations are obtained for full characterization of the process or simplifications limit the application of the model to fewer operating modes. Wavelet based models have been suggested earlier for identification of control relevant process identification [10] to extract information localized in both time and frequency. Auto Regressive with eXogenous input (ARX) or State-space models in time are popularly used for design of state feedback control. A class of multiscale state space models have been extensively studied in [15] for modeling stochastic phenomenon given noisy mesurements. There each scale is viewed as in the notion of system state for coarse-to-fine prediction using scale-to-scale linear recursion. It is well known that for a controllable system, poles can be placed at desired locations by full state feedback control [26]. However, in case of input-output modeling, system states are not visible (measureable) for full state feedback control. In this thesis, it is suggested that wavelet states of the identified model can be used for the purpose of control. The conditions for pole assignable design of a full wavelet state feedback controller are derived. The method of controller design is illustrated with an example of a multiscale model of a nuclear reactor. The proposed technique of system identification is validated by modeling a complex process in a large Pressurized Heavy Water Reactor (PHWR). Large thermal reactors such as the 540 MWe PHWR are inherently unstable with respect to xenon induced spatial oscillations due to loose neutronic coupling among different regions of the core and comparatively very small reactivity feedback effect of fuel. These spatial oscillations must be controlled, failing which, a potentially serious situation known as “flux tilting” may arise, i.e., one portion of the reactor may tend to produce more energy than other portion does. This can cause fuel in that region to exceed its thermal design limits and thus affecting fuel integrity [64, 77]. Xenon-induced spatial oscillations are kept from growing in large PHWRs, by means of Liquid Zone Control System (LZCS). The mathematical model of LZCS is useful in analysis work such as the study of reactor power variations during normal operational transients, besides in control system design [77]. A full scale test set up of the LZCS of 540 MWe PHWR has been developed at Bhabha Atomic Research Centre. Data from this test setup have been used in this work for identification and validation of models. In particular, an identification scheme of a SISO time varying ARX model of the LZCS with wavelets as basis is studied. The framework is that of uniformly sampled signals i.e. both output and input are sampled at the same rate. The system is identified using orthogonal and biorthogonal wavelets. An LTI model of a Zone Control Compartment (ZCC) is derived as a special case of time varying model. The new scheme of estimating system parameters based on consistent output prediction in wavelet domain is used for unbiased estimation of system parameters. Further a novel technique for estimating process time-delay is suggested. The estimate is shown by application to the LZCS to be insensitive to noise and is used as an input to the linear model. The proposed wavelet based model is compared with traditional and wavelet basis locally LTI, ARX type models for identifying the LZCS process. It is observed that even a low order wavelet based model gives excellent approximation from fewer numbers of input output data points, whereas the traditional models fail.

Thresholding gives rise to sparse representation in wavelet domain [20, 21]. While projecting the signal in time for approaching the solution by alternate projection, the problem can be cast as a non-linear signal approximation problem from Wavelet Representation with Missing Samples (WMSR). In literature, method of alternate projection is suggested to have a consistent estimate, based on modulus maxima or extrema representation. The same is used in this work for reconstruction from a wavelet representation with missing samples. As an extension of the existing schemes, a new scheme has been devised based on Wavelet Maximum Curvature Point Representation (WMCPR), which can be viewed as a generalization of the existing schemes for signal reconstruction from sparse representations. The algorithm seeks the solution to have minimum local error in wavelet coefficients by alternately projecting the solution in time and wavelet domain. The algorithm of alternate projection is used for denoising and compression of Non Destructive Testing (NDT) signals. An application of signal characterization using WMCPR is demonstrated using Magnetic Flux Leakage (MFL) signal. Estimated defect size based on the reconstruction is found to be acceptable.

The thesis is organized as follows. Chapter I introduce the topic and gives a brief overview of the work.

Chapter II gives an overview of the existing work on system identification using wavelet basis functions and places the proposed work in that backdrop indicating relevance and necessity for the development. Distinction and novelties of the proposed method $vis - \grave{a} - vis$ the existing modeling methods are clearly indicated.

Chapter III discusses the fundamentals of system modeling using wavelets as basis, establishes the admissibility of wavelet basis functions for system identification with time varying linear models and formulates the identification problem in wavelet domain. It has been shown that modeling with spline biorthogonal wavelets can be viewed as identification with prefiltered input and output.

Chapter IV presents the main contribution of this work *e.g.* estimation of parameters of a linear time varying model using consistent output prediction. The chapter also gives a brief introduction to the LZCS in a PHWR. Details of experiments conducted in a full scale test set up and results of the experiments are discussed. It is shown that an unbiased LTI model can be derived as a special case of LTV model and parameter estimates by consistent output prediction from noisy measurements are unbiased and bounded. Issues of signal representations in wavelet domain using Undecimated Discrete Wavelet Transform (UDWT) are discussed. A recipe of the algorithm for parameter estimation of an LTV system and reconstruction of consistent output by alternate projection is presented. Efficacy of modeling with orthogonal and biorthogonal wavelets is demonstrated by cross validation using data from a full scale test set up of the LZCS. Accuracy of the derived model in presence of noise is established.

Methodology of closed loop control design using output wavelet states is demonstrated in Chapter V. The notion of designing a control with output wavelet states is introduced and controllability of the open loop and closed loop system with output feedback is studied. The technique is demonstrated with multi scale simulation examples.

Chapter VI concludes the work indicating major achievements and future scope.

Apart from the work carried out to achieve the main objective, a few related theories have been developed during the course of the programme. These theories and the verification results using simulation and actual plant data have been documented in the appendices as stand alone supporting works. Data denoising and signal characterization based on wavelet maximum curvature representation, with a novel dual thresholding scheme has been presented in Appendix A. Appendix B presents a new robust technique for estimation of time delay in a process. The technique is used for estimating transport delay in the LZCS.

Chapter II

Literature Survey

The field of system identification is well developed with established and widely understood techniques. To formulate a problem of parametric identification, first it is necessary to choose the basis for approximation. Model parameters are estimated based on a criterion, such as that of LS, to match the output for the given input data. In this sense, the best identified system is only as good as the input-output pair. Model validation is the last step in identification and it aims at assessing objectively, whether the identified model agrees sufficiently well with the observed data. The generic model structure, given by Ljung [43, 69], can be extended to the family of linear difference equations, parameterized by the time varying linear filters A^t, B^t, C^t, D^t and F^t.

$$A^t(q)y(t) = \frac{B^t(q)}{F^t(q)}u(t) + \frac{C^t(q)}{D^t(q)}\varepsilon(t) \tag{2.1}$$

Here q denotes the shift operator, $y(t)$ measured output, $u(t)$ measured input and $\varepsilon(t)$ is called disturbance, error or noise. The most used ARX model ($F^t = C^t = D^t = 1$) relates the current output $y(t)$ to a finite number of past outputs $y(t-k), k = 1, 2, ...P$ and inputs $u(t-k-d), k = 1, 2, ...Q$. d denotes pure time-delay or the dead-time in terms of number of samples. There are several elaborations (referred as ARX type model, later in this thesis) of the basic ARX model, where different filter structures are introduced to provide flexibility. These include the well known Auto Regressive Moving Average with Exogenous input (ARMAX) ($F^t = D^t = 1$), Output-Error (OE) ($A^t = C^t = D^t = 1$), and Box-Jenkins (BJ) ($A^t = 1$) model.

Modeling systems in transform domain, in Fourier [63] or Gabor [25] space are single resolution alternatives to time domain representation. A comprehensive discussion on mathematical foundation of linear and nonlinear black box identification problem with different basis can be found in [35] where it has been observed that the quality of identification is a tradeoff between the number of parameters used to describe the model and the approximation error. Error is found to increase proportionally with the number of parameters. Hence a good approximation which minimizes number of parameters by suitable choice of basis function is important for good identification. This is a

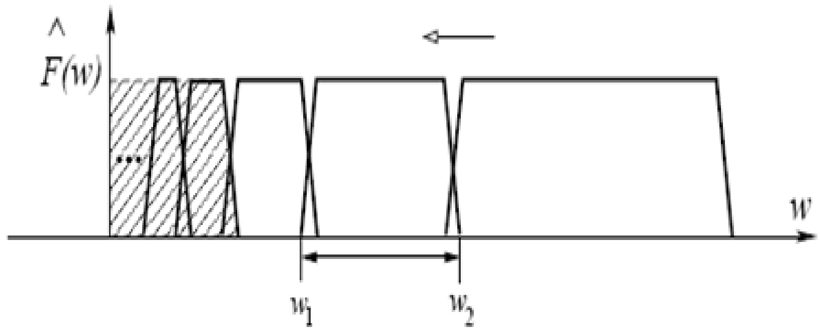

Figure 2.1: Constant Q wavelet filters

strong motivation in favour of compact support wavelet basis representation of a system which minimizes number of model parameters. The paper [35] introduces the wavelets as basis for identification and establishes its usefulness in system modeling.

2.1 Properties of wavelet transform

Wavelets are basis functions that span the space (L_2) of functions with finite energy *i.e.* any function in this space can be represented by wavelet basis. Wavelets are constant Q $\left(=\frac{\triangle\omega}{\omega}\right)$ band pass filters. As we keep dilating wavelets in time, $\triangle\,\omega \to 0$ as shown in Figure 2.1. Ideally (countable) infinite number of wavelets are needed to represent a function in L_2. Instead, in wavelet theory it is proved that dialates and translates of scaling function $\phi\,(t)$, combined with wavelets $\psi\,(t)$ span the whole frequency axis. A measurement in terms of wavelet basis can be expressed as follows.

$$y\,(t) = \sum_k c_k \phi_k\,(t) + \sum_{j,k} d_{jk}\psi_{jk}\,(t) \tag{2.2}$$

Wavelet basis functions have many interesting properties that make them ideal for function approximation.

1. Wavelets are known to provide near optimal non-linear estimates of signals [46]. A non linear estimation (with I basis) may use basis functions located at the indices of I largest projections and hence, $I \notin \mathbb{Z}$. As I increases, approximation error in case of non linear estimation decays faster than that of linear estimation. With a priori knowledge about the system or signal, an appropriate choice of basis can be made from a large range of wavelet functions to obtain a sparse representation needed non linear estimation.

2. Wavelets are natural choice for modeling time varying systems. They are local approximators with compact support (in time and frequency). Figure 2.2 shows compact support B-spline scaling functions and wavelets. These functions are ideal for projecting time limited signals or functions in shift invariant subspace of L_2 [81].

3. The correlation functions in wavelet domain decay faster than the correlation functions in time. Thus, higher-order systems in the measurement space reduce to lower-order models in the projection space.

4. The multi-resolution approximation abilities of wavelet basis functions naturally accommodate pre-filtering of data working in the sub-space of measurement space [52].

5. Wavelets efficiently isolate multiple time-scales in an LTI system because they use variable size time-frequency atoms as shown in the second grid of Figure 2.3 (first grid showing grid for Short Time Fourier Transform (STFT)).

6. Wavelet shrinkage (denoising) enables robust estimate of parameters.

2.2 Admissibility of wavelets as basis

An LTV system, in general, can be described by a parametric or non parametric linear time varying impulse response function $h(t,\tau)$. A time varying ARX (TVARX) model output using the continuous impulse response function is given by

$$\hat{y}(t) = h(t,\tau) \star u(t) \tag{2.3}$$

Noise term is considered to be zero for the time being and shall be included later in the formulation. Output of the system at discrete time t is

$$y[t] = \sum_{k\in\mathbb{Z}} h[t,k]\, u[t-k]\,,\ t \in \mathbb{Z} \tag{2.4}$$

If the discrete system is causal and discrete version of the impulse response function $h[t,k]$ is appropriately truncated, number of past inputs required to represent output at each time t is equal to the length of the impulse response function. In general, parametric representation of $h(t,\tau)$ is expressed as a weighted sum of infinite number of integer indexed basis functions. Let us define the shift invariant infinite dimension sub-space V of Hilbert space, $\mathcal{H} = L_2$ (space of all functions that are square-integrable in Lebesgue's sense) as

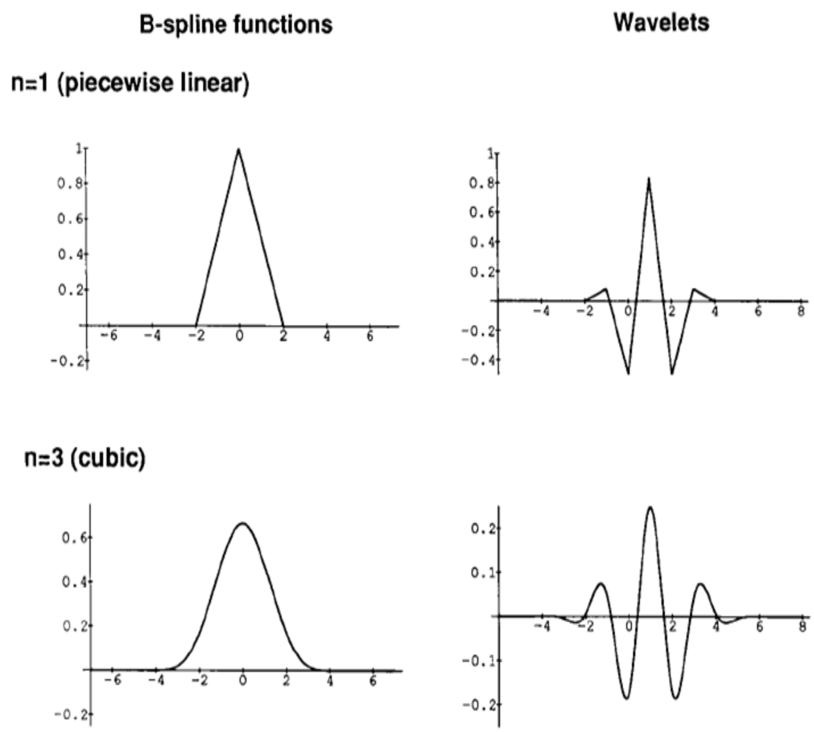

Figure 2.2: Compact support B-spline scaling functions and wavelets at resolution level - 1

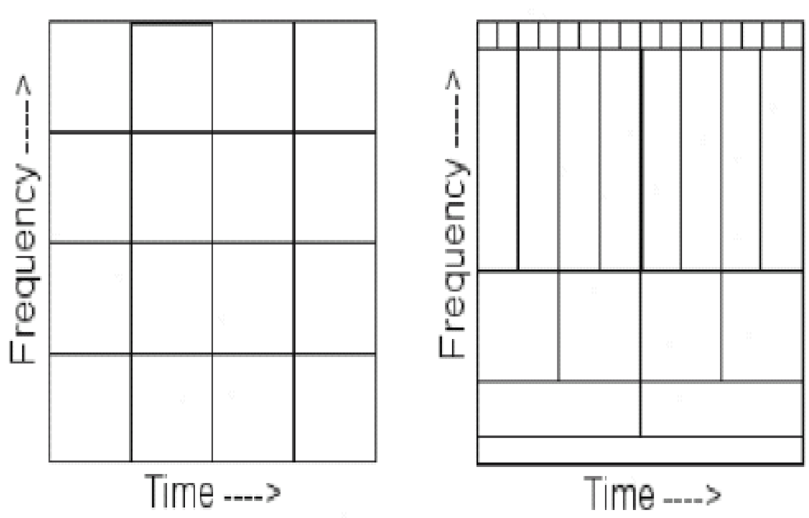

Figure 2.3: Variable size time-frequency atoms for wavelets

$$V(\varphi) = \left\{ h(t,\tau) = \sum_{i \in \mathbb{Z}} \hbar_i(t)\, \varphi_i(\tau) : \hbar_i \in l_2(\mathbb{Z}) \right\}, V \subset L_2 \tag{2.5}$$

Shift invariant basis functions $\varphi_i(t)$ and its time varying coefficient $\hbar_i(t)$, together constitute the discrete-continuous model of the time varying system.

$$y(t) = h(t,\tau) \star u(t) = \sum_{i \in \mathbb{Z}} \hbar_i(t)(\varphi_i \star u)(t) \tag{2.6}$$

Output of the system given by the discrete impulse response $\hbar_i \in l_2(\mathbb{Z})$, at discrete time t is given by

$$y[t] = \sum_{k \in \mathbb{Z}} \left(\sum_{i \in \mathbb{Z}} \hbar_i[k]\, \varphi_i[k] \right) u[t-k] \tag{2.7}$$

In case of uniformly sampled data in single resolution, it is easy to see that $i = k$.

For φ_i to qualify as basis, it is necessary that three conditions are satisfied *e.g.* the sequence of coefficients $\hbar_i$ must be square summable, the family of basis functions should form a Riesz basis of $V(\varphi)$ and the basis functions satisfy the partition of unity condition [81].

Let us define now Wavelet Transform (WT) given by an operator W. Applying W to the noisy measurements WT of $y(t)$ is obtained as

$$Wy[t] = w^y[t]$$

Right hand inequality of Riesz basis condition, stated below in (2.8), considering φ_i as wavelets ensures a stable reconstruction by the inverse WT operator because the energy in discrete and continuous domain satisfies the following condition.

$$A \left\| w^y[t] \right\|^2 \leq \left\| \sum_{i \in \mathbb{Z}} \hbar_i(t)(\varphi_i \star u)(t) \right\|^2 \leq B \left\| w^y[t] \right\|^2,\ 0 < A,\ B < \infty \tag{2.8}$$

By virtue of (2.8), a solution to the parametric identification problem could be obtained in wavelet domain as well, by minimizing the total energy (in LS sense) as discussed in chapter 4. $\hbar_i(t)$ are the parameters which are estimated satisfying LS criterion. The scaling functions of wavelet transform satisfying all the three conditions mentioned above are admissible as or basis functions. In addition they satisfy a two scale relation

$$\phi\left(\frac{t}{2}\right) = \sqrt{2} \sum_{k \in \mathbb{Z}} f_k \phi(t-k) \tag{2.9}$$

Table 2.1: Classification of spline wavelets

Wavelet type	Orthogonality	Compact support	Implementation
Orthogonal	Yes	No	IIR/IIR
Semi-orthogonal	Inter-scale	Analysis/ Synthesis	Recursive IIR/FIR
Shift-orthogonal	Intra-scale	No	IIR/IIR
Biorthogonal	No	Yes	FIR

where f_k is refinement filter. Again, to admit wavelet basis functions instead of a single space $V(\phi) = V_0$, a ladder of rescaled subspaces are considered. These subspaces are indexed by scale number j and are given by

$$V_j = \text{span}(\phi_{j,k})_{k\in\mathbb{Z}} \text{ with } \phi_{j,k}(t) = 2^{-\frac{j}{2}}\phi\left(\frac{t}{2^j} - k\right) \tag{2.10}$$

If ϕ satisfies (2.9) then these spaces are nested and form a multi-resolution analysis (MRA) [49] of L_2. Defining difference spaces $W_j = V_{j-1} - V_j$, wavelet basis functions ψ given by

$$\psi\left(\frac{t}{2}\right) = \sqrt{2}\sum_{k\varepsilon\mathbb{Z}} g_k\phi(t-k) \tag{2.11}$$

can be designed for the discrete-continuous models, such that they form the Riesz basis of difference spaces W_j, i.e. $W_j = span\,(\psi_{jk})_{k\in\mathbb{Z}}$.

Wavelets with underlying B-spline scaling functions are called spline wavelets. For a given order, the basis functions of spline wavelets are shortest (ideal for modeling time varying systems), most regular and have the best approximation properties among all known wavelets. Unlike most other wavelet bases, they have explicit formulae in both time and frequency domain allowing easy manipulation. Spline wavelets can be classified into four categories, primarily based on their orthogonality property as shown in Table 2.1 [80].

Although orthogonality condition is not satisfied, spline biorthogonal wavelets are very popular, primarily because of the ease with which they can be implemented. Properties like symmetric FIR filter structure and compact support are best suited for model identification based on time limited input-output signal. In this thesis, advantages of using spline biorthogonal wavelets as underlying basis for modeling have been studied in detail. The refinement filter in (2.9) in this case is given in frequency domain by

$$f(\omega) = \sqrt{2}exp\left(\frac{-i\epsilon\omega}{2}\right)\left(cos\frac{\omega}{2}\right)^p \tag{2.12}$$

with $\epsilon = 0$ for p even and $\epsilon = 1$ for p odd. The scaling function computed is then spline of degree $p-1$.

$$\phi(\omega) = exp\left(\frac{-i\epsilon\omega}{2}\right)\left(\frac{sin\frac{\omega}{2}}{\frac{\omega}{2}}\right)^p \tag{2.13}$$

Since ψ is a linear combination of splines given by (2.11), it is a compactly supported polynomial spline of same degree.

Representation of signals or functions with Undecimated Discrete Wavelet Transforms (UDWT) [1] is rigorously treated in the next sub-section.

2.3 Representation with undecimated discrete wavelet transfom

A brief introduction to UDWT [8, 17] is presented in the beginning of this section.

Let us assume that all signals are square summable, discrete time sequences $x[k], k\epsilon\mathbb{Z}\ i.e.$ the space is $l_2(\mathbb{Z})$. UDWT is a linear bounded operator W consisting of $J+1$ linear operators

$$W_j : l_2(\mathbb{Z}) \to l_2(I),\ l_2(I) = l_2(\{j = 1, 2, \ldots, J+1\} \times \mathbb{Z}) \tag{2.14}$$

In wavelet literature j is referred to as scale, as an alternative to frequency. One can compute UDWT of the discrete signal $x[k]$ with a bank of low-pass filters (f) and a high-pass filters (g). Filters f and g are finite impulse response (FIR) filters. The resulting sequence of discrete signals of the form $\{(W_j x), 1 \leq j \leq J,\ U_j\}$ are called the UDWT of the sequence $x[k]$. The operators W_j and U_j, for undecimated dyadic discrete wavelet transform (DWT), are the convolution operators giving details and approximation of a signal at the scale j. The impulse responses of the undecimated octave band filter bank with z-transform $F(z^{2^j})$ and $G(z^{2^j})$, i.e. f_j and g_j can be obtained by putting $2^j - 1$ zeros between two consecutive coefficients of f and g, respectively. The reconstruction operator W^*, the inverse of W, can also be implemented by a undecimated octave band filter bank $\widetilde{F}(z)$ and $\widetilde{G}(z)$. For perfect reconstruction of an arbitrary signal $x \in l_2(\mathbb{Z})$ from its wavelet transform, Wx, it is necessary and sufficient that there exist two FIR filters $\widetilde{F}(z)$ and $\widetilde{G}(z)$, satisfying

$$F(z)\widetilde{F}(z) + G(z)\widetilde{G}(z) = 1 \tag{2.15}$$

$\widetilde{F}(z)$ and $\widetilde{G}(z)$ are referred to as the undecimated synthesis octave band filter bank. In case of decimated filter bank [73] an equation very similar to (2.15)

[1] For the development of the proposed theory of identification, shift invariant UDWT is considered to be the underlying transform, throughout the thesis.

is obtained with the exception that the right hand side has a factor of $2z^{-l}$, i.e.

$$F_1(z)\widetilde{F}_1(z) + G_1(z)\widetilde{G}_1(z) = 2z^{-l} \tag{2.16}$$

Subscripts 1 denote filters for decimated analysis and synthesis. With the substitution $\widetilde{F}_1(z) = G_1(-z)$ and $\widetilde{G}_1(z) = -F_1(-z)$, (2.16) takes the following form:

$$z^l F_1(z)\widetilde{F}_1(z) - z^l F_1(-z)\widetilde{F}_1(-z) = 2 \tag{2.17}$$

Let us define the product filter $P(z) = \frac{z^l F_1(z)\widetilde{F}_1(z)}{2}$. Also the discussion is restricted to odd length of filters, i.e. to odd values of l. The max-flat product filter$P(z)$, satisfying (2.17), is designed as a symmetric polynomial with only odd powers of z. All even powers in $P(z)$ are zero except the constant term. With this substitution (2.17) becomes $P(z)+P(-z) = 1$, for odd values of l, because

$$P(-z) = \frac{(-z)^l F_1(-z)\widetilde{F}_1(-z)}{2} = \frac{-z^l F_1(-z)\widetilde{F}_1(-z)}{2} \tag{2.18}$$

It may be noted that all odd powers of $P(z)$ cancel those of $P(-z)$.

One obvious choice of undecimated filter bank would be given by

$$F(z)\widetilde{F}(z) = P(z) = \frac{z^l F_1(z)\widetilde{F}_1(z)}{2} \tag{2.19}$$

and

$$G(z)\widetilde{G}(z) = P(-z) = \frac{-z^l F_1(-z)\widetilde{F}_1(-z)}{2} = \frac{z^l G_1(z)\widetilde{G}_1(z)}{2} \tag{2.20}$$

The substitution of (2.19) and (2.20) in (2.15), which defines the perfect reconstruction relationship for undecimated filter bank, also takes the form of (2.16). Hence in undecimated case, the same FIR filter coefficients, odd length and symmetric, derived in decimated case, can be used after dividing each of them by $\sqrt{2}$ [54]. This allows us to test signals with the family of readily available, biorthogonal and reverse biorthogonal wavelets, in our work. The impulse responses of f_j and g_j, $\widetilde{f}_j$ and $\widetilde{g}_j$, used in this work, are either symmetric or anti-symmetric. Non-causal implementation of the filter banks ensures zero delay in the reconstructed signal.

Although a given system is identified in wavelet domain, reconstruction of the system output from a wavelet representation of past values of input and output is necessary for cross validation of the output estimate. Reconstruction from wavelet multi-scale maxima or zero crossing representations have been suggested in the beginning of nineties and stability and uniqueness issues were long deliberated [4, 17, 48]. Reconstruction and transforms are always performed on the regular sampling interval although sampling of representation in transform domain is usually irregular. Hence, the topic can be viewed as bordering the more general and classical topic of signal reconstruction from irregular samples. An overview of the status of contemporary work on sampling including irregular sampling may be obtained in an excellent depiction by Unser [81]. Generalized or multichannel sampling introduced by Papoulis in 1977 for signal reconstruction from measurements performed in structured manner, and its many variants such as interlaced and derivative sampling have become popular in the last decade. Design of robust and computationally simple reconstruction techniques are still active research topics [50, 78]. In this work, however, fixed rate sampling is used by employing undecimated wavelet transform. Recently the notion of compressed sensing (CS), put forward in [78], suggests that a signal could be accurately reconstructed from fewer samples than regular nominal numbers by solving for the transform coefficients consistent with measured data and having minimum l_1 norm. The results are demonstrated to work nicely on several synthetic experiments mimicking problems in imaging and spectroscopy. The method of CS seems to work in reverse direction as compared to earlier ones based on multi-scale edges. A number of reconstruction algorithms have been proposed based on consistent estimate and very accurate reconstructions are reported using alternate projection. Basic theory of alternate projection is discussed in next few paragraphs.

Let the input signal $x\,(t)$ belong to L_2 and $W_j x$ be its (dyadic) wavelet transform at scale j. A function $x^c(t)$ is defined such that a chosen wavelet representation of x^c equals that of x. The set of abscissa's where representation of Wx is significant and non zero is denoted by t_k^j. Two conditions are imposed on the wavelet representation $R_m x^c$ of $W_j x^c$.

1. At each scale j, at all t_k^j, $W_j x^c = W_j x$.

2. At each scale, projections in the representations are located only at t_k^j.

Let $\Im$ be the closure of the linear combination of all (dyadic) wavelet functions in L_2 and assume that the projections of x^c on $\Im$ is equal to the projection of x on $\Im$, for all abscissa t_k^j. Let O be the orthogonal compliment of $\Im$ in L_2. Thus, $\Im \bigoplus O = L_2$. Defining $x^o\,(t)$, $x^o \in \Im$, x can be expressed as $x = x^c + x^o$. Here, if $U = L_2$, then $O \in \{0\}$, which implies that $x = x^c$. As this is not in general true, condition 1 above does not uniquely characterize x. Condition 2 stated above is approximated using convex functions.

Let $\mathcal{L}$ be the affine space of sequence of functions $w_j^{x^c}\,(t)$ such that at any scale j and at all t_k^j, $w_j^{x^c} = W_j x$. The solution lies in both the spaces

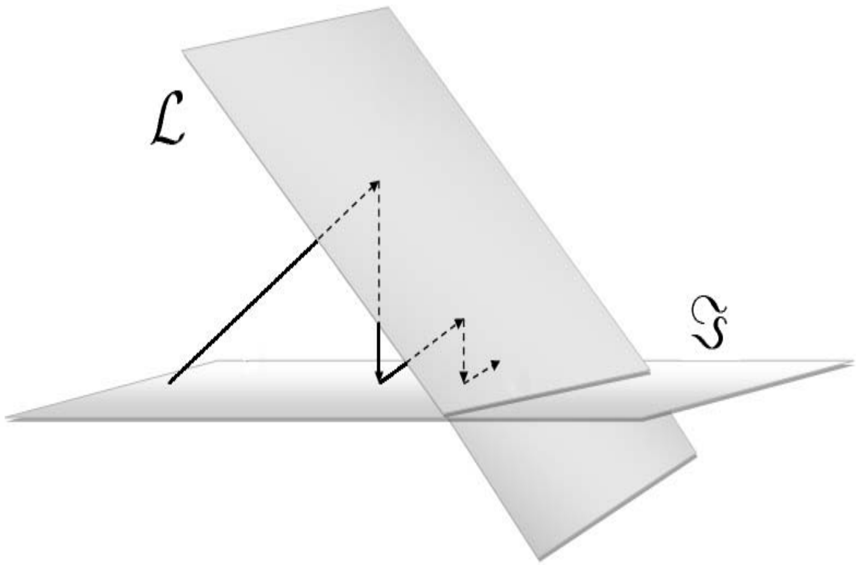

Figure 2.4: Alternate projection algorithm

$\Im$ and $\mathcal{L}$, which implies (dyadic) wavelet transform that satisfy condition 1 are sequences of functions that belong to $\Im \cap \mathcal{L}$. The objective is to find the element in $\Im \cap \mathcal{L}$ whose norm is minimum. This is achieved by alternate projections on $\Im$ and $\mathcal{L}$. If W denotes (dyadic) wavelet transform and W^* denotes inverse (dyadic) wavelet transform, then any (dyadic) wavelet transform, $w_j^{x^c}(t)$ remains invariant under the operator $P_\Im = WoW^*$. As $\Im$ is a space of all (dyadic) wavelet transforms, $P_\Im w \in \Im$ for any sequence $w = w_j^{x^c}$. Thus $P_\Im$ acts as a projector of w on $\Im$. The projector on $\mathcal{L}$, operator $P_\mathcal{L}$ is implemented by retaining the magnitude of representation at each scale at the appropriate abscissa's and convex interpolation. Convex constraints are imposed to suppress any spurious oscillations in the reconstructed wavelet transform. Defining a new operator, $P = P_\Im o P_\mathcal{L}$, P performs alternate projections in both the spaces. Since $\Im$ belongs to L_2 and $\mathcal{L}$ is the affine space, repeated application of alternate projection, $P^{(n)}$ reduces the normal distances between two spaces and $Lim_{n \longrightarrow +\infty} P^{(n)} w = P_{\Im \cap \mathcal{L}} w$. Alternate projections on $\Im$ and $\mathcal{L}$ strongly converge to the orthogonal projection on $\Im \cap \mathcal{L}$. The solution in fact converges to an element of $\Im \cap \mathcal{L}$ whose norm is minimum. The imposition that the norm of $w_j^{x^c}$ is as small as possible on average, generally creates significant projections locally at t_k^j's. Figure 2.4 shows a graphical representation of the alternate projection algorithm.

2.4 *Wavelets in identification: A review*

In last two decades, a number of papers suggested usefulness of wavelet bases, in particular, orthogonal wavelet bases for identification of systems with nonlinearities. Sureshbabu and Farrell [74] gave a strong motivation for using local orthogonal wavelet approximators for modeling non linear systems. Moreover, they suggested accuracy of approximation can be improved locally where system function changes rapidly by adding new basis functions to the existing model without affecting the existing model. They propose system identification based on compactly supported orthogonal wavelets such that sample distribution can be considered constant over the support and thresholding is permitted to zero out insignificant wavelet coefficients. It may be noted here that decimated filter bank implementation of an orthogonal wavelet transform is non-invariant in time. Time or shift invariance [46, 48] of a representation is a preferred property from the point of view of system identification. Conventionally, redundant wavelet representations are used to alleviate the problem of time non-invariance. Generalized classes of orthogonal wavelets with time invariance property have also been suggested for parameter estimation/detection of signals in [3, 62]. A recursive criterion for selecting different sets of orthonormal wavelet coefficients is devised in the paper to reflect the matching properties of a redundant representation to a signal. So, redundant as well as non redundant representation with orthogonal wavelet is possible and can be used for signal and system modeling.

The problem of nonlinear system modeling can also be reduced to scattered data interpolation problem using wavelet bases. Considering that input and output are sampled at regular time interval, non linear mapping of output with input as argument, would in general, be on an irregularly spaced grid. In [40] interpolants are computed that minimize a wavelet based norm associated with a Reproducing Kernel Hilbert Space (RKHS) subject to interpolatory constraints decided by the scattered data. Although wavelet functions are designed to form an orthogonal basis for L_2, the kernels are not shift invariant and provide spatially varying resolution suitable for irregularly distributed data samples. On contrary to this, our approach is to describe any system on a regular grid as a linear combination of basis function by embedding the nonlinearity, if present, in the basis. However, we work with sparse representations in wavelet domain and prove that a stable and consistent reconstruction of output in time can be obtained.

A methodology for identifying Nonlinear Auto-Regressive Moving Average with eXogenous input (NARMAX) models from noise corrupted data is introduced based on semi-orthogonal wavelet multi-resolution approximation in [16]. For a non linear model, in general output can be expressed in terms of nonlinear function f of the finite dimensional regressor obtained by the regression function g_k.

$$y[k] = f(g_k(y[k-1], \ldots, u[k-1] \ldots, \varepsilon[k-1] \ldots), \rho_k)$$

Best non linear parameterization of f is searched as a linear expansion in terms of the basis functions of non linear regressors g_k such that

$$f = \sum_{k \in \mathbb{Z}} \rho_k g_k$$

The family of non linear basis functions or regressors used for approximation of f are B-spline scaling and wavelet functions.

Broadly identification techniques fall in two categories.

1. Model Identification with input as argument, sampled at irregular grid points.

2. Model Identification with time as argument, both input and output sampled on a regular grid.

Most of the works mentioned above, address the problem of identification from input-output measurements, sampled at irregular grid points. But since most practical systems employ uniform or regular time sampling, this work is restricted to uniformly sampled signals and to shift invariant basis functions for identifying system impulse response function. For non linear systems, uniform time sampling may turn out to be inadequate when output is sampled with respect to input and it is directly mapped with respect to input. To get rid of this problem, the system function with time as argument is searched rather than input as argument.

Modeling of linear time varying systems using wavelet basis has been addressed in [24, 90]. Time varying impulse response of the system is modeled in these papers, either from input side or from output side as given below.

$$From\ input: \ y(t) = \sum_I a_I(t)(\eta_I \star u)(t) \tag{2.21}$$

$$From\ output: \ y(t) = \sum_I \xi_I(t)(a_I \star u)(t) \tag{2.22}$$

where a_I are time varying parameters of the system and $\{\eta_I(t)\}_I$ and $\{\xi_I(t)\}_I$ are wavelets used to expand time varying response function from input side and output side respectively. $I = (i,\ j)$, where i is the shifting and j is the scaling parameter. From either models however, it is possible to derive a model structure with constant parameters a_{IJ} in the following form [24, 86].

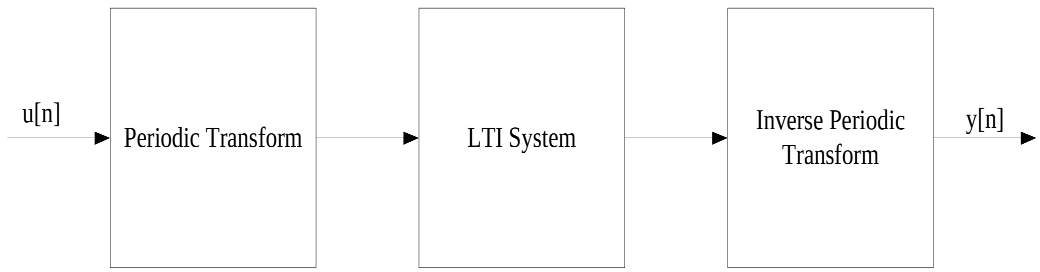

Figure 2.5: Generalized raised model with wavelet basis for identification of LPTV systems

$$y\left(t\right)=\sum_{I}\xi_{I}\left(t\right)\sum_{J}a_{IJ}\left(\zeta_{J}\star u\right)\left(t\right) \tag{2.23}$$

Modeling of LPTV systems is treated as a specific case in the time varying framework considering output functions $\left\{\xi_I(t)\right\}_I$ to be periodic. Generalized raised model with wavelet basis has the similar structure (Figure 2.5) and it has been found to be particularly suitable for adaptive identification of LPTV systems. It is possible to identify an LTI model in each frequency sub-band with the periodic transform and its inverse implemented using wavelet analysis and synthesis filter banks respectively. However, in [23] the problem has been reformulated using Wavelet Analysis Tree (WAT) and Wavelet Synthesis Tree (WST) (first and third block respectively) to equalize the sampling rates of all tree branches. The LTI system is identified as a Multiple Input Multiple Output (MIMO) FIR system.

Although, not the best choice, for identification of constant system parameters a_{IJ}'s, wavelet Least Mean Square (LMS) adaptive filtering algorithm has been suggested [23, 24]. Except under strong restrictions (sufficient order modeling with orthonormal wavelets, white noise input), the optimal parameter vector is time varying. Parameters a_{IJ}'s are estimated by minimizing mean squared error between the adaptive filter output and the output of the unknown system (Figure 2.6). For adaptation, error could be minimzed in

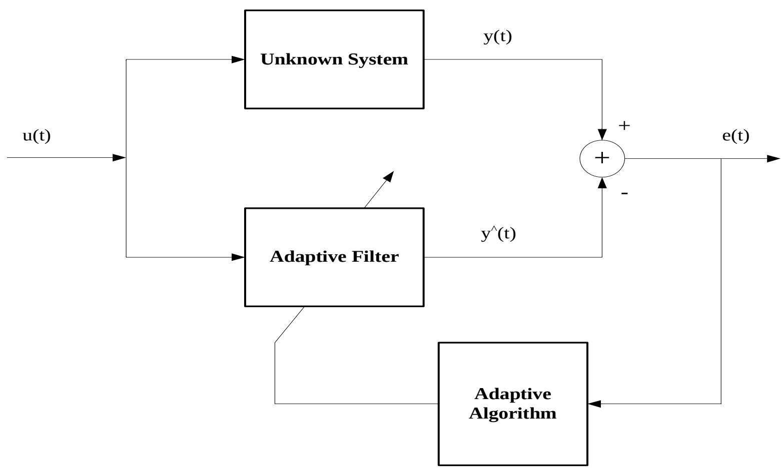

Figure 2.6: System identification by LMS algorithm

wavelet domain as well if orthonormal wavelets are used for modeling [23]. The wavelet LMS algorithm works only if the system is slowly time varying and variations are small. However, wavelet LMS algorithm converges faster than the FIR LMS algorithm and approaches optimality with lower error.

Identification of relatively fast changes in an LTV system, due to natural dynamics and/ or faulty behaviour has been addressed in the twin papers [90, 91]. LTV system modeling, in this work, has the similar structure as discussed in (2.21), (2.22) and (2.23). The work proposes a general function space approximation based framework for modeling and identification. The framework is rigorous and with the assumption of Banach space as the approximation space, it is shown that the approximation error converges to zero as the number of terms in the approximation increases. However, for the identification of system parameters, the time varying system is assumed to be locally time invariant, (ideally over the total length of filter, $L = \bigcup_{i=1,2,...I} (L_i)$, where I is the system order). Although, time variation is addressed by parameter updation at every time instant, the assumption puts serious limitation for approximating rapidly changing systems. It can be seen in "examples and discussion" section of the same paper, that the error in identification (using Daubechies length 4 filer) is high where time domain system has a unit step change. The formulation allows modeling with biorthogonal basis functions and provides mathematical tools for finding time-frequency localized bounds on the approximation error. However, the approximation error can be unacceptably high (ill conditioning problem) due to following reasons which are also interrelated.

1. Model order (number of basis function used for approximation) is insufficient.

2. The formulation precludes the possibility of non linear approximation.

3. Assumption of local time invariance is not in agreement with actual system dynamics.

It may be noted here for the sake of completion, that non parametric time frequency representation has the similar form as the parametric representation derived in the present work [68]. A technique to model LTV systems in frequency domain has been proposed in the above reference employing Continuous Wavelet Transform. The method is similar to Empirical Transfer Function Estimate (ETFE) [42]. The method works with time frequency representation such that a non linear LS estimator could be used to obtain the set of parameters of the model transfer function, optimized in multiresolution.

2.5 Motivation and relevance

The approach in this work is similar to the above mentioned works, only to the extent that the estimation of $a_I(t)$ is achieved by minimizing local error in LS sense. An adaptive filter or the technique proposed in [91] also looks for a local solution of the time varying estimation problem. The major difference of the identification method, proposed in this thesis, with the existing methods are as follows.

1. Instead of identifying the system in time domain, parameters are estimated in wavelet domain (in multiresolution) by projecting filtered output and its shifted version onto the same basis as that used for the input. The technique utilizes the parsimony in the representation, the real strength of wavelet basis.

2. Many existing techniques assume that it is possible to model a linear time varying system as a number of linear time invariant systems in the subspace. The proposed technique makes no such assumption and

identifies LTV models in the wavelet subspace. Moreover, the identified subband models are strictly time varying. The solution does not even need any explicit assumption of local time invariance because the identification is done truly in multi-resolution with decorrelated projections.

3. The proposed technique permits non linear approximation to alleviate ill conditioning problem (blowing up of local error) encountered in case of linear approximation of system model. Method of non linear approximation models a system with basis functions located at significant projections and has additional advantage of reducing model order considerably.

4. It may be noted here, that for modeling a practical system, a time varying ARX model may be better suited. Although it could be included in the ambit of the existing formulations, the TVARX formulation is not discussed explicitly. TVARX modeling is popular in application domain and an example of modeling by expanding time varying parameters with temporal basis functions can be found in [66], for detecting changes in the dynamic stiffness of human elbow following the onset of a broadband perturbation. This thesis particularly focuses on TVARX formulation with wavelet basis and related issues.

5. Let us assume that the coefficient function $a_I(t)$ can be expressed in terms of basis functions $\theta_J^*(t)$ and (2.21) can be written as

$$y(t) = \sum_{I \in \mathbb{Z}} \sum_{J \in \mathbb{Z}} a_{IJ} \theta_J^*(t) (\eta_I \star u)(t) \tag{2.24}$$

Projection of the output onto the flipped version of the basis *i.e.* $\theta_J(t)$ would be given by

$$\langle y(t), \theta_J(t) \rangle = \sum_{I \in \mathbb{Z}} a_{IJ} (\eta_I \star u)(t) \mid_{t=J} \tag{2.25}$$

since, $\langle \theta_I(t), \theta_J^*(t) \rangle = \delta_{I-J}$. The solution of the problem in general is underdetermined because it requires estimation of IJ parameters from J equations. Hence, the best set of a_{IJ} (containing less than or equal to J elements) is searched that would minimize error in output projections in least squares sense. For example, if summation is omitted in (2.25), a determinable set of a_{IJ} is obtained as solution. The result has a striking similarity to time frequency representation of an LTV model defined in continuous domain, expressed as ratio of Continuous Wavelet Transform of output and input [68],

$$TFR(j,k) = \frac{W_{(j,k)}\left[y(t)\right]}{W_{(j,k)}\left[u(t)\right]}$$

The technique however, is computationally expensive and gives a frequency domain description of the system. In contrast to the existing methods of identification where an analytical solution is obtained, the technique proposed in this thesis suggests an elegant, computationally efficient, algorithmic, alternative solution to the LS problem, amenable for discrete implementation.

6. In general, a model is tested and validated by reconstructing the output. The synthesis basis maps the function in approximation space to the output space using the inverse transform (complex sine, sinc, cardinal spline etc in single resolution case). In the proposed scheme with wavelets, mapping from the lower dimension approximation space to the higher dimension output space is implemented using alternate projection algorithm. The algorithm iterates the solution between the spaces imposing convex constraints in each iteration and converges to the true solution in the limit.

7. Projections of measurements onto wavelet basis space is equivalent to filtering the measurements using multi-rate filter banks [46]. Pre-filtering of data prior to identification is not uncommon. Extended instrumental variable methods which employ pre-filtering of the data is shown to be generically consistent [71]. From this viewpoint, the proposed method is reminiscent of the familiar technique of identification using pre-filtered data [43]. It is therefore worthwhile pointing out certain salient differences between the pre-filtering approach and the approach under study. Firstly, in the approach using pre-filters, the data is projected back in time followed by the usual identification exercise. In contrast, the models developed here work with projections. In this respect, the proposed method takes a different stance from that of the several existing methods using wavelets for identification. Secondly, the noise is handled in the proposed approach by means of elimination using the well-known thresholding strategy, which is equivalent to a non-linear estimation of the signal [20, 21] - while, the approach of pre-filtering traditionally carries out noise elimination using linear filters. In this study, it is assumed that the (white) noise directly enters the output. Consequently, the focus is on achieving the best predictions of the deterministic component of the output, but in the wavelet space.

2.6 Application in nuclear reactors

In application domain, very few actual applications [85] of wavelet based identification have been reported, particularly that are related to nuclear reactors. A demonstration of black box modeling using wavelet basis function, utilizing data obtained during two transients of Monju fast breeder reactor, can be found in [34]. A method was proposed in this paper that pairs the analytical ability of the wavelets and computational power of a radial basis function network. It may be observed from the surveyed literature that the subject of system identification using wavelets is still young and at an exploratory stage of maturity. This is evident from the varied research interest and in rare attempts to apply the theoretical techniques in real life problems. It would be worth mentioning that lot of work is taking place in the domain of mathematics to support the research activity in the field of control system analysis.

A nuclear reactor is a complex system with non linearities, integrating effects and multiple time scale behaviour. Modeling of a nuclear reactor and its sub-systems for control design, failure analysis and simulation is an extremely important research activity. It is particularly relevant in the present scenario considering the global concern about inherent safety of the working nuclear plants. The work in this thesis has a great potential to be used for efficient modeling and control design of a nuclear reactor. A point kinetic model of a nuclear reactor is used as a plant and the method of wavelet based identification is tested. The derived LTI model is used for control design of a nuclear reactor. The concept of modeling systems with projections as states is not new [14, 29, 51]. For example Gilbert's method suggests that a system can be modeled using projections on eigen vectors. As the system gets completely decoupled, control inputs can be applied along the directions of eigen vectors. A similar approach is taken here by projecting the system on wavelet basis for applying control along each basis.

The proposed technique of system identification is also validated by modeling a complex process called the LZCS in a large Pressurized Heavy Water Reactor (PHWR). Besides providing continuous fine control of reactor power level and power distribution in the core of a large PHWR, the LZCS also compensates for routine reactivity perturbations due to on-power refueling and other minor changes in parameters such as temperature etc. The LZCS in a 540 MWe PHWR consists of 14 individually controllable compartments in the reactor, called Zone Control Compartments (ZCC) (Figure 2.7). Control of the reactor power level and the core power distribution is achieved by the LZCS through variation of light water levels in the ZCCs. Water serves as a neutron absorber and its inflow to ZCCs can be individually varied by the reactor regulating system by maneuvering the control valves in the water

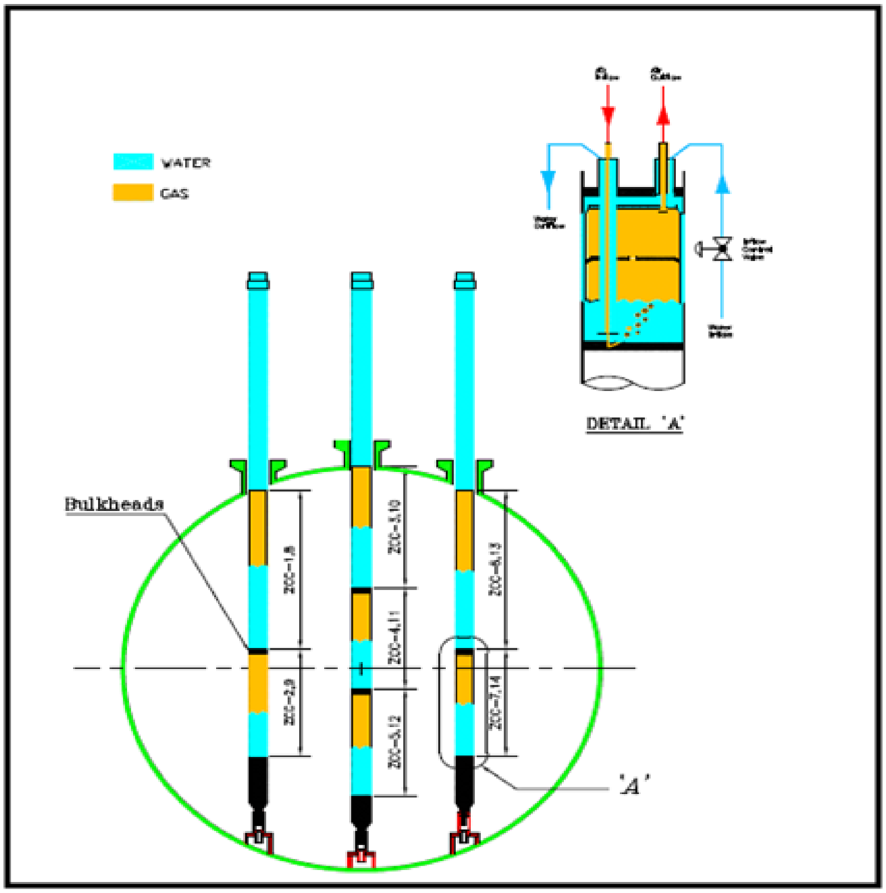

Figure 2.7: General assembly of the LZCS

inlet lines of each ZCC. However, a constant outflow of water takes place from each ZCC. Data from a full scale test setup [76] has been used in this work, for identification and validation of models. A full scale test set up of the LZCS of 540 MWe PHWR has been erected at Bhabha Atomic Research Centre. Data from this test setup have been used in this work for identification and validation of models.

2.7 Summary

A brief review of literature survey of related works on system identification using wavelet basis function has been presented in this chapter. Prior to embarking on the subject, admissibility of wavelets as basis for system modeling is established. As a system is identified in wavelet domain, consistent output prediction needs alternate projection algorithm (revisited in this chapter, based on Mallat's work) for reconstruction. In this context, the theory of shift invariant, undecimated wavelet transform is discussed and it is shown that decimated filterbanks, available freely in public domain, can even be used in undecimated case. The existing works in wavelet based modeling, in general, indicate advantages of using wavelets as basis over classical identification techniques for modeling non linearities and other complex system behaviours (widely modeled with time varying systems). In most of the works the formulation is reduced to that of LS minimization. Classical LS solution needs assumptions which are not practical for modeling strictly time varying systems and hence compromises model accuracy. Moreover, if the model is of insufficient order even for an LTI system, minimum point of the error surface varies with time and in general the optimal solution may never be reached. In specific, a very poor quality model of a nuclear reactor, a multiple time scale system, is obtained using classical ARX modeling technique. The model even fails to capture slow and fast modes with equal efficiency. Wavelet bases have compact support in both time and frequency and with their multi-resolution ability, there is a much higher chance to have a sufficient order model. As a consequence, approach to optimality is closer and faster (using LMS or RLS algorithm). This is the reason why with wavelet basis (and not with sine or sinc basis) it is possible to get consistent (true and unbiased and bounded) parameter estimates. In the hindsight, one may not capture individual poles but now there are singularities with a bundle of frequencies. Even with short support basis, existing wavelet models need the assumption of local time invariance which compromises model accuracy when dynamic behaviour of a system changes rapidly. Moreover, the existing techniques preclude non linear approximation. Motivation of the present work in relation to the existing works, pointing out major similarities and differences has been explained. Novelties of the proposed method vis-à-vis the existing modeling methods are clearly indicated.

Chapter III

Identification with Wavelet Basis

Identification of a dynamical system based on its input-output data can be grouped in two categories: parametric and nonparametric. Wavelet based parametric models are more popular for identification of linear and non linear systems and their formulation are considered in this chapter. In the following text, the symbol t in the parentheses indicates the time argument of a function and k in square bracket indicates the sample of the function at time $t = k$. A subscript used with a basis function denotes the basis function with time shift given by the subscript. For example, the basis function φ placed at k^{th} time index will be given by $\varphi_k = \varphi(t-k)$.

3.1 A discrete predictor model

For the purpose of identification, the TVARX model of a system can be expressed by partitioning input and output elements of the regression vector in terms of two impulse responses given by $h^{uy}(t,\tau)$ and $h^{yy}(t,\tau)$ respectively.

$$\hat{y}(t) = (h^{uy} \star u)(t) + (h^{yy} \star y)(t) \tag{3.1}$$

Let us choose two different finite length basis functions $\theta_i,\ i = 1, 2, \dots P$ and $\gamma_i,\ i = 1, 2, \dots Q$ for projecting outputs and inputs respectively in the approximation space. It may be noted here, that subscripts are used to denote discrete indices of the basis functions. Assuming a finite number of basis functions are needed for approximation of the time varying impulse responses following simplification is possible.

$$h^{uy}(t,\tau) = \sum_{i=1}^{Q} b_i(t)\,\gamma_i(\tau) \tag{3.2}$$

$$h^{yy}(t,\tau) = \sum_{i=1}^{P} a_i(t)\,\theta_i(\tau) \tag{3.3}$$

In general, a linear time varying function can be expressed in function space as a weighted sum of infinite number of integer indexed basis functions. The formulation does not require this restriction of using finite number of basis functions and for the time being, the generalization $i \in \mathbb{Z}$ (and $\mathbb{Z}$ shall be skipped from now on) is continued. One can see that output of a dynamical system can be estimated by linear filtering of projections (given by the convolutions) of past output and past inputs onto θ_i and γ_i respectively, in the following alternate description of the system derived from (3.1).

$$\hat{y}(t) = \sum_i a_i(t)(\theta_i \star y)(t) + \sum_i b_i(t)(\gamma_i \star u)(t) \tag{3.4}$$

This alternate description of the system shall be used throughout this thesis. The formulation allows the use of two different set of wavelet basis for output and input. This permits a better match with the respective wavelet basis leading to fewer coefficients of the modeled signal in the transform domain. Now, expressing $a_i(t) = \sum_l a_{il}\theta_i^*(t)$ and $b_i(t) = \sum_l b_{il}\theta_i^*(t)$ (3.4) can be written as

$$\hat{y}(t) = \sum_i \sum_l a_{il}\theta_l^*(t)(\theta_i \star y)(t) + \sum_i \sum_l b_{il}\theta_l^*(t)(\gamma_i \star u)(t) \tag{3.5}$$

in terms of constant parameters a_{il}'s and b_{il}'s. In general, θ_i and γ_i could be considered as shift invariant generating functions. $(\theta_i \star y)(t)$ and $(\gamma_i \star u)(t)$ are convolution of output and input with respective i^{th} basis and samples at $t = k$, $(\theta_i \star y)[k]$ and $(\gamma_i \star u)[k]$ are called generalized samples [82] of output and input respectively. It is important to note that for predictive modelling of a causal system, the discrete generalized samples $\sum_{\tau \in i=1,2,\ldots} \theta_i[\tau]\, u[k-\tau]$ and $\sum_{\tau \in i=1,2,\ldots} \gamma_i[\tau]\, y[k-\tau]$ are obtained using convolution (computed at $t = k$) of past values of input and output with respective basis functions. Hence without any loss of generality, a_{il}'s and b_{il}'s can be considered to be parameters of a discrete TVARX model [79] and in that case left hand side of (3.5) will turn out to be one-step-ahead prediction expressed as

$$\hat{y}[k+1] = \sum_i \sum_l a_{il}\theta_l^*[k](\theta_i \star y)[k] + \sum_i \sum_l b_{il}\theta_l^*[k](\gamma_i \star u)[k] \tag{3.6}$$

It may be noted here, that this formulation is similar to the input side modeling used in [24, 90, 91] the difference being adoptation of TVARX model instead of FIR and use of two different basis functions for approximating input and output. k^{th} generalized sample of the output projected [72] on i^{th} basis θ_i can be expressed in terms of inner product with the time reversed θ_i shifted to (placed at) $t = k$

$$(\theta_i \star y)[k] = \left\langle \theta_k^T, y \right\rangle, \; \theta_k^T(t) = \theta_k(-t)$$

Similarly, k^{th} generalized sample of the input is given by

$$(\gamma_i \star u)[k] = \left\langle \gamma_k^T, u \right\rangle, \; \gamma_k^T(t) = \gamma_k(-t)$$

For biorthogonal spline wavelets with symmetry $\theta_k(-t) = \theta_k(t)$ and $\gamma_k(-t) = \gamma_k(t)$. This implies

$$\hat{y}[k+1] = \sum_k \sum_l a_{kl} \langle \theta_k, y \rangle \theta_l^* [k] + \sum_k \sum_l b_{kl} \langle \gamma_k, u \rangle \theta_l^* [k] \qquad (3.7)$$

where a_{kl}'s and b_{kl}'s are time varying model parameters to be estimated. In case θ_i and γ_i are sinc functions, the convolutions in (3.6) pick-up time samples of input and output and the model reduces to classical ARX type model. If the output is approximated in multiresolution, the generalized basis functions are admissible scaling and wavelet functions, as has been explained in section 2.2.

3.1.1 Spline biorthogonal wavelets as basis

It may be noted here that projecting both input and output eventually onto the same basis in approximation space has two advantages.

1. If $\theta_i = \beta^m \star \beta_i^r$ and $\gamma_i = \beta^n \star \beta_i^r$ the approximation space is spanned by β_i^r. Hence, scalar addition of projections onto the analysing basis functions θ_i and γ_i in approximation space is admitted.

2. The correlation structure of filtered version of output and input (obtained by convolving with β^m and β^n respectively) is maintained in the approximation space.

Scalar addition of projections is a useful technique that could be used in consistent output prediction discussed later in the thesis. The structure of θ_i and γ_i suggests use of spline biorthogonal scaling and wavelet functions as basis for system identification because higher order spline functions are formed by successive convolution of the spline functions of order zero.

$$\beta^n(t) \;\; = \;\; \beta^{n-1} \star \beta^0(t) \qquad (3.8)$$

where $\beta^0(t)$ is the box function spline of order 0.

$$\beta^0(t) = \begin{cases} 1, & |t| < 1/2 \\ 0, & |t| \geq 1/2 \end{cases} \qquad (3.9)$$

The structure of β^n is same as that of θ_i and γ_i and hence projections on splines of different orders can be directly added. Moreover, box splines can be used as the underlying scaling function $\phi_i = \phi_k = \phi(t-k)$ for designing spline biorthogonal wavelets. As explained earlier, for approximating the measurement $y[k+1]$, projections onto the basis functions θ_i and γ_i could be weighted and directly added for every index k, if eventually both input and output are mapped into the space spanned by the same set of basis. Spline biorthogonal wavelets are popularly known as Reverse Biorthogonal (RBIO) Wavelets and are designated as *rbio* $p.\tilde{p}$ or *spline* $p.\tilde{p}$. p and $\tilde{p}$ are vanishing moments of reconstruction and anlysis wavelets. Associated analyzing scaling function is the spline of degree $p-1$.

3.1.2 Identification from pre-filtered input and output

From (3.8) it is easy to see that a spline of order n with integer time shift k, can be written as

$$\beta_k^n(t) = \beta^{n-1}(t) \star \beta_k^0(t) \tag{3.10}$$

Hence projecting output and input onto $\beta_k^m = \beta^{m-1} \star \beta_k^0$ and $\beta_k^n = \beta^{n-1} \star \beta_k^0$ respectively is equivalent to first pre-filtering output with β^{m-1} and input with β^{n-1} before projecting them in the approximation space spanned by β_k^0.

Box splines can be used as the underlying scaling function for designing spline biorthogonal wavelets. It may be noted that since $\psi(\frac{t}{2})$ is a linear combination of box splines $\phi(t-k)$, spline biorthogonal wavelets are also compactly supported polynomial splines of same degree as that of $\phi(t)$. Thus in (3.6), the approximation is no longer based on input and output but on the basis of low pass filtered version of input and output. In general input and output are vectors belonging to different spaces. Hence a mapping as such is necessary for linearly transforming a vector from input or output space to the approximation space where the parameters a_{kl} and b_{kl} are estimated.

Primary objective of pre-filtering however, is to decorrelate samples such that direct addition of projections for local fit can be considered as a possible solution. Moreover, thresholding decorrelated generalized samples obtained by pre-filtering makes the identified model insensitive to noise.

3.2 Modeling with wavelets

It is often useful to formulate the problem in terms of matrix-vector formulation using wavelet projections. In general, it follows from (3.7) that for a linear Single-Input-Single-Output (SISO) system, the estimate of the wavelet transform of one-step-ahead output of the system y_s, can be written as a difference equation in terms of wavelet projections of the time limited input-output data set

$$\left(\hat{W y_s}\right)[k] = \Theta_k w^y[k] + \Gamma_k w^u[k] \tag{3.11}$$

with

$$\Theta_k = \begin{bmatrix} d^y_{11} \\ \vdots \\ d^y_{1K} \\ \cdots \\ \vdots \\ \cdots \\ d^y_{J1} \\ \vdots \\ d^y_{JK} \\ \cdots \\ c^y_{J1} \\ \vdots \\ c^y_{JK} \end{bmatrix}', \quad \Gamma_k = \begin{bmatrix} d^u_{11} \\ \vdots \\ d^u_{1K} \\ \cdots \\ \vdots \\ \cdots \\ d^u_{J1} \\ \vdots \\ d^u_{JK} \\ \cdots \\ c^u_{J1} \\ \vdots \\ c^u_{JK} \end{bmatrix}'$$

$$w^y[k] = \begin{bmatrix} \left(\theta^h_{11} \star y\right)[k] \\ \vdots \\ \left(\theta^h_{1K} \star y\right)[k] \\ \cdots \\ \vdots \\ \cdots \\ \left(\theta^h_{J1} \star y\right)[k] \\ \vdots \\ \left(\theta^h_{JK} \star y\right)[k] \\ \cdots \\ \left(\theta^l_{J1} \star y\right)[k] \\ \vdots \\ \left(\theta^l_{JK} \star y\right)[k] \end{bmatrix}, \quad w^u[k] = \begin{bmatrix} \left(\gamma^h_{11} \star u\right)[k] \\ \vdots \\ \left(\gamma^h_{1K} \star u\right)[k] \\ \cdots \\ \vdots \\ \cdots \\ \left(\gamma^h_{J1} \star u\right)[k] \\ \vdots \\ \left(\gamma^h_{JK} \star u\right)[k] \\ \cdots \\ \left(\gamma^l_{J1} \star u\right)[k] \\ \vdots \\ \left(\gamma^l_{JK} \star u\right)[k] \end{bmatrix}$$

It is assumed that each of the sets $\left\{\theta^h_{jk}\right\}$ and $\left\{\gamma^h_{jk}\right\}$, $k = 1, 2, \ldots K$ independently span W_j and similarly each of the sets $\left\{\theta^l_{Jk}\right\}$ and $\left\{\gamma^l_{Jk}\right\}$, $k = 1, 2, \ldots K$ independently span V_J where K is the time limit. Assuming that the system is identified in subspaces W_j, $\left(\hat{W y_s}\right)[k]$ is a scalar. However, the model in general admits contribution from the basis functions at all scales. For the sake of brevity subscript j in the formulation is omitted. A similar vector-matrix formulation using orthogonal wavelet basis is given in [65].

To start with, let us assume here that $w^y[k]$ and $w^u[k]$ are obtained by applying a shift invariant wavelet transform e.g. UDWT suggested by Mallat and Zhang [48]. $w^y[k]$ and $w^u[k]$ can be considered as states of the identified system. To distinguish $w^y[k]$ and $w^u[k]$ from states of a system $x[k]$ let us call the former wavelet states. Collecting wavelet states for all k,

$$\begin{bmatrix} \vdots \\ \left(\hat{W}y_s\right)[k-1] \\ \left(\hat{W}y_s\right)[k] \\ \left(\hat{W}y_s\right)[k+1] \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \cdots \\ \cdots & \Theta_{k-1} & 0 & 0 & \cdots \\ \cdots & 0 & \Theta_k & 0 & \cdots \\ \cdots & 0 & 0 & \Theta_{k+1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ w^y[k-1] \\ w^y[k] \\ w^y[k+1] \\ \vdots \end{bmatrix} +$$

$$\begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \cdots \\ \cdots & \Gamma_{k-1} & 0 & 0 & \cdots \\ \cdots & 0 & \Gamma_k & 0 & \cdots \\ \cdots & 0 & 0 & \Gamma_{k+1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ w^u[k-1] \\ w^u[k] \\ w^u[k+1] \\ \vdots \end{bmatrix}$$

or in shorter form

$$\hat{W}y_s = \hat{w}^{y_s} = \overline{\Theta}w^y + \overline{\Gamma}w^u \tag{3.12}$$

where a bar on Θ and Γ indicate that they are matrices. Extending the SISO model for a Multi-Input-Multi-Output (MIMO) system, having p inputs and q outputs, following can be written.

$$\begin{bmatrix} \hat{w}^{y_{s1}}[k] \\ \vdots \\ \hat{w}^{y_{sq}}[k] \end{bmatrix} = \begin{bmatrix} (\Theta_{11})_k & (\Theta_{12})_k & \cdots & (\Theta_{1q})_k \\ \vdots & \vdots & \ddots & \vdots \\ (\Theta_{q1})_k & (\Theta_{q2})_k & \cdots & (\Theta_{qq})_k \end{bmatrix} \begin{bmatrix} w^{y_1}[k] \\ \vdots \\ w^{y_q}[k] \end{bmatrix} +$$

$$\begin{bmatrix} (\Gamma_{11})_k & (\Gamma_{12})_k & \cdots & (\Gamma_{1p})_k \\ \vdots & \vdots & \ddots & \vdots \\ (\Gamma_{q1})_k & (\Gamma_{q2})_k & \cdots & (\Gamma_{qp})_k \end{bmatrix} \begin{bmatrix} w^{u_1}[k] \\ \vdots \\ w^{u_p}[k] \end{bmatrix} \tag{3.13}$$

or in shorter form,

$$\left(\overline{\hat{W}y_s}\right)[k] = \overline{\Theta}_k\overline{w}^y[k] + \overline{\Gamma}_k\overline{w}^u[k] \tag{3.14}$$

Here bars on $w^y[k]$ and $w^u[k]$ indicate that the vectors are formed by concatenating wavelet coefficients corresponding to multiple inputs and outputs. $\overline{\Theta}_k$ has q rows and $(J+1)Kq$ columns and matrix $\overline{\Gamma}_k$ has q rows and $(J+1)Kp$ columns. $\overline{w}^y[k]$ and $\overline{w}^u[k]$ are column vectors having $(J+1)Kq$

and $(J+1)\,Kp$ rows respectively. It can be observed in the above formulation that there is a $J+1$ fold increase in the dimensionality when generalized samples are used. In the applications of multi-rate control, increase in sampling rate is used to ensure that all the poles of the system can be assigned. For example, in Periodic Output Feedback (POF) control input is sampled at a higher rate than the rate at which output is sampled. Minimum rate at which input could be sampled is decided by the controllability index of the system [11]. Similarly, control of a system identified with wavelet basis functions also could employ higher rate of sampling which justifies increase in dimensionality. Here, a few comments are necessary to clarify the formulation.

1. It may be noted that the number of samples are increased to $(J+1)\,K$ in case UDWT is used which is shift invariant. In case DWT (decimated filter bank implementation) is used, the dimensionality would still be K as sub-sampling is employed.

2. Dimensionality is drastically reduced in case biorthogonal spline wavelets, having compact support are employed for modeling. As the support of the basis functions is compact, a basis shifted far away from k would not contribute to $(Wy_s)\,[k]$. Hence, $\overline{\Theta}_k$ and $\overline{\Gamma}_k$ would be sparse.

3. Very often models of many large systems naturally possess interacting dynamic modes giving rise to widely separated clusters of eigen values. There are a plenty of such examples in nuclear engineering and econometrics which are modeled with two or three time scales. The ill conditioning problem in designing of POF and FOS controllers for two-time-scale systems is solved by transforming the system in block triangular form so that the clusters of eigen values are visible [77]. It is argued here that multiple time scale systems are ideal candidates for modeling with wavelets. As wavelets tend to naturally cluster the eigen values, parameter J is reduced to number of clusters.

4. Non linear approximation by thresholding wavelet coefficients also limit number of wavelet basis functions used for modeling. Hence in practice, a much lower order problem is solved.

3.3 State space model in wavelet domain

Let us consider a discrete, linear system S_d of order N, given by one-step-ahead state-space model

$$S_d: \quad x\,[k+1] = A_k x\,[k] + B_k u\,[k]\,, \forall k \tag{3.15}$$

$x[k]$ and $x[k+1]$ are the state vectors respectively at time $t = k$ and $t = k+1$ where x is a sequence in $\Re^n$. It would be interesting to find if the state equation can be expressed in terms of wavelet coefficients. Let us now examine if it is possible to write

$$w^x[k+1] = A_k^w w^x[k] + B_k^w w^u[k], \forall k \tag{3.16}$$

Reckoning that x and for this formulation w are sequences in $\Re^n$, one can write

$$w^x[k+1] = w_s^x[k] \tag{3.17}$$

$$x[k+1] = x_s[k] \tag{3.18}$$

where $w_s^x[k]$ is the k^{th} element of the sequence w^x, shifted in future by one time step. To get back the sequence of states in time, more precisely the relation given by (3.15), an implementation of inverse WT, is designated as an operator W^*. Substituting (3.17) in (3.16), collecting instances at all time

$$w_s^x = A^w w^x + B^w w^u \tag{3.19}$$

Taking inverse WT of both sides of (3.19)

$$\Rightarrow W^* W x_s = W^* A^w W x + W^* B^w W u \tag{3.20}$$

A^w and B^w matrices are formed by collecting A_k^w and B_k^w matrices respectively for all k. (3.19) is of the same form as (3.12) derived for SISO case (refer (3.12)) . Similarly A and B matrices can be formed by collecting A_k and B_k matrices respectively for all k. It can be seen that (3.15) is obtained from (3.20) if $W^*Wx_s = x_s$, $W^*A^wWx = Ax$ and $W^*B^wWu = Bu$. Hence under certain restrictive conditions one can work with a prediction model in wavelet domain as well. This has a very far reaching implication that now instead of applying the control law on the block of data, it could be applied at every instant k. In extension, a prediction-correction type algorithm similar to Kalman Filter [36, 37] can be designed for controlling a process in real time. A multiscale Kalman filter algorithm is investigated in [15] for optimal estimation of model parameters.

To complete the one-step-ahead modeling, output equation can be written in a similar way as the state equation (5.1)

$$y[k] = T_k x[k] \tag{3.21}$$

where,

$$T_k = [T_{1k}\, T_{2k}\, \cdots T_{Nk}] .$$

As earlier, collecting samples at all k,

$$y = Tx \tag{3.22}$$

Pre-multiplying (3.22) by W, we would like to write it in terms of wavelet projections as below.

$$w^y = WTx = T^w x \tag{3.23}$$

Matrix T has K rows and KN columns.

$$T = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \cdots \\ \cdots & T_{k-1} & 0 & 0 & \cdots \\ \cdots & 0 & T_k & 0 & \cdots \\ \cdots & 0 & 0 & T_{k+1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

In case of single output, wavelet operator matrix W would have KN rows and K columns. One can see T_k^w in (5.2) is a submatrix of $T^w = WT$ (having KN rows and KN columns) and needs to be of rank N.

3.4 Parameterization in wavelet domain

The state equation in terms of wavelet coefficients

$$w^x [k+1] = A_k^w w^x [k] + B_k^w w^u [k] \tag{3.24}$$

can be shown equivalent to state equation in time

$$x [k+1] = A_k x [k] + B_k u [k] \tag{3.25}$$

if

1. W is invertible and shift invariant *i.e.*

$$W^* w_s^x = W^* W x_s = x_s \Rightarrow W^* W = I \tag{3.26}$$

2. Parameters in measurement and transform domain are related as

$$W^* A^w W = A,\ W^* B^w W = B \tag{3.27}$$

For minimum memory consistent estimation of parameters proposed in this work, A_k^w and B_k^w and hence A^w and B^w are diagonal matrices. Hence (3.27) implies that columns of invertible transform W are eigen vectors of A and B and diagonal elements of A^w and B^w are corresponding eigen values. Let,

$$A^w = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \alpha_{\mathbb{N}} \end{bmatrix}$$

where, $\alpha_1, \alpha_2, \cdots \alpha_{\mathbb{N}}$ are non zero distinct eigen valus and $\mathbb{N} = NK$.

Any linear transformation W in single resolution, satisfying

$$A^w W = WA$$

can in general be given by [26]

$$W = VU$$

One possibility, given $U = Q^{-1}, Q = \begin{bmatrix} B, & AB, \cdots A^{\mathbb{N}-1}B \end{bmatrix}$ (implying S_d controllable) with $Q\widetilde{A} = AQ$ where,

$$\widetilde{A} = \begin{bmatrix} 0 & 0 & \cdots & -a_{\mathbb{N}} \\ 1 & 0 & \cdots & -a_{\mathbb{N}-1} \\ 0 & 1 & \cdots & -a_{\mathbb{N}-2} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -a_1 \end{bmatrix}$$

is of rank $\mathbb{N}$. Here $(a_1 \cdots a_{\mathbb{N}})$ are the coefficients of K characteristic equations representing the time varying system. The transformation V satisfying $VA^w = V\widetilde{A}$ then is given by

$$V = \begin{bmatrix} -\left(\frac{a_{\mathbb{N}}}{\alpha_1}\right) k_1 & -\left(\frac{a_{\mathbb{N}}}{\alpha_2}\right) k_2 & \cdots & -\left(\frac{a_{\mathbb{N}}}{\alpha_N}\right) k_{\mathbb{N}} \\ -\left(\frac{a_{\mathbb{N}}}{\alpha_1} + a_{\mathbb{N}-1}\right) k_1 & -\left(\frac{a_{\mathbb{N}}}{\alpha_2} + a_{\mathbb{N}-1}\right) k_2 & \cdots & -\left(\frac{a_{\mathbb{N}}}{\alpha_N} + a_{\mathbb{N}-1}\right) k_{\mathbb{N}} \\ \vdots & \vdots & \vdots & \vdots \\ k_1 & k_2 & \cdots & k_{\mathbb{N}} \end{bmatrix}$$

V is of rank $\mathbb{N}$ if eigen values are non zero and distinct. Note that if eigen values are non zero and distinct no coefficient of characteristic equations $(a_1 \cdots a_{\mathbb{N}})$ can be zero. Since both U and V are of rank $\mathbb{N}$, W is also of rank $\mathbb{N}$. Now the question is - "is it possible to find a multi-resolution version of U or V ?" Design of wavelets through this route is an open problem. If the answer to the question posed above is "yes" then it can be concluded that a controllable system can be modeled with wavelet states provided eigen values of A^w and B^w are non zero and distinct. A method that diagonalizes A is

MSPCA where system states are transformed to wavelet states followed by principal component analysis at each scale.

Let us however start with a wavelet filterbank W and impose restrictions on eigen values of A^w and B^w while identifying the model using the method of consistent prediction. Although the wavelets are not known to be exact eigen functions or principal components of any operator, they are approximate eigen functions of a large variety of operators [19, 87] and hence the imposed restrictions are justified. The output is synthesized by alternate projection that fits the best model (across the scales and time) in LS sense to a sparse wavelet representation of data.

3.5 Summary

This chapter starts with formulating linear TVARX predictor model in continuous and discrete domain, using generalized basis functions. Advantages of using spline biorthogonal wavelets as generalized basis functions are explained. Modeling with spline biorthogonal wavelets can be viewed as identification with pre-filtered input and output. Further, projections on biorthogonal spline wavelets can be weighted and directly added for estimating parameters using consistent output prediction. In specific, SISO as well as MIMO linear models for identification in wavelet domain are derived and wavelet states are defined. Issues related to reduction of dimensionality are discussed. It is shown that state space model with wavelet states can be derived if the transform is invertible and shift invariant. Further, under certain restrictive assumptions, a controllable system can be modeled with wavelet states. It is argued that the restrictive assumptions are justified because wavelets are approximate eigen functions and the restrictions can be imposed in the framework of consistent output prediction.

Chapter IV

Parameter Estimation by Consistent Output prediction

This chapter discusses the solution method of parameter estimation problem using consistent output prediction. First a general solution is derived for a time varying model without considering noise. An approximate LTI model is derived from the general solution under piecewise linear assumption. The LTI model is characterized considering output noise in equation error framework. An algorithmic solution (alternate projection) is proposed for model identification and output reconstruction.

4.1 General considerations

Denote the shifted version of measurement $y(t+T)$ as $y_s(t)$, where T is the sampling time. Discrete measurement $y\left[k+1\right]$ can be expressed in terms of projections of shifted version of the measurement $y_s(t)$ onto analyzing wavelet θ_l

$$y\left[k+1\right] \;=\; y_s\left[k\right] = \sum_l \langle \theta_l,\, y_s\rangle\, \theta_l^*\left[k\right] \tag{4.1}$$

where θ_l^* denote the reconstruction wavelet. Minimum error solution of the estimation problem in LS sense is obtained by minimizing the error functional,

$$\chi \;=\; \sum_k \left(y\left[k+1\right] - \hat{y}\left[k+1\right]\right)^2 \tag{4.2}$$

Let us rather work with an equivalent cost function in terms of wavelet projections

$$\chi_1 = \sum_k \left(e\left[k\right]\right)^2 \tag{4.3}$$

where,

$$e\left[k\right] = \sum_l \langle \theta_l,\, y_s\rangle\, \theta_l^*\left[k\right] - \sum_l \sum_k a_{kl}\, \langle \theta_k,\, y\rangle\, \theta_l^*\left[k\right] + \sum_l \sum_k b_{kl}\, \langle \gamma_k,\, u\rangle\, \theta_l^*\left[k\right]$$

As suggested earlier, in this work parameter estimation is attempted completely in wavelet domain. This is in deviation from traditional approaches of identification where model parameters are estimated by minimizing error functional in measurement space (in time). If orthogonal wavelets are used, energy of the signal is equal by Parseval's relation in both time and wavelet domain and hence error minimization in LS sense in either domain would give the same solution. The proposed method of estimation of model parameters based on the idea of consistent output prediction, as formally defined later, however does not necessarily need the assumption of strict orthogonality.

A brief comparison with the LS approach in the original measurement space (or the space of filtered measurements) is in place here. The motivation underlying the use of LS approach in the measurement space is that every data point in that space contains both the signal and noise, and therefore predictions should not completely explain every point but be as "close" as possible to the data points. Modeling based on wavelet projections as proposed here provides two advantages in this context.

1. The signal and noise are separated to a large extent, if not completely. Consequently, a large number of projections (coefficients) can be treated as zero. Usually for penalized local error minimization a suitable thresholding is applied to determine the subset of projections that correspond to noise (and setting them to zero), while the remaining subset of projections are adjusted so that they only contain the signal. Subsequently, the parameters are determined such that the predictions exactly match these adjusted non-zero projections, which are much fewer in number than the original number of samples.

2. The projections themselves are relatively much less correlated compared to the original measurements. This is attributed to the fact that the correlation functions in wavelet domain are known to decay faster than the correlation functions of the original signal in time [7, 75]. Thus, higher-order systems in the measurement space reduce to lower-order models in the projection space. This advantage is lost when modeling is carried out on the inverse transformed data, *i.e.*, the filtered data since we are back in the space of Shannon basis functions.

The second point discussed above is satisfied for several other basis and/or operator spaces as well. In fact this has been exploited in a few works on identification using orthonormal functions [60, 61]. The novelty in this work stems from the first point, which sets the tone for consistent output prediction. The following definitions are inspired by the notion of consistent estimate proposed by [17] and are necessary to begin the discussion on the proposed solution.

Definition 4.1: A signal representation in transform domain is an ordered collection of significant signal values (obtained by a nonlinear operation like maxima detection or thresholding).

Consequence of the above definition is that every element in the representation is a pair formed by the signal value associated with an index changing in ascending order. For example, thresholded wavelet coefficients of a measurement (of a noisy signal) form a representation of the signal in wavelet domain because thresholding removes noise coefficients. Hereafter, by representation we allude to wavelet representation of a signal. In this context, a consistent prediction is defined as follows.

Definition 4.2: A consistent prediction is such that the output signal and its prediction in measurement space have the same signal representation in the wavelet domain.

A consistent estimate of a signal from a measurement is one whose transform is identical to the transform of the signal component of the measurement. The idea is illustrated in Figure 4.1 with an application of signal denoising. First consider a synthetic signal marked as original signal in the top pane of the figure. A noisy version of the synthetic original signal is obtained by adding a coloured noise of signal-to-noise ratio (SNR) 30 dB. The denoised reconstructed signal shown in the third pane of the figure is obtained by thresholding wavelet projections of the noisy signal and applying iterative alternate projection algorithm to obtain consistent estimate of the original signal. Error plot in the bottom pane of the figure clearly shows strong convergence to the consistent estimate. The reconstruction here is based on significant projections (projections above a threshold) of the noisy signal.

Wavelet projections of the noisy signal and those of its reconstruction (indicated by circles and solid lines respectively) are shown in Figure 4.2 in frequency subbands $W2$ through $W4$. Frequency band $W1$ is found to be predominantly noise and is rejected in reconstruction. $W5$ shows the output of the low pass filter at the highest scale of the Haar filter bank. It is used without thresholding in this exercise for implementing undecimated wavelet decomposition. It is not necessary that all the significant projections of the reconstructed signal (obtained by applying inverse wavelet transform once) would match with those of the original signal because the reconstruction has been made from a sparse representation in wavelet domain. Ideally a reconstruction would qualify as a consistent estimate if it had the same wavelet representation. The same is forced by equating values of original signal and its reconstruction at all significant projections in iterative alternate projection algorithm described in details later.

Definition 4.3: Reconstruction set $\Re$, associated with the representation $R_m w^x$ is a set consisting of all the sequences $w^{x^c} \in l_2(\mathbb{Z})$, called consistent

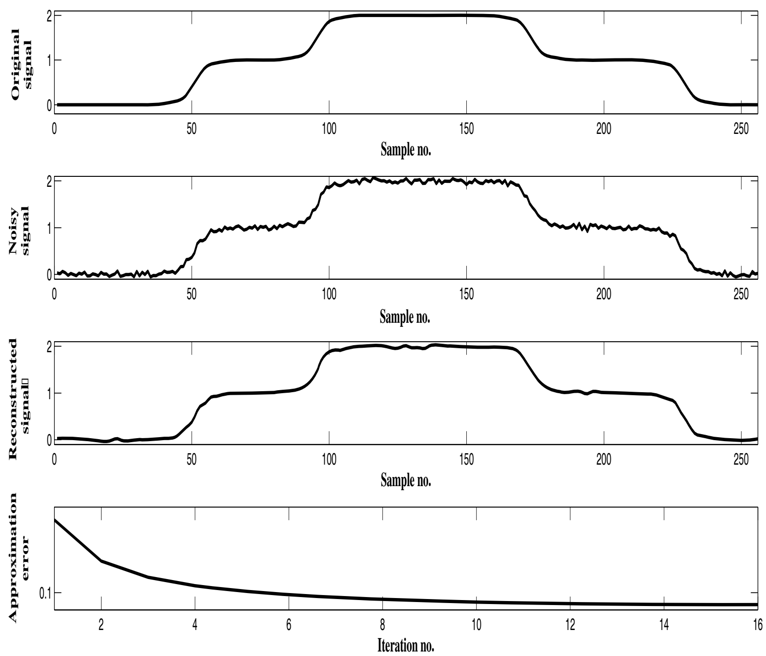

Figure 4.1: Consistent estimation of a function in wavelet projections

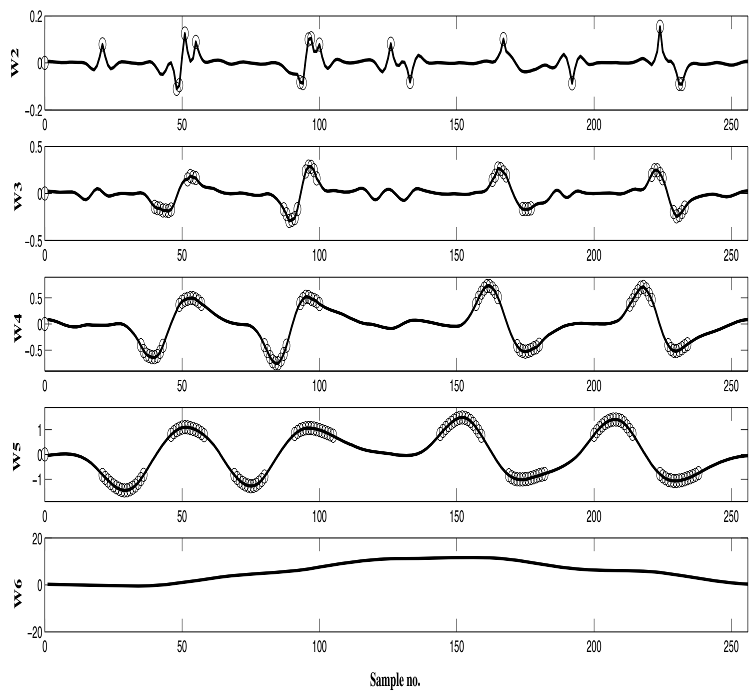

Figure 4.2: Consistent estimation of a function in wavelet projections

estimates which have the same representation $R_m w^x$ *i.e.*

$$\Re = \left\{ w^{x^c} \in l_2(z) : \ R_m w^{x^c} = R_m w^x \right\} \tag{4.4}$$

Reconstruction set is the set of signals which have the same wavelet representation. The definition is inspired by the notion of consistent estimate proposed by [17]. An estimate in the reconstruction set of the representation is obtained by the projection operator $P_{\Im} = WoW^*$. The method of consistent output estimate uses local error minimization in wavelet domain. The technique forces error at every significant projection locally to go to zero. The operator $P_{\mathcal{L}}$ extracts the values of projections at the representative points, sets the projections equal to the projection of measurements and projects in $\mathcal{L}$ space using convex constraints. Rest of the projections due to noise are set equal to zero. Repeated iteration of alternate projection operator, $P^{(n)}$, $P = P_{\Im} o P_{\mathcal{L}}$ reduces the normal distances between two spaces $\Im$ and $\mathcal{L}$ and the alternate projections on $\Im$ and $\mathcal{L}$ strongly converge to the orthogonal projection on $\Lambda = \Im \cap \mathcal{L}$. The iteration actually converges to the member of the reconstruction set that minimizes error norm giving the solution of the optimization problem.

As the generalized samples of the transformed system in wavelet domain are likely to have less memory compared to the time samples of the original system in measurement space, the solution of consistent output prediction using local error minimization will work better with generalized samples. The memory however is not lost. In fact it remains embedded in the properly designed wavelet basis [55]. Moreover, a residual dependency structure still remains between magnitudes of wavelet coefficients both across the scale and at neighbouring temporal locations. In this context a minimum memory system can be defined as follows.

Definition 4.4: A system for which output estimate at any instant is completely decided by the weighted sum of the measurements of input and output at that instant is called a minimum memory system.

It may be noted that the output estimate could be a prediction. Extending the arguments, the following conjecture may be stated.

Conjecture 4.1: Any dynamical system can be transformed into a minimum memory system by proper design of representative wavelet basis.

As enumerated in the two advantages above, representation of a noisy measurement in terms of its wavelet projections strips the noise components from signal (so that noise components can be removed by amplitude thresholding) as well as concentrates energy in fewer projections. As a result, a projection becomes less related to the other. Hence, proposition 4.1 follows.

Proposition 4.1: The consistent output prediction, obtained by local error minimization in wavelet domain, approaches the true solution as the transformed system (given by wavelet based representation) tends to become a minimum memory system.

Proof of proposition 4.1 can be constructed by arguing that the true solution for a minimum memory system shall have no dependence on the measurements at any time instant other than the ones at present time instant. If conjecture 4.1 is true *i.e.*, if it is possible to decorrelate wavelet coefficients of the signal to the extent that the estimated prediction at any time instant is solely dependent on the input and output measurement of a single (present) time instant, consistent prediction by local error minimization in wavelet domain shall give the true solution.

Assuming that the system remains time varying in transform domain it is wiser to minimize local error instead of global error (or error over a finitely long time interval). As local error minimization is admissible by virtue of Proposition 4.1, a near-true solution can be obtained here, using proposed method of consistent output prediction because wavelets are known to be approximate eigen functions of a large variety of operators. Wavelet coefficients may not be fully decorrelated and as a result, amplitude thresholding may fail to cut-off noise completely and the near-true solution is reached by alternate projection. Nevertheless, given the wavelet basis and the chosen thresholds for the modeling exercise, the solution is optimum in LS sense. Although not exactly true, the solution has certain useful advantages. The method identifies a complex system (for example with an integrator) with high order dynamics by a time varying model with fewer parameters.

4.2 Proposed LTV solution using consistent output prediction

Consider identifying a subsystem in wavelet subspace W_j with $t = k$ denoting the time index of a basis function θ_{jk} and γ_{jk} in W_j. For the sake of bravity j is skipped in the index of the basis. Such subsystems can be interconnected to get the original system. Theoretical background of the proposed solution is built with the assumption that the projections in transform domain are derived by an undecimated wavelet transform (for example, using the shift invariant algorithm suggested by [48]) such that l and k represent indices of the original measurement sampling grid. However, the theory will also hold good under mild assumptions, in case decimated wavelet transform is used. We need to estimate a_{kl}, b_{kl}, constant valued parameter of the wavelet model. As scalar summation is allowed in the space spanned by the wavelet basis, error in time $e[k]$ can be written in terms of errors in wavelet domain, $w^e[l]$

$$e[k] \;=\; \sum_l \left[\langle \theta_l,\, y_s \rangle - \sum_k a_{kl} \langle \theta_k,\, y \rangle - \sum_k b_{kl} \langle \gamma_k,\, u \rangle \right] \theta_l^* [k]$$

$$= \sum_l w^e[l]\theta_l^*[k] \tag{4.5}$$

where $w^e[l] = \langle \theta_l, y_s \rangle - \sum_k a_{kl} \langle \theta_k, y \rangle - \sum_k b_{kl} \langle \gamma_k, y \rangle$.

The parameters are now estimated by setting the partial derivatives of error to zero. For instance,

$$\frac{\partial \chi_1}{\partial a_{kl}} = 2 \sum_k e[k] \langle \theta_k, y \rangle \theta_l^*[k] = 0 \tag{4.6}$$

Assuming that there are K input-output pairs of data, both input and output being sampled exactly on the same time grid $k, l = 1, 2, \ldots, K$, the above problem is underdetermined because one needs to compute $2K^2$ coefficients from $2K$ equations. However, it is possible to search for the best set of coefficients (less than K) that would minimize $w^e[l]$ globally. The underlying assumption is that the system is time invariant in transform domain. The assumption is not necessarily true and hence the optimum global solution may not be a good solution.

A numerical example

A hypothetical numerical example can be considered to show the advantage of time varying modeling which is in general under determined in original measurement space. Let us take the case of modeling a 5^{th} order time varying system from 1000 input-output data samples. Strictly speaking, solution in original measurement space needs estimation of 5000 parameters and is clearly under determined. On the contrary, the regressor may have a few significant wavelet projections needed for modeling.

It can be seen from (4.6) that an optimum solution can be obtained by either setting the error in time, $e[k] = 0$ (for significant projections) or by setting projection $\langle \theta_k, y \rangle = 0$ (for noise projections). From (4.5), $e[k] = 0$ implies $w^e[l] = 0$ for all $l = k$ with the assumption $\theta_l^\star$ span the output error space. $w^e[l]$ can be set equal to zero by a wavelet basis locally linear time invariant model, estimating projections of the shifted measurement as

$$\hat{w}^{y_s}[l] = \langle \theta_l, y_s \rangle = \sum_k a_{kl} \langle \theta_k, y \rangle + \sum_k b_{kl} \langle \gamma_k, u \rangle \tag{4.7}$$

(4.7) is underdetermined and coefficients of the linear filters given by a_{kl} and b_{kl} can be estimated only with the assumption of local time invariance. A strictly time varying solution can however, be approached assuming the system is minimum memory (by omitting the summations over k in (4.7)). Validity of this assumption is discussed at length earlier. In summary the proposed method works as follows.$w^e[l]$ can be set equal to zero if coefficients of wavelet expansion of the shifted measurement at every time instant k, is set equal to the weighted sum of coefficients of wavelet expansion of output $\langle \theta_k, y \rangle$ and input $\langle \gamma_k, y \rangle$. Thresholding in wavelet domain is known to reduce noise and is a solution to the (penalized) least squares problem. Let λ_u and λ_y be two strictly positive values. In (penalized) LS minimization, only those wavelet coefficients of input and output are used which have modulus values more than λ_u and λ_y respectively. Let us define those as significant wavelet coefficients. Thresholding allows nonlinear estimation of the output and the method works as follows.

1. It sets the local error *i.e.* the error at the location of each significant wavelet coefficient of the shifted output to zero.

2. Rest of the projections are considered to be noise and are directly set equal to zero.

Reckoning that $w^e[l]$ for all $l = k$ is obtained by subtracting weighted scalar summation of wavelet coefficients at k^{th} instant from the wavelet coefficient of the shifted measurement at k^{th} instant (summation is omitted), l in the subscript of a and b can be dropped. Based on the above discussion solution of consistent output prediction can be written as

$$\begin{gathered} \langle \theta_k, y_s \rangle - a_k \langle \theta_k, y \rangle - b_k \langle \gamma_k, u \rangle = 0 \\ \forall k \ \in \{ I_u : |\langle \gamma_k, u \rangle| \geq \lambda_u \bigcap \ I_y : |\langle \theta_k, y \rangle| \geq \lambda_y \} \\ a_k = b_k = 0 \ \forall k \notin I_u \text{ and } \forall k \notin I_y \end{gathered} \tag{4.8}$$

If $Dim\ I_u \cap I_y = M$, the system is optimally identified by a_k and b_k, $\forall k \ \in \{I_u \bigcap I_y\}$, in $M - Dim$ subspace with $M \ll K$. At each k it is still required to find two parameters a_k and b_k from a single equation (4.8). An algorithmic solution for identification of an LTV model is derived in section 4.4.2. The derivation is motivated by Theorem 4.1 derived later in this chapter.

As the system is identified in wavelet domain, a method is needed to be devised to reconstruct the system output in time domain for model testing and cross validation. Note that the consistent output prediction is based on a subset of a sequence of projections. The reconstruction algorithm works by alternately projecting the output prediction on $\mathcal{L}$ and $\Im$ spaces. The output prediction is a wavelet transform of a function in time and belongs to $\Im$ if and only if reproducing kernel equations hold [47]. Projection operator on $\Im$, $P_\Im = WoW^*$ is invariant in $\Im$. For the algorithm to work, it is necessary that the projection operator, $P_\mathcal{L}$ takes the solution out of $\Im$ space by violating the kernel equations.

Application of the operator $P_{\mathcal{L}}$ followed by $P_{\Im}$ can be seen as a mapping from a space of lower dimension to a space of higher dimension and hence the reconstruction is not unique. $P_{\Im}$ projects an intermediate solution, in the affine space $\mathcal{L}$, back to time and then again the crude estimate of output prediction, in time, forth to transform domain. However, projection paths of $P_{\Im}$ on $\Im$ and $P_{\mathcal{L}}$ on $\mathcal{L}$ are different. The projections are orthogonal on respective spaces if wavelets are symmetric or anti-symmetric which is the case with spline biorthogonal wavelets. Minimum norm solution in $\Im \bigcap \mathcal{L}$ is approached iteratively by projecting the solution alternately on $\Im$ and $\mathcal{L}$ *i.e.* alternate projection converges to the member of the reconstruction set that minimizes the cost function in (4.2).

An important point to note here is that at indices k for which $\langle \theta_k, y \rangle = 0$, error in time $e[k]$ is not necessarily zero. At all other k, $e[k]$ is exactly equal to zero. The result shows close similarity to that obtained in compressed sensing [78] and indicates signal representation with fewer measurements at indices for which $e[k] = 0$. It may be noted that in the above derivation, no assumption of strict orthogonality has been made and hence biorthogonal spline basis of order higher than 0 are admissible. It is however, necessary to ensure stability of reconstruction which readily follows because coefficients in transform domain are bounded [90] from the assumption of Riesz basis (refer (2.9)).

The basic tenets of the proposed theory is validated experimentally as the results are documented in this and subsequent chapters. First an approximate LTI model shall be built in equation error framework considering noisy measurements.

4.3 Output prediction in presence of noise

Although the formulation of model estimator is in general time varying, it is useful to identify an LTI model in each subband when it is known a priori (from physics) that the process is approximately LTI (mildly non linear or time varying). One way to account for the un-modeled non linearity is to model it in a piecewise linear fashion by assuming contribution from input and output to be constant over each piecewise linear region. Based on the idea, a simplification is possible in the form

$$a_k = k_1 \alpha_k, \; b_k = k_2 \alpha_k, \tag{4.9}$$

where intuitively, for one step ahead prediction, k_1 and k_2 can be seen someway related to the output auto-correlation and input-output cross-correlation coefficients at lag one. Hence, the following theorem can be proved by substituting (4.9) in (4.8) considering the size of $I_u \bigcap I_y = M$.

Theorem 4.1: Assuming that the noise in the estimate is stationary, *iid* $N(0,\sigma^2)$ distributed, a_k and b_k are given by $a_k = k_1\alpha_j$, $b_k = k_2\alpha_j$ where k_1 and k_2 are two real valued constants independent of time, then the first order estimate of an LTI model parameters at scale j based on the consistent output estimate using local error minimization in wavelet domain is given by

$$\hat{\alpha}_j \;=\; \frac{1}{M}\sum_{k=I_u\bigcap I_y}\left[\frac{\langle\theta_k,\, y_s\rangle}{k_1\langle\theta_k,\, y\rangle + k_2\langle\gamma_k,\, u\rangle}\right] \tag{4.10}$$

4.3.1 Characterization of the LTI model in presence of noise

It can be assumed here, that in the identified time varying model, variations in the system parameters are only due to noise in the output. As discussed in the last subsection, for an LTI model, $a_k = k_1\alpha_j$ and $b_k = k_2\alpha_j$, k_1, k_2 and α_j remaining constant over time in the subband (scale) indexed by j. Let us assume that the output measurement is corrupted with stationary, *iid* $N(0,\sigma^2)$ distributed (Gaussian, white) noise. Let superscript s indicate signal component and superscript n indicate noise component in the output measurement. Then from (4.8), it can be seen that a parameter can be expressed as a sum of a deterministic and a random component.

$$\alpha_j = \overline{\alpha_j} + \nabla\alpha_j(t) \tag{4.11}$$

with

$$\overline{\alpha_j} = \frac{\langle\theta_k,\, y_s^s\rangle}{k_1\langle\theta_k,\, y^s\rangle + k_2\langle\gamma_k,\, u\rangle}$$

and

$$\nabla\alpha_j = \frac{\langle\theta_k,\, y_s^n\rangle}{k_1\langle\theta_k,\, y^s\rangle + k_2\langle\gamma_k,\, u\rangle}$$

$$\forall k \;\in \left\{I_u : |\langle\gamma_k,\, u\rangle| \geq \lambda_u \bigcap\; I_y : |\langle\theta_k,\, y\rangle| \geq \lambda_y\right\}$$

It may be noted that noise in the regressor given by $\langle\theta_k,\, y_s^n\rangle$ is considered to be removed by thresholding and hence the denominator of both the terms on the Right Hand Side (RHS) of (4.11) are deterministic. Under the assumption that signal and noise components are independent of each other, the uncertainty in the parameter, given by the second term on the RHS of (4.11) is also zero mean random because

$$E\left(\nabla\alpha_j\right) = \frac{E\left(\langle\theta_k,\, y_s^n\rangle\right)}{k_1\langle\theta_k,\, y^s\rangle + k_2\langle\gamma_k,\, u\rangle} = 0 \tag{4.12}$$

where E denotes expectation operator. The variance term of parameter error can be estimated as

$$\hat{P} = E\left((\nabla\alpha_j)^2\right) = \frac{E\left(\langle\theta_k,\, y_s^n\rangle^2\right)}{\left(k_1\langle\theta_k,\, y^s\rangle + k_2\langle\gamma_k,\, u\rangle\right)^2} = \frac{\sigma^2}{R^2}$$

$$where, \quad R = \left(k_1\langle\theta_k,\, y^s\rangle + k_2\langle\gamma_k,\, u\rangle\right)$$

Since, y^s and u are finite, R is also finite and decides the bound of parameter error.

$$Max\left(\hat{P}\right) = \frac{\sigma^2}{min\left(R^2\right)} = \frac{\sigma^2}{min\left((k_1\lambda_y)^2, (k_2\lambda_u)^2\right)}$$

It can be seen that the bound of parameter uncertainty can be reduced by increasing thresholds. Hence accuracy of the identified model depends on the level of threshold chosen by the designer. At the same time, higher threshold could possibly remove significant signal component thereby compromising usefulness of the identified model. A tradeoff in this regard is necessary to meet the design objective as well as quality of identification.

4.3.2 A discussion on the assumption of Gaussianity

The assumption that a measurement is corrupt with white Gaussian noise is possibly the most common assumption about the noise distribution because the assumption holds perfectly in large class of practical systems. Moreover the assumption makes mathematical tractability simpler since a Gaussian distribution can be completely characterized by its first two moments. For Gaussian white noise the largest observation is of size $O\left(\sqrt{log\,(K)}\right)$ and can be successfully removed for a large range of smooth signal by applying a wavelet transform, thresholding the wavelet coefficients and then inverting the transform. If the noise has mild form of non-Gaussianity with sufficiently many finite moments, level-dependent threshold (somewhat higher than the Gaussian case) works for a limited range of smoothness classes [31]. However, in certain applications such as in analog telephony, radar signal processing and laser radar imaging data exhibit strongly non-Gaussian noise distribution. A strong non-Gaussianity *e.g.* Cauchy distributed noise has no moments and is characterized by sudden very large deviations. In this case thresholding of linear wavelet transform coefficients fails to remove noise. Median-interpolating pyramid transform - a kind of nonlinear wavelet transform has been suggested for denoising measurements with such strong non-Gaussian noise [22]. Shrinkage rule for the class of strong non-Gaussian noise, specifying a prior distribution on wavelet coefficients has also been suggested based on wavelet based Bayesian approach[1].

4.4 *Signal reconstruction using consistent estimate*

The structure of (4.10) suggests that an iterative scheme can be formulated to find the LTV solution. Parameter updating happens twice in each iteration. In each updating either input or output (not both) is used for prediction such that solution of (4.8) reduces to estimation of only one parameter. For intutive understanding the iterative algorithm can be described in the following way. If we start modeling the system as an AR process, next in the same iteration the residue is modeled as an FIR process. The algorithm continues to iterate seeking the minimum norm solution by alternately projecting it on $\mathcal{L}$ and $\Im$ [55]. It is important that operator $P_{\mathcal{L}}$ takes the solution out of $\Im$ space. Since at every iteration, the solution is estimated once using the projections of the output followed by using those from the input, the operator $P_{\mathcal{L}}$ takes the solution out of space $\Im$, even when decimated Discrete Wavelet Transform (DWT) with orthogonal wavelet basis is used.

Before we embark on the method of alternate projection let us formally define different representations in wavelet domain.

Let $L_0 w^x$ be the set of arguments of a signal w^x such that the operator $P_{\mathcal{E}}$, extracts values of w^x at representative points

$$P_{\mathcal{E}} w^x = (L_0 w^x, V_0 w^x) \tag{4.13}$$

where

$$V_0 w^x = w^x[k], \ \forall k \in L_0 w^x$$

The difference between $P_{\mathcal{L}}$ and $P_{\mathcal{E}}$ is that $P_{\mathcal{L}}$ also includes projection (using convex constraints) of $P_{\mathcal{E}} w^x$ in $\mathcal{L}$ space. Using the definition of $P_{\mathcal{E}}$, a representation R_m is defined as

$$R_m w^x = (P_{\mathcal{E}}(W_j x)_{1\le j\le J}, V_0(W_j x)_{1\le j\le J}, V_0(V_{J+1} x)) \tag{4.14}$$

On an irregular grid, arguments of four wavelet representations are defined as below.

1. Wavelet Representation with Missing Samples (WMSR) (obtained by amplitude thresholding)

$$L_{0ms}(w^x) = \{k : |w^x[k]| > \lambda\} \tag{4.15}$$

2. Wavelet Modulus Maxima Representation (WMMR)

$$L_{0mm}(w^x) = \{k : |w^x[k+1]| < |w^x[k]|, \ |w^x[k-1]| \le |w^x[k]|\} \bigcap \{k : |w^x[k]| > \lambda\} \tag{4.16}$$

3. Wavelet Extrema Point Representation (WEPR)

$$L_{oe}(w^x) = \{k : w^x[k+1] < w^x[k],\ w^x[k-1] \leq w[k]\}$$

$$\bigcap \{k : w^x[k+1] > w^x[k],\ w^x[k-1] \geq w^x[k]\} \tag{4.17}$$

4. Wavelet Maximum Curvature Point Representation (WMCPR)

$$L_{0mc}(w^x) = \{k : C(w^x[k+1]) < C(w^x[k]), C(w^x[k-1]) \leq C(w^x[k])\}$$

$$\bigcap \{k : C(w^x[k]) > C_{th}\} \tag{4.18}$$

where C denotes curvature. Let us now define the representation R_m for WMSR as below

$$(L_{0ms}(W_j x)_{1\leq j\leq J}, V_0(W_j x)_{1\leq j\leq J}, V_0(W_{J+1}x)) = (L(w^x), V(w^x)) \tag{4.19}$$

4.4.1 Stability of reconstruction from wavelet representation with missing samples

Modeling from missing sample using wavelets is an active research topic [27]. To prove stability of reconstruction from WMSR, Theorem 1 of [4] is restated for the sake of completion.

Theorem 4.2: Any multi-scale maxima representation is an inherently bounded Adaptive Quasi Linear Representation (AQLR).

The theorem is valid as wavelets satisfy Riesz frame conditions.

It can be easily seen that

$$\|W_j x^c\| \leq K_1 \|V_{max} w^x\| \leq K_2 \|V w^x\|, K_1, K_2 > 0 \tag{4.20}$$

where V_{max} is the set of multi-scale maxima values in maxima representation. The first inequality is proved in Theorem 1 in [4] and the second inequality holds because $L_{0mm} \subset L_{0ms}$. This implies

$$\|W_j x^c\| \leq K \|V w^x\|, K > 0 \tag{4.21}$$

$$\|x^c\| \leq K \|V w^x\|, K > 0 \tag{4.22}$$

Hence, we state Corollary 4.1: WMSR is also an inherently bounded AQLR and reconstruction is stable.

4.4.2 Parameter estimation of the LTV model by alternate projection

The alternate projection algorithm starts with significant wavelet coefficients (or the wavelet representation with missing samples) and seeks the solution to have minimum local error norm by alternately projecting the solution orthogonally on $\mathcal{L}$ and $\Im$. At every iteration, the solution is estimated once using the projections of the input (path a in Figure 4.3) followed by using those from the output (path b in Figure 4.3) [55]. The alternate projection algorithm gives better results compared to that from exact implementation of 4.10 as restrictive constraints given by 4.9 are not needed in the implementation. An implementation of the algorithm works through the following steps in each iteration z with initial values of parameters $\hat{a}_{jk}^{0-}$ all set to zero.

1. Intermediate values of parameters are computed at each scale j and each time instant k given by $\hat{a}_{jk}^{Z-}$ using contribution from significant wavelet coefficients of output.

2. The output estimate is projected in time domain (first equality) and again back onto wavelet basis (second equality).

3. Final values of parameters for the iteration given by $\hat{a}_{jk}^{Z+}$ is computed using contribution from significant wavelet coefficients of the input.

The algorithm listed below continues till the mean squared errors between two consecutive iterations satisfy $\varepsilon^{z} - \varepsilon^{z-1} < \varepsilon_{th}$, where ε_{th} is the threshold of the incremental error.

$$1.\ \hat{a}_{jk}^{z-} = \hat{a}_{jk}^{(z-1)+} + \left[\frac{\langle \theta_k,\, y^s \rangle - \left\langle \theta_k,\, \hat{y^s}^{(z-1)+} \right\rangle}{\langle \theta_k,\, y \rangle} \right],$$

$$\forall k \in \{ I_y : |\langle \theta_k,\, y \rangle| \geq \lambda_y \}$$

$$\hat{a}_{jk}^{z-} = \hat{a}_{jk}^{(z-1)+},$$

$$\forall k \notin \{ I_y : |\langle \theta_k,\, y \rangle| \geq \lambda_y \}$$

$$2.\ \sum_{j,k} \hat{a}_{jk}^{z-} \langle \theta_{jk},\, y \rangle \, \theta_{jk}^{*}(t) = \hat{y^s}^{z-} = \sum_{j,k} \left\langle \theta_{jk},\, \hat{y^s}^{z-} \right\rangle \theta_{jk}^{*}(t)$$

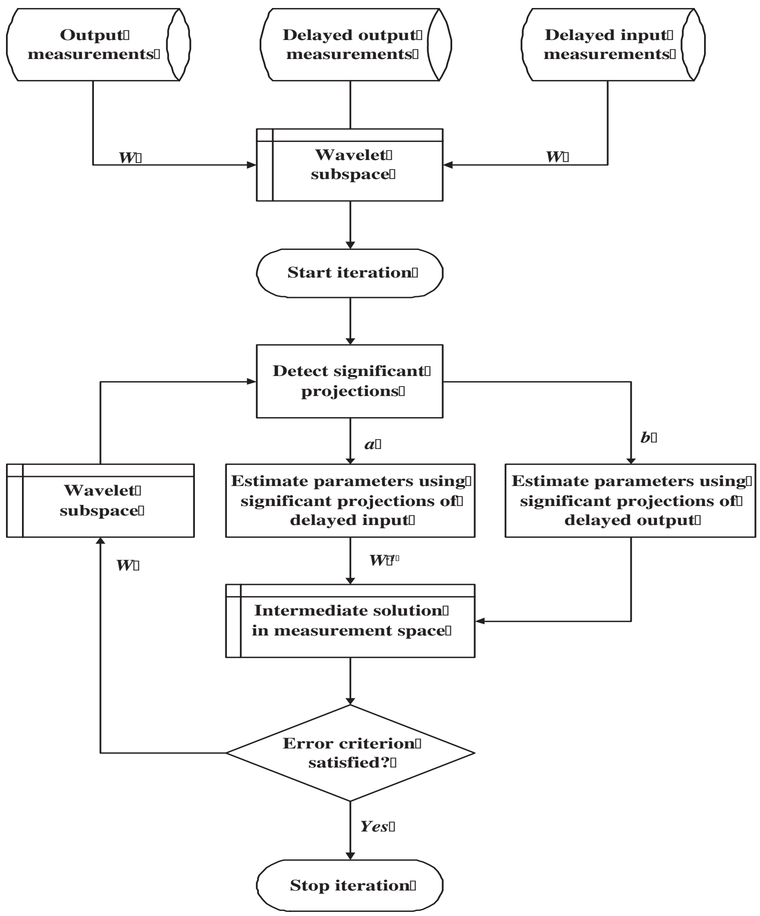

Figure 4.3: Alternate projection algorithm

$$3.\ \hat{b}_{jk}^{z+} = \hat{b}_{jk}^{z-} + \left[\frac{\langle \theta_k,\, y^s \rangle - \left\langle \theta_k,\, \hat{y^s}^{z-} \right\rangle}{\langle \gamma_k,\, u \rangle} \right],$$

$$\forall k \in \{ I_u : |\langle \gamma_k,\, u \rangle| \geq \lambda_u \}$$

$$\hat{b}_{jk}^{z+} = \hat{b}_{jk}^{z-},$$

$$\forall k \notin \{ I_u : |\langle \gamma_k,\, u \rangle| \geq \lambda_u \}$$

$$4.\ \sum_{j,k} \left[\hat{b}_{jk}^{z+} \langle \gamma_k,\, u \rangle + \left\langle \theta_{jk},\, \hat{y^s}^{z-} \right\rangle \right] \theta_{jk}^{*}(t) = \hat{y^s}^{z+} = \sum_{j,k} \left\langle \theta_{jk},\, \hat{y^s}^{z+} \right\rangle \theta_{jk}^{*}(t),$$

$$5.\ \varepsilon^Z = \frac{1}{K} \sum_k \left(\hat{y}_{k+1} - \hat{y}_{k+1}^{z+} \right)^2$$

4.5 Experiments in the liquid zone control system

The mathematical model of the LZCS is useful in analysis work such as the study of reactor power variations during normal operational transients, besides in control system design. The normal operation of the LZCS full scale test setup consists of maintaining ZCC water levels at the desired level set points and increasing it or decreasing it at controlled rates. Sets of input output data were collected from experiments conducted in LZCS test setup for obtaining mathematical model of the water level dynamics in both open loop and closed loop cases. A prediction error method will consistently estimate the system under certain conditions [43] even if the data is collected under feedback. Here, the closed loop system is identified from reference input to output. Open loop system can be retrieved form the identified system making use of the known regulator. In next chapter, a method of controller design, using parameters of the identified open and closed loop models, is discussed. As the ZCC water level depends essentially on the water inflow and the inflow varies, in turn, depending on the position of inlet control valve of the ZCC, the input data are the samples of the signal to the inlet control valve and the output data are the samples of ZCC water level signal.

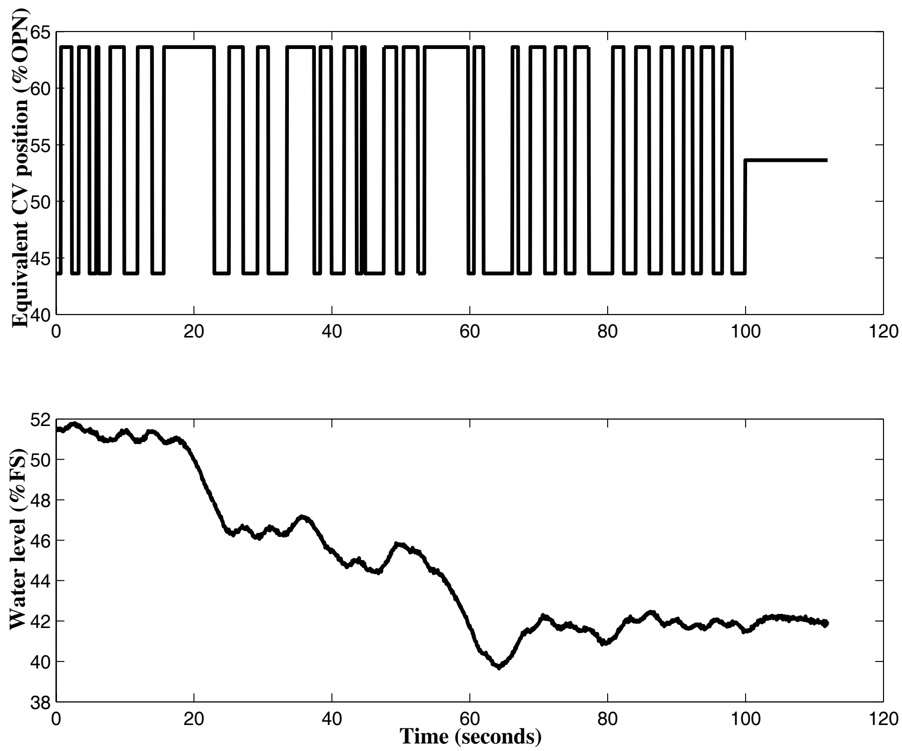

Figure 4.4: Open loop input-output data set for training.

4.5.1 Open loop case

Input output data for one ZCC at a time was collected by isolating the ZCC for a short duration from the automatic level controller and applying Pseudo Random Binary Sequence (PRBS) input to the inlet control valve of the ZCC. PRBS is popularly used to expose all the system modes in a system identification problem. PRBS comprises of a sequence of variable duty cycle square pulses which can assume two states. This ensures that the process component to be identified is persistently excited. Here, PRBS is used as the reference or control input to the valve. The amplitude of the input signal was equivalent to 10 % opening of the inlet Control Valve (CV) around the steady state position. This magnitude is large enough to obtain a good signal-to-noise ratio. Figures 4.4 and 4.5 depict the input output data sets, collected at 50 ms uniform interval, used later for model training and validation respectively. After the collection of data, the inflow control valve was switched back to the automatic level controller. The water level after the experiment is found to be slightly different from the steady water level before the experiment, but its effect on identification is expected to be insignificant. Hence, this deviation was ignored.

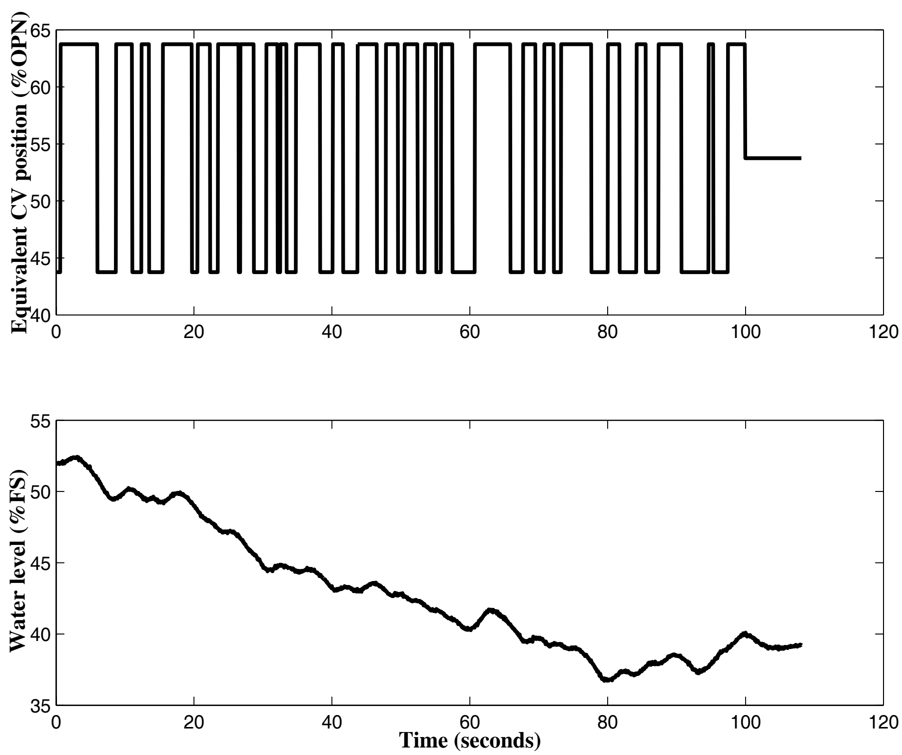

Figure 4.5: Open loop input-output data set for validation.

4.5.2 Closed loop case

Identification experiments in closed loop are performed to take care of the open loop system pole, located at origin. In the experiments, the water level in each ZCC was regulated by its level controller with initial level setpoint at 30%. Input output data for one ZCC at a time were collected by varying its level setpoint in the range of 30% to 80% and the level was allowed to stabilize at the new setpoint. In each experiment, the water level is brought back to 30% at the end of the experiment. The variations in level setpoint are designed to be large so as to obtain a good Signal-to-Noise Ratio (SNR) such that the variation of water level due to process noise is minimum. Such an input design for the purpose of identification would ensure that effect of noise on the parameter is negligible. Figures 4.6 through 4.9 depict four sets of input output data collected at 50 ms uniform interval, in four experiments (experiment 1 through 4), conducted on the LZCS test set-up. Input signal $u(k)$, is shown as the equivalent desired position of the CV in terms of percentage opening (%OPN). The output signal $y(k)$, is the level of water expressed as percentage of full scale (%FS). Full scale level means that the height of the water column is equal to the full height of the ZCC. The control input to the valve is changed from one level to the new level in step manner by the controller, on demand, to change the water level in the compartment.

4.6 Validation of assumptions

It would be interesting to see the correlation functions i.e. auto-correlation and cross-correlation functions of the input and output data set (refer figure 4.10, the same figure is used again in Appendix B for estimation of time delay of the process). As expected, system output is less strongly correlated with itself, as compared to that of input. But one can see that the plant is predominantly an AR process because the input-output cross-correlation is not very high.

Let us now check output auto-correlation in frequency sub-bands (Figure 4.11) generated by a realization of UDWT. It can be seen that except for approximation by $W_6\, LPF$, width of the main lobe of the auto-correlation function has greatly reduced. So it can be concluded that the system can be modeled as reduced memory systems in sub-bands. Lower frequency sub bands show stronger correlation justifying decimation.

4.7 Identification of the LZCS with wavelet basis

A simple model of the LZCS can be developed from first principles considering a zone control compartment (ZCC) as a tank in which the water level

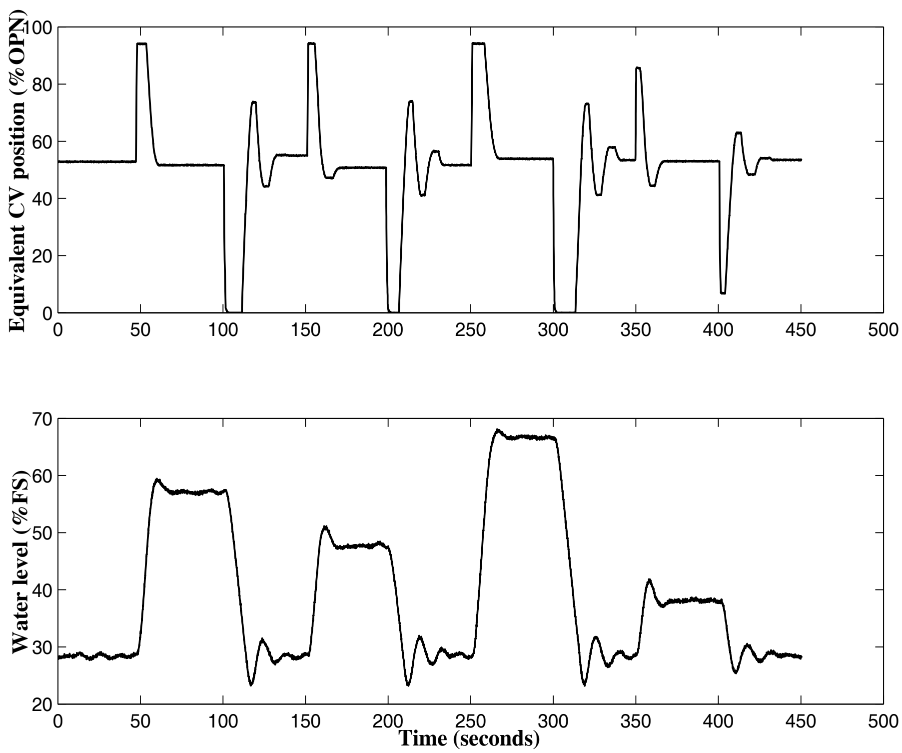

Figure 4.6: Input output data indicating the equivalent %OPN of the CV and level of water as %FS: Water level is raised from 30% FS and is broughtback to 30% FS fourtimes.

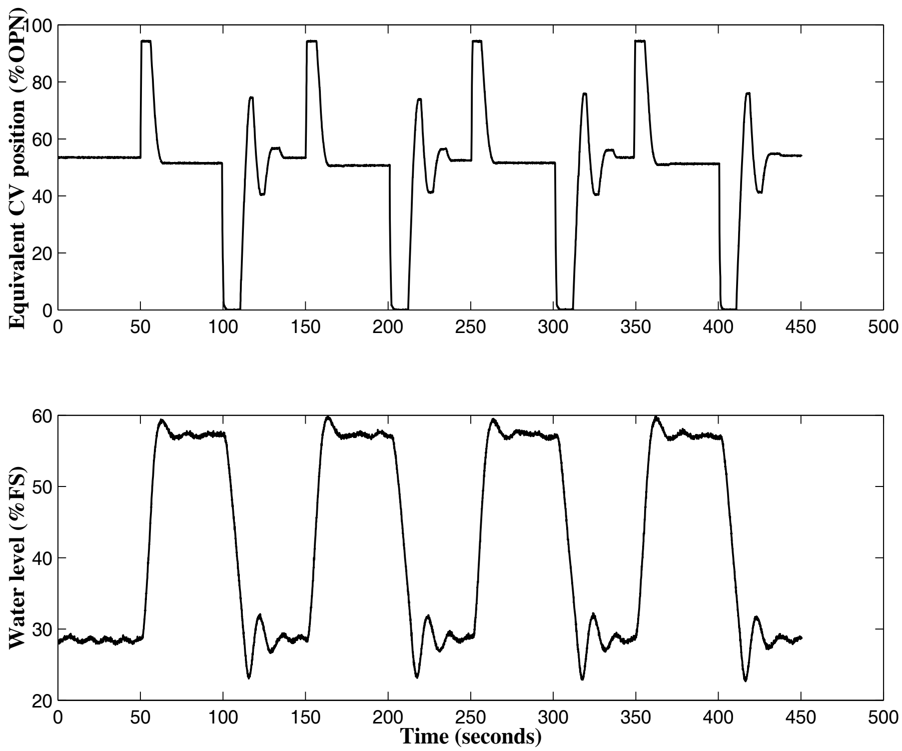

Figure 4.7: Input output data indicating the equivalent %OPN of the CV and level of water as %FS: Each time water level is changed from 30% FS to 60% FS and is allowed to settle there.

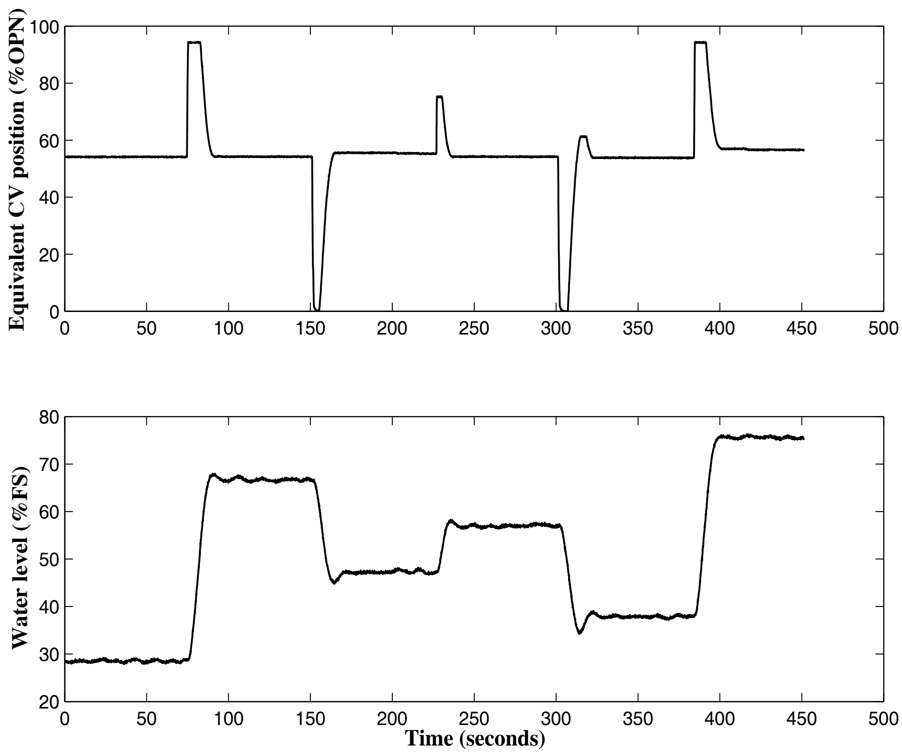

Figure 4.8: The input output data indicating the equivalent %OPN of the CV and level of water as %FS in LZCS: Water level is raised from 30% FS and is allowed to settle at five different levels.

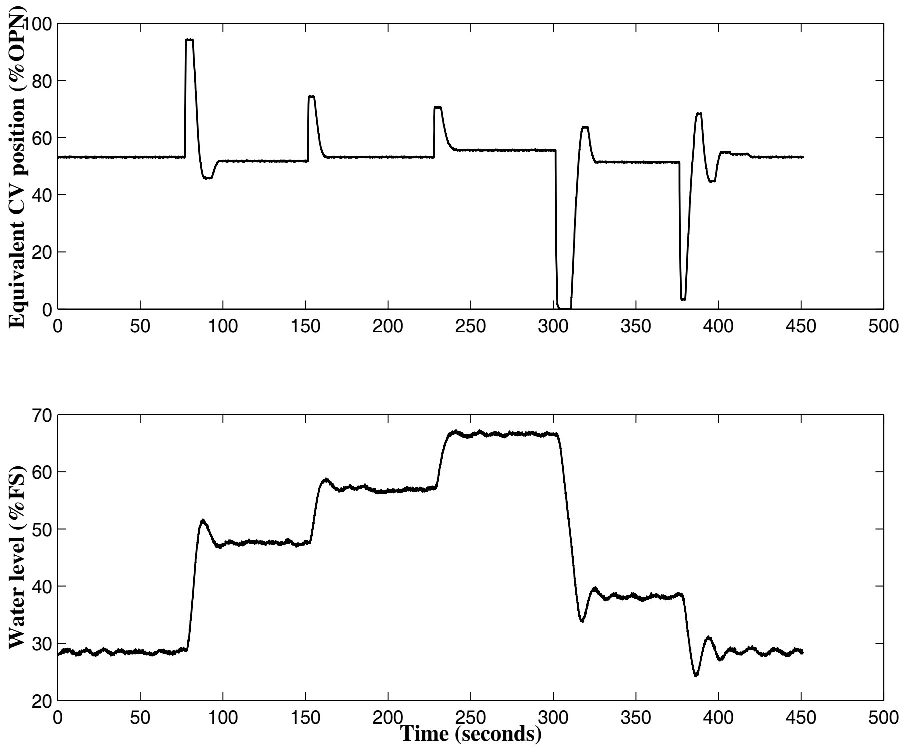

Figure 4.9: The input output data indicating the equivalent %OPN of the CV and level of water as %FS in LZCS: Water level is raised from 30% FS and is brought back to 30% level after excursion over four different steady levels.

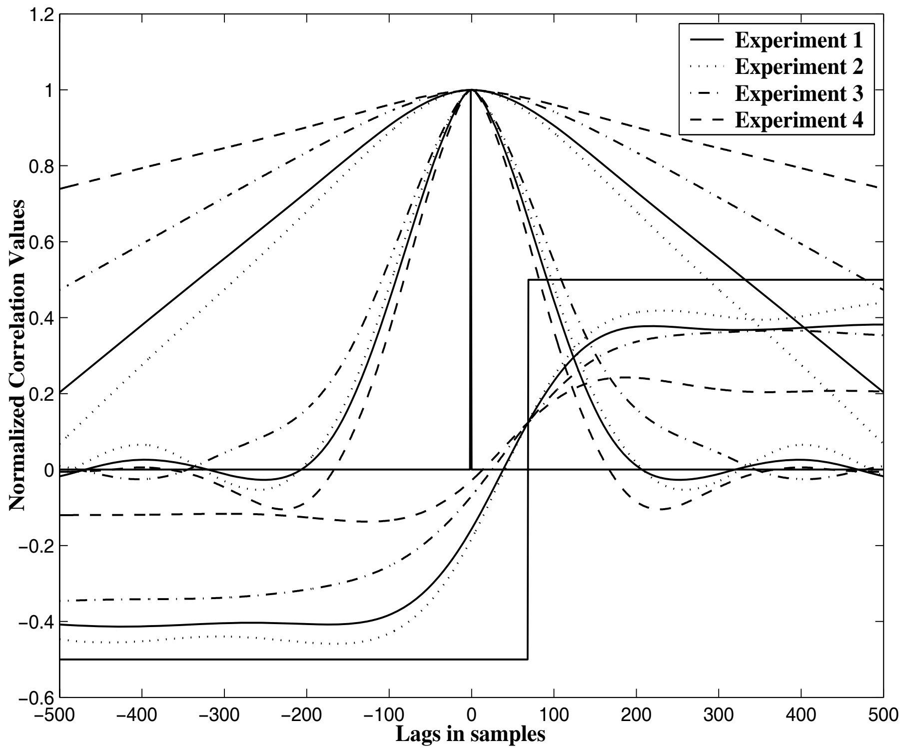

Figure 4.10: Input output correlation of ZCC plotted with lags. Cross correlation functions concur at a point whose abscissa gives time-delay.

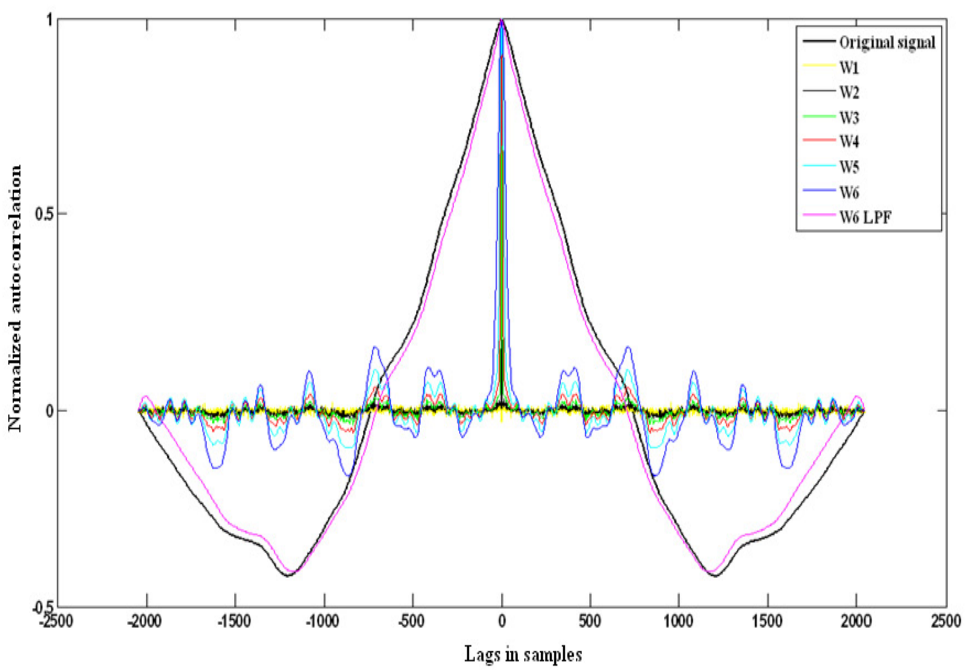

Figure 4.11: Auto-correlation functions in sub-bands

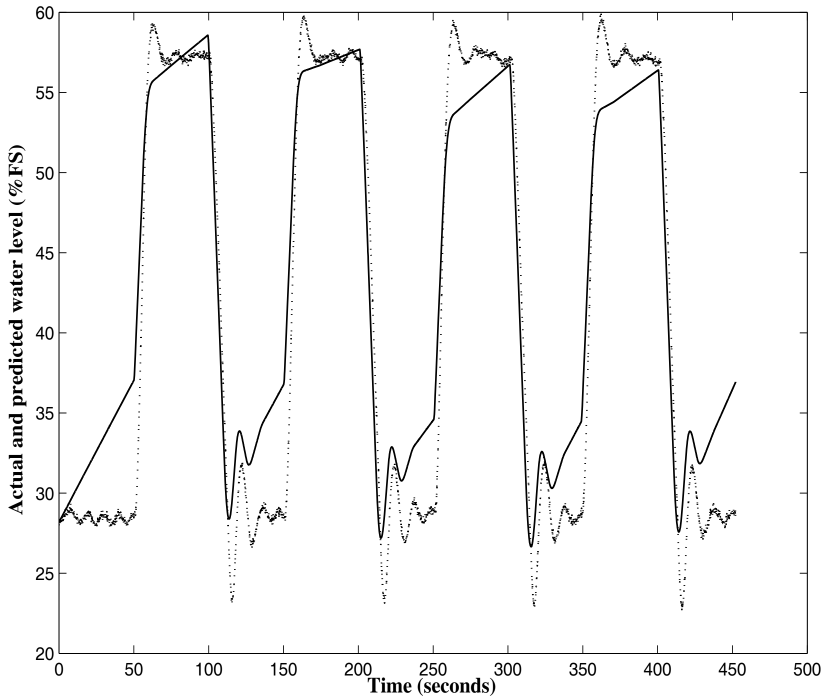

Figure 4.12: Validation result with a simple first order model.

variation is caused due to variations in inflow that occurs when position of control valve changes due to variations in input. Figure 4.12 compares the estimated output of such a simplistic model with actual output shown in Figure 4.7.

Although such a simplistic first order model is adequate for the initial design of control system, simulation needs rigorous models of LZCS. This would require knowledge of valve design data including the characteristics of its different accessories. Moreover, for obtaining a reasonably accurate model, it would also be necessary to account for the transport delay as the water level in the ZCC would vary after a finite delay and not instantaneously following a change in flow through the control valve. This is because of considerable length of piping between the control valve and ZCC. Similarly due to large length of tube from the purge flow regulator to ZCC and from the ZCC to gas outlet header, there could be delay in sensing the variation in ZCC level. In view of these difficulties, developing the mathematical model for ZCC water level dynamics employing a suitable method of identification from measurement of input and output is preferred.

Identification of LZCS with classical input-output models is attempted

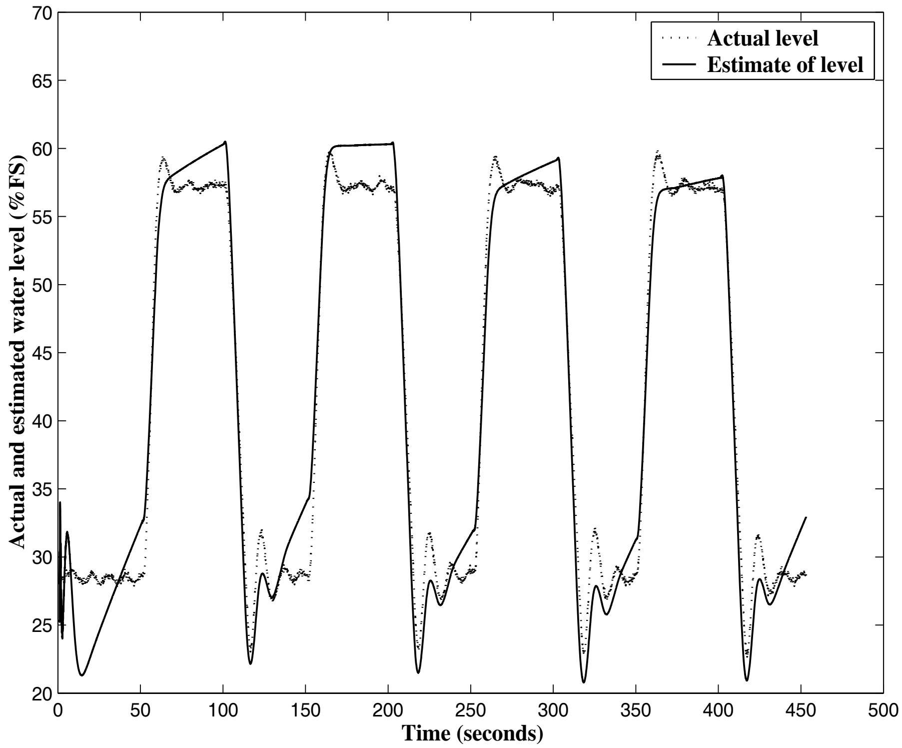

Figure 4.13: Validation result with a high order BJ model.

first. Initially, identification with sinc basis ARX model and its variants were attempted. ARX type models showed unacceptable mismatches between output of the model and actual output in all the experiments. Most of the lower order ARX type models offered very poor match clearly failing to capture dominant dynamic modes. Simulated BJ model identified shows best result where order of all filters is 9. The model is identified based on the data given in Figure 4.6 and cross validation result with the data given in Figure 4.7 is shown in Figure 4.13.

The mismatch does not reduce even by increasing the filter orders or fine tuning other parameters. Such a high order model in any case is not fit for the purpose of control system analysis. It was also observed from the pole-zero plot, that the identified system is almost always unstable, having poles outside the boundary of the unit circle. Even when the poles are inside, nearness to the boundary compromises the robustness of the system. The open-loop LZCS system is essentially an integrator with nonlinearities due to control valve and flow characteristic. The instability to the step input is due to the response of the integrator and associated nonlinearity which could not be captured appropriately by sinc basis used in the ARX type models. A common practice for identifying a system with an integrator is to differentiate the input prior to identification. Although it slightly improves the match, the noise is often amplified. In conclusion, the modeling exercise with sinc basis ARX model and its variants was not satisfactory. Hence, modeling based on wavelet basis was explored. Wavelet based models are expected to perform better due to their excellent local approximation property. In case of wavelets, systems are approximated in a shift invariant subspace and the idealized restriction of band limited approximation is loosened. As a result lower order models are expected with wavelet bases.

Traditionally, efficacy of a technique of system identification is demonstrated by unbiased estimate of system parameters. It is proposed here however, to check the local consistency in the output and verify the model by cross-validation. Unbiased estimates of system parameters are obtained in the process as suggested in Section 4.3.1. It is argued that for all practical purposes it is sufficient to cross validate a model exhaustively with different inputs for testing the consistency in the output.

The proposed method of system identification in wavelet domain based on consistent output estimate has been applied to the LZCS and the results are presented in next section. Decimated wavelet transform is not shift invariant and aliasing takes place in wavelet domain as very often crude filters are used. Detailed description of decimated wavelet transforms and design methods of filter banks could be found in [73]. Assuming no loss of significant signal due to thresholding the results however, would be commensurate with the theory developed for the shift invariant wavelet basis. The advantage of using decimated wavelet transform is that for predicting an output signal of length N with wavelets, numbers of basis functions used are also N. There are of course $2N$ numbers of wavelet coefficients involved when both input and output constitute the regressor. However, as the method of consistent estimate solves only N equations due to N basis, the proposed method of solution uses a linear combination of input and output at each of these N points. For modeling with wavelet basis, the decimated wavelet transform is used where past outputs are used in addition to past inputs (delayed) to predict the one-step-ahead estimate as given in (3.6). The idea is to get the prediction as a consistent estimate such that the significant coefficients in wavelet domain, of the shifted output signal and weighted sum of wavelet coefficients of the input and the output match at every instant. The estimated time-delay between the input and the output is of the order of $3.5s$ which was computed using the technique described in B.4.

4.8 LTI model of the LZCS with orthogonal wavelets

For identification of an LTI model of the LZCS with orthogonal wavelets, it is assumed that $\theta_i = \gamma_i$, such that working is possible in the output space. This is permitted, because for a SISO system, both input and output are scalars and belong to the same space. It may be noted that, given a certain number of basis functions, reconstruction from orthogonal projection gives minimum error in least squares sense. The results presented in this section are derived using orthogonal wavelet basis.

As mentioned earlier, θ_i is designed to minimize the number of coefficients of wavelet transform of the input by finding a good match of the signal with the basis and its dilated versions. The best basis, in this sense, would have maximum cross correlation with the signal and minimum with the noise, assuming that noise is uncorrelated with the signal. As input is a sequence of steps, $db1$ of Daubechies family of wavelets (Haar wavelet) is chosen to represent the signal in wavelet domain. A good match of the input signals with the Haar wavelet results relatively less number of significant coefficients in the decomposition. A sparse representation thus obtained, reduces number of basis functions used to represent a system.

4.8.1 Open loop model

In open loop case, first the results are presented considering the system to be piecewise linear, time invariant. Parameters are estimated by exactly implementing (4.10) with $k_1 = 0.99$ and $k_2 = 0.01$. Figure 4.14 shows the result of training, using data in Figure 4.4. The identified model is validated with the data in Figure 4.5 and the test results are shown in Figure 4.15. The result shows clear deviation of the estimate from the actual measurement at least on two occasions (time intervals). From this point onwards, we switch to iterative estimation algorithm based on alternate projection and the results presented in subsequent sections are based on the algorithm listed in section 4.4.2.

4.8.2 Closed loop model

In closed loop case, data of Figure 4.8 is used for identification of the model and data of Figure 4.9 is used for validation of the model. The proposed iterative algorithm estimates the time invariant parameters at each scale. Alternate projection scheme described in section 4.4.2 is used for reconstruction of the estimate of the shifted output. In each iteration, the optimum estimate of $\hat{a}_j^{z+}$as defined in section 4.4.2 is computed as the mean of all $\hat{a}_j^{z+}s$ over each scale, while changing the scale index from one to eight. As the algorithm converges, $\hat{a}_j^{z+}s$ for input and output may come out to be different

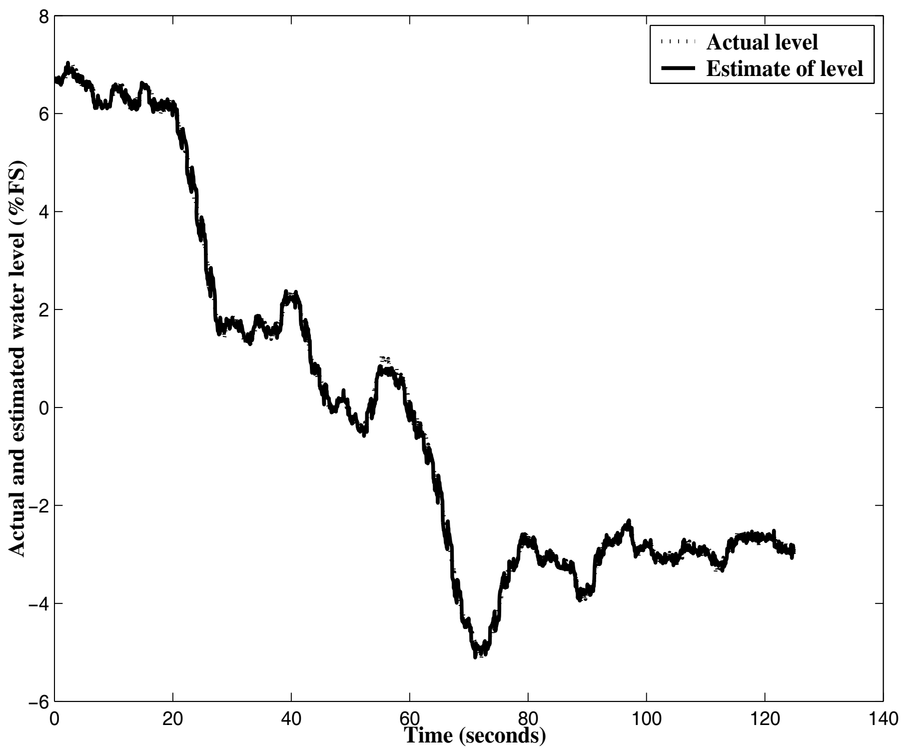

Figure 4.14: Training of the open loop, piecewise linear, time invariant model

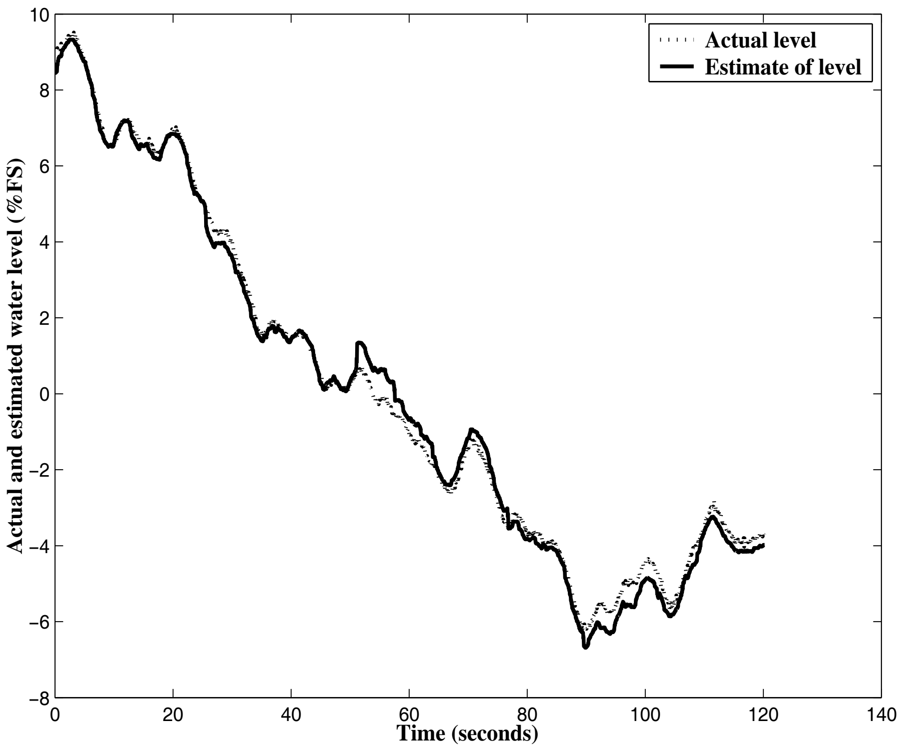

Figure 4.15: Validation of the open loop, piecewise linear, time invariant model

Table 4.1: Wavelet LTI Model of the LZCS

Scale index, j	1	2	3	4	5	6	7	8	9
a_j	0.00	0.18	0.06	0.95	0.98	0.99	1.02	0.98	1.00
b_j	-1.91	3.77	-0.96	-0.99	-4.16	-0.02	0.84	0.60	0.02

as no restrictive assumptions were made as in Theorem 4.3. The LTI model of the LZCS thus obtained, is given by two sets of multipliers as shown in Table 4.1 .

Note that low frequencies in the input contribute more in the estimate. However, similar conclusion cannot be made for the output.

The iterative scheme attempts to use minimum number of basis by thresholding wavelet coefficients for the reconstruction. Thresholds λ_u and λ_o are taken as 0.05% of the maximum absolute values of wavelet coefficients of input and output respectively at each scale. The choice is a balance between reducing the number of basis functions and quality of approximation.

The reconstructed water level output signal after 7 iterations and actual water level output signal are compared in Figure 4.16. Although approximated by a Haar wavelet which is derived from zero order spline, an excellent match is observed between the consistent prediction and the actual output.

The identified LTI model based on the input output data given in Figure 4.8, thus obtained, is now tested to check if actual output can be predicted, also for the input output data shown in Figure 4.9. The output in this case, is again measured by exciting the control valve with a different sequence of steps. The cross validation result is shown in Figure 4.17.

Note that an excellent visual match is observed between the model output and the actual output level of the ZCC. The result conclusively proves the validity of proposed method of parameter estimation based on consistent output estimate. In the next section LTV modeling of the LZCS with spline biorthogonal wavelets is investigated.

4.9 Modeling with biorthogonal spline wavelets: LTV and LTI models

The LTV modeling approach due to [90, 91] which assumes time invariance over the length of support of the spline biorthogonal basis functions, when applied to the LZCS, failed in cross validation. This is primarily because the approach precludes non linear operation such as thresholding, resulting locally unstable solution. The instability arises due to ill conditioning of the regressor correlation matrix and invalid assumption of local time invariance failing to model rapid changes in the response. In this section, the results

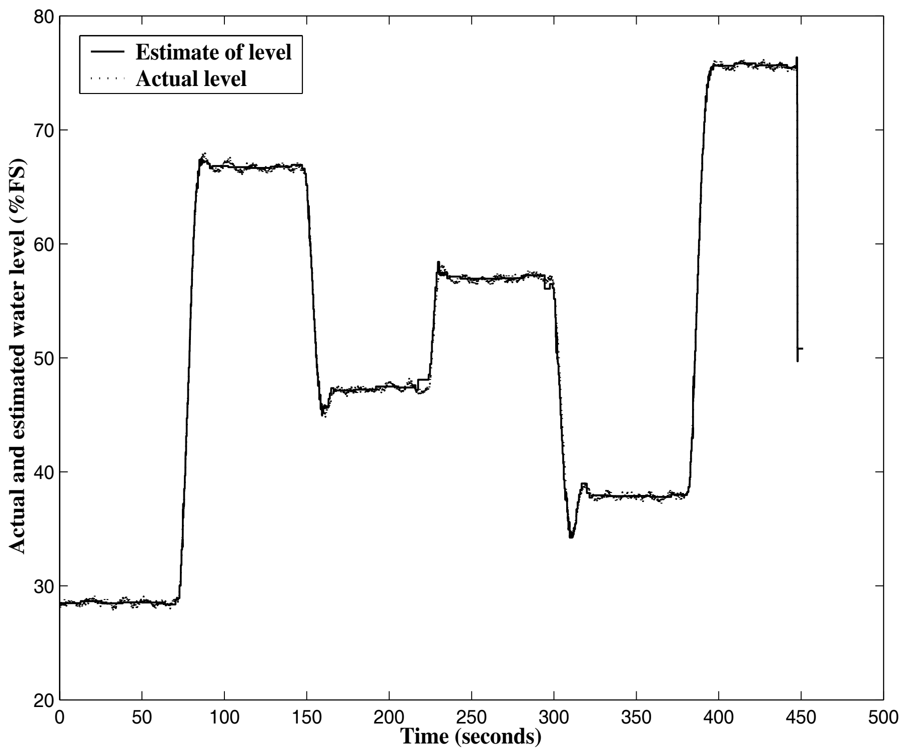

Figure 4.16: Result of model identification using data of Figure 4.8: Reconstructed ZCC water level after 7 iterations and actual water level output.

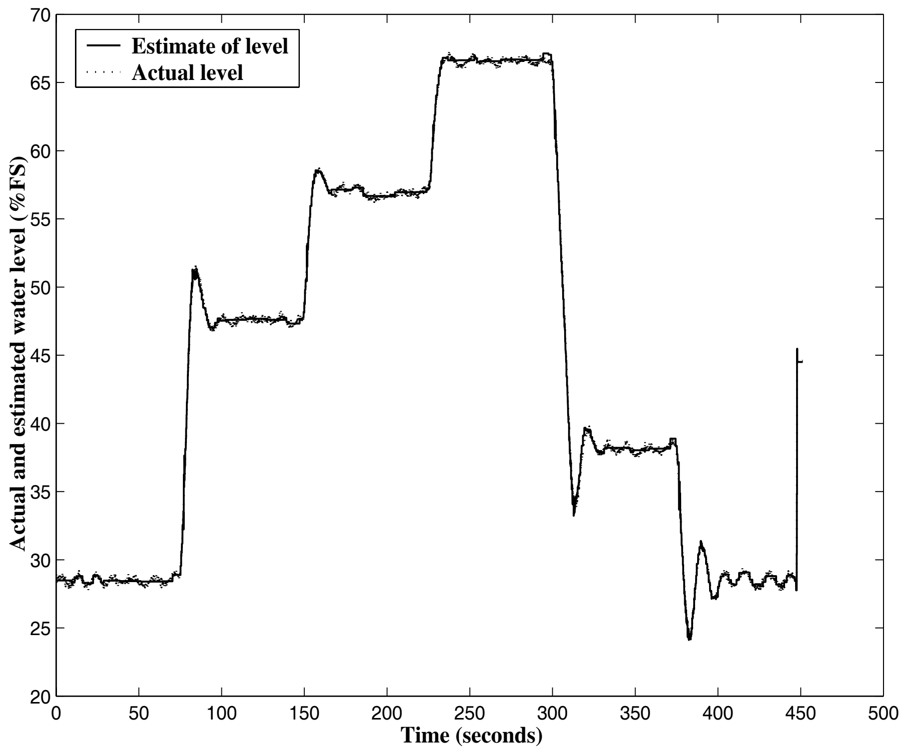

Figure 4.17: Result of cross validation of the identified model with a new input: Model output is compared with the actual output shown in Figure 4.9

of LTV modeling of the LZCS using consistent output prediction with spline biorthogonal wavelets are presented.

4.9.1 Discussion on the choice of spline biorthogonal wavelets

Two spline biorthogonal wavelets of different degrees are used, one for projecting the input and the other for projecting the output. It has been shown in chapter 3 that even input side modeling with spline biorthogonal wavelets permits use of two different wavelets for approximating input and output. The choices are motivated by the economy of representation. Wavelets can be seen as matched filters i.e. the guideline for choice is to pick up the wavelet which correlates best with the signal such that representation is maximally sparse. Wavelet RBIO1.5 is used for projecting or analyzing the input. As in $db1$ the analyzing scaling function of RBIO1.5 is a box function or box spline of degree zero. As has been observed earlier, projection of step or PRBS input on the scaling and wavelet functions of RBIO1.5 shall minimize number of significant wavelet coefficients.

As explained earlier a higher order spline can be considered as convolution of a low pass filter with a lower order spline. The low pass filter removes noise from output and the process can be viewed as identifying the dynamic modes using pre-filtered output. Output measurement is usually noisy and reflects excited dynamic modes of the system which limit output bandwidth. Approximation with a spline function of higher degree is more appropriate because it economizes number of model parameters as well as derives an accurate model insensitive to noise. RBIO2.4 has been used for projecting the output. The analyzing scaling function of RBIO2.4 is a triangular function or box spline of degree one.

4.9.2 Open loop model

In open loop case, an LTV model is trained with the data in Figure 4.4 to get the consistent output estimate (refer Figure 4.18) (so that the significant coefficients in wavelet domain match). The identified model is tested on the second set of data shown in Figure 4.5, again excited by a PRBS input. A threshold of 0.05 of the maximum coefficient value has been used for thresholding at each scale, both at input and output stage. A reasonable match is observed between the model output and the actual output when cross validation is tried with the new data set. The result is shown in Figure 4.19.

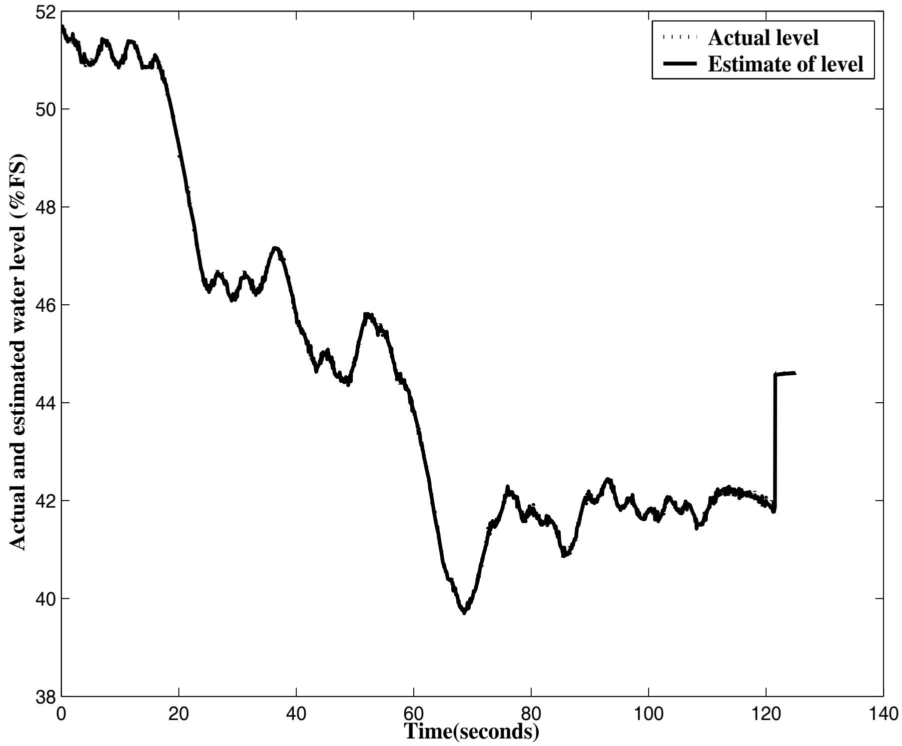

Figure 4.18: Training of the open loop model

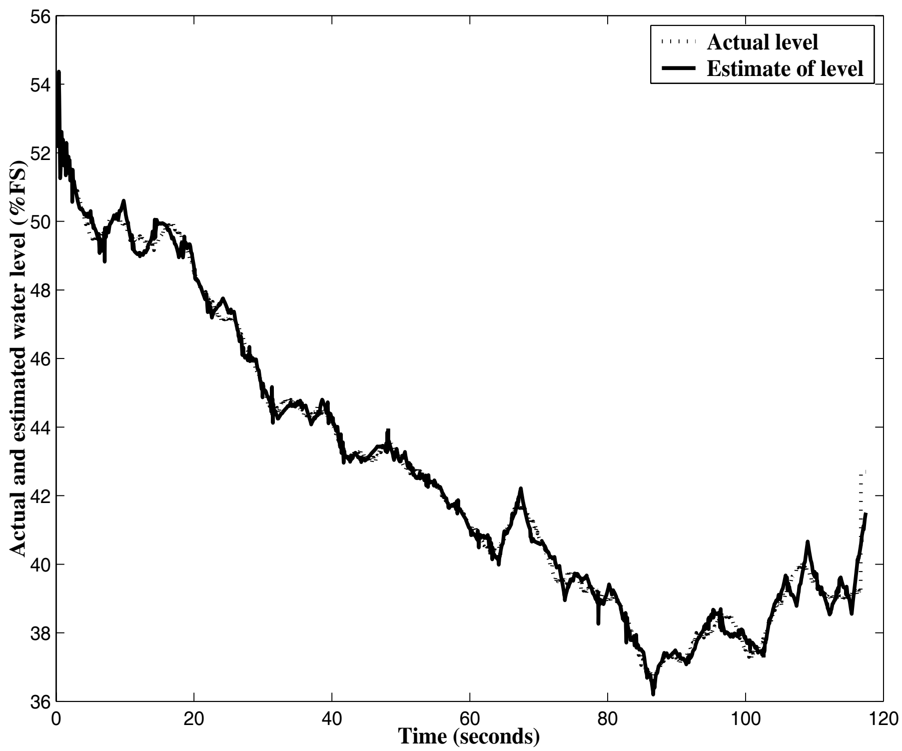

Figure 4.19: Validation of the open loop model

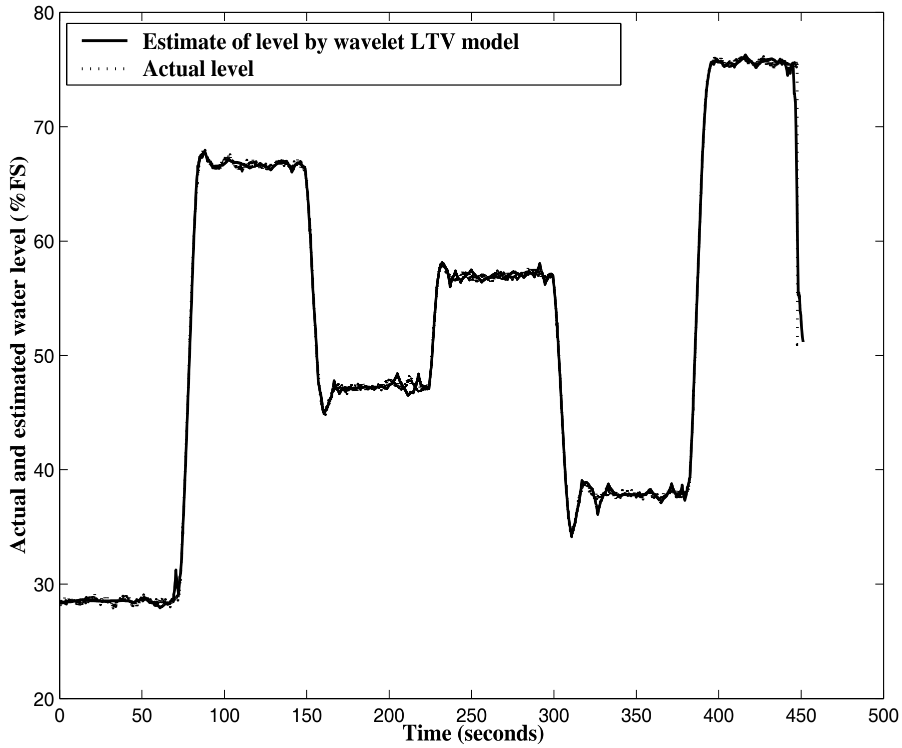

Figure 4.20: Result of model identification with spline biorthogonal wavelets using data of Figure 4.8

4.9.3 Closed loop model

As earlier, data of Figure 4.8 is used for identification of the model and data shown in Figure 4.9 is used for validation of the model. The proposed iterative algorithm estimates the time varying parameters at each scale. The reconstructed water level output signal of the training set and actual water level output signal are compared in Figure 4.20. One can observe an excellent match between the consistent prediction and the actual output. The identified LTV model, is now tested for the input output data shown in Figure 4.9. The cross validation result of the LTV model is shown in Figure 4.21. Again a good match is observed between the model output and the actual output level of the ZCC. Superimposed on the same figure is the cross validation result of the LTI model which shows a better match with the actual output indicating that the system is predominantly LTI..

Comparison of validation results in Figure 4.17 and 4.21 show that a better match is obtained in transient and steady state responses with spline biorthogonal wavelets. The results validate use of spline biorthogonal wavelets as basis for identification.

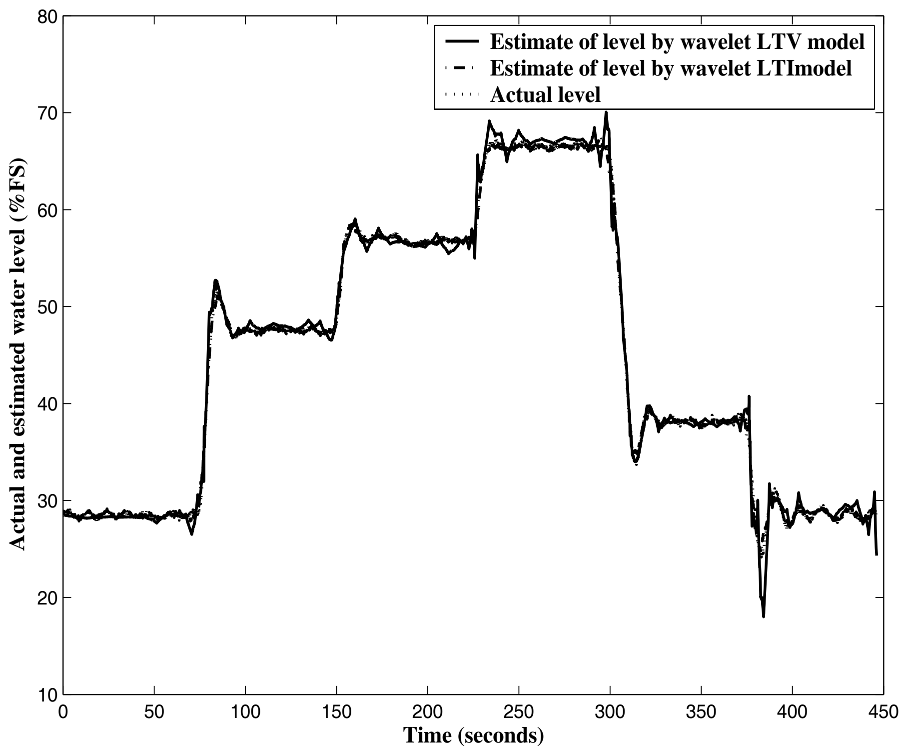

Figure 4.21: Result of cross validation of the identified spline biorthogonal wavelet model with a new input.

Table 4.2: Variation in Normalized Mean Square Error with Signal-to-Noise Ratio of the output (Input and output threshold is kept at 0.05% for both identification and validation)

Signal-to-Noise-Ratio	Normalized Mean Square Error
Measured Output (>50dB)	0.015
30 dB additive noise	0.015
20 dB additive noise	0.025
10 dB additive noise	0.045

As the method is validated on the data obtained from a full scale test set up of the LZCS, it can be considered sufficiently rugged and insensitive to process and measurement noise usually present in an actual system. However, a numerical study of the effect of an additive, synthetic output noise (white Gaussian) on identification and validation is summarized in Table 4.2. Normalized Mean Square error (NMSE) shown in the table is given by

$$NMSE = \frac{\sum_{k=1}^{K} \left(y\left[k\right] - y\,\hat{\left[k\right]} \right)^2}{\sum_{k=1}^{K} \left(y\left[k\right] \right)^2}$$

where $y\left[k\right]$ is the measurement and $\hat{y}\left[k\right]$ is its estimate.

For the same threshold of 0.05%, Normalized Mean Square Error (NMSE) in output estimate increases as SNR falls i.e. as the noise increases. For output with low SNR, NMSE can be improved by increasing threshold, obviously because higher threshold would pre-filter more noise. Choice of best threshold depends on the level of noise and is systematically studied in wavelet literature [20, 21]. For the present application of identification of the LZCS, threshold is decided such that NMSE in validation is low and parameters are insensitive to small change in threshold value.

4.10 Summary

This chapter presents the method of estimation of model parameters in wavelet domain, using consistent output prediction. It is established that the method is an efficient alternative to classical solution of least squares minimization problem. A strictly time varying LS minimization solution of a linear prediction model, with wavelet basis functions is derived. The solution is near true and optimum given the choice of the wavelet basis and the level of threshold. The method minimizes the local error and provides a solution consistent with the output measurement in wavelet domain. Validity of the assumptions made while deriving the new solution is justified based on the correlation structure of input and output data in transform domain. An approximate piecewise linear model is derived as a special case of general LTV model and it is proved that the parameter estimates of the linear model by consistent output prediction are unbiased and bounded. The estimator of the approximate LTI model has error bounds that can be controlled by the choice of threshold. The characterization considers additive noise in the output measurement and handles it in equation error framework. For testing and validation of the identified model, it is necessary to devise a method to reconstruct the model output. Reconstruction from a representation in wavelet domain is discussed. A proof of stability of the reconstruction based on WMSR is stated in corollary 4.1. The proof is necessary to support the non linear approximation used for the reconstruction of a signal from sparse wavelet representation. A recipe of the new algorithm for reconstruction of consistent LTV model output by alternate projection is proposed. The algorithm is motivated by the derivation of parameter estimates given by theorem 4.3. The chapter also presents a brief description of the LZCS used for power control in large PHWRs. Details of experiments conducted in open loop and closed loop cases in full scale test set up and the results of the experiments are discussed. Both open loop and closed loop cases are considered for modeling, as next chapter presents a method of controller design with wavelet states using both open loop and closed loop model parameters. Limitations of classical modeling techniques for identification of non linear complex systems such as the LZCS are demonstrated. Efficacy of the proposed modeling technique with orthogonal and biorthogonal wavelets is illustrated by the LZCS model identification, testing and cross validation. Accuracy of the derived models in presence of noise is established.

Chapter V

Controller design by output feedback

The modeling approach in section 3.2 establishes that an input-output model of a system can be identified which would have considerably less memory compared to the original system. Intuitively, control solution of a system identified with a minimum memory model may be searched in the class of admissible controls, also constrained to be minimum memory. In this chapter, answer to the question if it is possible to achieve desired dynamic response, using full wavelet state feedback control will be sought.

5.1 Controller design

Prerequisite of designing a full state feedback controller for the discrete, linear time varying system S_d of order N, given by one-step-ahead state-space model

$$S_d : \quad x[k+1] = A_k x[k] + B_k u[k], \forall k \tag{5.1}$$

modeled from input and output is that the system should be controllable and observable [13]. It is now necessary to examine under what conditions S_d, modeled with wavelet states would also be controllable and observable.

Since for an input-output model, only output measurements are available for applying control, all the system modes can be controlled if the open loop system is observable. From this point of view, wavelet states can also be seen as states of a linear observer [44, 45]. Let us assume, to start with, that wavelet states of S_d can be obtained from the states of identified system given in (3.14) by following linear transformations [32, 33].

$$\overline{w}^y[k] = H_k x[k] \tag{5.2}$$

$$x[k] = \overline{G}_k \overline{w}^u[k] \Longrightarrow \overline{w}^u[k] = G_k x[k], G_k = \overline{G}_k^* \tag{5.3}$$

Although more rigorous analysis would follow, intuitively it is easy to see that the relationship given in (5.2) may arise from observability condition and from the fact that no two wavelet states are completely redundant. Extending the same, (5.3) suggests that for a controllable system, it is expected that input wavelet states are related to the states of the system.

Here, a few comments are necessary regarding the structure of H_k and G_k. Note that a linear time varying system is completely described over K time samples by KN states (time or wavelet). For wavelet based identification, it is always possible to have a model structure such that

$$(J+1)qK > NK,\ (J+1)pK > NK \tag{5.4}$$

where p and q are dimension of input and output and J is the maximum scale index. Note that the matrix H_k has $(J+1)\,qK$ rows and N columns and the matrix G_k has $(J+1)pK$ rows and N columns. Further, it may be observed that it is possible to choose KN non redundant wavelet states of $\overline{w}^y$ ($\overline{w}^u$) for perfect reconstruction of y. Methods to construct $(\overline{w}^y)^N$ $((\overline{w}^u)^N)$ (non redundant wavelet state description) have been reported in the literature [3, 62]. Let us define an N-tuple, for each k consisting N projections or generalized samples as $(\overline{w}^y)^N[k]$. Let H_k^N be an NXN square submatrix of H_k such that

$$(\overline{w}^y)^N[k] = H_k^N x[k] \tag{5.5}$$

A simple example may be taken to clarify the idea. Let us consider a single output controllable and observable system. Elements of column vector $(\overline{w}^y)^N[k]$ are projections on wavelet basis and could be obtained by collecting k^{th} sample of UDWT of y at each of $J+1 = N$ scales. These projections have quantifiable redundancy [48] but not completely redundant. Then wavelet states can be obtained by weighting elements of the state vector x. Now, based on above discussion following proposition can be made.

Proposition 5.1

1. S_d is observable by wavelet states if rank of H_k is N.

2. S_d is controllable by wavelet states if rank of G_k is N.

Proof of Proposition 5.1, *part* 1

Consider $(\overline{w}^y)^N[k]$ is obtained by a linear filtering (using wavelets) with multiscale causal filterbank W_k^N [73, 83],

$$(\overline{w}^y)^N[k] = W_k^N y \tag{5.6}$$

where, $y\{= y[1], y[2]\cdots y[K]\}$ is a discrete sequence in $\mathbb{R}^n$. Existance of a W of rank KN for a controllable system with non zero and distinct eigen values is proved in section 3.4. W_k^N can be seen as a sub matrix of W.

Hence from (5.5),

$$x[k] = \left(H_k^N\right)^{-1} W_k^N y[k] \tag{5.7}$$

States of S_d can be computed from $y[k]$ if and only if $\left(H_k^N\right)^{-1}$ exists *i.e.* H_k^N is of rank N and hence S_d modeled with wavelet states is observable.

Part 2 *of Proposition* 5.1 can also be proved by observing

$$(\overline{w}^u)^N [k] = W_k^N u[k] \tag{5.8}$$

and arguing in the same manner.

Considering the class of admissible controls for S_d, constrained to be discrete time memory less linear functions of the output,

$$u[k] = C_k y[k] \tag{5.9}$$

Substituting (5.9) in (5.1) and using (5.7)

$$x[k+1] = A_k x[k] + B_k C_k \left(W_k^N\right)^{-1} H_k^N x[k] \tag{5.10}$$

$$x[k+1] = \left(A_k + B_k C_k \left(W_k^N\right)^{-1} H_k^N\right) x[k] \tag{5.11}$$

Note that $(A_c)_k = \left(A_k + B_k C_k \left(W_k^N\right)^{-1} H_k^N\right)$ is the closed loop state transition matrix.

The necessary and sufficient conditions for S_d to be pole assignable by output feedback is that $((A_c)_k, B_k)$ is controllable. That the closed loop system with state feedback x is controllable follows from the assumption that the open loop system with state feedback x is controllable. By virtue of *proposition* 5.1, S_d is controllable with wavelet states, if rank of G_k is N. Moreover, if the loop is closed with output wavelet state feedback then (5.11) implies that it is possible to control S_d with full wavelet state feedback, if rank of $(A_c)_k$ is N which in turn necessitates that rank of H_k should be N. Hence it is necessary that the matrices H_k and G_k are of rank N (i.e. open loop system in wavelet states is controllable and observable) for S_d to be pole assignable by output wavelet state feedback.

5.2 Modeling of a multiple time scale system

The proposed technique can be used very efficiently for modeling and control of multiple time scale systems characterized by fast transients superimposed on slowly varying quasi-steady states. Control design for multiple time scale systems is difficult. The challenge comes from efficiently handling high dimensionality and ill conditioning resulting from the interaction of slow and fast dynamic modes. Such systems often occur naturally and examples include nuclear reactors, power systems, economic models etc. Singular perturbation models are used when states can be grouped into slow and fast

Table 5.1: Parameters of the LTI Model

Scale index, j	1	2	3	4	5
a_j	-0.7	0.1	0.8	1.0	1.0
$b_j\ X10^{-4}$	1.5	1.3	-2.4	-5.1	5.8

ones. Modeling with wavelet basis turns out to be a natural choice for these systems as illustrated below.

Let us consider transfer function $T(s)$ of a 4^{th} order system having poles at $s = 0, -1, -8, -9$

$$T(s) = \frac{0.08(2s^2 + 18s + 9)}{s^4 + 18s^3 + 89s^2 + 72s}$$

Clearly dynamic modes can be grouped into two or three categories, one containing the fast ones due to poles at $s = -8, -9$ and the other(s) containing slower ones due to poles at $s = 0, -1$. Motivation for modeling with wavelet basis is to isolate these groups with minimum interaction. This is likely because in case of wavelets, frequency window is broader than that for Fourier analysis. An LTI model $h(\tau)$ is derived by exciting $T(s)$ with a train of impulses and by consistent prediction of the noisy output as shown in Figure 5.1. The identification now, per se amounts to estimating parameters of the system in multiresolution by assuming the parameters are constants across each scale j. The solution is obtained by iterating alternate projections till the mean square error diminishes. The parameters (Projections on θ_j *i.e.* a_j) given in the Table 5.1 show a definite separation of modes in the wavelet model, fast mode indicated by a_1 and the slower ones by a_2, a_3 and a_4. Both input and output are sampled at 0.1 s.

The model is cross validated by using a sinusoidal input and matching actual and predicted output as shown in Figure 5.2. It may be noted that decimated wavelet transform naturally isolates slow and fast operating modes with optimum resolution. However, number of parameters indicating a group shall depend on several factors such as width of the frequency window chosen for analysis vis-a-vis width of the group, sampling frequency, *etc.*

5.3 Design of a controller with output feedback for a nuclear reactor

In this section the method of computation of controller gains with output feedback is illustrated using an input-output wavelet model of a large nuclear reactor. In a nuclear reactor neutrons are produced, get absorbed or escape at different energies. The mathematical model of a practical nuclear

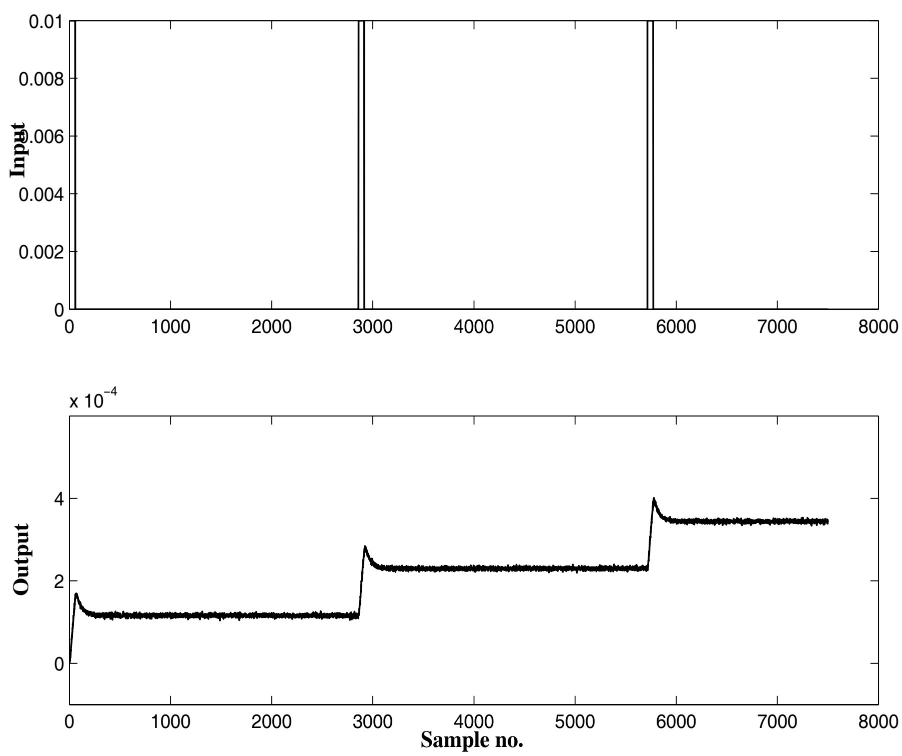

Figure 5.1: Training data for identification

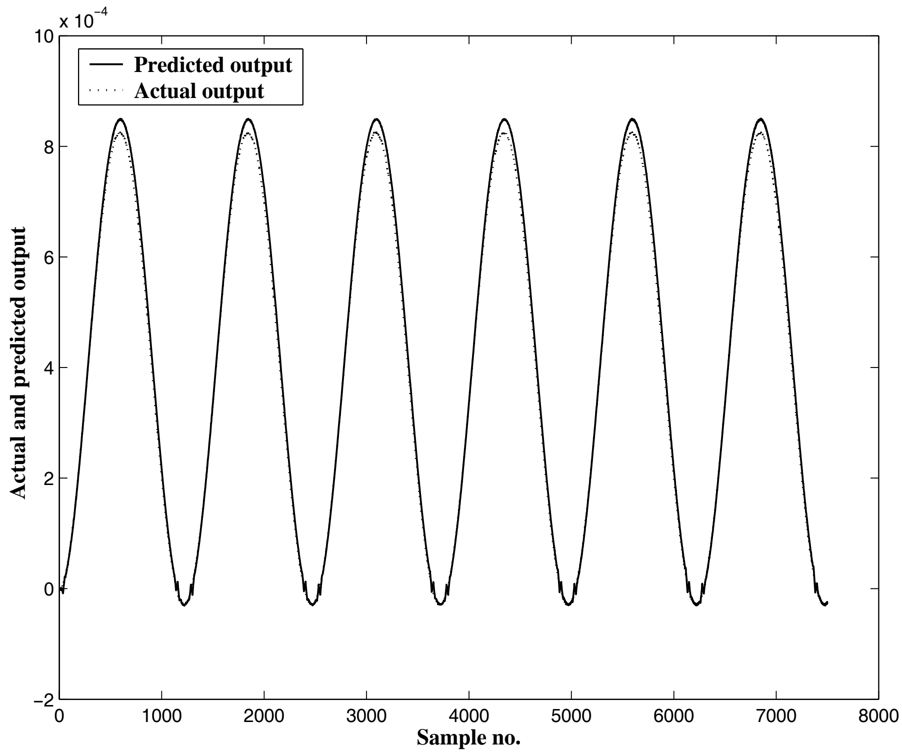

Figure 5.2: Validation with sine wave input

reactor is a complex, continuous function of neutron energy. However, representative results can be obtained by considering single neutron energy group model. The assumption is that generation, diffusion, absorption and leakage of neutrons take place at a single energy. The model also assumes that time variation of neutron flux at all points of the core is identical. This type of model is called point kinetic model [30, 67, 39, 41]. Neutronic power P in a nuclear reactor is given by the point kinetic model, described as below.

$$\frac{dP}{dt} = \left[\rho - \sum_{i=1}^{m_d} \beta_i - \frac{\overline{\sigma}_X X}{\Sigma_a}\right] \frac{P}{l^*} + \sum_{i=1}^{m_d} \lambda_i C_i \tag{5.12}$$

Xenon and Iodine concentrations are given by following two equations.

$$\frac{dI}{dt} = -\left[\lambda_I + \overline{\sigma}_I P\right] I + \gamma_I \Sigma_f P \tag{5.13}$$

$$\frac{dX}{dt} = -\left[\lambda_X + \overline{\sigma}_X P\right] X + \lambda_I I + \gamma_X \Sigma_f P \tag{5.14}$$

Parameters used in (5.12), (5.13) and (5.14) are defined in the begining of the document. It may be noted here that point kinetic model does not apply to situations involving three dimensional effects occurring due to Xenon oscillation and similar transients. Nevertheless, the simulation results based on point kinetic model clearly indicate potential benefits of a wavelet subband model. Extension of the wavelet based model for 3D description of a large power reactor is beyond the scope of this work.

The rate of formation of the delayed neutron precursor of the i^{th}group (representative group) is given by

$$\frac{dC_i}{dt} = \frac{\beta_i}{l^*} n - \lambda_i C_i \tag{5.15}$$

Addition of reactivity depends on reactivity control mechanism and in general can be expressed as a non linear function f of control input v.

$$\frac{d\rho}{dt} = f(v) \tag{5.16}$$

To study reactor kinetics, the system of non linear equations (5.12) - (5.16) are linearized around a steady state operating point $(P_0, C_{i0}, I_0, X_{0,}\rho_o\ )$ to obtain a fifth order state space representation in the following form.

$$\frac{dx}{dt} = Ax + Bu, \quad y = Cx \tag{5.17}$$

where,

Table 5.2: Typical parameter values in a PHWR

$l = 0.00079s$	$\lambda = 0.091s^{-1}$
$\Sigma_f = 1.262X10^{-3}cm^{-1}$	$\Sigma_a = 3.2341X10^{-3}cm^{-1}$
$\sigma_X = 1.2X10^{-18}cm^2$	$\sigma_I = 0.0cm^2$
$\gamma_X = 0.006$	$\gamma_I = 0.0618$
$\lambda_X = 2.1X10^{-5}s^{-1}$	$\lambda_I = 2.878X10^{-5}s^{-1}$

$$x = \begin{bmatrix} \frac{\delta I}{I_0} \\ \frac{\delta X}{X_0} \\ \frac{\delta C_i}{C_{i0}} \\ \delta\rho \\ \frac{\delta P}{P_0} \end{bmatrix}, \quad u = \delta v, \quad y = \delta P$$

and A, B and C are constant matrices. Typical values of physical data for a PHWR is given in Table 5.2 .

Based on the above data open loop transfer function $T(s)$ from control input v to reactor power output $\frac{\delta P}{P_0}$ can be obtained.

$$T(s) = \frac{-0.89s^3 - 0.008s^2 - 2.5X10^{-6s} - 6.5X10^{-11}}{s^5 + 9.6s^4 - 0.004s^3 - 0.0007s^2 + 1.7X10^{-9}s}$$

The transfer function model, mimics an actual nuclear reactor and output data is generated using standard inputs. This Input-output data is used for demonstration of identification and control with wavelets. Note that the poles of the transfer function model are clustered in two regions in the $s-$plane. One cluster having four poles at $s = 0, 2.6X10^{-6}, 8.5X10^{-3}, -8.1X10^{-3}$ is located close to origin. Two of the poles in this cluster are in the right half of s-plane and cause instability. This cluster is responsible for slow dynamic mode of the system. There is one more pole at $s = -9.6$ which contributes to the fast response of the system.

For complete control of the process all poles should be assignable. Let us assume that the following closed loop transfer function $T_c(s)$ satisfying the desired dynamic response is specified by the designer of the control system.

$$T_c(s) = \frac{-0.89s^3 - 0.008s^2 - 2.5X10^{-6}s - 6.5X10^{-11}}{s^5 + 13s^4 - 42s^3 - 58s^2 + 37s + 9}$$

Note that zeros of $T_c(s)$ are at the same locations as those of $T(s)$ whereas poles are shifted to $s = -1, -1, -1, -1, -9$ to get a stable response. Open loop stable pole at $s = -9.6$ is only shifted marginally closer to the other

Table 5.3: Open Loop Response Function

Scale index, j	1	2	3	4	5
a_j	-0.63	0.41	0.56	1.45	1.1
b_j	-0.03	-0.034	-0.077	-0.023	0.015

cluster. All unstable or marginally stable poles are shifted to -1 to improve stability and sharing of control action. To compare the responses of the open loop and the closed loop transfer functions, both the systems given by $T(s)$ and $T_c(s)$ respectively, are excited by the same input $v(t)$, a $[0,\ u_0]$ pulse train. In addition, to simulate measurement noise, an additive, zero mean, white Gaussian noise of Signal-to-Noise-Ratio (SNR) 100 dB is injected at the output. Both input and output are sampled at an interval of 0.2 s.

As explained in Chapter 4 an LTI open loop model with wavelet is derived. Here, idea is to show that a wavelet based model captures pole clusters, rather than individual poles and has a potential to reduce the order of a model very similar to the way time-scale methods [56] and singular perturbation methods [38] do. As input is a sequence of impulses, $db1$ of Daubechies family of wavelets (Haar wavelet) could be chosen to represent the signal in wavelet domain. A good match of the input signals with the Haar wavelet results sparse representation in wavelet domain and reduces number of basis functions used to represent systems from input to output. Wavelet RBIO1.5 is used in the modeling exercise shown in Figure 5.3 for projecting or analyzing the input. As in $db1$ the analyzing scaling function of RBIO1.5 is a box function or box spline of degree zero. The output is projected on RBIO2.4. The choice of wavelets are explained in details in section 4.9. Thresholds λ_u and λ_y are taken as 0.05% of the maximum absolute values of wavelet coefficients of input and output respectively at each scale. A very good match is observed between the predicted output and actual output as shown in Figure 5.3. The estimated model turns out to be AR as expected because location of zeros of $T(s)$ are close to origin. Identified parameters in open loop case are listed in Table 5.3 for input-to-output and output-to-output responses.

Let d^{yo} and d^{uo} be the input and output parameters respectively estimated for the open-loop system. The i^{th} elements of each vector in (3.14) is thus computed using consistent output estimate as given below.

$$\left(W\hat{y}_i^s\right)[k] = d_j^{yo} w_j^{y_i}[k] + d_j^{uo} w_j^{v_i}[k] \tag{5.18}$$

A close look at the distribution of estimated parameters confirm peaking at scale $j = 0$ and at $J = 2$ and above indicating two clusters of eigen values. Note that parameter d_1 is found to be at least an order less compared to

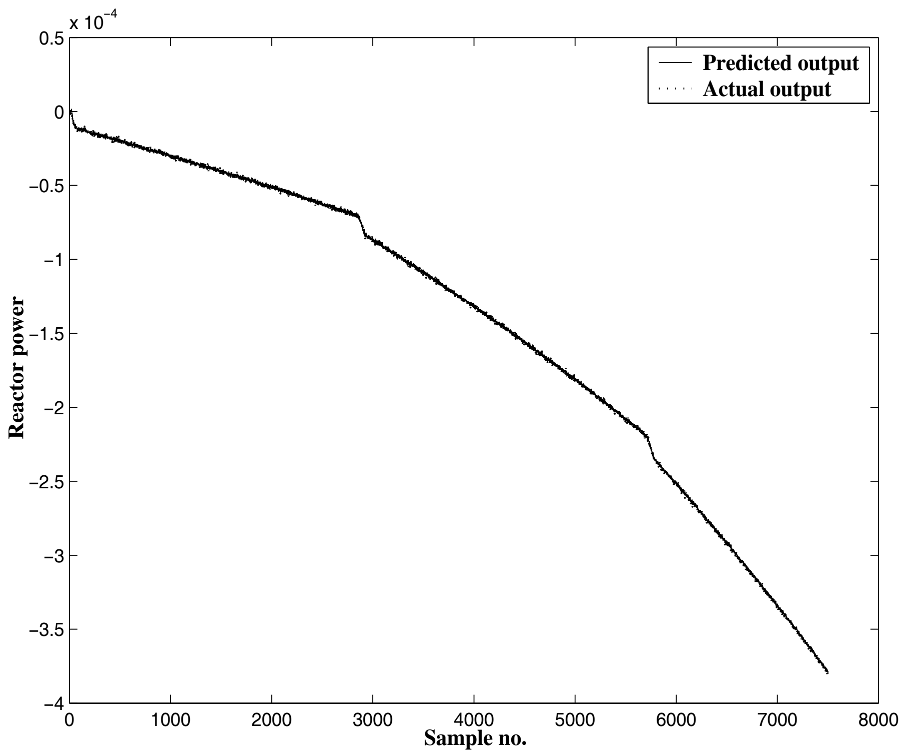

Figure 5.3: Variation of reactor power plotted in open loop case against predicted model output in response to $[0,\ u_0]$ input.

other parameters. Moreover, although the transfer functions are of order 5, working up to scale 4 seems to be sufficient for exposing the clusters.

Output of the closed loop system is the desired response and is used for estimating controller gains. Alternatively, a desired response could be obtained directly from the specification. It may be noted here that the linear control law defined in (5.9) for the wavelet model is memory less as depicted in Figure 5.4. Block D in the figure denotes one step delay and c_j^y are proportionality constants of the full wavelet state feedback controller. The figure shows the simulated plant and the basis structure of the controller or compensator with wavelet basis. It can be seen that the controller uses generalized samples of control input and open loop output as indicated in (5.9).

Let d^{yc} and d^{uc} be the input and output parameters respectively estimated for the close loop system using the same control input (as in open loop case) and the desired output. Hence in closed loop case, the i^{th} elements of each vector in (3.14) is computed as below.

$$\left(\hat{Wy_i^s}\right)[k] = d_j^{yc} w_j^{y_i}[k] + d_j^{uc} w_j^{v_i}[k] \tag{5.19}$$

Looking at the structure of the controller in Figure 5.4, (5.19) can be expressed in terms of open loop parameters as below.

$$\left(\hat{Wy_i^s}\right)[k] = c_j^y d_j^{uo} w_j^{y_i}[k] + c_j^u d_j^{uo} w_j^{v_i}[k] + d_j^{yo} w_j^{y_i}[k]$$

$$\Rightarrow \left(\hat{Wy_i^s}\right)[k] = \left(d_j^{yo} + c_j^y d_j^{uo}\right) w_j^{y_i}[k] + c_j^u d_j^{uo} w_j^{v_i}[k] \tag{5.20}$$

This implies that if predicted output is made to relate to the exogenous input in both open loop and closed loop models in a way such that

$$c_j^u = \frac{d_j^{uc}}{d_j^{uo}} \tag{5.21}$$

scalar gain of the controller would be given by

$$c_j^y = \frac{d_j^{yc} - d_j^{yo}}{d_j^{uo}} \tag{5.22}$$

Loosely speaking, c_j^u can also be thought as fixed part of the controller. Parameters of the closed loop system d_j^{uc}, $d_j^{yc}\,\forall j$ are estimated by consistent prediction of input and output of the closed loop transfer function $T_c(s)$. Both c_j^u as well as c_j^y can be computed using (5.21) and (5.22) by alternate projection algorithm as contributions from input and output are alternatively used for regression. Figure 5.5 shows the validation result comparing closed

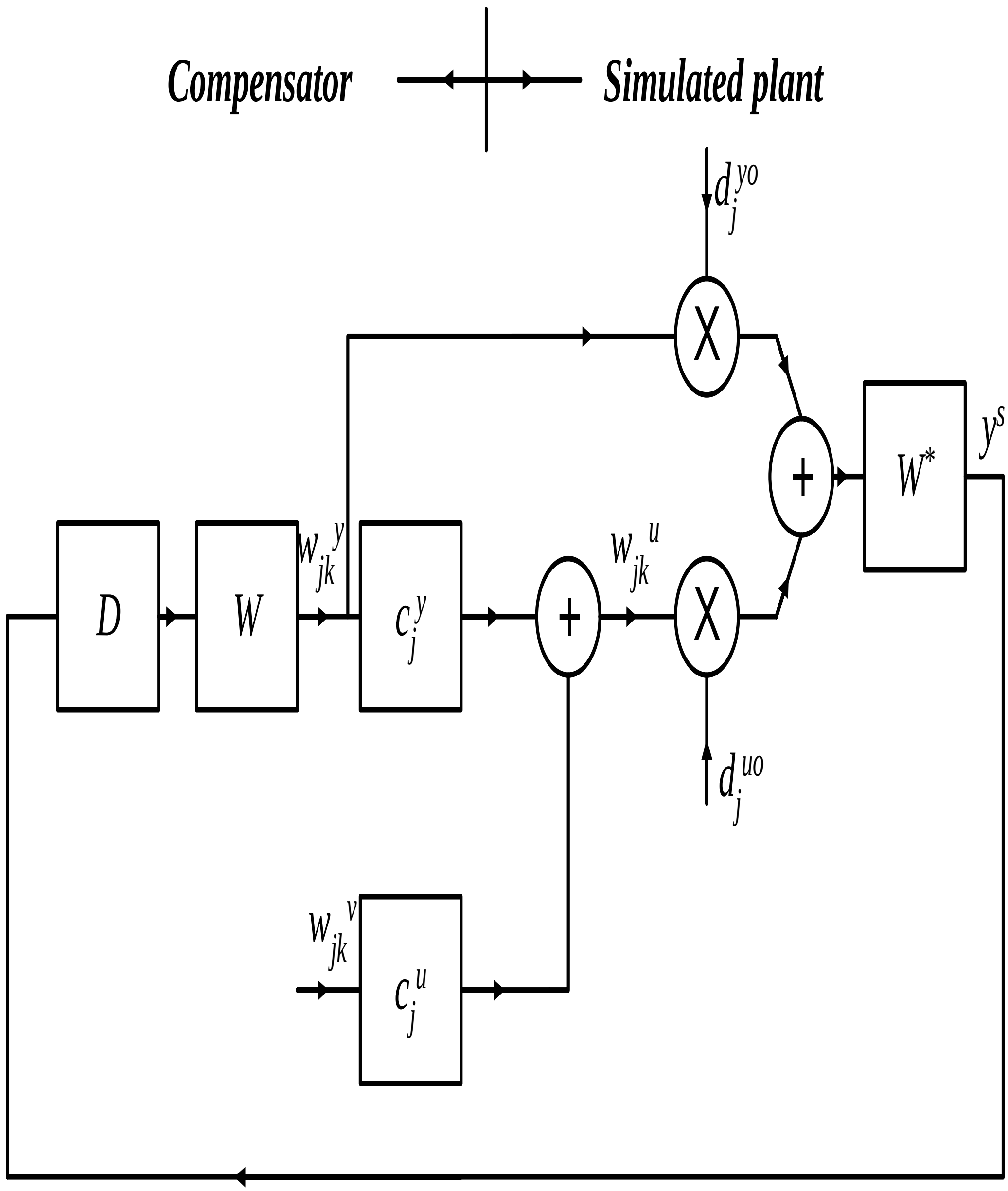

Figure 5.4: Structure of the controller with wavelet basis.

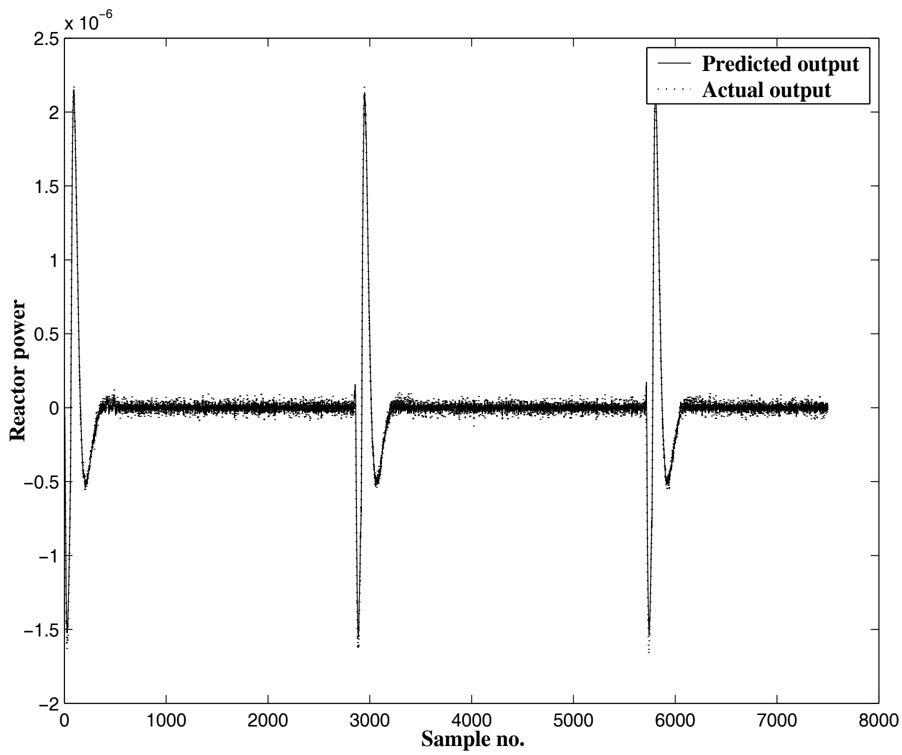

Figure 5.5: Closed loop wavelet state response (reactor power) to [0, u_0] output.

Table 5.4: Closed Loop Response function

Scale index, j	1	2	3	4	5
a_j	-0.47	0.39	0.8	1.04	0.85
b_j	-0.003	0.006	0.007	0.001	-0.001

Table 5.5: Controller Gains

Scale index, j	1	2	3	4	5
Output gain	5.4	0.6	-2.8	17.6	-1.6
Input gain	-0.1	-0.2	0.1	-0.06	-0.07

loop output of the wavelet model with the output of desired transfer function.

Again a nice match is observed between the actual and predicted outputs. The parameters of the model in closed loop case are estimated as earlier for output-to-output response function and are tabulated in Table 5.4. Now, controller gains are computed for each scale j using (5.22) and are tabulated in Table 5.5. A good design of controller would require small gains.

5.4 A note on using the model for real time control applications

In the last section, methodology of designing a controller with a wavelet based model is discissed. It is a common practice to apply wavelet transform on a block of data and the efficacy of the technique is demonstrated working on a data block. However, for implementing a real time control in wavelet domain It is desirable that the state equation (5.1) is written in terms of wavelet coefficients (projections). As has been observed in section 3.2 this would indeed be possible under certain restrictive conditions. And hence a Kalman filter type algorithm for designing an optimum (in presence of noise) wavelet state observer can be formulated for real time control applications.

5.5 Summary

The chapter presents the theory of controller design with wavelet states. The notion of designing a control with output wavelet states is introduced and controllability of the open loop and closed loop system with output wavelet state feedback is studied. Proposition 5.1 presents necessary conditions for wavelet state model to be controllable and observable. It is proved that the closed loop control with wavelet states is pole assignable if open loop system is controllable and observable. The technique is demonstrated with multi scale simulation examples. Simulation model of a large nuclear reactor is considered for design of full wavelet state feedback control. The methodology can be seen as designing a wavelet state observer. It is possible to use the observer states for designing a full state feedback compensator to achieve required closed loop performance. The controller gains can be found from the open and closed loop model parameters. Use of the proposed control methodology in a real time control application is discussed.

Chapter VI

Conclusion

The major contribution of this work comes in the form of designing a computationally efficient alternative to classical least squares minimization problem by penalized local error minimization in wavelet domain. We call this "method of consistent output prediction" and use it for estimation of model parameters. It is shown that estimates are unbiased and bounded. An LTI model by first order estimate of time varying model parameters (*Theorem* 4.1) is derived as a special case. An algorithmic technique for parameter estimation by alternate projection in time and wavelet domain is suggested. The algorithm works on sparse representation in wavelet domain. As an extension a new representation called "wavelet maximum curvature point representation" is proposed as a generalization of existing representations and reconstruction based on WMCPR is validated using an NDT signal.

The other novelty of work is in admitting two different basis functions matching input and output for identification of a system. Spline biorthogonal wavelet basis of different degrees which allow direct weighted addition of projections are used in this work for system modeling. This new class of wavelet basis provides a robust representation of a system with fewer parameters. The process of identification can be viewed as modeling with pre-filtered input and output and gives a model insensitive to noise.

The theory is validated with the help of a case study - identification of the LZCS in a large PHWR. A graphical method to compute process delay is devised and used for finding the delay in control action in LZCS. Results of the iterative alternate projection algorithm suggested for estimating process parameters show excellent match with the experimental data in cross validation. Results of modeling with orthogonal and spline biorthogonal wavelts are compared. The estimated model is validated on the data obtained from a full scale test set up of the LZCS and hence can be considered sufficiently accuarte and insensitive to process and measurement noise usually present in an actual system. However, a numerical study of the effect of noise on identification is included in the work.

Efficiency in the proposed algorithm is achieved by shifting memory to the wavelet basis and getting a low order model with local memory. Intuitively the class of admissible controls, also constrained to be memory-less would be best suited. The theory of controller design with output wavelet state feedback for pole assignment is developed. It is shown that under certain restrictive conditions the closed loop system with output wavelet states will be pole assignable. The technique is demonstrated by designing a control with a simulated point kinetic model of a nuclear reactor.

The theory developed in this work has a huge potential of proliferation in the area of signal and system modeling. For example, the method of consistent estimate could readily be used for learning the system function in a wavelet neural network. One-step-ahead formulation in wavelet domain strongly indicates a development of recursive algorithms as an extension in future.

References

[1] A. Antoniadis, D. Leporini, and J. C. Pesquet. Wavelet thresholding for some classes of non-gaussian noise. *Statistica Neerlandica*, 56 (4):434 –453, 2002.

[2] B. R. Bakshi. Multiscale PCA with application to multivariate statistical process monitoring. *AIChE Journal*, 44(7):1596–1610, 1998.

[3] S. A. Benno and J. F. Moura. Scaling functions robust to translation. *IEEE Transaction on Signal Processing*, 46(12):3269–3281, 1998.

[4] Z. Berman and S. Baras. Properties of multiscale maxima and zero crossing representation:. *IEEE Transaction on Signal Processing*, 41(12):3216–3231, 1993.

[5] S. Björklund and L. Ljung. A review of time-delay estimation techniques. In *Proceedings of the 42nd IEEE conference on decision & control*, 2003.

[6] S. Björklund and L. Ljung. An improved phase method for time-delay estimation. *Automatica*, 45(10):2467–2470, 2009.

[7] M. Breakspear, M. Brammer, and P. A. Robinson. Construction of multivariate surrogate sets from nonlinear data using the wavelet transform. *Physica D*, 182:1–22, 2003.

[8] C. Sidney Burrus, Ramesh A. Gopinath, and Haitao Guo. *Introduction to wavelets and wavelet transforms*. Prentice-Hall International, Inc., 1998.

[9] C. Cai and P. de B. Harrington. Different discrete wavelet transforms applied to denoising analystical data. *J. Chem. Inf. Comput. Sci.*, 38:1161–1170, 1998.

[10] J. F. Carrier and G. Stephanupoulos. Wavelet-based modulation in control-relavant process identification. *AIChE Journal*, 44(2):341–360, 1998.

[11] A. B. Chammas and C. T. Leondes. Pole assignment by piecewise constant output feedback. *International Journal of Control*, 29(1):31–38, 1979.

[12] X.W. Chang and L. Qu. Wavelet estimation of partially linear model. *Computational Statistics and Data Analysis*, 47(1):31–48, August 2004.

[13] C. T. Chen. *Linear system theory and design*. Holt, Rinehart and Winston, Holt-Saunders Japan, 1984.

[14] C. T. Chen and C. A. Desoer. Controllability and observability of composite systems. *IEEE Trans. on Automatic Control*, AC-12(4):402–409, 1967.

[15] K. C. Chou, A. S. Willsky, and A. Benveniste. Multiscale recursive estimation, data fusion, and regularization. *IEEE Trans. on Automatic Control*, 39 (3):464–478, 1994.

[16] D. Coca and S. A. Billings. Non-linear system identification using wavelet multiresolution models. *International Journal of Control*, 74(18):1718–1736, 2001.

[17] Z. Cvetković and M. Vetterli. Discrete time wavelet extrema representation: design and consistent reconstruction. *IEEE Transactions on Signal processing*, 43(3):681–693, March 1995.

[18] I. Daubechies. *Ten lectures on wavelets*. SIAM, 1992.

[19] R. W. Dijkerman and R. R. Majumdar. Wavelet representations of processes and multiresolution stochastic models. *IEEE Trans. on Signal Processing*, 42:1640, 1994.

[20] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. on Inf. Theory*, 41:613–627, 1995.

[21] D. L. Donoho and I.M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika, vol 81, pp. 425-455.*, 81:425–455, 1994.

[22] D. L. Donoho and Thomas P. Y. Yu. Nonliner pyramid transforms based on median-interpolation. *SIAM J. Math. Anal.*, 31 (5):1030–1061, 2000.

[23] Y. Dorfan, A. Feuer, and B. Porat. Model and identification of LPTV systems by wavelets. *Signal processing*, 84(8):1285–1297, 2004.

[24] M.I. Doroslovacki and H. Fan. Wavelet-based linear system modeling and adaptive filtering. *IEEE Transactions on Signal processing*, 44(5):1156–1165, 1996.

[25] S. Farkash and S. Raz. Linear systems in gabor time-frequency space. *IEEE Trans. on Signal Processing*, 42(3):611–617, 1994.

[26] B. Friedland. *Control System Design An Introduction to State Space Method*. Dover Publications, Inc., Mineola, New York, 2005.

[27] L. Geng, T. Zhang, D. Xiao, and J. Song. Wavelet based oe model identification with random missing data. In *Proceedings of the 17th World Congress, The IFAC, Seoul, Korea*, 2008.

[28] R. Ghanem and F. Romeo. A wavelet-based approach for the identification of linear time-varying dynamical systems. *Journal of Sound and Vibration*, 234(4):555–576, 2000.

[29] E. G. Gilbert. Controllability and observability in multivariable control system. *J. S. I. A. M. Control Ser. A*, 2(1):128–151, 1963.

[30] S. Glasstone and A. Sesonske. *Nuclear reactor engineering: reactor design basics, Vol 1.* Kluwer Academic Publishers, 1994.

[31] P. Hall and P. Patil. On the choice of smoothing parameter, threshold and truncation in nonparametric regression by nonlinear wavelet methods. *J. Roy. Statist. Soc. Ser. B*, 58:361–377, 1996.

[32] P. Halmos. *Linear Algebra*. Springer-Verlag Berlin and Heidelberg GM, 1993.

[33] K. Hoffman and R Kunze. *Linear Algebra, 2nd Edition*. Prentice Hall International, 1971.

[34] A. Ikonomopoulos and A. Endou. Wavelet decomposition and radial basis function network for system monitoring. *IEEE Transaction on Nuclear Science*, 45(5):2293–2301, 1998.

[35] A. Juditsky, H. Hjalmarsson, A. Benveniste, B. Delyon, J. SjÖberg L. Ljung, and Q. Zhang. Nonlinear black-box models in system identification: Mathematical foundations. *Automatica*, 31(12):1725–1750, December 1995.

[36] R. E. Kalman. A new approach to linear filteringand prediction problem. *Trans. ASME, J. Basic Eng.*, Series 82D:16–26, 1960.

[37] R. E. Kalman. Mathematical description of linear dynamical systems. *J. S. I. A. M. Control, Ser. A*, 1(2):152–192, 1963.

[38] P. V. Kokotovic, R. E. OMalley, and P. Sannuti. Singular perturbations and order reduction in control theoryan overview. *Automatica*, 12 (2):123–132, 1976.

[39] J. R. Lamarsh and A. J. Baratta. *Introduction to nuclear engineering*. Prentice Hall, 2001.

[40] W. Lawton, Y. Yu, S. L. Lee, S. Tan, and J. Vandewalle. Wavelet based modeling of nonlinear systems. In *Proceedings of the International workshop on advanced black-box techniques for non linear modeling with time-series prediction competition*, 1998.

[41] E. E. Lewis. *Fundamentals of Nuclear Reactor Physics*. Academic press, 2008.

[42] L. Ljung. On the estimate of transfer functions. *Automatica*, 21 (6):677–696, 1985.

[43] L. Ljung. *System Identification-Theory for the user*. Prentice Hall PTR, Upper Saddle River, New Jersey, USA, second edition, 1999.

[44] D. G. Luenberger. Observers for multivariable systems. *IEEE Trans. on Automatic Control*, AC-11:190–197, 1966.

[45] D. G. Luenberger. An introduction to observers. *IEEE Trans. on Automatic Control*, AC-16:596–602, 1971.

[46] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, second edition, 1999.

[47] S Mallat. Zero-crossings of a wavelet transform. *IEEE Trans. on Information Theory*, 37 (4):1019–1033, 19991.

[48] S. Mallat and S Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on PAMI*, 14(7)(3):710–732, March 1992.

[49] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE transaction on pattern analysis and machine intelligence*, 11(7):674–693, 1989.

[50] O. Michailović and A. Tanenbaum. On approximaton of smooth functions from samples of partial derivatives with applications to phase unwrapping. *Signal Processing*, 88(2):358–374, 2008.

[51] B. C. Moore. Principal component analysis in linear systems: Controllability, observability and model reduction. *IEEE Transaction on Automatic Control*, AC-26(1):17–32, 1981.

[52] S. Mukhopadhyay, U. Mahapatra, A. K. Tangirala, and A. P. Tiwari. Spline wavelets for system identification. In *Proceedings of DYCOPS 2010*, 2010.

[53] S. Mukhopadhyay and G. P. Srivastava. Characterisation of metal loss defects from magnetic flux leakage signals with discrete wavelet transform. *NDT & E International*, 33:57–65, 2000.

[54] S. Mukhopadhyay and A. P. Tiwari. Characterization of ndt signals: Reconstruction from wavelet transform maximum curvature representation:. *Signal Processing*, 90(1):261–268, 2010.

[55] S. Mukhopadhyay and A. P. Tiwari. Consistent output estimate with wavelets: An alternative solution of least squares minimization problem for identification of the lzc system of a large phwr. *Annals of Nuclear Energy*, 37(7):974–984, Accepted for publication.

[56] D. S. Naidu and A. K. Rao. *Singular Perturbation Analysis of Discrete Control Systems (Lecture Notes in Mathematics).* Springer-Verlag, Berlin, Germany, 1985.

[57] T. Thao Nguyen and M. Vetterli. Deterministic analysis of oversampled a/d conversion and decoding improvement based on consistent estimates. *IEEE Transaction on Signal Processing*, 42(3):519–531, 1994.

[58] G. Olsson. *Modelling and Identification of a Nuclear Reactor in System Identification: Advances and Case Studies , Mathematics in Science and Engineering, 126.* Academic Press, N.Y, 1976.

[59] H. M. Paiva, R. Kawakami, and H. Galvao. Wavelet-packet identification of dynamic systems in frequency subbands. *Signal Processing*, 86:2001–2008, 2006.

[60] S. C. Patwardhan, S. Manuja, S. Narasimhan, and S. L. Shah. From data to diagnosis and control using generalized orthonormal basis filters, part ii: Model predictive and fault tolerant control. *J. Process Control*, 16:157–175, 2006.

[61] S. C. Patwardhan and S. L. Shah. From data to diagnosis and control using generalized orthonormal basis filters, part i: Development of state observers. *J. Process Control*, 15:819–835, 2006.

[62] J.C. Pesquet, H. Krim, and H. Carfantan. Time invariant orthonormal wavelet representations. *IEEE Transaction on Signal Processing*, 44(8):1964 – 1977, 1996.

[63] R. Pintelon and J. Shoukens. Frequency domain system identification with missing data. *IEEE Trans. on Automatic Control*, 45(2):364–369, 2000.

[64] G. D. Reddy, B. Bandyopadhya, and A.P. Tiwari. Multrate output feedback based sliding mode control for a large phwr. *IEEE Transactions on Nuclear Sciences*, 54(6):2677 – 2686, December 2007.

[65] M. S. Reis. A multiscale empirical modeling framework for system identification. *Journal of Process Control*, 19:1546–1557, 2009.

[66] S. Sanyal, E. J. Perreault S. L. Kukreja, and D. T. Westwick. Identificaton of linear time varying systems using basis persuits. In *Proceedings of the 2005 IEEE, Engineering in Medicine and Biology 27th annual conference, shanghai, China*, September 1-4, 2005.

[67] M. A. Schultz. *Control of nuclear reactors and power plants.* McGraw-Hill Book Company, Inc., 1955.

[68] X Shan and J. B. Burl. Continuous wavelet based time-varying system identification. *Signal Processing*, 91(6):1476 – 1488, 2011.

[69] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.Y. Glorennec, H. Hjalmarsson, and A. Juditsky. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, December 1995.

[70] O. J. M. Smith. Closer control of loops with dead time. *Chemical Engineering Progress*, 53 (5):217–219, 1957.

[71] T. Soderstrom and P. Stoica. Instrumental variable methods for system identification. *Circuits systems signal processing*, 21(1):1–9, 2002.

[72] G. Strang. *Linear algebra and its applications*. Cengage Learning, 2006.

[73] G. Strang and T. Nguyen. *Wavelets and filter banks*. Wellesley-Cambridge Press, 1997.

[74] N. Sureshbabu and J. A. Farell. Wavelet based system identification for nonlinear control. *IEEE Transaction on Automatic Control*, 44(2):412–417, 1999.

[75] A. H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional brownian motion. *IEEE Transactions on Information Theory*, 38(2):904–909, 1992.

[76] A. P. Tiwari. Modeling control and simulation of 540 MWe PHWR and development of liquid zone control system. *Nuclear India*, 39(5-6), 2005.

[77] A. P. Tiwari, B. Bandopadhyay, and H. Warner. Spatial control of a large PHWR by piecewise constant periodic output feedback. *IEEE Transaction on Nuclear Science*, 47(2):389–402, 2000.

[78] Y. Tsaig and D. L. Donoho. Extension of compressed sensing. *Signal processing*, 86(3):549–571, 2006.

[79] M. K. Tsatsanis and G. B. Giannakis. Time-varying system identification and model validation using wavelets. *IEEE Trans. on Signal Processing*, 41 (12):3512–3523, 1993.

[80] M. Unser. Ten good reasons for using spline wavelets. In *Proceedings SPIE vol. 3169, Wavelets applications in signal and image procesing V*, 1997.

[81] M. Unser. Sampling-50 years after shannon. *Proceedings of the IEEE*, 88(4):569–587, April 2000.

[82] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. on Signal Processing*, 42(11):2915–2925, 1994.

[83] M. Vetterli and C. Harley. Wavelets and filter banks: theory and design. *IEEE Trans. on Signal Processing*, 40(9):2207–2232, 1992.

[84] M. Vetterli and J. Kovacevic. *Wavelets and subband coding*. Prentice Hall PTR, Englewood Cliffs, New Jersey, 1995.

[85] S. K. Wattamwar, S. Weiland, and T. Backx. Identification of low-order parameter-varying models for large-scale systems. *Journal of process control*, 20(2):158 – 172, 2010.

[86] H. L. Wei and S. A. Billings. Identification of time -varying systems using multiresolution wavelet models. *International journal of system science*, 33(15):1217–1222, 2002.

[87] G. W. Wornell. A Karhunen-Loeve-like expansion for 1/f process via wavelets. *IEEE Trans. on Inform. Theory*, IT-36(4):859, 1990.

[88] Y. Zhang, T. T. Lie, and C. B. Soh. Consistent parameter estimation of systems disturbed by correlated noise. *IEE Proc.-Control Theory Appl.*, 144(1):40–44, 1997.

[89] H. Zhao and J. Bentsman. Wavelet-based identification of fast linear time-varying systems using function space methods. In *Proceedings of the American Control Conference, Chicago*, pages 939–943, 2000.

[90] H. Zhao and J. Bentsman. Biorthogonal wavelet based identification of fast linear time-varying systems - part i: system representations. *Journal of dynamic systems, measurement, and control*, 123(4):585–592, 2001.

[91] H. Zhao and J. Bentsman. Biorthogonal wavelet based identification of fast linear time-varying systems - part ii: algorithms and performance analysis. *Journal of dynamic systems, measurement, and control*, 123(4):593–600, 2001.

Appendix A

Alternate projection for reconstruction of signals from wavelet maximum curvature point representation

An iterative algorithm, similar to those based on multi-scale edges, for the reconstruction of signal from maximum curvature point representation in wavelet domain is proposed. From the application point of view, it is desired that the reconstructed signal has all the properties of the original signal as described by the representation.

A.1 Maximum curvature point representation

A new scheme of signal reconstruction by alternating projections is proposed based on WMCPR. WMCPR is a convex representation as it includes the wavelet extrema points giving as many inequality constraints. Curvature of a function $w^x(t)$ at a sample point k is defined as

$$C\left(w^x\right)\big|_{t=k} = \frac{\left|\frac{d^2 w^x}{dt^2}\big|_{t=k}\right|}{\left[1+\left(\frac{dw^x}{dt}\big|_{t=k}\right)^2\right]^{\frac{3}{2}}} \tag{A.1}$$

$w^x(t)$ is assumed continuous around k. Numerical computation of first and second order derivatives from the sample points are needed for determining curvature at k. Maximum curvature points are detected as first order peaks of the curvature signal. Insignificant peaks may be ignored for denoising using threshold.

Inclusion of maximum curvature points is an attempt to include more significant points in the representation over and above the extrema points. It can also be interpreted as inclusion of left and right boundaries of a peak for capturing its shape. Maximum curvature representation is, however, a non-linear representation. To enforce linearity the curvature can be approximated as

$$C^1(w^x) = \left|\frac{d^2 w^x}{dt^2}\right| \tag{A.2}$$

At the extrema points, $\frac{dw^x}{dt}$ is exactly equal to zero. C^1 is now a linear operator giving some extra equality constraints at approximate maximum curvature points. As extrema points are included in the representation, the inequality relations shall be the same as those for extrema representation. The approximate wavelet maximum curvature representation (AWMCR) now can be treated as an inherently bounded AQLR as defined in [4]. All the results proved in [4] related to different properties including stability shall be valid for AWMCR also.

An alternative interpretation can be given in continuous case. A smoothing function $\varphi(t)$ is one which has integral equal to 1 and converges to 0 at infinity. It is also assumed that $\varphi(t)$ and its dilated versions are at least thrice differentiable. Then $\frac{d\varphi(t)}{dt}$, $\frac{d^2\varphi(t)}{dt^2}$ as well as $\frac{d^3\varphi(t)}{dt^3}$ can be considered to be wavelets as

$$\int_{-\infty}^{+\infty} \frac{d^n \varphi\left(t\right)}{dt^n} dt = 0,\ n = 1, 2, 3 \tag{A.3}$$

Hence, the approximate wavelet maximum curvature representation with the wavelet transform of x at scale s, given by

$$w^x\left(t\right) = s \frac{d\left(x \star \varphi_s\right)\left(t\right)}{dt} \tag{A.4}$$

can be considered as wavelet modulus maxima representation using wavelet $\frac{d^3\varphi_s(t)}{dt^3}$. Thus for proper choice of wavelet, maxima based scheme shall give similar result to that based on maximum curvature. Since a definite relationship exists between the peak in time domain and the evolution of wavelet modulus maxima, hence more than intuitive, there also exists an indirect relationship between the peak feature and wavelet maximum curvature point representation. It may be noted that AWMCR as discussed here is a representation at maximum curvature points of the Continuous Wavelet Transform (CWT). The discussion on use of CWT shall be limited to this extent. The discrete implementation and design of the structures of wavelet filter banks is beyond the scope of this work.

In case the values of the function at maximum curvature points instead of the curvature value are used then the resulting representation shall be quasi linear. We shall call this representation wavelet maximum curvature point representation. Wavelet maximum curvature point representation is an inherently bounded AQLR.

Let us define the representation R_m for WMCPR as below

$$(L_{0mc}(W_jx)_{1\leq j\leq J}, V_0(W_jx)_{1\leq j\leq J}, V_0(U_Jx)) = (L'(w^x), V'(w^x)) \tag{A.5}$$

Now it is further argued that

$$\|W_j x^c\| \leq K_1 \|V_{max} w^x\| \leq K_2 \|V' w^x\|, K_1, K_2 > 0 \tag{A.6}$$

where V_{max} is the set of multi-scale maxima values in maxima representation at arguments given by

$$L_{0mm}(w^x) = \{k : |w^x[k+1]| < |w^x[k]|, \ |w^x[k-1]| \leq |w^x[k]|\} \tag{A.7}$$

The first inequality is proved in Theorem 1 in [4] and the second inequality holds because $L_{0mm} \subset L_{0e} \subset L_{0mc}$. This implies

$$\|W_j x^c\| \leq K \|V' w^x\|, K > 0 \tag{A.8}$$

$$\left\|w^{x^c}\right\| \leq K \|V' w^x\|, K > 0 \tag{A.9}$$

We allow, in this work, non-linear operations such as thresholding [21, 20] and band (read scale) rejection of wavelet maximum curvature point representation to obtain $L\ \ (w^x) \neq L(w^x)$ prior to reconstruction. Let the reconstruction set of such signals be $P(R'_m w^x)$.

Assumption $A.1$

Signal modeling based on the reconstructed signal in $P(R'_m w^x) \neq P(R_m w^x)$ is valid, if it is ensured that there is no loss of useful information due to thresholding and band rejection.

This of course, calls for proper choice of threshold. As such, the reconstruction from a representation is not unique as null space of V' is not empty. Hence in the context of signal modeling, the following theorem is necessary for validation of the reconstruction set.

Theorem $A.1$

A signal in the reconstruction set $P(R'_m w^x)$ reconstructed based on wavelet maximum curvature point representation associated with specified non-linear operation such as thresholding and band rejection removing insignificant wavelet maximum curvature points, contains complete information for validation.

The tacit assumption as stated in Assumption $A.1$ is the non-linear operations are defined to retain the significant information. Although it sounds obvious a proof by constructive argument shall be given. In the extreme case if $M(W_j x)_{1 \leq j \leq J+1} = \{\emptyset\}$ then the reconstructed signal is also identically equal to zero. Even in the absence of useful signal, the conditions mentioned above are not encountered in general due to presence of noise. However, they can be forced by non-linear operations in the wavelet domain.

While analyzing NDT signals all the scales are rarely exhausted [54]. By virtue of Theorem $A.1$ it is permitted to do so. By assumption, the peaks in time contain complete information for characterization. Hence, the theorem can be proved.

Proof of Theorem A.1

If the signal reconstructed from the null space of $\Im$ does not contain any peak then any signal in $P\left(R'_m w^x\right)$contains complete information for characterization. Now indeed, the signals in the null space of wavelet maximum curvature point representation associated with specified non-linear operations do not contain any point where curvature is maximum, *i.e.* the signals are monotonic with constant slope. As inverse wavelet transform is a linear operation, the signal reconstructed from the kernel of the representation cannot contain any peak. The reverse, however, is not true as we allow quasi linear and non-linear operations in the forward direction.

A.2 Reconstruction by alternate projection from WMCPR

WMCPR is a set of convex constraints on the signal. The reconstruction from the representation recovers $x(t)$, the signal in time by inverse wavelet transform. The closure of the reconstruction set of $F = Wx$ is given by

$$\Re = \Im \bigcap \overline{\Im} \bigcap \left(\bigcap_{jk} N_{jk} \right) \tag{A.10}$$

where $\Im$ is the range of the wavelet transform, $\overline{\Im}$ is the closed convex set of all $G \in l_2(I)$ such that $C(F) = C(G)$ for all k, which are maximum curvature points across all scales and $N'_{jk}s$ are only applicable to those points k where curvature exceeds a specified threshold. N_{jk}, $k \epsilon K, j \epsilon J$ are the set of convex constraints given by (A.12). With convex constraints on $\overline{\Im}$ a subspace in $l_2(I)$, alternating projection of any initial point $F^0 \epsilon l_2(I)$ onto time and wavelet bases, shall converge to a consistent estimate in $P\left(R'_m w^x\right)$. The initial value $F^0 \epsilon l_2(I)$ can be given by

$$F_k^0 = V\left(w^x\right)\ if\ k = L'\left(w^x\right) \tag{A.11}$$

and

$$F_k^0 = 0\ otherwise$$

A.3 Denoising by thresholding

For denoising, a dual thresholding scheme is proposed. The basic assumption is that the noise is not only present in a particular band of frequency but also at all frequencies and it can be removed by thresholding. Original signal is first decomposed using UDWT. Curvatures of UDWT coefficients at each scale are computed and maximum curvature point representation is obtained at the following points.

$$L_{0mc} = \{k : C(w^x[k+1]) < C(w^x[k]), C(w^x[k-1]) \le C(w^x[k]), C(w^x[k]) \ge C_{th}\} \tag{A.12}$$

(A.12) contains curvature thresholding in addition to peak detection. Insignificant curvatures in the representation are ignored and coefficients in the UDWT decomposition are zeroed by non-linear hard thresholding operations. Hard thresholding keeps those coefficients in the representation which are more than the specified threshold and cuts those off which are lower than the threshold. This is first level of thresholding.

In the second level, signal power thresholding is used for denoising. The dual thresholding scheme proposed is more robust compared to amplitude thresholding as it takes into account the shape of a peak. The left and right boundaries of a peak are used to compute the signal power of a peak. The boundaries of a peak are already available as maximum curvature points and need not be computed separately. Recognizing that a peak is also detected as a maximum curvature point, if l and r are left and right boundaries (also at the maximum curvature points) on both sides of the extremum (peak) the signal power of the peak is defined as

$$P_k = \frac{\sum_{k=l}^{r} (w^x[k])^2}{(l_k - r_k + 1)} \tag{A.13}$$

and the signal power thresholding is given by

$$w^x[k] = w^x[l] + \frac{k-l}{r-l}(w^x[r] - w^x[l]) \;\; if \; P_k \le P_{th} \tag{A.14}$$

and

$$w^x[k] = w^x[k] \;\; if \; P_k$$

Level of threshold P_{th} is decided by estimating median absolute deviation of UDWT coefficients at the finest scale. It is assumed that the finest scale mainly comprises of noise. One method to determine P_{th} could be by deriving standard deviation of noise (σ_n) (by estimating maximum absolute deviation of the UDWT coefficients at the finest scale) and computing P_{th} as a function of σ_n. One can choose to reject UDWT coefficients at some scales if all the coefficients at those scales are considered to be noise.

A.4 An example application in NDT

In NDT applications, defects are detected and characterized by the parameters of peaks in the sensor output signals. Typically the amplitude, shape and location of a peak feature would have information regarding shape and size of a defect. Classification of the features based on the parameter vectors and associating a shape and size template to the class is the ultimate objective in any NDT problem.

Again as the data size increases it becomes necessary to characterize the defects in unsupervised fashion. On line compression of data reduces storage requirement and makes data handling less cumbersome also allowing unsupervised characterization. The technique proposed in this work could be used for characterization and compression of magnetic flux leakage (MFL) signal. MFL technique is frequently used in in-line inspection (ILI) tools for detection and characterization of metal loss defects. Any metal loss in the pipe wall causes significant geometrical deformation and acts as an area of increased magnetic reluctance. This causes a local perturbation of the magnetic field distribution [53]. Measurement of leakage flux density provides sizing and contour information. Simulation results for demonstrating denoising and reconstruction are demonstrated. The wavelet maximum curvature point representation contains the complete information for characterization in fewer coefficients. Hence it can be used for compression of MFL data. The technique could also be effectively used in other defect characterization problems as well as for biomedical applications. The technique could be used for sparse representation of systems as well.

With wavelets, as the frequency resolution, f, increases with frequency, time or space resolution, t, reduces. This property of UDWT is particularly useful for the analysis of NDT signals. A common feature of NDT signal is the appearance of sharp peaks (having high frequency components locally at the edges of defects), closely spaced in time or space (due to defects having small length or for two defects with trailing edge of one very close to the leading edge of the other). As t becomes finer at higher frequency it is easier to detect these sharp peaks with wavelets.

Figure A.1 shows an artificially generated signal $S1$ as could be picked up by an MFL sensor placed radially under a set of defects in presence of additive noise. A synthetic signal is used to validate the proposed technique. Radial component of flux density changes polarity under a defect. Unipolar peaks in $S1$ could indicate beginning and end of a thicker section. The exact location and amplitude of peaks are corrupted by noise. This is a typical representation of MFL signal. The exact method of synthesis is trivial and hence omitted. It is possible to obtain a complete wavelet representation using a filter-bank h_j and g_j, $j = 1, 2, \ldots, J$ by adopting what popularly known as $bior2.2$ filter bank. In general it is possible to use $bior2.x(x = 2, 4, \ldots)$ filter banks as indicated in Section (2.3). The maximum curvature point representation at scales 1–6 after first level of curvature thresholding given by (A.12) is shown in Figure A.2. Low amplitudes maximum curvature

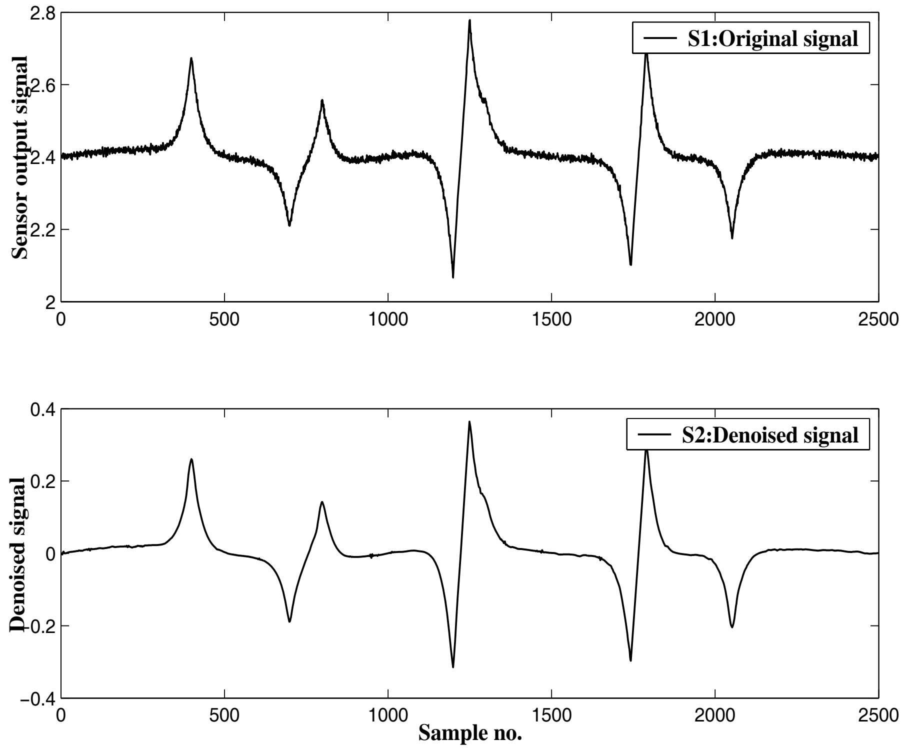

Figure A.1: Artificially generated signal simulating MFL signal over pipe features and defects and result of denoising using maximum curvature point representation.

points at finer scales are predominantly due to noise. The numbers of points at finer scales are more compared to coarser scales. However, it is clearly seen that the information content is more at coarser scales.

The sparse representation of information indicates that the information can be stored in compressed form. However, thresholding makes the compression scheme a lossy one. The loss is determined by the levels of thresholding. Search of optimum thresholds is problem dependent. A formal search can be designed in the application domain minimizing a metric in the space of defect features.

The low pass filtered signal at scale J, $U_J x$ is not used for reconstruction (also not shown in Figure A.2), as the peaks due to metal loss defects are insignificantly small at this scale with a specified thresholding condition. Theorem $A.1$ justifies the rejection of the low pass filtered signal at scale J, $U_J x$. The desired denoised signal can be directly obtained by alternate projection as described in Section (A.2). High frequency noise at the finer scales and the low frequency undulations at scale J are removed in the de-

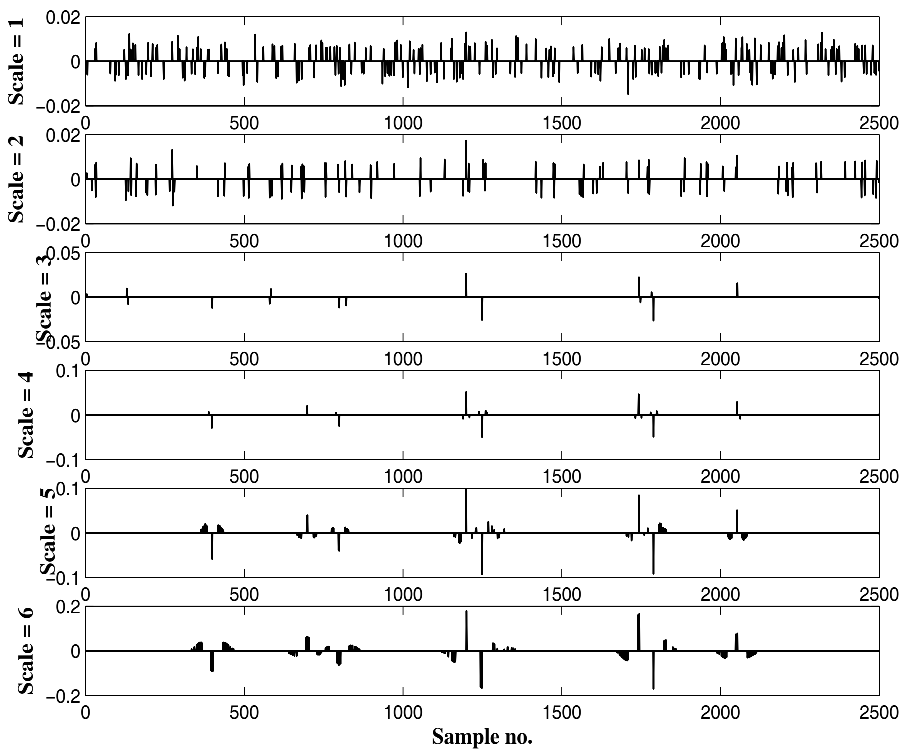

Figure A.2: Maximum curvature point representation of the signal in Fig. 5.1.

noised signal. Noise at other scales is removed by signal power thresholding as described in (A.14). The reconstructed denoised signal using $J = 6$, after 20 iterations is also shown in Figure A.1 as S2. The DC shift in the denoised signal is due to rejection of $W_{J+1}x$ as explained earlier. It may be noted that location and relative value of the peaks over the base level have been preserved in the reconstructed signal and hence characterization based on the reconstructed signal shall be very accurate.

Figure A.3 shows similar results for the actual MFL signal acquired by a hall-effect sensor over three slot type defects (marked as $D1, D2$ and $D3$) of different dimension created on a pipe spool. The pipe spool with defects is joined to other spools by flanges on both sides (marked as $F1$ and $F2$) and supported by metallic supports ($S1$ and $S2$). As mentioned earlier, high frequency noise and the low frequency undulations are removed in the denoised signal. Noise at other scales is removed by signal power thresholding. The denoised signal has been reconstructed from maximum curvature point representation using $J = 6$, after 20 iterations. The distinct signatures of different pipe features can be clearly seen in the reconstructed denoised signal. The location and amplitude of peak features are also preserved in the reconstructed signal.

A.5 Summary

Data denoising and signal characterization based on wavelet maximum curvature representation, with a novel dual thresholding scheme has been presented here. The number of scales to which a signal needs to be decomposed using wavelet transform depends on the concentration of significant information in frequency. Theorem $A.1$ proves that for specific applications, the part of the signal that does not contain significant information can be rejected using curvature based thresholding scheme. The scheme can as well be viewed as a generalization of the similar schemes already existing for signal reconstruction from sparse representations. The technique was tested on the synthetic as well as actual Magnetic Flux Leakage (MFL) data. Characterization of defects based on the reconstructed signal was proper and the results were found satisfactory. The mapping of signal parameters to the defect parameters is however application specific. The work strongly indicates the possibility of characterizing defects by inverse mapping (signal features to defect parameters) directly (without reconstruction) from the representation in wavelet domain since it is proved that the representation contains complete information for characterization.

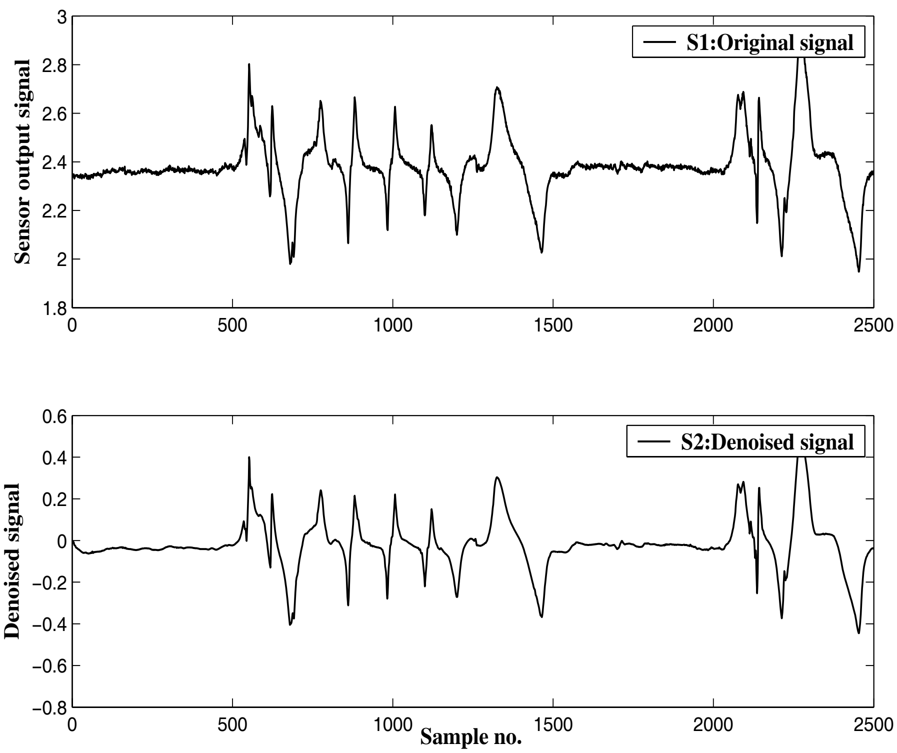

Figure A.3: Denoising of actual MFL signal from three defects in a pipe section using maximum curvature point representation.

Appendix B

Estimation of process time-delay

All industrial processes have inherent time-delays due to process dynamics and instrumentation. For example, transport lags in piping, delays in the instrument response etc can significantly contribute to time-delay in responses. For control design, Smith predictor invented by O. J. M. Smith in 1957 [70] is used as a predictive controller for a system with pure time delay. It is a common practice to model time-delays with higher order dynamics, although in principle, time-delays are different from system order. And it is impractical and inappropriate to fit a higher order parametric model to account for the time-delay. Due to complexity of any industrial process, estimation of time-delay of a process often turns out to be easier from input-output data rather than from the physics of the process. However, the measurement of input and output could be noisy and problem of Time-Delay Estimation (TDE) needs to be addressed in a statistical framework. Although TDE has been a much studied problem [5, 6], yet a clear agreement for the best method does not exist. This appendix specifically focuses on time domain approximation methods and proposes a new approximation method in time domain.

B.1 Estimation of process delay from noisy measurements

Let the impulse response of a process having $u(t)$ as input and $y(t)$ as output be $h(t)$ with unknown process delay d. Let us assume that the output measurement is corrupted with measurement noise $w(t)$ which is zero mean and uncorrelated with input. The output $y(t)$ can be expressed as

$$y(t) = u(t-d) \star h(t) + w(t) \tag{B.1}$$

This implies,

$$\int_{-\infty}^{+\infty} y\,(t+\tau)\,u\,(\tau)\,d\tau = \int_{-\infty}^{+\infty} h\,(t) \star u\,(t+\tau-d)\,u\,(\tau)\,d\tau + \int_{-\infty}^{+\infty} w\,(t+\tau)\,u\,(\tau)\,d\tau \tag{B.2}$$

As the noise is zero mean and uncorrelated with input the second term on the right hand side in (B.2) vanishes. Hence, in terms of cross correlation function $A_{yu}(t)$ and auto correlation function $A_{uu}(t)$, (B.2) can be expressed in time lag (t) – correlation amplitude (A) plane ($t - A\,plane$) as follows.

$$A_{yu}(t) = h(t) \star A_{uu}(t-d) \tag{B.3}$$

In case the system is excited by pure white noise having signal power λ, $A_{uu}(t)$ becomes an impulse and (B.3) becomes

$$A_{yu}(t) = \lambda h(t-d) \tag{B.4}$$

It can be seen from (B.4) that the impulse response of the system can be estimated from the cross correlation function and time-delay can be measured by the delay for the impulse response to start. For low Signal to Noise Ratio (SNR) the estimated impulse response (estimated from input output data) could be noisy due to deviation of noise pattern from the assumed model. And hence it could be difficult to find the exact start of the impulse response.

B.2 Concurrent cross correlation method

An elegant technique insensitive to noise is devised in this work for estimating time-delay of a process. It can be seen that if the process is excited by any input other than the white noise, the cross correlation function will be given by (B.3). It is always possible to find some inputs (different from each other) such that the cross correlation functions concur at a fixed point in the first quadrant of $t - A$ plane. To start with let us assume that indeed the cross correlation functions due to two different inputs u_1 and u_2 intersect at a fixed point $t = t_1$. From (B.3) following can be written.

$$\int_{-\infty}^{+\infty} A_{u_1 u_1}(t_1 - d + \tau)\, h(\tau)\, d\tau = \int_{-\infty}^{+\infty} A_{u_2 u_2}(t_1 - d + \tau)\, h(\tau)\, d\tau \tag{B.5}$$

Evidently, $h(t-d)$, the cross correlation function with pure white noise input will also pass through the same point. As time-delay is a parameter which is independent of dynamic modes of the system, time-delay d is then expected to be given by the abscissa of this point of concurrence. Substituting $t_1 = d$,

$$\int_{-\infty}^{+\infty} A_{u_1 u_1}(\tau)\, h(\tau)\, d\tau = \int_{-\infty}^{+\infty} A_{u_2 u_2}(\tau)\, h(\tau)\, d\tau \tag{B.6}$$

It can be seen from (B.6) that a condition for cross correlation functions to intersect at a point with abscissa $t = d$ is that inputs should have same area under the curve $A_{uu}(\tau)h(\tau)$. The proposed method does not depend on any thresholding scheme for detection of the start of cross correlation curve making it immune to choice of threshold. Hence there is no need for estimating uncertainty as required by some methods. Again, as the method of estimating time delay implicitly looks for a definite point (point of concurrence) obtained by intersection of minimum two cross correlation curves, the estimate becomes to a great extent insensitive to noise.

B.3 Simulation example

The proposed technique of time delay estimation by concurrent cross correlation method is demonstrated on simulated examples as well as on actual data collected from LZCS full scale model.

In this section two simulation examples are presented.

1. Simulation of a first order system having pole at -1 with 1 second time-delay.

2. Simulation of a second order system having poles at 0 and -1 with 2 seconds time-delay.

Simulation example 1: In the first simulation example following transfer function is considered.

$$T(s) = \frac{e^{-s}}{s+1} \tag{B.7}$$

Input is excited by a band limited white noise with auto correlation function shown dotted in Figure B.1. The cross correlation function is plotted in red. Ideally for the system the impulse response is an exponential function. It can be seen that the cross correlation function also shows exponential behavior. The deviation from the ideal exponential behavior is due to the fact that the input is not a pure white noise. The estimated impulse response given by the cross correlation function starts at $t = 1$ second. So, the time-delay of 1 second can be estimated from the start of the cross correlation function.

Simulation example 2: For the second simulation example let us consider a second order linear time invariant system with transfer function having time-delay of 2 seconds which in amplitude is significant when comparable with process time constant.

$$T(s) = \frac{e^{-2s}}{s(s+7)} \tag{B.8}$$

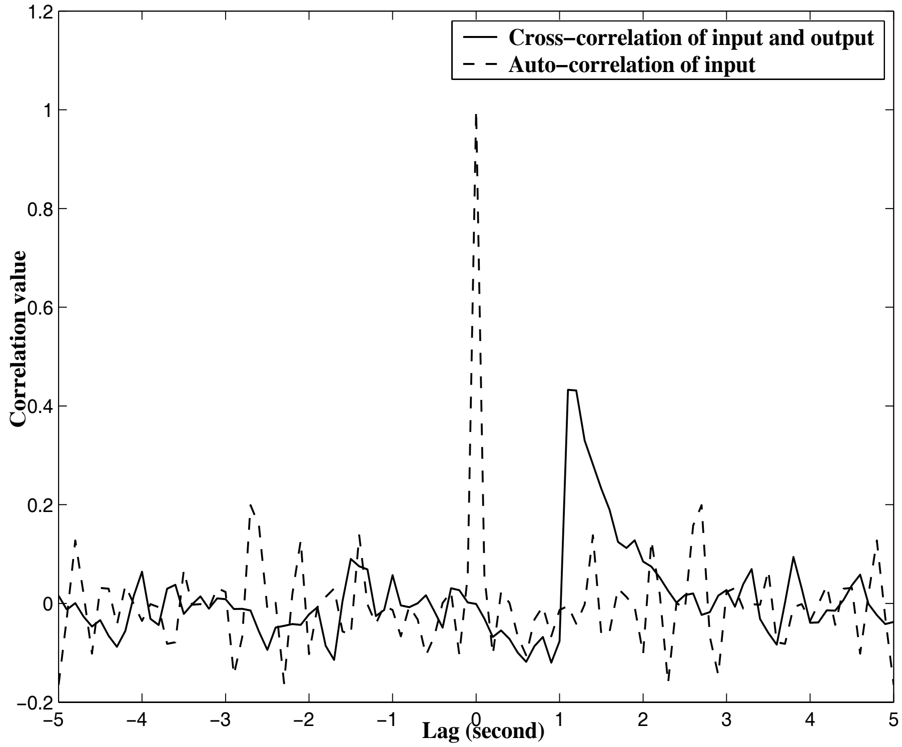

Figure B.1: Simulation example 1: The cross correlation function due to band limited white noise input starts at time-delay 1 s.

Table B.1: Parameters of input 1 and input 2.

Input	t_1	t_2	t_3	t_4
1	2	3	6	7
2	2	2.5	6	6.5

In this simulation, the system is excited three times with three different inputs. Each input spanned over 10 seconds and was uniformly sampled at 100 ms intervals. First two inputs are designed as follows.

$$u(t) = \begin{cases} 0 & for\ 0 \le t < t_1 \\ 1 & for\ t_1 \le t < t_2 \\ 0 & for\ t_2 \le t < t_3 \\ -1 & for\ t_3 \le t < t_4 \\ 1 & for\ t_4 \le t < 10 \end{cases} \tag{B.9}$$

The parameters t_1 through t_4 for first two inputs are as given in Table B.1. The third input is a band-limited white noise input. The estimates of normalized cross correlation of input and corresponding output are shown in Figure B.2.

It can be seen that the estimates of cross correlation functions concur at a point with abscissa approximately equal to 2 seconds corresponding to the process delay. However, the cross correlation function due to band-limited white noise which is also the estimated impulse response turns out to be noisy and there is an ambiguity regarding its start point. The correct start actually corresponds to the third peak of the estimated impulse response and a suitable thresholding criterion needs to be designed to designate the third peak as the first significant value or the start of the impulse response. On the contrary, given the well designed inputs, the proposed method does not suffer from such ambiguity.

B.4 Estimation of time delay of the LZCS

Data from a full scale test setup of LZCS has been used in this work, for estimation of time delay between application of control action and response observed in change of water level. The time-delay is primarily due to transport lags in the piping as water is pumped from the delay tank kept physically away from the ZCCs into the water inlet header of ZCCs, through a heat exchanger.

The ZCC can basically be viewed as an integrator. However, the level variation would start taking place after a delay, past any variation in the control valve position. The delay in the plant can be computed from the cross

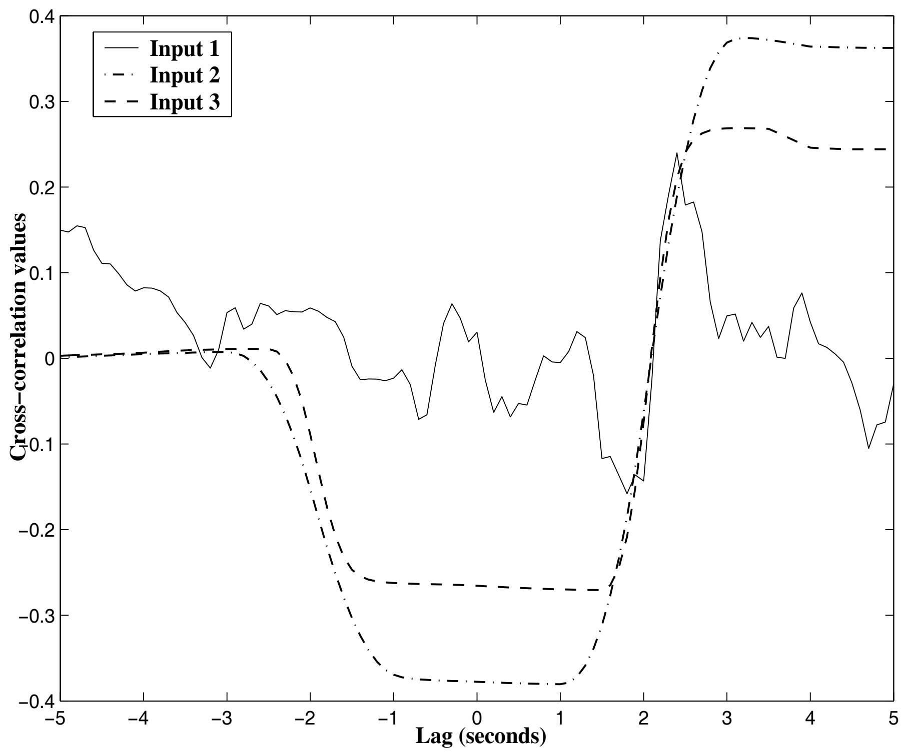

Figure B.2: Results of simulation example 2: The cross correlation functions corresponding to three inputs concur at a point having ordinate, indicating estimation of time-delay.

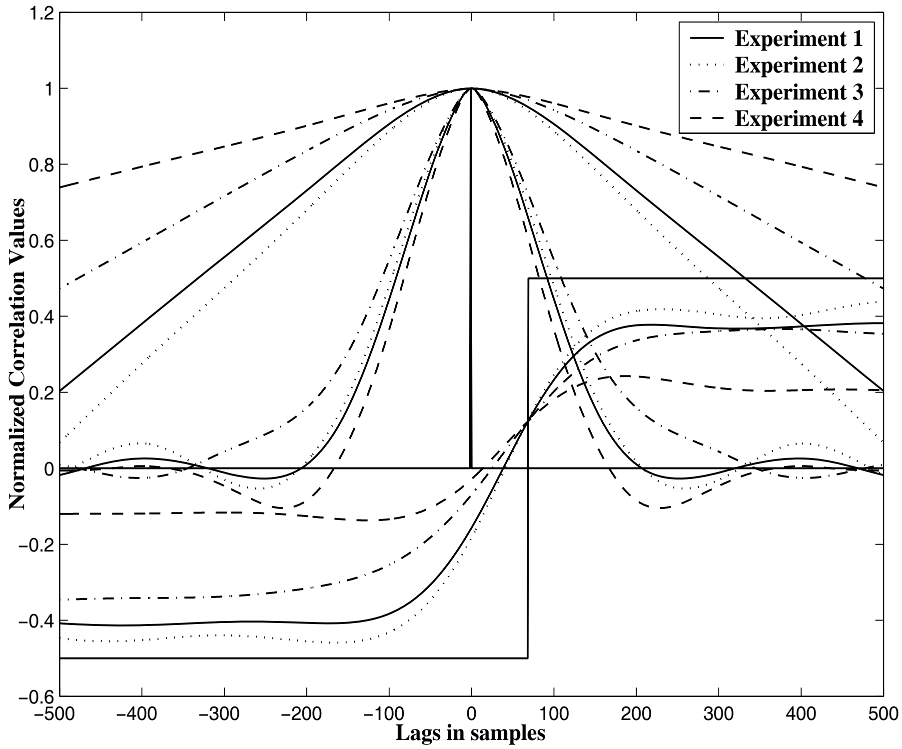

Figure B.3: Input output correlation of ZCC plotted with lags. Cross correlation functions concur at a point whose abscissa gives time-delay.

correlation functions of input and output as described below. To start with, the input and the output are de-trended. Three parameters, the normalized autocorrelation values of the input, the normalized autocorrelation values of the output and the normalized input-output cross correlation values are computed at lags up to 500 and plotted in Figure B.3, for four data sets shown in Figure 4.6 through 4.9.

The general nature of these plots can be understood with the help of simplified model of the ZCC level variation, i.e.,

$$y(t) = \mu \int_{-\infty}^{+\infty} u\,(t-d)\,dt \tag{B.10}$$

where μ is a scalar. From this following is obtained

$$\int_{-\infty}^{+\infty} y\,(t+\tau)\,u\,(\tau)\,d\tau = \mu \int_{-\infty}^{t} \int_{-\infty}^{+\infty} \left[u\,(t+\tau-d)\,u\,(\tau)\,d\tau\right] dt \tag{B.11}$$

And thus the following relationship between autocorrelation and cross correlation coefficients is obvious.

$$A_{yu}\left(t\right) = \mu \int_{-\infty}^{+\infty} A_{uu}\left(t - d\right) dt \tag{B.12}$$

It can be seen from the cross-correlation plots in Figure B.3 that the four inputs depending on their richness are exciting different dynamic modes in the process. But very interestingly, all the cross correlation plots are concurrent. So this point of concurrence has to be related to a system parameter which is invariant to the richness of the input. Now, let us consider a hypothetical situation, in which the system is excited by a pure white noise input. In such a case, it can be seen that for the ideal system given by (B.10), as the auto-correlation of input is a unit impulse, the input-output cross correlation is a step with transition at a lag of d. Since the point of concurrence is invariant to the richness of the input, the transition of the input-output cross-correlation curve would also be through the same point of concurrence. So it can be concluded that the x-coordinate (lag in samples) of the point of concurrence is the delay d of the system [55]. Delay computed for the LZCS test set up using this technique tallies with the expected value of delay which is due to transport delay in the piping.

B.5 Summary

The appendix presents a new robust technique for estimation of time delay in a process. The proposed technique estimates delay in time domain in presence of measurement noise. The technique is used for estimating transport delay in the LZCS. The delay is fed as an input parameter (implicit parameter) to the linear wavelet model. The report proposes a new method based on correlation analysis. The delay is estimated from the point of concurrence of cross correlation plots. It is established by illustrative examples that the method is less sensitive to noise and does not depend on a thesholding scheme depending on the measure of uncertainty. However, one needs to carefully design inputs for effective use of the new method. A condition is derived for design of such inputs. The method is validated using simulation examples and applying it on the data collected from a full scale model of LZCS used in large PHWRs. Systematic study of design of inputs and error in the estimate of time-delay could be taken up in future.

Courses and Publications

Courses and Credit Seminars Completed

1. Course 1: Process Modeling, Simulation and Optimization

2. Course 2: Process Dynamics and Control

3. Credit seminar 1: Signal Approximation in Shift Invariant Subspace

4. Credit Seminar 2: Estimation of Time-Delay from Noisy measurements

Publications Based on This Work

1. S. Mukhopadhyay, U. Mahapatra, A. K. Tangirala and A. P. Tiwari, Spline wavelets for system identification, Proceedings of DYCOPS 2010, Leuven, Belgium, 2010.

2. S. Mukhopadhyay and A. P. Tiwari, Characterization of NDT signals: Reconstruction from wavelet transfor maximum curvature representation, Signal Processing, 90(1), 261-268, 2010.

3. S. Mukhopadhyay and A. P. Tiwari, Consistent output estimate with wavelets: An alternative solution of least squares minimization problem for identification of the LZC system of a large PHWR, Annals of Nuclear Engineering, 37(7), 974-984, 2010.