IDENTIFICATION OF DISCRETE TIME BILINEAR SYSTEMS THROUGH EQUIVALENT LINEAR MODELS

NECDET BERK HIZIR

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2010

UMI Number: 3447850

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 3447850
Copyright 2011 by ProQuest LLC.
All rights reserved. This edition of the work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

©2010
NECDET BERK HIZIR
All Rights Reserved

ABSTRACT

IDENTIFICATION OF DISCRETE TIME BILINEAR SYSTEMS THROUGH EQUIVALENT LINEAR MODELS

NECDET BERK HIZIR

This work discusses 3 methods for identification of discrete-time, time-invariant, single-input, multi-output, state-space bilinear models using the steady-state portion of a single input/output time-history recorded by exciting the system with a linear combination of sines and cosines of user-selected frequencies enriched by a subtle amount of random component. The proposed methods rely on conversion of the bilinear system into an equivalent linear model (ELM) by an accurate approximation of the state in the bilinear term using a set of sine and cosine basis functions whose frequencies are obtained using an N^{th} order perturbation theory approximation of the state.

In the first proposed method, Observer/Kalman Filter Identification (OKID), a linear time invariant (LTI) system identification algorithm, is used to identify the aforementioned ELM from which the original bilinear model is recovered.

In the second proposed method, the aforementioned ELM is converted to an overparameterized ELM by replacing its original state time-history with its past input/output time-history. The order of the new ELM is reduced by discarding its unobservable states. Since the state time-history of the ELM is (known and) common to the bilinear model, the original bilinear system is recovered via a least-square fit.

On the other hand, in the third proposed method, the aforementioned ELM is identified using a direct method of LTI system identification. Perturbation theory is used to show that the identified ELM can be used to identify the steady-state portion of the state time-history corresponding to $(N+1)th$ order perturbation theory approximation of the state.

An iterative process is defined whereby the steady-state portion of the state time-history is identified exactly. The bilinear system matrices are identified via a least-square fit using the steady-state portion of the identified state-time history.

Numerical examples are also provided.

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to thank Prof. Raimondo Betti for guiding me throughout my graduate education, opening me the doors to a whole new field which I enjoyed exploring tremendously, and setting me up in a research team every meeting of which was such a great pleasure and privilege to attend. His contribution and support have been immense, my gratitude is endless. Grazie mille, Professore!

I would like to extend my deepest gratitudes to Prof. Richard W. Longman from whom I acquired most of the quantitative skills that I needed to complete this work. His teaching, experience and guidance have been most valuable; our discussions most rewarding.

My most cordial thanks are due to Prof. Minh Q. Phan. I have benefited vastly from his innovative approach to research and problem-solving. He has always been a great source of inspiration as a scholar and one of the main reasons why this work exists in the first place.

My thanks are also due to Prof. George Deodatis and Prof. Andrew W. Smyth. I have benefited from their teaching and advices significantly.

Last, but surely not least, I would like to thank my comrades at Columbia University who have made my experience the most enjoyable and special. My thanks are due to Gunes Demet Senturk, Efe Karanci, Cengiz Ucbenli, Virginia Mosquera, Ah Lum Hong, Jie Xu and Eleni Chatzi. Their friendship and support meant a lot to me.

Berk Hizir

Columbia University,

June 2010

"I think there is a bug in your prgoarm..." *program*

Raimondo Betti

Chapter 1

Introduction

1.1 What is a Bilinear System?

In the context discussed in this thesis, *system identification* involves the use of mathematical tools on measured data to develop a dynamical model that characterizes the time-dependent behavior of a system such that the identified system can be used to predict the response of the system due to any excitation.

The fields of applications of system identification are as numerous as the fields where dynamical models are needed, ranging from Engineering to Social Sciences. In many of these applications, the systems to be identified behave linearly or almost linearly, and linearity introduces the simplifications that come with the *principle of superposition of effects.* These reasons made *linear system identification* a very active research topic in the past and currently a well-established field with various, highly-accurate methods available in the literature, Ref.[1–13]. However, not all systems behave linearly, and there arises a need for more complicated models, yet still have relatively simple dynamic behavior to analyze. This makes bilinear systems, which stand out as a transition step between linear and nonlinear systems, a tempting option to use (see frequently cited survey papers Ref.[14–16]).

A discrete-time, time-invariant, multi-input, multi-output bilinear system can be written

in the state-space form, as in Eq.(1.1) and (1.2).

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{N}\mathbf{u}(k) \otimes \mathbf{x}(k) \tag{1.1}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \tag{1.2}$$

where k denotes the k^{th} time-step, while τ is the sampling interval such that $\mathbf{x}(k)$ implies $\mathbf{x}(k\tau)$. The $n \times 1$ vector $\mathbf{x}(k) \in \mathbb{R}^n$ is the state vector, $\mathbf{y}(k) \in \mathbb{R}^r$ is the output vector, while $\mathbf{u}(k) \in \mathbb{R}^m$ is the input (or control) vector with components $u_i(k) \in \mathbb{R}$ $(i = 1, 2, \cdots, m)$. The $n \times n$ matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the state matrix, $\mathbf{B} \in \mathbb{R}^{n \times m}$ represents the input matrix, while $\mathbf{C} \in \mathbb{R}^{r \times n}$ and $\mathbf{D} \in \mathbb{R}^{r \times m}$ indicate the output and direct transmission matrices, respectively. The matrix $\mathbf{N} \in \mathbb{R}^{n \times mn}$ is the matrix that weighs the bilinear term $\mathbf{u}(k) \otimes \mathbf{x}(k)$, where $\otimes$ represents the *kronecker product*. The matrix $\mathbf{N}$ is of the form $\begin{bmatrix} \mathbf{N}_1 & \mathbf{N}_2 & \cdots & \mathbf{N}_m \end{bmatrix}$, where each matrix $\mathbf{N}_i \in \mathbb{R}^{n \times n}$ weighs the bilinear term $\mathbf{x}(k)u_i(k)$ associated with the corresponding input. Eq.(1.1) can be viewed in another commonly used, yet less compact form by expanding it as follows:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \sum_{i=1}^{m} \mathbf{N}_i \mathbf{x}(k) u_i(k) \tag{1.3}$$

The bilinear systems of the form in Eq.(1.4), where the input matrix $\mathbf{B} = 0$ are called *homogeneous* bilinear systems.

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \sum_{i=1}^{m} \mathbf{N}_i \mathbf{x}(k) u_i(k) \tag{1.4}$$

Moreover, if also the state matrix $\mathbf{A} = 0$ as in Eq.(1.5), such systems are called *symmetric*.

$$\mathbf{x}(k+1) = \sum_{i=1}^{m} \mathbf{N}\mathbf{x}(k) u_i(k) \tag{1.5}$$

Eq.(1.3) reduces to the form in Eq.(1.6) for a *single-input* bilinear system, which will be the main focus of this work.

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \mathbf{N}\mathbf{x}(k)u(k) \tag{1.6}$$

By definition being linear in the state and the input separately, bilinear systems can be viewed as linear, time-varying systems as shown in Eq.(1.7).

$$\mathbf{x}(k+1) = [\mathbf{A} + \sum_{i=1}^{m} \mathbf{N}_i u_i(k)]\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \tag{1.7}$$

where $[\mathbf{A} + \sum_{i=1}^{m} \mathbf{N}_i u_i(k)]$ represents the time-varying state matrix. One way of interpreting Eq.(1.7) is by noting that the state matrix is obtained by weighing the matrix $\mathbf{A}$ by 1 and the matrices $\mathbf{N}_i$ by the inputs $u_i(k)$ at each time step. Inputting an excitation $\mathbf{u}(k)$ such that $\|\mathbf{u}(k)\| \ll 1$ makes the system behave close to a linear, time-invariant system as in Eq.(1.8)

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \tag{1.8}$$

As a quick illustration of this behavior, consider a bilinear system of the form in Eq.(1.6) where the system matrices are defined as

$$\begin{aligned} &\mathbf{A} = \begin{bmatrix} 0.9924 & 0.0118 \\ -1.1778 & 0.8747 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.0126 \\ 0.0042 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}, \\ &\mathbf{C} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \mathbf{D} = 0 \end{aligned} \tag{1.9}$$

Feeding a random input that takes random values from a uniform distribution defined in the interval $[0, 0.01]$, the bilinear system outputs a response virtually identical to the output of the equivalent linear system of the form in Eq.(1.8), as shown in Fig. 1.1

Furthermore, when the applied input is constant, a bilinear system becomes a linear, time-invariant system, in accordance with the form in Eq.(1.7). This property was used by various researchers working on stabilization, control and identification of bilinear systems, examples of which include Ref.[16–22].

Having discussed various ways to interpret bilinear systems, one wonders about the conditions needed for determining their stability. The derivation of a stability criterion for bilinear systems is not as straightforward as that for linear systems, since there exists no known closed-form solutions as stated in Ref.[23], although a *Volterra Series* expansion can be used to serve as an approximation to the solution as discussed in Ref.[24–26]. However, due to the fact that bilinear systems can be considered as linear, time-varying systems, *Floquet's Theorem* can be employed to derive conditions for the stability of bilinear systems subject to periodic input excitations (see Ref.[27–29] for a more generic and detailed discussion or examples).

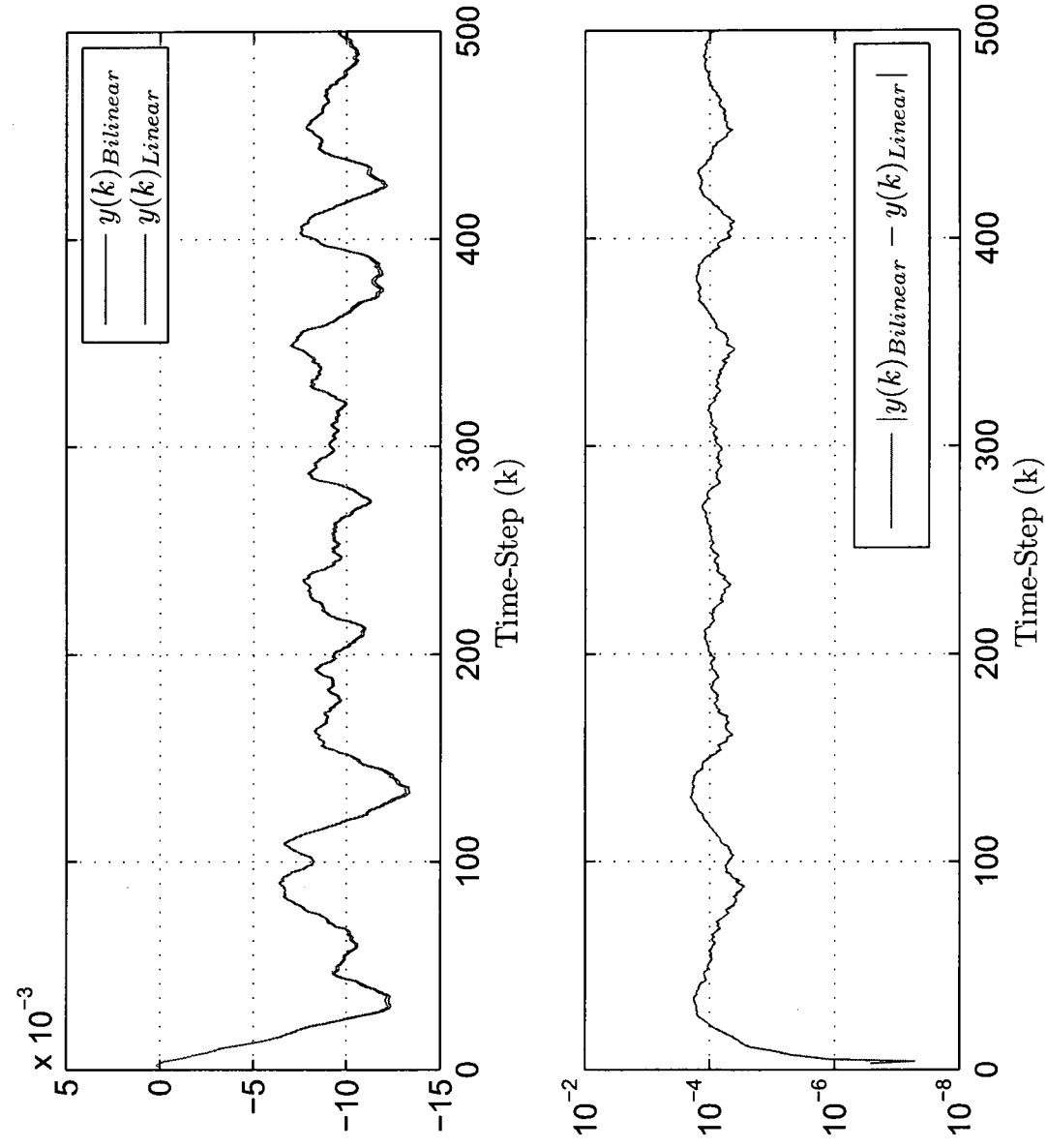


Figure 1.1: Output time-histories generated by a small amplitude input

1.1.1 Floquet's Theorem and Stability

Consider the following linear time-varying system whose state matrix $\mathcal{A}(k)$ varies with time with a period T:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathcal{A}(k)\mathbf{x}(k) + \mathcal{B}\boldsymbol{u}(k) \\ \mathbf{y}(k) &= \mathcal{C}\mathbf{x}(k) + \mathcal{D}\mathbf{u}(k) \end{aligned} \tag{1.10}$$

where $\mathcal{A}(k) = \mathcal{A}(k+T)$. The stability of Eq.(1.10) does not depend on the input $\mathbf{u}(k)$, therefore, for convenience, $\mathcal{B}$ can be assumed to be zero. Then, the state time-history can be written as:

$$\mathbf{x}(1) = \mathcal{A}(0)\mathbf{x}(0) \tag{1.11}$$

$$\mathbf{x}(2) = \mathcal{A}(1)\mathcal{A}(0)\mathbf{x}(0) \tag{1.12}$$

$$\vdots$$

$$\mathbf{x}(T) = [\mathcal{A}(T-1)\mathcal{A}(T-2)\cdots\mathcal{A}(0)]\mathbf{x}(0) = \mathbf{M}\mathbf{x}(0) \tag{1.13}$$

$$\mathbf{x}(T+1) = [\mathcal{A}(T)\mathbf{M}]\mathbf{x}(0) = [\mathcal{A}(0)\mathbf{M}]\mathbf{x}(0) \tag{1.14}$$

$$\mathbf{x}(T+2) = [\mathcal{A}(1)\mathcal{A}(0)\mathbf{M}]\mathbf{x}(0) \tag{1.15}$$

$$\vdots$$

$$\mathbf{x}(2T) = \mathbf{M}^2\mathbf{x}(0) \tag{1.16}$$

$$\vdots$$

$$\mathbf{x}(nT) = \mathbf{M}^n\mathbf{x}(0) \tag{1.17}$$

Assuming the matrix $\mathbf{M}$ has the eigenvalue decomposition $\mathbf{M} = \mathbf{S}\Lambda\mathbf{S}^{-1}$, where $\mathbf{S}$ is the matrix of eigenvectors and Λ is the diagonal matrix of eigenvalues, $\mathbf{M}^n$ can be expressed as $\mathbf{S}\Lambda^n\mathbf{S}^{-1}$. Using Eq.(1.17), it can be concluded that, if the eigenvalues of $\mathbf{M}$ are inside the unit circle, then the system defined in Eq.(1.10) must be stable since, as n goes to infinity $\mathbf{M}^n$, and therefore, $\mathbf{x}(nT)$ tends to zero. Note that this formulation does not define a contraction mapping such as $\|\mathbf{x}(k)\| > \|\mathbf{x}(k+1)\|$ or even $\|\mathbf{x}(nT)\| > \|\mathbf{x}((n+1)T)\|$ which requires the singular values of $\mathbf{M}$ to lie within the unit circle. It merely suggests that, as n

increases, $\mathbf{x}(nT)$ approaches zero (not necessarily monotonically), and thus, $\mathbf{x}(k\tau)$ should approach zero.

In the presence of a periodic input excitation with period T (i.e. $u(k) = u(k+T)$), the bilinear system defined in Eq.(1.7) takes the form in Eq.(1.18).

$$\mathbf{x}(k+1) = \tilde{\boldsymbol{\mathcal{A}}}(k)\mathbf{x}(k) + \boldsymbol{\mathcal{B}}u(k) \tag{1.18}$$

where $\tilde{\boldsymbol{\mathcal{A}}}(k) = [\mathbf{A} + \mathbf{N}u(k)]$ and $\tilde{\boldsymbol{\mathcal{A}}}(k) = \tilde{\boldsymbol{\mathcal{A}}}(k+T)$. Hence, the condition for asymptotic stability of the bilinear system becomes the spectral radius of the matrix $[\tilde{\boldsymbol{\mathcal{A}}}(T-1)\tilde{\boldsymbol{\mathcal{A}}}(T-2)\cdots\tilde{\boldsymbol{\mathcal{A}}}(0)]$ be less than 1.

In the case of an aperiodic excitation input, a time-invariant bilinear system can be treated as a linear, time-varying system whose stability can be proved by satisfying one of the various sufficient conditions for stability of linear time-varying systems available in the literature such as Ref.[30; 31].

1.1.2 Observability and Controllability

The concepts of observability and controllability of dynamical systems were first introduced by Ref.[32–34]. *Complete observability* of a system implies that, given a finite-length measurement of the output, one can track back the full state vector at the start of the measurement. The concept of observability is especially useful in system identification, since it tells whether the unmeasurable states can be constructed using the output measurements, which may possibly contain contributions from only some of the states.

On the other hand, *complete controllability* of a system implies that there exists an input signal $\mathbf{u}(t)$ that can transfer any initial state $\mathbf{x}(0) \in \mathbb{R}^n$, where n is the order of the system, to any final state $\mathbf{x}(l) \in \mathbb{R}^n$ in the finite time interval $0 \leq t \leq l$. The importance of observability and controllability stem from the fact that the identification of a system yields only its controllable and observable part, leaving out its uncontrollable and/or unobservable parts.

One of the commonly used, necessary and sufficient conditions for observability and controllability of continuous-time, time-invariant, linear systems of the form in Eq.(1.19)

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \tag{1.19}$$

is that both the controllability matrix $\boldsymbol{\zeta}$ and the observability matrix $\boldsymbol{\Theta}$, defined in Eq.(1.20) and (1.21),

$$\boldsymbol{\zeta} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{1.20}$$

$$\boldsymbol{\Theta} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{1.21}$$

be full rank, respectively. These conditions are also called *Kalman rank conditions.*

However, the above conditions are not applicable to bilinear systems, although as presented in Ref.[35–38] there exists sufficient (but not necessary) conditions to prove that a nonlinear system is observable and controllable, and these conditions can be adapted to bilinear systems as well, as performed in Ref.[39–42].

Consider a continuous-time, nonlinear system of the form

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) &= \mathbf{h}(\mathbf{x}(t)) \end{aligned} \tag{1.22}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the state vector, $\mathbf{u}(t) \in \mathbb{R}^m$ is the input vector, while $\mathbf{y}(t) \in \mathbb{R}^r$ represents the output vector. Each of these vectors have $x_i(t) \in \mathbb{R}$, $u_i(t) \in \mathbb{R}$ and $y_i(t) \in \mathbb{R}$, respectively. The functions $\mathbf{f}(\mathbf{x}, \mathbf{u}) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\mathbf{h}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^r$ are vector functions comprised of the components $f_i(\mathbf{x}(t), \mathbf{u}(t))$ and $h_i(\mathbf{x}(t))$, respectively, and let's recall the *Lie derivative* of h_j with respect to $\mathbf{f}$, defined as:

$$L_{\mathbf{f}} h_j = \frac{\partial h_j}{\partial \mathbf{x}^T}\mathbf{f} = \sum_{i=1}^{n} \frac{\partial h_j}{\partial x_i} f_i \tag{1.23}$$

and its partial derivative of order k with respect to $\mathbf{x}^T$:

$$\frac{\partial L_{\mathbf{f}}^k h_j}{\partial \mathbf{x}^T} = \begin{bmatrix} \frac{\partial L_{\mathbf{f}}^k h_j}{\partial x_1} & \frac{\partial L_{\mathbf{f}}^k h_j}{\partial x_2} & \dots & \frac{\partial L_{\mathbf{f}}^k h_j}{\partial x_n} \end{bmatrix} \tag{1.24}$$

According to the Hermann-and-Krener Theorem, Ref.[36], a continuous-time, nonlinear system is said to be *strongly (globally) observable*, if there exists an input $u(t)$ for which the rank of the matrix in Eq.(1.25) is equal to the order of the system for $\forall \mathbf{x} \in \mathbb{R}^n \backslash \mathbf{0}$, where $\mathbb{R}^n \backslash \mathbf{0}$ denotes the n dimensional space of real numbers excluding the origin. If the rank condition is satisfied only for a small neighborhood of $\mathbf{x} \in \mathbb{R}^n$, then the system is said to be *weakly (locally) observable*. Note that for nonlinear systems observability is not solely a system property but also depends on the input.

$$\begin{bmatrix} \begin{bmatrix} \frac{\partial L_{\mathbf{f}}^0 h_1}{\partial \mathbf{x}^T} \\ \vdots \\ \frac{\partial L_{\mathbf{f}}^0 h_r}{\partial \mathbf{x}^T} \end{bmatrix} \\ \begin{bmatrix} \frac{\partial L_{\mathbf{f}}^1 h_1}{\partial \mathbf{x}^T} \\ \vdots \\ \frac{\partial L_{\mathbf{f}}^1 h_r}{\partial \mathbf{x}^T} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \frac{\partial L_{\mathbf{f}}^{n-1} h_1}{\partial \mathbf{x}^T} \\ \vdots \\ \frac{\partial L_{\mathbf{f}}^{n-1} h_r}{\partial \mathbf{x}^T} \end{bmatrix} \end{bmatrix} \tag{1.25}$$

The rank-condition for observability is consistent with the Kalman-rank condition for linear systems, and has already been generalized to discrete-time nonlinear systems in Ref.[43].

As a demonstration of the rank-condition, consider the single-input, single-output bilinear system of the form in Eq.(1.26)

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \begin{bmatrix} 0 & 1 \\ -100 & -10 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}(t)u(t) \\ \mathbf{y}(t) &= \begin{bmatrix} 1 & 2 \end{bmatrix} \mathbf{x}(t) \end{aligned} \tag{1.26}$$

where $f_1 = x_2 + u + x_1 u$, $f_2 = -100x_1 - 10x_2 + u + x_2 u$ and $h = x_1 + 2x_2$. Then, the rank

condition becomes as in Eq.(1.27).

$$rank\begin{bmatrix} \frac{\partial h}{\partial x^T} \\ \frac{\partial L_f^1 h}{\partial x^T} \\ \vdots \\ \frac{\partial L_f^0 h}{\partial x^T} \end{bmatrix} = rank\begin{bmatrix} 1 & 2 \\ x_2 + u + x_1 u & -200x_1 - 20x_2 + 2u + 2x_2 u \end{bmatrix} \stackrel{?}{=} 2 \qquad (1.27)$$

For instance, for $x_1 = x_2 = 1$, the rank condition is satisfied and the bilinear system is concluded to be locally observable about this point. On the other hand, for $x_1 = x_2 = 0$ the rank condition is not satisfied, yet, this alone does not imply that the system is globally unobservable, since (i) the rank condition is a sufficient but not necessary condition for observability, and (ii) being locally unobservable at the origin does not imply global unobservability, Ref.[16].

Similar results have been derived for controllability of bilinear systems as well. The conditions derived for the controllability of nonlinear systems in Ref.[41] require that for homogeneous bilinear systems of the form in Eq.(1.4) to be controllable (i) a Lie-rank condition be satisfied, (ii) the eigenvalues of the **A** matrix be imaginary and distinct. Moreover, Ref.[42] presents sufficient conditions for controllability of symmetric bilinear systems of the form in Eq.(1.5).

1.1.3 Applications

Bilinear systems find applications in a variety of fields. In chemical engineering, *continuous stirred-tank reactors* (CTSR), *nuclear fission reactors* and *distillation columns* are modeled and controlled by bilinear systems. For example, Ref.[44] describes a method to approximate distillation processes via bilinear models. Ref.[45] uses the aforementioned method to accurately approximate a 3-input (flow rates and heating power), single-output (head concentration), 11-state nonlinear distillation column using a 3-state bilinear model. The identified reduced order model is shown to remain robust even under large input, output variations.

In biology and ecology, as presented in Ref.[46], microbial cell growth and product formation of various waste treatment and fermentation systems can be modeled using bilinear

systems. An estimate of the state is used to compensate for the lack of accurate biological sensors. The state estimate can also be used for the on-line control of these biological processes.

In demography, bilinear control models are used in various demographic processes to examine birth rates, death rates, migration rates, differentiation rates, etc. for their socioeconomic implications. Ref.[47] generalizes these discussions from human demographics to general biological species and cells with particular emphasis to compartmental processes where migration between regions or compartments occur.

Within the aforementioned framework, in immunology, bilinear systems are utilized to examine the behavior of white blood cells called lymphocytes. Ref.[48] presents a model for compartmental control of the immune process. Ref.[49] develops such models for a specific type of lymphocytes called T-cells and B-cells to inspect their generation of antibody to bind alien antigens.

Even economy is among the fields of applications of bilinear systems, Ref.[50]. In Ref.[51], several bilinear versions of the Harrod-Domar macroeconomical growth model are derived, and their sensitivity is analyzed.

1.2 Bilinear System Identification Methods

Although the literature contains various methods for stabilization and control of bilinear systems, their identification is not as well understood. The currently available methods either put restrictions on the type of input excitations used or demand a high number of experiments that grow with the size of the system or are valid only for specific types of bilinear systems.

The bilinear system identification methods can be roughly categorized as exact methods and approximate methods, such as those proposed in this thesis. The exact methods aim to identify a realization of the bilinear system matrices directly whereas approximate methods convert the bilinear system identification problem to a linear one. Most of the methods from both categories use either the fact that bilinear systems can be treated as linear,

time-varying systems or the fact that they become linear, time-invariant systems when the applied input excitation is constant. In general, the approximate methods rely on the conversion of the bilinear systems to linear ones by employing various series expansions on the inputs, outputs and/or states.

Ref.[11–13] present the theory and several versions of the algorithms of the *subspace identification* for linear systems, where 4 subspaces are defined using the measurements of input/output time-histories: *past input* U_p, *past output* Y_p, *future input* U_f and *future output* Y_f. The state space X is shown to be contained by $\{U_p \cup Y_p\}$ alone and by $\{U_f \cup Y_f\}$ alone. Since X lies in both of these subspaces, it must be contained by their intersection $\{U_p \cup Y_p\} \cap \{U_f \cup Y_f\}$. Then, *Grassmann's Dimension Theorem*, which states the dimension of the space formed by the union of any 2 spaces is equal to the sum of the dimensions of each of the 2 spaces minus the dimension of the space formed by the intersection of the 2 spaces, is used to prove that the dimension of the space $\{U_p \cup Y_p\} \cap \{U_f \cup Y_f\}$, which contains X, is equal to the order of the system, which is equal to the dimension of X, as shown in Eq.(1.28)

$$\begin{aligned} dim(\{U_p \cup Y_p\} \cap \{U_f \cup Y_f\}) &= dim(U_p \cup Y_p) + dim(U_f \cup Y_f) \\ &\quad - dim(\{U_p \cup Y_p\} \cup \{U_f \cup Y_f\}) \\ &= dim(X) \end{aligned} \tag{1.28}$$

where $dim(\bullet)$ represents the dimension of the space $\bullet$. In other words, the subspace $\{U_p \cup Y_p\} \cap \{U_f \cup Y_f\}$ and the state-space X have equivalent sets of basis vectors. Thus, a matrix whose rows are the basis vectors for $\{U_p \cup Y_p\} \cap \{U_f \cup Y_f\}$ is in fact a realization of the state time-history. Once the state time-history is determined, a least-square fit is performed to obtain the system matrices. The extraction of the basis vectors of X can be accomplished by first forming $\bar{Y}_f$ by orthogonally projecting the basis vectors that span Y_f on the space formed by the union of U_f, U_p, Y_p; then, subtracting the oblique projection of $\bar{Y}_f$ on U_f from $\bar{Y}_f$ to obtain X. Ref.[52–54], extends the above discussion for linear systems to bilinear systems by treating them as time-varying linear systems. However, an important constraint, which hammers the practicality of the algorithm, is imposed by bounding the excitation input to be white noise.

Ref.[18] discusses a method for identification of *continuous-time* bilinear systems using noise corrupted output measurements by approximating the input and output time-histories using block-pulse functions. Such a method is computationally costly, and could only be shown to work for a single-input, single-output, first order system. Similarly, benefiting from the fact that bilinear systems behave as linear, time-invariant systems under constant input, Ref.[19; 20] discuss the identification of *continuous-time* bilinear systems using a family of unit-pulses of fixed amplitude but varying widths as the inputs in a series of experiments the number of which is proportional to the order of the system. Moreover, Ref.[21; 22] show that neither step inputs nor single pulses, but fixed amplitude-varying width pulses suffice for identification of bilinear systems. Therefore, the practical applicability of the current form of this method is greatly hampered by the need for the high number of experiments for large systems and the strict requirements on the type of the input utilized. Besides, the identified continuous-time model needs to be discretized for its integration to a digital environment such as a digital controller or computer, and depending on the method of discretization this process may be a potential source of error.

Another alternative for converting bilinear systems to linear systems is using Walsh functions, which are basically a train of pulse functions of varying widths with amplitudes equal to either +1 or -1 (see Fig.1.2 for an example). Walsh functions form an orthonormal basis function set for square-integrable functions on the unit interval (see Ref.[55; 56] for a more detailed description). In other words, in close resemblance to Fourier Series expansion, periodic wave forms can be expressed in terms of Walsh functions via Walsh-Hadamard Transform. Ref.[57] discusses solutions to linear systems using Walsh functions, and Ref.[58] extends this reasoning to optimal control of time-varying linear systems. A later work, Ref.[59; 60] presents a method for approximating solutions to bilinear systems in a computationally efficient way using *butterfly networks* similar to those used for Fast Fourier Transforms and Fast Walsh-Hadamard Transforms. As far as the identification of bilinear systems is concerned, Ref.[61] presents a method of identification for continuous-time, *time-variant* bilinear systems that relies on an accurate approximation of the input and output signals by a finite set of Walsh functions. However, the method is only applicable to systems that have the same number of outputs as the number of states. In addition,

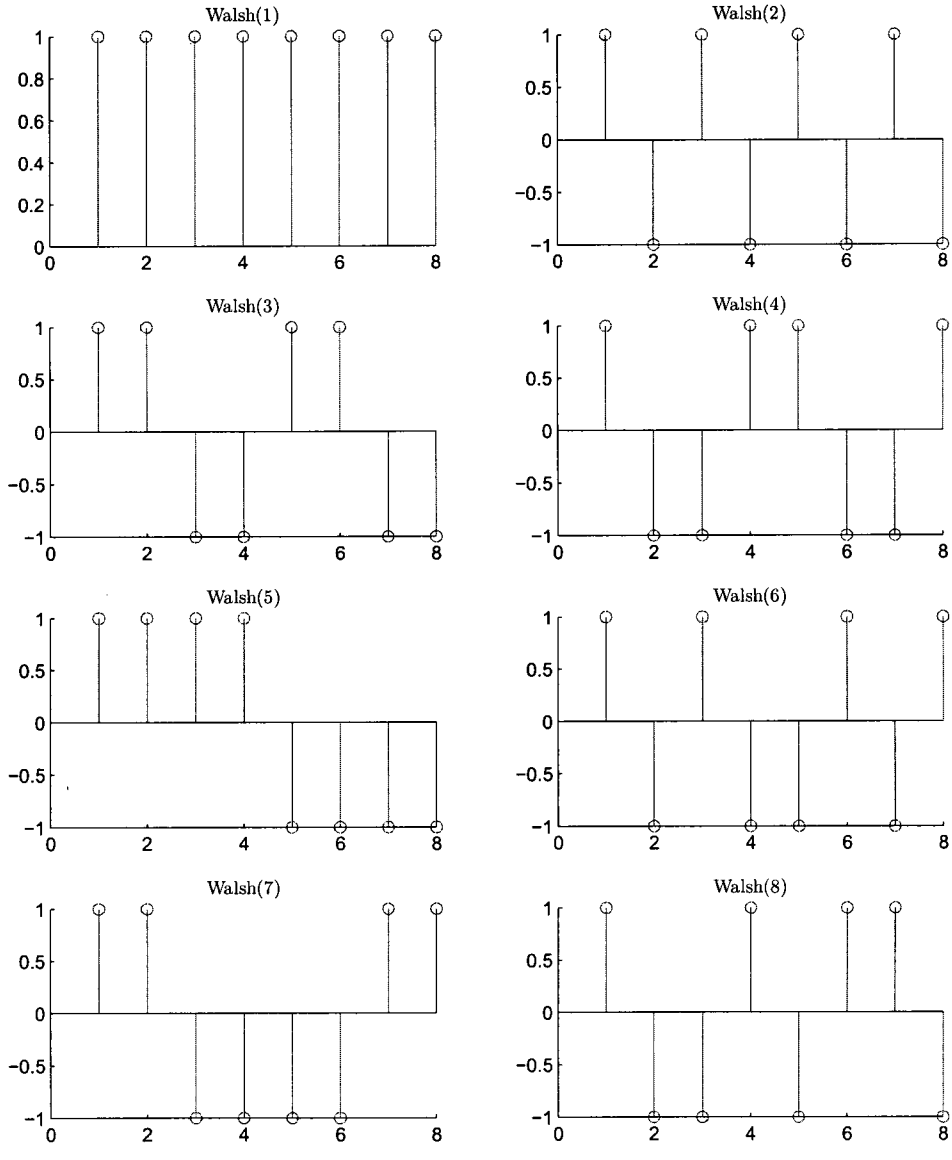


Figure 1.2: The complete set of 8-point, discrete-time Walsh Functions as an example

the difficulty of knowing in practice the system order in advance and the growth of the computational cost with this number contribute to make the method impractical, despite the promise of the method's applicability to more general problems than merely bilinear ones.

Another transform used in the identification of bilinear systems is the *Hartley Transform.* The Hartley Transform of a function, $f(t)$ is calculated as in Eq.(1.29).

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)\text{cas}(\omega t)\text{dt} \tag{1.29}$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} H(w)\text{cas}(\omega t)\text{d}\omega \tag{1.30}$$

where $\text{cas}(\omega t) = \cos(\omega t) + \sin(\omega t)$. As an alternative to Fourier Transform, Hartley Transform maps real-valued functions to real-valued functions, and is an involution as shown in Eq.(1.30) (see Ref.[62] for a more rigorous description and properties). Ref. [63; 64] discuss the identification of continuous-time, bilinear systems in the observability canonical form, where the bilinear system is treated as a linear, time-varying system. Taking the inner product of the input and output functions with Hartley Modulating functions and using their various properties, the model is brought to a form ready for identification of the model parameters via regression. Finally, a frequency weighted least squares algorithm is used to identify the system.

In the identification of *causal* nonlinear systems, one approach has used Volterra Series expansion shown in Eq.(1.33), which can be thought as a combination of the Taylor series expansion for *memoryless* nonlinear systems, as shown in Eq.(1.31) and the *convolution integral* for *causal* linear systems, as shown in Eq.(1.32),

$$y(t) = \sum_{i=0}^{\infty} c_n [u(t)]^n \tag{1.31}$$

$$y(t) = \int_{-\infty}^{\infty} u(t-\tau)k_n(\tau)\text{d}\tau \tag{1.32}$$

$$y(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \frac{1}{n!} \int_{-\infty}^{\infty} \text{d}\tau_1 \cdots \int_{-\infty}^{\infty} \text{d}\tau_n k_n(\tau_1, \cdots, \tau_n) \prod u(t-\tau_n) \tag{1.33}$$

where c_n are the Taylor Series coefficients, $u(t)$ is the input, $y(t)$ is the output and $k_n(\tau)$ are the *Volterra kernels* (see Ref.[65] for more details and examples on Volterra series

expansions). Ref.[66] proves that there exists an infinite number of non-vanishing Volterra kernels for bilinear systems, and Ref.[67–70] discuss some of the algorithms used in the estimation of the first few of the kernels from the measured input/output time-histories. Though an accurate estimation of the first few kernels are shown to be sufficient to predict the system response with reasonable accuracy, the algorithms are prone to measurement noise.

Lastly, the literature contains recursive methods used in the identification of bilinear systems. Again, exploiting the property that bilinear systems can be treated as linear, time-varying systems, various recursive least square algorithms have been proposed by researchers most of which yield unbiased estimates for the system parameters, Ref.[71–74].

In conclusion, there exists a need for an algorithm that can accurately identify a bilinear system such that (i) it allows for the design of the input excitation without putting strict restrictions on the type of excitation used, (ii) requires a single experiment, and (iii) is applicable to all types of bilinear systems.

1.3 Observer/Kalman Filter Identification (OKID)

The proposed methods for the identification of time-invariant bilinear systems rely on the accurate identification of equivalent linear models (ELM). Therefore, the methods involve linear, time-invariant (LTI) system identification algorithms. Although the first method can be coupled with any LTI system identification method, Observer/Kalman Filter Identification (OKID) has been preferred for its high robustness against noise. In this section, OKID is briefly introduced with due emphasis on the parts closely related to the proposed method to familiarize the reader to its judicious usage in the chapters to follow. Further details regarding OKID can be found in Ref.[2; 3].

Consider a finite dimensional, discrete-time, linear, time-invariant system as shown in Eq.(1.34).

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \tag{1.34}$$

where $k = 0, 1, 2, ..., l-1$ and all other the symbols maintain their predefined meanings. For zero initial conditions, $\mathbf{x}(0) = \mathbf{0}$, one can express the output time-history as in Eq.(1.35)

$$\mathcal{Y} = \hat{\mathbf{Y}}\hat{\mathbf{U}} \tag{1.35}$$

where

$$\mathcal{Y} = \begin{bmatrix} \mathbf{y}(0) & \mathbf{y}(1) & \mathbf{y}(2) & \cdots & \mathbf{y}(p) & \cdots & \mathbf{y}(l-1) \end{bmatrix} \tag{1.36}$$

$$\hat{\mathbf{Y}} = \begin{bmatrix} \mathbf{D} & \mathbf{CB} & \mathbf{CAB} & \cdots & \mathbf{CA}^{p-1}\mathbf{B} & \cdots & \mathbf{CA}^{l-2}\mathbf{B} \end{bmatrix} \tag{1.37}$$

$$\hat{\mathbf{U}} = \begin{bmatrix} \mathbf{u}(0) & \mathbf{u}(1) & \mathbf{u}(2) & \cdots & \mathbf{u}(p) & \cdots & \mathbf{u}(l-1) \\ & \mathbf{u}(0) & \mathbf{u}(1) & \cdots & \mathbf{u}(p-1) & \cdots & \mathbf{u}(l-2) \\ & & \mathbf{u}(0) & \cdots & \mathbf{u}(p-2) & \cdots & \mathbf{u}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \mathbf{u}(0) & \cdots & \mathbf{u}(l-p-1) \\ & & & & & \ddots & \vdots \\ & & & & & & \mathbf{u}(0) \end{bmatrix} \tag{1.38}$$

The matrix $\hat{\mathbf{Y}}$ in Eq.(1.37) contains the *Markov parameters* of the LTI system defined in Eq.(1.34). For a finite dimensional linear system, the Markov parameters are unique. Once the Markov parameters are identified, the extraction of the system matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ can be accomplished as described in Ref.[75; 76] by the *Eigensystem Realization Algorithm* (ERA) or *Eigensystem Realization Algorithm using Data Correlation* (ERA-DC), which are methods derived from the System Realization Algorithm proposed in Ref.[1].

Eq.(1.35) contains more unknowns ($r \times ml$) than equations ($r \times l$), therefore, it cannot be readily used to solve for the system Markov parameters $\hat{\mathbf{Y}}$. However, for an asymptotically stable LTI system, the spectral radius of the matrix $\mathbf{A}$ is less than 1, therefore, $\mathbf{CA}^k\mathbf{B} \approx 0$, $\forall k \geq p$, for large enough p. Thus, the first p Markov parameters $\mathbf{Y}$ can be approximated using

$$\mathbf{Y} = \mathcal{Y}\mathbf{U}^{+} \tag{1.39}$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{D} & \mathbf{CB} & \mathbf{CAB} & \cdots & \mathbf{CA}^{p-1}\mathbf{B} \end{bmatrix} \tag{1.40}$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}(0) & \mathbf{u}(1) & \mathbf{u}(2) & \cdots & \mathbf{u}(p) & \cdots & \mathbf{u}(l-1) \\ & \mathbf{u}(0) & \mathbf{u}(1) & \cdots & \mathbf{u}(p-1) & \cdots & \mathbf{u}(l-2) \\ & & \mathbf{u}(0) & \cdots & \mathbf{u}(p-2) & \cdots & \mathbf{u}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \mathbf{u}(0) & \cdots & \mathbf{u}(l-p-1) \end{bmatrix} \tag{1.41}$$

and $^+$ denotes the Moore-Penrose pseudoinverse. For lightly damped structures such as Aerospace or Civil Engineering structures, the length of the required input-output measurement l and the number of non-negligible Markov parameters p are still so high that the size of the matrix $\mathbf{U}$ becomes too large to calculate its pseudo-inverse. Moreover, for a unique solution, the matrix $\mathbf{U}$ should be full-rank, and if $\mathbf{u}(0) = \mathbf{0}$, this cannot be satisfied. Besides, the formulation is valid only for the case of zero initial conditions.

In close resemblance to an *observer equation*, the following algebraic manipulation yields an equivalent LTI system with more inputs and the same outputs, but has a customizable decay rate which can be used to shrink the size of the matrix U:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{M}\mathbf{y}(k) - \mathbf{M}\mathbf{y}(k) \\ &= (\mathbf{A}+\mathbf{M}\mathbf{C})\mathbf{x}(k) + (\mathbf{B}+\mathbf{M}\mathbf{D})\mathbf{u}(k) - \mathbf{M}\mathbf{y}(k) \\ &= (\mathbf{A}+\mathbf{M}\mathbf{C})\mathbf{x}(k) + \begin{bmatrix}(\mathbf{B}+\mathbf{M}\mathbf{D}) & -\mathbf{M}\end{bmatrix} \begin{bmatrix}\mathbf{u}(k) \\ \mathbf{y}(k)\end{bmatrix} \\ &= \bar{\mathbf{A}}\mathbf{x}(k) + \bar{\mathbf{B}}\mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \tag{1.42}$$

where $\mathbf{M}$ is an arbitrary $n \times q$ matrix, and

$$\bar{\mathbf{A}} = \mathbf{A} + \mathbf{M}\mathbf{C} \tag{1.43}$$

$$\bar{\mathbf{B}} = \begin{bmatrix}\mathbf{B}+\mathbf{M}\mathbf{D} & -\mathbf{M}\end{bmatrix} \tag{1.44}$$

$$\mathbf{v}(k) = \begin{bmatrix}\mathbf{u}(k) \\ \mathbf{y}(k)\end{bmatrix} \tag{1.45}$$

The manipulation above is really an effort to form a deadbeat observer to place all eigenvalues of $\bar{\mathbf{A}}$ at the origin, as detailed in Ref.[8]. Thus, $\mathbf{C}\bar{\mathbf{A}}^i\bar{\mathbf{B}} = 0$ for $i \geq p$. The

so-called *observer Markov parameters*, which are the Markov parameters of the equivalent system defined in Eq.(1.42), can be determined from Eq.(1.46)

$$\bar{\mathbf{Y}} = \mathcal{Y}\mathbf{V}^{+} \tag{1.46}$$

where

$$\mathcal{Y} = \begin{bmatrix} \mathbf{y}(0) & \mathbf{y}(1) & \mathbf{y}(2) & \cdots & \mathbf{y}(p) & \cdots & \mathbf{y}(l-1) \end{bmatrix} \tag{1.47}$$

$$\bar{\mathbf{Y}} = \begin{bmatrix} \mathbf{D} & \mathbf{C}\bar{\mathbf{B}} & \mathbf{C}\bar{\mathbf{A}}\bar{\mathbf{B}} & \cdots & \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}} \end{bmatrix} \tag{1.48}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{u}(0) & \mathbf{u}(1) & \mathbf{u}(2) & \cdots & \mathbf{u}(p) & \cdots & \mathbf{u}(l-1) \\ & \mathbf{v}(0) & \mathbf{v}(1) & \cdots & \mathbf{v}(p-1) & \cdots & \mathbf{v}(l-2) \\ & & \mathbf{v}(0) & \cdots & \mathbf{v}(p-2) & \cdots & \mathbf{v}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \mathbf{v}(0) & \cdots & \mathbf{v}(l-p-1) \end{bmatrix} \tag{1.49}$$

Once the observer Markov parameters are determined, the system Markov parameters can be determined using a recursive formulation. For non-zero initial conditions, the input/output relationship for the system defined in Eq.(1.42) can be expressed as in Eq.(1.50)

$$\bar{\mathcal{Y}} = \mathbf{C}\bar{\mathbf{A}}^{p}\mathcal{X} + \bar{\mathbf{Y}}\bar{\mathbf{V}} \tag{1.50}$$

where

$$\bar{\mathcal{Y}} = \begin{bmatrix} \mathbf{y}(p+1) & \mathbf{y}(p+2) & \cdots & \mathbf{y}(l-1) \end{bmatrix} \tag{1.51}$$

$$\mathcal{X} = \begin{bmatrix} \mathbf{x}(0) & \mathbf{x}(1) & \cdots \mathbf{x}(l-p-2) \end{bmatrix} \tag{1.52}$$

$$\bar{\mathbf{Y}} = \begin{bmatrix} \mathbf{D} & \mathbf{C}\bar{\mathbf{B}} & \mathbf{C}\bar{\mathbf{A}}\bar{\mathbf{B}} & \cdots & \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}}) \end{bmatrix} \tag{1.53}$$

$$\bar{\mathbf{V}} = \begin{bmatrix} \mathbf{u}(p+1) & \mathbf{u}(p+2) & \cdots & \mathbf{u}(l-1) \\ \mathbf{v}(p) & \mathbf{v}(p+1) & \cdots & \mathbf{v}(l-2) \\ \mathbf{v}(p-1) & \mathbf{v}(p) & \cdots & \mathbf{v}(l-3) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}(0) & \mathbf{v}(1) & \cdots & \mathbf{v}(l-p-2) \end{bmatrix} \tag{1.54}$$

Since, by definition $\bar{\mathbf{A}}^{p} \approx 0$ and $\{\mathbf{x}(k) : \|\mathbf{x}(k)\| < \infty, \forall k \geq 0\}$, the term $\mathbf{C}\bar{\mathbf{A}}^{p}\mathcal{X}$ in Eq.(1.50) can be neglected as in Eq.(1.55).

$$\bar{\mathcal{Y}} = \bar{\mathbf{Y}}\bar{\mathbf{V}} \tag{1.55}$$

The observer Markov parameters can be obtained using Eq.(1.56).

$$\bar{\mathcal{Y}}\bar{\mathbf{V}}^{+} = \bar{\mathbf{Y}} \tag{1.56}$$

Note that, for a *unique* solution, the matrix $\bar{\mathbf{V}}$ must have unique inverse, which is only possible if it is full rank. Otherwise, although the identified Markov parameters will be able to reproduce the measured output via Eq.(1.55), they will not be able to predict the output time-history correctly, when the input is different. As in the case of the proposed methods of bilinear system identification, if the input excitation is sinusoidal (i.e. not rich), due to the row shifting involved, the matrix $\bar{\mathbf{V}}$ has a tendency to have some of its rows linearly dependent. One way of handling this problem numerically is by adding a subtle amount of White Gaussian Noise (WGN) on the input so that $\bar{\mathbf{V}}$ is full rank, but the change in $\bar{\mathbf{Y}}$ is negligible. Another approach is to minimize the number of row-shifts by keeping the integer p to a minimum. The lower bound on p can be derived from a relationship between the observer Markov parameters and system Markov parameters.

For this purpose, the observer Markov parameters can be partitioned as follows:

$$\bar{\mathbf{Y}} = \begin{bmatrix} \bar{\mathbf{Y}}_{-1} & \bar{\mathbf{Y}}_0 & \bar{\mathbf{Y}}_1 \cdots & \bar{\mathbf{Y}}_{p-1} \end{bmatrix} \tag{1.57}$$

where

$$\bar{\mathbf{Y}}_k = \mathbf{C}\bar{\mathbf{A}}^k\bar{\mathbf{B}} \tag{1.58}$$

$$= \begin{bmatrix} \mathbf{C}(\mathbf{A}+\mathbf{KC})^k(\mathbf{B}+\mathbf{KD}) & -\mathbf{C}(\mathbf{A}+\mathbf{KC})^k\mathbf{M} \end{bmatrix} \tag{1.59}$$

$$= \begin{bmatrix} \bar{\mathbf{Y}}_k^{(1)} & \bar{\mathbf{Y}}_k^{(2)} \end{bmatrix}; \quad k = 0, 1, 2... \tag{1.60}$$

$$\bar{\mathbf{Y}}_k^{(-1)} = \mathbf{D} \tag{1.61}$$

As also presented in Ref.[8], the system Markov parameters can be obtained from the recursive relationship:

$$\mathbf{Y}_k = \bar{\mathbf{Y}}_k^{(1)} + \sum_{i=0}^{k-1} \bar{\mathbf{Y}}_i^{(2)}\mathbf{Y}_{k-i-1} + \bar{\mathbf{Y}}_k^{(2)}\mathbf{D} \tag{1.62}$$

To impose the lower bound on p, the relationship between observer Markov parameters and system Markov parameters should be viewed as follows:

$$\bar{\mathbf{Y}}^{(2)}\mathbf{H} = \tilde{\mathbf{Y}} \tag{1.63}$$

where

$$\bar{\mathbf{Y}}^{(2)} = \begin{bmatrix} -\bar{\mathbf{Y}}_{p-1}^{(2)} & -\bar{\mathbf{Y}}_{p-2}^{(2)} & \cdots & -\bar{\mathbf{Y}}_{0}^{(2)} \end{bmatrix} \tag{1.64}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 & \cdots & \mathbf{Y}_N \\ \mathbf{Y}_2 & \mathbf{Y}_3 & \cdots & \mathbf{Y}_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}_p & \mathbf{Y}_{p+1} & \cdots & \mathbf{Y}_{N+p+1} \end{bmatrix} \tag{1.65}$$

$$\tilde{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y}_{p+1} & \mathbf{Y}_{p+1} & \cdots & \mathbf{Y}_{p+1} \end{bmatrix} \tag{1.66}$$

and N is a sufficiently large arbitrary integer, while $\mathbf{H}$ is the generalized Hankel matrix. The rank of a sufficiently large $\mathbf{H}$ is equal to the order of the both controllable and observable part of the system to be identified n. The matrix $\mathbf{H}$ is of size $qp \times Nm$. Assuming $Nm > qp$, the rank of $\mathbf{H}$ is at most qp. Thus, in order to be able to solve for $\bar{\mathbf{Y}}^{(2)}$ uniquely, p should be chosen such that $qp \geq n$. This condition imposes the lower bound on the smallest possible p as the number of outputs times p must be greater than or equal to the order of the system.

1.4 Organization of the Thesis

In this chapter, the reader is first introduced to bilinear systems. Various definitions, a stability criterion for the case of periodic excitations, sufficient conditions for observability and controllability, several fields of applications and frequently cited bilinear system identification methods are presented. Lastly, an LTI system identification algorithm called OKID is introduced for reference in Chapter 2.

Chapter 2 discusses the first proposed method for the identification of single-input, multi-output, discrete-time, bilinear systems that uses the steady-state portion of a single input/output time-history obtained by exciting the system with a linear combination of sine and cosine functions of user-selected frequencies enriched by a subtle amount of white Gaussian noise. The method relies on the conversion of the bilinear system into an equivalent linear model (ELM) by an accurate approximation of the state in the bilinear term using a set of sine and cosine basis functions. The frequencies of these functions are obtained using perturbation theory. The equivalent linear model can be identified using one of the

available algorithms for the identification of LTI systems (e.g. the OKID algorithm), and from this identified ELM, the original bilinear system can be recovered. The method can be generalized to multi-input system by increasing the number of experiments needed.

Chapter 3 presents the second proposed method for the identification of single-input, multi-output, discrete-time, bilinear systems using the steady-state portion of a single input/output time-history measurement. Similar to the first approach, the bilinear system is excited by a user-designed input in the form of a linear combination of sine and cosine functions, enriched by a subtle amount of highband random component, where the user is given the freedom of selecting the frequencies and amplitudes of the sine and cosine functions besides the amount of the added random component. The obtained ELM is converted to a new, overparameterized ELM using its past input/output time-history as its original state time-history. The order of the new ELM is reduced by discarding its unobservable states. Since the state time-history of the ELM is (known and) common to the original bilinear model, the original bilinear system matrices are recovered via a least-square fit.

Chapter 4 explains the third proposed method for the identification of single-input, multi-output, discrete-time, bilinear systems using the steady-state portion of a single input/output time-history measurement. As in the previous method, the bilinear system is fed a linear combination of sine and cosine functions, enriched by a subtle amount of highband random component. By choosing the frequencies of the sine and cosine functions and the amount of highband random component, the user designs the excitation input. To be valid for the specific input/output time-history, the bilinear model is converted to an ELM by representing the state in the bilinear term as a linear combination of sine and cosine basis functions, where the frequencies of these basis functions are selected via perturbation theory. The identified ELM is used to generate an approximation of the time-history of the state. Then, perturbation theory is used to generate a recursive relationship in every iteration of which the error in the identified state drops exponentially. Once the error in the identified state drops to a sufficiently negligibly small number, the corresponding ELM is used to identify the bilinear system.

Chapter 5 is devoted to the conclusions and remarks on future research.

Chapter 2

Identification of Bilinear Systems Through Equivalent Linear Models via Observer/Kalman Filter Identification (OKID) Algorithm

This chapter presents the first proposed method for the identification of single-input, multi-output, time-invariant, discrete-time bilinear systems using the steady-state portion of single-input, multi-output time-histories obtained by exciting the system with a linear combination of sine and cosine functions, of user-selected frequencies, enriched by a subtle amount of white Gaussian noise (WGN). The method relies on the conversion of the bilinear model into an equivalent linear model (ELM) by an accurate approximation of the state in the bilinear term using a set of sine and cosine basis functions whose frequencies are obtained from perturbation theory. The addition of the random component (RC), along with a step in which the columns of the ELM matrices which correspond to the dependent rows of the input vector are compressed into single columns, serves for obtaining an input vector with linearly independent rows. Lastly, the ELM is identified using an LTI system identification algorithm called Observer Kalman Filter Identification (OKID), and the bilinear model is retrieved from the identified ELM.

2.1 Notation and Basic Formulation

Consider a single-input, multi-output, time-invariant, discrete-time bilinear system represented by a state-space model of order n as follows:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \mathbf{N}\mathbf{x}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}u(k), \quad k = 0, 1, ..., s, s+1, ..., l \end{aligned} \tag{2.1}$$

where k denotes the time-step, τ is the time *sampling interval* such that $\mathbf{x}(k)$ implies $\mathbf{x}(k\tau)$, $\mathbf{x}(k) \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}$ is the excitation input, while $\mathbf{y}(k) \in \mathbb{R}^r$ represents the $r \times 1$ output vector. $\mathbf{A} \in \mathbb{R}^{n \times n}$ represents the state matrix, $\mathbf{B} \in \mathbb{R}^n$ is the $n \times 1$ input matrix, $\mathbf{C} \in \mathbb{R}^{r \times n}$ is the output matrix, $\mathbf{D} \in \mathbb{R}^r$ is the direct transmission matrix and $\mathbf{N} \in \mathbb{R}^{n \times n}$ represents the matrix that weighs the bilinear term $\mathbf{x}(k)u(k)$. The steady state portion of the signal is assumed to start at the s^{th} time-step and the record of input/output time-histories end at the l^{th} time-step.

2.2 Linearization of The Bilinear Problem

Let's first define the excitation input to be used in the identification as

$$u(k) = \sum_{i=0}^{f} a_i \cos(\omega_i k\tau) + \sum_{i=1}^{f} b_i \sin(\omega_i k\tau) \tag{2.2}$$

where $a_i, b_i \in \mathbb{R}$ and $k = 0, 1, ..., l$. The amplitudes a_i, b_i and frequencies ω_i $(i = 0, 1, .., f)$ are chosen by the user, and the frequencies can be grouped in a set denoted by $\Omega_0 = \{\omega_0, \omega_1, ..., \omega_f\}$, where $\omega_0 = 0$. The input defined in Eq.(2.2) is applied to the bilinear system defined in Eq.(2.1), and the time-history of the output vector is recorded.

For this specific input excitation, Eq.(2.1) can be reduced to an ELM capable of reproducing the steady-state portion of the recorded output time-history. Such an ELM is obtained by approximating the state $\mathbf{x}(k)$ in the bilinear term $\mathbf{N}\mathbf{x}(k)u(k)$ using a set of sine and cosine basis functions whose frequencies are obtained by *perturbation theory*.

2.2.1 Perturbation Theory

Perturbation theory can be used to find an approximate solution to a problem that has no analytic solution by adding corrections to the closed-form solution of a related problem if the related problem is convertible to the original one by addition of a perturbation term. The approximate solution is written as a truncated power series in terms of the parameter $\epsilon \in \mathbb{R}$, which controls the magnitude of the added perturbation term and corrections.

In the case of the bilinear system defined in Eq.(2.1), a closed-form solution to the state exists for the linear system associated with it; and a perturbation $\epsilon \mathbf{N}\mathbf{x}(k)u(k)$ can be added to it as in Eq.(2.3). Thus, the solution to the state of the bilinear system can be approximated in the form of a power series in the parameter ϵ, as in Eq.(2.4)

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \epsilon \mathbf{N}\mathbf{x}(k)u(k) \tag{2.3}$$

$$\mathbf{x}(k) = \sum_{i=0}^{\infty} \epsilon^i \boldsymbol{\chi}_i(k) \approx \sum_{i=0}^{N} \epsilon^i \boldsymbol{\chi}_i(k) \tag{2.4}$$

where $\boldsymbol{\chi}_0(k) \in \mathbb{R}^n$ is the solution to the linear problem obtained by setting $\epsilon = 0$; $\boldsymbol{\chi}_i(k)$ $(i = 1, 2, ..., N)$ are the basis functions in terms of which the state $\mathbf{x}(k)$ is expressed; the term $\epsilon^i \boldsymbol{\chi}_i(k)$ is the i^{th} order correction, and $\sum_{i=0}^{N} \epsilon^i \boldsymbol{\chi}_i(k)$ is the N^{th} order approximation to the solution of the state vector $\mathbf{x}(k)$ in Eq.(2.1).

Plugging Eq.(2.4) into Eq.(2.3) and equating the terms with equal powers of ϵ yield the set of equations in Eq.(2.5), which will be used to obtain a set of basis functions $\boldsymbol{\chi}_i(k)$ to approximate the state.

$$\begin{aligned}
\boldsymbol{\chi}_0(k+1) &= \mathbf{A}\boldsymbol{\chi}_0(k) + \mathbf{B}u(k) \\
\boldsymbol{\chi}_1(k+1) &= \mathbf{A}\boldsymbol{\chi}_1(k) + \mathbf{N}\boldsymbol{\chi}_0(k)u(k) \\
\boldsymbol{\chi}_2(k+1) &= \mathbf{A}\boldsymbol{\chi}_2(k) + \mathbf{N}\boldsymbol{\chi}_1(k)u(k) \\
&\vdots \\
\boldsymbol{\chi}_N(k+1) &= \mathbf{A}\boldsymbol{\chi}_N(k) + \mathbf{N}\boldsymbol{\chi}_{N-1}(k)u(k)
\end{aligned} \tag{2.5}$$

Starting from the first and moving to the next, a closed-form solution can be written for each linear equation. Recalling the form of the input excitation $u(k)$ in Eq.(2.2), one can

write $\chi_0(k)$ $(k \geq s)$ as a linear combination of sines and cosines of the set of frequencies Ω_0. Furthermore, using the trigonometric identities in Eq.(2.6), where $\alpha, \beta \in \mathbb{R}$ are any two scalars, the particular solution of $\chi_1(k)$ can be written as a linear combination of sines and cosines of a set of frequencies Ω_1, where Ω_1 contains the sums and differences of the frequencies in Ω_0 with each other.

$$\begin{aligned}
\cos(\alpha)\cos(\beta) &= [\cos(\alpha-\beta)+\cos(\alpha+\beta)]/2 \\
\sin(\alpha)\sin(\beta) &= [\cos(\alpha-\beta)-\cos(\alpha+\beta)]/2 \\
\sin(\alpha)\cos(\beta) &= [\sin(\alpha+\beta)+\sin(\alpha-\beta)]/2 \\
\cos(\alpha)\sin(\beta) &= [\sin(\alpha+\beta)-\sin(\alpha-\beta)]/2
\end{aligned} \tag{2.6}$$

Similarly, for $k \geq s$, $\chi_2(k)$ can be written as a linear combination of sines and cosines of a set of frequencies Ω_2, where Ω_2 contains the sums and differences of the frequencies in Ω_0 with the frequencies in Ω_1. This reasoning can be extended to $\chi_N(k)$ and Ω_N to conclude that, if Eq.(2.4) is true, then the steady-state portion of the state vector $\mathbf{x}(k)$ can be approximated by a linear combination of sines and cosines of the frequencies from the set $\Omega = \{\Omega_0 \cup \Omega_1 \cup ... \cup \Omega_N\} = \{0, \hat{\omega}_1, \hat{\omega}_2, ..., \hat{\omega}_q\}$. Hence, the state vector $\mathbf{x}(k)$ for the bilinear system in Eq.(2.1) can be represented in the steady-state condition as follows:

$$\mathbf{x}(k) = \begin{pmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{pmatrix} \approx \mathbf{\Lambda\Phi}(k) = \begin{bmatrix} \lambda_{10} & \lambda_{11} & \cdots & \lambda_{1(2q)} \\ \lambda_{20} & \lambda_{21} & \cdots & \lambda_{2(2q)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n0} & \lambda_{n1} & \cdots & \lambda_{n(2q)} \end{bmatrix} \begin{pmatrix} \phi_0(k) \\ \phi_1(k) \\ \vdots \\ \phi_{2q}(k) \end{pmatrix}, \quad k = s, s+1, ..., l \tag{2.7}$$

where $x_i(k) \in \mathbb{R}$ is the i^{th} component of the state vector $\mathbf{x}(k)$ (i.e. the i^{th} state), $\phi_0(k) = 1$, $\phi_i(k) = \sin(\hat{\omega}_i k\tau)$ while $\phi_{q+i}(k) = \cos(\hat{\omega}_i k\tau)$ for $i = 1, 2, .., q$. The vector $\mathbf{\Phi}(k) \in \mathbb{R}^{(2q+1)}$ represents the vector of the aforementioned basis functions, and $\mathbf{\Lambda} \in \mathbb{R}^{n\times(2q+1)}$ is the matrix containing the coefficients multiplying the basis functions.

Substituting Eq.(2.7) into the state $\mathbf{x}(k)$ in the bilinear term $\mathbf{N}\mathbf{x}(k)u(k)$ of Eq.(2.1)

yields the following ELM:

$$\begin{aligned}
\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \begin{bmatrix}\mathbf{B} & \mathbf{N}\end{bmatrix}\begin{pmatrix} u(k) \\ \mathbf{x}(k)u(k)\end{pmatrix} \\
&= \mathbf{A}\mathbf{x}(k) + \begin{bmatrix}\mathbf{B} & \mathbf{N}\mathbf{\Lambda}\end{bmatrix}\begin{pmatrix} u(k) \\ \mathbf{\Phi}(k)u(k)\end{pmatrix} \\
\mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \begin{bmatrix}\mathbf{D} & \mathbf{0}\end{bmatrix}\begin{pmatrix} u(k) \\ \mathbf{\Phi}(k)u(k)\end{pmatrix}, \quad k = s, s+1, ..., l
\end{aligned} \tag{2.8}$$

where $\mathbf{0} \in \mathbb{R}^{r\times(2q+1)}$.

2.3 Identification of the ELM

Although the ELM in Eq.(4.15) is capable of reproducing, for the same input, the same output as the bilinear system after the steady-state is reached, such an ELM cannot be identified *uniquely* because, at least, the first two rows of its input vector are linearly dependent, as shown in Eq.(2.9)

$$\begin{pmatrix} u(k) \\ \mathbf{\Phi}(k)u(k)\end{pmatrix} = \begin{pmatrix} u(k) \\ \phi_0(k)u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k)\end{pmatrix} = \begin{pmatrix} u(k) \\ u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k)\end{pmatrix} \tag{2.9}$$

However, this problem can be fixed by compressing the columns of the system matrices

that correspond to the repeated rows of the input vector into single columns as follows:

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \begin{bmatrix}\mathbf{B}+\mathbf{N}\boldsymbol{\Lambda}_1 & \mathbf{N}\boldsymbol{\Lambda}_2\end{bmatrix}\begin{pmatrix} u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k)\end{pmatrix} \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \begin{bmatrix}\mathbf{D} & \mathbf{0}\end{bmatrix}\begin{pmatrix} u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k)\end{pmatrix}, \quad k = s, s+1, ..., l\end{aligned} \tag{2.10}$$

where $\mathbf{0} \in \mathbb{R}^{r\times(2q)}$, $\boldsymbol{\Lambda}_1 \in \mathbb{R}^n$ is the first column of the coefficient matrix $\boldsymbol{\Lambda}$, and $\boldsymbol{\Lambda}_2 \in \mathbb{R}^{n\times(2q)}$ is the matrix of the remaining columns so that $\boldsymbol{\Lambda} = \begin{bmatrix}\boldsymbol{\Lambda}_1 & \boldsymbol{\Lambda}_2\end{bmatrix}$. To identify the ELM in Eq.(2.10), one needs to build the input vector time-history numerically and feed it, along with the recorded output-time histories, to a linear system identification tool. This tool should demonstrate robustness against noise (both measurement and process noise), as there will be a certain amount of numerical noise that depends on the value of N chosen in truncating the power series in Eq.(2.4), even if the record of the input/output time-histories contain no measurement noise.

In the bilinear system identification approach proposed here, the Observer Kalman Filter Identification (OKID) is used for its robustness against measurement and process noise. The details regarding OKID were already discussed in the section 1.3, on page 15. One of the steps of the OKID algorithm requires the matrix $\bar{\mathbf{V}}$ in Eq.(1.50), which corresponds to the matrix $\mathbf{V}$ in Eq.(2.11), to be full rank

$$\mathbf{V} = \begin{pmatrix} \boldsymbol{\Phi}(s+p+1)u(s+p+1) & \boldsymbol{\Phi}(s+p+2)u(s+p+2) & \cdots & \boldsymbol{\Phi}(l-1)u(l-1) \\ \boldsymbol{\Psi}(s+p) & \boldsymbol{\Psi}(s+p+1) & \cdots & \boldsymbol{\Psi}(l-2) \\ \boldsymbol{\Psi}(s+p-1) & \boldsymbol{\Psi}(s+p) & \cdots & \boldsymbol{\Psi}(l-3) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}(s) & \boldsymbol{\Psi}(s+1) & \cdots & \boldsymbol{\Psi}(l-p-2) \end{pmatrix} \tag{2.11}$$

where

$$\mathbf{\Psi}(k) = \begin{pmatrix} \mathbf{\Phi}(k)u(k) \\ \mathbf{y}(k) \end{pmatrix}, \tag{2.12}$$

as shown before, the parameter $p \in \mathbb{Z}^+$ must be greater than or equal to the product of the number of outputs r and the user's guess for the order of the system. This requirement is typically satisfied by using an input "rich" enough in terms of frequency content, together with sufficiently long time-history records. Yet, in the proposed formulation, it is very difficult to satisfy the richness condition in order to uncorrelate the rows of $\mathbf{V}$ in Eq.(2.11). Increasing the frequency content of the excitation input leads to the generation of new basis function frequencies $\hat{\omega}_i$, in accordance with Eq.(2.5); consequently, it increases the number of rows of $\mathbf{\Phi}(k)$, and therefore, the number of rows of $\mathbf{V}$. Thus, due to the identities in Eq.(2.6), some of the rows of the matrix $\mathbf{V}$ may remain or become linearly dependent or numerically close to being so.

In order to maintain the linear dependence between rows of $\mathbf{V}$, it is enough to add a subtle amount of random component (RC) in the form of white Gaussian noise (WGN) to the input time-history $u(k)$ to make sure that the matrix $\mathbf{V}$ is full rank. Hence, the final form of the excitation input becomes

$$\hat{u}(k) = u(k) + \text{wgn}(k) \tag{2.13}$$

where $\text{wgn}(k)$ stands for the faint amount of WGN added. The precise amount of the WGN depends on the specifics of the problem in question such as on the parameter p in the matrix $\mathbf{V}$ that controls the number of shifts, which inherently depends on the number of outputs r and on the order of the system to be identified, n, through the requirement $p \geq rn$, on the number of basis functions preferred, on the frequency ranges the excitation input and basis functions cover, and so forth. Although Eq.(2.13) adversely affects the error made in approximating the state in Eq.(2.7), with the assumption that the excitation input is of the form given in Eq.(2.2), the increase in the error is not significant, since the amount of added RC is quite small. In addition, the robustness of OKID in handling noisy data and the ability of the basis functions to absorb the unaccounted changes in the state vector due to the added RC contribute to keep the overall identification error low.

In summary, the bilinear system to be identified is excited with an input $\hat{u}(k)$ of the form elaborated in Eq.(2.13) and (2.2) (i.e. a combination of sinusoidal functions with a small random component), and the output time-history is recorded. Next, the input time-histories for $\mathbf{\Phi}(k)\hat{u}(k)$ are numerically generated to serve as the input vector of the ELM. Finally, the steady-state portions of the numerically generated input time-histories and of the recorded output time-history are fed into OKID to identify the corresponding ELM from which the bilinear system will then be recovered.

2.4 Recovery of the Bilinear System from the Identified ELM

Let $\mathbf{A}_{ok} \in \mathbb{R}^{n\times n}$, $\mathbf{B}_{ok} \in \mathbb{R}^{n\times(2q+1)}$, $\mathbf{C}_{ok} \in \mathbb{R}^{r\times n}$, $\mathbf{D}_{ok} \in \mathbb{R}^{r\times 1}$, be the identifed ELM matrices obtained using OKID, and $\mathbf{x}_{ok}(k) \in \mathbb{R}^n$ be the corresponding state vector so that

$$\begin{aligned} \mathbf{x}_{ok}(k+1) &= \mathbf{A}_{ok}\mathbf{x}_{ok}(k) + \mathbf{B}_{ok}\mathbf{\Phi}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C}_{ok}\mathbf{x}_{ok}(k) + \mathbf{D}_{ok}\mathbf{\Phi}(k)u(k) \end{aligned} \tag{2.14}$$

Let also $\mathbf{T} \in \mathbb{R}^{n\times n}$ be an invertible coordinate transformation matrix as in Eq.(2.15), which allows the transformation from the state vector of the ELM identified by OKID to the original state vector of the bilinear system.

$$\mathbf{x}_{ok}(k) = \mathbf{T}\mathbf{x}(k) \tag{2.15}$$

Then, the bilinear system in Eq.(2.1) can be transformed into Eq.(2.16):

$$\begin{aligned} \mathbf{x}_{ok}(k+1) &= \tilde{\mathbf{A}}\mathbf{x}_{ok}(k) + \tilde{\mathbf{B}}u(k) + \tilde{\mathbf{N}}\mathbf{x}_{ok}(k)u(k) \\ \mathbf{y}(k) &= \tilde{\mathbf{C}}\mathbf{x}_{ok}(k) + \mathbf{D}u(k), \quad k = 0, 1, ..., s, s+1, ..., l \end{aligned} \tag{2.16}$$

where $\tilde{\mathbf{A}} \in \mathbb{R}^{n\times n}$, $\tilde{\mathbf{B}} \in \mathbb{R}^{n\times 1}$, $\tilde{\mathbf{N}} \in \mathbb{R}^{n\times n}$, $\tilde{\mathbf{C}} \in \mathbb{R}^{r\times n}$ are the system matrices in the same coordinates as the identified ELM, and

$$\begin{aligned} \mathbf{A} &= \mathbf{T}^{-1}\tilde{\mathbf{A}}\mathbf{T} \\ \mathbf{B} &= \mathbf{T}^{-1}\tilde{\mathbf{B}} \\ \mathbf{N} &= \mathbf{T}^{-1}\tilde{\mathbf{N}}\mathbf{T} \\ \mathbf{C} &= \tilde{\mathbf{C}}\mathbf{T} \end{aligned} \tag{2.17}$$

Note that both Eq.(2.1) and (2.16) define the same input/output mapping, and in accordance with Eq.(2.7), the state vector of the identified ELM $\mathbf{x}_{ok}(k)$ can also be expressed as follows:

$$\mathbf{x}_{ok}(k) \approx \tilde{\mathbf{\Lambda}}\mathbf{\Phi}(k) \tag{2.18}$$

where $\tilde{\mathbf{\Lambda}} = \mathbf{T}\mathbf{\Lambda}$ is the matrix of the coefficients that multiply the vector of basis functions $\mathbf{\Phi}(k)$.

Then, to conclude the identification, the set of equations in Eq.(2.19) can be used to recover the bilinear system in Eq.(2.16) from the identified ELM

$$\begin{aligned}
\tilde{\mathbf{A}} &= \mathbf{A}_{ok} \\
\tilde{\mathbf{N}} &= (\mathbf{B}_{ok})_2 \tilde{\mathbf{\Lambda}}_2^{+} \\
\tilde{\mathbf{B}} &= (\mathbf{B}_{ok})_1 - \tilde{\mathbf{N}}\tilde{\mathbf{\Lambda}}_1 \\
\tilde{\mathbf{C}} &= \mathbf{C}_{ok} \\
\mathbf{D} &= (\mathbf{D}_{ok})_1
\end{aligned} \tag{2.19}$$

where $\tilde{\mathbf{\Lambda}}_1$, $(\mathbf{B}_{ok})_1$, $(\mathbf{D}_{ok})_1$ are the vectors formed by the first columns of the matrices $\tilde{\mathbf{\Lambda}}$, $\mathbf{B}_{ok}$, $\mathbf{D}_{ok}$ respectively, while $\tilde{\mathbf{\Lambda}}_2$, $(\mathbf{B}_{ok})_2$ are the matrices formed by their remaining columns such that $\tilde{\mathbf{\Lambda}} = \begin{bmatrix} \tilde{\mathbf{\Lambda}}_1 & \tilde{\mathbf{\Lambda}}_2 \end{bmatrix}$, $\mathbf{B}_{ok} = \begin{bmatrix} (\mathbf{B}_{ok})_1 & (\mathbf{B}_{ok})_2 \end{bmatrix}$. The superscript $^{+}$ indicates the Moore-Penrose pseudoinverse. To utilize Eq.(2.19), $\tilde{\mathbf{\Lambda}}$ should be determined first. This can be achieved by using the identified ELM as a state generator for $\mathbf{x}_{ok}(k)$ by feeding it back the same input time-history it was identified with:

$$\tilde{\mathbf{\Lambda}} = \begin{bmatrix} \mathbf{x}_{ok}(s) & \mathbf{x}_{ok}(s+1) & \cdots & \mathbf{x}_{ok}(l) \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}(s) & \mathbf{\Phi}(s+1) & \cdots & \mathbf{\Phi}(l) \end{bmatrix}^{+} \tag{2.20}$$

Note that both the identified bilinear system and the original bilinear system represent the same input/output mapping, even though the identified matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\tilde{\mathbf{N}}$, $\tilde{\mathbf{C}}$ are not the same as the original system matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{N}$, $\mathbf{C}$, respectively.

2.5 The Effect of The Parameters The User Can Choose

In the implementation of such an identification approach, certain parameters have to be selected by the user. In this section, the analyses of the effects of each parameter on the

identification results are presented.

2.5.1 Length of Input/Output Time-Histories

The effect of using a longer input/output time-history record is two-fold. On one hand, a longer data set helps reducing the ill-conditioning of the matrix $\mathbf{V}$ in Eq.(2.11), and therefore, enhances the accuracy of the identification by diminishing the required amount of WGN to be added to the input. On the other hand, the approximation error made in Eq.(2.7) gets larger because the truncation in Eq.(2.4) leaves out some of the frequencies that should actually be in the set of basis function frequencies. Therefore, as the data length increases, the limited ability of the basis functions to compensate for the missing basis functions becomes more evident. This phenomenon can be better understood through a simple, non-dimensional example. Consider the following signal:

$$y(k) = \sin(2k\tau) + \sin(3k\tau) + \sin(4k\tau) \tag{2.21}$$

where the sampling interval $\tau = 0.01$. The signal is approximated by $\hat{y}(k)$ which is a linear combination of sines and cosines of the set of frequencies $\{1, 2, 4, 5, 6, 7\}$ obtained by performing a least square fit over the following data lengths: $l = \{400, 500, 1000, 2000\}$. Fig.2.1 shows that the quality of the fit deteriorates as the length of the data set increases. The declining accuracy in the fit is due to the fact $\sin(3k\tau)$ is present in $y(k)$ but not present in the set of sine and cosine basis functions from which $\hat{y}(k)$ is formed, and as the data length increases this fact becomes more and more evident from the increasing equation error in the least square fit.

2.5.2 Amount of Random Component

Similar to the effects of the data length, the effect of adding a certain amount of RC in the form of WGN to the sinusoidal input $u(k)$ is twofold. In fact, while the addition of a higher amount of RC makes the matrix $\mathbf{V}$ in Eq.(2.11) less ill-conditioned, and therefore, improves the quality of the identification, it also increases the approximation error made in

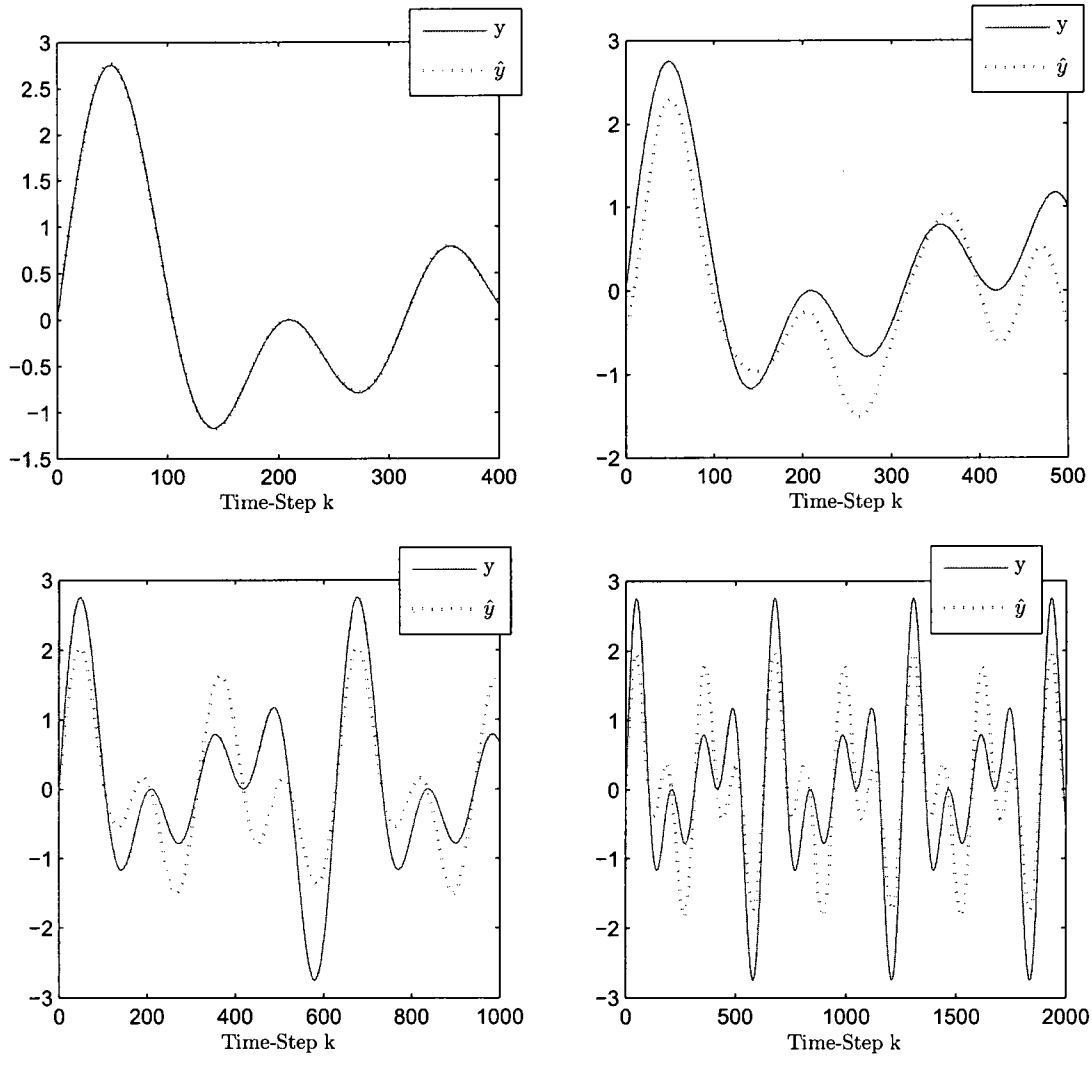


Figure 2.1: The Performance of Least-Square Fitting For Different Data Lengths

Eq.(2.7), since the finite set of basis function frequencies cannot include all of the frequencies contained by the RC.

Let's illustrate this point through a dimensionless example. Consider a single-input, single-output (SISO), single-state bilinear system of the form in Eq.(2.1), where the system matrices are as defined in Eq.(2.22).

$$\mathbf{A} = 0.85, \quad \mathbf{B} = 0.2, \quad \mathbf{N} = 0.2, \quad \mathbf{C} = 1, \quad \mathbf{D} = 0.1 \tag{2.22}$$

For comparison purposes, 4 different states $\mathbf{x}_{RC}(k)$ and outputs $\mathbf{y}_{RC}(k)$ are simulated by feeding 4 corresponding inputs $\hat{u}_{RC}(k)$ constructed in accordance with Eq.(2.23) and (2.24).

$$u(k) = \sin(5k\tau) + \sin(7k\tau) + \sin(12k\tau) \tag{2.23}$$

$$\hat{u}_{RC}(k) = u(k) + \text{wgn}_{RC}(k) \tag{2.24}$$

where $k = 1, 2, ..., 5000$, the time-step at which the steady-state starts is $s = 305$, and the time-step at which the input/output record ends is $l = 5000$, with a sampling period $\tau = 0.01$. The function $u(k)$ is the base input signal before enrichment and it is modified by adding WGN via the function $\text{wgn}_{RC}(k)$, with the percentage RC representing the ratio of the root-mean-square (RMS) of the added WGN noise to the RMS of the base signal $u(k)$. In this example, the values of RC is equal to $0\%, 0.0005\%, 0.05\%, 5\%$ have been used. RC is used as a subscript to denote the associated input, output and state with the corresponding amount of RC. The matrix $\mathbf{V}$ is formed for each case for perturbation order $N = 2$ and the number of row-block shifts $p = 2$.

Fig.2.2 shows the plots of the singular values of the matrix $\mathbf{V}$ from which the reduction in the level of ill-conditioning can easily be deduced by observing that the singular values close to zero (i.e. those causing ill-conditioning) start approaching the highest singular values as the added RC is increased. This observation can be quantified by calculating, the *2-norm* condition number of $\mathbf{V}$, which is defined as the ratio of the highest singular value to the smallest singular value and used it as a quantitative measure of the ill-conditioning. Table 2.1 shows the 2 norm condition numbers of $\mathbf{V}$, denoted by $K_2(\mathbf{V})$, for the cases presented in Fig.2.2.

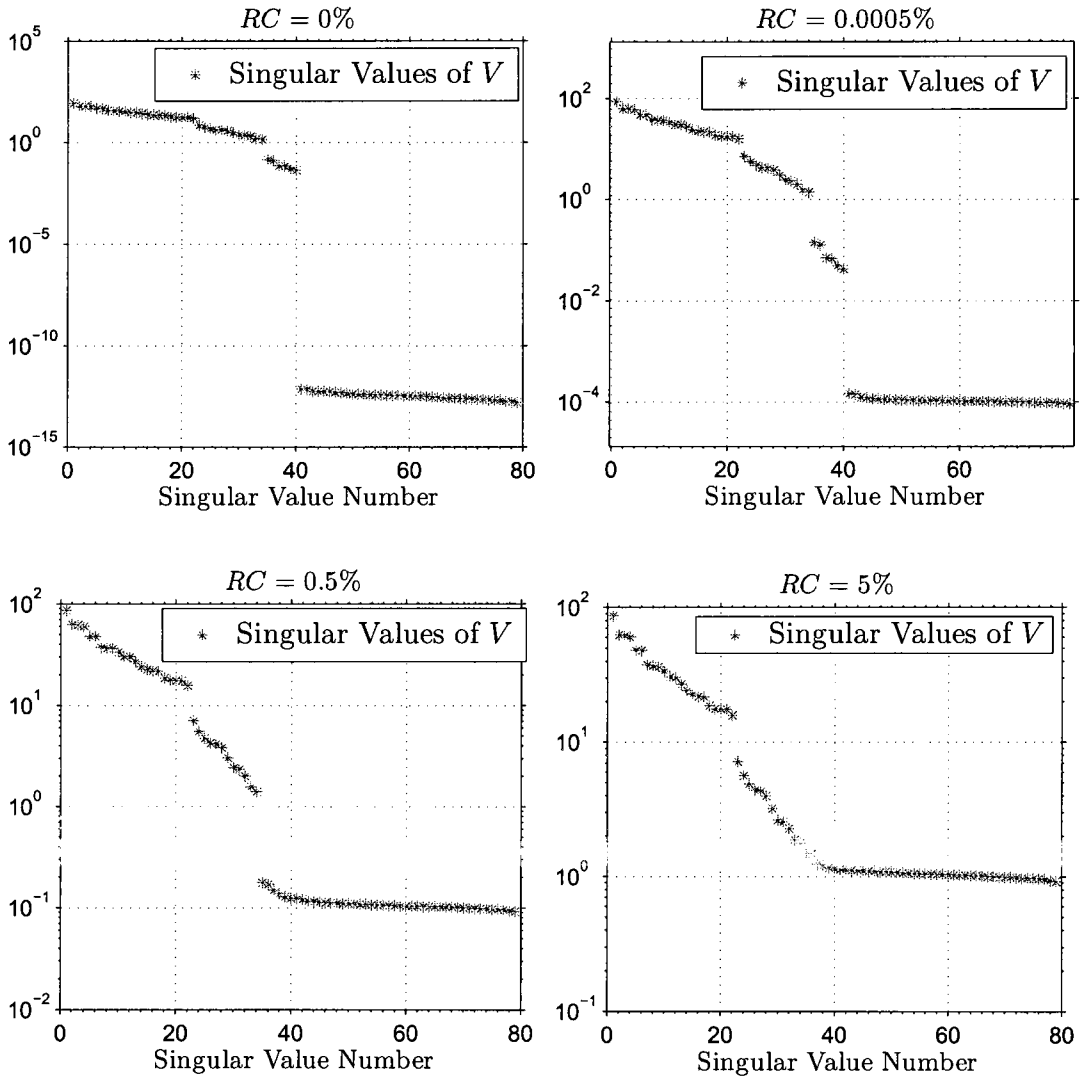


Figure 2.2: The effect of the enrichment of the input by addition of different amounts of WGN on the matrix **V**

RC	$K_2(\mathbf{V})$
0%	5.4053e14
0.0005%	9.3999e5
0.05%	9.4021e2
5%	9.4216e1

Table 2.1: The effect of different amounts of RC on the condition number of $K_2\mathbf{V}$

The adverse effects of the RC on the approximation error of the state in Eq.(2.7) can be quantified by comparing the steady-state portion of the simulated state $\mathbf{x}_{RC}$ to its representation in terms of the basis functions, $\hat{\mathbf{x}}_{RC\%}$, through an error function defined as in Eq.(2.25)

$$error(\mathbf{x}_{RC}, \hat{\mathbf{x}}_{RC}) = \frac{\text{RMS}(\mathbf{x}_{RC} - \hat{\mathbf{x}}_{RC})}{\text{RMS}(\mathbf{x}_{RC})} \tag{2.25}$$

where

$$\hat{\mathbf{x}}_{RC} = \mathbf{x}_{RC}\mathbf{\Phi}^{+}\mathbf{\Phi} \tag{2.26}$$

since

$$\begin{aligned} \hat{\mathbf{x}}_{RC} &= \hat{\mathbf{\Lambda}}\mathbf{\Phi} \\ \mathbf{x}_{RC} &= \mathbf{\Lambda}\mathbf{\Phi} \end{aligned} \tag{2.27}$$

and

$$\begin{aligned} \mathbf{x}_{RC} &= \begin{bmatrix} \mathbf{x}_{RC}(s) & \mathbf{x}_{RC}(s+1) & \cdots & \mathbf{x}_{RC}(l) \end{bmatrix} \\ \mathbf{\Phi} &= \begin{bmatrix} \mathbf{\Phi}(s) & \mathbf{\Phi}(s+1) & \cdots & \mathbf{\Phi}(l) \end{bmatrix} \end{aligned} \tag{2.28}$$

It can be inferred from Table 2.2 that, as the amount of RC increases, the state approximation error increases as well.

2.5.3 OKID Related Parameters

The most important OKID parameter that affects the computational cost and the performance of this method is the number of time-steps p, since the matrix $\mathbf{V}$, which needs to be invertible for a unique solution, contains $p+1$ many shifted row-blocks, each of which may

RC	$error(\mathbf{x}_{RC\%}, \hat{\mathbf{x}}_{RC\%})$
1%	3.71%
2%	3.76%
3%	3.84%
4%	3.96%
5%	4.10%

Table 2.2: The effect of different amounts of RC on the state approximation error

contain rows that are linearly dependent due to the identities in Eq.(2.6). Hence, keeping p to a minimum is essential to avoid ill-conditioning of the matrix $\mathbf{V}$. As discussed earlier in subsection 1.3, the requirement for minimum p is such that the number of inputs times p is greater than or equal to the order of the system to be identified. As an illustration of this discussion, Fig.2.3 provides the singular value plots of the matrix $\mathbf{V}$, for various values of p, for the example problem discussed above, setting $RC = 1\%$ and perturbation order $N = 2$. It can be observed that as p increases, (i) although the largest singular value does not change significantly, the smallest singular value gets smaller in value, and (ii) the number of singular values close to the smallest singular value increases. Both of these observations imply that $\mathbf{V}$ gets more ill-conditioned, as p increases. The corresponding 2-norm condition numbers of $\mathbf{V}$, denoted by $K_2(\mathbf{V})$, are tabulated in Table 2.3.

p	$K_2(\mathbf{V})$
1	322.2
2	417.3
3	493.2
4	560.4

Table 2.3: The effect of p on the condition number of $\mathbf{V}$

On the other hand, a higher value of p allows OKID to remain robust under higher amounts of noise, which also includes the process noise due to the approximation error in

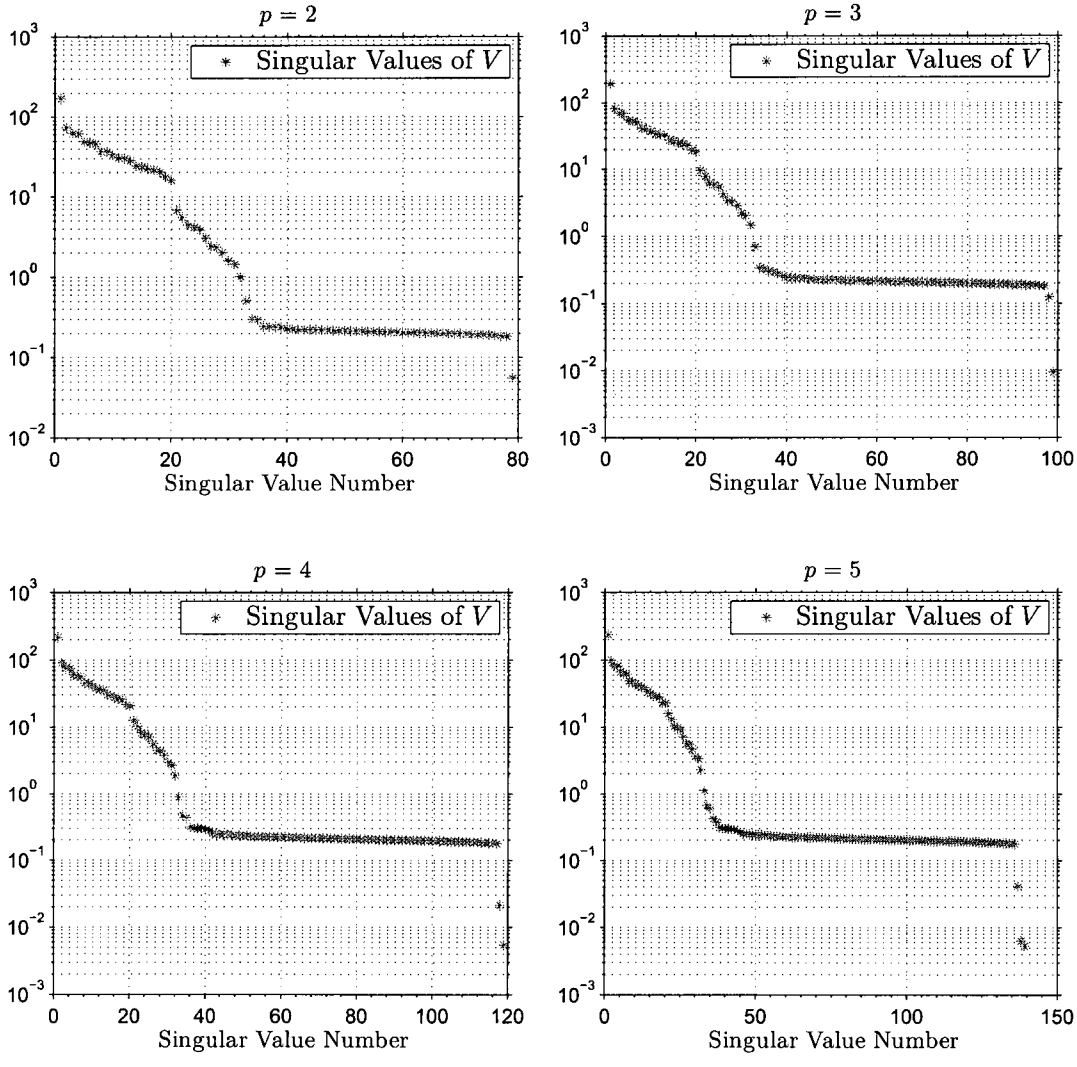


Figure 2.3: The effect of p on the ill-condition of the matrix $\mathbf{V}$

Eq.(2.7).

2.5.4 Perturbation Order N

The use of a higher perturbation order N improves the quality of the approximation of the state in Eq.(2.7) by increasing the number of basis functions used. For the example problem given above, setting $RC = 1\%$, the steady-state portion of the simulated state $\mathbf{x}_{1\%}$ can be compared to its representation in terms of the basis functions, $\hat{\mathbf{x}}_{1\%}$, through the error function defined in Eq.(2.25), for different perturbation orders $N = 2, 3, 4, 5$. Fig.2.4 and Table 2.4 present the result of such a comparison, indicating that a higher perturbation order N significantly reduces the approximation error of the state in Eq.(2.7). Furthermore, the reduction in error is not due to the enhancement of the quality of the approximation of the state over a specific time interval but over the entire time interval in a global sense.

N	$error(\mathbf{x}_{1\%}, \hat{\mathbf{x}}_{1\%})$
1	36.69%
2	12.87%
3	3.63%
4	0.95%

Table 2.4: The effect of the perturbation order N on the state approximation error

On the other hand, a higher perturbation order N may cause the matrix $\mathbf{V}$ to become more ill-conditioned by leading to the generation of more basis functions. The vector of basis functions $\Phi(k)$ is multiplied by the input, which is formed by taking a linear combination of a subset of the set of basis functions, to form a row block, and the matrix $\mathbf{V}$ is formed by shifting the row-blocks. As a result, in accordance with Eq.(2.6), as the number of basis functions increases, the possibility of having linearly dependent rows in $\mathbf{V}$ also increases. Extending the previous example to illustrate this reasoning, the singular value plots of the matrix $\mathbf{V}$ for $RC = 1\%$ and $p = 2$ are formed for the perturbation orders $N = 1, 2, 3, 4$ respectively, as shown in Fig.2.5. The 2-norm condition numbers $K_2(\mathbf{V})$ are presented in

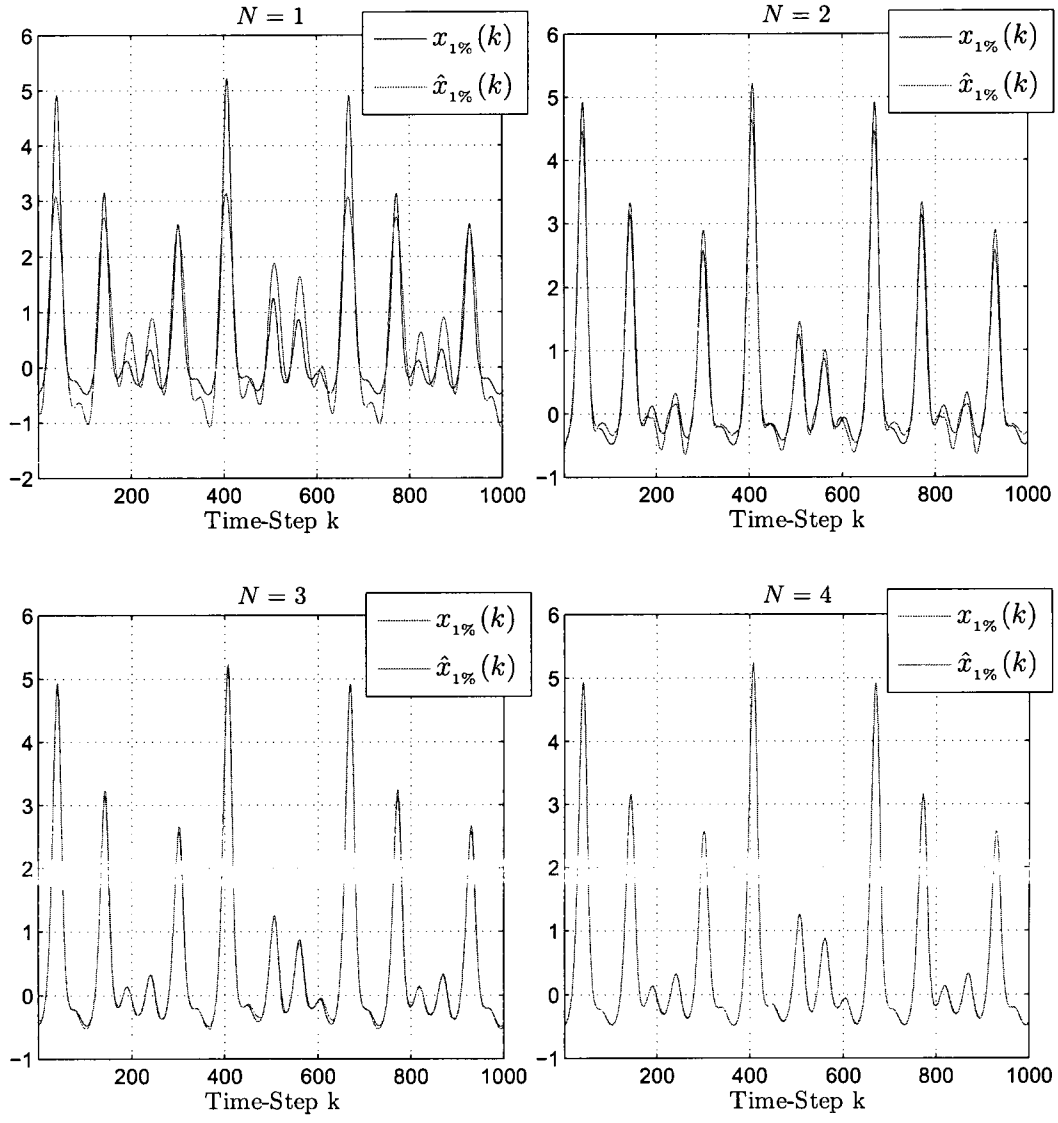


Figure 2.4: The effect of the perturbation order N on the state approximation via basis functions

Table 2.5 for each of these cases. The results indicate that, as the perturbation order N is increased, the matrix $\mathbf{V}$ becomes more ill-conditioned.

N	$K_2(\mathbf{V})$
1	256.0
2	371.2
3	348.9
4	380.7

Table 2.5: The effect of the perturbation order N on 2-norm condition number of the matrix $\mathbf{V}$

2.5.5 Choice of Ω_0

The input excitation frequencies should be chosen from a range that would allow the system dynamics to be observed, as in the case of linear system identification problems. However, unlike linear systems, increasing the number of excitation frequencies does not only make the input look "richer" (i.e. less periodic and so more desirable), but also increases the number of basis function frequencies obtained from perturbation theory.

Moreover, a set of Ω_0 formed by randomly selected frequencies is found to be more successful in avoiding the numerical ill-conditioning of the matrix $\mathbf{V}$ than one consisting of evenly spaced frequencies. Thus, Ω_0 should ideally be comprised of randomly selected frequencies. This behavior is due to the fact that a signal comprised of evenly spaced frequencies does not exhibit a rich behavior as much as a signal formed by randomly spaced frequencies. As an illustration of this discussion, consider two signals $u_r(k)$ and $u_e(k)$ formed by

$$\begin{aligned} u_r(k) &= \sum_{i=1}^{10} \sin(R_i k\tau) \\ u_e(k) &= \sum_{i=1}^{10} \sin(E_i k\tau) \end{aligned} \tag{2.29}$$

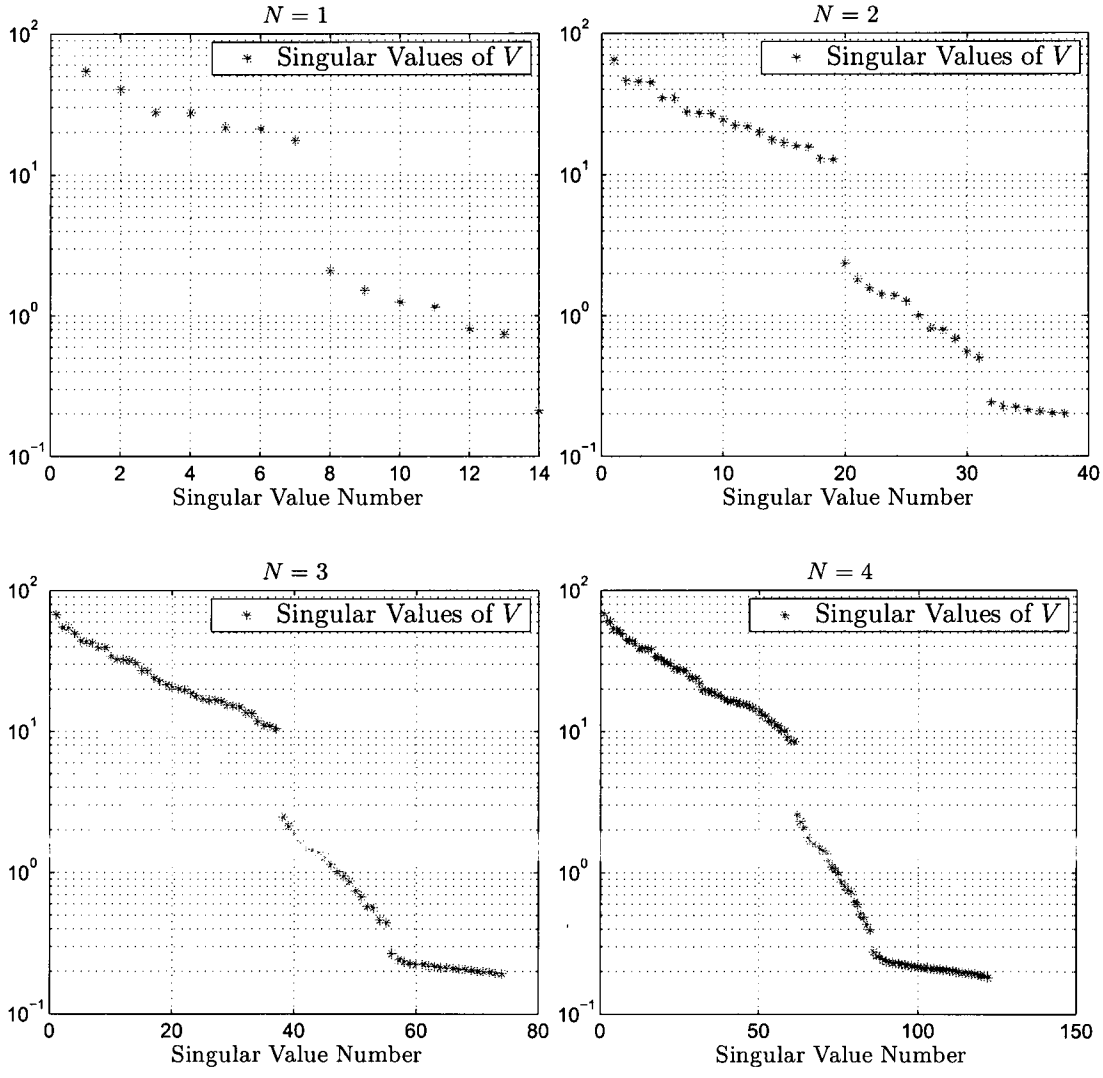


Figure 2.5: The singular value plot of the matrix $\mathbf{V}$, for various values of the perturbation order N

where $k = 1, 2, ..., 2000$, and the sampling period is $\tau = 0.01$. The number set $E = \{5.0000,$ 8.3333, 11.6667, 15.0000, 18.3333, 21.6667, 25.00004, 28.3333, 31.6667, 35.0000$\}$ is a set of evenly spaced frequencies from the interval $[5, 35]$, while $R = \{6.8175,$ 11.8369, 12.4375, 13.3475, 14.0865, 15.1725, 19.6034, 20.8193, 21.8381, 33.3887$\}$ represents a set of frequencies selected randomly from a uniform distribution defined in the same interval $[5, 35]$. Fig.2.6 depicts that $u_r(k)$ exhibits higher richness than $u_e(k)$.

2.6 Numerical Example

Let's illustrate the proposed method through the analysis of a non-dimensional example. Consider a bilinear system of the form in Eq.(2.1), where

$$\mathbf{A} = \begin{bmatrix} 0.9924 & 0.0118 \\ -1.1778 & 0.8747 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.0126 \\ 0.0042 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 0.2000 & 0 \\ 0 & 0.2000 \end{bmatrix}, \qquad (2.30)$$
$$\mathbf{C} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \mathbf{D} = 1.2000$$

the number of inputs is $m = 1$; the number of outputs is $r = 1$, and the order of the system considered is equal to $n = 2$. Hence, the minimum choice for the OKID parameter p is 2, since $rp \geq n$. In addition, choosing $p = 2$ will allow us to keep the dimensions of the matrices small, in this numerical example.

Let's choose the sampling period τ to be equal to 0.01 seconds for a set of input/output measurement time-histories of length $20,000$ time-steps. The steady-state portion of the response is assumed to start at time-step $s = 748$, since $\prod_{k=1}^{j} (\mathbf{A} + \mathbf{N}u(k)) < 1e - 20$ for $s \leq j \leq 20,000$. In an actual identification task some a priori knowledge of the system is necessary for such an assumption to be valid. The set of frequencies Ω_0 and the corresponding amplitudes a_i, b_i used in constructing the base input excitation signal $u(k)$ in accordance with Eq.(2.2) are tabulated in Table 2.6.

The excitation input $\hat{u}(k)$ is obtained by adding a RC in the form of WGN to the previously defined $u(k)$ as described in Eq.(2.13). The amount of the added RC is chosen such that the ratio of the RMS of the added WGN to the RMS of $u(k)$ is 3%. Previously, it was discussed that the main motivation behind the addition of the RC was to eliminate the

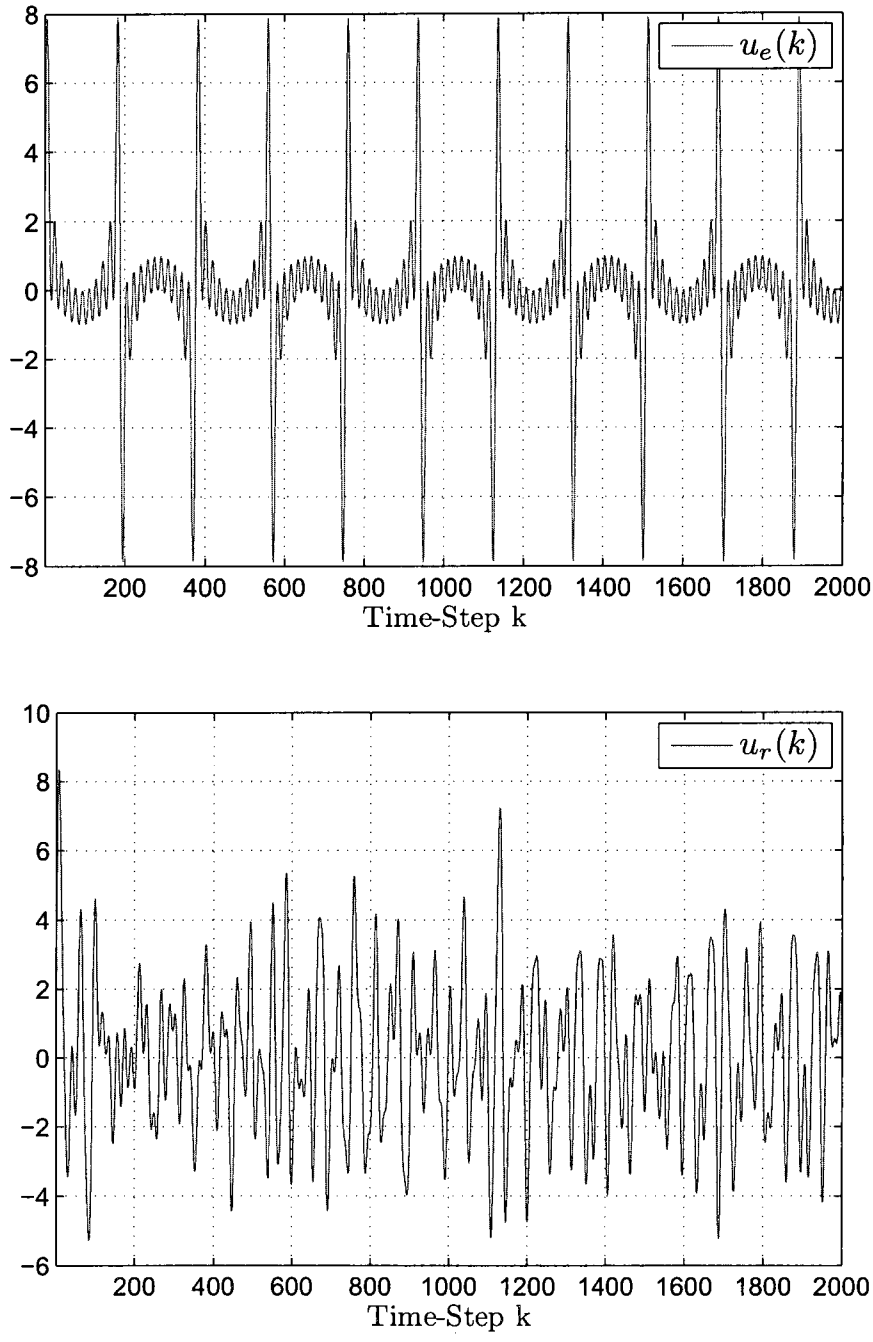


Figure 2.6: Signals constructed from a set of evenly spaced frequencies and randomly spaced frequencies

i	a_i	b_i	w_i(rad/second)
0	0.0000		0.0000
1	0.0936	0.0633	6.5624
2	0.0836	0.0828	7.3933
3	0.0992	0.0681	14.5528
4	0.0745	0.0890	15.9935
5	0.0774	0.0645	19.8812
6	0.0704	0.0525	20.7882
7	0.0786	0.0679	26.4501
8	0.0639	0.0918	28.7872
9	0.0637	0.0934	32.3738
10	0.0676	0.0891	32.7644

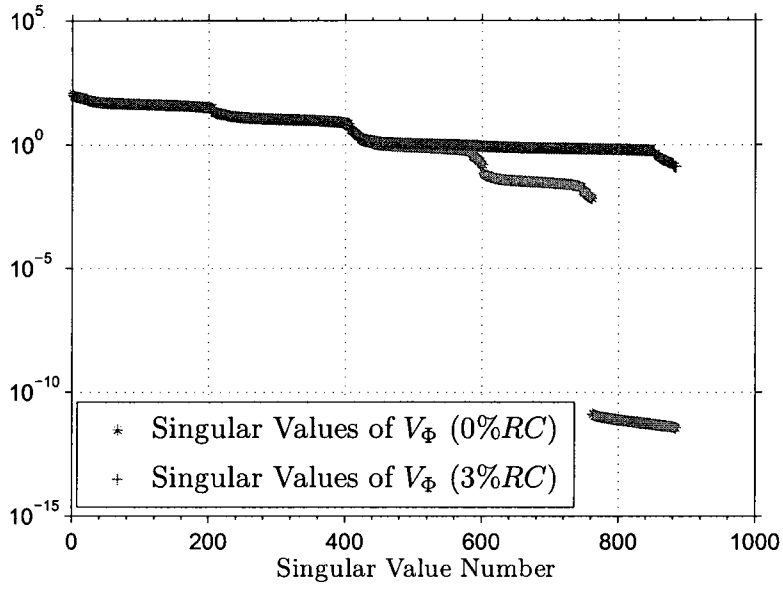
Table 2.6: The value of the parameters used in the construction of the base input signal $u(k)$ described in Eq.(2.2)

ill-conditioning in the matrix $\mathbf{V}$ shown in Eq.(2.11). Therefore, the choice of the amount of RC strongly depends on $\mathbf{V}$. However, in order to construct $\mathbf{V}$, the output $\mathbf{y}(k)$ needs to be known in advance. A more careful inspection of the structure of $\mathbf{V}$ hints a way out of this problem. The main source of the ill-conditioning is not due to the rows involving $\mathbf{y}(s+p-i)$ $(i=0,1,...,p)$, but rather due to the rows containing $\mathbf{\Phi}(s+p-i)u(s+p-i)$ $(i=-1,0,...,p)$ which may become linearly dependent due the identities in Eq.(2.6). In addition, the number of rows related to the output make up only a very small percentage of the total number of rows in $\mathbf{V}$ (e.g. for this example, even for the perturbation order $N=2$ and the OKID parameter $p=2$, they make up merely 3 of the 887 rows). Thus, a modified version of the matrix $\mathbf{V}$, where the rows related to the output are eliminated, can be used to predict a sufficient amount of RC rather accurately. While Fig.2.7a shows the plots of the singular values of such a matrix $\mathbf{V}_{\mathbf{\Phi}}$, defined in Eq.(2.31), before and after addition of the RC, Fig. 2.7b shows the same plots for the matrix $\mathbf{V}$ that is formed by OKID. Two results are immediately conceivable from these plots. First, prior to the addition of the RC

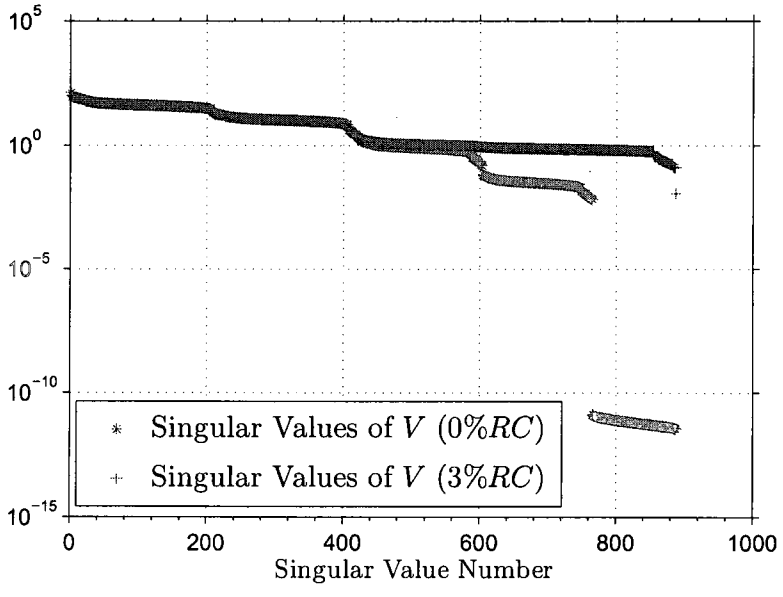
both $\mathbf{V}_{\mathbf{\Phi}}$ and $\mathbf{V}$ suffer from the same order of ill-conditioning, which can be quantitatively observed by comparing their condition numbers $3.2362e{+}013$ and $3.7096e{+}013$ respectively. Secondly, the addition of the 3% RC is sufficient for eliminating the ill-conditioning from both $\mathbf{V}_{\mathbf{\Phi}}$ and $\mathbf{V}$, as reflected by their condition numbers calculated after the enrichment of the input as $0.9223e+004$ and $1.1657e+004$ respectively.

$$\mathbf{V}_{\mathbf{\Phi}} = \begin{pmatrix} \mathbf{\Phi}(s+p+1)u(s+p+1) & \mathbf{\Phi}(s+p+2)u(s+p+2) & \cdots & \mathbf{\Phi}(l-1)u(l-1) \\ \mathbf{\Phi}(s+p)u(s+p) & \mathbf{\Phi}(s+p+1)u(s+p+1) & \cdots & \mathbf{\Phi}(l-2)u(l-2) \\ \mathbf{\Phi}(s+p-1)u(s+p-1) & \mathbf{\Phi}(s+p)u(s+p) & \cdots & \mathbf{\Phi}(l-3)u(l-3) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Phi}(s)u(s) & \mathbf{\Phi}(s+1)u(s+1) & \cdots & \mathbf{\Phi}(l-p-2)u(l-p-2) \end{pmatrix} \tag{2.31}$$

At this point, after being sure that the selected input excitation input wil not make $\mathbf{V}$ (or $\mathbf{V}_{\mathbf{\Phi}}$) ill-conditioned, the designed input excitation $\hat{u}(k)$, presented in Fig.2.8, is fed to the bilinear system, and the output $\mathbf{y}(k)$ is obtained. Using perturbation theory, the vector of basis functions $\mathbf{\Phi}(k)$ is formed for the perturbation order $N = 2$, and the input time-history for the ELM $\mathbf{\Phi}(k)\hat{u}(k)$ is generated numerically. This choice can be justified by the fact that the basis functions generated for $N = 2$ can represent the steady-state portion of the measured output time-history with an accuracy within the acceptable limit that allows OKID to identify the ELM robustly. Although the basis functions are used actually to represent the state, the output can still be used to roughly estimate the required number of the perturbation order, since it contains contributions from at least one of the states which is coupled with the rest of the states in the difference equation representing the bilinear model. The contribution coming to the output from the term $\mathbf{D}\hat{u}(k)$ does not violate this reasoning, as the set of frequencies present in $u(k)$ is a subset of those present in the state $\mathbf{x}(k)$ and the added RC is subtle. As a demonstration of the above discussion, Fig.2.9 depicts the steady-state portion of the time histories of the output $\mathbf{y}(k)$ and its representation $\hat{\mathbf{y}}(k)$ in terms of the basis functions $\mathbf{\Phi}(k)$. Moreover, the amount of mismatch between $\mathbf{y}(k)$ and $\hat{\mathbf{y}}(k)$ can be quantified using the error function defined in Eq.(2.25) as $error(\mathbf{y}(k), \hat{\mathbf{y}}(k)) = 4.8775\%$. The validity of the discussion in terms of the output $\mathbf{y}(k)$ can be better observed from Fig.2.10, where the plots of the steady-state portions of time-histories of the second state $\mathbf{x}(k)$ and its



(a)



(b)

Figure 2.7: The plots of the singular values of the matrix $\mathbf{V}_\Phi$ and $\mathbf{V}$ formed both before and after the addition of RC

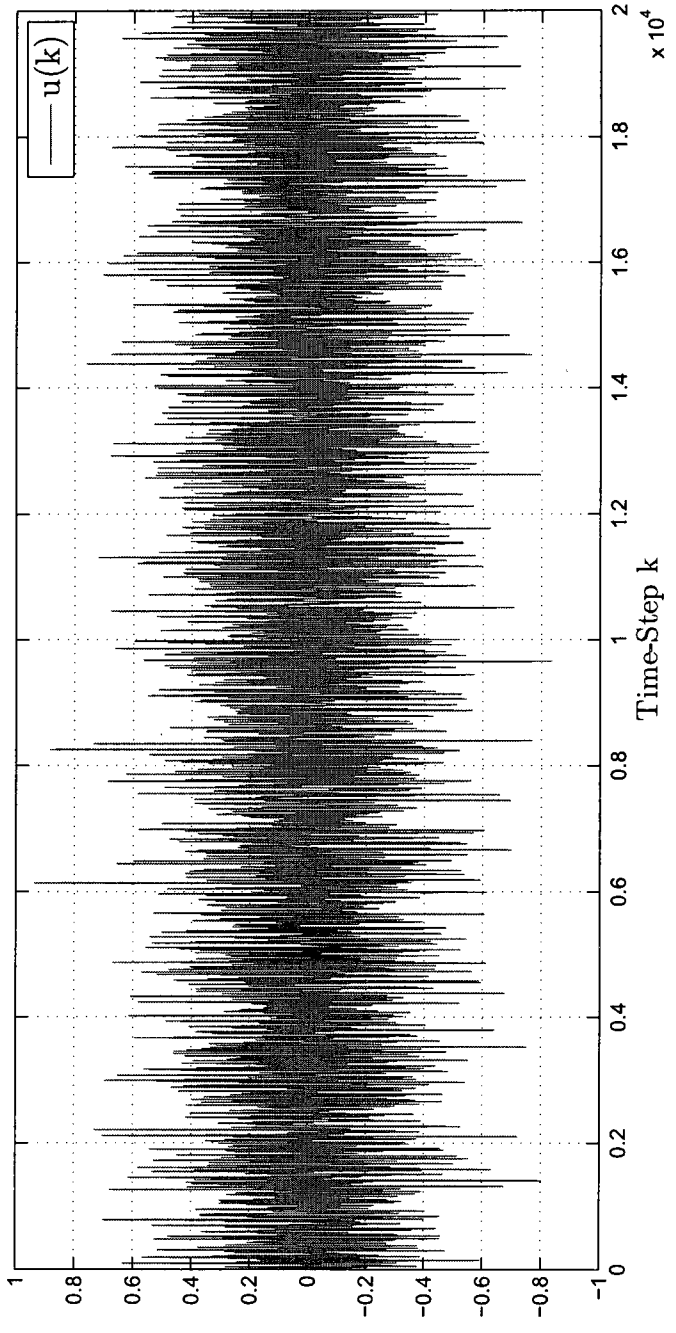


Figure 2.8: The performance of least-square fitting for different data lengths

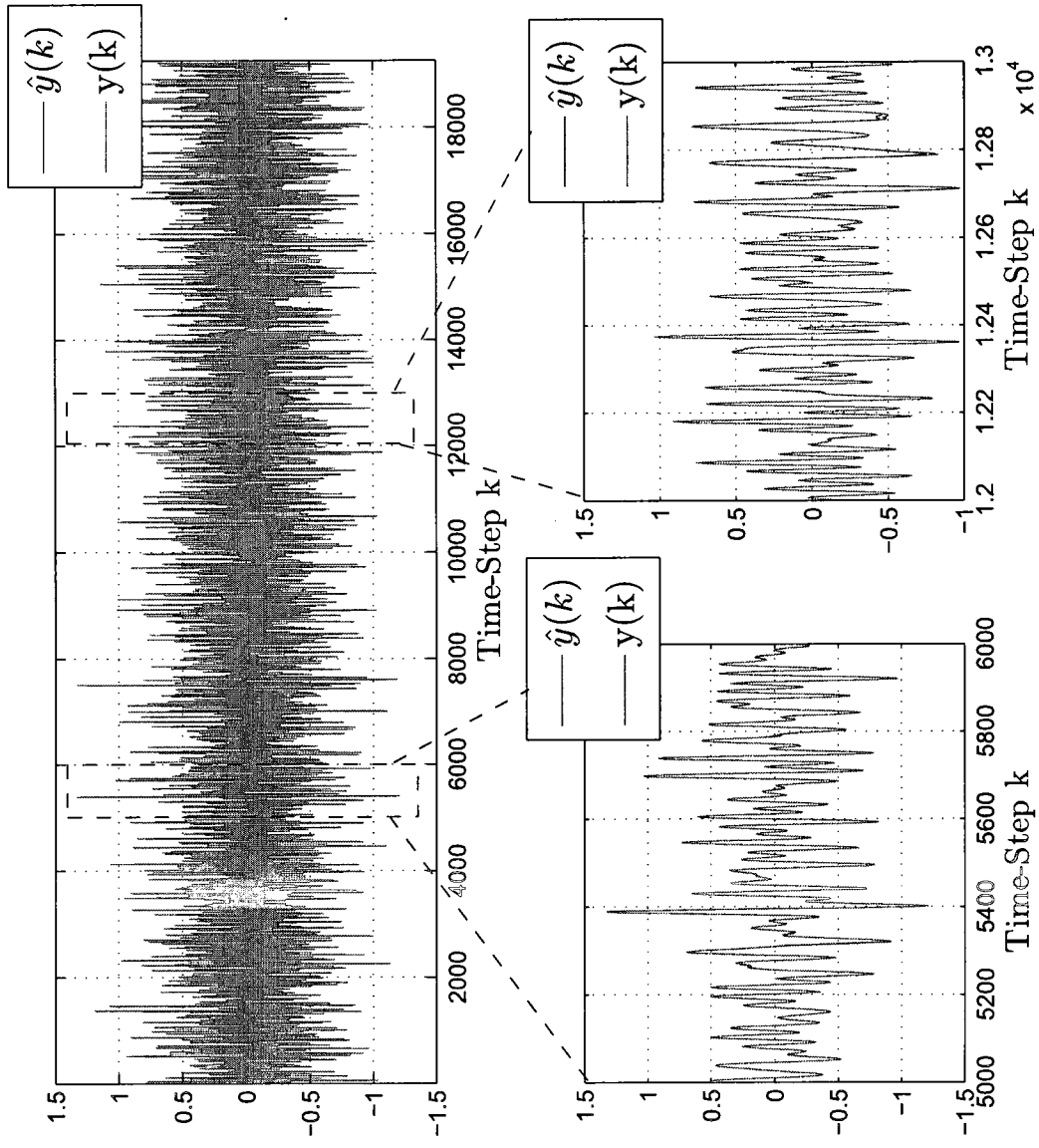


Figure 2.9: The measured output time-history $\mathbf{y}(k)$ and its representation $\hat{\mathbf{y}}(k)$ in terms basis functions $\phi(k)$

representation $\hat{\mathbf{x}}(k)$ in terms of the basis functions are provided. The amount of mismatch between them is calculated using $error(\mathbf{x}(k), \hat{\mathbf{x}}(k))$ to be 6.6747% and 8.6145% for the first and second states respectively, which are reasonably close to that for the output. The steady-state portion of the time-histories of the numerically generated input for the ELM, $\mathbf{\Phi}(k)u(k)$, and the measured output $\mathbf{y}(k)$ is fed to OKID to acquire the system matrices of the corresponding ELM shown in Eq.(2.14). The time-history of the state $\mathbf{x}_{ok}(k)$ and output $\mathbf{y}_{ok}(k)$ of the bilinear system in the coordinates of the identified ELM can be generated by simulating the identified ELM using $\mathbf{\Phi}(k)u(k)$. Fig.2.11 depicts the steady-state portion of the measured output time-history against its reproduction by the ELM. The approximation error between the two is calculated using $error(\mathbf{y}(k), \mathbf{y}_{ok}(k))$ to be 1.9850%. As shown both visually and numerically, the reproduced output time-history is virtually identical to the measured output time-history. Then, the matrix of the coefficients $\tilde{\mathbf{\Lambda}}$ is solved using Eq.(2.20), and the identification of the bilinear system is concluded by retrieving the system matrices using Eq.(2.19) as

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0.9673 & -0.0784 \\ 0.1473 & 0.9008 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.3370 \\ -0.8113 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 0.1993 & -0.0095 \\ -0.0061 & 0.2039 \end{bmatrix}, \\ \mathbf{C} &= \begin{bmatrix} -0.1013 & -0.0776 \end{bmatrix}, \mathbf{D} = 1.2000 \end{aligned} \tag{2.32}$$

As a measure of the accuracy of the identification, a new random input of length 20,000 time-steps is generated from a Gaussian distribution with mean 0 and standard deviation 1. The generated test input $u_t(k)$ is fed both to the original and the identified bilinear systems and the corresponding output time histories, $\mathbf{y}_t(k)$ and $\mathbf{y}_p(k)$ respectively, are obtained. Fig.2.12 shows the plot of the predicted output time-history $\mathbf{y}_p(k)$ against the actual output time-history $\mathbf{y}_t(k)$ with a prediction error of 0.9437% between the two. This clearly shows that the predicted output $\mathbf{y}_p(k)$ is a very accurate estimate of the output $\mathbf{y}_t(k)$ of the bilinear system. Furthermore, the original bilinear system matrices $\mathbf{A}$ and $\mathbf{N}$ are related to the identified bilinear system matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{N}}$ respectively, through a similarity transform as shown in Eq.(2.15) and (2.16). Therefore the eigenvalues of $\mathbf{A}$, $\tilde{\mathbf{A}}$ and $\mathbf{N}$, $\tilde{\mathbf{N}}$ need to be the same. For the example above, these eigenvalues are found to be $\{0.9336 \pm 0.1022i\}$, $\{0.9340 \pm 0.1022i\}$ and $\{0.2000, 0.2000\}$, $\{0.1937, 0.2096\}$, respectively. Being invariant against coordinate transformations, the direct transmission matrix associated with

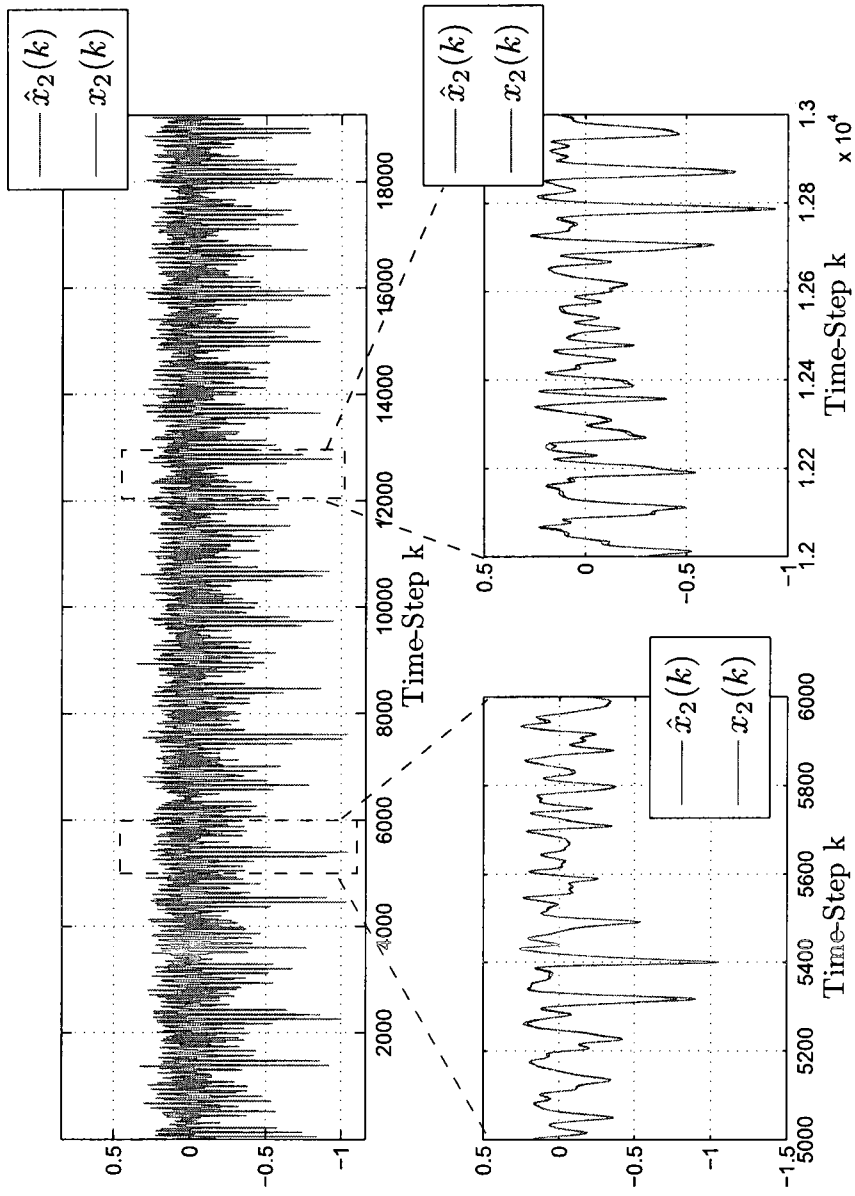


Figure 2.10: The time-history of the second state of the original bilinear system $\mathbf{x}_2(k)$ and its representation $\hat{\mathbf{x}}_2(k)$ in terms of the basis functions $\phi(k)$

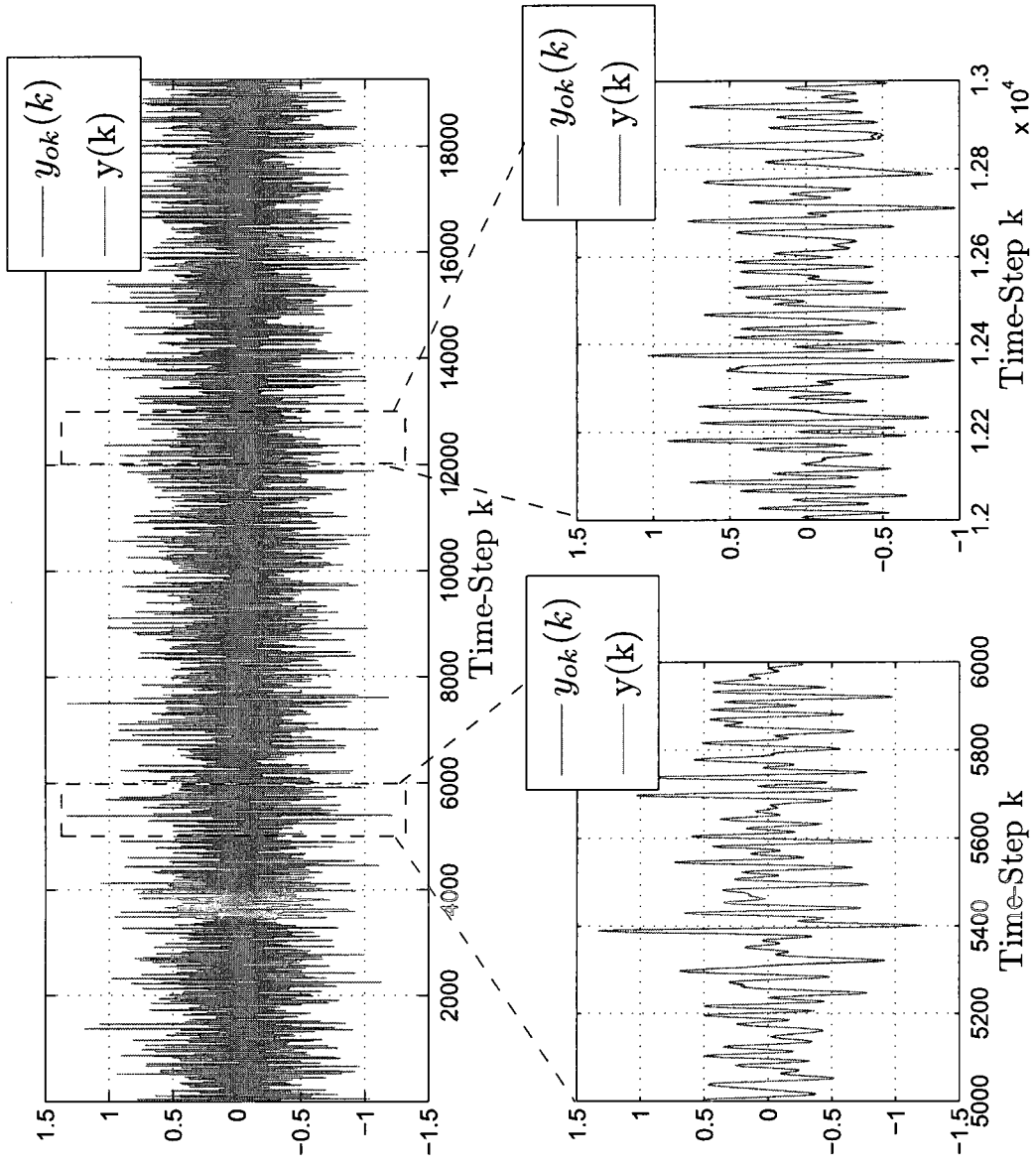


Figure 2.11: The measured output time-history $\mathbf{y}(k)$ and its reproduction by the ELM $\mathbf{y}_{ok}(k)$ in terms basis functions $\phi(k)$

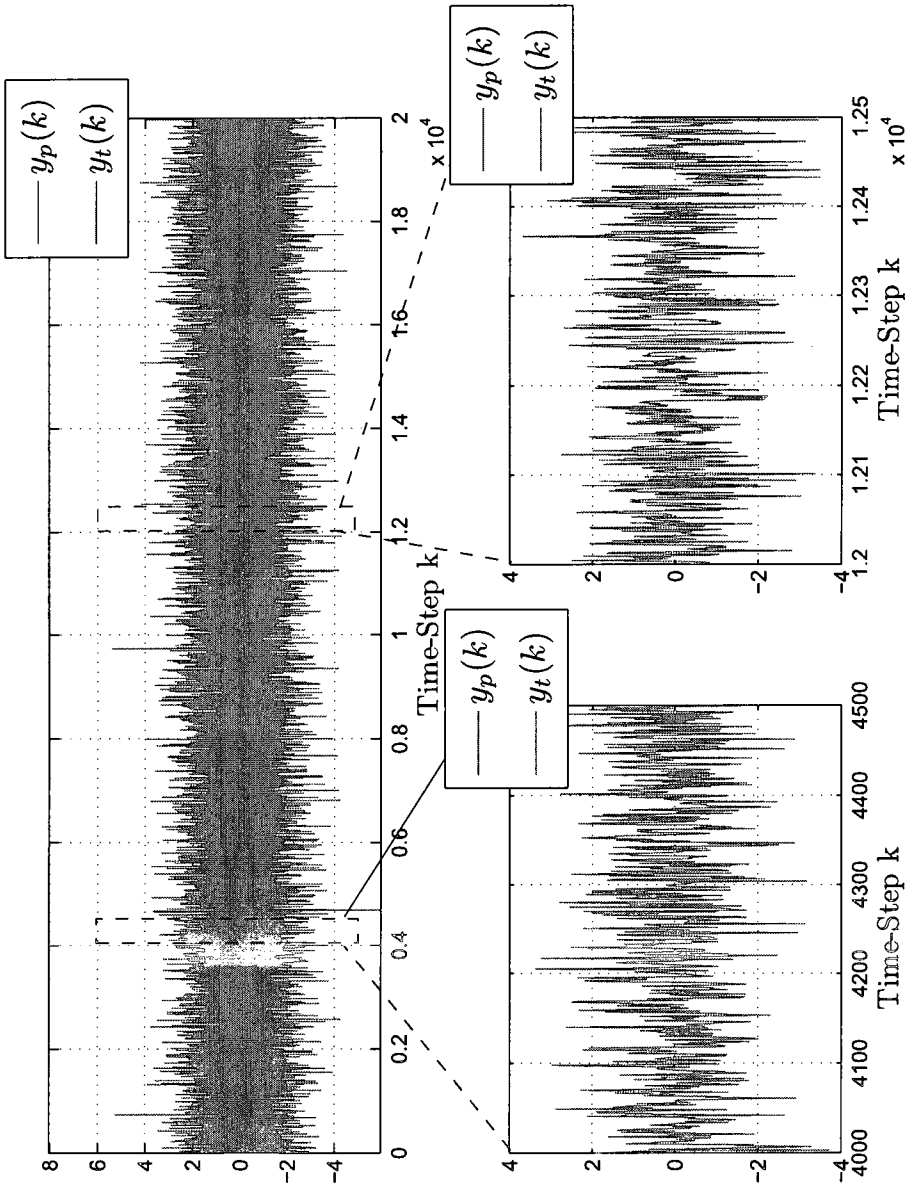


Figure 2.12: The predicted output time-history $\mathbf{y}_p(k)$ and the simulated test output time-history $\mathbf{y}_{\mathbf{\Phi}}(k)$

the original system $\mathbf{D} = 1.2000$ and identified bilinear system $\mathbf{D}_{ok} = 1.2000$ should be the same, when there is no identification error. The "almost" perfect match between the eigenvalues of the identified and original system matrices and the exact estimation of the direct transmission term confirms the success of the identification.

2.7 Conclusions

As presented in the above example and the preceding explanation, the identification of a discrete-time, time-invariant, state-space bilinear model can be converted to an ELM that can reproduce the same output-time history as the original bilinear system by expressing the steady-state portion of the state in the bilinear term as a linear combination of sine and cosine functions. The frequencies of the basis functions are selected in accordance with perturbation theory. While the ELM is valid only for the specific set of input/output time-histories, it can be used to extract the system matrices of the bilinear model. The accuracy of the identified bilinear model is verified by the fact that it can accurately predict the response of the original bilinear system subject to a new, randomly generated test input. The aforementioned method differs significantly from the existing methods in the literature by allowing for the design of the excitation input and by requiring only a single set of input/output measurement time-histories. Furthermore, the effect of the parameters selected by the user during the implementation of the method were discussed. It was outlined that as the data length is increased, the state approximation error increases and the ill-conditioning of the matrix $\mathbf{V}$ reduces. The amount of random component should be high enough eliminate the ill-conditioning in $\mathbf{V}$, but small enough to justify the use of perturbation theory. The OKID parameter p should be kept as small as possible to curtail ill-conditioning, although a higher p augments the robustness of OKID against measurement and process noise. The choice of a randomly selected excitation frequencies were shown to depict higher richness in comparison to evenly spaced input frequencies.

Chapter 3

Identification of Bilinear Systems Through Equivalent Linear Models via Data Driven Basis Functions

This chapter presents the second proposed method for the identification of single-input, multi-output, time-invariant, discrete-time bilinear systems using the steady-state portion of a single input/output time-history obtained by exciting the system with a linear combination of sine and cosine functions of user-selected frequencies enriched by a subtle amount of highband random component (HRC). The method relies on the conversion of the bilinear model into an equivalent linear model by an accurate approximation of the state in the bilinear term using a set of sine and cosine basis functions whose frequencies are obtained from perturbation theory. The columns of the ELM matrices which correspond to the dependent rows of the input vector are compressed into single columns for obtaining an input vector with linearly independent rows. Then, the states of the ELM are replaced by the past input/output time-history measurements to create an overparameterized model with known states. The system matrices of this model are obtained by a least-squares fit and its unobservable states are removed to obtain a minimal realization of the ELM. Finally, the original bilinear system is retrieved using the steady-state portion of the state time-history generated by simulating the identified ELM with the input used in its identification.

3.1 Introduction

The proposed method addresses the identification of the bilinear system described in Eq.(2.1). As in the former method, the bilinear system is excited by a linear combination of sine and cosine functions of user-selected frequencies. Then, the bilinear model can be converted to an equivalent linear model by approximating the state $\mathbf{x}(k)$ in the bilinear term $\mathbf{x}(k)u(k)$ in Eq.(2.1) using a linear combination of sine and cosine basis functions whose frequencies are obtained by perturbation theory. The details regarding this procedure can be found in Section 2.2. In order to obtain a unique ELM, the columns of the system matrices of the ELM that correspond to the linearly dependent inputs are compressed into single columns as elaborated in Section 2.3. The resulting ELM, which is only valid for the specific input time-history, is shown in Eq.(3.1)

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B_e}\mathbf{\Phi}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}u(k), \quad k = s, s+1, ..., l \end{aligned} \tag{3.1}$$

where $\mathbf{B_e} = \begin{bmatrix} \mathbf{B} + \mathbf{N}\mathbf{\Lambda}_1 & \mathbf{N}\mathbf{\Lambda}_2 \end{bmatrix}$, while all the other symbols maintain their predefined meanings in Sections 2.1 and 2.2.

3.2 A "Pick Your Own States" Approach

The approximation of the state by a linear combination of sine and cosine basis functions, as shown in Eq.(2.7), can also be performed for the remainder of the terms in Eq.(3.1) as follows:

$$\begin{aligned} \mathbf{\Phi}(k+1) &= \mathbf{A}_{\mathbf{\Phi}}\mathbf{\Phi}(k) + (\mathbf{B_e})_{\mathbf{\Phi}}\mathbf{\Phi}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C}_{\mathbf{\Phi}}\mathbf{\Phi}(k) + \mathbf{D}u(k), \quad k = s, s+1, ..., l \end{aligned} \tag{3.2}$$

where

$$\begin{aligned} \mathbf{A}_{\mathbf{\Phi}} &= \mathbf{\Lambda}^{+}\mathbf{A}\mathbf{\Lambda} \\ (\mathbf{B_e})_{\mathbf{\Phi}} &= \mathbf{\Lambda}^{+}\mathbf{B_e} \\ \mathbf{C}_{\mathbf{\Phi}} &= \mathbf{C}\mathbf{\Lambda} \end{aligned} \tag{3.3}$$

This approximation implies that the vector of basis functions $\mathbf{\Phi}(k)$ can be used as the state vector of the ELM. Such a choice reduces the bilinear system to an ELM with a known state time-history. Thus, unlike in the previous method, where the user determines only the frequency content of the state by choosing the excitation input frequencies and the order of the perturbation theory approximation, in this method the user also picks the time-history that the state of the ELM follows.

Eq.(3.2) can be rewritten as follows:

$$\begin{aligned} \mathbf{\Phi}(k+1) &= \begin{bmatrix} \mathbf{A}_{\mathbf{\Phi}} & (\mathbf{B_e})_{\mathbf{\Phi}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}(k) \\ \mathbf{\Phi}(k)u(k) \end{bmatrix} \\ \mathbf{y}(k) &= \begin{bmatrix} \mathbf{C}_{\mathbf{\Phi}} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}(k) \\ u(k) \end{bmatrix}, \quad k = s, s+1, ..., l \end{aligned} \tag{3.4}$$

The recursive relationship in Eq.(3.4) can be cast into matrix form as follows:

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} \mathbf{A}_{\mathbf{\Phi}} & (\mathbf{B_e})_{\mathbf{\Phi}} \end{bmatrix} \mathbf{Q} \\ \mathbf{Y} &= \begin{bmatrix} \mathbf{C}_{\mathbf{\Phi}} & \mathbf{D} \end{bmatrix} \mathbf{L} \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} \mathbf{\Phi}(s+1) & \mathbf{\Phi}(s+2) & \cdots & \mathbf{\Phi}(l) \end{bmatrix} \\ \mathbf{Q} &= \begin{bmatrix} \mathbf{Q_1} \\ \mathbf{Q_2} \end{bmatrix} \\ \mathbf{L} &= \begin{bmatrix} \mathbf{Q_1} \\ \mathbf{L_2} \end{bmatrix} \\ \mathbf{Y} &= \begin{bmatrix} \mathbf{y}(s) & \mathbf{y}(s+1) & \cdots & \mathbf{y}(l-1) \end{bmatrix} \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} \mathbf{Q_1} &= \begin{bmatrix} \mathbf{\Phi}(s) & \mathbf{\Phi}(s+1) & \cdots & \mathbf{\Phi}(l-1) \end{bmatrix} \\ \mathbf{Q_2} &= \begin{bmatrix} \mathbf{\Phi}(s)u(s) & \mathbf{\Phi}(s+1)u(s+1) & \cdots & \mathbf{\Phi}(l-1)u(l-1) \end{bmatrix} \\ \mathbf{L_2} &= \begin{bmatrix} u(s) & u(s+1) & \cdots & u(l-1) \end{bmatrix} \end{aligned} \tag{3.7}$$

Thus, the system matrices of the ELM $\mathbf{A_\Phi}$ and $(\mathbf{B_e})_\Phi$ can be directly solved through the least-square approach as follows:

$$\begin{aligned} \begin{bmatrix} \mathbf{A_\Phi} & (\mathbf{B_e})_\Phi \end{bmatrix} &= \mathbf{M}\mathbf{Q}^+ \\ \begin{bmatrix} \mathbf{C_\Phi} & \mathbf{D} \end{bmatrix} &= \mathbf{Y}\mathbf{L}^+ \end{aligned} \tag{3.8}$$

where $^+$ indicates the Moore-Penrose pseudoinverse. For a unique identification of the ELM, it is required that the matrices $\mathbf{Q}$ and $\mathbf{L}$ have unique pseudoinverses. This condition will be satisfied by first replacing the sine and cosine basis functions with an equivalent set of orthogonal basis functions to ensure that the upper submatrix $\mathbf{Q_1}$, common to both $\mathbf{Q}$ and $\mathbf{L}$, is full-rank. Secondly, the linearly dependent rows of the lower submatrices $\mathbf{Q_2}$ and $\mathbf{L_2}$ are discarded through the process of input reduction. Finally, through the process of input absorption, those rows of the submatrices $\mathbf{Q_2}$ and $\mathbf{L_2}$, which are linearly dependent on the remaining rows of $\mathbf{Q_1}$ after the application of input reduction process, are discarded. Consequently, the resulting $\mathbf{Q}$ and $\mathbf{L}$ matrices are full-rank and, therefore have unique pseudoinverses.

3.2.1 Orthogonal Basis Functions

Even though $\Phi(k)$ is comprised of sine and cosine basis functions of distinct frequencies, the submatrix $\mathbf{Q_1}$ is full-rank only if the data length l is long enough. In Section 2.5.1, the relationship between the data length and the ability of a set of sine and cosine basis functions to mimic a sinusoidal function not contained by the set of basis functions was demonstrated. As a corollary of that discussion, for a short data length, some of the rows of $\mathbf{Q_1}$ are expected to be expressible in terms of its remaining rows. In such a case, the submatrix $\mathbf{Q_1}$ and, thus, the matrix $\mathbf{Q}$ become rank deficient. This problem can be remedied either by using a longer data length or by replacing the sine and cosine basis functions with a set of orthogonal basis functions by performing a singular value decomposition on a matrix whose columns form the steady-state portion of the time-history of the basis function vector

$\mathbf{\Phi}(k)$ as follows:

$$\begin{aligned}\left[\mathbf{\Phi}(s) \quad \mathbf{\Phi}(s+1) \quad \cdots \quad \mathbf{\Phi}(l)\right] &= \mathbf{U_\Phi \Sigma_\Phi V_\Phi}^T \\ &= \left[(\mathbf{U_\Phi})_1 \quad (\mathbf{U_\Phi})_2\right] \begin{bmatrix} \mathbf{S_\Phi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \left[(\mathbf{V_\Phi})_1 \quad (\mathbf{V_\Phi})_2\right]^T \end{aligned} \tag{3.9}$$

where ^{T} denotes transpose, $\mathbf{\Sigma_\Phi}$ indicates the matrix of the singular values, $\mathbf{U_\Phi}$ and $\mathbf{V_\Phi}$ are unitary matrices, while $\mathbf{S_\Phi}$ represents the matrix of non-zero singular values. Consequently, $(\mathbf{U_\Phi})_1$ and $(\mathbf{V_\Phi})_1$ are the matrices formed by those columns of the matrices $\mathbf{U_\Phi}$ and $\mathbf{V_\Phi}$ that correspond to the non-zero singular values contained by $\mathbf{S_\Phi}$, while $(\mathbf{U_\Phi})_2$ and $(\mathbf{V_\Phi})_2$ indicate the matrices formed by the remaining columns of $\mathbf{U_\Phi}$ and $\mathbf{V_\Phi}$, respectively. The submatrix ${(\mathbf{V_\Phi})_1}^T$ has the same rowspace as the original matrix it was obtained from; in addition, its rows are orthogonal. Thus, its rows can be used as the steady-state portion of the time-history of an alternative basis function vector. Such a choice for the vector of basis functions ensures that the submatrix $\mathbf{Q_1}$ is full-rank.

3.2.2 Input Reduction

In accordance with the trigonometric identities in Eq.(2.6), if the product of $\mathbf{\Phi}(k)$ and $u(k)$ produces sine and cosine functions of identical frequencies, the submatrix $\mathbf{Q_2}$ may become numerically ill-conditioned. Similarly as before, a singular value decomposition (SVD) can be performed on a matrix whose columns form the steady-state portion of the time-history of the input of the ELM, $\mathbf{\Phi}(k)u(k)$, in order to obtain a set of orthogonal inputs, as follows:

$$\begin{aligned}\left[\mathbf{\Phi}(s)u(s) \quad \mathbf{\Phi}(s+1)u(s+1) \quad \cdots \quad \mathbf{\Phi}(l)u(l)\right] &= \bar{\mathbf{U}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^T \\ &= \left[\bar{\mathbf{U}}_1 \quad \bar{\mathbf{U}}_2\right] \begin{bmatrix} \bar{\mathbf{S}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \left[\bar{\mathbf{V}}_1 \quad \bar{\mathbf{V}}_2\right]^T \end{aligned} \tag{3.10}$$

where $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ are unitary matrices that multiply the matrix of the singular values $\bar{\mathbf{\Sigma}}$ from the left and right, while $\bar{\mathbf{S}}$ denotes the matrix of non-zero singular values. The submatrices $\bar{\mathbf{U}}_1$ and $\bar{\mathbf{V}}_1$ are formed by the columns of $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ that correspond to the non-zero singular

values contained by the matrix $\bar{\mathbf{S}}$, while $\bar{\mathbf{U}}_\mathbf{2}$ and $\bar{\mathbf{V}}_\mathbf{2}$ represent the submatrices formed by the remaining columns of $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$, respectively.

The matrix $\bar{\mathbf{V}}_\mathbf{1}^T$ has the same rowspace as the original matrix it is obtained from, but it is comprised of orthogonal rows. Thus, it can be used as the input for the ELM as follows:

$$\begin{aligned} \mathbf{\Phi}(k+1) &= \mathbf{A}_\mathbf{\Phi}\mathbf{\Phi}(k) + (\bar{\mathbf{B}}_\mathbf{e})_\mathbf{\Phi}\nu(k) \\ \mathbf{y}(k) &= \mathbf{C}_\mathbf{\Phi}\mathbf{\Phi}(k) + \mathbf{D}u(k) \quad k = s, s+1, ..., l \end{aligned} \tag{3.11}$$

where $(\bar{\mathbf{B}}_\mathbf{e})_\mathbf{\Phi} = \mathbf{B}_\mathbf{\Phi}(\bar{\mathbf{U}}_\mathbf{1}\bar{\mathbf{S}})$, and $\nu(k)$ is the $(k-s+1)st$ column of the matrix $\bar{\mathbf{V}}_\mathbf{1}$. After the application of the aforementioned process of input reduction, the submatrix $\mathbf{Q_2}$ corresponding to the ELM in Eq.(3.11) becomes:

$$\mathbf{Q_2} = \begin{bmatrix} \nu(s) & \nu(s+1) & \cdots & \nu(l-1) \end{bmatrix}, \tag{3.12}$$

and it is now full-rank.

3.2.3 Input Absorption

Even if the submatrices $\mathbf{Q_1}$ and $\mathbf{Q_2}$ are full rank separately, some of the rows in $\mathbf{Q_1}$ may be linearly dependent on some of the rows in $\mathbf{Q_2}$, causing the matrix $\mathbf{Q}$ to be rank deficient. This problem can be avoided either by addition of a subtle amount of random component to the excitation input $u(k)$ or through the process of input absorption. The addition of a RC uncorrelates the rows of the matrix $\mathbf{Q_2}$ at the expense of increasing the approximation error in Eq.(2.7) as discussed earlier. On the other hand, the process of input absorption suggests that those rows of $\mathbf{Q_2}$, which are linearly dependent on the rows of $\mathbf{Q_1}$, can simply be discarded and, therefore, the resulting $\mathbf{Q}$ matrix is full-rank without the need of adding a RC. The discarding of the linearly dependent rows is valid because it implies that $\mathbf{Q_1}$ absorbs the contribution of the eliminated rows of $\mathbf{Q_2}$.

Consider a 2-state system described by Eq.(3.13)

$$\begin{aligned} \begin{bmatrix} \phi_1(k+1) \\ \phi_2(k+1) \end{bmatrix} &= \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{bmatrix} \phi_1(k) \\ \phi_2(k) \end{bmatrix} + \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} \\ y(k) &= \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{bmatrix} \phi_1(k) \\ \phi_2(k) \end{bmatrix} + du_1(k) \end{aligned} \tag{3.13}$$

and assume that

$$u_1(k) = \alpha\phi_1(k) + \beta\phi_2(k) + \gamma u_2(k) \tag{3.14}$$

Plugging Eq.(3.14) into Eq.(3.13) produces the following set of equations in the state space form:

$$\begin{aligned} \begin{bmatrix} \phi_1(k+1) \\ \phi_2(k+1) \end{bmatrix} &= \begin{pmatrix} a_1 + b_1\alpha & a_2 + b_1\beta \\ a_3 + b_3\alpha & a_4 + b_3\beta \end{pmatrix} \begin{bmatrix} \phi_1(k) \\ \phi_2(k) \end{bmatrix} + \begin{pmatrix} b_2 + b_1\gamma \\ b_4 + b_3\gamma \end{pmatrix} u_2(k) \\ y(k) &= \begin{pmatrix} c_1 + d\alpha & c_2 + d\beta \end{pmatrix} \begin{bmatrix} \phi_1(k) \\ \phi_2(k) \end{bmatrix} + d\gamma u_2(k) \end{aligned} \tag{3.15}$$

As it can be seen from this example, the first input $u_1(k)$ is absorbed by the basis functions and the second input $u_2(k)$. Thus, the process of input absorption has been used to eliminate those input functions that are linearly dependent on the basis functions, thereby yielding an ELM, where the basis functions and the remaining inputs form a set of linearly independent functions.

3.2.4 Rotation Equation

Let's assume that the following ELM is obtained by the application of the input reduction and absorption steps to the ELM in Eq.(3.2):

$$\mathbf{\Phi}(k+1) = \tilde{\mathbf{A}}_{\mathbf{\Phi}}\mathbf{\Phi}(k) + (\tilde{\mathbf{B}}_{\mathbf{e}})_{\mathbf{\Phi}}\mu(k) \tag{3.16}$$

where $\tilde{\mathbf{A}}_{\mathbf{\Phi}}$, $(\tilde{\mathbf{B}}_{\mathbf{e}})_{\mathbf{\Phi}}$ are, respectively, the final forms of the system matrices $\mathbf{A}$ and $\mathbf{B}$ of the ELM in Eq.(3.2), while $\mu(k)$ is the final form of the input $u(k)$, after the application of the input reduction and absorption processes. One may now expect to identify the ELM in Eq.(3.16) uniquely using the following equation:

$$\begin{bmatrix} \tilde{\mathbf{A}}_{\mathbf{\Phi}} & (\tilde{\mathbf{B}}_{\mathbf{e}})_{\mathbf{\Phi}} \end{bmatrix} = \tilde{\mathbf{M}}\tilde{\mathbf{Q}}^{+} \tag{3.17}$$

where

$$\begin{aligned} \tilde{\mathbf{M}} &= \begin{bmatrix} \mu(s+1) & \mu(s+2) & \cdots & \mu(l) \end{bmatrix} \\ \tilde{\mathbf{Q}} &= \begin{bmatrix} \mathbf{\Phi}(s) & \mathbf{\Phi}(s+1) & \cdots & \mathbf{\Phi}(l-1) \\ \mu(s) & \mu(s+1) & \cdots & \mu(l-1) \end{bmatrix} \end{aligned} \tag{3.18}$$

since the steps of input reduction and input absorption ensure that the matrix $\tilde{\mathbf{Q}}$ has a unique pseudoinverse. However, such an attempt would still fail, as it yields a solution in which the identified $(\tilde{\mathbf{B}}_\mathbf{e})_\mathbf{\Phi} = \mathbf{0}$ and $\tilde{\mathbf{A}}_\mathbf{\Phi} = \mathbf{R}$, where

$$\mathbf{\Phi}(k+1) = \mathbf{R}\mathbf{\Phi}(k) \tag{3.19}$$

Let's call the matrix $\mathbf{R} \in \mathbb{R}^{(2q+1)\times(2q+1)}$ as the *rotation matrix*, while Eq.(3.19) as the *rotation equation*. The rotation equation implies that $\mathbf{\Phi}(k+1)$ can be written as a linear combination of the rows of $\mathbf{\Phi}(k)$. This relationship stems from the trigonometric identities in Eq.(2.6). Consider the following example as a demonstration of the rotation equation, where

$$\mathbf{\Phi}(k) = \begin{pmatrix} \sin(3k\tau) \\ \sin(5k\tau) \\ \cos(3k\tau) \\ \cos(5k\tau) \end{pmatrix} \tag{3.20}$$

and τ is the sampling period. Then, in accordance with Eq.(2.6), the matrix $\mathbf{R}$, which satisfies Eq.(3.19), becomes

$$\mathbf{R} = \begin{bmatrix} \cos(3\tau) & 0 & \sin(3\tau) & 0 \\ 0 & \cos(5\tau) & 0 & \sin(5\tau) \\ -\sin(3\tau) & 0 & \cos(3\tau) & 0 \\ 0 & -\sin(5\tau) & 0 & \cos(5\tau) \end{bmatrix} \tag{3.21}$$

Hence, Eq.(3.17) cannot be used to identify the ELM in Eq.(3.16), since it yields a trivial solution where $(\tilde{\mathbf{B}}_\mathbf{e})_\mathbf{\Phi} = \mathbf{0}$ and $\tilde{\mathbf{A}}_\mathbf{\Phi} = \mathbf{R}$.

3.3 The Proposed Method

3.3.1 Data Driven Basis Functions

The choice of the vector of sine and cosine basis functions $\mathbf{\Phi}(k)$ as the state is a particularly convenient one for the reasons discussed above. However, Eq.(3.19) does not allow the identification of the ELM as it leads to a trivial solution, where $(\tilde{\mathbf{B}}_\mathbf{e})_\mathbf{\Phi} = 0$ which implies

falsely that the input has no effect on the state. On the other hand, this choice of basis basis functions is not unique, and another alternative, which does not lead to an equation equivalent to the rotation equation, is to use the past input/output time-histories as the states. The validity of this choice can be shown conveniently using the interaction matrices.

Consider the LTI system in Eq.(1.34), which underwent the algebraic manipulation described in Eq.(1.42). As discussed previously, such a manipulation allows for the placement of the poles of the matrix $\bar{\mathbf{A}}$ at the origin such that $\bar{\mathbf{A}}^p \approx 0$, where $p \geq nr$, on condition that the system is controllable and observable. The solution to the state in Eq.(1.42) can be written as

$$\mathbf{x}(k+p) = \bar{\mathbf{A}}^p\mathbf{x}(k) + \bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}}\mathbf{v}(k) + \cdots + \bar{\mathbf{A}}\bar{\mathbf{B}}\mathbf{v}(k+p-2) + \bar{\mathbf{B}}\mathbf{v}(k+p-1) \tag{3.22}$$

where $k \geq 0$, $\mathbf{x}(k)$ represents the state vector and $\mathbf{v}(k)$ denotes the vector of input and output such that

$$\mathbf{v}(k) = \begin{bmatrix} \mathbf{u}(k) \\ \mathbf{y}(k) \end{bmatrix} \tag{3.23}$$

Recalling $\bar{\mathbf{A}}^p \approx 0$, it is clear from Eq.(3.22) that the state can be expressed as a linear combination of the past inputs $\{\mathbf{u}(k), \mathbf{u}(k+1), \cdots, \mathbf{u}(k+p-1)\}$ and past outputs $\{\mathbf{y}(k), \mathbf{y}(k+1), \cdots, \mathbf{y}(k+p-1)\}$.

The same reasoning can be extended to the ELM in Eq.(3.1) such that the state $\mathbf{x}(k)$ can be written for $s+p \leq k \leq l$ as

$$\mathbf{x}(k) = \mathbf{T}_\mathbf{z}\mathbf{z}(k) \tag{3.24}$$

where

$$\mathbf{z}(k) = \begin{bmatrix} \mathbf{y}(k-1) \\ \mathbf{y}(k-2) \\ \vdots \\ \mathbf{y}(k-p) \\ \mathbf{\Phi}(k-1)u(k-1) \\ \mathbf{\Phi}(k-2)u(k-2) \\ \vdots \\ \mathbf{\Phi}(k-p)u(k-p) \end{bmatrix} \tag{3.25}$$

and $\mathbf{T_z} \in \mathbb{R}^{2\times(pr+2pq)}$ is a transformation matrix. Substituting Eq.(3.24) into Eq.(3.1) yields the following ELM

$$\begin{aligned} \mathbf{z}(k+1) &= \mathbf{A_z}\mathbf{z}(k) + (\mathbf{B_e})_\mathbf{z}\mathbf{\Phi}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C_z}\mathbf{z}(k) + \mathbf{D}u(k) \end{aligned} \tag{3.26}$$

where

$$\begin{aligned} \mathbf{A_z} &= \mathbf{T_z}^{+}\mathbf{A}\mathbf{T_z} \\ (\mathbf{B_e})_\mathbf{z} &= \mathbf{B}\mathbf{T_z} \\ \mathbf{C_z} &= \mathbf{C}\mathbf{T_z} \end{aligned} \tag{3.27}$$

In this formulation, the state vector $\mathbf{z}(k)$ is still known as in the formulation where the state vector was $\mathbf{\Phi}(k)$; however, unlike $\mathbf{\Phi}(k)$, $\mathbf{z}(k)$ does not satisfy the rotation equation.

For the above formulation to be valid, the transformation matrix $\mathbf{T_z}$ must have a unique pseudoinverse $\mathbf{T_z}^{+}$. This is possible only if the matrix

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}(s+p) & \mathbf{z}(s+p+1) & \cdots & \mathbf{z}(l) \end{bmatrix} \tag{3.28}$$

is full-rank. Yet, as discussed in Section 2.3, the time-shifted rows of $\mathbf{\Phi}(k)u(k)$ contained by $\mathbf{Z}$ are not linearly independent. A full-rank matrix $\mathbf{W}$, whose rows span the same rowspace as $\mathbf{Z}$ and, thus, whose columns $\mathbf{w}(k)$ $(s+p \leq k \leq l)$ can be be used as the steady-state portion of the state, is formed by performing a singular value decomposition on the matrix $\mathbf{Z}$ as follows:

$$\begin{aligned} \mathbf{W} &= \begin{bmatrix} \mathbf{w}(s+p) & \mathbf{w}(s+p+1) & \cdots & \mathbf{w}(l) \end{bmatrix} \\ &= \mathbf{V_1}^T \end{aligned} \tag{3.29}$$

where

$$\begin{aligned} \mathbf{Z} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V^T} \\ &= \begin{bmatrix} \mathbf{U_1} & \mathbf{U_2} \end{bmatrix} \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V_1} & \mathbf{V_2} \end{bmatrix}^T \end{aligned} \tag{3.30}$$

and the banded $\mathbf{\Sigma}$ contains the singular values of $\mathbf{Z}$ on its main diagonal, while its submatrix $\mathbf{S}$ represents the diagonal matrix of non-zero singular values. The matrices $\mathbf{U}$ and $\mathbf{V}$ are

unitary matrices comprised of the submatrices $\mathbf{U_1}$, $\mathbf{U_2}$ and $\mathbf{V_1}$, $\mathbf{V_2}$, respectively, where $\mathbf{U_1}$ and $\mathbf{V_1}$ are comprised of those columns of $\mathbf{U}$ and $\mathbf{V}$ corresponding to non-zero singular values, while $\mathbf{U_2}$ and $\mathbf{V_2}$ contain the remaining columns. The columns of the matrix $\mathbf{W}$ form the steady-state portion of the time-history of the vector $\mathbf{w}(k)$. The relationship between

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(s+p) & \mathbf{x}(s+p+1) & \cdots & \mathbf{x}(l) \end{bmatrix} \tag{3.31}$$

and $\mathbf{W}$ can be written as follows:

$$\begin{aligned} \mathbf{X} &= \mathbf{T_z Z} \\ &= \mathbf{T_z}(\mathbf{U\Sigma V}^T) \\ &= (\mathbf{T_z U_1 S}){\mathbf{V_1}}^T \\ &= (\mathbf{T_z U_1 S})\mathbf{W} \end{aligned} \tag{3.32}$$

in accordance with Eq.(3.29). For $s+p \leq k \leq l$, the above relationship implies that

$$\mathbf{x}(k) = \mathbf{T_w w}(k) \tag{3.33}$$

where $\mathbf{T_w} = \mathbf{T_z U_1 S}$ and $\mathbf{w}(k)$ denotes to the $(k-s-p+1)st$ column of $\mathbf{W}$.

3.3.2 Identification of the ELM

Plugging Eq.(3.33) into Eq.(3.1) yields the following ELM:

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{A_w w}(k) + (\mathbf{B_e})_\mathbf{w} \mathbf{\Phi}(k) u(k) \\ \mathbf{y}(k) &= \mathbf{C_w w}(k) + \mathbf{D}u(k) \end{aligned} \tag{3.34}$$

where $s+p \leq k \leq l-1$ and

$$\begin{aligned} \mathbf{A_w} &= {\mathbf{T_w}}^+ \mathbf{A T_w} \\ (\mathbf{B_e})_\mathbf{w} &= {\mathbf{T_w}}^+ \mathbf{B_e} \\ \mathbf{C_w} &= \mathbf{C T_w} \end{aligned} \tag{3.35}$$

The recursive relationship in Eq.(3.34) can be converted to a matrix equation as follows:

$$\begin{aligned} \mathcal{W} &= \begin{bmatrix} \mathbf{A_w} & \mathbf{B_w} \end{bmatrix} \bar{\mathbf{Q}} \\ \mathcal{Y} &= \begin{bmatrix} \mathbf{C_w} & \mathbf{D} \end{bmatrix} \bar{\mathbf{L}} \end{aligned} \tag{3.36}$$

where

$$\begin{aligned}
\mathcal{W} &= \begin{bmatrix} \mathbf{w}(s+p+1) & \mathbf{w}(s+p+2) & \cdots & \mathbf{w}(l) \end{bmatrix} \\
\bar{\mathbf{Q}} &= \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{Q}}_2 \end{bmatrix} \\
\mathcal{Y} &= \begin{bmatrix} \mathbf{y}(s+p) & \mathbf{y}(s+p+2) & \cdots & \mathbf{y}(l-1) \end{bmatrix} \\
\bar{\mathbf{L}} &= \begin{bmatrix} \bar{\mathbf{Q}}_1 \\ \bar{\mathbf{L}}_2 \end{bmatrix}
\end{aligned} \tag{3.37}$$

and

$$\begin{aligned}
\bar{\mathbf{Q}}_1 &= \begin{bmatrix} \mathbf{w}(s+p) & \mathbf{w}(s+p+1) & \cdots & \mathbf{w}(l-1) \end{bmatrix} \\
\bar{\mathbf{Q}}_2 &= \begin{bmatrix} \mathbf{\Phi}(s+p)u(s+p) & \mathbf{\Phi}(s+p+1)u(s+p+1) & \cdots & \mathbf{\Phi}(l-1)u(l-1) \end{bmatrix} \\
\bar{\mathbf{L}}_2 &= \begin{bmatrix} u(s+p) & u(s+p+1) & \cdots & u(l-1) \end{bmatrix}
\end{aligned} \tag{3.38}$$

An attempt to identify the ELM above using:

$$\begin{aligned}
\begin{bmatrix} \mathbf{A}_\mathbf{w} & \mathbf{B}_\mathbf{w} \end{bmatrix} &= \mathcal{W}\bar{\mathbf{Q}}^+ \\
\begin{bmatrix} \mathbf{C}_\mathbf{w} & \mathbf{D} \end{bmatrix} &= \mathcal{Y}\bar{\mathbf{L}}^+
\end{aligned} \tag{3.39}$$

would not return a unique solution because the matrix $\bar{\mathbf{Q}}$, which needs to be inverted, contains shifted rows of $\mathbf{\Phi}(k)u(k)$, therefore, it is in general not full-rank. As in the previous method, the issue is overcome by enrichment of the excitation input $u(k)$ by a subtle amount of RC.

However, in this method, in order to lessen the error due to the approximation of the state $\mathbf{x}(k)$ by the basis functions $\phi_i(k)$ $(i = 0, 1, \cdots, 2q)$ in Eq.(2.7), the random component to be added to the excitation input $u(k)$ is chosen to be a highband random signal obtained by passing a randomly generated WGN signal through a high-pass filter. Since any physical system intrinsically behaves as a low-pass filter, the change in the state of the bilinear system due to the addition of RC to the excitation input can be kept to a minimum by using a highband random component (HRC). Thus, the final form of the excitation input $u(k)$ becomes

$$u(k) = \sum_{i=0}^{f} a_i \cos(\omega_i k\tau) + \sum_{i=1}^{f} b_i \sin(\omega_i k\tau) + hrc(k) \tag{3.40}$$

where $hrc(k)$ stands for the faint amount of HRC added. The precise amount of HRC depends on how ill-conditioned the matrices $\bar{\mathbf{Q}}$, $\bar{\mathbf{L}}$ in Eq.(3.39) are for the specific problem under consideration. The level of ill-conditioning depends on the parameter p, which depends on the number of outputs r and on the order of the system to be identified n through the requirement $p \geq rn$, in addition to the number of basis functions preferred, the frequency ranges the excitation input and basis functions cover, and so forth.

In summary, the bilinear system to be identified is excited with an input $u(k)$ of the form elaborated in Eq.(3.40) (i.e. a linear combination of sine and cosine functions, of user-selected frequencies, enriched by a small amount of HRC), and the output time-history is recorded. Next, the input time-history of $\mathbf{\Phi}(k)u(k)$ is numerically generated to serve as the input vector of the ELM. The steady-state portions of the numerically generated input time-history and the recorded output time-history are fed into Eq.(3.39) to identify the system matrices of the ELM. A minimum order realization of the ELM that has the same order as the original bilinear system will be derived by eliminating its unobservable states.

The reduction of the model order is accomplished by forming the observability matrix in Eq.(3.41)

$$\mathbf{\Theta} = \begin{bmatrix} \mathbf{C_w} \\ \mathbf{C_w A_w} \\ \vdots \\ \mathbf{C_w A_w}^{\bar{n}} \end{bmatrix} \tag{3.41}$$

where $\bar{n}$ is the order of the overparameterized ELM, and then, by performing an SVD on $\mathbf{\Theta}$ to obtain a full-rank transformation matrix

$$\mathbf{T_\Theta} = (\mathbf{U_\Theta})_1 \mathbf{S_\Theta} (\mathbf{V_\Theta})_1{}^T \tag{3.42}$$

where

$$\begin{aligned} \mathbf{\Theta} &= \mathbf{U_\Theta \Sigma_\Theta V_\Theta}^T \\ &= \begin{bmatrix} (\mathbf{U_\Theta})_1 & (\mathbf{U_\Theta})_2 \end{bmatrix} \begin{bmatrix} \mathbf{S_\Theta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} (\mathbf{V_\Theta})_1 & (\mathbf{V_\Theta})_2 \end{bmatrix}^T \end{aligned} \tag{3.43}$$

and the submatrices $(\mathbf{U_\Theta})_1$, $(\mathbf{V_\Theta})_1$ are respectively formed by the columns of the unitary matrices $\mathbf{U_\Theta}$, $\mathbf{V_\Theta}$ which correspond to the diagonal matrix of non-zero singular values $\mathbf{S_\Theta}$,

which itself is a submatrix of the banded singular value matrix $\mathbf{\Sigma_\Theta}$. On the other hand, $(\mathbf{U_\Theta})_2$, $(\mathbf{V_\Theta})_2$ are comprised of the remaining columns of $\mathbf{U_\Theta}$, $\mathbf{V_\Theta}$, respectively. Lastly, the overparameterized ELM is transformed to the coordinates in which the unobservable states are eliminated as follows:

$$\begin{aligned} \mathbf{r}(k+1) &= \mathbf{A_r}\mathbf{r}(k) + (\mathbf{B_e})_\mathbf{r}\mathbf{\Phi}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C_r}\mathbf{r}(k) + \mathbf{D}u(k) \end{aligned} \tag{3.44}$$

where $s+p \leq k \leq l-1$ and $\mathbf{r}(k) \in \mathbb{R}^n$ represents the vector of observable states, while

$$\begin{aligned} \mathbf{A_r} &= \mathbf{T_\Theta}^+\mathbf{A_w}\mathbf{T_\Theta} \\ \mathbf{B_r} &= \mathbf{T_\Theta}^+(\mathbf{B_e})_\mathbf{w} \\ \mathbf{C_r} &= \mathbf{C_w}\mathbf{T_\Theta} \end{aligned} \tag{3.45}$$

3.3.3 Recovery of the Bilinear System

The reduced order ELM in Eq.(3.44) has the same order n as the original bilinear system. In order to recover the system matrices of the original bilinear model, the steady-state portion of the state time-history of the reduced order model $\mathbf{r}(k)$ $(s+p \leq k \leq l)$ is generated by feeding $\mathbf{\Phi}(k)u(k)$ to the reduced order model in Eq.(3.44), and the identification of the bilinear system matrices is concluded using:

$$\begin{aligned} \begin{bmatrix} \mathbf{A_r} & \mathbf{B_r} & \mathbf{N_r} \end{bmatrix} &= \mathbf{R}\hat{\mathbf{Q}}^+ \\ \begin{bmatrix} \mathbf{C_r} & \mathbf{D_r} \end{bmatrix} &= \mathcal{Y}\hat{\mathbf{L}}^+ \end{aligned} \tag{3.46}$$

where

$$\begin{aligned} \mathbf{R} &= \begin{bmatrix} \mathbf{r}(s+p+1) & \mathbf{r}(s+p+2) & \cdots & \mathbf{r}(l) \end{bmatrix} \\ \hat{\mathbf{Q}} &= \begin{bmatrix} \mathbf{r}(s+p) & \mathbf{r}(s+p+1) & \cdots & \mathbf{r}(l-1) \\ \mathbf{\Phi}(s+p)u(s+p) & \mathbf{\Phi}(s+p+1)u(s+p+1) & \cdots & \mathbf{\Phi}(l-1)u(l-1) \end{bmatrix} \\ \mathcal{Y} &= \begin{bmatrix} \mathbf{y}(s+p+1) & \mathbf{y}(s+p+2) & \cdots & \mathbf{y}(l) \end{bmatrix} \\ \hat{\mathbf{L}} &= \begin{bmatrix} \mathbf{y}(s+p) & \mathbf{y}(s+p+1) & \cdots & \mathbf{y}(l-1) \\ u(s+p) & u(s+p+1) & \cdots & u(l-1) \end{bmatrix} \end{aligned} \tag{3.47}$$

The original bilinear system is related to the identified bilinear system through the following coordinate transformation:

$$\mathbf{r}(k) = \mathbf{T_r}\mathbf{x}(k) \tag{3.48}$$

such that

$$\begin{aligned}
\mathbf{A} &= \mathbf{T_r}^{-1}\mathbf{A_r}\mathbf{T_r} \\
\mathbf{B} &= \mathbf{T_r}^{-1}\mathbf{B_r} \\
\mathbf{N} &= \mathbf{T_r}^{-1}\mathbf{N_r}\mathbf{T_r} \\
\mathbf{C} &= \mathbf{C_r}\mathbf{T_r} \\
\mathbf{D} &= \mathbf{D_r}
\end{aligned} \tag{3.49}$$

where $s + p \leq k \leq l - 1$ and $\mathbf{T_r}$ is an invertible matrix.

3.4 The Effect of the Parameters the User Can Choose

In the implementation of such an identification approach, certain parameters have to be selected by the user. In this section, the analyses of the effects of each parameter on the identification results are presented. The effects of the choice of the set of excitation input frequencies $\mathbf{\Omega_0}$, perturbation order N, and data length l were already discussed in Section 2.5, and are the same for this method, as well. As a summary of the main conclusions, (i) the excitation input frequencies should preferably be selected randomly rather than evenly spaced in order to avoid ill-conditioning, (ii) the perturbation order N should be selected high enough to ensure that the approximation in Eq.(2.7) is valid (On the other hand, choosing a too high value of N causes numerical ill-conditioning in the matrices that contain shifted rows of $\mathbf{\Phi}(k)u(k)$), (iii) while a long data length l helps eliminate similar numerical ill-conditioning, it also increases the approximation error in Eq.(2.7).

3.4.1 The Parameter p

The parameter p should be chosen such that $rp \geq n$. In practice, it is usually not possible to know the system order n a priori, therefore, a high enough value of p should

be chosen. As discussed previously, the higher the choice of p, the more ill-conditioned the matrix $\bar{\mathbf{Q}}$ gets because its lower submatrix $\bar{\mathbf{Q}}_2$ contains time-shifted row blocks of $\mathbf{\Phi}(k)u(k)$, and due to the trigonometric identities in Eq.(2.6) some of these shifted rows may be linearly dependent. The addition of the HRC resolves the ill-conditioning issue at the expense of increasing the state approximation error in Eq.(2.7), as not all of the frequencies introduced to the state by the addition of HRC can be captured by the basis functions. Therefore, p should be kept as small as possible, as in the previous method but large enough to satisfy $rp \geq n$.

3.4.2 Cutoff Frequency of HRC

The generation of HRC is accomplished by passing a WGN signal through a high-pass filter that attenuates the amplitudes of the frequencies below a chosen cutoff frequency ω_c. The farther apart ω_c is chosen from the Nyquist frequency, the richer the HRC gets. As the HRC gets richer, its frequency content includes more of the lower frequencies. Therefore, on the one hand, since all physical systems act as low-pass filters, the state approximation error in Eq.(2.7) increases, for the same amount of HRC. On the other hand, a smaller amount of HRC suffices to uncorrelate the linearly dependent rows of $\bar{\mathbf{Q}}_2$.

3.4.3 The Amount of HRC

In contrast to the algorithm described in Chapter 2, where there are $p+2$ row blocks of $\mathbf{\Phi}(k)u(k)$ and $p+1$ row blocks of time-shifted $\mathbf{y}(k)$, in this algorithm there are merely $p+1$ row blocks of $\mathbf{\Phi}(k)u(k)$ and p row blocks of $\mathbf{y}(k)$. Therefore, the amount of RC necessary to avoid the ill-conditioning of $\bar{\mathbf{Q}}$ should be less compared to the previous formulation, if the added RC were in the form of WGN. Instead, the excitation input is enriched using a highband RC. Though not as rich as WGN, a small amount of HRC still suffices to obtain a full-rank $\bar{\mathbf{Q}}$ as illustrated in the succeeding example, when the choice of the cutoff frequency is not too close to the Nyquist frequency.

3.5 Numerical Example

Let's illustrate the proposed method through the analysis of the same non-dimensional example used in Section 2.6, where

$$\mathbf{A} = \begin{bmatrix} 0.9924 & 0.0118 \\ -1.1778 & 0.8747 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.0126 \\ 0.0042 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 0.2000 & 0 \\ 0 & 0.2000 \end{bmatrix}, \qquad (3.50)$$
$$\mathbf{C} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \mathbf{D} = 1.2000$$

the number of inputs $m = 1$, the number of outputs $r = 1$, the order of the system $n = 2$. According to the requirement $rp \geq n$, the minimum choice for the parameter p is 2,therefore, p will be chosen to be 2 as in the former example for direct comparison.

Let's keep the sampling period τ to be equal to 0.01 seconds for a set of input/output measurement time-history of length equal to $l = 5,000$ time-steps or 50 seconds. The excitation input $u(k)$ is constructed in accordance with Eq.(3.40). Tabulated in Table 2.6, the excitation frequencies ω_i and the amplitudes a_i, b_i are chosen to be the same as in the former numerical example. The highband random component $hrc(k)$ is obtained by passing a randomly generated WGN, whose RMS is 5% of the RMS of the base signal, through a highpass 10^{th} order Butterworth filter, which has a cut-off frequency equal to 45Hz. The cut-off frequency corresponds to 90% of the Nyquist frequency, which is 50Hz. Fig.3.1 presents the Bode plot of the aforementioned filter, while Fig.3.2a shows the designed excitation input $u(k)$ before and after the addition of HRC, shown solely in Fig.3.2b. It can be inferred from these plots that the change in $u(k)$ due to the addition of $hrc(k)$ is very small. The steady-state of portion of the state time-history is assumed to start at the time-step $s = 748$, since $\prod_{k=1}^{j} (\mathbf{A} + \mathbf{N}u(k)) < 1e-20$ for $s \leq j \leq l$.

The vector of sine and cosine basis functions $\mathbf{\Phi}(k)$ is created using a 2^{nd} order perturbation theory approximation. Although the basis functions are actually used to represent the state, the output can still be used to roughly estimate the required number of the perturbation order, since it contains the contributions from at least one of the states, which is coupled with the rest of the states in the difference equation representing the bilinear model. The contribution coming to the output $\mathbf{y}(k)$ from the term $\mathbf{D}u(k)$ does not violate

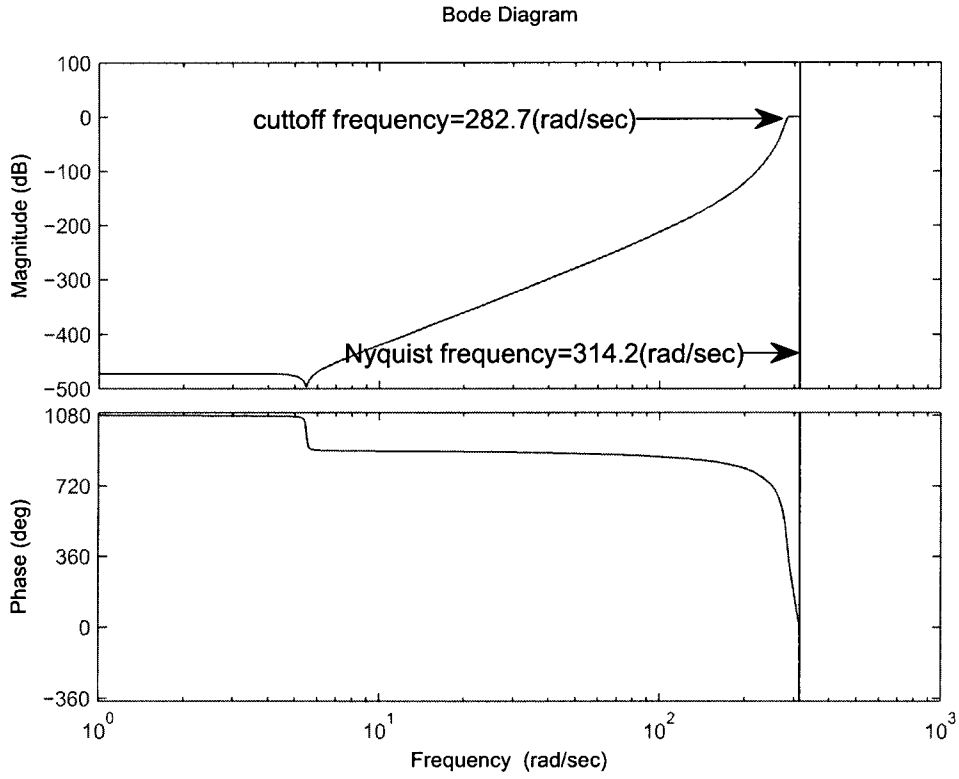
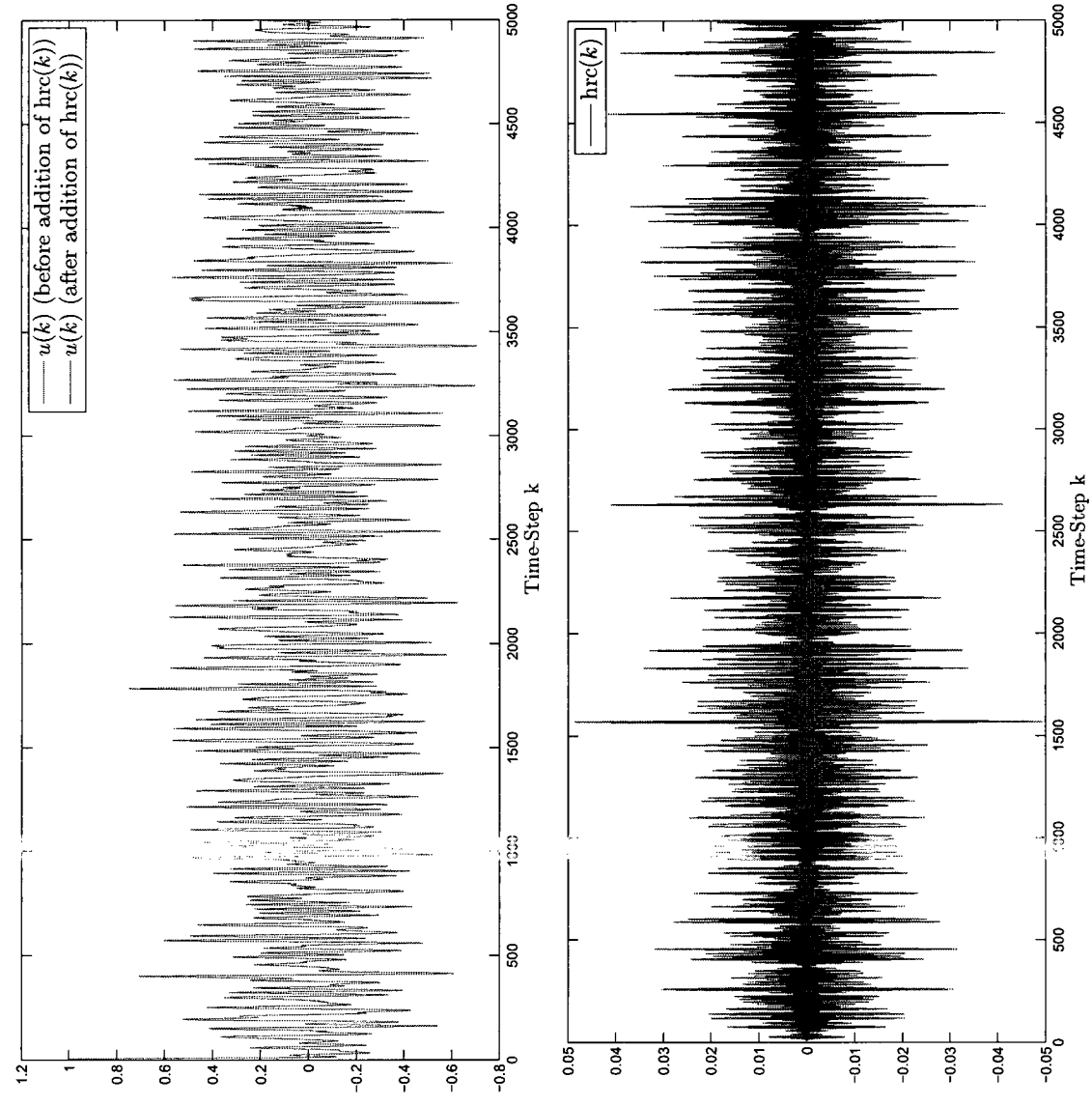


Figure 3.1: The Bode plot of the 10^{th} order Butterworth filter with cutoff frequency 45Hz used in obtaining the HRC from a randomly generated WGN signal



(a) The plot of designed excitation input $u(k)$ before and after addition of the HRC

(b) The plot of added HRC, $hrc(k)$

Figure 3.2

this reasoning, as the set of frequencies present in $u(k)$ is a subset of those present in the state $\mathbf{x}(k)$. As a demonstration of the above discussion, Fig.3.3 depicts the steady-state portion of the time-histories of the output $\mathbf{y}(k)$ and its representation $\hat{\mathbf{y}}(k)$ in terms of the basis functions $\phi_i(k)(i = 0, 1, \cdots, 2q)$, where

$$\begin{bmatrix} \hat{\mathbf{y}}(s) & \hat{\mathbf{y}}(s+1) & \cdots & \hat{\mathbf{y}}(l) \end{bmatrix} = \begin{bmatrix} \mathbf{y}(s) & \mathbf{y}(s+1) & \cdots & \mathbf{y}(l) \end{bmatrix} \begin{bmatrix} \phi(s) \\ \phi(s+1) \\ \vdots \\ \phi(l) \end{bmatrix}^{+} \begin{bmatrix} \phi(s) \\ \phi(s+1) \\ \vdots \\ \phi(l) \end{bmatrix} \tag{3.51}$$

Moreover, the amount of the mismatch between $\mathbf{y}(k)$ and its representation $\hat{\mathbf{y}}(k)$ can be quantified using the error function defined in Eq.(2.25) as $error(\mathbf{y}(k), \hat{\mathbf{y}}(k)) = 5.7451\%$. The validity of the discussion in terms of the output $\mathbf{y}(k)$ can be better observed from Fig.3.4, where the plots of the steady-state portions of the time-histories of the second state $x_2(k)$ and its representation $\hat{x}_2(k)$ in terms of the basis functions, where

$$\begin{bmatrix} \hat{x}_2(s) & \hat{x}_2(s+1) & \cdots & \hat{x}_2(l) \end{bmatrix} = \begin{bmatrix} x_2(s) & x_2(s+1) & \cdots & x_2(l) \end{bmatrix} \begin{bmatrix} \phi(s) \\ \phi(s+1) \\ \vdots \\ \phi(l) \end{bmatrix}^{+} \begin{bmatrix} \phi(s) \\ \phi(s+1) \\ \vdots \\ \phi(l) \end{bmatrix} \tag{3.52}$$

are provided. The amount of mismatch between the two is calculated, using $error(\mathbf{x}(k), \hat{\mathbf{x}}(k))$, to be 5.9394% and 8.2060% for the first and second states, respectively, which are reasonably close to the mismatch calculated for the output.

At this point, it is also worthwhile to mention that, as tabulated in Table 3.1, the mismatch between $\mathbf{x}(k)$ and $\hat{\mathbf{x}}(k)$ would decrease if the excitation input $u(k)$ were not enriched at all, and would increase if the excitation input were enriched using the generated WGN without passing it through the designed Highpass Butterworth Filter. Thus, the enrichment of $u(k)$ should be performed by addition of HRC.

The time-history of the vector $\mathbf{z}(k)$, whose rows span the state-space of the ELM, is

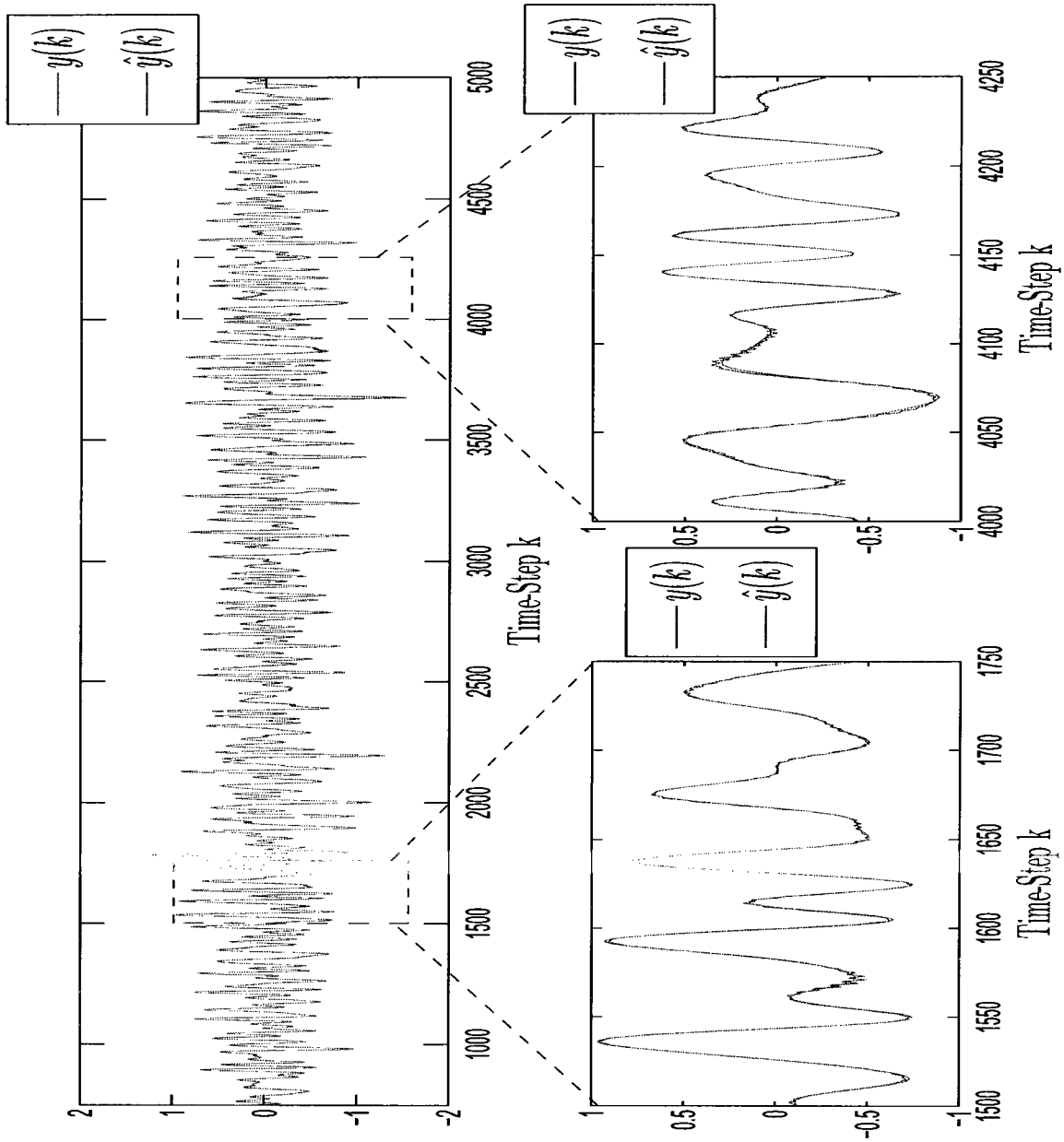


Figure 3.3: The plot of the steady-state portion of the time-history of $\mathbf{y}(k)$ and its representation $\hat{\mathbf{y}}(k)$, in terms of the basis functions

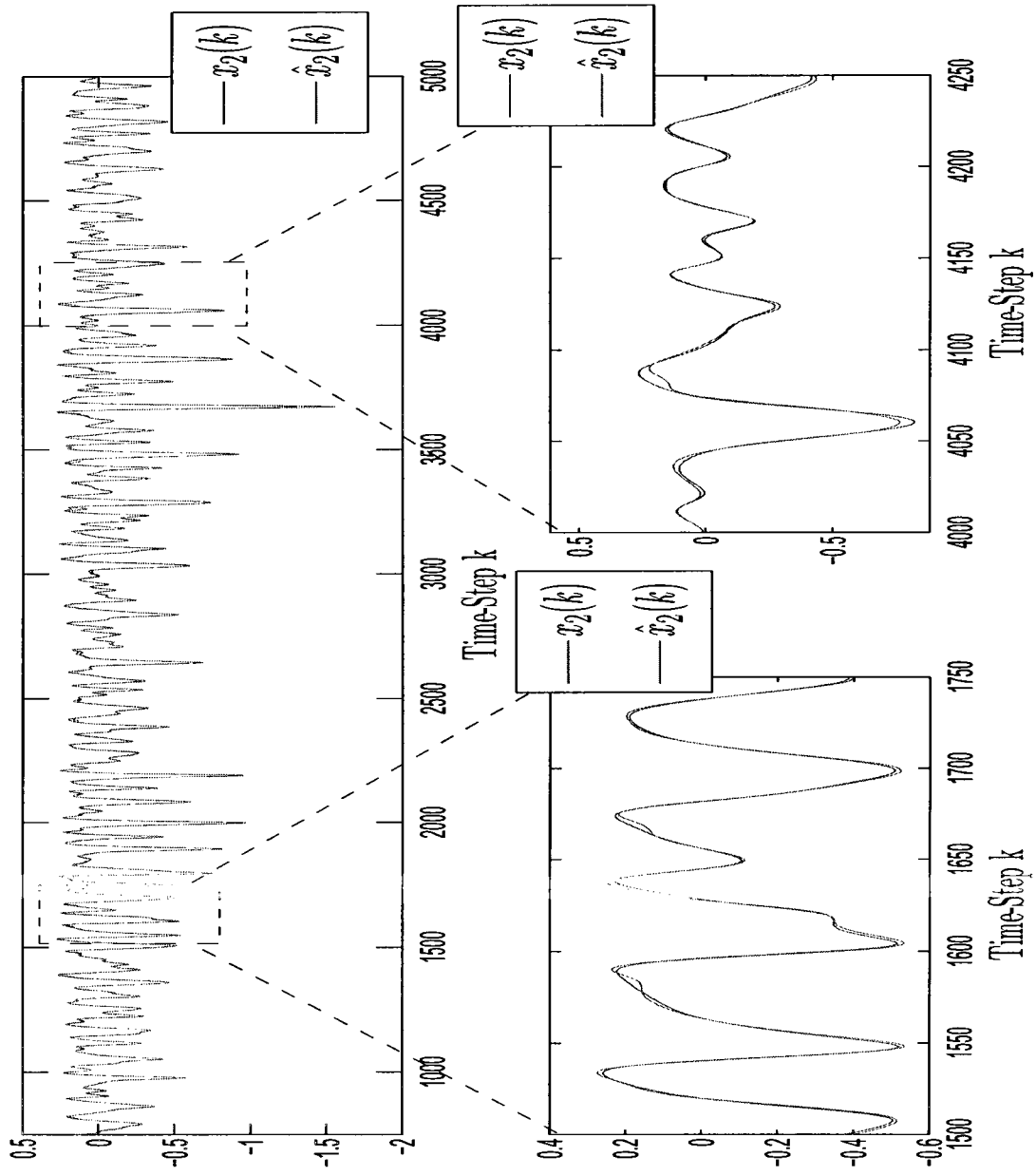


Figure 3.4: The plot of the steady-state portion of the time-history of the second state $x_2(k)$ and its representation $\hat{x}_2(k)$, in terms of the basis functions

Enrichment By	$error(x_1(k), \hat{x}_1(k))$	$error(x_2(k), \hat{x}_2(k))$
None	5.9322%	8.2045%
HRC	5.9394%	8.2060%
WGN	6.1225%	8.3861%

Table 3.1: The effect of the type of enrichment of $u(k)$ on the mismatch between the first and second states, $x_1(k)$ and $x_2(k)$, and their representations, denoted by $\hat{x}_1(k)$ and $\hat{x}_2(k)$, in terms of the sine and cosine basis functions

generated using the past input and output time-histories through the following relationship:

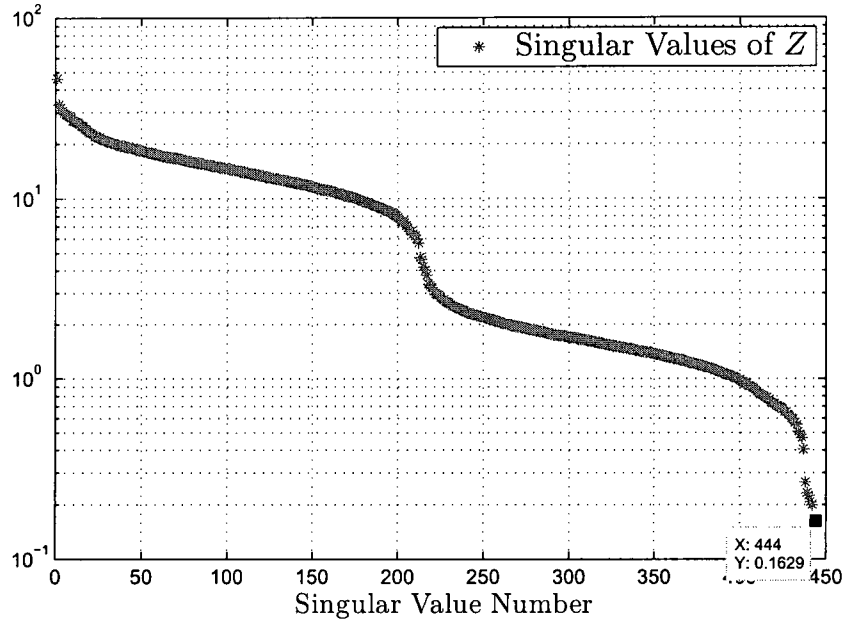
$$\mathbf{z}(k) = \begin{bmatrix} \mathbf{y}(k-1) \\ \mathbf{y}(k-2) \\ \mathbf{\Phi}(k-1)u(k-1) \\ \mathbf{\Phi}(k-2)u(k-2) \end{bmatrix} \tag{3.53}$$

where $s+p \leq k \leq l$. The generated time-history is cast into the matrix form as follows:

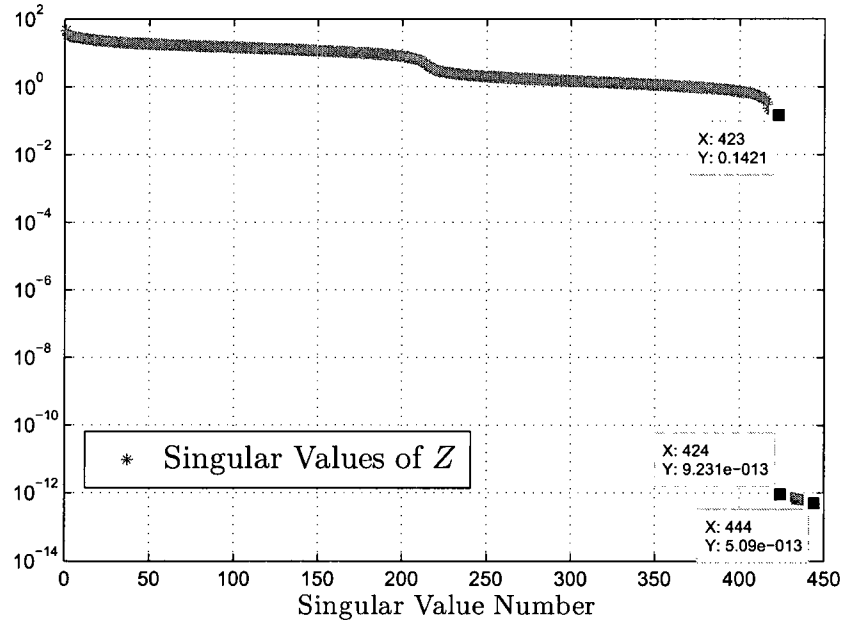
$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}(s+p) & \mathbf{z}(s+p) & \cdots & \mathbf{z}(l) \end{bmatrix} \tag{3.54}$$

where the columns of the matrix $\mathbf{Z}$ form the time-history of the vector $\mathbf{z}(k)$. The matrix $\mathbf{W}$, whose columns form the time-history of the vector of basis functions $\mathbf{w}(k)$, is obtained by performing a singular value decomposition to $\mathbf{Z}$ as described in Eq.(3.29) and (3.30). Fig.3.5a exhibits the singular value plot of $\mathbf{Z}$, while Fig.3.5b presents the same plot if the excitation input contained no HRC. It is clear from these plots that the addition of HRC causes $\mathbf{Z}$ to become full-rank.

Then, the matrix $\mathbf{W}$ is obtained as in Eq.(3.29). Although no rows of $\mathbf{Z}$ are linearly dependent, which implies that $\mathbf{z}(k)$ could be used as the state of the ELM, $\mathbf{W}$ has a better condition number than $\mathbf{Z}$ (the 2-norm condition number of $\mathbf{W}$ is 1.0000, while the 2-norm condition number of $\mathbf{Z}$ is $2.8150 \cdot 10^2$); thus, as the state of the ELM, the usage of $\mathbf{w}(k)$ is preferred over $\mathbf{z}(k)$, though both are valid choices. The aforementioned reasoning can be demonstrated by calculating the mismatch between the state $\mathbf{x}(k)$ and its representation in



(a) The plot of the singular values of the matrix **Z**



(b) The plot of the singular values of the matrix **Z**, if no HRC were added to the excitation input $u(k)$

Figure 3.5

terms of the rows of $\mathbf{Z}$, $\hat{\mathbf{x}}_{\mathbf{z}}(k)$, where

$$\left[\hat{\mathbf{x}}_{\mathbf{z}}(s+p) \quad \hat{\mathbf{x}}_{\mathbf{z}}(s+p+1) \quad \cdots \quad \hat{\mathbf{x}}_{\mathbf{z}}(l)\right] = \left[\mathbf{x}(s+p) \quad \mathbf{x}(s+p+1) \quad \cdots \quad \mathbf{x}(l)\right] \mathbf{Z}^{+}\mathbf{Z} \tag{3.55}$$

using $error(\mathbf{x}(k), \hat{\mathbf{x}}_{\mathbf{z}}(k))$ as 1.7267% and 0.4698% for the first and second states respectively. Moreover, Fig.3.6 indicates that the time-histories of the steady-state portions of the second original state $x_2(k)$ and its representation $(\hat{x}_z)_2(k)$ match "almost" perfectly. In the light of the above findings, it is concluded that the steady-state portions of the past input and output time-histories form a set of basis functions in terms of which the steady-state portion of the state time-history can be represented accurately. Recalling that the rowspace of $\mathbf{W}$ is the same as the rowspace of $\mathbf{Z}$, the same conclusion is also valid for the set of basis functions formed by the rows of $\mathbf{W}$. In fact, it has been shown both numerically and visually that the past input and output time-histories outperform the sine and cosine basis functions in representing the state more accurately.

The identification of an overparameterized ELM is accomplished by Eq.(3.39). As discussed previously, a unique identification of the ELM is contingent upon the pseduoinversion of $\bar{\mathbf{Q}}$ uniquely, which is only possible if the excitation input $u(k)$ is enriched by a subtle amount of HRC. Fig.3.7a presents the plot of the singular values of $\bar{\mathbf{Q}}$, while Fig.3.7b shows the same plot if $u(k)$ were not enriched. The plots indicate that $\bar{\mathbf{Q}}$ contains linearly dependent rows, unless $u(k)$ is enriched.

As shown in Fig.3.8, where the norm of the eigenvalues of the matrix $\mathbf{A_w}$ are presented, only 2 out of the 444 eigenvalues of $\mathbf{A_w}$ are non-negligible. Hence, the order of the overparameterized model can be reduced to 2. The redundant states of the ELM can be discarded through a model reduction process, where the unobservable states are eliminated. Fig.3.9a presents the plot of the singular values of the observability matrix $\mathbf{\Theta}$ defined as follows:

$$\mathbf{\Theta} = \begin{bmatrix} \mathbf{C_w} \\ \mathbf{C_w A_w} \\ \mathbf{C_w A_w}^2 \\ \vdots \\ \mathbf{C_w A_w}^{443} \end{bmatrix} \tag{3.56}$$

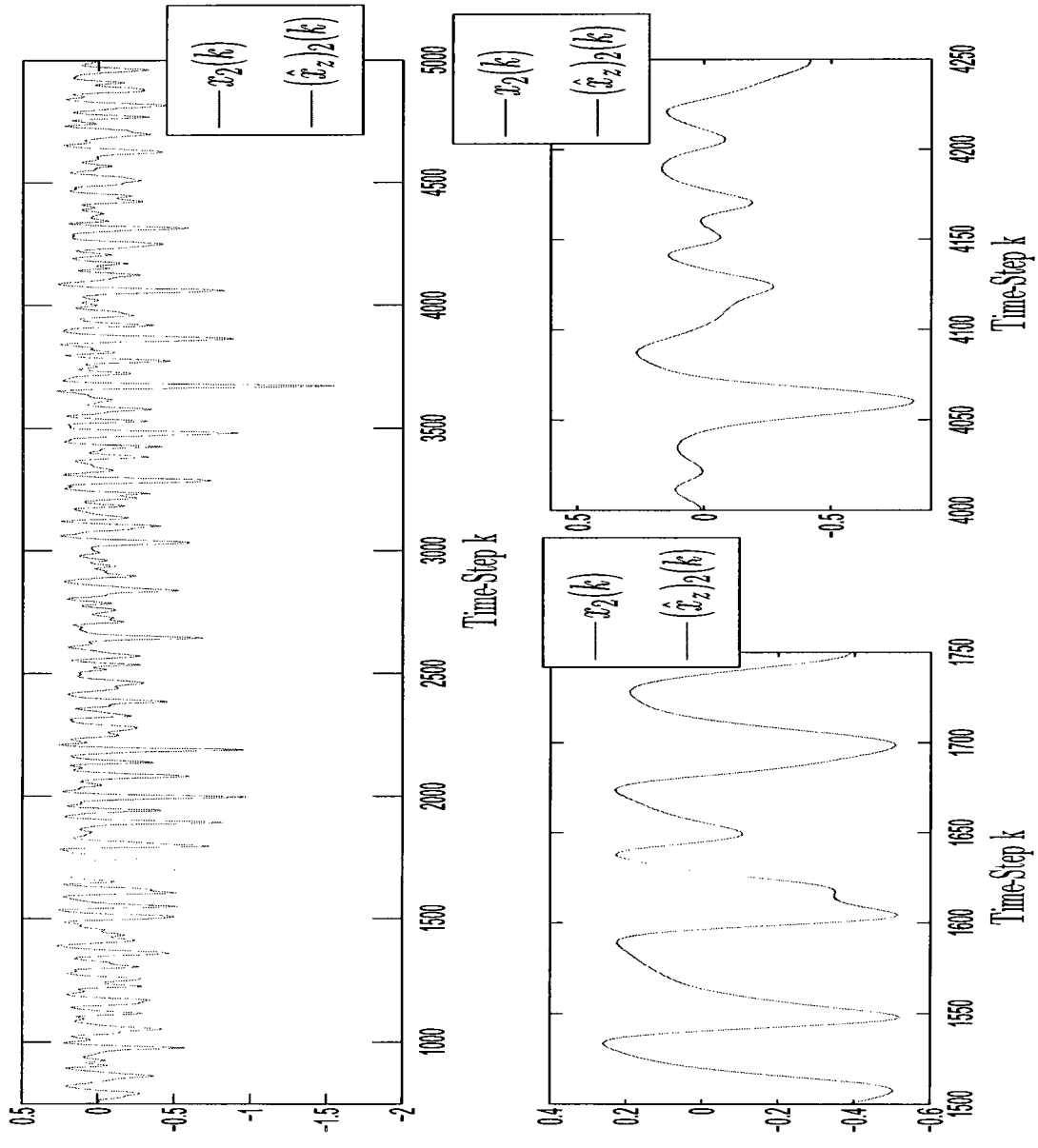
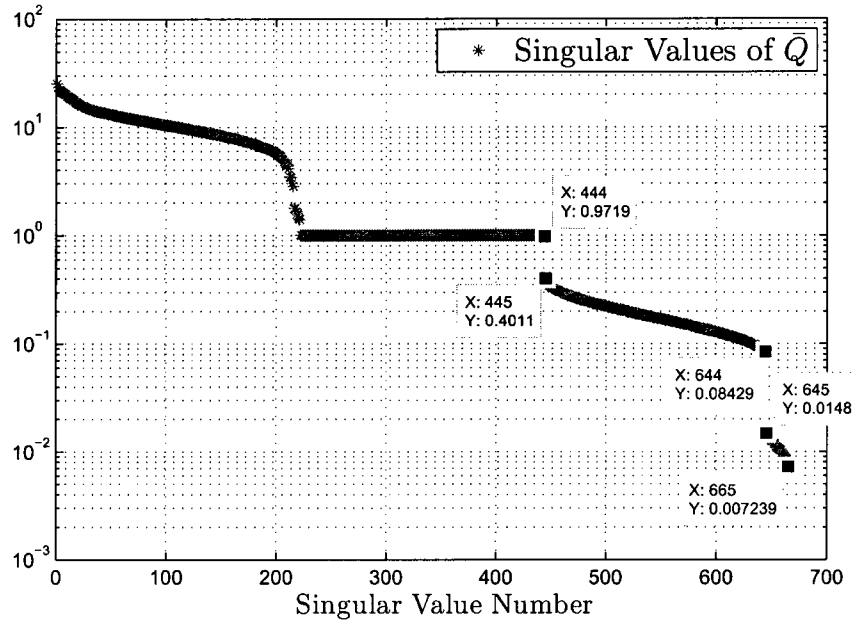
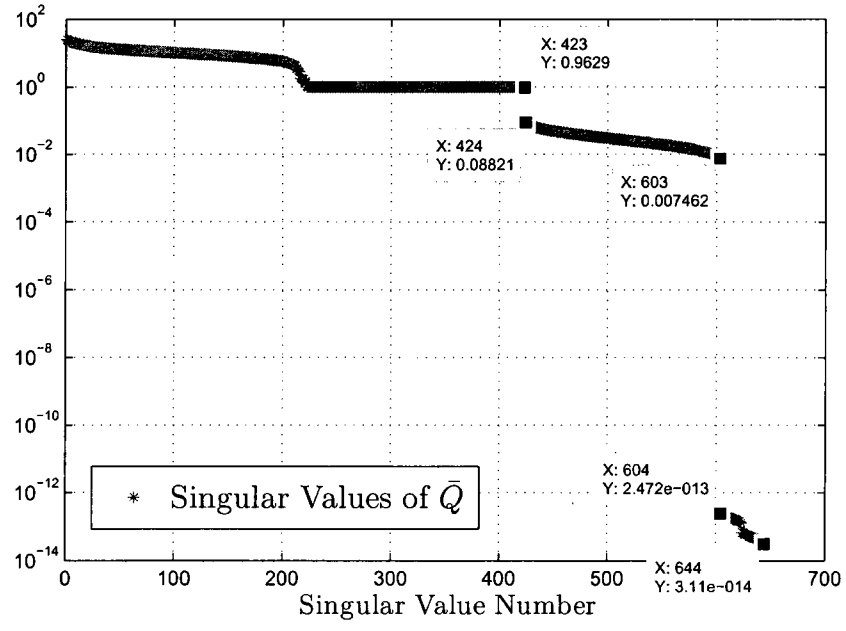


Figure 3.6: The plot of the steady-state portion of the time-history of the second state $x_2(k)$ and its representation $(\hat{x}_z)_2(k)$, in terms of the basis functions



(a) The plot of the singular values of the matrix $\bar{\mathbf{Q}}$



(b) The plot of the singular values of the matrix $\bar{\mathbf{Q}}$, if no HRC were added to the excitation input $u(k)$

Figure 3.7

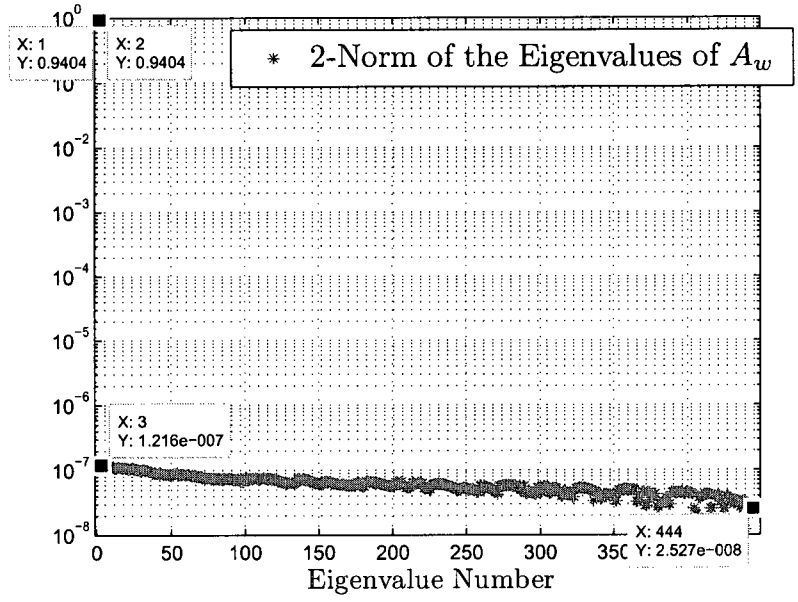


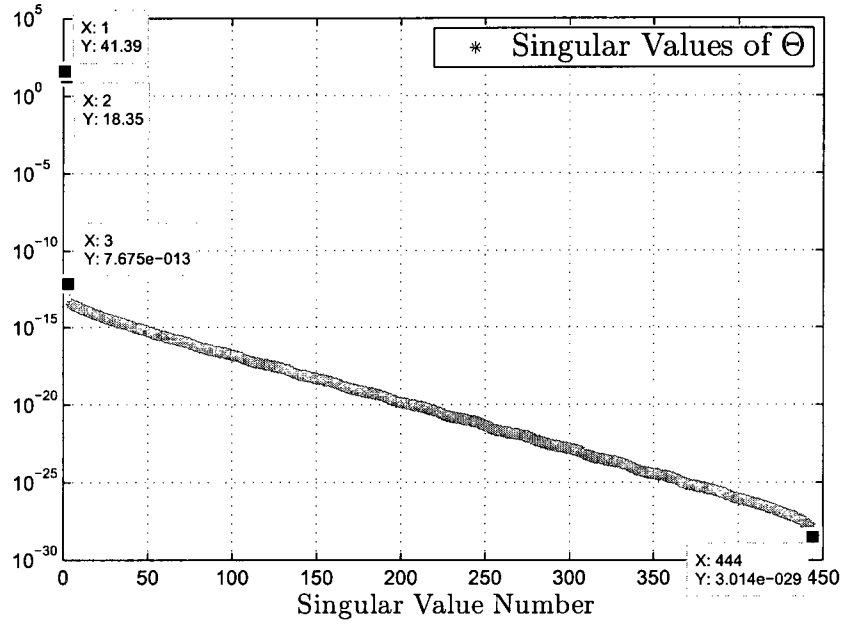
Figure 3.8: The plot of the 2-norm of the eigenvalues of the matrix $\mathbf{A_w}$

while Fig.3.9b presents the same plot for the controllability matrix $\mathcal{C}$ defined as follows:

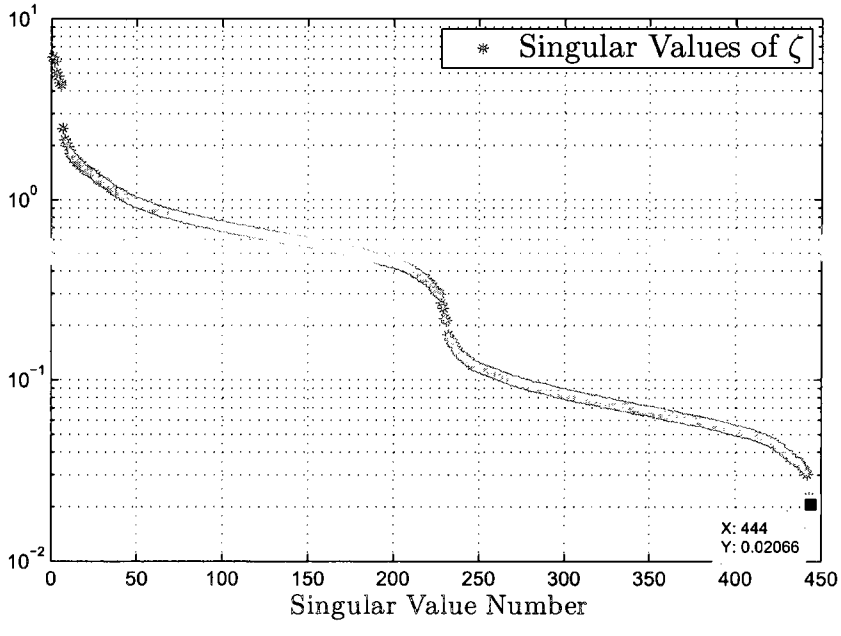
$$\zeta = \begin{bmatrix} \mathbf{B_w} & \mathbf{A_w B_w} & \cdots & \mathbf{A_w}^{443}\mathbf{B_w} \end{bmatrix} \tag{3.57}$$

It can be inferred from these plots that while all states are controllable, there exist only 2 observable states. Since the identified model must contain only the states which are both observable and controllable, the 442 unobservable states are eliminated to obtain a minimal realization of order 2 as described in Eq.(3.44).

Both of the ELMs represented by Eq.(3.34) and (3.44) are capable of reproducing the steady-state portion of the output $\mathbf{y}(k)$, when fed the input $\Phi(k)u(k)$. In fact, while Fig.3.10a presents the steady-state portion of the output reproduced by the ELM in Eq.(3.34), denoted by $\mathbf{y_w}(k)$, against the original output $\mathbf{y}(k)$, Fig.3.10b shows the steady-state portion of the output reproduced by the ELM in Eq.(3.44), denoted by $\mathbf{y_r}(k)$, against $\mathbf{y}(k)$. It can be inferred from these figures that there is a strong agreement between the reproduced and original outputs. Furthermore, the mismatch between the reproduced outputs and the original output can be quantified by using the error function defined in Eq.(2.25) as



(a) The plot of the singular values of the observability matrix Θ



(b) The plot of the singular values of the controllability matrix ζ

Figure 3.9

$error(\mathbf{y}(k), \mathbf{y_w}(k)) = 1.0313\%$ and $error(\mathbf{y}(k), \mathbf{y_r}(k)) = 1.0313\%$. The fact that both error functions yield the same value can be explained by recalling that the unobservable states does not contribute to the output. Moreover, the strong agreement between the reproduced and original outputs is indicative of the quality of the agreement between the reproduced and original states, even though they cannot be compared directly, as they are related to each other through a coordinate transformation.

Finally, the steady-state portion of the reproduced state $\mathbf{x_r}(k)$ is used to identify the bilinear system matrices as in Eq.(3.46) to conclude the identification, where the identified system matrices are as follows:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0.9672 & 0.0745 \\ -0.1593 & 0.9016 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.0035 \\ 0.0085 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 0.1995 & 0.0094 \\ 0.0058 & 0.1882 \end{bmatrix}, \\ \mathbf{C} &= \begin{bmatrix} -9.7764 & 7.3912 \end{bmatrix}, \mathbf{D} = 1.1999 \end{aligned} \tag{3.58}$$

Fig.3.11 presents the singular values of the matrix $\hat{\mathbf{Q}}$, which needs to be full-rank for a unique identification. As it can be seen from the plot, $\hat{\mathbf{Q}}$ is uniquely invertible, since the ratio of the largest singular value to the smallest singular value, namely its 2-norm condition number, is $7.0998 \cdot 10^1$.

As a measure of the accuracy of the identification, a new random input of length $20,000$ time-steps is generated from a Gaussian distribution with mean equal to 0 and standard deviation equal to 1. The generated test input $\mathbf{u}_t(k)$ is fed both to the original and to the identified bilinear systems and the corresponding output time histories, $\mathbf{y}_t(k)$ and $\mathbf{y}_p(k)$ respectively, are obtained. Fig.3.12 shows the plot of the predicted output time-history $\mathbf{y}_p(k)$ against the actual output time-history $\mathbf{y}_t(k)$ with a prediction error of 1.7409% between the two. The small prediction error clearly indicates that the predicted output $\mathbf{y}_p(k)$ is a very accurate estimate of the output $\mathbf{y}_t(k)$ of the bilinear system.

Furthermore, the original bilinear system matrices $\mathbf{A}$ and $\mathbf{N}$ are related to the identified bilinear system matrices $\mathbf{A}_r$ and $\mathbf{N}_r$, respectively, through a similarity transform as shown in Eq.(3.49). Therefore, the eigenvalues of $\mathbf{A}$, $\mathbf{A}_r$ and $\mathbf{N}$, $\mathbf{N}_r$ need to be the same. For the example above, these eigenvalues are found to be $\{0.9336 \pm 0.1022i\}$, $\{0.9344 \pm 0.1039i\}$ and

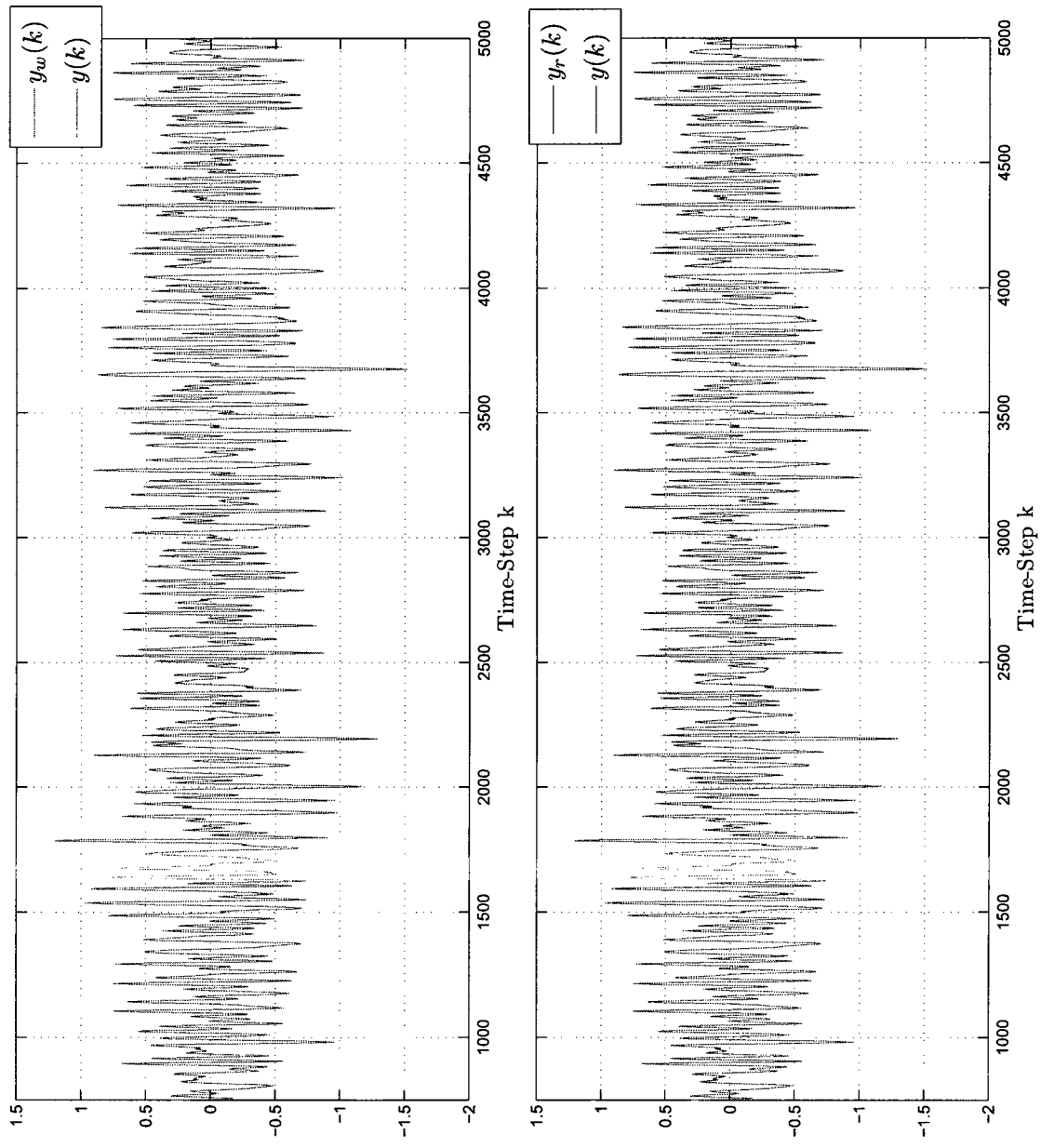


Figure 3.10: The plot of the steady-state portion of the time history of the reproduced outputs $\mathbf{y}_w(k)$ and $\mathbf{y}_r(k)$ against the original output $\mathbf{y}(k)$

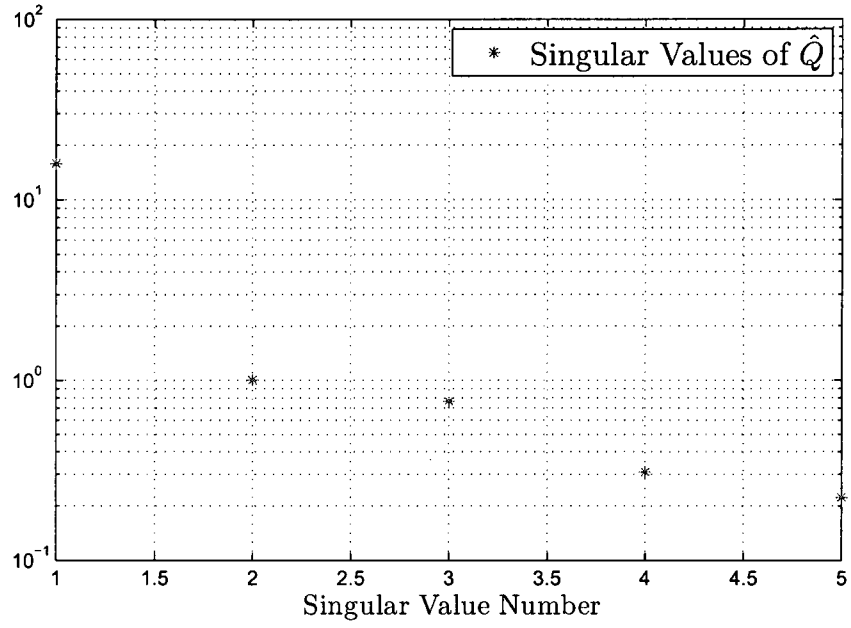


Figure 3.11: The plot of the singular values of the matrix $\hat{\mathbf{Q}}$

$\{0.2000, 0.2000\}$, $\{0.2032, 0.1846\}$, respectively. Being invariant against coordinate transformations, the direct transmission matrix associated with the original system $\mathbf{D} = 1.2000$ and identified bilinear system $\mathbf{D}_r = 1.1999$ should be the same, when there is no identification error. The "almost" perfect match between the eigenvalues of the identified and original system matrices and the "almost" exact estimation of the direct transmission term confirms the success of the identification.

3.6 Conclusions

In the previous chapter, it was shown that a bilinear system could be converted to an ELM by representing its steady-state portion of its state as a linear combination of sine and cosine basis functions. Since the bilinear system matrices are extracted from the identified ELM, it is important that the conversion from the bilinear system to the ELM is accurate. In this chapter, it was shown using the interaction matrix formulation that the state could be expressed in terms of past input and output time-histories for a more accurate representation

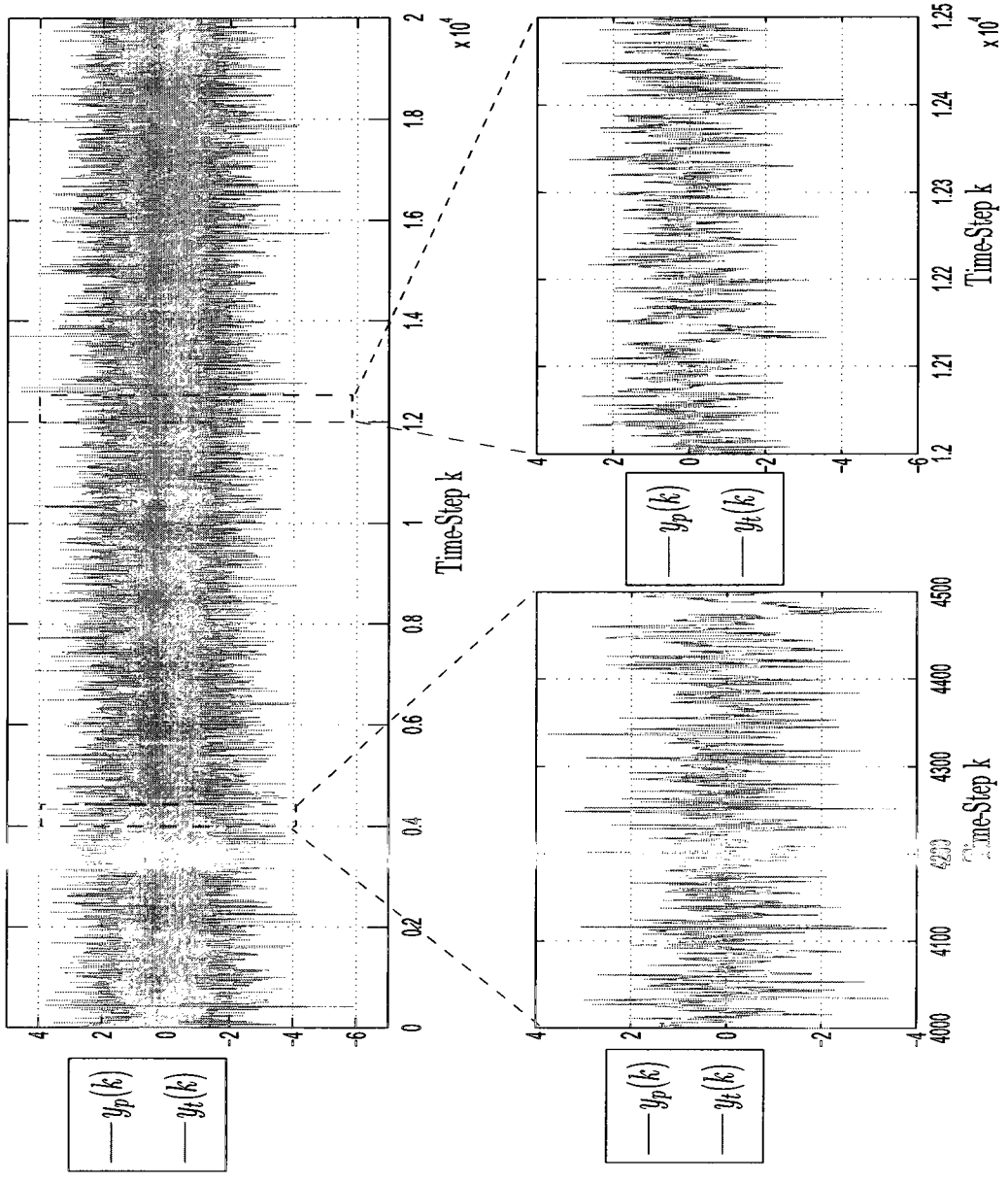


Figure 3.12: The predicted output time-history $\mathbf{y}_p(k)$ and the simulated test output time-history $\mathbf{y}_t(k)$

of the state and, thus, for a more accurate conversion. By discarding the unobservable states of the resulting overparameterized ELM, it was illustrated that a new, minimal order ELM could be obtained. The bilinear system matrices are, then, extracted from the reduced order ELM. The accuracy of the identified bilinear system was confirmed by the fact that it was able to predict the response of the original bilinear system due to a new, randomly generated test input. During the implementation of the proposed algorithm the user is given the freedom of selecting certain parameters. The effect of choosing a high parameter p was shown to cause ill-conditioning. Thus, it should be kept as small as possible. It was illustrated that as the cutoff frequency of the HRC increases, the state approximation error decreases; while the state approximation error increases, as the amount of HRC increases. The concepts of input reduction and input absorption, which are processes to obtain a unique ELM, were also introduced.

Chapter 4

Identification of Bilinear Systems Through Equivalent Linear Models via an Iterative Method

This chapter presents the third proposed method for the identification of single-input, multi-output, time-invariant, discrete-time bilinear systems using the steady-state portion of single-input, multi-output time-histories obtained by exciting the system with a linear combination of sine and cosine functions. These functions of user-selected frequencies are enriched by a subtle amount of highband random component (HRC). The method relies on the conversion of the bilinear model into an equivalent linear model (ELM) by an accurate approximation of the state in the bilinear term using a set of sine and cosine basis functions whose frequencies are obtained from perturbation theory for an $N)^{th}$ order of approximation. The addition of the highband random component (HRC), along with a step in which the columns of the ELM matrices that correspond to the dependent rows of the input vector are compressed into single columns, is needed so to obtain an input vector with linearly independent rows. The aforementioned ELM is identified using a direct method of LTI system identification, and the steady-state portion of the state-time history of the bilinear model is reproduced by feeding the identified ELM the input used in its identification. The steady-state portion of the generated state time-history corresponds to the $(N+1)^{th}$ order approximation of the steady-state portion of the state time-history of the bilinear system.

In this approach, a new ELM is obtained by rearranging the terms of the perturbation theory based approximation equation, whose state is the $(N+2)th$ order approximation to

the state of the bilinear system and whose input vector is comprised of the rows $u(k)$ and $\mathbf{x}_{N+1}(k)u(k)$, where $\mathbf{x}_{N+1}(k)$ is the $(N+1)^{th}$ order approximation to the state. The new ELM is identified using a direct method of LTI system identification, and the steady-state portion of the state time-history of the bilinear model is reconstructed by exciting it with the input used in its identification. The steady-state portion of the generated state time-history corresponds to the $(N+2)th$ order approximation to the steady-state portion of the state time-history. The process is repeated iteratively until the steady-state portions of the consecutively reproduced state time-histories differ only by a negligibly small amount. At this point, the bilinear system is retrieved using the steady-state portion of the identified state time-history.

4.1 Introduction

The proposed method addresses the identification of the bilinear system

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \mathbf{N}\mathbf{x}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}u(k), \quad 0 \leq k \leq l \end{aligned} \tag{4.1}$$

As in the former method, the bilinear system is excited by $u(k)$, a linear combination of sine and cosine functions of user-selected frequencies $\omega_i (i = 0, 1, \cdots, f)$, as shown below:

$$u(k) = \sum_{i=0}^{f} a_i \cos(\omega_i k\tau) + \sum_{i=1}^{f} b_i \sin(\omega_i k\tau), \quad 0 \leq k \leq l \tag{4.2}$$

where $b_i (i = 1, 2, \cdots, f)$ and $a_i (i = 0, 1, \cdots, f)$ are respectively the amplitudes of the sine and cosine functions, while τ represents the sampling period. Then, the bilinear model can be converted to an equivalent linear model by approximating the steady-state portion of the state $\mathbf{x}(k)$ in the bilinear term $\mathbf{x}(k)u(k)$ in Eq.(4.1) using a linear combination of sine and cosine basis functions, whose frequencies are obtained by perturbation theory. The details regarding this procedure can be found in Section 2.2, where it is also stated that the state of the bilinear system in Eq.(4.3)

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \epsilon\mathbf{N}\mathbf{x}(k)u(k), \quad 0 \leq k \leq l \tag{4.3}$$

can be represented by the power series:

$$\mathbf{x}(k) = \sum_{i=0}^{\infty} \epsilon^i \boldsymbol{\chi}_i(k), \qquad s \leq k \leq l \tag{4.4}$$

and the N^{th} order approximation $\mathbf{x}_N(k)$ to the state $\mathbf{x}(k)$ is the truncated power series:

$$\mathbf{x}_N(k) = \sum_{i=0}^{N} \epsilon^i \boldsymbol{\chi}_i(k), \qquad s \leq k \leq l \tag{4.5}$$

where $\chi_i(k)(i = 1, 2, \cdots, N)$ are not only the solutions to Eq.(4.6)

$$\begin{aligned}
\boldsymbol{\chi}_0(k+1) &= \mathbf{A}\boldsymbol{\chi}_0(k) + \mathbf{B}u(k) \\
\boldsymbol{\chi}_1(k+1) &= \mathbf{A}\boldsymbol{\chi}_1(k) + \mathbf{N}\boldsymbol{\chi}_0 u(k) \\
\boldsymbol{\chi}_2(k+1) &= \mathbf{A}\boldsymbol{\chi}_2(k) + \mathbf{N}\boldsymbol{\chi}_1 u(k) \\
&\vdots \\
\boldsymbol{\chi}_N(k+1) &= \mathbf{A}\boldsymbol{\chi}_N(k) + \mathbf{N}\boldsymbol{\chi}_{N-1} u(k)
\end{aligned} \tag{4.6}$$

but also the basis functions in terms of which $\mathbf{x}_N(k)$ is expressed.

4.2 The Proposed Method

4.2.1 Perturbation Theory Revisited

The parameter ϵ, which controls the amount of perturbation (bilinearity) added to the corresponding linear system

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) \tag{4.7}$$

is set to 1 so that Eq.(4.3) takes the form of the bilinear system to be identified, shown in Eq.(4.1). Thus, the expression for the state in Eq.(4.4) becomes

$$\mathbf{x}(k) = \sum_{i=0}^{\infty} \boldsymbol{\chi}_i(k), \qquad s \leq k \leq l \tag{4.8}$$

while the N^{th} order approximation $\mathbf{x}_N(k)$ to the state $\mathbf{x}(k)$ reduces to the truncated power series:

$$\mathbf{x}_N(k) = \sum_{i=0}^{N} \boldsymbol{\chi}_i(k), \qquad s \leq k \leq l \tag{4.9}$$

Let's add both sides of Eq.(4.6) to obtain a single equation:

$$\sum_{i=0}^{N} \boldsymbol{\chi}_i(k) = \mathbf{A}\sum_{i=0}^{N} \boldsymbol{\chi}_i(k) + \mathbf{B}u(k) + \mathbf{N}\sum_{i=0}^{N-1} \boldsymbol{\chi}_i(k)u(k), \quad s \leq k \leq l \tag{4.10}$$

where the left-hand-side of the equation is equal to the series in Eq.(4.9). The series on the right-hand-side of Eq.(4.10) can also be rewritten in the light of Eq.(4.9) as:

$$\mathbf{x}_N(k+1) = \mathbf{A}\mathbf{x}_N(k) + \mathbf{B}u(k) + \mathbf{N}\mathbf{x}_{N-1}(k)u(k), \quad s \leq k \leq l \tag{4.11}$$

thereby leading to a recursive relationship in terms of the approximation order N. Replacing N in Eq.(4.11) by $N+1$, the recursive relationship becomes:

$$\mathbf{x}_{N+1}(k+1) = \mathbf{A}\mathbf{x}_{N+1}(k) + \mathbf{B}u(k) + \mathbf{N}\mathbf{x}_N(k)u(k), \quad s \leq k \leq l \tag{4.12}$$

Eq.(4.12) can be used to generate the steady-state portion of the $(N+1)th$ order approximation to the state $\mathbf{x}(k)$, if the time-history of the steady-state portion of the N^{th} order approximation $\mathbf{x}_N(k)$ and the bilinear system matrices $\mathbf{A}, \mathbf{B}, \mathbf{N}$ are known. In fact, Eq.(4.12) can be rewritten in the form of an ELM as follows, which exhibits the validity of this statement more evidently:

$$\mathbf{x}_{N+1}(k+1) = \mathbf{A}\mathbf{x}_{N+1}(k) + \begin{bmatrix}\mathbf{B} & \mathbf{N}\end{bmatrix}\begin{pmatrix} u(k) \\ \mathbf{x}_N(k)u(k)\end{pmatrix}, \quad s \leq k \leq l \tag{4.13}$$

As elaborated in Section 2.2, the steady-state portion of $\mathbf{x}_N(k)$ can be approximated using sine and cosine basis functions of frequencies obtained by perturbation theory for a choice of approximation order N, as shown in Eq.(2.7) and repeated below for convenience:

$$\mathbf{x}_N(k) = \boldsymbol{\Lambda}\boldsymbol{\Phi}(k), \quad s \leq k \leq l \tag{4.14}$$

where $\boldsymbol{\Lambda}$ is the matrix of coefficients that multiplies the vector of basis functions $\boldsymbol{\Phi}(k)$, comprised of rows $\phi_i(k)(i = 0, 1, \cdots, 2q)$ such that $\phi_0(k) = 1$, $\phi_i(k) = \sin(\hat{\omega}_i k\tau)$ while $\phi_{q+i}(k) = \cos(\hat{\omega}_i k\tau)$ for $i = 1, 2, .., q$. The sampling period is denoted by τ, and $\hat{\omega}_i$ indicates the basis function frequencies obtained by perturbation theory, while q is the number of nonzero basis function frequencies.

4.2.2 Iterative Method of Identification

Substituting Eq.(4.14) into Eq.(4.13) yields the following ELM:

$$\begin{aligned} \mathbf{x}_{N+1}(k+1) &= \mathbf{A}\mathbf{x}_{N+1}(k) + \begin{bmatrix} \mathbf{B} & \mathbf{N}\boldsymbol{\Lambda} \end{bmatrix} \begin{pmatrix} u(k) \\ \boldsymbol{\Phi}(k)u(k) \end{pmatrix} \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}_{N+1}(k) + \begin{bmatrix} \mathbf{D} & \mathbf{0} \end{bmatrix} \begin{pmatrix} u(k) \\ \boldsymbol{\Phi}(k)u(k) \end{pmatrix}, \quad s \le k \le l \end{aligned} \tag{4.15}$$

where $\mathbf{0} \in \mathbb{R}^{r\times(2q+1)}$ and r is the number of outputs.

Such an ELM cannot be identified uniquely because, at least, the first two rows of its input vector are linearly dependent, as shown in Eq.(4.16)

$$\begin{pmatrix} u(k) \\ \boldsymbol{\Phi}(k)u(k) \end{pmatrix} = \begin{pmatrix} u(k) \\ \phi_0(k)u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k) \end{pmatrix} = \begin{pmatrix} u(k) \\ u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k) \end{pmatrix} \tag{4.16}$$

However, this problem can be fixed by compressing the columns of the system matrices that correspond to the repeated rows of the input vector into single columns as follows:

$$\begin{aligned} \mathbf{x}_{N+1}(k+1) &= \mathbf{A}\mathbf{x}_{N+1}(k) + \begin{bmatrix} \mathbf{B}+\mathbf{N}\boldsymbol{\Lambda}_1 & \mathbf{N}\boldsymbol{\Lambda}_2 \end{bmatrix} \begin{pmatrix} u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k) \end{pmatrix} \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}_{N+1}(k) + \begin{bmatrix} \mathbf{D} & \mathbf{0} \end{bmatrix} \begin{pmatrix} u(k) \\ \phi_1(k)u(k) \\ \vdots \\ \phi_{2q}(k)u(k) \end{pmatrix}, \quad s \le k \le l \end{aligned} \tag{4.17}$$

where $\mathbf{0} \in \mathbb{R}^{r\times(2q)}$, $\boldsymbol{\Lambda}_1 \in \mathbb{R}^n$ is the first column of the coefficient matrix $\boldsymbol{\Lambda}$, and $\boldsymbol{\Lambda}_2 \in \mathbb{R}^{n\times(2q)}$ is the matrix of the remaining columns so that $\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \boldsymbol{\Lambda}_2 \end{bmatrix}$, while n denotes the order of the bilinear model.

Eq.(4.17) can be written in a more compact form as:

$$\begin{aligned} \mathbf{x}_{N+1}(k+1) &= \mathbf{A}\mathbf{x}_{N+1}(k) + \begin{bmatrix} \mathbf{B}+\mathbf{N}\boldsymbol{\Lambda}_1 & \mathbf{N}\boldsymbol{\Lambda}_2 \end{bmatrix} \boldsymbol{\Phi}(k)u(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}_{N+1}(k) + \begin{bmatrix} \mathbf{D} & \mathbf{0} \end{bmatrix} \boldsymbol{\Phi}(k)u(k), \quad s \le k \le l \end{aligned} \tag{4.18}$$

To identify the ELM in Eq.(4.18), one needs to build the input vector time-history numerically and feed it, along with the recorded output-time histories, to a linear system identification tool. This tool should demonstrate robustness against noise (both measurement and process noise), as there will be a certain amount of numerical noise that depends on the value of N chosen in truncating the power series in Eq.(4.5), even if the record of the input/output time-histories contains no measurement noise.

In the bilinear system identification approach proposed here, a direct method of LTI system identification (DirectID), presented in Ref.[9], is preferred for its robustness against measurement and process noise. One of the steps of the DirectID algorithm requires a matrix $\mathbf{V}$, composed of only the past input and output data time-histories as shown in Eq.(4.19):

$$\mathbf{V} = \begin{pmatrix} \mathbf{y}(s) & \mathbf{y}(s+1) & \cdots & \mathbf{y}(l-p-1) \\ \mathbf{y}(s+1) & \mathbf{y}(s+2) & \cdots & \mathbf{y}(l-p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}(s+p-1) & \mathbf{y}(s+p) & \cdots & \mathbf{y}(l-2) \\ \boldsymbol{\Phi}(s)u(s) & \boldsymbol{\Phi}(s+1)u(s+1) & \cdots & \boldsymbol{\Phi}(l-p-1)u(l-p-1) \\ \boldsymbol{\Phi}(s+1)u(s+1) & \boldsymbol{\Phi}(s+2)u(s+2) & \cdots & \boldsymbol{\Phi}(l-p)u(l-p) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}(s+p)u(s+p) & \boldsymbol{\Phi}(s+p+1)u(s+p+1) & \cdots & \boldsymbol{\Phi}(l-1)u(l-1) \end{pmatrix} \tag{4.19}$$

where $rp \ge n$. The algorithm requires that this matrix is full-rank. As discussed previously, the matrix $\mathbf{V}$ in Eq.(4.19) contains time-shifted rows of $\boldsymbol{\Phi}(k)u(k)$, and thus, it is generally not full-rank, unless the excitation input $u(k)$ is enriched by a subtle amount of random component. The random component to be added to the excitation input $u(k)$ is chosen to be a highband random signal obtained by passing a randomly generated WGN signal through a high-pass filter. Since any physical system intrinsically behaves as a low-pass filter, the

change in the state of the bilinear system due to the addition of RC to the excitation input can be kept to a minimum by using a highband random component (HRC). Thus, the final form of the excitation input $u(k)$ becomes:

$$u(k) = \sum_{i=0}^{f} a_i \cos(\omega_i k\tau) + \sum_{i=1}^{f} b_i \sin(\omega_i k\tau) + \mathrm{hrc}(k) \tag{4.20}$$

where f represents the number of excitation input frequencies and $hrc(k)$ is the HRC used in the enrichment of $u(k)$. In order to reproduce the steady-portion of the state time-history of the ELM $\mathbf{x}_{N+1}(k)$, the numerically generated input time history, which was used in its identification, is fed to it. Any choice of the initial condition $\mathbf{x}_{N+1}(0)$ is acceptable for this task, since we are interested merely in the steady-state portion of the state time-history. At this point, it is worthwhile to remark that, although the identified ELM is based on the assumption that $\mathbf{x}(k) \approx \mathbf{x}_N(k)(s \leq k \leq l)$, the time-history of the reproduced state corresponds to $\mathbf{x}_{N+1}(k)$. The preceding statement may appear contradictory because the DirectID algorithm is an LTI system identification tool and LTI systems have the property that the frequencies contained by their inputs are also present in their states and outputs. However, in the case discussed above, the steady-state portion of the measured output time-history, obtained by exciting the original bilinear system with the excitation input $u(k)$, contains frequencies not contained by the numerically generated input $\mathbf{\Phi}(k)u(k)$, used in its identification. Thus, the input/output time-histories used in the identification of the ELM violates a property shared by all LTI systems, leading to an ambiguity on what the frequency content of the steady-state portion of its state time-history would be. It is observed that the DirectID algorithm yields an ELM whose steady-state portion of the reproduced state time-history contains only the frequencies present in the input time-history used in the identification of the ELM, in accordance with the previous reasoning which stated that the steady-state portion of the reproduced time-history of the identified ELM corresponds to $\mathbf{x}_{N+1}(k)$.

The identified state time-history corresponding to $\mathbf{x}_{N+1}(k)$ still contains numerical errors, since the output time-history, used in its identification, is obtained from the original bilinear system, thus, it contains frequencies which should not be present in $\mathbf{x}_{N+1}(k)$. However, the identified state time-history corresponding to $\mathbf{x}_{N+1}(k)$ is still a better approxima-

tion of the original state of the bilinear system $\mathbf{x}(k)$ than that corresponding to $\mathbf{x}_N(k)$, as it is a higher order approximation. Consequently, in the light of this discussion, an iterative process can be defined, where the identified state is improved at each iteration, using Eq.(4.13) as follows:

1. Identify the ELM in Eq.(4.18) acquired by performing an N^{th} order approximation to the steady-state portion of the state time-history using sine and cosine basis functions of frequencies obtained by perturbation theory.

2. Reproduce the state of the identified ELM by feeding it the input time-history used in its identification (i.e. $\mathbf{\Phi}(k)u(k)$). The reproduced state will be denoted by $\mathbf{z}_{N+1}(k)$ and the steady-portion of its state time-history corresponds to the $(N+1)th$ order approximation of the state, $\mathbf{x}_{N+1}(k)$.

3. Identify the ELM in Eq.(4.13), where N is replaced by $N+1$, using the steady-state portion of the time-history of $\mathbf{z}_{N+1}(k)$, obtained in the former step.

4. Reproduce the state of the identified ELM by feeding it the input time-history used in its identification. The reproduced state will be denoted by $\mathbf{z}_{N+2}(k)$ and the steady-portion of its state time-history corresponds to the $(N+2)th$ order approximation of the state, $\mathbf{x}_{N+2}(k)$.

5. Identify the ELM in Eq.(4.13), where N is replaced by $N+2$, using the steady-state portion of the time-history of $\mathbf{z}_{N+2}(k)$, obtained in the former step.

6. Repeat steps 4-5 by incrementing the approximation order by 1 at each iteration, until the change in the steady-state portions of the reproduced state time-histories in consecutive iterations is negligibly small.

Note that the identified ELMs in the aforementioned procedure do not have to be in the same coordinates. On the contrary, in general, they are linked to each other through linear coordinate transformations as follows:

$$\mathbf{z}_{N+1}(k) = \mathbf{T}_N \mathbf{x}_N(k), \quad s \leq k \leq l \tag{4.21}$$

where $\mathbf{T}_N$ is the coordinate transformation matrix relating $\mathbf{z}_{N+1}(k)$ to $\mathbf{x}_N(k)$. Similarly the steady-state portion of the identified state time-history that results at the end of the last iteration is linked to the state of the original bilinear system as follows:

$$\mathbf{z}_M(k) = \mathbf{T}_x\mathbf{x}(k), \quad s \le k \le l \tag{4.22}$$

where M denotes the highest order of approximation achieved in the last step of iteration, while $\mathbf{T}_x$ is the coordinate transformation matrix relating $\mathbf{z}_M(k)$ to $\mathbf{x}(k)$

Finally, after the last iteration, the bilinear system matrices can be retrieved using:

$$\begin{aligned} \begin{bmatrix} \mathbf{A}_M & \mathbf{B}_M & \mathbf{N}_M \end{bmatrix} &= \mathbf{Z}_M\mathbf{Q}_{M-1}^{+} \\ \begin{bmatrix} \mathbf{C}_M & \mathbf{D}_M \end{bmatrix} &= \mathbf{Y}\mathbf{R}_M^{+} \end{aligned} \tag{4.23}$$

where

$$\begin{aligned} \mathbf{Z}_M &= \begin{bmatrix} \mathbf{z}_M(s+1) & \mathbf{z}_M(s+2) & \cdots & \mathbf{z}_M(l) \end{bmatrix} \\ \mathbf{Q}_{M-1} &= \begin{bmatrix} \mathbf{z}_M(s) & \mathbf{z}_M(s+1) & \cdots & \mathbf{z}_M(l-1) \\ u(s) & u(s+1) & \cdots & u(l-1) \\ \mathbf{z}_{M-1}(s)u(s) & \mathbf{z}_{M-1}(s+1)u(s+1) & \cdots & \mathbf{z}_{M-1}(l-1)u(l-1) \end{bmatrix} \\ \mathbf{Y} &= \begin{bmatrix} \mathbf{y}(s+1) & \mathbf{y}(s+2) & \cdots & \mathbf{y}(l) \end{bmatrix} \\ \mathbf{R}_M &= \begin{bmatrix} \mathbf{z}_M(s+1) & \mathbf{z}_M(s+2) & \cdots & \mathbf{z}_M(l) \\ u(s+1) & u(s+2) & \cdots & u(l) \end{bmatrix} \end{aligned} \tag{4.24}$$

and the identified bilinear system matrices are related to the original bilinear system matrices through:

$$\begin{aligned} \mathbf{A} &= \mathbf{T}_x^{-1}\mathbf{A}_M\mathbf{T}_x \\ \mathbf{B} &= \mathbf{T}_x^{-1}\mathbf{B}_M \\ \mathbf{N} &= \mathbf{T}_x^{-1}\mathbf{N}_M\mathbf{T}_x \\ \mathbf{C} &= \mathbf{C}_M\mathbf{T}_x \\ \mathbf{D} &= \mathbf{D}_M \end{aligned} \tag{4.25}$$

4.3 Illustrative Numerical Example

Let's illustrate the proposed method through the analysis of a simple non-dimensional example. Consider a bilinear system of the form in Eq.(4.1), where

$$\mathbf{A} = 0.3 \quad \mathbf{B} = 0.2 \quad \mathbf{N} = 0.3 \quad \mathbf{C} = 1 \quad \mathbf{D} = 0 \tag{4.26}$$

the number of inputs is $m = 1$, the number of outputs is $r = 1$, and the order of the system considered is equal to $n = 1$. The minimum choice for the DirectID parameter p is 1, since $rp \geq n$ and such a choice allows us to keep, in this numerical example, the dimensions of the matrices small.

Let's choose the sampling period τ to be equal to 0.01 seconds for a set of input/output measurement time-history of length 2,000 time-steps. The steady-state portion of the response is assumed to start at time-step $s = 43$, since $\prod_{k=1}^{j} (\mathbf{A} + \mathbf{N}u(k)) < 1e - 20$ for $s \leq j \leq 20,000$. In an actual identification task some a priori knowledge of the system is necessary for such an assumption to be valid. The excitation input $u(k)$ is defined as below:

$$u(k) = 1.8541 \sin(30k\tau) \tag{4.27}$$

and it is not enriched by the addition of RC. Recall that the addition of RC is necessary to avoid ill-conditioning of matrices formed during application of DirectID algorithm. However, due to the simplicity of the above example, there is no need for enrichment of $u(k)$.

The N^{th} order approximation of the state $\mathbf{x}_N(k)$ is formed for $N = 1, 2, 3, 4, 5, 10$ using Eq.(4.9). Fig.4.1 contains the power spectral density (PSD) plots of the steady-state portions of the time-histories of the reconstructed state $\mathbf{x}_N(k)$ and that of the state $\mathbf{x}(k)$ of the bilinear system. The PSD plot of $\mathbf{x}(k)$ contains evenly spaced peaks of decreasing amplitudes. As discussed previously, the frequencies contained by the steady-state portion of the time-history of $\mathbf{x}(k)$ can be obtained using perturbation theory as the sums and differences of the input frequencies, which correspond to the integer multiples of the excitation input frequency (also referred to as the harmonic frequencies or harmonics). It can also be observed that the amplitudes of the peaks corresponding to higher frequencies are lower than those corresponding to low frequencies. This is due to the fact that bilinear

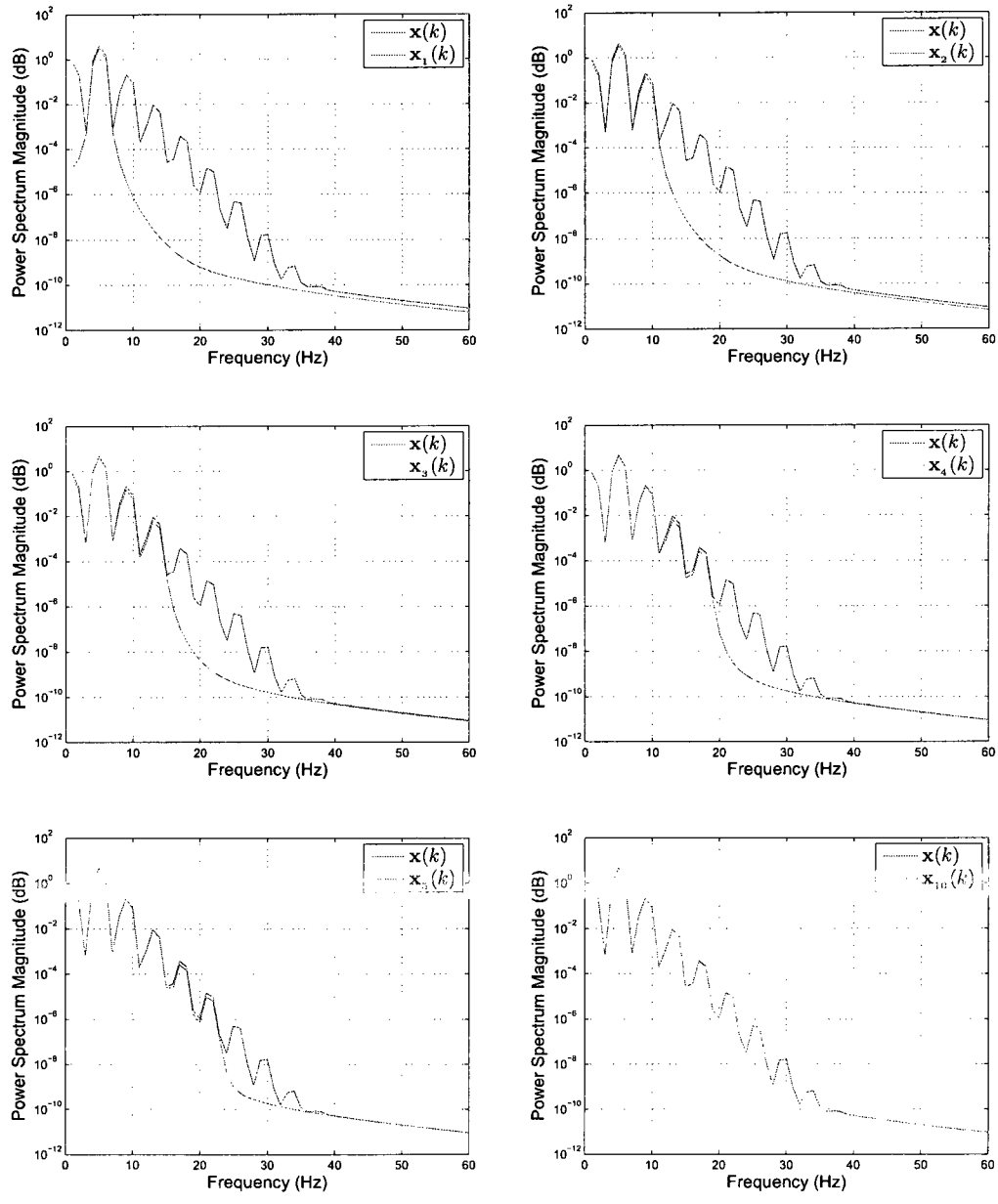


Figure 4.1: The PSD plots of the steady-state portions of the time-histories of $\mathbf{x}(k)$ and $\mathbf{x}_N(k)$ for $N = 1, 2, 3, 4, 5, 10$

systems, like other physical systems, act as low-pass filters. The PSD plot of $\mathbf{x}_1(k)$ contains a single peak that corresponds to the excitation input frequency, whereas $\mathbf{x}_2(k)$ contains 3 peaks: the one corresponding to the input frequency, double the input frequency and zero. Similarly, as the value of N increases by 1, one more peak is added to its PSD plot, which corresponds to the next harmonic.

The vector of basis functions $\mathbf{\Phi}(k)$ is formed for the perturbation order $N = 1$, and the input time-history for the ELM $\mathbf{\Phi}(k)u(k)$ is generated numerically. The choice of $N = 1$ implies that there will be no basis functions to capture the bilinear effects on the state due to the term $\mathbf{N}\mathbf{x}(k)u(k)$ in accordance with Eq.(2.5). Fig.4.2 presents the plots of the steady-state portions of the time-histories of the state $\mathbf{x}(k)$ and its representation $\hat{\mathbf{x}}(k)$ in terms of the basis functions. The amount of mismatch between them is calculated using $error(\mathbf{x}(k), \hat{\mathbf{x}}(k))$ to be 21.5%, which is large to start with, but will shrink through the iterative process.

The corresponding ELM is identified by inputing to the DirectID algorithm the steady-state portion of the output time-history $\mathbf{y}(k)$ and the generated input time-history $\mathbf{\Phi}(k)u(k)$. The identified ELM is fed $\mathbf{\Phi}(k)u(k)$ to reproduce the state of the ELM $\mathbf{z}_2(k)$. The iterative procedure described in Section 4.2.2 is carried out to obtain the steady-state portions of the time histories for $\mathbf{z}_3(k), \mathbf{z}_4(k), \cdots, \mathbf{z}_{40}(k)$. Fig.4.3 presents the PSD plots of the steady-state portions of the time-histories of the reproduced state $\mathbf{z}_i(k)(i = 2, 3, 4, 5)$, the state of the original bilinear system $\mathbf{x}(k)$, excitation input $u(k)$ and i^{th} order perturbation theory approximation of the state $\mathbf{x}_i(k)(i = 2, 3, 4, 5)$. It can be observed from these plots that the peaks of the PSD plots of $\mathbf{z}_i(k)$ are aligned with those of $\mathbf{x}_i(k)$, indicating that the reproduced states at the end of each iteration corresponds to an order of iteration.

Furthermore, the mismatch between the original output and the reproduced output $\mathbf{y}_{\mathbf{z}_i}(k)$ at each iteration can be quantified using $error(\mathbf{y}(k), \mathbf{y}_{\mathbf{z}_i}(k))$, where

$$\mathbf{y}_{\mathbf{z}_i}(k) = \mathbf{C}_i\mathbf{z}_i(k) + \mathbf{D}_i u(k) \tag{4.28}$$

Fig.4.4 shows the calculated value of $error(\mathbf{y}(k), \mathbf{y}_{\mathbf{z}_i}(k))$ against the approximation order i. It can be observed from this plot that the error drops almost exponentially to a small number below which the computational precision is lost, in about 30 iterations, and no

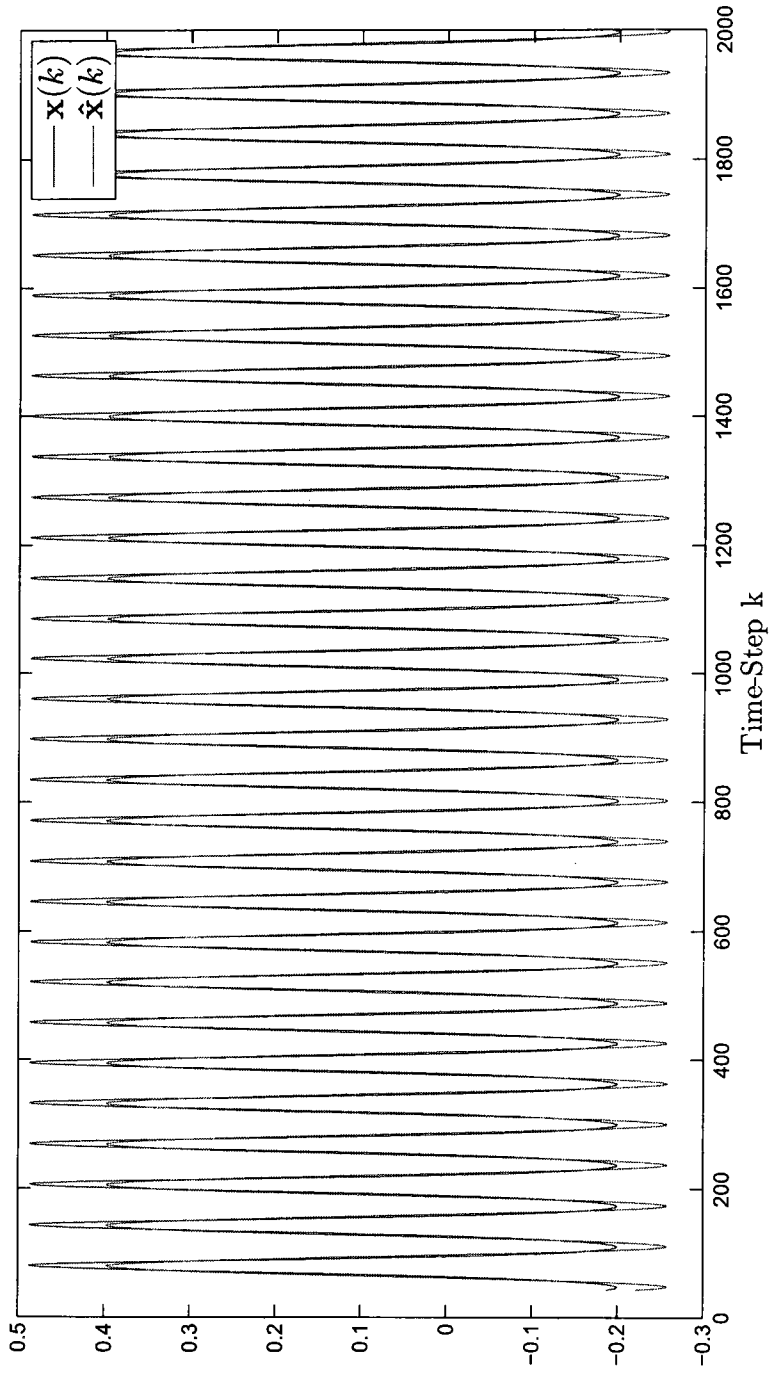


Figure 4.2: The plot of the steady-state portion of the state $\mathbf{x}(k)$ and its 1^{st} order approximation $\hat{\mathbf{x}}(k)$

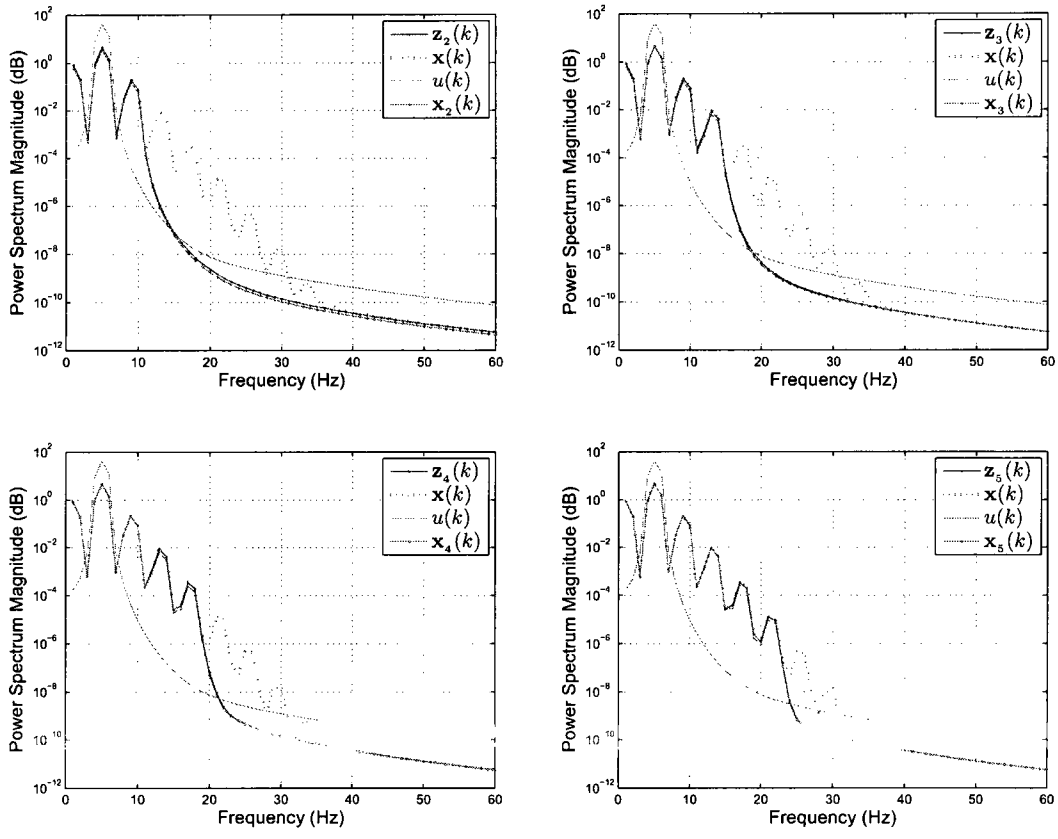


Figure 4.3: The PSD plots of the steady-state portions of the time-histories of the reproduced state $\mathbf{z}_i(k)(i = 2, 3, 4, 5)$, the state of the original bilinear system $\mathbf{x}(k)$, excitation input $u(k)$ and i^{th} order perturbation theory approximation of the state $\mathbf{x}_i(k)(i = 2, 3, 4, 5)$

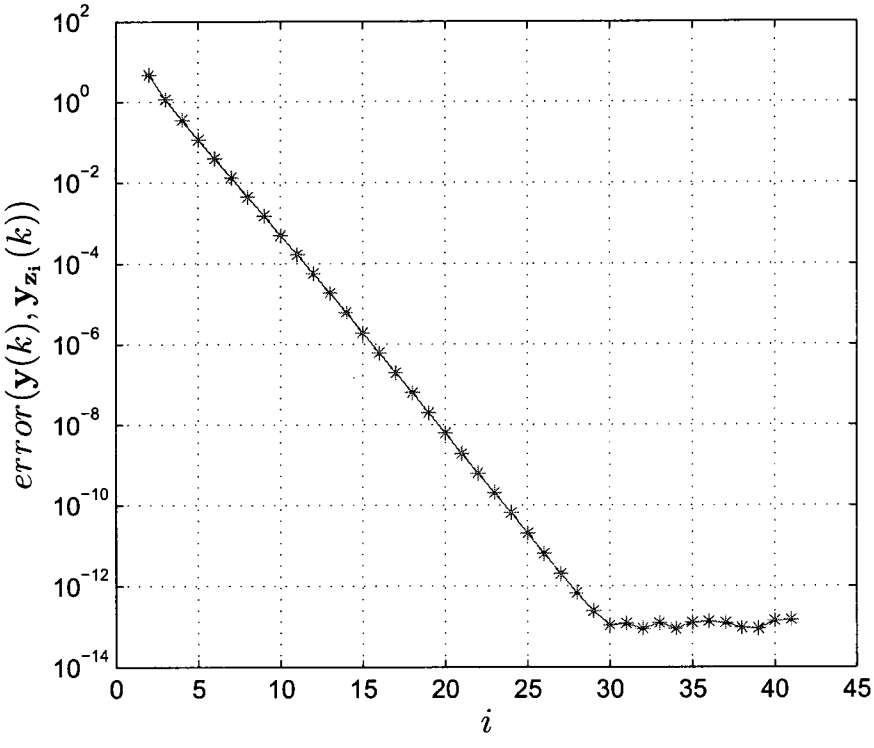


Figure 4.4: The plot of $error(\mathbf{y}(k), \mathbf{y}_{\mathbf{z}_i}(k))$ versus the approximation order i

further improvement is observed beyond it. The fluctuations beyond the 30^{th} iteration are due to the loss of computational precision.

The bilinear system matrices are identified using Eq.(4.23) to be:

$$\mathbf{A}_{30} = 0.3000 \quad \mathbf{B}_{30} = 0.2000 \quad \mathbf{N}_{30} = 0.3000 \quad \mathbf{C}_{30} = 1.0000 \quad \mathbf{D}_{30} = 0.0000 \tag{4.29}$$

which are exactly equal to the original bilinear system matrices.

4.3.1 Generic Numerical Example

Having illustrated the major concepts through a simple illustrative example, let's analyze a second order dimensionless example. Consider a bilinear system of the form in Eq.(4.1),

where

$$\mathbf{A} = \begin{bmatrix} 0.9924 & 0.0118 \\ -1.1778 & 0.8747 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.0126 \\ 0.0042 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} 0.2000 & 0 \\ 0 & 0.2000 \end{bmatrix}, \quad (4.30)$$
$$\mathbf{C} = \begin{bmatrix} 1 & 2 \end{bmatrix}, \mathbf{D} = 0.0000$$

Thus, the number of inputs and outputs remain equal to 1. Unlike the bilinear model used in the illustrative example, this model contains 2 states, which are coupled through the non-diagonal $\mathbf{A}$ matrix. The parameter p is increased to 2 in accordance with the requirement that $rp \geq n$.

Let's also keep the sampling period τ to be 0.01 seconds, while increasing the length of the set of input/output measurement time-history to $4,000$ time-steps. The steady-state portion of the response is assumed to start at time-step $s = 749$, since $\prod_{k=1}^{j} (\mathbf{A} + \mathbf{N}u(k)) < 1e{-}20$ for $s \leq j \leq 20,000$. The base excitation input $u(k)$ is constructed as in Eq.(4.2) using Table 2.6. However, it is enriched via the addition of a HRC obtained by passing a randomly generated WGN, whose RMS is 5% of the RMS of the base signal, through a highpass 10^{th} order Butterworth filter, which has a cut-off frequency equal to 45Hz. The cut-off frequency corresponds to 90% of the Nyquist frequency, which is 50Hz. This is the same filter whose Bode plot was presented in Fig.3.1. Moreover, Fig.4.5 shows the designed excitation input $u(k)$ before and after the addition of HRC. The addition of the HRC is necessary to eliminate the ill-conditioning of the matrix $\mathbf{V}$ in Eq.(4.19). Since the construction of $\mathbf{V}$ requires the a priori knowledge of $\mathbf{y}(k)$, the minimum amount of enrichment necessary cannot be precisely determined. However, the amount of HRC needed can be roughly estimated by slightly increasing the amount that would make the lower submatrix of $\mathbf{V}$, which only contains the time-shifted rows of $\mathbf{\Phi}(k)u(k)$, full-rank. In fact, the choice of 5% for the amount of enrichment yields a condition number of $1.3637e3$ for the lower submatrix of $\mathbf{V}$ and $1.4502e3$ for $\mathbf{V}$ itself, implying that both of them are full-rank. On the other hand, if $u(k)$ were not enriched by HRC, the condition number for the lower submatrix of $\mathbf{V}$ and $\mathbf{V}$ itself would be $1.1049e14$ and $1.1350e14$, respectively.

The vector of basis functions $\mathbf{\Phi}(k)$ is formed for the perturbation order $N = 2$, and the input time-history for the ELM $\mathbf{\Phi}(k)u(k)$ is generated numerically. The amount of

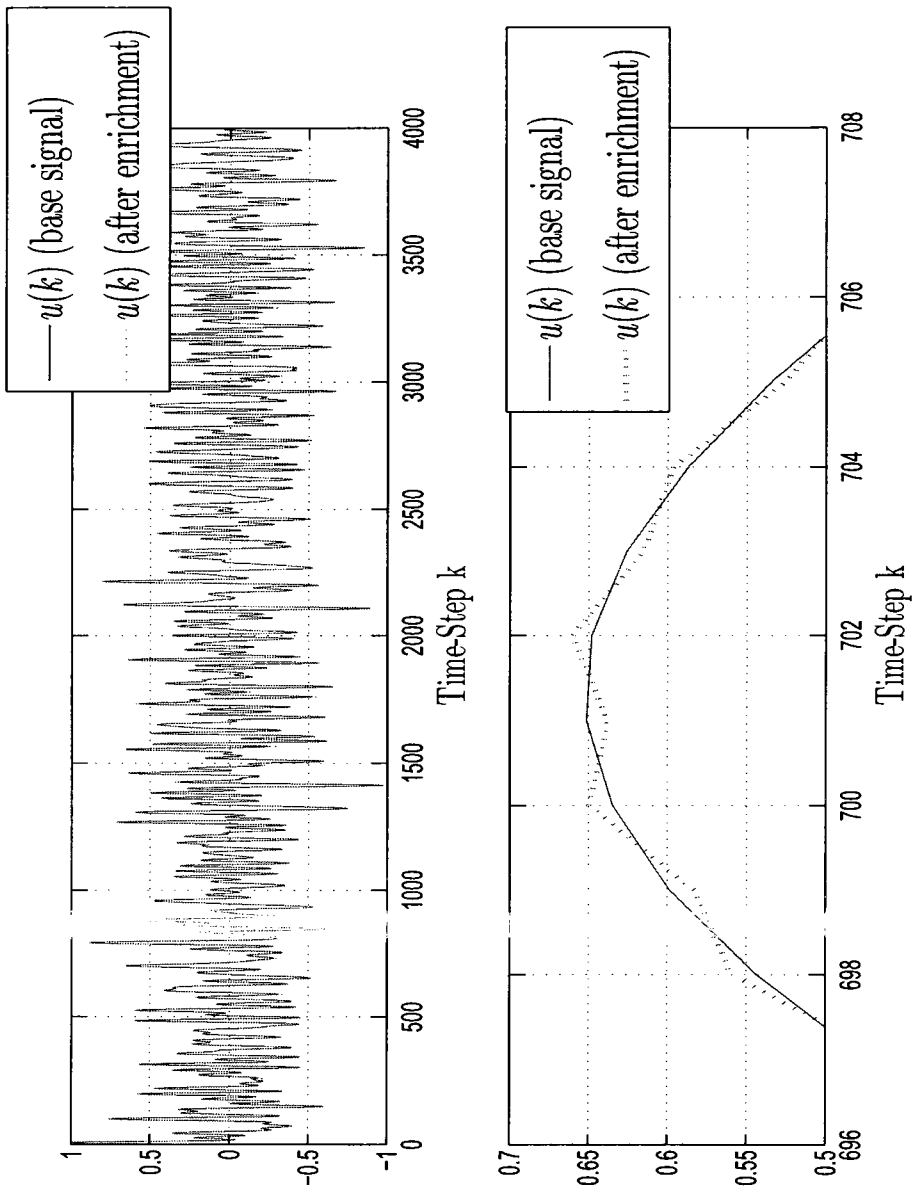


Figure 4.5: The plot of the time-history of excitation input $u(k)$ before and after enrichment by $hrc(k)$

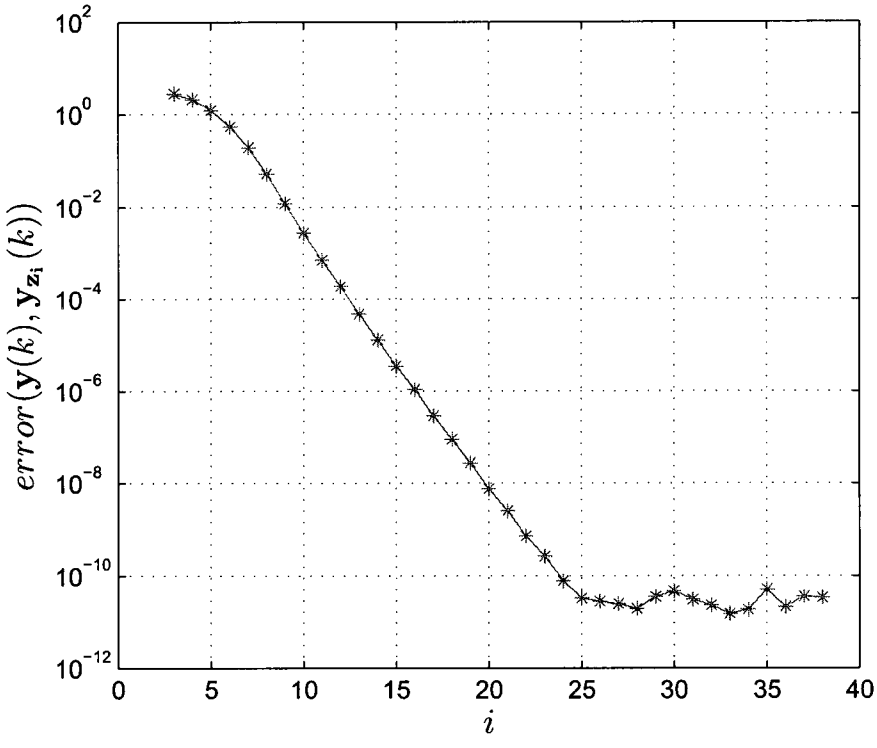


Figure 4.6: The plot of $error(\mathbf{y}(k), \mathbf{y}_{\mathbf{z}_i}(k))$ versus the approximation order i

mismatch between the steady-state portions of the time-histories of the state $\mathbf{x}(k)$ and its representation $\hat{\mathbf{x}}(k)$ in terms of the basis functions is calculated using $error(\mathbf{x}(k), \hat{\mathbf{x}}(k))$ to be 8.0914% and 10.0118%, for the first and second state, respectively.

The corresponding ELM is identified by inputing the steady-state portion of the output time-history $\mathbf{y}(k)$ and the generated input time-history $\mathbf{\Phi}(k)u(k)$ to the DirectID algorithm. Then, the iterative process described in Section 4.2.2 is carried out, and the state time-histories of the identified ELMs $\mathbf{z}_3(k), \mathbf{z}_4(k), \cdots, \mathbf{z}_{38}(k)$ and the time-histories of the outputs reproduced by them, $\mathbf{y}_{\mathbf{z}_3}(k), \mathbf{y}_{\mathbf{z}_4}(k), \cdots, \mathbf{y}_{\mathbf{z}_{38}}(k)$, are obtained. Fig.4.6 shows the amount of mismatch between $\mathbf{y}(k)$ and $\mathbf{y}_{\mathbf{z}_i}(k)(i = 3, 4, \cdots, 38)$, quantified by $error(\mathbf{y}(k), \mathbf{y}_{\mathbf{z}_i}(k))$. It is clear from the plot that the output reproduction error drops almost exponentially to a small number below which the computational precision is lost, in merely 26 iterations.

Needless to say, the ELMs obtained beyond the 26^{th} iteration are also equally valid. In fact, the identified bilinear system matrices can be extracted from the last ELM as follows:

$$\begin{aligned}
\mathbf{A}_{38} &= \begin{bmatrix} 1.3818e-14 & 1.0000 \\ -8.8195e-1 & 1.8671 \end{bmatrix} \\
\mathbf{B}_{38} &= \begin{bmatrix} 2.1000e-002 \\ -9.7793e-003 \end{bmatrix} \\
\mathbf{N}_{38} &= \begin{bmatrix} 2.0000e-1 & -1.9105e-14 \\ 8.9016e-14 & 2.0000e-1 \end{bmatrix} \\
\mathbf{C}_{38} &= \begin{bmatrix} 1 & 7.2642e-18 \end{bmatrix} \\
\mathbf{D}_{38} &= 0.0000
\end{aligned} \tag{4.31}$$

Since the identified bilinear system matrices $\mathbf{A}_{38}$ and $\mathbf{N}_{38}$ are related to the original bilinear system matrices $\mathbf{A}$ and $\mathbf{N}$ through the similarity transformation described in Eq.(4.25), their respective eigenvalues need to be the same. In fact, the eigenvalues of both $\mathbf{A}_{38}$ and $\mathbf{A}$ are both calculated to be identically equally to $0.09336 \mp i0.1022$, while the eigenvalues of both $\mathbf{N}_{38}$ and $\mathbf{N}$ are equal to 0.2000. The exact match between the identified and actual eigenvalues of the system matrices is indicative of remarkable performance of the proposed identification method.

As another measure of the quality of the identification, a new random input of length $20,000$ time-steps is generated from a Gaussian distribution with mean 0 and standard deviation 1. The generated test input $\mathbf{u}_t(k)$ is fed to both the original and the identified bilinear systems and the corresponding output time histories, $\mathbf{y}_t(k)$ and $\mathbf{y}_p(k)$ respectively, are obtained. Fig.4.7 shows the plot of the predicted output time-history $\mathbf{y}_p(k)$ against the actual output time-history $\mathbf{y}_t(k)$. The prediction error between the two is calculated by $error(\mathbf{y}_p(k), \mathbf{y}_t(k))$ as $7.1254e-11\%$. The smallness of the prediction error clearly illustrates that the predicted output $\mathbf{y}_p(k)$ is a very accurate estimate of the output $\mathbf{y}_t(k)$ of the bilinear system.

The perfect match between the eigenvalues of the identified and original system matrices, and the "almost" perfect prediction of the output produced by the test input confirm the success of the proposed identification algorithm.

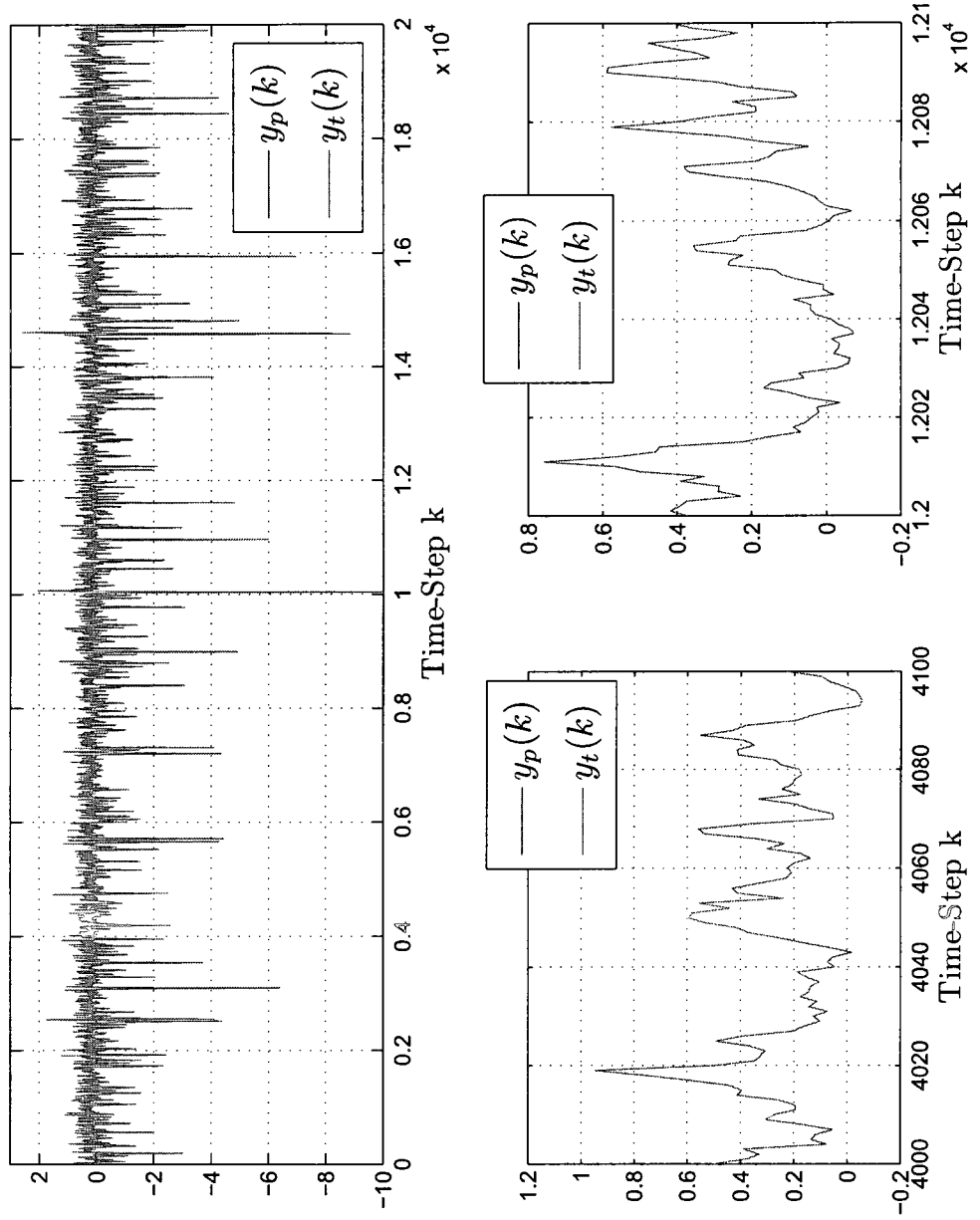


Figure 4.7: The predicted output time-history $\mathbf{y}_p(k)$ and the simulated test output time-history $\mathbf{y}_t(k)$

4.4 Conclusions

The proposed method in Chapter 2 first casts a bilinear model into an ELM by expressing the steady-state portion of its state in terms of sine and cosine basis functions for an N^{th} order perturbation theory approximation, then, identifies the ELM, and finally extracts the bilinear system matrices from the identified ELM. In this chapter, it is shown that the steady-state portion of state time-history reproduced by the identified ELM corresponds to the $(N+1)th$ order perturbation theory approximation. It uses this fact to generate a recursive algorithm where the order of the perturbation theory approximation is increased over each iteration until the steady-state portion of the state time-history is exactly reproduced. The steady-state portion of the resulting state time-history is, then, used to identify the bilinear system matrices exactly. The proposed method is distinguished from the existing methods in the literature most strikingly by requiring only a single set of input/output measurements and allowing for the design of the excitation input, while still yielding an exact identification.

Chapter 5

Conclusions and Future Research

5.1 Conclusions

This work presents an attempt at the identification of single-input, multi-output, discrete-time, state-space bilinear models. While the existing methods in the literature impose strict restrictions either by requiring a specific type of excitation input such as pulses of varying durations, white inputs, etc. or by entailing a multiple number of experiments, the proposed methods relax these requirements by allowing the design of the excitation input and by requiring only a single experiment.

The proposed methods focus on approximation of the bilinear model by an equivalent linear model (ELM), thereby converting a bilinear model identification task to a linear model identification task at the expense of introducing minor approximation errors. Such a conversion is accomplished by exciting the bilinear system with a linear combination of sine and cosine functions of user-selected frequencies and amplitudes, so that the steady-state portion of the state time-history of the bilinear system can be expressed as a linear combination of sine and cosine basis functions.

Perturbation theory is used to show that the frequencies of the aforementioned sine and cosine basis functions can be chosen efficiently as the sums and differences of the input frequencies to obtain a small state approximation error using a small number of basis functions. One of the basis functions generated in this fashion has to be a constant basis

function, unless the perturbation order used in approximating the state is 1 and the DC gain of the excitation input is equal to zero. Thus, the resulting ELM was shown to be non-unique due to the presence of 2 linearly dependent rows in the input vector, indicating that it could not be identified. This problem was remedied by compressing the linearly dependent rows of the input vector and the corresponding columns of the system matrices they multiply into single columns.

In the first proposed method, the resulting ELM is identified using an LTI system identification algorithm OKID. However, one of the steps of the OKID algorithm requires that a matrix containing time-shifted rows of input and output time-histories be inverted. The inversion of such a matrix cannot be accomplished, since in general the matrix is ill-conditioned as some of its rows are linearly dependent. In order to uncorrelate the linearly dependent rows of the aforementioned matrix, the user-designed excitation input is enriched by addition of a subtle amount of random component in the form of white Gaussian noise, thereby giving the excitation input to be fed to the bilinear system its final form. Although the exact amount of the random component necessary requires the knowledge of the output time-history, it was shown that the required amount can be predicted by slightly increasing the amount of random component sufficient to remove the ill-conditioning in the submatrix containing the time-shifted rows of the input-time history (i.e. excluding the rows containing the time-shifted rows of output time-history).

It was shown that the addition of the random component allows the identification of the ELM via OKID by enabling the aforementioned matrix inversion at the expense of increasing the state approximation error because the basis function frequencies selected by perturbation theory do not contain all of the frequencies present in the random component. However, since the amount of the added random component is subtle compared to the user-designed portion of the excitation input, the increase it causes on the state approximation error was shown to remain only minor. Moreover, it was depicted that if the user prefers to choose the excitation input frequencies randomly as opposed to evenly, the excitation input exhibits higher richness. Hence, a smaller amount of random component suffices to enable the inversion of the matrix. The ill-conditioning problem also reveals the importance of the usage of perturbation theory to keep the number of basis functions small due to the fact

that the higher the number of basis functions used, the more ill-conditioned the matrix to be inverted becomes.

Even though perturbation theory allows for picking the basis function frequencies efficiently, it was explained why the choice of perturbation order should still be performed carefully. On the one hand, the choice of a too small perturbation order does not allow an accurate approximation of the state in terms of the basis functions. On the other hand, if the perturbation order is chosen too high, it was shown to augment the ill-conditioning and, thus, require a higher amount of random component, which itself increases the state approximation error.

Apart from the choice of perturbation order, the effect of the choice of data length was also discussed. It was shown that even though the basis functions do not necessarily include all of the frequencies by present in the state, the lack of the missing basis functions can be compensated by the rest of the basis functions for short data lengths. On the other hand, it was also outlined that for longer data lengths, the ill-conditioning diminish.

Lastly, the effect of the choice of the OKID parameter p, which is proportional to the number of times the input and output time-histories are shifted to construct the matrix whose inversion is problematic, with regard to the ill-conditioning issue was evaluated. It was shown that for higher values of p, the condition number of the ill-conditioned matrix grows larger; therefore, it should be kept minimal, despite the fact that a higher value of p also enhances the robustness of the OKID algorithm against both process and measurement noise.

A set of recovery equations are presented to extract the bilinear system matrices from the identified ELM. The identified bilinear system matrices are related to the original bilinear system matrices through a linear coordinate transformation. The high quality of the identified bilinear model was attributed to the small difference between the output time-histories obtained by inputting a new, randomly generated test signal to the identified and original bilinear systems.

As an improvement to the first proposed method, the second proposed method focuses on curtailing the state approximation error using a different set of basis functions to represent

the steady-state portion of the state time-history more accurately, keeping the excitation input the same, except for the form of the added random component. For higher accuracy, in this method, the excitation input is enriched by a highband random component. It was shown that as the cutoff frequencies of the highband random component is increased the state approximation error decreases. Such a reduction is explained by the fact that all physical systems act as low-pass filters and the effect of the addition of highband random component on the output is attenuated by the bilinear system.

Along with the second proposed method, the concepts of input reduction, the process of eliminating the linearly dependent rows of the input vector, and input absorption, the process of eliminating the linearly dependent rows of the input vector on the basis functions, were also introduced. As in the first proposed method, the second proposed method initially yields an ELM using the sine and cosine basis functions to represent the steady-state portion of the state time-history and compressing the dependent rows of the input vector and the columns of the system matrices they multiply. However, rather than identifying the arising ELM via an LTI system identification algorithm, the second proposed method uses the interaction matrix formulation to depict that the ELM can be converted to an overparameterized ELM using the steady-state portions of the past input and output time-histories as the steady-state portion of the state time-history. In addition, the superiority of the data basis functions to the sine and cosine basis functions are shown. However, the resulting ELM is in general not unique, since some of the rows of its state time-history are linearly-dependent. After eliminating the redundant rows of the steady-state portion of the state time-history, a new, unique and overparameterized ELM is identified by performing a least-square fit, since the state-time history is already known. It is shown that the overparameterized ELM cannot be identified uniquely, unless the identified ELM is enriched by a random component.

In order to obtain a minimal order realization of the overparameterized ELM, its unobservable states are discarded. The resulting minimal order ELM is used to reproduce the steady-state portion of the state time-history, which can then be used to recover the bilinear system matrices, since the steady-state portion of the state time-history is shared by both the bilinear model and the minimal order ELM. The identified bilinear system matrices are

linked to the original bilinear system matrices through a linear coordinate transformation. As in the first method, the validity of identified bilinear model was illustrated by the fact that both the identified and original bilinear systems produce the "almost" exactly the same output, when both are fed a new, randomly generated test input signal.

The third proposed method is an exact, iterative method of bilinear system identification with an exponential speed of convergence. It uses the same excitation input as the second proposed method, which is obtained by enriching a linear combination of sine and cosine functions of user-selected frequencies and amplitudes with a highband random component. As in the first proposed method, the bilinear system is cast into an ELM by expressing the state in its bilinear term in terms of sine and cosine basis functions. The frequencies of the basis functions are derived for an N^{th} order perturbation theory approximation. A unique ELM is obtained by compressing the columns of the system matrices corresponding to the dependent rows of the input vector. The resulting ELM is identified using an LTI system identification algorithm DirectID, and the steady-state portion of its state time-history is reproduced. By rearranging the terms in the perturbation theory based state approximation equation, it was shown that the steady-state portion of the reproduced state time-history corresponds to the $(N+1)th$ order perturbation theory approximation. Moreover, the same equation is modified to generate an ELM whose state corresponds to the $(N+2)th$ order perturbation theory approximation to the state of the bilinear system and whose input vector contains the rows of the excitation input and the $(N+1)th$ order approximation to the state. The new ELM is identified using DirectID and its state reproduced to recover the $(N+2)th$ order approximation to the steady-state portion of the state time-history. The aforementioned process is repeated iteratively until the reproduced state time-history does not change significantly, marking the convergence of the iterations. Finally, it was explained that the bilinear system matrices could be recovered through a least square fit using the steady-state portion of the reproduced state time-history.

5.2 Future Research

Despite the fact that the proposed methods are capable of identifying bilinear systems rather accurately or even exactly. Future research should address the 2 main questions of how the performance of the methods can be enhanced and how the domain of problems the methods are applicable to can be expanded.

Firstly, all proposed methods make use of sine and cosine basis functions to convert the bilinear system identification task to a linear system identification task. However, the choice of sine and cosine basis functions is not unique, even though rather effective when the system is excited by a linear combination of sine and cosine functions. There may exist other pairs of excitation inputs and corresponding sets of basis functions that allow a more accurate approximation of the state for a specific number of basis functions. Further research should be carried out to identify such superior pairs.

The proposed methods require that the excitation input is enriched by a subtle amount of random component (RC), which contributes by a small amount to the approximation error of the state in terms of the sine and cosine basis functions whose frequencies are selected in accordance with perturbation theory, neglecting the frequencies introduced by incorporating the RC. Future research should investigate whether the need for RC can be eliminated. Through the relation of RC with the issue of ill-conditioning described in the thesis, such a research effort can return rewarding outcomes that can be used to expand the complexity of problems the proposed methods are applicable to.

The third proposed method paves the way for development of an exact, iterative identification algorithm. Although the sufficient set of conditions necessary to ensure the convergence of the method needs to be addressed in a future research effort, it was shown in this thesis that the state approximation error can be completely eliminated through a recursive algorithm. The significance of the method also stems from the fact that the increase in the state approximation error due to the addition of RC becomes trivial, as the recursive algorithm yields an exact state time-history anyway. The most outstanding aspects of the proposed method in comparison to the existing methods in the literature are that (i) it entails a single set of input/output time-histories, (ii) it allows for the design of the excitation

input, and (iii) it identifies the bilinear model exactly.

The fact that the proposed methods make use of only the steady-state portion of the measured data may hamper the identification of bilinear systems with a considerably long transient state time-history. Future research should be conducted to query whether a new formulation, such as the interaction matrix formulation, can be adapted to reformulate the problem so that the entire measurement time-history, including its transient portion, can be benefited from.

It should be noted that the proposed methods in this work are confined to single-input bilinear systems. Future research should be conducted on how the methods can be generalized to the case of multi-input bilinear systems or can be adapted to other nonlinear problems.

Bibliography

[1] B. Ho and R. Kalman, "Effective construction of linear state-variable models from input/output functions," *Regelungstechnik*, vol. 14, pp. 545–548, December 1966.

[2] J.-N. Juang, M. Phan, L. Horta, and R. Longman, "Identification of observer/kalman filter markov parameters: Theory and experiments," Tech. Rep. NAS 1.15:104069, National Aeronautics and Space Administration, Hampton, VA. Langley Research Center., United States, June 1991.

[3] J.-N. Juang, M. Phan, L. Horta, and R. Longman, "Identification of observer/kalman filter markov parameters: theory and experiments," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 2, pp. 320–9, 1993.

[4] J. N. Juang, *Applied System Identification.* Prentice Hall, Inc., Englewood Cliffs, New Jersey 07632, 1994.

[5] B. D. Moor, P. V. Overschee, and W. Favoreel, *Numerical Algorithms for Subspace State Space System Identification - An overview.* Boston, Massachusetts: Birkhauser, 1998.

[6] J.-N. Juang and M. Q. Phan, *Identification and Control of Mechanical Systems.* Cambridge University Press, New York, NY 10011-4211,, 2001.

[7] M. Phan, J.-N. Juang, and R. W. Longman, "On markov parameters in system identification," in *Proceedings of the International Modal Analysis Conference - IMAC*, vol. 2, (Florence, Italy), pp. 1415–1421, Proceedings of the 9th International Modal Analysis Conference Part 2 (of 2), Union Coll, Schenectady, NY, United States, April 1991.

[8] M. Q. Phan, L. G. Horta, J. N. Juang, and R. W. Longman, "Linear system identification via an asymptotically stable observer," *Journal of Optimization Theory and Applications*, vol. 79, pp. 59–86, October 1993.

[9] M. Q. Phan, J. A. Solbeck, and L. R. Ray, "A direct method for state-space model and observer/kalman filter gain identification," in *Collection of Technical Papers - AIAA Guidance, Navigation, and Control Conference*, vol. 5, (Providence, RI, United states), pp. 3317–3333, Collection of Technical Papers - AIAA Guidance, Navigation, and Control Conference, American Institute of Aeronautics and Astronautics Inc., August 2004.

[10] M. Q. Phan and R. W. Longman, "Extracting mass, stiffness, and damping matrices from identified state-space models," in *Collection of Technical Papers - AIAA Guidance, Navigation, and Control Conference*, vol. 5, (Providence, RI, United states), pp. 3334–3352, Collection of Technical Papers - AIAA Guidance, Navigation, and Control Conference, American Institute of Aeronautics and Astronautics Inc., August 2004.

[11] M. Verhaegen and P. Dewilde, "Subspace model identification - part: 1. the output-error state-space model identification class of algorithms," *International Journal of Control*, vol. 56, pp. 1187–210, November 1992.

[12] P. van Overschee and B. Moor, *Subspace Identification for Linear Systems: Theory - Implementation - Applications*. Springer, 1 ed., May 1996.

[13] W. Favoreel, B. De Moor, and P. Van Overschee, "Subspace state space system identification for industrial processes," in *IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS-5)*, vol. 1, (Corfu, Greece), pp. 319–27, Proceedings of 5th IFAC Symposium. Dynamics and Control of Process Systems 1998, Elsevier Science, June 1999.

[14] C. Bruni, G. DiPillo, and G. Koch, "Bilinear systems: An appealing class of nearly linear systems in theory and applications," *Ricerche Di Automatica*, vol. 19, no. 4, pp. 334–348, 1974.

[15] R. Mohler and W. Kolodziej, "An overview of bilinear system theory and applications," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-10, pp. 683–8, October 1980.

[16] D. L. Elliot, "Bilinear systems," in *Encyclopedia of Electrical Engineering* (J. Wiley and Sons, eds.), John Webster, 1999.

[17] R. Luesink and H. Nijmeijer, "On the stabilization of bilinear systems via constant feedback," *Linear Algebra and Its Applications*, vol. 122-124, pp. 457–74, November 1989.

[18] H. Dai and N. Sinha, "Robust identification of systems using block-pulse functions," *IEE Proceedings D: Control Theory and Applications*, vol. 139, pp. 308–316, May 1992.

[19] J. N. Juang, "Continuous-time bilinear system identification," Tech. Rep. NASA/TM-2003-212646; L-19004, NASA Langley Research Center, United States, September 2003.

[20] J.-N. Juang, "Continuous-time bilinear system identification," *Nonlinear Dynamics*, vol. 39, pp. 79–94, January 2005.

[21] E. Sontag, Y. Wang, and A. Megretski, "Input classes for identifiability of bilinear systems," *IEEE Transactions on Automatic Control*, vol. 54, pp. 195–207, February 2009.

[22] E. D. Sontag, Y. Wang, and A. Megretski, "Remarks on input classes for identification of bilinear systems," in *Proceedings of the American Control Conference*, (New York, NY, United States), pp. 4345–4350, 2007 American Control Conference, ACC, Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ 08855-1331, United States, July 2007.

[23] W. Rugh, *Nonlinear System Theory.* Baltimore, Maryland: John Hopkins University Press, 1981.

[24] G. Koch, "Volterra series expansion for stochastic bilinear systems," Tech. Rep. R.2-07, Univ. Rome, Italy, April 1972.

[25] V. Karanam, C. Hsu, and R. Mohler, "On the volterra series of bilinear systems. i," *Automatic Control Theory and Applications*, vol. 8, pp. 75–9, September 1980.

[26] A. Bertuzzi, C. Bruni, A. Gandolfi, and A. Germani, "Volterra series expansion for bilinear systems in a nonlinear feedback loop," *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, vol. 103, pp. 84–8, June 1981.

[27] G. Floquet, "Sur les quations diffrentielles linaires coefficients priodiques," *Annales Scientifiques de l'cole Normale Suprieure*, vol. 12, pp. 47–88, 1883.

[28] E. Reithmeier, *Periodic solutions of nonlinear dynamical systems : numerical computation, stability, bifurcation, and transition to chaos.* Springer-Verlag New York, Inc., 1991.

[29] N. Fameli, F. Curzon, and S. Mikoshiba, "Floquet's theorem and matrices for parametric oscillators: Theory and demonstrations," *American Journal of Physics*, vol. 67, pp. 127–32, February 1999.

[30] D. P.-W. A. Ilchmann, D. H. Owens, "Sufficient conditions for stability of linear time-varying systems," *Systems & Control Letters*, vol. 9, pp. 157–163, August 1987.

[31] S. T. Germain Garcia, Pedro L. D. Peres, "Assessing asymptotic stability of linear continuous time-varying systems by computing the envelope of all trajectories," *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, vol. 55, pp. 998–1003, April 2010.

[32] Y. Ho, "What constitutes a controllable system?," *Institute of Radio Engineers Transactions on Automatic Control*, vol. AC-7, p. 76, April 1962.

[33] E. Gilbert, "Controllability and observability in multivariable control systems," *Society of Industrial and Applied Mathematics - Journal on Control Series A*, vol. 1, no. 2, pp. 128–151, 1963.

[34] R. Kalman, "Mathematical description of linear dynamical systems," *Society of Industrial and Applied Mathematics - Journal on Control Series A*, vol. 1, no. 2, pp. 151–191, 1963.

[35] G. W. Haynes and H. Hemes, "Non-linear controllability via lie theory," *SIAM J. Contr.*, vol. 8, pp. 450–460, 1970.

[36] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. AC-22, pp. 728–40, October 1977.

[37] A. Kugi, K. Schlacher, and R. Novak, "Symbolic computation for the analysis and synthesis of nonlinear control systems," in *Software for Electrical Engineering Analysis and Design. Peter P Silvester Memorial Conference. Fourth International Conference on Software for Electrical Engineering Analysis and Design. ELECTROSOFT 99*, (Seville, Spain), pp. 255–64, Proceedings of ELECTROSOFT 99, May 1999.

[38] J. Pan and R. Wang, "Nonlinear observability in the structural dynamic identification," in *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 5765, (San Diego, CA, USA), pp. 1045–52, Smart Structures and Materials 2005: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems, SPIE-Int. Soc. Opt. Eng, March 2005.

[39] R. Rink and R. Mohler, "Complete controllable bilinear systems," *SIAM Journal on Control*, vol. 6, no. 3, pp. 477–86, 1968.

[40] A. Isidori, "Direct construction of minimal bilinear realizations from nonlinear input-output maps," *IEEE Transactions on Automatic Control*, vol. AC-18, pp. 626–31, December 1973.

[41] V. Jurdjevic and J. Quinn, "Controllability and stability," *Journal of Differential Equations*, vol. 28, pp. 381–9, June 1978.

[42] W. Boothby, "Some comments on positive orthant controllability of bilinear systems," *SIAM Journal on Control and Optimization*, vol. 20, pp. 634–44, September 1982.

[43] H. Nijmeijer, "Observability of autonomous discrete time non-linear systems: A geometric approach," *International Journal of Control*, vol. 36, pp. 867–874, November 1982.

[44] M. M. Espana and I. D. Landau, "Bilinear approximation of the distillation processes," *Ricerche di A utomatica*, vol. 6, pp. 20–46, July 1975.

[45] M. Espana and I. Landau, "Reduced order bilinear models for distillation-columns," *Automatica*, vol. 14, no. 4, pp. 345–355, 1978.

[46] D. Williamson, "Observation of bilinear systems with application to biological control," *Automatica*, vol. 13, pp. 243–54, May 1977.

[47] R. R. Mohler and P. A. Frick, "Bilinear demographic control processes," *Policy Analysis and Information Systems*, vol. 2, pp. 57–70, January 1979.

[48] R. Mohler and C. Barton, "Compartmental control model of the immune process," in *Proceedings of the 8th IFIP Conference on Optimization Techniques*, (Wurzburg, West Germany), pp. 421–30, Proceedings of the 8th IFIP Conference on Optimization Techniques, Springer-Verlag, September 1978.

[49] R. Mohler and C. Hsu, "T-b cell control processes in immunology," in *Proceedings of the 1978 IEEE Conference on Decision and Control Including the 17th Symposium on Adaptive Processes*, (San Diego, CA, USA), pp. 755–60, Proceedings of the 1978 IEEE Conference on Decision and Control Including the 17th Symposium on Adaptive Processes, IEEE, January 1979.

[50] M. Aoki, "Some examples of dynamic bilinear models in economics," in *Variable Structure Systems with Application to Economics and Biology* (A. Ruberti and R. R. Mohler, eds.), (New York), pp. 163–169, Springer, 1975.

[51] P. D'Alessandro, "Bilinearity and sensitivity in macroeconomy," Report review R74-02, University of Rome, Italy, January 1974.

[52] W. Favoreel, B. D. Moor, and P. V. Overschee, "Subspace identification of bilinear systems subject to white inputs," Tech. Rep. ESAT-SISTA/TR 1996-531, Katholieke Universiteit Leuven, Leuven, Belgium, 1996.

[53] W. Favoreel, B. De Moor, and P. Van Overschee, "A bilinear extension of subspace identification for systems subject to white inputs," in *Proceedings of the 1997 American*

Control Conference (Cat. No.97CH36041), vol. 1, (Albuquerque, NM, USA), pp. 607–11, Proceedings of 16th American CONTROL Conference, American Autom. Control Council, June 1997.

[54] W. Favoreel, B. De Moor, and P. Van Overschee, "Subspace identification of bilinear systems subject to white inputs," *IEEE Transactions on Automatic Control*, vol. 44, pp. 1157–65, June 1999.

[55] J. L. Walsh, "A closed set of normal orthogonal functions," *American Journal of Mathematics*, vol. 45, pp. 5–24, January 1923.

[56] N. J. Fine, "The generalized walsh function," *Transactions of the American Mathematical Society*, vol. 69, pp. 66–77, July 1950.

[57] C. Chen and C. Hsiao, "A state-space approach to walsh series solution of linear systems," *International Journal of Systems Science*, vol. 6, pp. 833–58, September 1975.

[58] W.-L. Chen and Y.-P. Shih, "International journal of control," *International Journal of Control*, vol. 27, pp. 917–32, June 1978.

[59] F. Lewis, V. Mertzios, G. Vachtsevanos, and M. Christodoulou, "Analysis of bilinear systems using walsh functions," *IEEE Transactions on Automatic Control*, vol. 35, pp. 119–123, January 1990.

[60] F. Lewis and D. Fountain, "Walsh function of analysis of linear and bilinear discrete time systems," *International Journal of Control*, vol. 53, pp. 847–53, April 1991.

[61] V. Karanam, P. Frick, and R. Mohler, "Bilinear system identification by walsh functions," *IEEE Transactions on Automatic Control*, vol. AC-23, pp. 709–13, August 1978.

[62] R. Bracewell, *The Hartley Transform.* New York: Oxford University Press, 1986.

[63] S. Daniel-Berbe and H. Unbehauen, "State space identification of bilinear continuous-time canonical systems via batch scheme hartley modulating functions approach," in *Proceedings of the 37th IEEE Conference on Decision and Control (Cat.*

No.98CH36171), vol. 4, (Tampa, FL, USA), pp. 4482–7, Proceedings of the 37th IEEE Conference on Decision and Control, IEEE, December 1998.

[64] S. Daniel-Berhe and H. Unbehauen, "Bilinear continuous-time systems identification via hartley-based modulating functions," *Automatica*, vol. 34, pp. 499–503, April 1998.

[65] J. A. Cherry, *Distortion analysis of weakly nonlinear filters using Volterra series.* PhD thesis, Carleton University, Canada, February 1995.

[66] C. Bruni, G. Di Pillo, and G. Koch, "On the mathematical models of bilinear systems," *Ricerche Di Automatica*, vol. 2, p. 1971, January 1971.

[67] R. Baheti, R. Mohler, and H. Spang III, "New cross correlation algorithm for volterra kernel estimation of bilinear systems," *IEEE Transactions on Automatic Control*, vol. AC-24, pp. 661–664, August 1979.

[68] M. Inagaki and R. Kamiya, "Bilinear system identification by estimated volterra kernels.," *Electrical Engineering in Japan (English translation of Denki Gakkai Ronbunshi)*, vol. 101, pp. 110–116, May-June 1981.

[69] M. Inagaki and H. Mochizuki, "Bilinear system identification by volterra kernels estimation," *IEEE Transactions on Automatic Control*, vol. AC-29, pp. 746–749, August 1984.

[70] R. G. Kvaternik and W. A. Silva, "Computational procedure for identifying bilinear representations of nonlinear systems using volterra kernels," Tech. Rep. NASA/TM-2008-215320; L-19461, NASA Langley Research Center, United States, June 2008.

[71] F. Fnaiech and L. Ljung, "Recursive identification of bilinear systems," *International Journal of Control*, vol. 45, pp. 435–70, February 1987.

[72] J. Ralston and B. Boashash, "Identification of bilinear systems using bandlimited regression," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (Cat. No.97CB36052)*, vol. 5, (Munich, Germany), pp. 3925–8, IEEE International Conference on Acoustics, Speech, and Signal Processin, IEEE Comput. Soc. Press, 1997.

[73] S. Meddeb, J. Y. Tourneret, and F. Castanie, "Unbiased parameter estimation for the identification of bilinear systems," in *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (Cat. No.00TH8496)*, (Pennsylvania, PA, USA), pp. 176–180, Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing, IEEE, Los Alamitos, CA, United States, August 2000.

[74] N. Kalouptsidis, P. Koukoulas, and V. Mathews, "Blind identification of bilinear systems," *IEEE Transactions on Signal Processing*, vol. 51, pp. 484–99, February 2003.

[75] J.-N. Juang and R. Pappa, "An eigensystem realization algorithm for modal parameter identification and model reduction," *Journal of Guidance, Control, and Dynamics*, vol. 8, pp. 620–7, September-October 1985.

[76] J.-N. Juang, J. Cooper, and J. Wright, "Eigensystem realization algorithm using data correlations (era/dc) for modal parameter identification," *Control, theory and advanced technology*, vol. 4, pp. 5–14, March 1988.