Model Selection, Identification and Robust Control for Dynamical Systems

Thesis by

Ka-Veng Yuen

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2002

(Defended April 17, 2002)

CALIFORNIA INSTITUTE OF TECHNOLOGY

EARTHQUAKE ENGINEERING RESEARCH LABORATORY

MODEL SELECTION IDENTIFICATION AND ROBUST CONTROL FOR DYNAMICAL SYSTEMS

BY

KA-VENG YUEN

REPORT NO. EERL 2002-03

PASADENA, CALIFORNIA

APRIL 2002

Acknowledgements

I would like to express my sincere gratitude and appreciation to my advisor, Prof. James L. Beck, for his enthusiastic guidance and his support throughout the course of my research. His continuous availability along with many open-minded discussions, are greatly appreciated. I would also like to thank my committee members, Prof. Joel W. Burdick, Prof. John F. Hall, Prof. Wilfred D. Iwan and Prof. Erik A. Johnson, for their valuable comments and suggestions.

My studies are supported by the Caltech teaching assistantship, the Harold Hellwig Fellowship and the George W. Housner Fellowship. These generous supports are gratefully acknowledged.

Special thanks are due to Prof. Lambros S. Katafygiotis, Prof. Costas Papadimitriou and Prof. Guruswami Ravichandran for their continuous encouragement and support.

My sincere appreciation goes to Ivan Au, Paul Lam, Swami Krishnan, Arash Yavari, Mortada Mehyar, Fuling Yang, Victor Kam, Lawrence Cheung and Donald Sze for their warm friendships, which have made my graduate life colorful and lively.

Finally, I am deeply grateful to my parents and my beloved fiancee, Vanessa Cheong, to whom this thesis is dedicated to, for their love and patience in the pursuit of my academic goals.

Abstract

To fully exploit new technologies for response mitigation and structural health monitoring, improved system identification and controller design methodologies are desirable that explicitly treat all the inherent uncertainties. In this thesis, a probabilistic framework is presented for model selection, identification and robust control of smart structural systems under dynamical loads, such as those induced by wind or earthquakes. First, a probabilistic based approach is introduced for selecting the most plausible class of models for a dynamical system using its response measurements. The proposed approach allows for quantitatively comparing the plausibility of different classes of models among a specified set of classes.

Then, two probabilistic identification techniques are presented. The first one is for modal identification using nonstationary response measurements and the second one is for updating nonlinear models using incomplete noisy measurements only. These methods allow for updating of the uncertainties associated with the values of the parameters controlling the dynamic behavior of the structure by using noisy response measurements only. The probabilistic framework is very well-suited for solving this nonunique problem and the updated probabilistic description of the system can be used to design a robust controller of the system. It can also be used for structural health monitoring.

Finally, a reliability-based stochastic robust control approach is used to design the controller for an active control system. Feedback of the incomplete response at earlier time steps is used, without any state estimation. The optimal controller is chosen by minimizing the robust failure probability over a set of possible models for the system. Here, failure means excessive levels of one or more response quantities representative of the performance of the structure and the control devices. When calculating the robust failure probability, the plausibility of each model as a representation of the system's dynamic behavior is quantified by a probability distribution over the set of

possible models; this distribution is initially based on engineering judgement, but it can be updated using the aforementioned system identification approaches if dynamic data become available from the structure. Examples are presented to illustrate the proposed controller design procedure, which includes the procedure of model selection, identification and robust control for smart structures.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

The goal of this work is to develop a complete probabilistic procedure for robust controller design for smart structures that treats all the inherent uncertainties, and includes new system identification techniques that allow the robust controller design to be improved if dynamic data from a structure is available.

1.1 System Identification

The problem of system identification of structural or mechanical systems using dynamic data has received much attention over the years because of its importance in response prediction, control and health monitoring (Natke and Yao 1988; Housner et al. 1997; Ghanem and Sture (Eds.) 2000).

The first question is: Which structural model class should be used for identification? In practice, it is not possible to use directly the finite-element model from the structural drawing for identification because there are too many uncertain parameters, which will lead to an unidentifiable case (Beck and Katafygiotis 1998; Katafygiotis and Beck 1998). However, the problem of model class selection has not been well explored in system identification. It is obvious that a more complicated model can 'fit' the data better than a less complicated one which has fewer adjustable (uncertain) parameters. Therefore, if the optimal model class is chosen by minimizing some measure of the error between the data and the corresponding predictions of the optimal model in each class, the optimal model class will always be the most complicated one. For example, in modal identification, using a 20-mode model would always be better than using a 10-mode model because the former one would fit the data better, although the improvement might be negligible. This approach therefore leads to over-fitting the data. When an over-fitted model is used for future prediction, it will very likely lead to poor results because the model depends too much on the details of

the data and the noise in the data might have an important role in the data fitting. Therefore, in model class selection, it is necessary to penalize a complicated model.

This was recognized early on by Jeffreys (1961) who did pioneering work on the application of Bayesian methods. He pointed out the need for a quantitative expression of the very old philosophy of 'Ockham's razor' which in this context implies that simpler models are more preferable than unnecessarily complicated ones, that is, the selected class of models should accurately describe the behavior of the system but be as simple as possible. Box and Jenkins (1970) also emphasize the same principle when they refer to the need for parsimonious models in time-series forecasting, although they do not give a quantitative expression of their principle of parsimony. Akaike (1974) recognized that maximum likelihood estimation is insufficient for model order selection in time-series forecasting using ARMA models and came up with another term to be added to the logarithm of the likelihood function that penalizes against parameterization of the models. This was later modified by Akaike (1976) and by Schwarz (1978).

In recent years, there has been a re-appreciation of the work of Jeffreys (1961) on the application of Bayesian methods, especially due to the expository publications of Jaynes (1983). In particular, the Bayesian approach to model selection has been further developed by showing that the *evidence* for each model class provided by the data, i.e., the probability of getting the data based on the whole model class, automatically enforces a quantitative expression of a principle of model parsimony or of Ockham's razor (Gull 1988; Mackay 1992; Sivia 1996). There is no need to introduce ad-hoc penalty terms as done in some of the earlier work on model selection.

In Chapter 2, the Bayesian approach is expounded and applied to select the most plausible class of dynamic models representing a structure from within some specified set of model classes by using its response measurements. The model class selection procedure is explained in detail. Examples are presented using a single-degree-of-freedom bilinear hysteretic system, a linear two-story frame and a linear ten-story shear building, all of which are subjected to seismic excitation.

Chapter 3 is devoted to the modal identification using nonstationary response

measurements. Much attention has been devoted to the identification of modal parameters of linear systems without measuring the input time history, such as in the case of ambient vibrations. In an ambient vibration survey, the naturally occurring vibrations of the structure (due to wind, traffic, micro-tremors, etc.) are measured and then a system identification technique is used to identify the small-amplitude modal frequencies and modeshapes of the lower modes of the structure. The assumption usually made is that the input excitation is a broadband stochastic process adequately modeled by stationary white noise. Many time-domain methods have been developed to tackle this problem. One example is the random decrement technique (Asmussen et al. 1997) which is based on curve-fitting of the estimated random decrement functions corresponding to various triggering conditions. Several methods are based on fitting directly the correlation functions using least-squares type of approaches (Beck et al. 1994). Different ARMA based methods have been proposed, e.g., Gersch and Foutch (1974); Gersch et al. (1976); Pi and Mickleborough (1989); and Andersen and Kirkergaard (1998). Methods based on the extended Kalman filter method have been proposed to estimate dynamic properties such as natural frequencies, modal damping coefficients and participation factors, of a linear multiple-degree-of-freedom (MDOF) system (Gersch and Foutch 1974; Beck 1978; Hoshiya and Saito 1984; Quek et al. 1999; Shi et al. 2000).

A common assumption in modal identification using response measurements only is that the responses are stationary. However, there are many cases where the response measurements are better modeled as nonstationary, e.g., a series of wind gusts or in the case of measured seismic response. In the literature, there are very few approaches which tackle modal identification using nonstationary response data, e.g., Safak (1989); Sato and Takei (1997). These methods rely on a forgetting factor formulation, which has been demonstrated to be difficult to choose. A bad choice of this forgetting factor will lead to poor results.

The results of system identification studies are usually restricted to the "optimal" estimates of the model parameters, whereas there is additional information related to the uncertainty associated with these estimates which is very important. For ex-

ample, how precisely are the values of the individual parameters pinned down by the measurements made on the system? Probability distributions may be used to describe this uncertainty quantitatively and so avoid misleading results (Beck and Katafygiotis 1998). Also, if the identification results are used for damage detection, this probability distribution for the identified model parameters may be used to compute the probability of damage (Vanik et al. 2000).

A Bayesian probabilistic system identification framework has been presented for the case of measured input (Beck and Katafygiotis 1998). In Chapter 3, a Bayesian time-domain approach is presented for the general case of linear MDOF systems using nonstationary response measurements. The proposed approach allows for the direct calculation of the probability density function (PDF) of the modal parameters which can be then approximated by an appropriately selected multi-variate Gaussian distribution. The importance of considering the response to be nonstationary is also discussed.

System identification using linear models is appropriate for the small-amplitude ambient vibrations of a structure that are continuously occurring. There is, however, a number of cases in recent years where the strong-motion response of a structure has been recorded but not the corresponding seismic excitation. In some cases this is because of inadequate instrumentation of the structure and in other cases it is because the free-field or base sensors malfunctioned during the earthquake. For example, the seismic response was recorded in several steel-frame buildings in Los Angeles which were damaged by the 1994 Northridge earthquake, but analysis of these important records has been hampered by the fact that the input (base motions) were not recorded and also because of the strong nonlinear response.

A literature search reveals relatively few papers that deal with system identification using nonlinear models (Hoshiya and Saito 1984; Loh and Tsaur 1988; Peng and Iwan 1992; Loh and Chung 1993; Roberts et al. 1995; Zeldin and Spanos 1998). In Chapter 4 this subject is tackled using a stochastic model for the uncertain input and a Bayesian probabilistic approach to quantify the uncertainties in the model parameters. This Bayesian probabilistic system identification framework is an exten-

sion of the case of measured input (Beck and Katafygiotis 1998; Katafygiotis et al. 1998). The proposed spectral-based approach utilizes important statistical properties of the Fast Fourier Transform (FFT) and their robustness with respect to the probability distribution of the response signal, e.g., regardless of the stochastic model for this signal, its FFT is approximately Gaussian distributed. The method allows for the direct calculation of the probability density function (PDF) for the parameters of a nonlinear model conditional on the measured response. The formulation is first presented for single-degree-of-freedom (SDOF) systems and then for multiple-degree-of-freedom systems. Examples using simulated data for a Duffing oscillator, an elasto-plastic system and a four-story yielding structure are presented to illustrate the proposed approach.

1.2 Structural Control

Because complete information about a dynamical system and its environment are never available, system and excitation parameters can not be determined exactly but can be given probabilistic descriptions which give a measure of how plausible the possible parameter values are (Cox 1961; Beck 1996; Beck and Katafygiotis 1998). Classical control methods based on a single nominal model of the system may fail to create a controller which can provide satisfactory performance for the system. Robust control methods, e.g., $\mathcal{H}_2$, $\mathcal{H}_\infty$ and μ-synthesis, etc., were therefore proposed so that the optimal controller can provide robust performance and stability for a set of 'possible' models of the system (Doyle et al. 1989; Doyle et al. 1992; Paganini 1996; Zhou and Doyle 1996; Johnson et al. 1998). In the proposed probabilistic robust control approach, an additional "dimension" is introduced by using probabilistic descriptions of all the possible models when selecting the controller to achieve optimal performance; these probability distributions are obtained from engineering judgement or system identification techniques. Specifically, a more probable model is given a high weighting for calculating the optimal gains, which is in contrast to standard robust control algorithms which give equal weighting to all possible models.

Over the last decade, there has been increasing interest in probabilistic, or stochastic, robust control theory. Monte Carlo simulations methods were used to synthesize and analyze control systems for uncertain systems (Stengel and Ray 1991; Marrison and Stengel 1995). In Spencer and Kaspari (1994); Spencer et al. (1994); Field et al. (1994); and Field et al. (1996), first- and second-order reliability methods were incorporated to compute the probable performance of linear-quadratic-regulator controllers (LQR). On the other hand, an efficient asymptotic expansion (Papadimitriou et al. 1997a) was used to approximate the probability integrals that are needed to determine the optimal parameters for a passive tuned mass damper (Papadimitriou et al. 1997b) and the optimal gains for an active mass driver (May and Beck 1998) for robust structural control. In May and Beck (1998), the proposed controller feeds back output measurements at the current time only, where the output corresponds to certain response quantities that need not be the full state vector of the system. However, there is additional information from past output measurements which may improve the performance of the control system.

In Chapter 5, the reliability-based methodology proposed in May and Beck (1998) is extended to allow feed back of the output (partial state) measurements at previous time steps. It is noted that in traditional linear-quadratic-Gaussian (LQG) control with partial state measurements, the optimal controller can be achieved by estimating the full state using a Kalman filter combined with the optimal LQG controller for full state feedback. However, in our case the separation principle does not apply and no state estimation is needed. The method presented for reliability-based robust control design may be applied to any system represented by linear state-space models but the focus here is on robust control of structures (Soong 1990; Housner et al. 1997; Caughey (Ed.) 1998).

In Chapter 5, the augmented vector formulation is presented for treating the output history feedback. Then, the statistical properties of the response quantities are calculated using the Lyapunov equation in discrete form. The robust control method is introduced which is based on choosing the feedback gains to minimize the robust failure probability (Papadimitriou et al. 2001). Examples using a shear building

model and a benchmark structure are given to illustrate the proposed approach.

1.3 Overview of this Thesis

1. Chapter 2 introduces a probabilistic approach for selecting the most plausible class of models for a structure using dynamic data.

2. Chapters 3 and 4 introduce two identification techniques for linear systems using nonstationary response measurements and for nonlinear systems with uncertain input.

3. Chapter 5 introduces a stochastic robust control methodology, with consideration of modeling uncertainty, structure-actuator interaction and time delay of the controller.

4. Chapter 6 illustrates the proposed robust controller design framework using a 20-DOF four-story structural frame.

5. Chapter 7 concludes this thesis and indicates possible future work.

Chapter 2 Model Selection

2.1 Overview

A Bayesian probabilistic approach is presented for selecting the most plausible class of models for a structure within some specified set of model classes, based on structural response data. The crux of the approach is to rank the classes of structural models based on their probabilities conditional on the response data which can be calculated based on Bayes' Theorem and an asymptotic expansion for the evidence for each model class. The approach provides a quantitative expression of a principle of model parsimony or of Ockham's razor which in this context can be stated as simpler models are to be preferred over unnecessarily complicated ones. Examples are presented to illustrate the method using a single-degree-of-freedom bilinear hysteretic system, a linear two-story frame and a ten-story shear building, all of which are subjected to seismic excitation.

2.2 Model Class Selection

Let $\mathcal{D}$ denote the input-output or output-only dynamical data from a structure. The goal is to use $\mathcal{D}$ to select the most plausible class of models representing the structure out of N_M given classes of models $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_{N_M}$. Since probability may be interpreted as a measure of plausibility based on specified information (Cox 1961), the probability of a class of models conditional on the set of dynamic data $\mathcal{D}$ is required. This can be obtained by using Bayes' Theorem as follows:

$$P(\mathcal{M}_j|\mathcal{D},\mathcal{U}) = \frac{p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})P(\mathcal{M}_j|\mathcal{U})}{p(\mathcal{D}|\mathcal{U})}, \quad j = 1, 2, \ldots, N_M \tag{2.1}$$

where $p(\mathcal{D}|\mathcal{U}) = \sum_{j=1}^{N_M} p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})P(\mathcal{M}_j|\mathcal{U})$ by the theorem of total probability and $\mathcal{U}$ expresses the user's judgement on the initial plausibility of the model classes, expressed as a prior probability $P(\mathcal{M}_j|\mathcal{U})$ on the model classes $\mathcal{M}_j, j = 1, ..., N_M$, where $\sum_{j=1}^{N_M} P(\mathcal{M}_j|\mathcal{U}) = 1$. The factor $p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})$ is called the *evidence* for the model class $\mathcal{M}_j$ provided by the data $\mathcal{D}$. Note that $\mathcal{U}$ is irrelevant in $p(\mathcal{D}|\mathcal{M}_j,\mathcal{U})$ and so it can be dropped in the notation because it is assumed that $\mathcal{M}_j$ alone specifies the probability density function (PDF) for the data, that is, it specifies not only a class of deterministic structural models but also the probability descriptions for the prediction error and initial plausibility for each model in the class $\mathcal{M}_j$ (Beck and Katafygiotis 1998). Eqn. 2.1 shows that the most plausible model class is the one that maximizes $p(\mathcal{D}|\mathcal{M}_j)P(\mathcal{M}_j|\mathcal{U})$ with respect to j.

Note that $P(\mathcal{M}_j|\mathcal{D},\mathcal{U})$ can be used not only for selection of the most probable class of models, but also for response prediction based on all the model classes. Let u denote a quantity to be predicted, e.g., first story drift. Then, the PDF of u given the data $\mathcal{D}$ can be calculated from the theorem of total probability as follows: $p(u|\mathcal{D},\mathcal{U}) = \sum_{j=1}^{N_M} p(u|\mathcal{D},\mathcal{M}_j)P(\mathcal{M}_j|\mathcal{D},\mathcal{U})$, rather than just using only the best model for prediction. However, if $P(\mathcal{M}_{best}|\mathcal{D},\mathcal{U})$ for the best model class is much larger than others, then the above expression is approximated by $p(u|\mathcal{D},\mathcal{U}) = p(u|\mathcal{D},\mathcal{M}_{best})$ and it is sufficient to just use the best model class.

The evidence for $\mathcal{M}_j$ provided by the data $\mathcal{D}$ is given by the theorem of total probability:

$$p(\mathcal{D}|\mathcal{M}_j) = \int_{\boldsymbol{\Theta}_j} p(\mathcal{D}|\boldsymbol{\theta}_j,\mathcal{M}_j)p(\boldsymbol{\theta}_j|\mathcal{M}_j)d\boldsymbol{\theta}_j, \quad j = 1, 2, \ldots, N_M \tag{2.2}$$

where $\boldsymbol{\theta}_j$ is the parameter vector in a parameter space $\boldsymbol{\Theta}_j \subset \mathbb{R}^{N_j}$ that defines each model in $\mathcal{M}_j$, the prior PDF $p(\boldsymbol{\theta}_j|\mathcal{M}_j)$ is specified by the user and the likelihood $p(\mathcal{D}|\mathcal{M}_j,\boldsymbol{\theta}_j)$ is calculated using the methods introduced in Section 2.3, Chapter 3 and Chapter 4.

In *globally identifiable* cases (Beck and Katafygiotis 1998), the updated (posterior) PDF for $\boldsymbol{\theta}_j$ given a large amount of data $\mathcal{D}$ may be approximated accurately by a

Gaussian distribution, so $p(\mathcal{D}|\mathcal{M}_j)$ can be approximated by using Laplace's method for asymptotic approximation (Papadimitriou et al. 1997a):

$$p(\mathcal{D}|\mathcal{M}_j) \approx p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j) p(\hat{\boldsymbol{\theta}}_j|\mathcal{M}_j) (2\pi)^{\frac{N_j}{2}} |\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)|^{-\frac{1}{2}}, \quad j = 1, 2, \ldots, N_M \tag{2.3}$$

where N_j is the number of uncertain parameters for the model class $\mathcal{M}_j$, the optimal parameter vector $\hat{\boldsymbol{\theta}}_j$ is the most probable value (it is assumed to maximize $p(\boldsymbol{\theta}_j|\mathcal{D}, \mathcal{M}_j)$ in the interior of $\boldsymbol{\Theta}_j$) and $\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)$ is the Hessian matrix of the function $-\ln[p(\mathcal{D}|\boldsymbol{\theta}_j, \mathcal{M}_j) p(\boldsymbol{\theta}_j|\mathcal{M}_j)]$ with respect to $\boldsymbol{\theta}_j$ evaluated at $\hat{\boldsymbol{\theta}}_j$. For *unidentifiable* cases (Beck and Katafygiotis 1998), the evidence $p(\mathcal{D}|\mathcal{M}_j)$ can be calculated by using an extension of the asymptotic expansion used in Eqn. 2.3 (Beck and Katafygiotis 1998; Katafygiotis et al. 1998) or by using a Markov chain Monte Carlo simulation technique (Beck and Au 2002) on Eqn. 2.2. The discussion here will focus on the *globally identifiable* case.

The likelihood factor $p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j)$ in Eqn. 2.3 will be higher for those model classes $\mathcal{M}_j$ that make the probability of the data $\mathcal{D}$ higher, that is, that give a better 'fit' to the data. For example, if the likelihood function is Gaussian, then the highest value of $p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j)$ will be given by the model class $\mathcal{M}_j$ that gives the smallest least-squares fit to the data. As mentioned earlier, this likelihood factor favors model classes with more uncertain parameters. If the number of data points N in $\mathcal{D}$ is large, the likelihood factor will be the dominant one in Eqn. 2.3 because it increases exponentially with N, while the other factors behave as N^{-1}, as shown below.

The remaining factors $p(\hat{\boldsymbol{\theta}}_j|\mathcal{M}_j)(2\pi)^{\frac{N_j}{2}} |\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)|^{-\frac{1}{2}}$ in Eqn. 2.3 are called the *Ockham factor* by Gull (1988). The Ockham factor represents a penalty against parameterization (Gull 1988; Mackay 1992), as we demonstrate in the following discussion.

We wish to show that the Ockham factor decreases exponentially with the number of uncertain parameters in the model class. For this purpose, consider an alternative expression for it, derived as follows. It is known that for a large number N of data points in $\mathcal{D}$, the updated (posterior) PDF $p(\boldsymbol{\theta}_j|\mathcal{D}, \mathcal{M}_j)$ is well approximated by a Gaussian PDF with mean $\hat{\boldsymbol{\theta}}_j$ and covariance matrix given by the inverse of the Hessian

matrix $\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)$. The principal posterior variances for $\boldsymbol{\theta}_j$, denoted by $\sigma_{j,i}^2$ with $i = 1, 2, \ldots, N_j$, are therefore the inverse of the eigenvalues of this Hessian matrix. The determinant factor $|\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)|^{-\frac{1}{2}}$ in the Ockham factor can therefore be expressed as the product of all the $\sigma_{j,i}$ for $i = 1, 2, \ldots, N_j$. Assume that the prior PDF $p(\boldsymbol{\theta}_j|\mathcal{M}_j)$ is Gaussian with mean (most probable value *a priori*) $\bar{\boldsymbol{\theta}}_j$ and a diagonal covariance matrix with variances $\rho_{j,i}^2$ with $i = 1, 2, \ldots, N_j$. The logarithm of the Ockham factor for the model class $\mathcal{M}_j$, denoted by β_j, can therefore be expressed as

$$\beta_j = -\sum_{i=1}^{N_j} \ln \frac{\rho_{j,i}}{\sigma_{j,i}} - \frac{1}{2} \sum_{i=1}^{N_j} \left(\frac{\hat{\theta}_{j,i} - \bar{\theta}_{j,i}}{\rho_{j,i}} \right)^2 \tag{2.4}$$

Since the prior variances will always be greater than the posterior variances if the data provides any information about the model parameters in the model class $\mathcal{M}_j$, all the terms in the first sum in Eqn. 2.4 will be positive and so will the terms in the second sum unless the posterior most probable value $\hat{\theta}_{j,i}$ just happens to coincide with the prior most probable value $\bar{\boldsymbol{\theta}}_{j,i}$. Thus, the log Ockham factor β_j will decrease if the number of parameters N_j for the model class $\mathcal{M}_j$ is increased. Furthermore, since the posterior variances are known to be inversely proportional to the number of data points N in $\mathcal{D}$, the dependence of the log Ockham factor on N is

$$\beta_j = -\frac{1}{2} \ln N N_j + R_j \tag{2.5}$$

where the remainder R_j depends primarily on the choice of prior PDF and is $\mathcal{O}(1)$ for large N. It is not difficult to show that this result holds for even more general forms of the prior PDF than the Gaussian PDF used here.

It follows from Bayes' Theorem that we have the exact relationship:

$$p(\mathcal{D}|\mathcal{M}_j) = p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j) p(\hat{\boldsymbol{\theta}}_j|\mathcal{M}_j) / p(\hat{\boldsymbol{\theta}}_j|\mathcal{D}, \mathcal{M}_j) \tag{2.6}$$

A comparison of this equation and Eqn. 2.3 shows that the Ockham factor is approximately equal to the ratio $p(\hat{\boldsymbol{\theta}}_j|\mathcal{M}_j)/p(\hat{\boldsymbol{\theta}}_j|\mathcal{D}, \mathcal{M}_j)$ which is always less than unity if the

data provides any information about the model parameters in the model class $\mathcal{M}_j$. Indeed, for large N, the negative of the logarithm of this ratio is an asymptotic approximation of the information about $\boldsymbol{\theta}_j$ provided by data $\mathcal{D}$ (Kullback 1968). Therefore, the log Ockham factor β_j removes the amount of information about $\boldsymbol{\theta}_j$ provided by $\mathcal{D}$ from the log likelihood $\ln p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j)$ to give the log evidence $p(\mathcal{D}|\mathcal{M}_j)$.

The Ockham factor may also be interpreted as a measure of robustness of the model class $\mathcal{M}_j$. If the updated PDF for the model parameters for the given model class is very peaked, then the ratio $p(\hat{\boldsymbol{\theta}}_j|\mathcal{M}_j)/p(\hat{\boldsymbol{\theta}}_j|\mathcal{D}, \mathcal{M}_j)$, and so the Ockham factor, is very small. But a narrow peak implies that response predictions using this model class will depend too sensitively on the optimal parameters $\hat{\boldsymbol{\theta}}_j$. Small errors in the parameter estimation will lead to large errors in the results. Therefore, a class of models with a small Ockham factor will not be robust to noise in the data during parameter estimation, that is, during selection of the optimal model within the class. Note that $\ln p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j)$ and the log Ockham factor β_j are approximately proportional to N and $\ln N$, respectively, where N is the number of data points in $\mathcal{D}$. Therefore, as N becomes larger, the contribution of the log Ockham factor becomes less important. This is reasonable because the uncertainty in the values of the model parameters becomes smaller as the number of data points grows, that is, the parameters can be estimated more precisely if more data points are available. In this case, the model class can be less robust since we are more confident about the values of the parameters of the model class.

To summarize, in the Bayesian approach to model selection, the model classes are ranked according to $p(\mathcal{D}|\mathcal{M}_j)P(\mathcal{M}_j|\mathcal{U})$ for $j = 1, \ldots, N_M$, where the best class of models representing the system is the one which gives the largest value of this quantity. The evidence $p(\mathcal{D}|\mathcal{M}_j)$ may be calculated for each class of models using Eqn. 2.3. The prior distribution $P(\mathcal{M}_j|\mathcal{U})$ over all the model classes $\mathcal{M}_j$, $j = 1, \ldots, N_M$, must be specified. In this work, a uniform prior distribution is chosen, leaving the Ockham factor alone to penalize model classes with increased numbers of parameters.

2.2.1 Comparison with Akaike's Approach

In the case of Akaike's information criterion (Akaike 1974), the best model class among the $\mathcal{M}_j$ for $j = 1, 2, \ldots, N_M$ is chosen by maximizing an objective function $\mathrm{AIC}(\mathcal{M}_j|\mathcal{D})$ over j that is defined by

$$\mathrm{AIC}(\mathcal{M}_j|\mathcal{D}) = \ln p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j) - N_j \tag{2.7}$$

where the log-likelihood function is roughly proportional to the number of data points N in $\mathcal{D}$, while the penalty term is taken to be N_j, the number of adjustable parameters in the model class $\mathcal{M}_j$. (Akaike actually stated his criterion as minimizing -2(AIC) but the equivalent form is more appropriate here). When the number of data points is large, the first term will dominate. Akaike (1976) and Schwarz (1978) later developed independently another version of the objective function, denoted BIC, that is defined by

$$\mathrm{BIC}(\mathcal{M}_j|\mathcal{D}) = \ln p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j) - \frac{1}{2} \ln N N_j \tag{2.8}$$

where now the penalty term increases with the number of data points N.

BIC can be compared directly with the logarithm of the evidence from Eqn. 2.3:

$$\ln p(\mathcal{D}|\mathcal{M}_j) \approx \ln p(\mathcal{D}|\hat{\boldsymbol{\theta}}_j, \mathcal{M}_j) + \beta_j \tag{2.9}$$

where the logarithm of the Ockham factor β_j is given by Eqn. 2.4 or Eqn. 2.5. The latter shows that for large N, the BIC agrees with the leading order terms in the logarithm of the evidence and so in this case it is equivalent to the Bayesian approach using equal priors for all of the $P(\mathcal{M}_j|\mathcal{U}_j)$.

2.3 Model Updating Using a Bayesian Framework

A general Bayesian framework for structural model updating was proposed in Beck and Katafygiotis (1998) and Katafygiotis et al. (1998). It was originally pre-

sented using input-output measurements. In this section, this Bayesian approach for linear/nonlinear model updating is presented. For details, see Beck and Katafygiotis (1998). The case of using output only measurements is covered later in Chapter 3 (linear models) and Chapter 4 (nonlinear models).

Consider a system with N_d degrees of freedom (DOFs) and equation of motion

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{f}_s(\mathbf{x}, \dot{\mathbf{x}}; \boldsymbol{\theta}_s) = \mathbf{T}\mathbf{f}(t) \tag{2.10}$$

where $\mathbf{M} \in \mathbb{R}^{N_d \times N_d}$ is the mass matrix, $\mathbf{f}_s \in \mathbb{R}^{N_d}$ is the nonlinear restoring force characterized by the structural parameters $\boldsymbol{\theta}_s$, $\mathbf{T} \in \mathbb{R}^{N_d \times N_f}$ is a force distributing matrix and $\mathbf{f}(t) \in \mathbb{R}^{N_f}$ is an external excitation, e.g., force or ground acceleration, which is assumed to be measured.

Assume now that discrete response data are available for $N_o (\leq N_d)$ measured DOFs. Let Δt denote the sampling time step. Because of measurement noise and modeling error, referred to hereafter as prediction error, the measured response $\mathbf{y}(n) \in \mathbb{R}^{N_o}$ (at time $t = n\Delta t$) will differ from the model response $\mathbf{L}_0\mathbf{x}(n)$ corresponding to the measured degrees of freedom where $\mathbf{L_0}$ denotes an $N_o \times N_d$ observation matrix, comprised of zeros and ones. Herein, it is assumed that this difference between the measured and model response can be adequately represented by a discrete zero-mean Gaussian white noise vector process $\boldsymbol{\eta}(n) \in \mathbb{R}^{N_o}$:

$$\mathbf{y}(n) = \mathbf{L}_0\mathbf{x}(n) + \boldsymbol{\eta}(n) \tag{2.11}$$

where the discrete process $\boldsymbol{\eta}$ satisfies

$$E[\boldsymbol{\eta}(n)\boldsymbol{\eta}^T(p)] = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}\delta_{np} \tag{2.12}$$

where $E[.]$ denotes expectation, δ_{np} denotes the Kronecker delta function, and $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ denotes the $N_o \times N_o$ covariance matrix of the prediction error process $\boldsymbol{\eta}$.

Let $\boldsymbol{\theta}$ denote the parameter vector for identification and it includes the following parameters: 1) the structural parameters $\boldsymbol{\theta}_s$; 2) parameters defining the structural

mass distribution; 3) the elements of the force distributing matrix $\mathbf{T}$; and 4) the elements of the upper right triangular part of the prediction-error covariance matrix $\mathbf{\Sigma}_\eta$ (symmetry defines the lower triangular part of this matrix). Herein, it is assumed that the mass distribution can be modeled sufficiently accurately from structural drawings and so it is not part of the model parameters to be identified.

If the data $\mathcal{D}$ consists of the measured time histories at N discrete times of the excitation and observed response, then it is easily shown that the most probable values $\hat{\boldsymbol{\theta}}$ of the model parameters are calculated by minimizing the mean square error between the measured and computed model response at the observed DOFs because of the assumed probability model for the prediction error. Assuming that the prediction errors have equal variance σ_η but are independent for different channels of measurements, the updated PDF of the model parameters $\boldsymbol{\theta}$ given dynamic data $\mathcal{D}$ and model class $\mathcal{M}$ is given by

$$p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}) = c_1 p(\boldsymbol{\theta}|\mathcal{M})(2\pi)^{-\frac{NN_o}{2}}\sigma_\eta^{-NN_o}\exp\left(-\frac{NN_o}{2\sigma_\eta^2}J_1(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})\right) \tag{2.13}$$

where c_1 is a normalizing constant and $p(\boldsymbol{\theta}|\mathcal{M})$ is the prior PDF of the model parameters $\boldsymbol{\theta}$ expressing the user's judgement about the relative plausibility of the values of the model parameters before data is used. The objective function $J_1(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$ is given by

$$J_1(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}) = \frac{1}{NN_o}\sum_{k=1}^{N}\left|\left|\mathbf{L}_0\mathbf{x}(k\Delta t;\boldsymbol{\theta},\mathcal{M}) - \mathbf{y}(k\Delta t)\right|\right|^2 \tag{2.14}$$

where $\mathbf{x}(k\Delta t;\boldsymbol{\theta},\mathcal{M})$ is the calculated response based on the assumed class of models and the parameter set $\boldsymbol{\theta}$ and $\mathbf{y}(k\Delta t)$ is the measured response at time $k\Delta t$, respectively. Furthermore, $||.||$ denotes the 2-norm of a vector. The most probable model parameters $\hat{\boldsymbol{\theta}}$ are obtained by maximizing $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$ in Eqn. 2.13. For large N, this is equivalent to minimizing $J_1(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$ in Eqn. 2.14 over all parameters in $\boldsymbol{\theta}$ that it depends on, because this factor dominates in Eqn. 2.13. The most probable value of the prediction-error variance in $\hat{\boldsymbol{\theta}}$ is $\hat{\sigma_\eta^2} = \min J_1(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$. In the globally identifiable

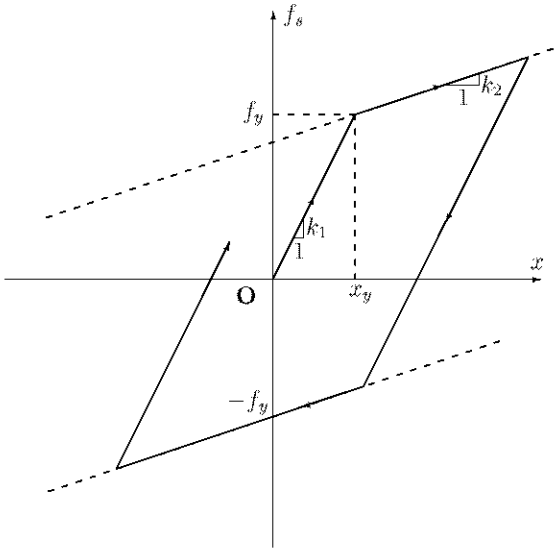

Figure 2.1: Relationship between the restoring force and the displacement of the bilinear hysteretic system (Example 2-1).

case (Beck and Katafygiotis 1998), it turns out that $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$ is well approximated by a Gaussian distribution with mean $\hat{\boldsymbol{\theta}}$ and covariance matrix equal to inverse of the Hessian of $-\ln[p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})]$ at $\hat{\boldsymbol{\theta}}$.

2.4 Illustrative Examples

2.4.1 Example 2-1: Single-degree-of-freedom Nonlinear Oscillator under Seismic Excitation

In this example, a bilinear hysteretic oscillator with linear viscous damping is considered:

$$m\,\ddot{x} + c\dot{x} + f_h(x; k_1, k_2, x_y) = f(t) \tag{2.15}$$

where m is the mass, c is the damping coefficient and $f_h(x; k_1, k_2, x_y)$ is the hysteretic restoring force, whose behavior is shown in Fig. 2.1. Here, $m = 1$kg is assumed known. The parameters $\tilde{\boldsymbol{\theta}} = [\tilde{c}, \tilde{k}_1, \tilde{k}_2, \tilde{x}_y]^T$ used to generate the data are: $\tilde{c} = 0.02$ Ns/m,

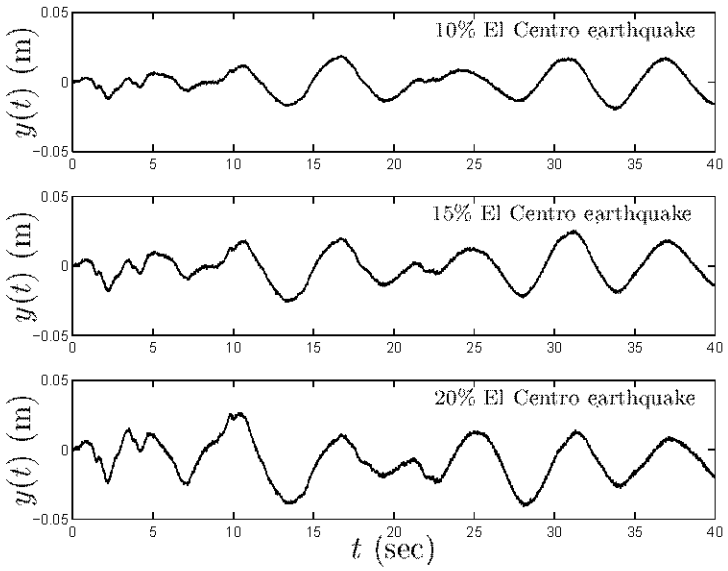

Figure 2.2: Response measurements of the oscillator for the three levels of excitation (Example 2-1).

$\tilde{k}_1 = 1.0$ N/m, $\tilde{k}_2 = 0.1$N/m, $\tilde{x}_y = 0.02$ m, which gives a small-amplitude natural frequency of $\frac{1}{2\pi}$Hz.

The oscillator is assumed to be excited by 10%, 15% and 20% of the 1940 El Centro earthquake record. The duration of measurement is $T = 40$ sec with sampling frequency 60Hz, so that the number of data points is $N = 2400$. It is assumed that the earthquake excitation and response displacement are measured to give the data $\mathcal{D}$ where 5% rms noise is imposed on the structural response measurements, i.e., the measurement noise is 5% of the rms of the noise-free response. Fig. 2.2 shows the measurements for the three levels of excitation and Fig. 2.3 shows the corresponding hysteresis loops. It can be seen that the oscillator behaved linearly (did not yield) when subjected to 10% of the El Centro earthquake record. Three classes of models are considered. They all use zero-mean Gaussian discrete white noise as the prediction-error model.

Model Class 1 ($\mathcal{M}_1$): Linear oscillators with damping coefficient $c > 0$, stiffness parameter $k_1 > 0$ and predictive-error standard deviation σ_η;

Model Class 2 ($\mathcal{M}_2$): Elasto-plastic oscillators, i.e., bilinear hysteretic but with

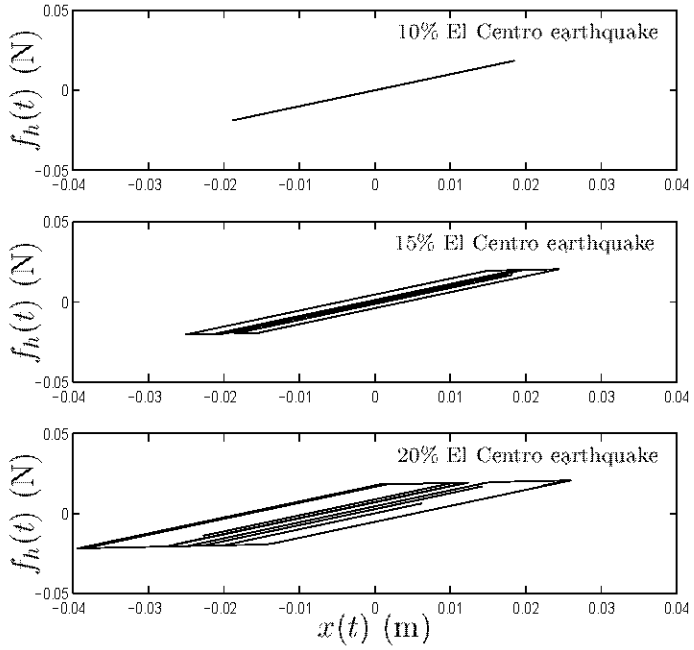

Figure 2.3: Hysteresis loops of the oscillator for the three levels of excitation (Example 2-1).

$k_2 = 0$, with stiffness parameter $k_1 > 0$, yielding level x_y and predictive-error standard deviation σ_η; and no viscous damping.

Model Class 3 ($\mathcal{M}_3$): bilinear hysteretic oscillators with pre-yield stiffness $k_1 > 0$, after yielding stiffness $k_2 > 0$, yielding level x_y and predictive error parameter σ_η. Note that this class of models does not include the exact model since linear viscous damping is not included.

Independent uniform prior distributions are assumed for the parameters c, k_1, k_2, x_y and σ_η over the range (0,0.5)N sec/m, (0,2)N/m, (0,0.5)N/m, (0,0.1)m, (0,0.01)m, respectively. Table 2.1 shows the optimal parameters of each class of models for the three levels of excitation. 'UN' indicates that the parameter is unidentifiable. For example, in $\mathcal{M}_2$ with 10% El Centro earthquake, x_y is unidentifiable because the oscillator behaves perfectly linearly (Fig. 2.3). In fact, the optimal parameters of $\mathcal{M}_1$ are very close to their target values in this level of excitation. For higher levels of excitation, the optimal linear model in $\mathcal{M}_1$ has lower stiffness and higher values of its damping coefficient to represent the increased flexibility and energy dissipation

Excitation Level	Model Class	c	k_1	k_2	x_y	σ_η
10% El Centro earthquake	1	0.0204	1.000	----	----	0.0005
	2	----	1.019	----	UN	0.0013
	3	----	1.019	UN	UN	0.0013
15% El Centro earthquake	1	0.0902	0.989	----	----	0.0020
	2	----	1.0179	----	0.0214	0.0017
	3	----	1.001	0.108	0.0197	0.0007
20% El Centro earthquake	1	0.1928	0.956	----	----	0.0098
	2	----	0.9936	----	0.0211	0.0051
	3	----	0.9942	0.0924	0.0200	0.0011

Table 2.1: Optimal (most probable) parameter values in each model class representing the oscillator (Example 2-1).

Excitation level	$P(\mathcal{M}_1\|\mathcal{D},\mathcal{U})$	$P(\mathcal{M}_2\|\mathcal{D},\mathcal{U})$	$P(\mathcal{M}_3\|\mathcal{D},\mathcal{U})$
10% El Centro earthquake	1.0	3.1×10^{-1217}	3.1×10^{-1217}
15% El Centro earthquake	4.4×10^{-1174}	3.2×10^{-957}	1.0
20% El Centro earthquake	6.4×10^{-2303}	5.7×10^{-1609}	1.0

Table 2.2: Probabilities of different model classes based on data (Example 2-1).

due to yielding.

Table 2.2 shows the values of $P(\mathcal{M}_j|\mathcal{D},\mathcal{U})$, $j = 1, 2, 3$ for the three levels of excitation that are calculated from Eqn. 2.1 using the evidence for each model from Eqn. 2.3 and equal priors $P(\mathcal{M}_j|\mathcal{U}) = \frac{1}{3}$. Note that in all three cases, the optimal model class has probability near 1.0, implying that the other model classes can be discarded for response prediction. In the case of 10% scaling of the El Centro earthquake record, it is not surprising that $P(\mathcal{M}_1|\mathcal{D},\mathcal{U})$ is the largest since the oscillator behaves linearly (Fig. 2.3). However, for higher levels of excitation, $P(\mathcal{M}_3|\mathcal{D},\mathcal{U})$ is the largest. Although $\mathcal{M}_3$ does not include linear viscous damping, the hysteretic behavior can be captured well by this model. More interestingly, $\mathcal{M}_2$ out-performs $\mathcal{M}_1$ at these two levels of excitation. Although $\mathcal{M}_2$ can not capture the viscous damping mechanism, the energy dissipated by the hysteretic behavior for 15% and 20% scaling of the El Centro earthquake record is much more significant than the contribution from the

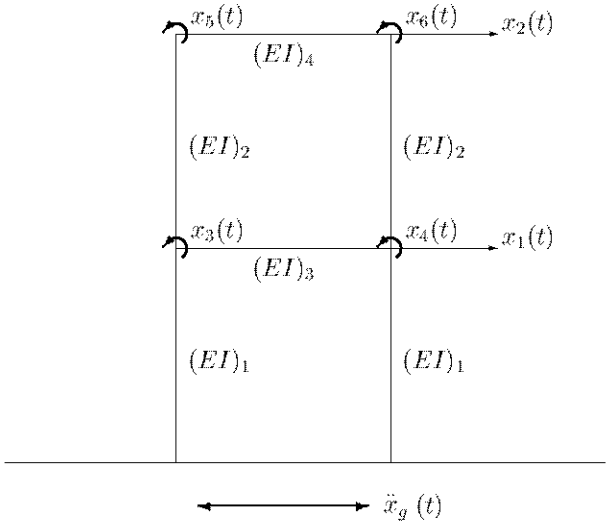

Figure 2.4: Linear two-story structural frame (Example 2-2).

viscous damping, as can be seen by the large increase in the optimal damping ratio for the corresponding "equivalent" linear systems $\mathcal{M}_1$ in Table 2.1. Furthermore, the restoring force behavior for $\mathcal{M}_2$ is more correct than for $\mathcal{M}_1$, although it is still not exact.

This example illustrates an important point in system identification. In reality, there is no exact class of models for a real structure and the best class depends on the circumstances. If we wish to select between the linear models ($\mathcal{M}_1$) and the elasto-plastic models ($\mathcal{M}_2$), then $\mathcal{M}_2$ is better for high levels of excitation while $\mathcal{M}_1$ is better for lower levels of excitation.

2.4.2 Example 2-2: Linear Two-story Frame under Seismic Excitation

The second example refers to a 6-DOF two-story structural frame with story height $H = 2.5$m and width $W = 4.0$m, as shown in Fig. 2.4. All the chosen model classes are linear. All members are assumed to be rigid in their axial direction. For each member, the mass is uniformly distributed along its length. The rigidity-to-mass ratio is chosen to be $\frac{\widetilde{EI}_1}{m} = \frac{\widetilde{EI}_2}{m} = \frac{\widetilde{EI}_3}{10m} = \frac{\widetilde{EI}_4}{10m} = 2252\text{m}^4\,\text{sec}^{-2}$, where m denotes the

mass per unit length of all members. As a result, the first two natural frequencies of this structure are 2.000Hz and 5.144Hz. Furthermore, a Rayleigh damping model is assumed, i.e., the damping matrix $\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}$, where $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices, respectively. In this case, the nominal values of the damping coefficients $\tilde{\alpha}$ and $\tilde{\beta}$ are chosen to be $0.182\,\mathrm{sec}^{-1}$ and $0.442 \times 10^{-3}\,\mathrm{sec}$ so that the damping ratios for the first two modes are 1.00%.

Three classes of structural models are considered. Independent zero-mean discrete Gaussian white noise is used for the prediction-error model, with spectral intensity $S_{n1} = 0.027\mathrm{m}^2\,\mathrm{sec}^{-3}$ and $S_{n2} = 0.059\mathrm{m}^2\,\mathrm{sec}^{-3}$ at the two observed degrees of freedom. In order to have better scaling, the damping parameters are parameterized as follows: $\alpha = \phi_1\tilde{\alpha}$ and $\beta = \phi_2\tilde{\beta}$.

Model Class 1 ($\mathcal{M}_1$): Assumes a class of two-story shear buildings with nominal interstory stiffness $\bar{k}_1 = \bar{k}_2 = 2 \times \frac{12\widetilde{EI}_1}{H^3}$. In order to have better scaling, the stiffness are parameterized as follows: $k_j = \theta_j\bar{k}_j$, $j = 1, 2$. Therefore, the uncertain parameters are θ_j, ϕ_j, S_{nj}, $j = 1, 2$.

Model Class 2 ($\mathcal{M}_2$): Assumes the actual class of models except that due to modeling error, $EI_1 = \theta_1\widetilde{EI}_1$, $EI_2 = \theta_2\widetilde{EI}_2$, $EI_3 = 0.5\theta_1\widetilde{EI}_3$ and $EI_4 = 0.5\theta_2\widetilde{EI}_4$, where the nominal values were given earlier. Therefore, the uncertain parameters are θ_j, ϕ_j and S_{nj}, $j = 1, 2$.

Model Class 3 ($\mathcal{M}_3$): Assumes that $EI_1 = \theta_1\widetilde{EI}_1$, $EI_2 = \theta_2\widetilde{EI}_2$ and $EI_j = \theta_3\widetilde{EI}_j, j = 3, 4$. Therefore, the uncertain parameters are: θ_1, θ_2, θ_3, ϕ_1, ϕ_2, S_{n1} and S_{n2}. Note that the true model lies in this set.

The structure is assumed to be excited by a white noise ground motion, which is not measured. The spectral intensity of the ground motion is taken to be $S_0 = 1.0 \times 10^{-5}\mathrm{m}^2\,\mathrm{sec}^{-3}$. The data $\mathcal{D}$ consists of the absolute accelerations with 10% measurement noise at the 1^{st} and 2^{nd} DOFs over a time interval of 100 sec, using a sampling interval of 0.01 sec. Identification was performed using the Bayesian spectral density approach of Chapter 4 with the same set of data for each of the three classes of models.

The prior distributions $p(\boldsymbol{\theta}_j|\mathcal{M}_j), j = 1, 2, 3$ are assumed to be an independent

uniform distribution over the interval $(0, 2)$ for θ_1, θ_2, θ_3, ϕ_1, ϕ_2 and over the interval $(0, 0.5)\text{m}^2\,\text{sec}^{-3}$ for S_{n1} and S_{n2}.

Parameter	ϕ_1	ϕ_2	θ_1	θ_2	θ_3	S_{n1}	S_{n2}
Case 1	1.131	1.007	0.913	0.879	—	0.158	0.159
Case 2	1.057	1.536	1.130	1.130	—	0.150	0.063
Case 3	1.027	1.093	0.988	1.001	1.024	0.085	0.080

Table 2.3: Optimal (most probable) structural parameter values in each model class representing the structural frame (Example 2-2).

Mode	1	2
Actual	2.000	5.144
Case 1	2.048	5.009
Case 2	2.000	5.323
Case 3	1.995	5.142

Table 2.4: Natural frequencies (in Hz) of the best model in each class (Example 2-2).

Table 2.3 shows the optimal structural parameters in each class of models. It is not surprising that both θ_1 and θ_2 in Case 1 are less than unity because the shear building models assume a rigid floor but the floors of the actual structure are not. Table 2.4 shows the associated natural frequencies with the actual frame and the optimal models. Note that the optimal model in $\mathcal{M}_3$ can fit both frequencies very well since the exact model is in this class. On the other hand, $\mathcal{M}_1$ and $\mathcal{M}_2$ can not fit the frequency of the second mode as well as $\mathcal{M}_3$. Fig. 2.5 - 2.7 show the estimated spectrum using the measurements (zigzag curve) with the best fitting spectrum (smoother curve) for the three classes of models, respectively. One can see that the best model in $\mathcal{M}_2$ provides a better fit to the first mode than $\mathcal{M}_1$, but it is the opposite for the second mode. The best model in $\mathcal{M}_3$ gives excellent matching with the estimated spectrum for both modes.

Table 2.5 shows the values of $P(\mathcal{M}_j|\mathcal{D}, \mathcal{U})$ for $j = 1, 2, 3$, calculated from Eqn. 2.1 using the evidence for each model from Eqn. 2.3 and equal priors $P(\mathcal{M}_j|\mathcal{U}) = \frac{1}{3}$.

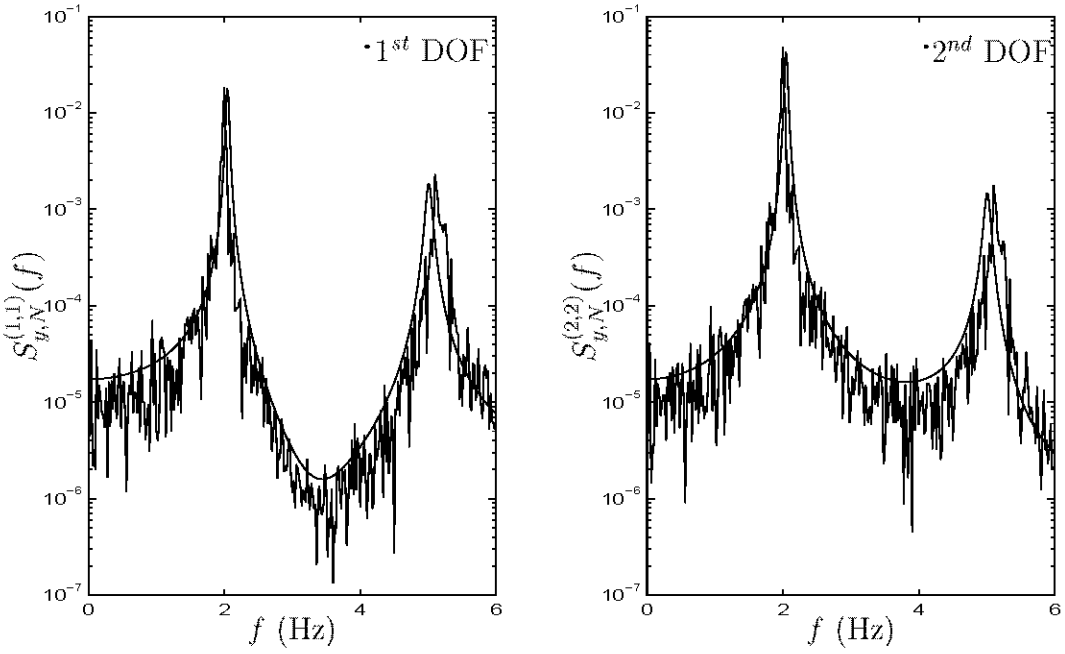

Figure 2.5: Response spectrum estimated by the measurements and the best fitting curve using Model Class 1 (Example 2-2).

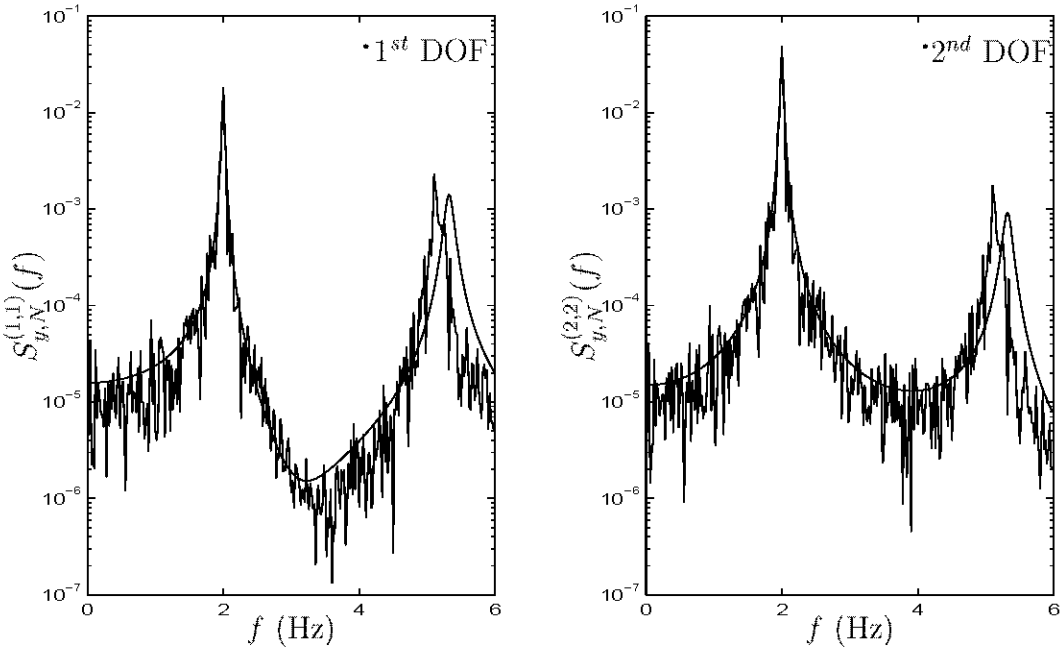

Figure 2.6: Response spectrum estimated by the measurements and the best fitting curve using Model Class 2 (Example 2-2).

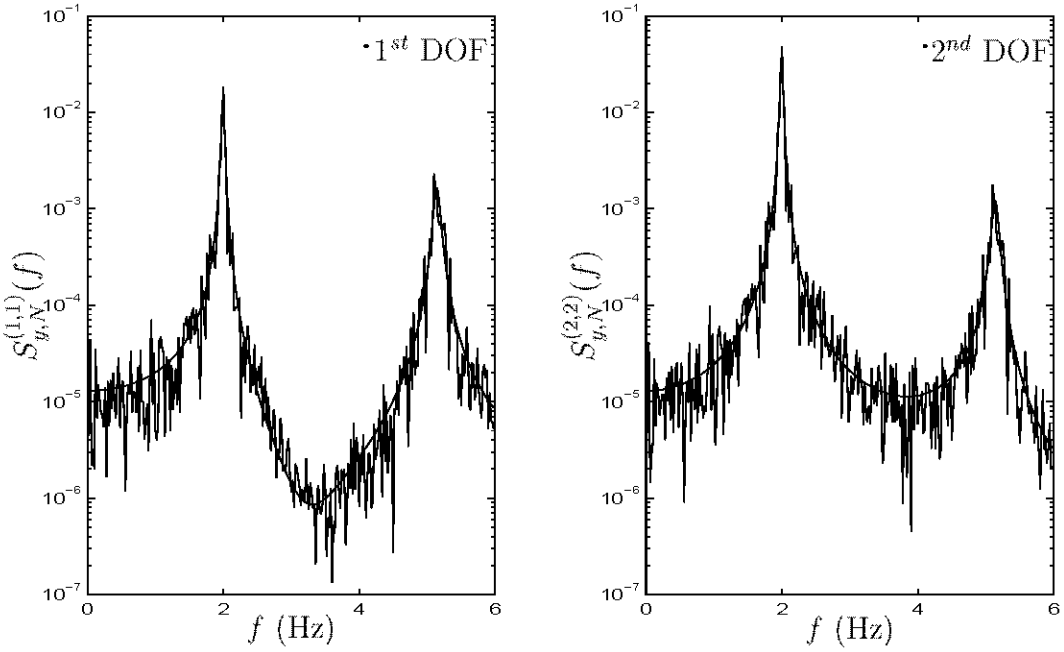

Figure 2.7: Response spectrum estimated by the measurements and the best fitting curve using Model Class 3 (Example 2-2).

As expected, $P(\mathcal{M}_3|\mathcal{D},\mathcal{U})$ is the largest among the three classes of models because it contains the actual model. On the other hand, $P(\mathcal{M}_1|\mathcal{D},\mathcal{U})$ is the smallest one. Although it gives a better fit for the second mode than $\mathcal{M}_2$, it does not fit the first mode as well as the best model in $\mathcal{M}_2$ and the contribution of the first mode to the structural response is one order of magnitude larger than the second mode. This implies that although $\mathcal{M}_2$ has significant modeling error for the beams (about 50%), it is still a better class of models than the shear building models.

$P(\mathcal{M}_1\|\mathcal{D},\mathcal{U})$	$P(\mathcal{M}_2\|\mathcal{D},\mathcal{U})$	$P(\mathcal{M}_3\|\mathcal{D},\mathcal{U})$
2.6×10^{-23}	1.7×10^{-15}	1.0

Table 2.5: Probabilities of different model classes based on data (Example 2-2).

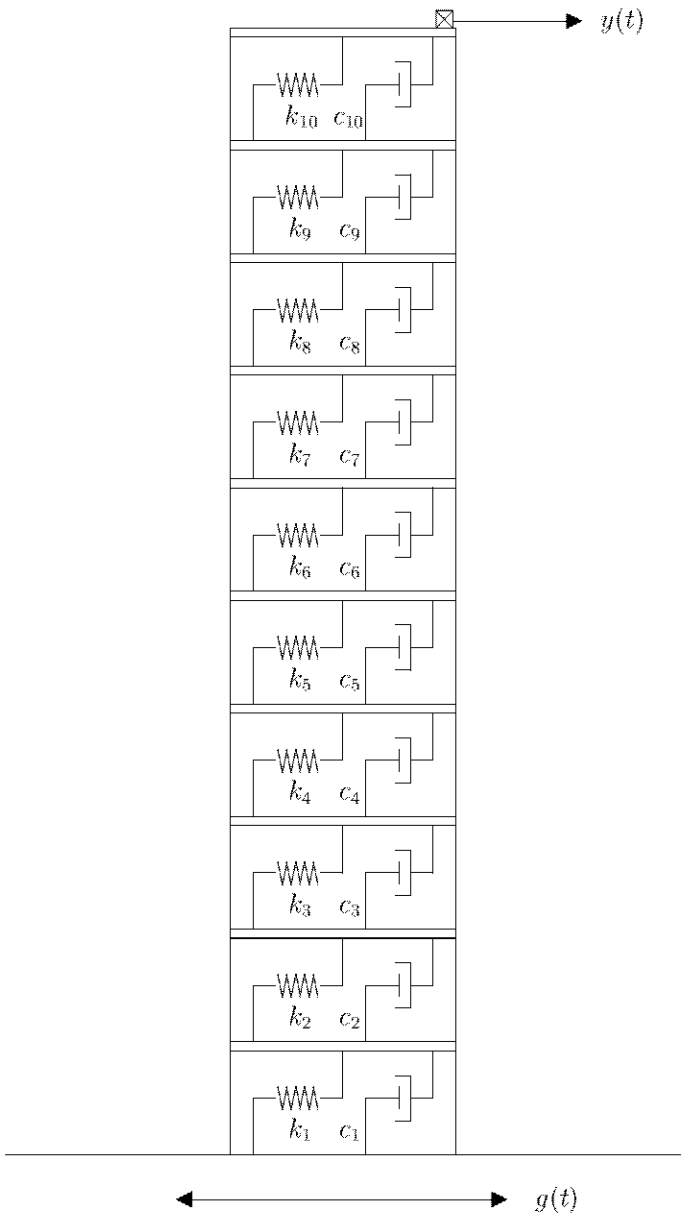

Figure 2.8: Ten-story shear building (Example 2-3).

2.4.3 Example 2-3: Ten-story Shear Building under Seismic Excitation

The third example uses response measurements from the ten-story building shown in Fig. 2.8. The Bayesian approach is applied to select the optimal number of modes for a linear model. It is assumed that this building has a uniformly distributed floor mass and story stiffness over its height. The stiffness to mass ratios $\frac{\tilde{k}_j}{m_j}, j = 1, \ldots, 4$, are chosen to be $1500\,\mathrm{sec}^{-2}$ so that the fundamental frequency of the building is 0.9213 Hz. Rayleigh damping is assumed, i.e., the damping matrix $\mathbf{C}$ is given by $\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}$, where $\alpha = 0.0866\,\mathrm{sec}^{-1}$ and $\beta = 0.0009\,\mathrm{sec}$. The structure is assumed to be subjected to a wide-band random ground motion, which can be adequately modeled as a Gaussian white noise with spectral intensity $S_{f0} = 0.02\mathrm{m}^2\,\mathrm{sec}^{-3}$. Note that the matrix $\mathbf{T}$ in Eqn. 2.10 is equal to the matrix $-[m_1, ..., m_{10}]^T$ in this case. Each model class $\mathcal{M}_j$ $(j = 1, ..., 8)$ consists of a linear modal model (Beck 1996) with j modes and the uncertain parameters are the natural frequency, damping ratio and modal participation factor for each mode; and the spectral intensity S_n of the prediction error at the measured degree of freedom.

The data $\mathcal{D}$ consists of the absolute accelerations at the top floor with 5% measurement noise over a time interval $T = 30\,\mathrm{sec}$, using a sampling interval $\Delta t = 0.01\,\mathrm{sec}$. The measurement noise is simulated using a spectral intensity $S_n = 1.94 \times 10^{-4}\mathrm{m}^2\,\mathrm{sec}^{-3}$. The Bayesian spectral density approach of Chapter 4 is used for the identification. The number of data points N is taken to be 600 because only the estimated spectrum up to 20.0 Hz is used.

Independent prior distributions for the parameters are taken as follows: Gaussian distribution for the natural frequencies with mean $5.5(2j-1)$ rad/sec and coefficient of variation 0.05 for the j^{th} mode. Furthermore, the damping ratios, modal participation factor and the spectral intensity of the modeling error are assumed to be uniformly distributed over the range $(0, 0.05)$, $(0, 2)$ and $(0, 0.01)\mathrm{m}^2\,\mathrm{sec}^{-3}$, respectively.

Table 2.6 shows the identified (most probable) natural frequencies for considering one mode to eight modes. Table 2.7 shows the values of the log-evidence $\ln p(\mathcal{D}|\mathcal{M}_m)$,

Number of modes	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8
Exact	5.789	17.24	28.30	38.73	48.30	56.78	64.00	69.79
1	6.946	.....	.....	.....	.....	.....	.....	.....
2	5.799	20.68	.....	.....	.....	.....	.....	.....
3	5.814	17.16	33.96	.....	.....	.....	.....	.....
4	5.842	17.18	27.94	43.82	.....	.....	.....	.....
5	5.848	17.19	27.97	38.06	50.58	.....	.....	.....
6	5.849	17.19	27.97	38.09	48.10	56.72	.....	.....
7	5.849	17.19	27.97	38.09	48.13	56.34	64.18	.....
8	5.849	17.19	27.97	38.09	48.13	56.34	64.18	69.41

Table 2.6: Identified natural frequencies in rad/sec of the building (Example 2-3).

Number of modes m	1	2	3	4
$\ln p(\mathcal{D}\|\mathcal{M}_m)$	1.894×10^3	2.251×10^3	2.511×10^3	2.619×10^3
$\ln \beta_m$	-43.7	-56.4	-68.9	-69.2
$P(\mathcal{M}_m\|\mathcal{D},\mathcal{U})$	3.0×10^{-336}	2.2×10^{-186}	6.4×10^{-79}	2.4×10^{-32}
Number of modes m	5	6	7	8
$\ln p(\mathcal{D}\|\mathcal{M}_m)$	2.682×10^3	2.714×10^3	2.723×10^3	2.723×10^3
$\ln \beta_m$	-75.9	-91.2	-109	-121
$P(\mathcal{M}_m\|\mathcal{D},\mathcal{U})$	1.0×10^{-7}	1.0	1.7×10^{-4}	1.3×10^{-9}

Table 2.7: Probabilities of models with different number of modes based on data (Example 2-3).

the log-Ockham factor $\ln \beta_m$ and $P(\mathcal{M}_j|\mathcal{D},\mathcal{U})$ $(j = 1, ..., 8)$ for the cases of model classes with one mode to eight modes, calculated from Eqn. 2.1 using the evidence for each model from Eqn. 2.3 and equal priors $P(\mathcal{M}_j|\mathcal{U}) = \frac{1}{8}$. It implies that using six modes is optimal. It is found that the seven-mode and eight-mode models give poor estimation of the damping ratios although the estimated natural frequencies are satisfactory, as shown in Table 2.6. The estimated (most probable) damping ratios of the seventh mode are 15.3% and 17.2%, using the seven-mode and eight-mode models, respectively. The eight-mode model gives 25.9% for the most probable damping ratio for the eighth mode. Note that the actual values of the damping ratios of the seventh and eighth mode are 2.86% and 3.10%, respectively.

Fig. 2.9 shows the estimated spectrum from the data (zigzag curve) and the best

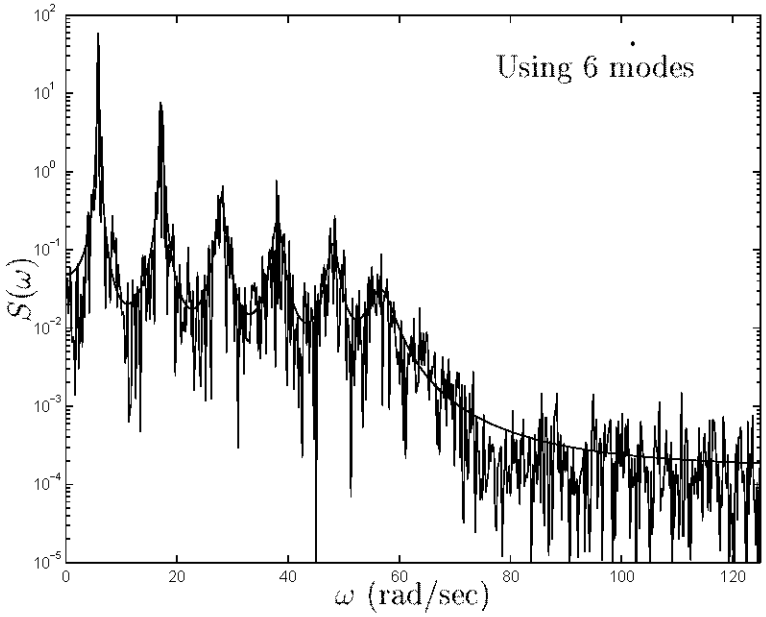

Figure 2.9: Response spectrum estimated by the measurements and the best fitting curve using six modes (Example 2-3).

fitting curve using six modes (smoother curve). One can see that the optimal model using six modes can fit the measured spectrum very well. Furthermore, all the six identified natural frequencies are very close to their target values, which is not the case for using two to five modes. It was found that if AIC is used, eight modes is optimal because the penalty term is too small compared to the changing of the log likelihood term in Eqn. 2.7. On the other hand, if BIC in Eqn. 2.8 is used, then six modes are optimal, agreeing with the Bayesian approach using the evidence for the various modal models.

2.5 Conclusion

A Bayesian probabilistic approach for model selection is presented and numerical examples are given to illustrate the method. The optimal class of models is taken to be the most plausible one based on the data, that is, it possesses the largest probability conditional on the data among the model classes. This probability depends on the

evidence for the model class provided by the data and the user's choice of prior probability distribution over the classes of models. The methodology can handle input-output and output-only data for linear and nonlinear dynamical systems. This is further illustrated in Chapters 3, 4 and 6.

The optimal class of models is taken to be the most plausible one based on the data, that is, it possesses the largest probability conditional on the data among the model classes. This probability depends on the evidence

Chapter 3 Modal Identification Using Nonstationary Noisy Measurements

3.1 Overview

This chapter addresses the problem of identification of the modal parameters for a structural system using measured nonstationary response time histories only. A Bayesian time-domain approach is presented which is based on an approximation of the probability distribution of the response to a nonstationary stochastic excitation. It allows one to obtain not only the most probable values of the updated modal parameters and stochastic excitation parameters but also their associated uncertainties using only one set of response data. It is found that the updated probability distribution can be well approximated by a Gaussian distribution centered at the most probable values of the parameters. Examples are presented to illustrate the proposed method.

3.2 Formulation for Modal Identification

3.2.1 Random Vibration Analysis

Consider a system with N_d degrees of freedom (DOF) and equation of motion:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\,\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{T}_o\mathbf{F}(t) \tag{3.1}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the mass, damping and stiffness matrices, respectively; $\mathbf{T}_o \in \mathbb{R}^{N_d \times N_F}$ is a force distributing matrix; and $\mathbf{F}(t) \in \mathbb{R}^{N_F}$ is a zero-mean Gaussian

nonstationary stochastic process which is modeled by

$$\mathbf{F}(t) = A(t)\mathbf{g}(t) \tag{3.2}$$

where $\mathbf{g(t)}$ is a Gaussian stationary stochastic process with zero mean and spectral density matrix $\mathbf{S}_g(\omega) \in \mathbb{R}^{N_F \times N_F}$ and $A(t) \in \mathbb{R}$ is a modulation function. Then, the autocorrelation function of $\mathbf{F}$ is given by

$$\mathbf{R}_F(t, t+\tau) = A(t)A(t+\tau)\mathbf{R}_g(\tau) \tag{3.3}$$

where $\mathbf{R}_g(\tau)$ is the autocorrelation function for the stationary process $\mathbf{g}(t)$.

Assuming classical damping, i.e., $\mathbf{CM^{-1}K} = \mathbf{KM^{-1}C}$ (Caughey and O'Kelly 1965), the uncoupled modal equations of motion by using modal analysis are given by

$$\ddot{q}_r\ (t) + 2\zeta_r\omega_r\ \dot{q}_r\ (t) + \omega_r^2 q_r(t) = A(t)f_r(t), \qquad r = 1, \ldots, N_d \tag{3.4}$$

where $\mathbf{q}(t) = [q_1(t), \ldots, q_{N_d}(t)]^T$ and $\mathbf{f}(t) = [f_1(t), \ldots, f_{N_d}(t)]^T$ are the modal coordinate vector and the modal forcing vector, respectively. The transformation between the original coordinates (forces) and the modal coordinates (forces) is given by

$$\mathbf{x}(t) = \mathbf{\Phi} \cdot \mathbf{q}(t) \quad \text{and} \quad \mathbf{f}(t) = (\mathbf{M\Phi})^{-1}\mathbf{T}_o\mathbf{g}(t) \tag{3.5}$$

where $\mathbf{\Phi}$ is the modeshape matrix, comprised of the modeshape vectors $\boldsymbol{\phi}^{(r)}$ which are assumed to be normalized so that

$$\boldsymbol{\phi}_{i_r}^{(r)} = 1, \qquad r = 1, ..., N_d \tag{3.6}$$

where i_r is a measured DOF which is not a node of the r^{th} mode. The modal forcing vector $\mathbf{f}(t)$ is a Gaussian stationary stochastic process with zero mean, spectral density

matrix

$$\mathbf{S}_f(\omega) = (\mathbf{M}\mathbf{\Phi})^{-1}\mathbf{T}_o\mathbf{S}_g(\omega)\mathbf{T}_o^T(\mathbf{M}\mathbf{\Phi})^{-T} \tag{3.7}$$

and autocorrelation matrix function

$$\mathbf{R}_f(\tau) = \int_{-\infty}^{\infty}\mathbf{S}_f(\omega)e^{i\omega\tau}d\omega \tag{3.8}$$

It is known that the response $\mathbf{x}(t)$ is a Gaussian process with zero mean, correlation function between x_j and x_l (Lutes and Sarkani 1997):

$$R_x^{(j,l)}(t,t+\tau) \simeq \sum_{r=1}^{N_m}\sum_{s=1}^{N_m}\phi_j^{(r)}\phi_l^{(s)}\int_0^t\int_0^{t+\tau}A(u)A(v)h_r(t-u)h_s(t+\tau-v)R_f^{(r,s)}(u-v)dudv \tag{3.9}$$

and with spectral density

$$S_x^{(j,l)}(t,\omega) \simeq \frac{1}{2\pi}\int_{-\infty}^{\infty}R_x^{(j,l)}(t,t+\tau)e^{-i\omega\tau}d\tau \tag{3.10}$$

where $h_r(.)$ denotes the modal unit impulse response function for the displacement of the r^{th} mode. Here, it is assumed that only N_m lower modes contribute significantly to the displacement response.

Assume that discrete data at times $t_k = k\Delta t, k = 1, \ldots, N$, are available at $N_o(\leq N_d)$ measured DOFs. Also, assume that due to measurement noise and modeling error there is prediction error, i.e., a difference between the measured response $\mathbf{y}(k) \in \mathbb{R}^{N_o}$ and the model response at time $t_k = k\Delta t$ corresponding to the measured degrees of freedom. The latter is given by $\mathbf{L}_o\mathbf{x}(k\Delta t)$ where $\mathbf{L_o}$ is an $N_o \times N_d$ observation matrix, comprised of zeros and ones, that is,

$$\mathbf{y}(k) = \mathbf{L}_o\mathbf{x}(k\Delta t) + \mathbf{n}(k) \tag{3.11}$$

It is assumed that the prediction error can be adequately represented by discrete

zero-mean Gaussian white noise $\mathbf{n}(k) \in \mathbb{R}^{N_o}$ with the following $N_o \times N_o$ covariance matrix

$$E[\mathbf{n}(m)\mathbf{n}^T(p)] = \mathbf{\Sigma}_n \delta_{m,p} \tag{3.12}$$

where $\delta_{m,p}$ is the Kroneker Delta function.

Note that $\mathbf{y}(k)$ is a discrete zero-mean Gaussian process with autocorrelation matrix function $\mathbf{R}_y$ given by

$$\begin{aligned} \mathbf{R}_y(m,p) &= E[\mathbf{y}(m)\mathbf{y}^T(p)] \\ &= \mathbf{L}_o \mathbf{R}_x(m\Delta t, p\Delta t)\mathbf{L}_o^T + \mathbf{\Sigma}_n \delta_{m,p} \end{aligned} \tag{3.13}$$

where $\mathbf{R}_x$ denotes the autocorrelation matrix function of the model response $\mathbf{x}(t)$ given by Eqn. 3.9, and $\mathbf{\Sigma}_n$ is the noise covariance.

3.2.2 Parameter Identification Using Bayes' Theorem

Since it is assumed that only N_m lower modes contribute significantly to the response, only the modal parameters corresponding to these modes are identified. Specifically, the parameter vector $\mathbf{a}$ for identification is comprised of: 1) the modal parameters $\omega_r, \zeta_r, r = 1, ..., N_m$ in Eqn. 3.4; 2) the modeshape components $\phi_j^{(r)}$ at the observed DOF $j = 1, \ldots, N_o$ for the modes $r = 1, \ldots, N_m$, except those elements which were used for the normalization of the modeshapes (which are assumed constant and equal to one); thus, a total of $N_m(N_o - 1)$ unknown modeshape parameters are to be identified; 3) the parameters prescribing the spectral density matrix $\mathbf{S}_g(\omega)$ and the modulation function $A(t)$ and 4) the elements of the upper right triangular part of $\mathbf{\Sigma}_n$ (symmetry defines the lower triangular part of this matrix).

Recall that here the scaling of each modeshape is chosen such that one of its components corresponding to a measured DOF is equal to unity. However, such scaling is arbitrary and therefore the above vectors can be identified only up to a constant scaling factor. A different modeshape normalization will cause all identified

components of the r^{th} modeshape to be scaled by some constant c_r; at the same time the values of the elements $S_f^{(r,s)}$ of the modal forcing spectral density matrix will be scaled by $(c_r c_s)^{-1}$.

Let the vector $\mathbf{Y}_{m,p}$ denote the zero-mean random vector comprised of the response measurements from time $m\Delta t$ to $p\Delta t$ $(m \leq p)$ in a time-descending order, that is,

$$\mathbf{Y}_{m,p} = [\mathbf{y}^T(p) \cdots \mathbf{y}^T(m)]^T, \qquad m \leq p \tag{3.14}$$

Using Bayes' theorem, the expression for the updated PDF of the parameters $\mathbf{a}$ given some measured response $\mathbf{Y}_{1,N}$ is

$$p(\mathbf{a}|\mathbf{Y}_{1,N}) = c_2 p(\mathbf{a}) p(\mathbf{Y}_{1,N}|\mathbf{a}) \tag{3.15}$$

where c_2 is a normalizing constant such that the integral of the right-hand side of Eqn. 3.15 over the domain of $\mathbf{a}$ is equal to unity. The factor $p(\mathbf{a})$ in Eqn. 3.15 denotes the prior PDF of the parameters and is based on previous knowledge or engineering judgement; in the case where no prior information is available, this is treated as a constant. $p(\mathbf{Y}_{1,N}|\mathbf{a})$ is the dominant factor in the right-hand side of Eqn. 3.15 reflecting the contribution of the measured data in establishing the posterior distribution. This can be expanded into a product of conditional probabilities as follows:

$$p(\mathbf{Y}_{1,N}|\mathbf{a}) = p(\mathbf{Y}_{1,N_p}|\mathbf{a}) \prod_{k=N_p+1}^{N} p(\mathbf{y}(k)|\mathbf{a}; \mathbf{Y}_{1,k-1}) \tag{3.16}$$

In order to improve computational efficiency, the following approximation is introduced:

$$p(\mathbf{Y}_{1,N}|\mathbf{a}) \simeq p(\mathbf{Y}_{1,N_p}|\mathbf{a}) \prod_{k=N_p+1}^{N} p(\mathbf{y}(k)|\mathbf{a}; \mathbf{Y}_{k-N_p,k-1}) \tag{3.17}$$

The conditional probability factors depending on more than N_p previous data points are approximated by conditional probabilities depending on only the last N_p data

points. The sense of this approximation is that data points belonging too far in the past do not have a significant effect on the statistical behavior of the present point. Of course, one expects this to be true, especially if N_p is so large that all the correlation functions have decayed to very small values. However, it is found that a value for N_p of the order of $\frac{T_1}{\Delta t}$ is sufficient, where T_1 is the fundamental period of the system and Δt is the sampling time step. For example, assuming a time step $\Delta t = \frac{1}{25}T_1$, it follows that a value of $N_p \approx 25$ is sufficient. The explanation for this behavior can be understood with the following simple example. Consider three random variables x, y and z and assume that one is interested in the conditional probability $p(x|y,z)$ (so $p(z|y) > 0$). Obviously, if x is independent of z given y, one can write $p(x|y,z) = p(x|y)$ $\Big($because in general, $p(x,z|y) = p(x|y,z)p(z|y)$ but for independence, $p(x,z|y) = p(x|y)p(z|y)\Big)$. Now, let x be dependent on z. If y and z are fully dependent, then one can still write the above equation $p(x|y,z) = p(x|y)$. If y and z are almost fully dependent, then this equation still holds approximately. The point of this example is that when considering conditional probabilities, some of the conditioning information may be redundant and can be omitted without significantly affecting accuracy. This argument can be applied to our case since measurements one period apart are highly correlated. Using larger values of N_p leads to significant increase of the computational effort without significantly further improving the accuracy of the identification. This was verified by numerous simulations.

The factor $p(\mathbf{Y}_{1,N_p}|\mathbf{a})$ follows an N_oN_p-variate Gaussian distribution with zero mean and covariance matrix $\mathbf{\Sigma}_{Y,N_p}$:

$$\begin{aligned}\mathbf{\Sigma}_{Y,N_p} &= E[\mathbf{Y}_{1,N_p}\mathbf{Y}_{1,N_p}^T] \\ &= \begin{bmatrix} \mathbf{\Gamma}_{N_p,N_p} & \cdots & \mathbf{\Gamma}_{N_p,1} \\ \vdots & \ddots & \vdots \\ \mathbf{\Gamma}_{1,N_p} & \cdots & \mathbf{\Gamma}_{1,1} \end{bmatrix}\end{aligned} \tag{3.18}$$

where each of the submatrices $\mathbf{\Gamma}_{m,p}, 1 \leq m,p \leq N_p$, has dimension $N_o \times N_o$. Based

on Eqn. 3.13, the (j,l) element of the matrix $\mathbf{\Gamma}_{m,p}$ is given by

$$\begin{aligned}\Gamma_{m,p}^{(j,l)} &= E[y_j(m)y_l(p)] \\ &= \sum_{r,s=1}^{N_o} L_o^{(j,r)} L_o^{(l,s)} R_x^{(r,s)}(m\Delta t, p\Delta t) + \Sigma_n^{(j,l)}\delta_{m,p}\end{aligned} \tag{3.19}$$

where $\delta_{m,p}$ is the Kronecker Delta function, $R_x^{(j,l)}$ denotes the (j,l) element of the auto-correlation function $\mathbf{R}_x(t_1,t_2)$ of the model response $\mathbf{x}(t)$ given by Eqn. 3.9, and $\Sigma_n^{(j,l)}$ is the (j,l) element of the noise covariance matrix defined in Eqn. 3.12.

Therefore, the joint probability distribution $p(\mathbf{Y}_{1,N_p}|\mathbf{a})$ is given by

$$p(\mathbf{Y}_{1,N_p}|\mathbf{a}) = \frac{1}{(2\pi)^{\frac{N_o N_p}{2}}|\mathbf{\Sigma}_{Y,N_p}|^{\frac{1}{2}}} \exp(-\frac{1}{2}\mathbf{Y}_{1,N_p}^T \mathbf{\Sigma}_{Y,N_p}^{-1}\mathbf{Y}_{1,N_p}) \tag{3.20}$$

Next, the general expression for the conditional probability involving α previous points $p(\mathbf{y}(k)|\mathbf{a};\mathbf{Y}_{k-\alpha,k-1})$ in Eqn. 3.17 is derived, where it is assumed that $k > \alpha \geq 1$. First, note that the covariance matrix $\mathbf{\Sigma}_{\mathbf{Y}_{k-\alpha,k}}$ of the random vector $\mathbf{Y}_{k-\alpha,k}$ is given by

$$\begin{aligned}\mathbf{\Sigma}_{\mathbf{Y}_{k-\alpha,k}} &= E[\mathbf{Y}_{k-\alpha,k}\mathbf{Y}_{k-\alpha,k}^T] \\ &= \begin{bmatrix} \mathbf{\Gamma}_{k,k} & \cdots & \mathbf{\Gamma}_{k,k-\alpha} \\ \vdots & \ddots & \vdots \\ \mathbf{\Gamma}_{k-\alpha,k} & \cdots & \mathbf{\Gamma}_{k-\alpha,k-\alpha} \end{bmatrix}\end{aligned} \tag{3.21}$$

where each of the submatrices $\mathbf{\Gamma}_{m,p}, k-\alpha \leq m,p \leq k$, is given by Eqn. 3.19.

Next, the matrix $\mathbf{\Sigma}_{\mathbf{Y}_{k-\alpha,k}}$ is partitioned as follows:

$$\mathbf{\Sigma}_{\mathbf{Y}_{k-\alpha,k}} = \begin{bmatrix} \mathbf{\Sigma}_{11}(k,\alpha) & \mathbf{\Sigma}_{12}(k,\alpha) \\ \mathbf{\Sigma}_{12}^T(k,\alpha) & \mathbf{\Sigma}_{22}(k,\alpha) \end{bmatrix} \tag{3.22}$$

where $\mathbf{\Sigma}_{11}(k,\alpha)$, $\mathbf{\Sigma}_{12}(k,\alpha)$ and $\mathbf{\Sigma}_{22}(k,\alpha)$ have dimensions $N_o \times N_o$, $N_o \times N_o\alpha$ and $N_o\alpha \times N_o\alpha$, respectively.

Since the measured response is assumed to have zero mean, the best estimator

$\mathbf{e}_\alpha(k)$ of $\mathbf{y}(k)$ given $\mathbf{Y}_{k-\alpha,k-1}$ $(k > \alpha)$ is (Brockwell and Davis 1991)

$$\begin{aligned}\mathbf{e}_\alpha(k) &\equiv E[\mathbf{y}(k)|\mathbf{Y}_{k-\alpha,k-1}] \\ &= \mathbf{\Sigma}_{12}(k,\alpha)\mathbf{\Sigma}_{22}^{-1}(k,\alpha)\mathbf{Y}_{k-\alpha,k-1}\end{aligned} \tag{3.23}$$

and the covariance matrix $\mathbf{\Sigma}_{\epsilon,\alpha}(k)$ of the prediction error $\boldsymbol{\epsilon}_\alpha(k) = \mathbf{y}(k) - \mathbf{e}_\alpha(k)$ is given by

$$\begin{aligned}\mathbf{\Sigma}_{\epsilon,\alpha}(k) &\equiv E[\boldsymbol{\epsilon}_\alpha(k)\boldsymbol{\epsilon}_\alpha^T(k)] \\ &= \mathbf{\Sigma}_{11}(k,\alpha) - \mathbf{\Sigma}_{12}(k,\alpha)\mathbf{\Sigma}_{22}^{-1}(k,\alpha)\mathbf{\Sigma}_{12}^T(k,\alpha)\end{aligned} \tag{3.24}$$

In conclusion, the conditional probability $p(\mathbf{y}(k)|\mathbf{a};\mathbf{Y}_{k-\alpha,k-1})$ follows an N_o-variate Gaussian distribution with mean $\mathbf{e}_\alpha(k)$ given by Eqn. 3.23 and covariance matrix $\mathbf{\Sigma}_{\epsilon,\alpha}(k)$ given by Eqn. 3.24:

$$p(\mathbf{y}(k)|\mathbf{a};\mathbf{Y}_{k-\alpha,k-1}) = \frac{1}{(2\pi)^{\frac{N_o}{2}}|\mathbf{\Sigma}_{\epsilon,\alpha}(k)|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}[\mathbf{y}(k)-\mathbf{e}_\alpha(k)]^T\mathbf{\Sigma}_{\epsilon,\alpha}^{-1}(k)[\mathbf{y}(k)-\mathbf{e}_\alpha(k)]\right\} \tag{3.25}$$

The proposed modal identification approach can be summarized as follows: Eqn. 3.15 is utilized with $p(\mathbf{Y}_{1,N}|\mathbf{a})$ being calculated through the approximation in Eqn. 3.17. The factor $p(\mathbf{Y}_{1,N_p}|\mathbf{a})$ can be calculated using Eqn. 3.20 along with Eqn. 3.18 and 3.19 and each conditional probability factor in Eqn. 3.17 can be calculated from Eqn. 3.25 along with Eqn. 3.21 - 3.24.

The most probable parameter values $\hat{\mathbf{a}}$ are obtained by minimizing $J(\mathbf{a}) = -\ln[p(\mathbf{a}|\mathbf{Y}_{1,N})]$. It is found that the updated PDF of the parameters $\mathbf{a}$ can be well approximated by a Gaussian distribution $N(\hat{\mathbf{a}}, \mathbf{H}(\hat{\mathbf{a}})^{-1})$ with mean $\hat{\mathbf{a}}$ and covariance matrix $\mathbf{H}(\hat{\mathbf{a}})^{-1}$, where $\mathbf{H}(\hat{\mathbf{a}})$ denotes the Hessian of $J(\mathbf{a})$ calculated at $\mathbf{a} = \hat{\mathbf{a}}$.

Although the above formulation was presented for the particular case where the measured response is assumed to consist of displacement histories, it can be easily modified to treat velocity or acceleration measurements by using the corresponding modal impulse response functions for velocity or acceleration in Eqn. 3.9.

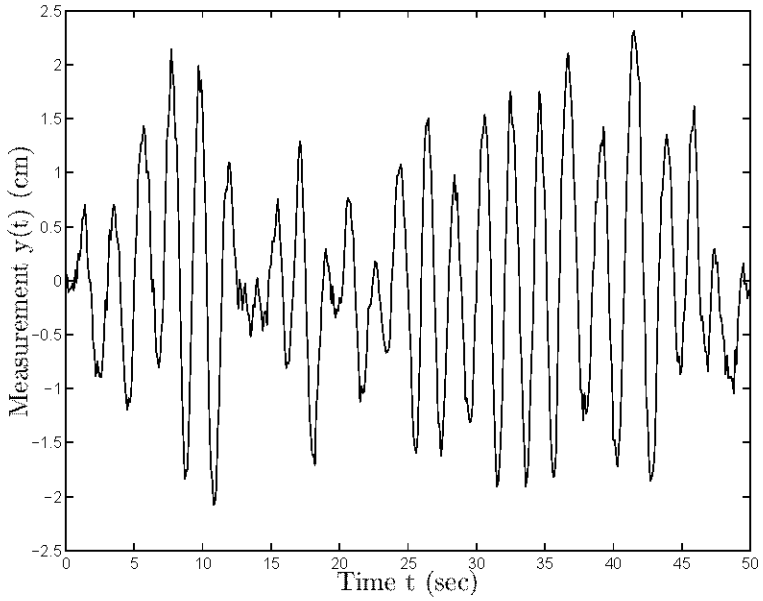

Figure 3.1: Measured time history (Example 3-1).

Note that if the right-hand side of Eqn. 3.2 is replaced by $\sum_{k=1}^{K} A_k(t|\boldsymbol{\theta}_A)\mathbf{g}_k(t)$, the proposed methodology can handle excitations having different modulation functions, e.g., ambient vibrations with a series of wind gusts.

3.3 Numerical Examples

3.3.1 Example 3-1: Transient Response of SDOF Linear Oscillator

In this example, the identification of a SDOF system from simulated noisy transient displacement response data shown in Fig. 3.1 is considered. Here, in Eqn. 3.4, $A(t) = U(t)$, the Heaviside unit step function, and $f(t)$ is white-noise with spectral intensity S_{fo}. The parameters $\tilde{\mathbf{a}} = [\tilde{\omega}_o, \tilde{\zeta}, \tilde{S}_{fo}, \tilde{\sigma}_n]^T$ used to generate the simulated data are: $\tilde{\omega}_o = 3.0$ rad/sec, $\tilde{\zeta} = 0.04$, $\tilde{S}_{fo} = 1.0\ \text{cm}^2\,\text{sec}^{-3}$ and $\tilde{\sigma}_n = 0.0959\ \text{cm}^2$. The chosen value of $\tilde{\sigma}_n$ corresponds to a 10% rms prediction-error level, i.e., the noise is 10% of the rms of the noise-free response. The time step used to generate the data is 0.01 sec. However, a much larger sampling time step was chosen ($\Delta t = 0.1$ sec) and the total time interval is $T = 50$ sec, so that the number of data points is $N = 500$.

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	Standard Deviation σ	$\lvert\frac{\sigma}{\hat{a}}\rvert$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
ω_o	3.0000	2.9525	0.0585	0.020	0.81
ζ	0.0400	0.0566	0.0210	0.371	0.79
S_{fo}	1.0000	0.8543	0.1168	0.137	1.25
σ_n	0.0959	0.0891	0.0038	0.042	1.81

Table 3.1: Identification results for one set of data and $N_p = 20$ (Example 3-1).

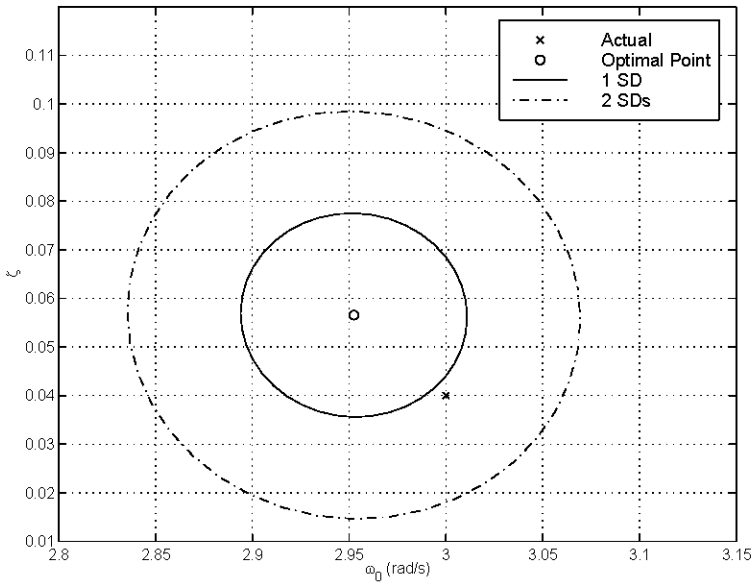

Figure 3.2: Contour of the updated joint PDF of frequency ω_o and damping ratio ζ (Example 3-1).

Table 3.1 refers to the identification results using a single set of displacement measurements $\hat{\mathbf{Y}}_{1,N}$. It shows the most probable values $\hat{\mathbf{a}} = [\hat{\omega}_o, \hat{\zeta}, \hat{S}_{fo}, \hat{\sigma}_n]^T$, the calculated standard deviations σ_{ω_o}, σ_ζ, $\sigma_{S_{fo}}$ and σ_{σ_n}, the coefficient of variation for each parameter and the value of a "normalized error" β for each parameter. The parameter β represents the absolute value of the difference between the identified optimal value and exact value, normalized with respect to the corresponding calculated standard deviation. Here, the value $N_p = 20$ (corresponding to one period of the oscillator) was used in Eqn. 3.17. Repeating the identification with a value of $N_p = 40$ yielded identical results, verifying that using $N_p = \frac{T_1}{\Delta t}$ is sufficient.

Fig. 3.2 shows contours in the (ω_o, ζ) plane of the marginal updated PDF $p(\omega_o, \zeta|\hat{\mathbf{Y}}_{1,N})$

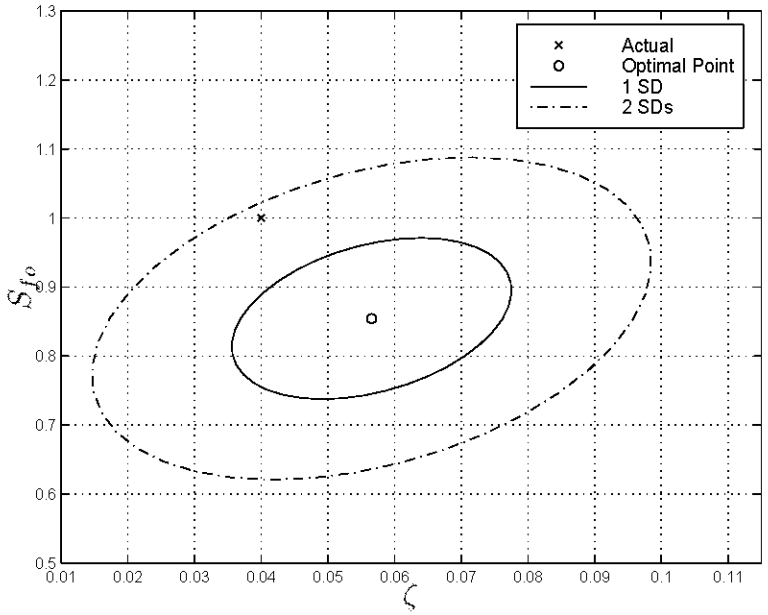

Figure 3.3: Marginal updated joint PDFs of the damping ratio ζ and the spectral intensity S_{fo} (Example 3-1).

calculated for the set of simulated data used for Table 3.1. Fig. 3.3 shows contours in the (ζ, S_{fo}) plane of the marginal updated PDF $p(\zeta, S_{fo}|\hat{\mathbf{Y}}_{1,N})$ calculated for the set of simulated data used for Table 3.1. One can see that the estimates of the damping ratio and the spectral intensity are quite correlated, as expected, because a larger value of the spectral intensity with a larger value of the damping ratio corresponds to a similar autocorrelation function and hence a similar probability given the data. On the contrary, as seen in Fig. 3.2, the estimates of ω_o and ζ can be considered as being uncorrelated.

Fig. 3.4 shows a comparison between the conditional PDFs $p(\omega_o|\hat{\mathbf{Y}}_{1,N}, \hat{\zeta}, \hat{S}_{fo}, \hat{\sigma}_n)$ and $p(\zeta|\hat{\mathbf{Y}}_{1,N}, \hat{\omega}_o, \hat{S}_{fo}, \hat{\sigma}_n)$, respectively, obtained from: i) Eqn. 3.15 (crosses) and ii) the Gaussian approximation $N(\hat{\mathbf{a}}, \mathbf{H}(\hat{\mathbf{a}})^{-1})$ described in Section 2.3.2 (solid line). It can be seen that the proposed Gaussian approximation is very accurate. Thus, the inverse Hessian matrix $\mathbf{H}(\hat{\mathbf{a}})^{-1}$ can be used to calculate the covariance matrix for the uncertainty in the value of the parameter $\mathbf{a}$, given the data $\hat{\mathbf{Y}}_{1,N}$; in particular, this gives the variance $\sigma^2(\mathrm{a}_i|\hat{\mathbf{Y}}_{1,N})$ for each parameter a_i of $\mathbf{a}$.

Next, one hundred sets of independent time histories were generated using the

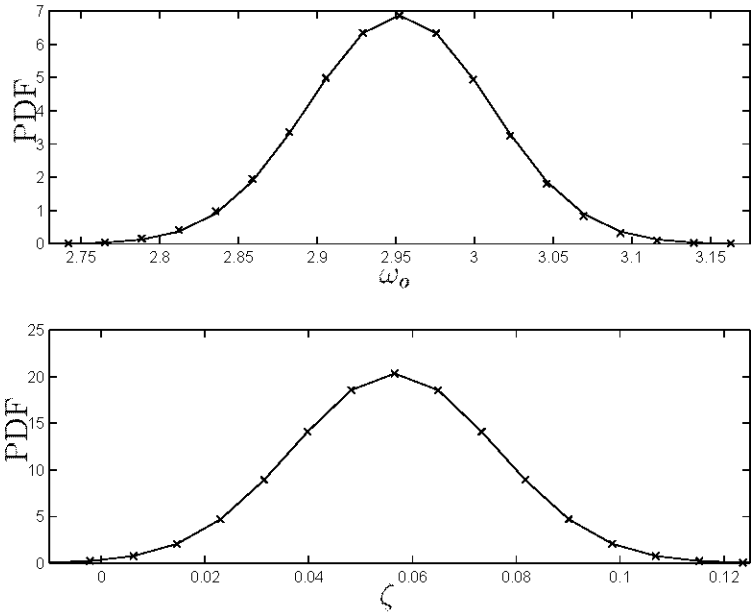

Figure 3.4: Conditional PDFs of the natural frequency and damping ratio obtained from: i) Eqn. 3.15 - **cross**; and ii) Gaussian approximations - **solid**. The remaining parameters are fixed at their optimal values (Example 3-1).

Parameter	Actual $\tilde{a}$	$\bar{\hat{a}}$	var($\hat{a}$)	$\bar{\sigma^2}$	$\bar{\beta^2}$
ω_0	3.0000	3.0027	0.0037	0.0039	0.9596
ζ	0.0400	0.0409	0.0004	0.0004	1.0610
S_{f0}	1.0000	0.9932	0.0164	0.0168	0.9767
σ_n	0.0959	0.0964	1.52×10^{-5}	1.68×10^{-5}	0.9573

Table 3.2: Identification results using 100 sets of data and $N_p = 20$ (Example 3-1).

same parameters as discussed in the beginning of this example. The optimal (most probable) parameter values $\hat{\mathbf{a}}^{(m)}, m = 1, \ldots, 100$ using each set of data were calculated separately. Then, the mean value and the covariance matrix of the optimal parameters were calculated from the set $\{\hat{\mathbf{a}}^{(m)}, m = 1, \ldots, 100\}$. The obtained mean values and variances of the optimal parameters are shown in the third and fourth columns, respectively, of Table 3.2. The fifth column in this table shows the mean value of the one hundred different variances where each variance is calculated using the inverse Hessian matrix $\mathbf{H}(\hat{\mathbf{a}})^{-1}$ derived from each set of data separately. Finally, based on the 100 samples, the mean square values of the normalized error parameter β, described

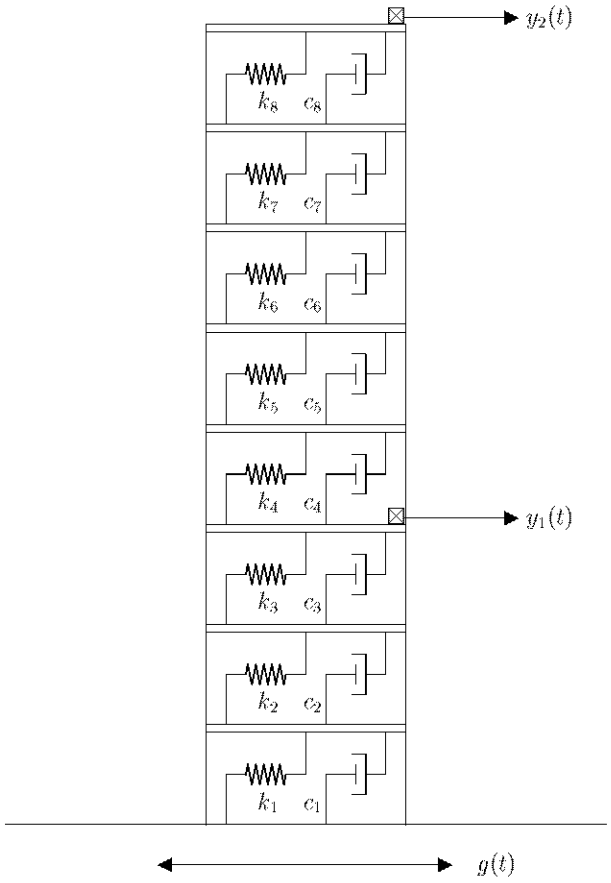

Figure 3.5: Eight-story shear building model (Example 3-2).

earlier for Table 3.1, are shown in the last column. It can be seen that the fourth and the fifth columns look similar, implying that the uncertainties calculated from a single sample are representative of the uncertainties of the optimal parameters obtained from several independent sets of data of equal length. Furthermore, the values in the last column are all approximately equal to unity. This verifies that the calculated uncertainties from our proposed approach using one set of data are reasonable and representative of the true uncertainties in the identification process.

3.3.2 Example 3-2: Eight-story Shear Building Subjected to Non-stationary Ground Excitation

The second example uses simulated response data from the shear building shown in Fig. 3.5. It is assumed that this building has a uniformly distributed floor mass

and story stiffness over its height and the stiffness to mass ratio is chosen to be 1160 sec^{-2} so that the first four modal frequencies are 2.0000 Hz, 5.9318 Hz, 9.6616 Hz and 13.0624 Hz. The damping ratio is assumed to be 2% for all modes. It is assumed that the displacements at the 4^{th} and 9^{th} floor were measured over a time interval $T = 40\,\text{sec}$, using a sampling interval $\Delta t = \frac{1}{40}\,\text{sec}$. Therefore, the total number of measured time points is $N = 1600$ and corresponds to 80 fundamental periods. Note that a much smaller time interval ($\frac{1}{400}\,\text{sec}$) was used for the data simulation so that the signal contains high frequency content which simulates a realistic situation. The structure is assumed to be subjected to a base acceleration given by stationary white noise of spectral intensity $S_{go} = 0.25\ \text{m}^2\,\text{sec}^{-3}$ modulated by $A(t) = \frac{t}{t_m} e^{1-\frac{t}{t_m}} U(t)$, where $U(t)$ denotes the Heaviside unit step function. Note that the envelope function has its maximum at $t = t_m$ equal to unity. The measurement noise for the response is taken to be 10%, i.e., the rms of the measurement noise for a particular channel of measurement is equal to 10% of the rms of the noise-free response at the corresponding DOF. Modal identification using the proposed approach is carried out for the lowest three modes of the structure. A value of $N_p = 20$ was used which corresponds to using previous data points over one fundamental period as the conditioning information at each time step in Eqn. 3.17.

Fig. 3.6 shows the Fourier amplitude spectra of the displacements measurements at the 4^{th} and 9^{th} floor. Table 3.3 shows the identification results. The second column in this table corresponds to the actual values used for generation of the simulated measurement data; the third and fourth columns correspond to the identified optimal parameters and the corresponding standard deviations, respectively; the fifth column lists the coefficient of variation for each parameter; and the last column shows the normalized error β described in Example 3-1. The first group of rows in the table corresponds to modal frequencies, followed by the modal damping ratios, the ratios of the modeshape components between the 4^{th} and the 9^{th} floor, the elements of the modal forcing spectral matrix $\mathbf{S}_{fo}$, the elements of the prediction error covariance matrix $\mathbf{\Sigma}_n$ and finally the time of maximum input intensity. Note that in this case, $S_{fo}^{(j,l)} = \sqrt{S_{fo}^{(j,j)} S_{fo}^{(l,l)}}, j, l = 1, \ldots, N_m$. Therefore, only the diagonal elements of $\mathbf{S}_{fo}$

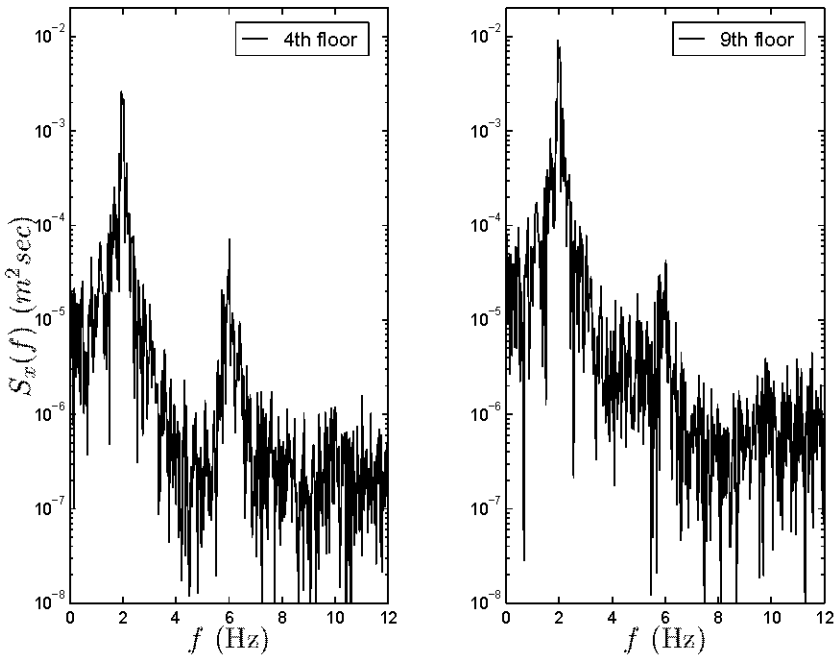

Figure 3.6: Displacement spectral density estimates for the 4^{th} and 9^{th} floor (Example 3-2).

are identified. The modeshapes are normalized so that the modeshape components at the 9^{th} floor are equal to unity for each of the modes considered.

It is worth noting that in all cases the coefficients of variation for the frequencies are much smaller than those of the damping ratios, indicating that frequencies are identified much better than dampings. An additional result observed, but not tabulated here, is that the modal damping ratios exhibit significant correlation with the corresponding modal forcing spectral intensities.

Fig. 3.7 shows the contours in the (ω_1, ω_2) plane of the marginal updated PDF of ω_1 and ω_2. One observes that in all cases the actual parameters are at reasonable distances, measured in terms of the estimated standard deviations, from the identified optimal parameters, i.e., the values of β are around zero to two. This shows the calculated uncertainties are consistent.

Fig. 3.8 is a typical plot showing comparisons between the conditional PDFs of ω_1 and ω_2 (keeping all other parameters fixed at their optimal values) obtained from: i) Eqn. 3.15 (crosses) and ii) the Gaussian approximation $N(\hat{\mathbf{a}}, \mathbf{H}(\hat{\mathbf{a}})^{-1})$ described

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	S.D. σ	$\lvert\frac{\sigma}{\hat{a}}\rvert$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
ω_1	2.0000	1.9903	0.0135	0.007	0.71
ω_2	5.9318	5.9507	0.0257	0.004	0.74
ω_3	9.6616	9.8400	0.1460	0.015	1.22
ζ_1	0.0200	0.0200	0.0062	0.312	0.01
ζ_2	0.0200	0.0199	0.0029	0.145	0.04
ζ_3	0.0200	0.0362	0.0144	0.398	1.12
$\phi_4^{(1)}/\phi_9^{(1)}$	0.5287	0.5312	0.0020	0.004	1.25
$\phi_4^{(2)}/\phi_9^{(2)}$	-1.0353	-1.0653	0.0575	0.054	0.52
$\phi_4^{(3)}/\phi_9^{(3)}$	0.4035	0.3731	0.1310	0.351	0.23
$S_{fo}^{(1,1)}$	0.3996	0.3856	0.0335	0.087	0.42
$S_{fo}^{(2,2)}$	0.0396	0.0345	0.0047	0.119	1.09
$S_{fo}^{(3,3)}$	0.0112	0.0227	0.0127	1.136	0.91
$\sigma_n^{(1)}$ (4th floor)	0.0035	0.0036	0.0001	0.028	0.37
$\sigma_n^{(2)}$ (9th floor)	0.0067	0.0066	0.0002	0.023	0.49
t_m	10.0000	10.1049	0.2991	0.030	0.35

Table 3.3: Identification results for the eight-story shear building using nonstationary approach (Example 3-2).

at the end of Section 2.3.2 (solid line). It can be seen that the proposed Gaussian approximation is very accurate.

Another identification was performed using absolute acceleration measurements with the same structure and the same excitation. The identification results are summarized in Table 3.4. Again, the proposed approach successfully identified the first three modes of the structure. Furthermore, the actual parameters are at reasonable distances from the optimal parameters, compared to the calculated standard deviations.

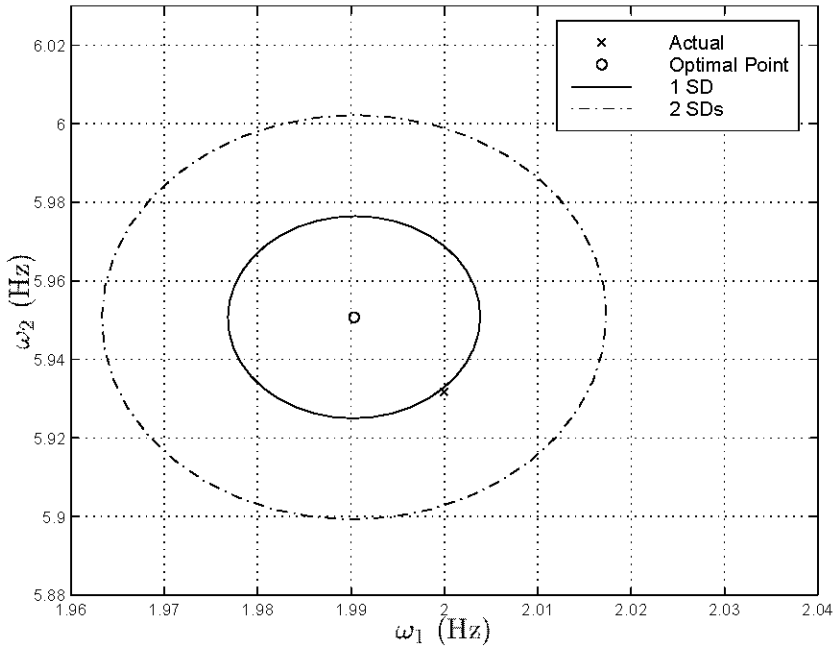

Figure 3.7: Marginal updated joint PDF of natural frequencies ω_1 and ω_2 (Example 3-2).

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	S.D. σ	$\lvert\frac{\sigma}{\tilde{a}}\rvert$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
ω_1	2.0000	2.0095	0.0182	0.009	0.52
ω_2	5.9318	5.9612	0.0252	0.004	1.17
ω_3	9.6616	9.7049	0.0224	0.002	1.93
ζ_1	0.0200	0.0167	0.0020	0.099	1.67
ζ_2	0.0200	0.0239	0.0036	0.179	1.09
ζ_3	0.0200	0.0224	0.0012	0.061	1.97
$\phi_4^{(1)}/\phi_9^{(1)}$	0.5287	0.5268	0.0086	0.016	0.22
$\phi_4^{(2)}/\phi_9^{(2)}$	-1.0353	-1.0385	0.0062	0.006	0.52
$\phi_4^{(3)}/\phi_9^{(3)}$	0.4035	0.3970	0.0135	0.033	0.49
$S_{fo}^{(1,1)}$	0.3996	0.4482	0.0499	0.125	0.97
$S_{fo}^{(2,2)}$	0.0396	0.0433	0.0028	0.071	1.34
$S_{fo}^{(3,3)}$	0.0112	0.0105	0.0004	0.035	1.86
$\sigma_n^{(1)}$ (4th floor)	0.0137	0.0139	0.0004	0.026	0.72
$\sigma_n^{(2)}$ (9th floor)	0.0096	0.0090	0.0003	0.029	2.18
t_m	10.0000	10.3907	0.2345	0.024	1.67

Table 3.4: Identification results for the eight-story shear building using nonstationary approach with acceleration measurements (Example 3-2).

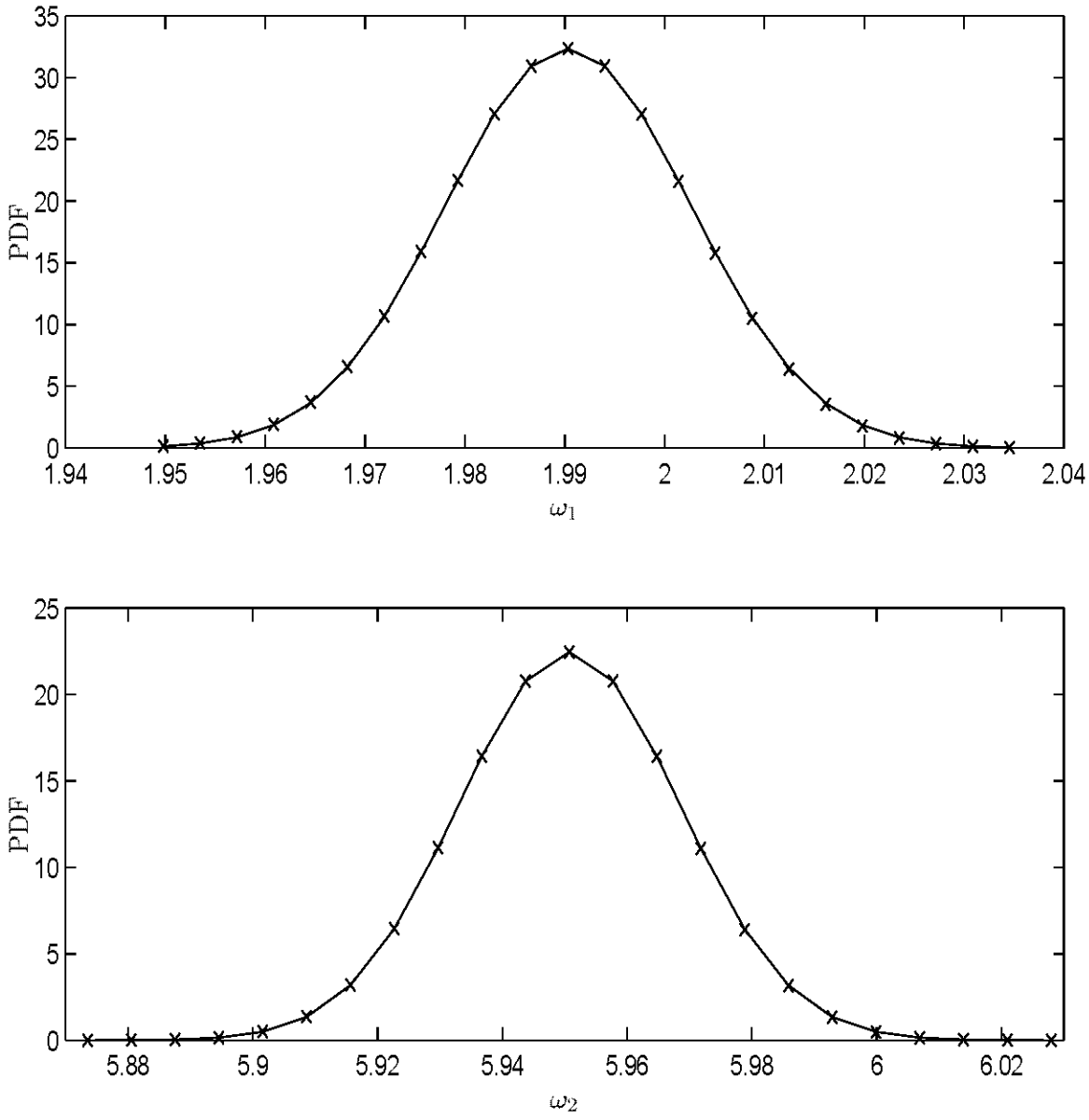

Figure 3.8: Conditional PDFs of the lower two natural frequencies obtained from: i) Eqn. 3.15 - **cross**; and ii) Gaussian approximations - **solid**. The remaining parameters are fixed at their optimal values (Example 3-2).

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	S.D. σ	$\lvert\frac{\sigma}{\hat{a}}\rvert$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
ω_1	2.0000	1.9759	0.0135	0.007	1.79
ω_2	5.9318	6.0226	0.0141	0.002	6.46
ω_3	9.6616	8.0513*	N.A.	N.A.	N.A.
ζ_1	0.0200	0.0199	0.0069	0.347	0.01
ζ_2	0.0200	0.0020*	N.A.	N.A.	N.A.
ζ_3	0.0200	0.7000*	N.A.	N.A.	N.A.
$\phi_4^{(1)}/\phi_9^{(1)}$	0.5287	0.5312	0.0023	0.004	1.08
$\phi_4^{(2)}/\phi_9^{(2)}$	-1.0353	-1.1162	0.0695	0.062	1.17
$\phi_4^{(3)}/\phi_9^{(3)}$	0.4035	0.0035*	N.A.	N.A.	N.A.
$S_{fo}^{(1,1)}$	0.3996	0.1235	0.0034	0.028	81.37
$S_{fo}^{(2,2)}$	0.0396	0.0088	0.0013	0.150	23.10
$S_{fo}^{(3,3)}$	0.0112	0.0900*	N.A.	N.A.	N.A.
$\sigma_n^{(1)}$ (4th floor)	0.0035	0.0042	0.0001	0.024	7.39
$\sigma_n^{(2)}$ (9th floor)	0.0067	0.0084	0.0002	0.018	11.40

Table 3.5: Identification results for the eight-story shear building using stationary approach (Example 3-2).

Importance of explicitly treating the response as nonstationary The same set of data for Example 3-2 was analyzed assuming that the response is stationary, i.e., using fixed $A(t) = 1, \forall\, t \in [0, T]$ during the identification. Results are shown in Table 3.5. The identified values which are marked with an asterisk ('$*$') do not converge and hit the boundaries of the optimization. For example, none of the parameters corresponding to the 3^{rd} mode converge. This is not surprising since Fig. 3.6 shows that the Fourier spectrum does not have any obvious peak at the frequency of the third mode of the structure. Furthermore, there are many model parameters with β values much larger than unity (Table 3.5) implying that the estimation of such parameters is biased. Therefore, consideration of the nonstationarity of the response is important.

3.4 Conclusion

A Bayesian time-domain approach for identification of the modal parameters and stochastic excitation parameters of MDOF linear systems using nonstationary noisy response data was presented. The updated PDF of the parameters can be accurately approximated by a multi-variate Gaussian distribution. The calculated mean and covariance matrix of this distribution offer an estimate of the most probable values of the parameters and their associated uncertainties. The uncertainties in the identified modal parameters are useful, for example, if one plans to proceed with the updating of a theoretical finite element model.

The presented methodology simultaneously utilizes the response histories at all measured DOFs, although only one observed degree of freedom is necessary to identify the modal frequencies and damping ratios. The approach proceeds without any difficulty by directly using the noisy measured response data. The calculation of the uncertainties does not require calculating parameter estimates from a number of different data sets and then calculating the statistics of these estimates. Instead, it follows directly from the methodology applied to a single set of measurements. Finally, the proposed methodology is expected to lead to improved modal identification using ambient vibration data where nonstationarity is evident.

Chapter 4 Updating Properties of Nonlinear Dynamical Systems with Uncertain Input

4.1 Overview

A spectral density approach is presented for the identification of nonlinear dynamical systems using only incomplete noisy response measurements. A stochastic model is used for the uncertain input and a Bayesian probabilistic approach is used to quantify the uncertainties in the model parameters. The proposed spectral-based approach utilizes important statistical properties of the Fast Fourier Transform and their robustness with respect to the probability distribution of the response signal in order to calculate the updated probability density function for the parameters of a nonlinear model conditional on the measured response. This probabilistic approach is well suited for the identification of nonlinear systems and does not require huge amounts of dynamic data. The formulation is first presented for single-degree-of-freedom systems and then for multiple-degree-of freedom systems. Examples using simulated data for a Duffing oscillator, an elasto-plastic system and a four-story yielding structure are presented to illustrate the proposed approach.

4.2 Introduction

Roberts et al. (1995) introduces a spectral method for identification of single-degree-of-freedom nonlinear dynamical systems using response measurements only. It was found that the parameters estimated from a single set of response measurement might be very unreliable. For example, consider a Duffing oscillator with linear

damping with random excitation:

$$m\,\ddot{x} + c\,\dot{x} + k_1 x + k_3 x^3 = f(t) \tag{4.1}$$

Assume that a group of many sets of response measurements corresponding to the same level of excitation are available and identification is performed for each of these sets. Fig. 4.1(a) shows the distribution for these estimates in the (k_1, k_3) plane schematically. It can be seen that it is unable to give an "optimal" estimation for k_1 and k_3 since these individual estimates are very scattered. Note that the slope of the best fitting line is approximately $-3\sigma_x^2$ because the equivalent linear system has linear stiffness $k_1 + 3\sigma_x^2 k_3$, where σ_x is the standard deviation of the structural response.

Therefore, Roberts et al. (1995) suggested that if another group of data, which corresponds to another level of excitation, can be obtained and identification is performed for each of these sets (circles in Fig. 4.1(b)). Then, least squares fit can be performed for the two groups of data. Finally, the optimal parameters can be obtained by finding the intersection of the two lines.

One of the main drawbacks of this approach is that it requires huge amount of data. First, many sets of data are needed for the least squares fit. Second, all sets of data corresponding to the same group have to correspond to the same level of excitation. Another main drawback is that the proposed approach gives equal weighting to the two groups of data. However, they might correspond to different number of data sets, different duration of observation and/or different level of noise, etc. In order to overcome these difficulties, a probabilistic approach is introduced in this chapter. This approach requires a reasonable amount of data, e.g., it requires only two sets of data in this case. Furthermore, the weighting of different sets of data is taken care of automatically by the probabilistic framework. The associated uncertainty of the model parameters can be directly computed by the proposed approach.

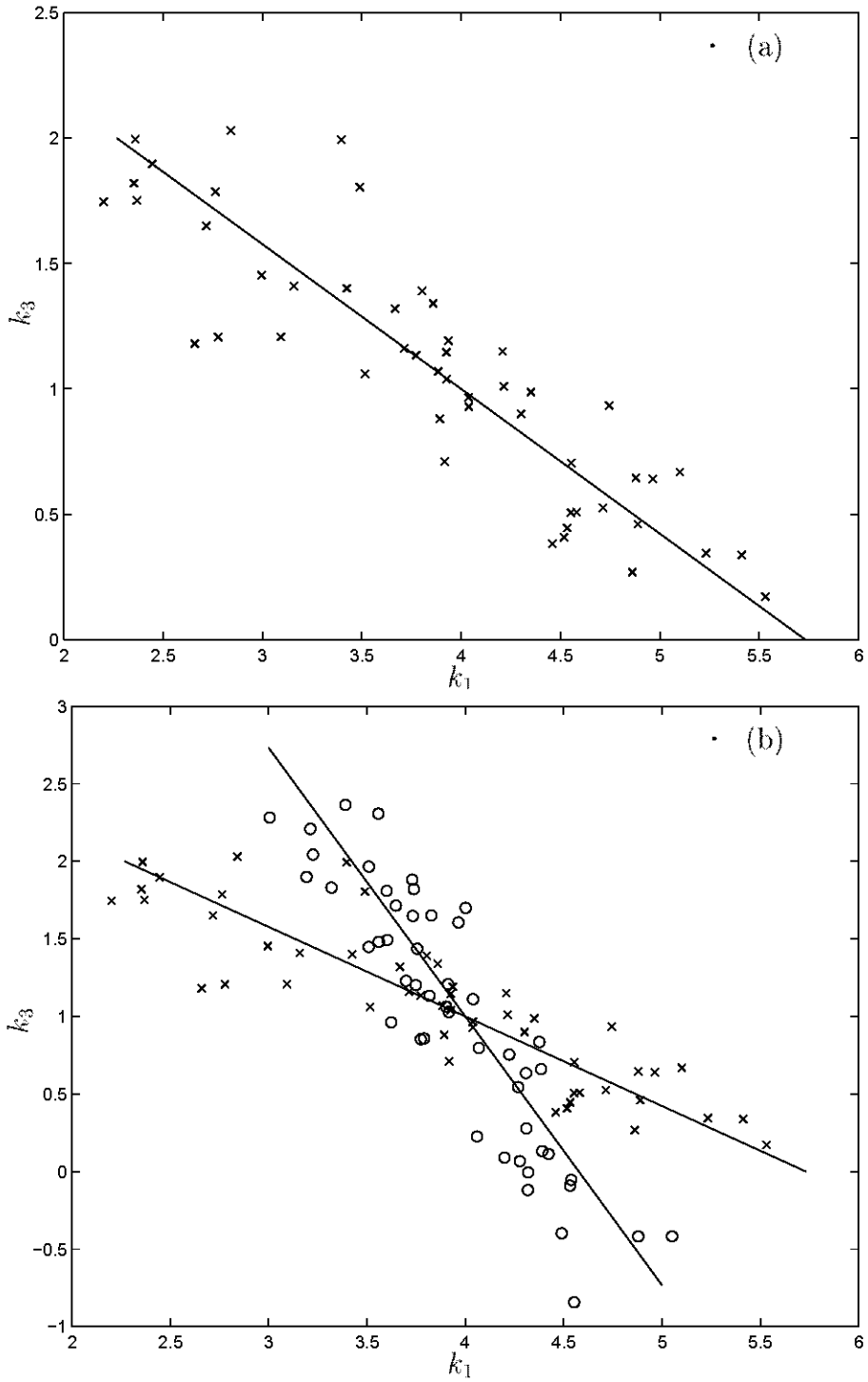

Figure 4.1: Schematic plots for identification of Duffing oscillator using the approach by Roberts et al. (1995): (a) Data from same level of excitation; and (b) Data from two different levels of excitation (+ as in (a) and o new level).

4.3 Single-degree-of-freedom Systems

4.3.1 Bayesian System Identification Formulation

Consider a structural or mechanical system whose displacement response x is modeled using a SDOF oscillator with equation of motion:

$$m\,\ddot{x} + f_s(x, \dot{x}; \boldsymbol{\theta}_s) = f(t) \tag{4.2}$$

where m, $\boldsymbol{\theta}_s$ and $f_s(x, \dot{x}; \boldsymbol{\theta}_s)$ are the mass, the model parameters and the nonlinear restoring force of the oscillator, respectively. Furthermore, the uncertain system input is modeled as a zero-mean stationary Gaussian random process f with spectral density function $S_f(\omega; \boldsymbol{\theta}_f)$, where $\boldsymbol{\theta}_f$ denotes the parameters of the stochastic process model for the excitation $f(t)$. The observed system response y is assumed to be stationary and is modeled by

$$y(t) = x(t) + \eta(t) \tag{4.3}$$

where the prediction error η accounts for modeling errors (differences between the system behavior and the model) as well as measurement noise. The uncertain prediction error is modeled as independent zero-mean Gaussian white noise, so

$$S_y(\omega) = S_x(\omega) + S_{\eta o} \tag{4.4}$$

where S_y, S_x and $S_{\eta o}$ are the spectral densities for the system response, model response and the prediction error. The spectral density function S_x, or the corresponding autocorrelation function R_x, can be approximated by equivalent linearization methods (Roberts and Spanos 1990; Lutes and Sarkani 1997) or by simulations.

Let $\hat{\mathbf{Y}}_N = [\hat{y}(0), \hat{y}(1), \ldots, \hat{y}(N-1)]^T$ denote a vector consisting of observed response data sampled at a time step Δt, where $\hat{y}(n) \equiv \hat{y}(n\Delta t)$, $n = 0, \ldots, N-1$. Herein, updating the uncertainty regarding the values of the model parameters $\mathbf{a} = [\boldsymbol{\theta}_s^T, \boldsymbol{\theta}_f^T, \sigma_{\eta o}]^T$ by using the data $\hat{\mathbf{Y}}_N$ is concerned, where $\sigma_{\eta o}^2 = \frac{2\pi}{\Delta t} S_{\eta o}$. From Bayes'

Theorem, the updated (posterior) PDF of the model parameters $\mathbf{a}$ given the data $\hat{\mathbf{Y}}_N$ is

$$p(\mathbf{a}|\hat{\mathbf{Y}}_N) = c_1 p(\mathbf{a}) p(\hat{\mathbf{Y}}_N|\mathbf{a}) \tag{4.5}$$

where c_1 is a normalizing constant and $p(\mathbf{a})$ denotes the prior PDF describing our initial belief about the uncertain parameter values. Note that $p(\mathbf{a}|\hat{\mathbf{Y}}_N)$ can be used to give the relative plausibility between two values of $\mathbf{a}$ based on measured data $\hat{\mathbf{Y}}_N$ which does not depend on the normalizing constant c_1. Also, the most probable value of $\mathbf{a}$, denoted by $\hat{\mathbf{a}}$ (the "optimal" parameter values), is given by maximizing $p(\mathbf{a})p(\hat{\mathbf{Y}}_N|\mathbf{a})$. For large N, $p(\hat{\mathbf{Y}}_N|\mathbf{a})$ is the dominant factor on the right-hand side of Eqn. 4.5.

A difficulty with implementing this approach is establishing the joint distribution $p(\hat{\mathbf{Y}}_N|\mathbf{a})$ for the response of the nonlinear system. Note that the response is not Gaussian distributed but the FFT of the response is approximately. This property is utilized to obtain a response PDF in the next section.

4.3.2 Bayesian Spectral Density Approach

Consider the stationary stochastic process $y(t)$ and the discrete estimator of its spectral density $S_y(\omega)$:

$$S_{y,N}(\omega_k) = \frac{\Delta t}{2\pi N} \left| \sum_{n=0}^{N-1} \exp(-i\omega_k n \Delta t) y(n) \right|^2 \tag{4.6}$$

where $\omega_k = k\Delta\omega$, $k = 0, \ldots, N_1 - 1$ with N_1 =INT$(N/2)$, $\Delta\omega = \frac{2\pi}{T}$, and $T = N\Delta t$. Here, INT denotes integer part. It can be shown that the estimator $S_{y,N}(\omega_k)$ is asymptotically unbiased, that is,

$$\lim_{N\to\infty} E[S_{y,N}(\omega_k)] = S_y(\omega_k) \tag{4.7}$$

where $E[.]$ denotes expectation (Yaglom 1987). However, for finite N, this estimator is biased. Calculating the expectation of the estimator in Eqn. 4.6 yields

$$E[S_{y,N}(\omega_k)] = \frac{\Delta t}{2\pi N} \sum_{n=0}^{N-1} \gamma_n R_x(n\Delta t) \cos(n\omega_k \Delta t) + S_{\eta o} \tag{4.8}$$

where R_x is the autocorrelation function of the response $x(t)$ and γ_n is given by

$$\begin{aligned} \gamma_n &= N, & n &= 0 \\ \gamma_n &= 2(N-n), & n &\geq 1 \end{aligned} \tag{4.9}$$

Note that the right-hand side of Eqn. 4.8 can be calculated using the FFT of the sequence $\gamma_n R_x(n\Delta t), n = 0, 1, \ldots, N-1$.

Based on the Central Limit Theorem, the real and imaginary part of the FFT are Gaussian distributed as $N \to \infty$. Therefore, the estimator $S_{y,N}(\omega_k)$, $k = 1, \ldots, N_1 - 1$, has the following asymptotic behavior:

$$\lim_{N\to\infty} S_{y,N}(\omega_k) = \frac{1}{2} S_y(\omega_k) \chi_2 \tag{4.10}$$

where χ_2 is a random variable having Chi-square distribution with two degrees of freedom (Yaglom 1987). Therefore, the PDF of the random variable $Y(\omega_k) = \lim_{N\to\infty} S_{y,N}(\omega_k)$ is asymptotically given by

$$p(Y(\omega_k)|\mathbf{a}) = \frac{1}{S_y(\omega_k)} \exp\left[-\frac{Y(\omega_k)}{S_y(\omega_k)}\right] \tag{4.11}$$

In the case of finite N, it can be shown using simulations that for $k << N_1$ the PDF of $S_{y,N}(\omega_k)$ can be accurately approximated by a Chi-square distribution in analogy to Eqn. 4.11 except that the mean $S_y(\omega_k)$ is replaced by $E[S_{y,N}(\omega_k)]$ given by Eqn. 4.8. Note that this approximation is very accurate even if $y(n\Delta t)$, $n = 1, \ldots, N$, is not Gaussian distributed. This is due to the robustness of the probability distribution of the FFT with respect to the probability distribution of the response signal.

Furthermore, it is shown in Appendix A that the random variables $S_{y,N}(\omega_k)$ and

$S_{y,N}(\omega_l)$ with $k \neq l$ and $k, l << N_1$, are uncorrelated asymptotically as $N \to \infty$. Note that uncorrelated Chi-square random variables are independent (Yaglom 1987). For large N, this property is approximately correct in a certain frequency range. In particular, for a sufficiently small number $K < N_1$, one can assume that the random vector $\mathbf{S}_{y,N}^K = [S_{y,N}(\omega_1), \ldots, S_{y,N}(\omega_K)]^T$ has all its elements approximately independently Chi-square distributed. Therefore, its joint PDF can be approximated as follows:

$$p(\mathbf{S}_{y,N}^K|\mathbf{a}) \simeq \prod_{k=1}^{K} \frac{1}{E[S_{y,N}(\omega_k)]} \exp\left(-\frac{S_{y,N}(\omega_k)}{E[S_{y,N}(\omega_k)]}\right) \tag{4.12}$$

In practice, ω_K can be chosen in range of $[1.5\omega_\rho, 2.0\omega_\rho]$ where ω_ρ is the frequency at which the peak of the spectral estimates $\hat{\mathbf{S}}_{y,N}(\omega_k)$ occurs. A more detailed discussion will be given in the numerical examples.

Given the observed data $\hat{\mathbf{Y}}_N$, one may substitute it in Eqn. 4.6 to calculate the corresponding observed spectral estimate $\hat{\mathbf{S}}_{y,N}^K = [\hat{S}_{y,N}(\omega_1), \ldots, \hat{S}_{y,N}(\omega_K)]^T$. Using Bayes' Theorem, the updated PDF of the model parameters $\mathbf{a}$ given the data $\hat{\mathbf{S}}_{y,N}^K$ follows from an analogy to Eqn. 4.5:

$$p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^K) = c_2 p(\mathbf{a}) p(\hat{\mathbf{S}}_{y,N}^K|\mathbf{a}) \tag{4.13}$$

where c_2 is a normalizing constant, and $p(\hat{\mathbf{S}}_{y,N}^K|\mathbf{a})$ is given by Eqn. 4.12 where each $S_{y,N}(\omega_k)$ is replaced by $\hat{S}_{y,N}(\omega_k)$, $E[S_{y,N}(\omega_k|\mathbf{a})]$ is calculated from Eqn. 4.8 and $R_x(n\Delta t) = R_x(n\Delta t|\mathbf{a})$ may be calculated by equivalent linearization methods or by simulation. The optimal parameters $\hat{\mathbf{a}}$ are obtained by minimizing an objective function $J(\mathbf{a}) = -\ln[p(\mathbf{a})p(\hat{\mathbf{S}}_{y,N}^K|\mathbf{a})]$. For the results in this chapter, this optimization is done using a MATLAB function 'fmins'.

In the case where several independent time histories $\hat{\mathbf{Y}}_N^{(1)}, \ldots, \hat{\mathbf{Y}}_N^{(M)}$ are available, the estimation can proceed by calculating the corresponding estimates $\hat{\mathbf{S}}_{y,N}^{K,(1)}, \ldots, \hat{\mathbf{S}}_{y,N}^{K,(M)}$

and then calculating the updated PDF

$$p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^{K,(1)},\ldots,\hat{\mathbf{S}}_{y,N}^{K,(M)}) = c_3 p(\mathbf{a}) \prod_{n=1}^{M} p(\hat{\mathbf{S}}_{y,N}^{K,(n)}|\mathbf{a}) \tag{4.14}$$

Note that in the proposed approach, each set of data can be corresponding to a different time duration T and different sampling time interval Δt and Eqn. 4.14 automatically takes care of the weighting for different sets of data.

4.4 Multiple-degree-of-freedom Systems

4.4.1 Model Formulation

Consider a system with N_d degrees of freedom (DOFs) and equation of motion:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{f}_s(\mathbf{x}, \dot{\mathbf{x}}; \boldsymbol{\theta}_s) = \mathbf{T}\mathbf{f}(t) \tag{4.15}$$

where $\mathbf{M} \in \mathbb{R}^{N_d \times N_d}$ is the mass matrix, $\mathbf{f}_s \in \mathbb{R}^{N_d}$ is the nonlinear restoring force characterized by the structural parameters $\boldsymbol{\theta}_s$, $\mathbf{T} \in \mathbb{R}^{N_d \times N_f}$ is a force distribution matrix and $\mathbf{f}(t) \in \mathbb{R}^{N_f}$ is an external excitation, e.g., force or ground acceleration, modeled by a stationary Gaussian process with zero mean and spectral density matrix function characterized by the excitation parameters $\boldsymbol{\theta}_f$:

$$\mathbf{S}_f(\omega) = \mathbf{S}_f(\omega; \boldsymbol{\theta}_f) \tag{4.16}$$

Assume now that discrete response data are available for $N_s (\leq N_d)$ observed DOFs. Let Δt denote the sampling time step. Because of measurement noise and modeling errors, the measured response $\mathbf{y}(n) \in \mathbb{R}^{N_s}$ (at time $t = n\Delta t$) will differ from the model response $\mathbf{q}(n)$, e.g., model displacement or model acceleration, calculated at the observed DOFs from Eqn. 4.15. This difference between the measured and model response, called prediction error, is modeled as a discrete zero-mean Gaussian

white noise vector process $\boldsymbol{\eta}(n) \in \mathbb{R}^{N_s}$ so

$$\mathbf{y}(n) = \mathbf{q}(n) + \boldsymbol{\eta}(n) \tag{4.17}$$

where the discrete process $\boldsymbol{\eta}$ is independent of $\mathbf{q}$ and satisfies

$$E[\boldsymbol{\eta}(n)\boldsymbol{\eta}^T(p)] = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}\delta_{np} \tag{4.18}$$

where $E[.]$ denotes expectation, δ_{np} denotes the Kronecker delta function, and $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ denotes the $N_s \times N_s$ covariance matrix of the prediction-error process $\boldsymbol{\eta}$.

Let $\mathbf{a}$ denote the parameter vector for identification; it includes the following parameters: 1) the structural parameters $\boldsymbol{\theta}_s$; 2) the excitation parameters $\boldsymbol{\theta}_f$; and 3) the elements of the upper right triangular part of $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ (symmetry defines the lower triangular part of this matrix). As in the SDOF case, Bayes' Theorem is applied to update the uncertainty regarding the values of the model parameters $\mathbf{a}$ based on the spectral density estimates.

4.4.2 Spectral Density Estimator and its Statistical Properties

Consider the stochastic vector process $\mathbf{y}(t)$ and a finite number of discrete data $\mathbf{Y}_N = \{\mathbf{y}(n), n = 0, \ldots, N-1\}$. Based on $\mathbf{Y}_N$, one can calculate the following discrete estimator of the spectral density matrix of the stochastic process $\mathbf{y}(t)$:

$$\mathbf{S}_{y,N}(\omega_k) = \boldsymbol{\mathcal{Y}}_N(\omega_k)\bar{\boldsymbol{\mathcal{Y}}}_N^T(\omega_k) \tag{4.19}$$

where $\bar{z}$ denotes the complex conjugate of a complex variable z and $\boldsymbol{\mathcal{Y}}_N(\omega_k)$ denotes the (scaled) Fourier Transform of the vector process $\mathbf{y}$ at frequency ω_k, as follows:

$$\boldsymbol{\mathcal{Y}}_N(\omega_k) = \sqrt{\frac{\Delta t}{2\pi N}} \sum_{n=0}^{N-1} \mathbf{y}(n) e^{-i\omega_k n \Delta t} \tag{4.20}$$

where $\omega_k = k\Delta\omega$, $k = 0, \ldots, N_1 - 1$ with $N_1 = \text{INT}(N/2)$, $\Delta\omega = \frac{2\pi}{T}$, and $T = N\Delta t$. Note that Eqn. 4.6 is a special case of Eqn. 4.19 and 4.20.

Using Eqn. 4.17 and taking expectation of Eqn. 4.19 (noting that $\mathbf{q}$ and $\boldsymbol{\eta}$ are independent) yields

$$E[\mathbf{S}_{y,N}(\omega_k)|\mathbf{a}] = E[\mathbf{S}_{q,N}(\omega_k)|\mathbf{a}] + E[\mathbf{S}_{\boldsymbol{\eta},N}(\omega_k)|\mathbf{a}] \tag{4.21}$$

where $\mathbf{S}_{q,N}(\omega_k)$ and $\mathbf{S}_{\boldsymbol{\eta},N}(\omega_k)$ are defined in a manner similar to that described by Eqn. 4.19 and 4.20. It easily follows from Eqn. 4.18 and 4.19 that

$$E[\mathbf{S}_{\boldsymbol{\eta},N}(\omega_k)|\mathbf{a}] = \frac{\Delta t}{2\pi}\boldsymbol{\Sigma}_{\boldsymbol{\eta}} \equiv \mathbf{S}_{\boldsymbol{\eta}0} \tag{4.22}$$

The term $E[\mathbf{S}_{q,N}(\omega_k)|\mathbf{a}]$ in Eqn. 4.21 can be also easily calculated by noting that $\mathbf{S}_{q,N}(\omega_k)$ has elements

$$S_{q,N}^{(j,l)}(\omega_k) = \frac{\Delta t}{2\pi N}\sum_{n,p=0}^{N-1} q_j(n)q_l(p)e^{-i\omega_k(n-p)\Delta t} \tag{4.23}$$

Grouping together terms having the same value of $(p - n)$ in Eqn. 4.23, and taking expectation, one obtains the following expression:

$$E[S_{q,N}^{(j,l)}(\omega_k)|\mathbf{a}] = \frac{\Delta t}{4\pi N}\sum_{n=0}^{N-1}\gamma_n[R_q^{(j,l)}(n\Delta t|\mathbf{a})e^{-i\omega_k n\Delta t} + R_q^{(j,l)}(-n\Delta t|\mathbf{a})e^{i\omega_k n\Delta t}] \tag{4.24}$$

where γ_n is given by Eqn. 4.9 and $R_q^{(j,l)}$ is the cross-correlation functions between the j^{th} and l^{th} component of the model quantity $\mathbf{q}$. However, it is usually not possible to obtain the correlation functions theoretically. In this case, for given $\mathbf{a}$, one can simulate samples of the response using Eqn. 4.15 and 4.16 and hence calculate their spectral density estimates in a similar manner to that described in Eqn. 4.19 and 4.20. Then, rather than using Eqn. 4.24, the expected values of the spectral estimates can be approximated by the average of the spectral density estimators obtained from the samples.

Next, the statistical properties of the estimator $\mathbf{S}_{y,N}(\omega_k)$ are discussed. Denote by $\mathcal{Y}_{N,R}(\omega_k)$ and $\mathcal{Y}_{N,I}(\omega_k)$ the real and imaginary part, respectively, of $\mathcal{Y}_N(\omega_k)$, that is, $\mathcal{Y}_N(\omega_k) = \mathcal{Y}_{N,R}(\omega_k) + i\mathcal{Y}_{N,I}(\omega_k)$. Since $\mathcal{Y}_N$ is a zero-mean Gaussian process, both $\mathcal{Y}_{N,R}(\omega_k)$ and $\mathcal{Y}_{N,I}(\omega_k)$, $k = 1, \ldots, N_1 - 1$ are zero-mean Gaussian vectors. Furthermore, in the limit when $N \to \infty$, the covariance matrix of the vector $[\mathcal{Y}^T_{N,R}(\omega_k), \mathcal{Y}^T_{N,I}(\omega_k)]^T$ has the form (Yuen 1999):

$$\mathbf{C}_N(\omega_k) = \begin{bmatrix} \mathbf{C}_{N,1}(\omega_k) & \mathbf{C}_{N,2}(\omega_k) \\ -\mathbf{C}_{N,2}(\omega_k) & \mathbf{C}_{N,1}(\omega_k) \end{bmatrix} \tag{4.25}$$

Eqn. 4.25 states that the real and imaginary part of $\mathcal{Y}_N(\omega_k)$ have equal covariance matrices $\mathbf{C}_{N,1}(\omega_k)$ for $k = 1, \ldots, N_1 - 1$, i.e., excluding the zero and Nyquist frequencies. Also, it states that the cross-covariance between the real and imaginary part has the property $\mathbf{C}^T_{N,2}(\omega_k) = -\mathbf{C}_{N,2}(\omega_k)$, i.e., $E[\mathcal{Y}^{(j)}_{N,R}(\omega_k)\mathcal{Y}^{(l)}_{N,I}(\omega_k)] = -E[\mathcal{Y}^{(l)}_{N,R}(\omega_k)\mathcal{Y}^{(j)}_{N,I}(\omega_k)]$. The latter property implies also that the diagonal elements of $\mathbf{C}_{N,2}$ are equal to zero, i.e., $E[\mathcal{Y}^{(j)}_{N,R}(\omega_k)\mathcal{Y}^{(j)}_{N,I}(\omega_k)] = 0$, for every j and ω_k. Because of Eqn. 4.25 the complex vector $\mathcal{Y}_N(\omega_k)$ is said to have a complex multivariate Normal distribution (Krishnaiah 1976) as $N \to \infty$.

Assume now that there is a set of independent, identically distributed, time histories $\mathbf{Y}^{(1)}_N, \ldots, \mathbf{Y}^{(M)}_N$. As $N \to \infty$, the corresponding Fourier Transforms $\mathcal{Y}^{(n)}_N(\omega_k), n = 1, \ldots, M$ are independent and follow an identical complex N_s-variate Normal distribution with zero mean. Then, if $M \geq N_s$, the average spectral density estimate

$$\mathbf{S}^M_{y,N}(\omega_k) = \frac{1}{M}\sum_{n=1}^{M} \mathbf{S}^{(n)}_{y,N}(\omega_k) = \frac{1}{M}\sum_{n=1}^{M} \mathcal{Y}^{(n)}_N(\omega_k)\bar{\mathcal{Y}}^{(n)^T}_N(\omega_k) \tag{4.26}$$

follows a central complex Wishart distribution of dimension N_s with M degrees of freedom and mean $E[\mathbf{S}^M_{y,N}(\omega_k)] = E[\mathbf{S}_{y,N}(\omega_k)] = 2[\mathbf{C}_{N,1}(\omega_k) - i\mathbf{C}_{N,2}(\omega_k)]$ as $N \to \infty$ (Krishnaiah 1976). The PDF of this distribution is given by

$$p(\mathbf{S}^M_{y,N}(\omega_k)) = c_4 \frac{|\mathbf{S}^M_{y,N}(\omega_k)|^{M-N_s}}{|E[\mathbf{S}_{y,N}(\omega_k)]|^M} \exp(-M\ tr\{E[\mathbf{S}_{y,N}(\omega_k)]^{-1}\mathbf{S}^M_{y,N}(\omega_k)\}) \tag{4.27}$$

where c_4 is a normalizing constant and $|\mathbf{A}|$ and $tr[\mathbf{A}]$ denote the determinant and the trace, respectively, of a matrix $\mathbf{A}$. Note that this approximation is very accurate even if $\mathbf{y}(n\Delta t)$, $n = 0, \ldots, N-1$, is not Gaussian. Again, this is due to the robustness of the Gaussian approximation of the FFT irrespective of the probability distribution of the response signal.

Also, note that in the special case of a SDOF oscillator or in the case of a MDOF system with only one set of data at one measured DOF ($M = 1$, $N_s = 1$), the distribution in Eqn. 4.27 becomes a Chi-square distribution with two degrees of freedom and so reduces to Eqn. 4.11.

Furthermore, when $N \to \infty$, the vectors $[\mathcal{Y}_{N,R}^T(\omega_k), \mathcal{Y}_{N,I}^T(\omega_k)]^T$ and $[\mathcal{Y}_{N,R}^T(\omega_l), \mathcal{Y}_{N,I}^T(\omega_l)]^T$ with $\omega_k \neq \omega_l$ are independent (Appendix A). This causes the complex vectors $\mathcal{Y}_N(\omega_k)$ and $\mathcal{Y}_N(\omega_l)$ to be independent (as $N \to \infty$). As a result, the matrices $\mathbf{S}_{y,N}^M(\omega_k)$ and $\mathbf{S}_{y,N}^M(\omega_l)$ are independently Wishart distributed for $k \neq l$:

$$p[\mathbf{S}_{y,N}^M(\omega_k), \mathbf{S}_{y,N}^M(\omega_l)] = p[\mathbf{S}_{y,N}^M(\omega_k)]p[\mathbf{S}_{y,N}^M(\omega_l)] \tag{4.28}$$

where the two right-hand factors are given by Eqn. 4.27. Although Eqn. 4.27 and 4.28 are correct only asymptotically as $N \to \infty$, it was shown by simulations that these are indeed very accurate approximations in a certain bandwidth of frequencies for the case where N is finite. In the case of displacements (or accelerations), such range of frequencies corresponds to the lower (or higher) frequency range $\omega_k \in [\omega_1, \omega_K]$ (or $[\omega_K, \omega_{N_1-1}]$).

4.4.3 Identification Based on Spectral Density Estimates

Based on the above discussion regarding the statistical properties of the average spectral estimator $\mathbf{S}_{y,N}^M(\omega_k)$, a Bayesian approach for updating the PDF of the uncertain parameter vector $\mathbf{a}$ is proposed as follows: Given $M \geq N_s$ independent sets of observed data $\hat{\mathbf{Y}}_N^{(n)}, n = 1, \ldots, M$, one may calculate the corresponding observed spectral estimate matrices $\hat{\mathbf{S}}_{y,N}^{(n)}, n = 1, \ldots, M$ using Eqn. 4.19 and 4.20. Next, one can calculate the average matrix estimates $\hat{\mathbf{S}}_{y,N}^M(\omega_k)$ using Eqn. 4.26 and then form

the set $\hat{\mathbf{S}}_{y,N}^{M,K} = \{\hat{\mathbf{S}}_{y,N}^{M}(k\Delta\omega), k = 1, \ldots, K\}$. Using Bayes' Theorem, the updated PDF of the model parameters $\mathbf{a}$ given the data $\hat{\mathbf{S}}_{y,N}^{M,K}$ is then given by

$$p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^{M,K}) = c_5 p(\mathbf{a}) p(\hat{\mathbf{S}}_{y,N}^{M,K}|\mathbf{a}) \tag{4.29}$$

where c_5 is a normalizing constant such that the integral of the right-hand side of Eqn. 4.29 over the domain of $\mathbf{a}$ is equal to one. The factor $p(\mathbf{a})$ in the above equation represents the prior PDF, which expresses the relative plausibilities of different values of $\mathbf{a}$ based on prior information and engineering judgement. The factor $p(\hat{\mathbf{S}}_{y,N}^{M,K}|\mathbf{a})$ expresses the contribution of the observed data. Based on Eqn. 4.27 and 4.28, this factor can be calculated as follows:

$$p(\hat{\mathbf{S}}_{y,N}^{M,K}|\mathbf{a}) \simeq c_6 \prod_{k=1}^{K} \frac{|\hat{\mathbf{S}}_{y,N}^{M}(\omega_k)|^{M-N_s}}{|E[\mathbf{S}_{y,N}(\omega_k)|\mathbf{a}]|^{M}} \exp(-M\ tr\{E[\mathbf{S}_{y,N}(\omega_k)|\mathbf{a}]^{-1}\hat{\mathbf{S}}_{y,N}^{M}(\omega_k)\}) \tag{4.30}$$

where $E[\mathbf{S}_{y,N}(\omega_k)|\mathbf{a}]$ is given by Eqn. 4.21 and 4.22 with $E[\mathbf{S}_{x,N}(\omega_k)|\mathbf{a}]$ estimated by simulation as explained earlier. It is suggested to choose ω_K such that the frequency range just includes all the peaks of the spectral density estimates. A more detailed discussion will be given in the third example.

The most probable parameters $\hat{\mathbf{a}}$ are obtained by minimizing an objective function $J(\mathbf{a}) = -\ln[p(\mathbf{a})p(\hat{\mathbf{S}}_{y,N}^{M,K}|\mathbf{a})]$. Furthermore, the updated PDF $p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^{M,K})$ can be approximated by a Gaussian distribution centered at the optimal point $\hat{\mathbf{a}}$ if it is globally identifiable (Beck and Katafygiotis 1998). The corresponding covariance matrix $\mathbf{\Sigma}_a$ is equal to the inverse of the Hessian matrix of the function $J(\mathbf{a}) = -\ln[p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^{M,K})]$ calculated at $\mathbf{a} = \hat{\mathbf{a}}$, i.e., $\mathbf{\Sigma}_a = \mathbf{H}(\hat{\mathbf{a}})^{-1}$ where $H_{jl}(\hat{\mathbf{a}}) = \left.\frac{\partial^2 J(\mathbf{a})}{\partial a_j \partial a_l}\right|_{\mathbf{a}=\hat{\mathbf{a}}}$. For the presented results, this Hessian matrix is calculated using a finite difference method. This property provides a very efficient way for the quantification of the uncertainty for the model parameters without evaluating high dimensional integrals. However, it is not always a very accurate approximation, e.g., in unidentifiable cases. One check is to assume that the Gaussian approximation is accurate and calculate some lower dimensional conditional PDFs and compare with the values calculated from Eqn. 4.29. If

they match well, then the approximation can be used. If they do not match, simulation methods may be used, e.g., Beck and Au (2002), to calculate the associated uncertainties for the parameters.

4.5 Numerical Examples

4.5.1 Example 4-1: Duffing Oscillator

In this example, a SDOF Duffing oscillator of known mass is considered, which is subjected to zero-mean stationary Gaussian white noise $f(t)$ with spectral intensity S_{fo}:

$$m\,\ddot{x}\,(t) + c\dot{x}(t) + k_1 x(t) + k_3 x^3(t) = f(t) \tag{4.31}$$

The simulated stationary response history $\hat{\mathbf{Y}}_N^{(1)}$ was generated with parameters $\tilde{\mathbf{a}} = [\tilde{c}, \tilde{k}_1, \tilde{k}_3, \tilde{S}_{fo}^{(1)}, \tilde{\sigma}_{\eta o}^{(1)}]^T$ where $m = 1$ kg, $\tilde{c} = 0.1$ kg/s, $\tilde{k}_1 = 4.0$ N/m, $\tilde{k}_3 = 1.0$ N/m^3, $\tilde{S}_{fo}^{(1)} = 0.01N^2s$ and $\tilde{\sigma}_{\eta o}^{(1)} = 0.0526m$ (20% noise). The sampling interval is $\Delta t = 0.1\,\text{sec}$, with total time $T = 1000\,\text{sec}$, so $N = 10000$.

Multiplying Eqn. 4.31 with $x(t-\tau)$ and taking expectation yields

$$mR_x''(\tau) + cR_x'(\tau) + k_1 R_x(\tau) + k_3 E[x(t-\tau)x^3(t)] = 0 \tag{4.32}$$

where $R_x(\tau) \equiv E[x(t-\tau)x(t)], \forall t \in \mathbb{R}$. The term $E[x(t-\tau)x^3(t)]$ can be approximated by neglecting the fourth cumulant term, that is, $E[x(t-\tau)x^3(t)] \approx 3\sigma_x^2 R_x(\tau)$, where $\sigma_x^2 = R_x(0)$ is the variance of the response (Lutes and Sarkani 1997). Therefore, a differential equation for an approximation of the response autocorrelation function can be readily obtained:

$$mR_x''(\tau) + cR_x'(\tau) + (k_1 + 3\sigma_x^2 k_3)R_x(\tau) = 0 \text{ with } R_x(0) = \sigma_x^2 \text{ and } R_x'(0) = 0 \tag{4.33}$$

Eqn. 4.33 is a second-order ODE with constant coefficients, which can be solved

analytically. Then, $E[S_{y,N}(\omega_k)]$ can be obtained for a given parameter vector $\mathbf{a}$ by using Eqn. 4.8. Finally, the updated PDF $p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^{K,(1)})$ is readily obtained using Eqn. 4.6, 4.12 and 4.13, where $p(\mathbf{a})$ is taken as constant over the region where $p(\hat{\mathbf{S}}_{y,N}^{K,(1)}|\mathbf{a})$ is large, i.e., a locally non-informative prior PDF (Box and Tiao 1973).

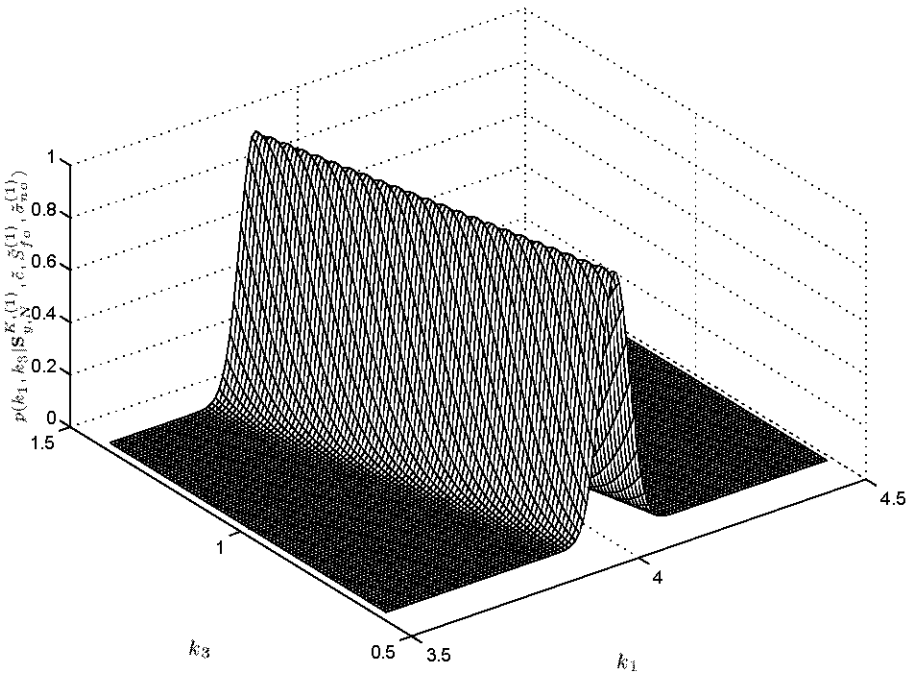

Figure 4.2: Conditional updated PDF $p(k_1, k_3|\hat{\mathbf{S}}_{y,N}^{K,(1)}, \tilde{c}, \tilde{S}_{fo}^{(1)}, \tilde{\sigma}_{\eta o}^{(1)})$ (Example 4-1).

Fig. 4.2 shows the conditional posterior PDF $p(k_1, k_3|\hat{\mathbf{S}}_{y,N}^{K,(1)}, \tilde{c}, \tilde{S}_{fo}^{(1)}, \tilde{\sigma}_{\eta o}^{(1)})$ normalized in such a way that the peak value is unity, which is obtained by utilizing only the spectral estimates up to frequency $\omega_K = 1.0$Hz ($K = 1000$). Note that the small-amplitude natural frequency of the oscillator is $\frac{1}{\pi}$Hz ≈ 0.32Hz. It is obvious that this case is unidentifiable, i.e., given one set of dynamic data, the estimates of k_1 and k_3 suffer from large uncertainty as there are infinitely many combinations of k_1 and k_3 which give similar values for the posterior PDF.

Another time history data set $\hat{\mathbf{Y}}_N^{(2)}$ was generated for the same oscillator (same $\tilde{c}$, $\tilde{k}_1$ and $\tilde{k}_3$) but with $\tilde{S}_{fo}^{(2)} = 0.04N^2s$ and $\tilde{\sigma}_{\eta o}^{(2)} = 0.1092m$ (20% noise). This case is, again, unidentifiable. However, if one plots these two posterior PDFs together

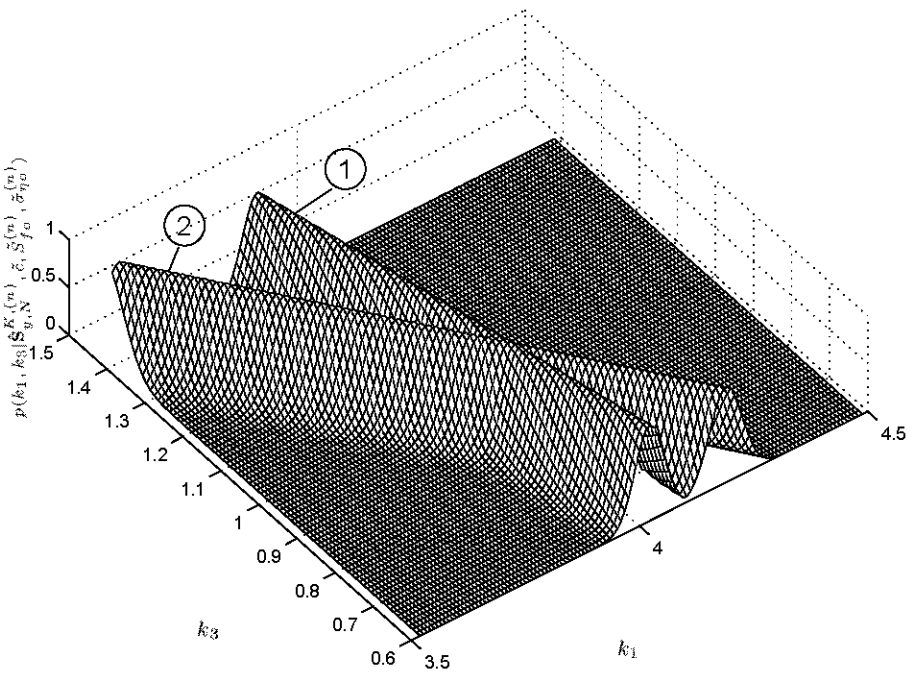

Figure 4.3: Conditional updated PDFs $p(k_1, k_3|\hat{\mathbf{S}}_{y,N}^{K,(n)}, \tilde{c}, \tilde{S}_{fo}^{(n)}, \tilde{\sigma}_{\eta o}^{(n)})$,$n = 1, 2$ (Example 4-1).

(shown in Fig. 4.3), the peak trajectories in the (k_1, k_3) plane have different slope. By Eqn. 4.33, the equivalent linear system has stiffness $k_1 + 3\sigma_x^2 k_3$. Therefore, the auto-correlation coefficients depend on σ_x and hence the level of excitation S_{fo}, showing that different levels of excitation lead to different slopes of the peak trajectories in the (k_1, k_3) plane. Since the coefficient $3\sigma_x^2$ is always positive, the slope of the peak trajectories in the (k_1, k_3) plane is always negative. This is expected because a larger value of k_1 can compensate for a smaller value of k_3, and vice versa.

Fig. 4.3 suggests that if one uses the two dynamic data sets $\hat{\mathbf{Y}}_N^{(1)}$ and $\hat{\mathbf{Y}}_N^{(2)}$ together, uncertainty in k_1 and k_3 can be significantly reduced. Table 4.1 shows the estimated optimal values $\hat{\mathbf{a}} = [\hat{c}, \hat{k}_1, \hat{k}_3, \hat{S}_{fo}^{(1)}, \hat{S}_{fo}^{(2)}, \hat{\sigma}_{\eta o}^{(1)}, \hat{\sigma}_{\eta o}^{(2)}]^T$ and the calculated standard deviations σ_c, σ_{k_1}, σ_{k_3}, $\sigma_{S_{fo}^{(1)}}$, $\sigma_{S_{fo}^{(2)}}$, $\sigma_{\sigma_{\eta o}^{(1)}}$ and $\sigma_{\sigma_{\eta o}^{(2)}}$ obtained using both data sets $\hat{\mathbf{Y}}_N^{(1)}$ and $\hat{\mathbf{Y}}_N^{(2)}$. It also gives the coefficient of variation (COV) for the parameter estimates and a "normalized error" β. This normalized error parameter represents the absolute value of the difference between the identified optimal value and exact

value, normalized with respect to the corresponding calculated standard deviation. The COVs in Table 4.1 are all quite small, showing the parameter values are pinned down rather precisely by the data. The normalized errors β in Table 4.1 are the order of 2 or less, suggesting that the procedure is not producing "biased" estimates, that is, the errors are not unusually large.

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	Standard Deviation σ	COV $\alpha = \frac{\sigma}{\tilde{a}}$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
c	0.1000	0.1021	0.0108	0.108	0.20
k_1	4.0000	3.9420	0.0463	0.012	1.25
k_3	1.0000	0.9868	0.1295	0.130	0.10
$S_{fo}^{(1)}$	0.0100	0.0098	0.0005	0.046	0.41
$S_{fo}^{(2)}$	0.0400	0.0454	0.0020	0.051	2.64
$\sigma_{\eta o}^{(1)}$	0.0526	0.0514	0.0022	0.042	0.55
$\sigma_{\eta o}^{(2)}$	0.1092	0.1025	0.0045	0.041	1.49

Table 4.1: Comparison of the actual parameters versus the optimal estimates and their statistics for the Duffing oscillator (Example 4-1).

Fig. 4.4 shows the conditional updated PDFs $p(k_1|\hat{\mathbf{S}}_{y,N}^{K,(1)}, \hat{\mathbf{S}}_{y,N}^{K,(2)}, \hat{c}, \hat{k}_3, \hat{S}_{fo}^{(1)}, \hat{S}_{fo}^{(2)}, \hat{\sigma}_{\eta o}^{(1)}, \hat{\sigma}_{\eta o}^{(2)})$ and $p(k_3|\hat{\mathbf{S}}_{y,N}^{K,(1)}, \hat{\mathbf{S}}_{y,N}^{K,(2)}, \hat{c}, \hat{k}_1, \hat{S}_{fo}^{(1)}, \hat{S}_{fo}^{(2)}, \hat{\sigma}_{\eta o}^{(1)}, \hat{\sigma}_{\eta o}^{(2)})$, obtained from: (i) Eqn. 4.14 (crosses) and (ii) the Gaussian approximation (solid line). It can be seen that the Gaussian approximation is very accurate. This property provides a very efficient way for the quantification of the uncertainty for the model parameters without evaluating high dimensional integrals.

Fig. 4.5 shows nearly elliptical contours (solid lines) in the (k_1, k_3) plane of the conditional updated PDF $p(k_1, k_3|\hat{\mathbf{S}}_{y,N}^{K,(1)}, \hat{\mathbf{S}}_{y,N}^{K,(2)}, \hat{c}, \hat{S}_{fo}^{(1)}, \hat{S}_{fo}^{(2)}, \hat{\sigma}_{\eta o}^{(1)}, \hat{\sigma}_{\eta o}^{(2)})$ calculated using Eqn. 4.14 (keeping all the other parameters fixed at their optimal values). These contours correspond to the parameter sets, which give $80\%, 60\%, 40\%, 20\%, 10\%$ and 5% of the conditional PDF values at its peak. Furthermore, by using the Gaussian approximation, the one standard deviation and two standard deviations contours can be calculated, which are shown by a dotted line and a dashed line, respectively. One can see that the orientation of the ellipses is the same for the two groups of contours, showing that the Gaussian approximation is very accurate in this case. Note that

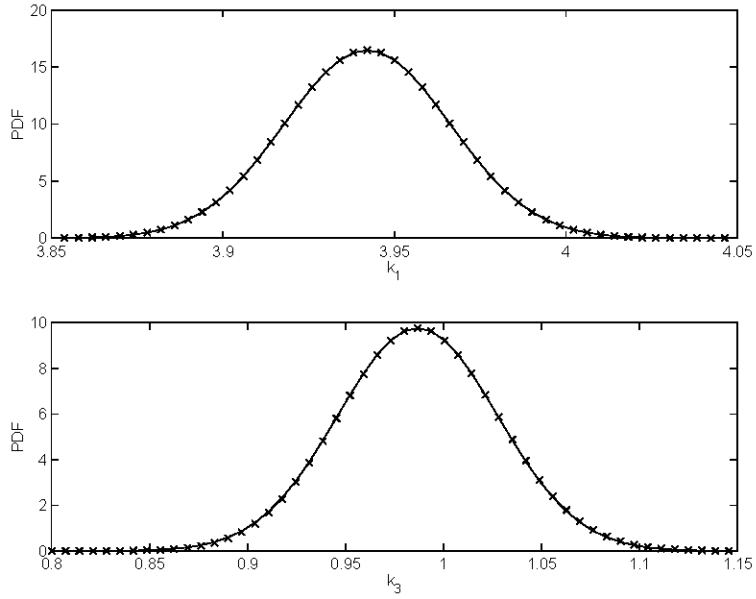

Figure 4.4: Conditional PDFs of k_1 and k_3 calculated using: i) Eqn. 4.14 - **crosses**; and ii) Gaussian approximation - **solid**. The remaining parameters are fixed at their optimal values (Example 4-1).

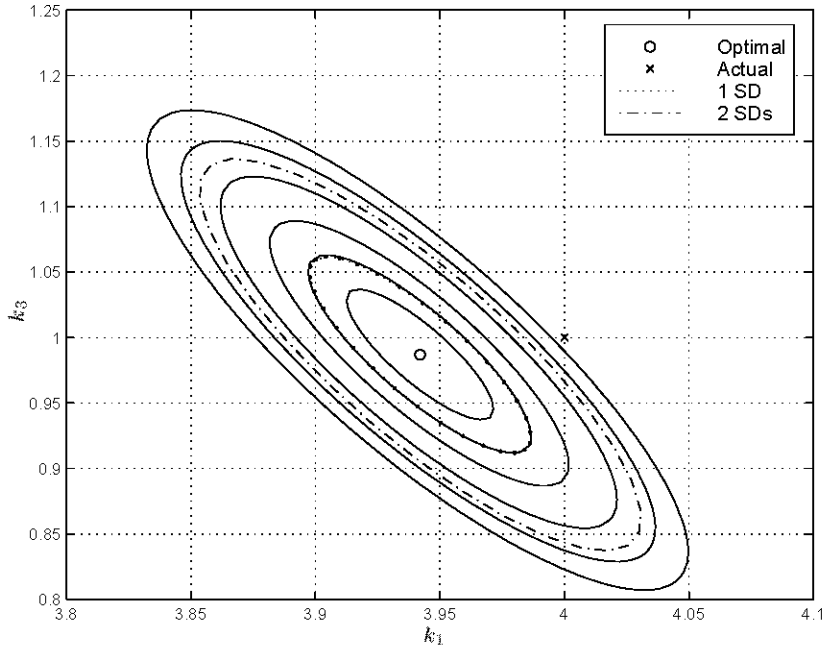

Figure 4.5: Contours in the (k_1, k_3) plane of conditional updated PDF $p(k_1, k_3|\hat{\mathbf{S}}_{y,N}^{K,(1)}, \hat{\mathbf{S}}_{y,N}^{K,(2)}, \hat{c}, \hat{S}_{fo}^{(1)}, \hat{S}_{fo}^{(2)}, \hat{\sigma}_{\eta o}^{(1)}, \hat{\sigma}_{\eta o}^{(2)})$ (Example 4-1).

the optimal parameter values seems to be more than two standard deviations away from their actual values because this figure shows the conditional PDF but not the marginal PDF.

Note that the estimation of the model parameters $\boldsymbol{\theta}_s$ is not sensitive to the choice of the cutoff frequency ω_K as long as it is larger than the frequency at which the peak of the response spectral density estimates occurs. Identification using the same sets of data was also carried out with $\omega_K = 5.0$Hz (the Nyquist frequency in this case). The results were virtually the same as those using $\omega_K = 1.0$Hz except that there were significant reductions in the uncertainty of the noise levels, i.e., utilizing a larger ω_K gives better estimates for the noise level only. Therefore, it is suggested to choose an ω_K ranging from $1.5\omega_\rho$ to $2\omega_\rho$ where ω_ρ is the frequency at which the peak of the spectral estimates $\hat{\mathbf{S}}_{y,N}(\omega_k)$ occurs. It is computationally efficient to use such values of ω_K without sacrificing the quality of the identification for the model parameters $\boldsymbol{\theta}_s$.

4.5.2 Example 4-2: Elasto-plastic Oscillator

In this example, an elasto-plastic SDOF oscillator of known mass is considered, which is subjected to zero-mean stationary Gaussian white noise $f(t)$ with spectral intensity S_{fo}:

$$m\,\ddot{x}\,(t) + f_s(x(t)) = f(t) \tag{4.34}$$

where $f_s(x(t))$ is the restoring force of the system. The restoring force-displacement relationship is shown in Fig. 4.6. The simulated stationary response history $\hat{\mathbf{Y}}_N$ was generated with parameters $\tilde{\mathbf{a}}_o = [\tilde{k}_1, \tilde{x}_y, \tilde{S}_{fo}, \tilde{\sigma}_{\eta o}]^T$ where $m = 1$ kg, $\tilde{k}_1 = 16.0$ N/m, $\tilde{x}_y = 1.0$ m, $\tilde{S}_{fo} = 0.15N^2s$ and $\tilde{\sigma}_{\eta o} = 0.1206m$ (20% noise). The sampling rate interval is $\Delta t = 0.05\,$sec, with a total time $T = 200\,$sec, that is, $N = 4000$. The hysteresis loops of the simulated data are shown in Fig. 4.7. Note that these hysteresis loops are not assumed to be measured; they are shown here only for illustrative purposes.

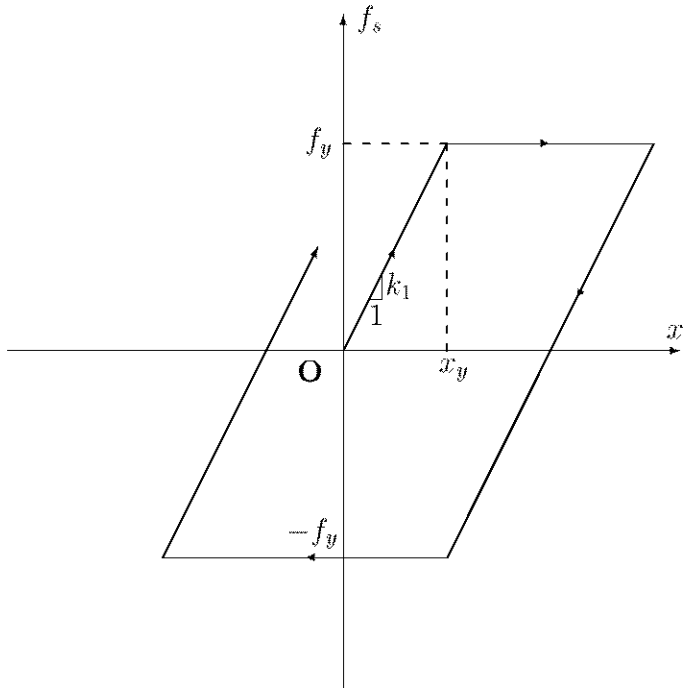

Figure 4.6: Relationship between the restoring force and the displacement of the system (Example 4-2).

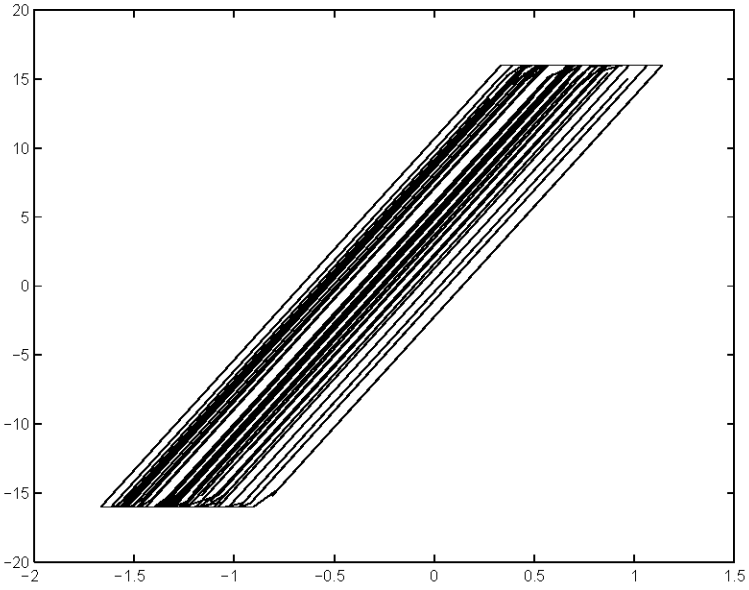

Figure 4.7: Hysteresis loops of the simulated data (Example 4-2).

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	Standard Deviation σ	COV $\alpha = \frac{\sigma}{\tilde{a}}$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
k_1	16.000	15.827	0.1162	0.007	1.49
x_y	1.0000	1.3493	0.4818	0.482	0.72
σ_x	0.6029	0.5762	0.1437	0.238	0.19
$\sigma_{\eta o}$	0.1206	0.1376	0.0209	0.173	0.82

Table 4.2: Identification results for the elasto-plastic system with the theoretical spectrum estimated by equivalent linearization (Example 4-2).

The equivalent linear system has the following equation of motion:

$$m\,\ddot{x}\,(t) + b_2\dot{x}(t) + b_1 x(t) = f(t) \tag{4.35}$$

where b_1 and b_2 are given by (Iwan and Lutes 1968; Lutes and Sarkani 1997)

$$\begin{aligned} b_1 &= k_1\left\{1 - \frac{8}{\pi}\int_1^\infty \left[\frac{1}{z^3} + \frac{x_y^2}{2\sigma_x^2 z}\sqrt{z-1}\exp(-\frac{x_y^2 z^2}{2\sigma_x^2})\right]dz\right\} \\ b_2 &= \sqrt{\frac{m}{2\pi b_1}}\frac{k_1 x_y}{\sigma_x}\left[1 - erf(\frac{x_y}{\sqrt{2}\sigma_x})\right] \end{aligned} \tag{4.36}$$

Note that the calculation of b_1 and b_2 requires σ_x^2, the variance of the response. Although σ_x can be determined from the spectral intensity of the excitation S_{fo}, it will be computationally more efficient to include σ_x directly instead of S_{fo} in the parameter set $\mathbf{a}$. Therefore, the parameter set $\mathbf{a} = [k_1, x_y, \sigma_x, \sigma_{\eta o}]^T$ is identified instead of $\mathbf{a}_o$ in this case. Then, $E[S_{y,N}(\omega_k)]$ can be obtained given a parameter set $\mathbf{a}$ by using Eqn. 4.8 where $R_x(n\Delta t)$ is approximated by the autocorrelation function for the equivalent linear system given by Eqn. 4.35 and 4.36. Finally, the updated PDF $p(\mathbf{a}|\hat{\mathbf{S}}_{y,N}^K)$ is readily obtained using Eqn. 4.6, 4.12 and 4.13. Note that a locally noninformative prior distribution is used, as in Example 4-1.

Table 4.2 shows the estimated optimal values $\hat{\mathbf{a}} = [\hat{k}_1, \hat{x}_y, \hat{\sigma}_x, \hat{\sigma}_{\eta o}]^T$ and the calculated standard deviations σ_{k_1}, σ_{x_y}, σ_{σ_x} and $\sigma_{\sigma_{\eta o}}$ obtained using the single data set $\hat{\mathbf{Y}}_N$. Fig. 4.8 shows contours in the (k_1, x_y) plane of the marginal updated PDF $p(k_1, x_y|\hat{\mathbf{S}}_{y,N}^K)$ calculated for one set of simulated data using Eqn. 4.14 (keeping all

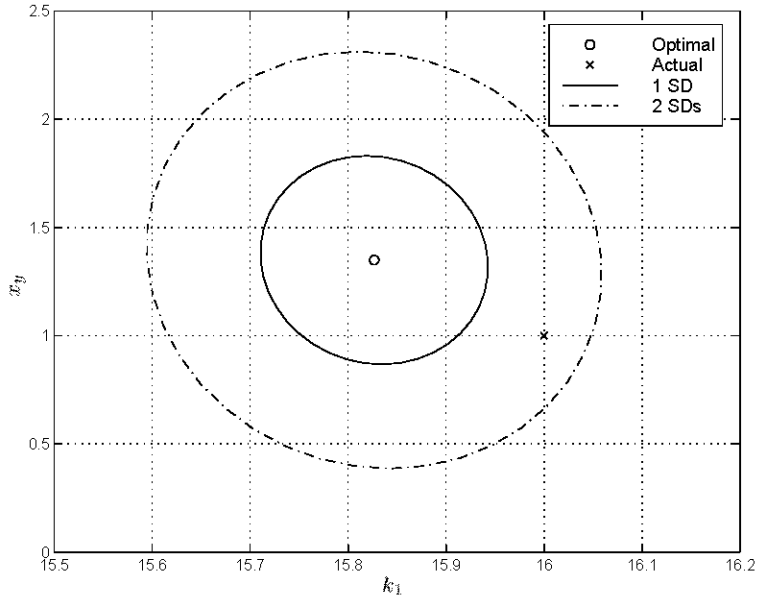

Figure 4.8: Contours of marginal updated PDF $p(k_1, x_y|\hat{\mathbf{S}}_{y,N}^K)$ with the theoretical spectrum estimated by equivalent linearization (Example 4-2).

the other parameters fixed at their optimal values) and utilizing only the spectral estimates up to frequency $\omega_K = 1.25$Hz (K=250). Note that the small-amplitude frequency of the oscillator is $\frac{2}{\pi}$Hz ≈ 0.63Hz. Again, ω_K can be chosen between $1.5\omega_\rho$ and $2.0\omega_\rho$, as in Example 4-1, where from Fig. 4.9, $\omega_\rho \approx 0.65$Hz.

Fig. 4.10 shows a similar plot of Fig. 4.8 but in the (x_y, σ_x) plane. It can be seen that the contours are very thin lying on the line $\sigma_x = \alpha_1 x_y + \alpha_2$, where $\alpha_1 \approx 0.28$ and $\alpha_2 \approx 0.2$, showing that the estimates of these parameters are very correlated. This is because b_1 and b_2 in Eqn. 4.35 depends on m, k_1 and x_y/σ_x only. The only factor that makes x_y and σ_x identifiable comes from the amplitude of the spectrum, which is proportional to σ_x^2. This also explains why the uncertainty for x_y and σ_x is so large when utilizing equivalent linearization. Note that although the actual values of the parameters $\tilde{x}_y$ and $\tilde{\sigma}_x$ are within two standard deviations from their optimal values $\hat{x}_y$ and $\hat{\sigma}_x$, respectively, the actual parameters in the (x_y, σ_x) plane lies far outside the two standard deviations contour, showing that this estimate is biased.

Table 4.3 shows the identification results using the same set of data with the

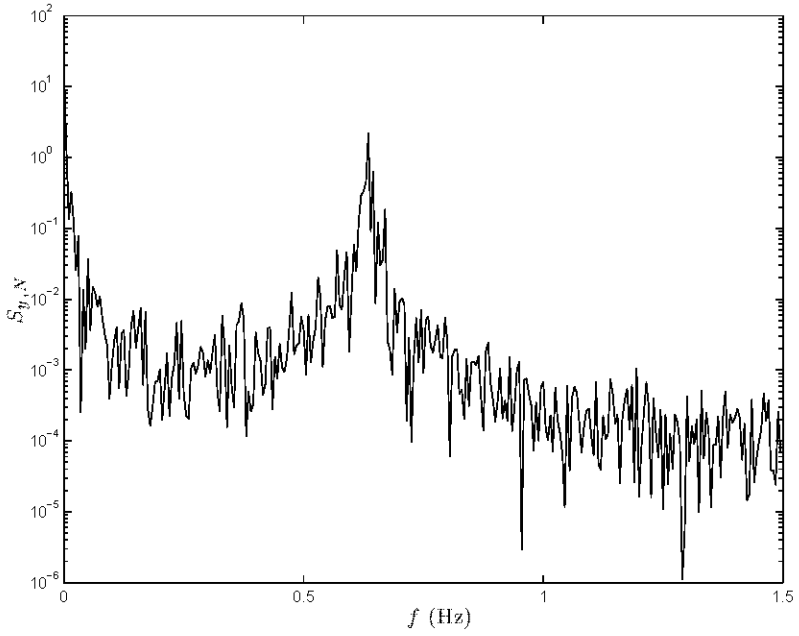

Figure 4.9: Spectral estimates using the measurements (Example 4-2).

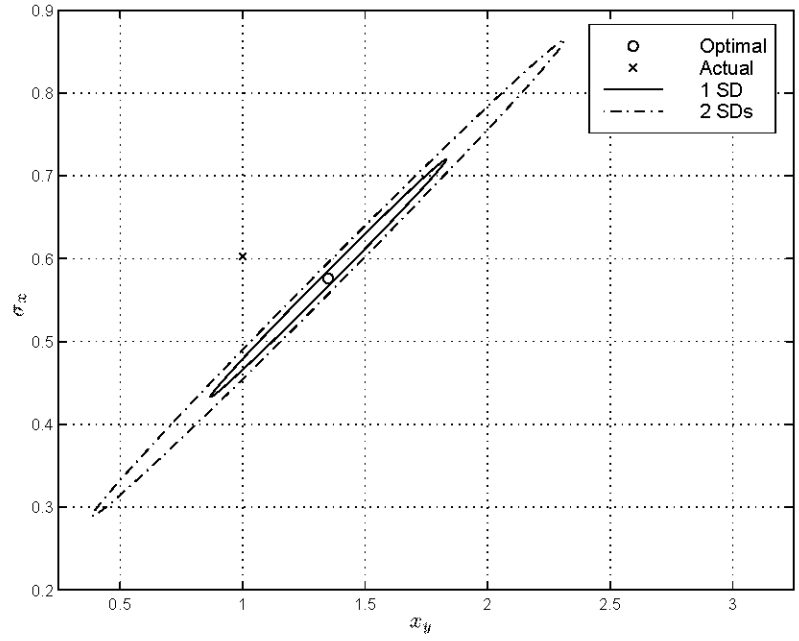

Figure 4.10: Contours of marginal updated PDF $p(x_y, \sigma_x|\hat{\mathbf{S}}^K_{y,N})$ with the theoretical spectrum estimated by equivalent linearization (Example 4-2).

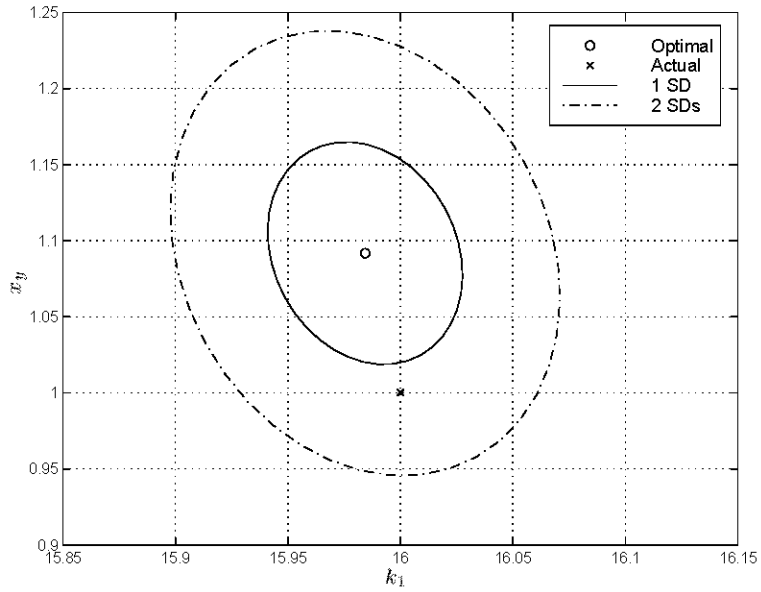

Figure 4.11: Contours of marginal updated PDF $p(k_1, x_y|\hat{\mathbf{S}}^K_{y,N})$ with the theoretical spectrum estimated by simulation (Example 4-2).

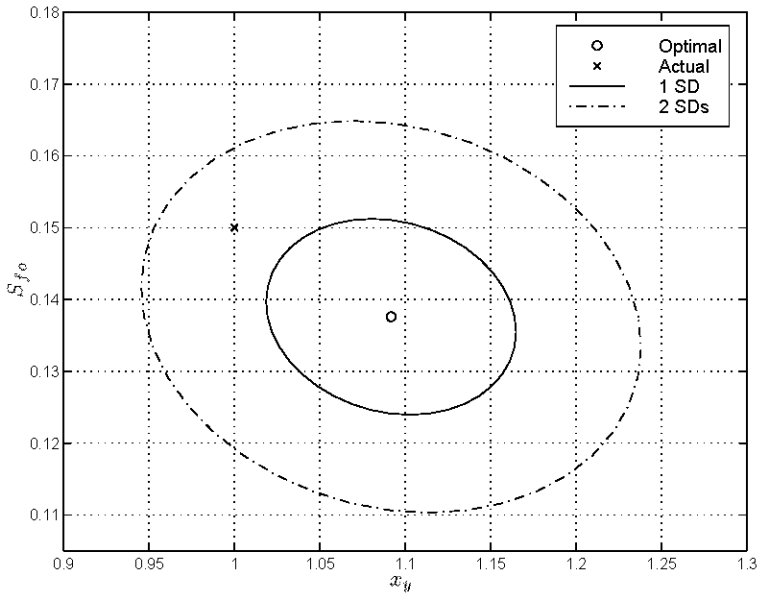

Figure 4.12: Contours of marginal updated PDF $p(x_y, S_{fo}|\hat{\mathbf{S}}^K_{y,N})$ with the theoretical spectrum estimated using simulation (Example 4-2).

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	Standard Deviation σ	COV $\alpha = \frac{\sigma}{\tilde{a}}$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
k_1	16.000	15.984	0.0433	0.003	0.36
x_y	1.0000	1.0918	0.0732	0.073	1.25
S_{fo}	0.1500	0.1376	0.0136	0.091	0.91
$\sigma_{\eta o}$	0.1206	0.1359	0.0201	0.166	0.76

Table 4.3: Identification results for the elasto-plastic system with the theoretical spectrum estimated by simulation (Example 4-2).

theoretical spectrum estimated by simulation, rather than by using Eqn. 4.35 and 4.36. Note that in this case, the uncertain parameter set is $\mathbf{a}_o = [k_1, x_y, S_{fo}, \sigma_{\eta o}]$, i.e., it includes the spectral intensity of the excitation instead of the rms of the response, because this is more efficient for the simulation of the system response. Here, one hundred samples of spectral estimates are simulated and the theoretical spectrum is approximated by the average of them. One can see that it gives more precise optimal parameter values than those in Table 4.2 by comparing the respective COVs. This is because the equivalent linear system can not capture completely the dynamics of the nonlinear oscillator. Therefore, the results obtained by using an equivalent linear system lose some information from the data, suggesting that for the identification of highly nonlinear systems, the simulation approach is the preferred one.

Fig. 4.11 and Fig. 4.12 show contours of the marginal updated PDF $p(k_1, x_y|\hat{\mathbf{S}}^K_{y,N})$ and $p(x_y, S_{fo}|\hat{\mathbf{S}}^K_{y,N})$, respectively. It can be seen that the optimal parameter set is within two standard deviations away from the actual parameter set in both (k_1, x_y) and (x_y, S_{fo}) plane.

4.5.3 Example 4-3: Four-story Yielding Structure

The third example uses simulated response data for a four-story yielding structure shown in Fig. 4.13. The nonlinear springs have the same behavior as described in Fig. 4.6 in Example 4-2. The structure has a uniformly distributed floor mass and story stiffness over its height. The linear stiffness to mass ratios $\frac{\tilde{k}_j}{m_j}, j = 1, \ldots, 4$, are chosen to be $1310\,\mathrm{sec}^{-2}$ so that the small amplitude fundamental frequency is

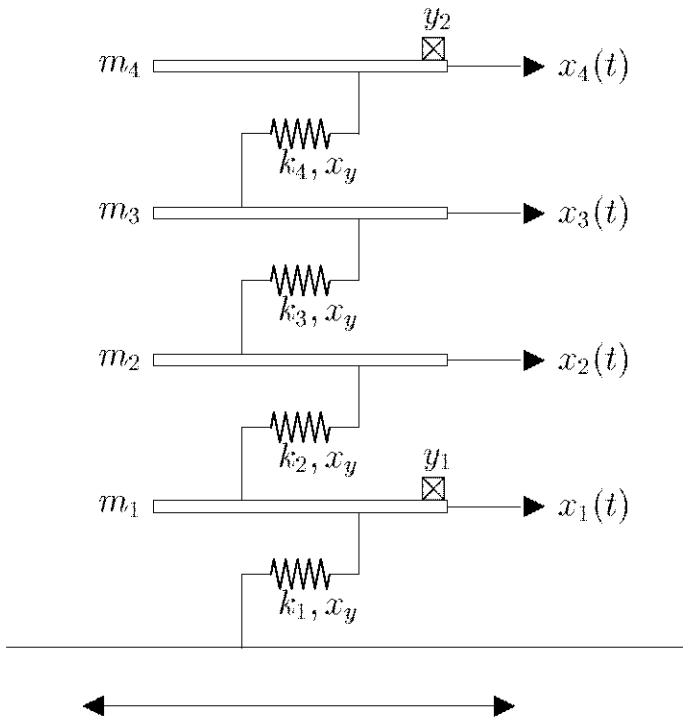

Figure 4.13: Four-story yielding structure (Example 4-3).

2.00Hz. Furthermore, the yielding level is chosen to be $\tilde{x}_y = 0.015$m for each story, which corresponds to 0.5% drift if the story height is 3.0m. For better scaling in the identification process, the stiffness and yielding parameters are parameterized by $k_j = \theta_j \tilde{k}_j, j = 1, \ldots, 4$ and $x_y = \theta_y \tilde{x}_y$, where $\tilde{k}_j = 2.10 \times 10^5$kN/m and $\tilde{x}_y = 0.015$m are the nominal values for the linear stiffness of the j^{th} story and the nominal yielding level for all four stories. Displacements at the 2^{nd} and 5^{th} floor were measured over a time interval $T = 25\,\text{sec}$, using a sampling interval $\Delta t = 0.01\,\text{sec}$. Therefore, the total number of measured time points is $N = 2500$. The structure is assumed to be subjected to a white noise base acceleration f with spectral intensity $S_{fo} = 0.006\ \text{m}^2\,\text{sec}^{-3}$. Note that the matrix $\mathbf{T}$ in Eqn. 4.15 is equal to the 4×1 matrix $-[m_1\ m_2\ m_3\ m_4]^T$ in this case. The noise added to the simulated response has a noise-to-signal ratio of 10%, i.e., the rms of the noise for a particular channel is equal to 10% of the rms of the noise-free response at the corresponding DOF.

Fig. 4.14 shows the simulated measured model displacement time histories at the

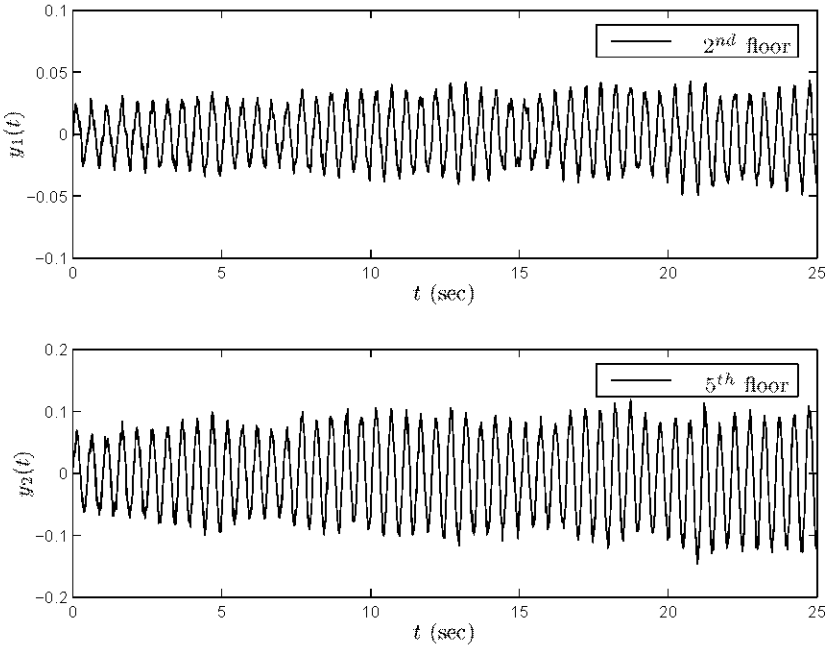

Figure 4.14: Displacement measurements at the 2^{nd} and 5^{th} floor (Example 4-3).

2^{nd} and 5^{th} floor and Fig. 4.15 shows the hysteresis loops for the fourth story, that is, the restoring force $f_{s4}(t)$ normalized by m_4 versus the interstory displacement $x_4(t) - x_3(t)$. Note that these hysteresis loops are not assumed to be measured; they are shown only for the purpose of illustrating the level of nonlinearity. Note also that the nonlinearity in the other stories is even higher. The time histories were separated into five segments ($M = 5$) with equal length in order to obtain five sets of spectral estimates. The expected value of the spectral density estimator is obtained by simulation. Fig. 4.16 shows the comparison between the spectral estimates $S_{y,N}$ (solid lines) and their expected values $E[S_{y,N}]$ (dashed lines) for the 2^{nd} and 5^{th} floor. One can see that $E[S_{y,N}]$ fits all the peaks of the measurements for both floors. Note that $E[\mathbf{S}_{y,N}]$ is obtained by the following procedure. First, simulate one hundred system responses. Then, by using Eqn. 4.19 and 4.20, one hundred samples of the spectral estimates can be obtained. By averaging these hundred samples for each discrete frequency, one obtains an estimate of the expected spectrum $E[\mathbf{S}_{y,N}]$.

Table 4.4 shows the identification results utilizing the spectral estimates up to

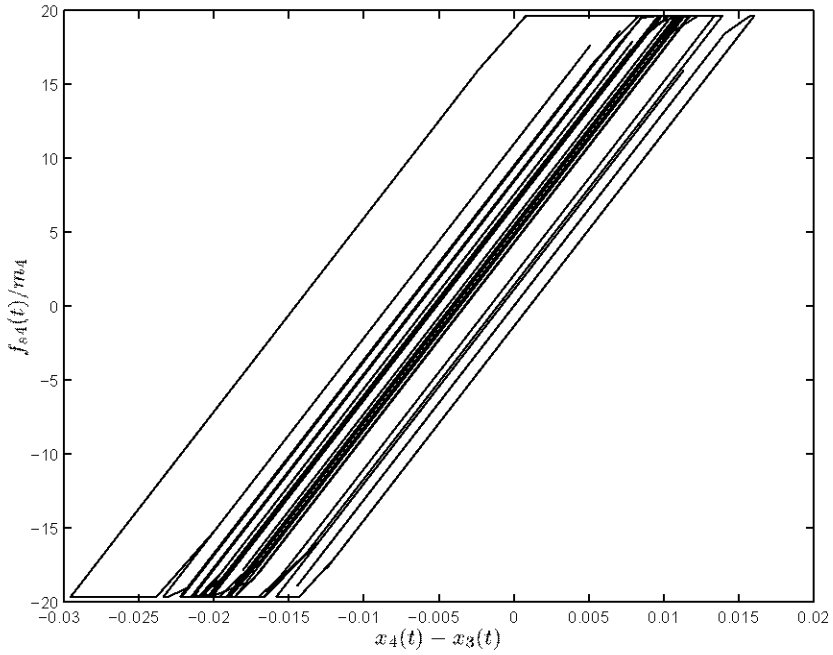

Figure 4.15: Hysteresis loops for the fourth story (Example 4-3).

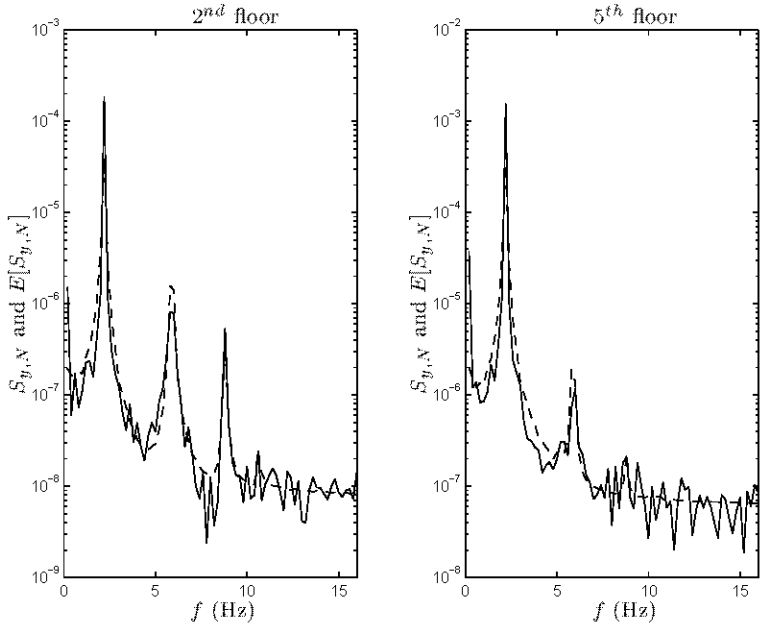

Figure 4.16: Spectral estimates and their expected values (Example 4-3).

Parameter	Actual $\tilde{a}$	Optimal $\hat{a}$	Standard Deviation σ	COV $\alpha = \frac{\sigma}{\tilde{a}}$	$\beta = \frac{\lvert\tilde{a}-\hat{a}\rvert}{\sigma}$
θ_1	1.0000	1.0122	0.0097	0.010	1.26
θ_2	1.0000	0.9907	0.0089	0.009	1.04
θ_3	1.0000	0.9903	0.0103	0.010	0.95
θ_4	1.0000	0.9947	0.0078	0.008	0.69
θ_y	1.0000	0.9577	0.0533	0.053	0.79
S_{fo}	0.0060	0.0076	0.0008	0.132	2.03
$\sigma_{\eta 1}$	0.0022	0.0022	0.0001	0.047	0.03
$\sigma_{\eta 2}$	0.0063	0.0062	0.0002	0.040	0.41

Table 4.4: Identification results for the four-story yielding building (Example 4-3).

$\omega_K = 16.0$Hz ($K = 80$). Again, a noninformative prior distribution for the model parameters is used. The second column in this table corresponds to the actual values used for generation of the simulated measurement data; the third and fourth columns correspond to the identified optimal parameters and the corresponding standard deviations, respectively; the fifth column lists the coefficient of variation for each parameter; and the last column shows the normalized error β, which is the difference between the actual and optimal parameters normalized by the calculated standard deviation. The first group of rows in the table corresponds to the stiffness parameters $\theta_j, j = 1, \ldots, 4$, followed by the yielding parameter θ_y, the forcing spectral intensity S_{fo} and the standard deviations of the prediction error, $\sigma_{\eta j}, j = 1, 2$, for the noise in the 2^{nd} and 5^{th} floor measured displacements. As shown by the small COVs, all the parameter values are pinned down rather precisely by the data. Also, the normalized errors β are the order of 2 or less, suggesting that the procedure is not producing "biased" estimates.

Fig. 4.17 shows the contours in the (θ_1, θ_2) plane of the marginal updated PDF of θ_1 and θ_2 (keeping all other parameters fixed at their optimal values). One observes that the actual parameters are at a reasonable distance, measured in terms of the estimated standard deviations, from the identified optimal parameters.

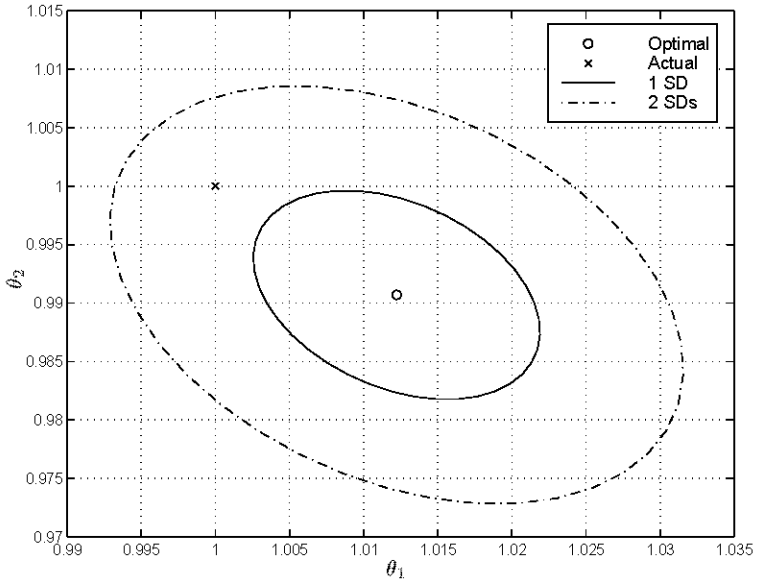

Figure 4.17: Contours of marginal updated PDF in the (θ_1, θ_2) plane (Example 4-3).

4.6 Conclusion

A Bayesian system identification approach is presented for updating the PDF of the model parameters for nonlinear systems using response data only. The proposed spectral-based approach relies on the robustness of the Gaussian approximation for the Fast Fourier Transform with respect to the probability distribution of the response signal in order to calculate the updated probability density function for the parameters of a nonlinear model conditional on the measured response. It does not require huge amounts of dynamic data, which is in contrast to most other published system identification methods for nonlinear models and unknown input. The approach provides not only the optimal estimates of the parameters but also the relative plausibilities of all values of the parameters based on the data. This probabilistic description is very important and can avoid misleading results, especially in unidentifiable cases. For the examples presented, the updated PDFs for the model parameters are well approximated by a multi-variate Gaussian distribution and so the precision with which the parameters are specified by the system response data are readily calculated.

Chapter 5 Stochastic Robust Control

5.1 Overview

A reliability-based output feedback control methodology is presented for controlling the dynamic response of systems that are represented by linear state-space models. The design criterion is based on a robust failure probability for the system. This criterion provides robustness for the controlled system by considering a probability distribution over a set of possible system models with a stochastic model of the excitation so that robust performance is expected. The control force can be calculated using incomplete response measurements at previous time steps without requiring state estimation. Examples of robust structural control using a shear building model and a benchmark structure are presented to illustrate the proposed method.

5.2 Stochastic Response Analysis

Consider a linear model of a structural system with N_d degrees-of-freedom (DOFs) and equation of motion:

$$\mathbf{M}(\boldsymbol{\theta}_s)\ddot{\mathbf{x}}(t) + \mathbf{C}(\boldsymbol{\theta}_s)\dot{\mathbf{x}}(t) + \mathbf{K}(\boldsymbol{\theta}_s)\mathbf{x}(t) = \mathbf{T}\cdot\mathbf{f}(t) + \mathbf{T}_c\cdot\mathbf{f}_c(t) \tag{5.1}$$

where $\mathbf{M}(\boldsymbol{\theta}_s)$, $\mathbf{C}(\boldsymbol{\theta}_s)$ and $\mathbf{K}(\boldsymbol{\theta}_s)$ are the $N_d \times N_d$ mass, damping and stiffness matrix, respectively, parameterized by the structural parameters $\boldsymbol{\theta}_s$ of the system; $\mathbf{f}(t) \in \mathbb{R}^{N_f}$ and $\mathbf{f}_c(t) \in \mathbb{R}^{N_{fc}}$ are the external excitation and control force vector, respectively, and $\mathbf{T} \in \mathbb{R}^{N_d \times N_f}$ and $\mathbf{T}_c \in \mathbb{R}^{N_d \times N_{fc}}$ are their distribution matrices. A control law is given later that specifies $\mathbf{f}_c$ by feedback of the measured output.

The uncertain excitation $\mathbf{f}(t)$ could be earthquake ground motions or wind forces, for example, and it is modeled by a zero-mean stationary filtered white-noise process

described by

$$\begin{aligned} \dot{\mathbf{w}}_f(t) &= \mathbf{A}_{wf}(\boldsymbol{\theta}_f)\mathbf{w}_f(t) + \mathbf{B}_{wf}(\boldsymbol{\theta}_f)\mathbf{w}(t) \\ \mathbf{f}(t) &= \mathbf{C}_{wf}(\boldsymbol{\theta}_f)\mathbf{w}_f(t) \end{aligned} \tag{5.2}$$

where $\mathbf{w}(t) \in \mathbb{R}^{N_w}$ is a Gaussian white-noise process with zero mean and unit spectral intensity matrix (here, spectral intensity is defined in the manner as in Eqn. 4.19); $\mathbf{w}_f(t) \in \mathbb{R}^{N_{wf}}$ is an internal filter state and $\mathbf{A}_{wf}(\boldsymbol{\theta}_f) \in \mathbb{R}^{N_{wf} \times N_{wf}}$, $\mathbf{B}_{wf}(\boldsymbol{\theta}_f) \in \mathbb{R}^{N_{wf} \times N_w}$ and $\mathbf{C}_{wf}(\boldsymbol{\theta}_f) \in \mathbb{R}^{N_f \times N_{wf}}$ are the parameterized filter matrices governing the properties of the filtered white noise. A vector $\boldsymbol{\theta}$ is introduced, which combines the structural parameter vector and the excitation parameter vector, i.e., $\boldsymbol{\theta} = [\boldsymbol{\theta}_s^T, \boldsymbol{\theta}_f^T]^T \in \mathbb{R}^{N_\theta}$. The dependence on $\boldsymbol{\theta}$ will be left implicit hereafter in this section.

Denote the state vector as: $\mathbf{y}(t) = [\mathbf{x}(t)^T, \dot{\mathbf{x}}(t)^T]^T$. Eqn. 5.1 can be rewritten in the state-space form as follows:

$$\dot{\mathbf{y}}(t) = \mathbf{A}_y \mathbf{y}(t) + \mathbf{B}_y \mathbf{f}(t) + \mathbf{B}_{yc} \mathbf{f}_c(t) \tag{5.3}$$

where $\mathbf{A}_y = \begin{bmatrix} \mathbf{0}_{N_d \times N_d} & \mathbf{I}_{N_d} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}$, $\mathbf{B}_y = \begin{bmatrix} \mathbf{0}_{N_d \times N_f} \\ \mathbf{M}^{-1}\mathbf{T} \end{bmatrix}$ and $\mathbf{B}_{yc} = \begin{bmatrix} \mathbf{0}_{N_d \times N_{fc}} \\ \mathbf{M}^{-1}\mathbf{T}_c \end{bmatrix}$. Here, $\mathbf{0}_{a \times b}$ and $\mathbf{I}_a$ denote the $a \times b$ zero and $a \times a$ identity matrix, respectively.

In order to allow more choices of the output to be fed back or to be controlled, an output vector $\mathbf{y}_f \in \mathbb{R}^{N_{yf}}$ is introduced that is modeled by the following state equation:

$$\dot{\mathbf{y}}_f(t) = \mathbf{A}_{yf}\mathbf{y}_f(t) + \mathbf{B}_{yf}\mathbf{y}(t) \tag{5.4}$$

where $\mathbf{A}_{yf} \in \mathbb{R}^{N_{yf} \times N_{yf}}$, $\mathbf{B}_{yf} \in \mathbb{R}^{N_{yf} \times 2N_d}$ are the matrices that characterizes the output filter. Note that the output vector can represent many choices of feedback. For example, it can handle displacement, velocity or acceleration measurements if the matrices $\mathbf{A}_{yf}$ and $\mathbf{B}_{yf}$ are chosen appropriately (Ivers and Miller 1991). Accelerations can be obtained approximately by passing the velocities in the state vector through

a filter with the transfer function $H_d(s)$:

$$H_d(s) = \frac{\omega_0^2 s}{s^2 + \sqrt{2}\omega_0 s + \omega_0^2} \tag{5.5}$$

This filter can approximate differentiation if ω_0 is chosen larger than the upper limit of the frequency band of interest. On the other hand, one can model the sensor dynamics for displacements or velocities measurements by using a low-pass filter with the transfer function $H_l(s) = \frac{\omega_0^2}{s^2+\sqrt{2}\omega_0 s+\omega_0^2}$. Another advantage of introducing the output vector $\mathbf{y}_f$ is that it allows for the modeling of the actuator. More details are given in the next chapter.

If the full state vector $\mathbf{v}(t) = [\mathbf{w}_f(t)^T, \mathbf{y}(t)^T, \mathbf{y}_f(t)^T]^T$ is introduced, then Eqn. 5.2 - 5.4 can be combined as follows:

$$\dot{\mathbf{v}}(t) = \mathbf{A}\mathbf{v}(t) + \mathbf{B}\mathbf{w}(t) + \mathbf{B}_c\mathbf{f}_c(t) \tag{5.6}$$

where the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{B}_c$ are given by

$$\mathbf{A} \equiv \begin{bmatrix} \mathbf{A}_{wf} & \mathbf{0}_{N_{wf}\times 2N_d} & \mathbf{0}_{N_{wf}\times N_{yf}} \\ \mathbf{B}_y\mathbf{C}_{wf} & \mathbf{A}_y & \mathbf{0}_{2N_d\times N_{yf}} \\ \mathbf{0}_{N_{yf}\times N_{wf}} & \mathbf{B}_{yf} & \mathbf{A}_{yf} \end{bmatrix}, \quad \mathbf{B} \equiv \begin{bmatrix} \mathbf{B}_{wf} \\ \mathbf{0}_{2N_d\times N_w} \\ \mathbf{0}_{N_{yf}\times N_w} \end{bmatrix} \text{ and } \mathbf{B}_c \equiv \begin{bmatrix} \mathbf{0}_{N_{wf}\times N_{fc}} \\ \mathbf{B}_{yc} \\ \mathbf{0}_{N_{yf}\times N_{fc}} \end{bmatrix} \tag{5.7}$$

By treating $\mathbf{w}$ as constant over each subinterval $[k\Delta t, k\Delta t + \Delta t)$, where Δt is the sampling time interval that is small enough to capture the dynamics of the structure, Eqn. 5.6 yields the following discrete-time equation:

$$\mathbf{v}[k+1] = \bar{\mathbf{A}}\mathbf{v}[k] + \bar{\mathbf{B}}\mathbf{w}[k] + \bar{\mathbf{B}}_c\mathbf{f}_c[k] \tag{5.8}$$

where $\mathbf{v}[k] \equiv \mathbf{v}(k\Delta t)$, $\bar{\mathbf{A}} \equiv e^{\mathbf{A}\Delta t}$, $\bar{\mathbf{B}} \equiv \mathbf{A}^{-1}(\bar{\mathbf{A}} - \mathbf{I}_{N_{wf}+2N_d+N_{yf}})\mathbf{B}$ and $\bar{\mathbf{B}}_c \equiv \mathbf{A}^{-1}(\bar{\mathbf{A}} - \mathbf{I}_{N_{wf}+2N_d+N_{yf}})\mathbf{B}_c$, and $\mathbf{w}[k]$ is Gaussian discrete white noise with zero mean and covariance matrix $\mathbf{\Sigma}_w = \frac{2\pi}{\Delta t}\mathbf{I}_{N_w}$.

Assume that discrete-time response data, with sampling time interval Δt, is available for N_o components of the output state, that is, the measured output is given by

$$\mathbf{z}[k] = \mathbf{L}_o \mathbf{v}[k] + \mathbf{n}[k] \tag{5.9}$$

where $\mathbf{L}_o \in \mathbb{R}^{N_o \times (N_{wf}+2N_d+N_{yf})}$ is the observation matrix and $\mathbf{n}[k] \in \mathbb{R}^{N_o}$ is the uncertain prediction error which accounts for the difference between the actual measured output from the structural system and the predicted output given by the model defined by Eqn. 5.8; it includes both modeling error and measurement noise. The prediction error is modeled as a stationary Gaussian discrete white noise process with zero mean and covariance matrix $\mathbf{\Sigma}_n$; this choice gives the maximum information entropy (greatest uncertainty) in the absence of any additional information about the unmodeled dynamics or output noise.

Now, choose a linear control feedback law using the current and the previous N_p output measurements,

$$\mathbf{f}_c[k] = \sum_{p=0}^{N_p} \mathbf{G}_p \mathbf{z}[k-p], \tag{5.10}$$

where $\mathbf{G}_p, p = 0, 1, \ldots, N_p$ are the gain matrices, which will be determined in the next section. It is worth noting that if the matrices $\mathbf{G}_p, p = 0, \ldots, N_p^\star$ $(N_p^\star < N_p)$ are fixed to be zero, the controller at any time step only utilizes output measurements from time steps that are more than $N_p^\star \Delta t$ back in the past. Furthermore, by choosing a value of $N_p^\star$ such that $N_p^\star \Delta t$ is larger than the reaction time of the control system (data acquisition, online calculation of the control forces and actuator reaction time), it is possible to avoid the instability problem caused by time-delay effects.

Substituting Eqn. 5.10 into Eqn. 5.8

$$\mathbf{v}[k+1] = (\bar{\mathbf{A}} + \bar{\mathbf{B}}_c \mathbf{G}_0 \mathbf{L}_o)\mathbf{v}[k] + \bar{\mathbf{B}}\mathbf{w}[k] + \bar{\mathbf{B}}_c \sum_{p=1}^{N_p} \mathbf{G}_p \mathbf{z}[k-p] + \bar{\mathbf{B}}_c \mathbf{G}_0 \mathbf{n}[k] \tag{5.11}$$

Now define an augmented vector $\mathbf{U}_{N_p}[k]$ as follows:

$$\mathbf{U}_{N_p}[k] \equiv [\mathbf{v}[k]^T, \mathbf{z}[k-1]^T, \ldots, \mathbf{z}[k-N_p]^T]^T \tag{5.12}$$

Then, Eqn. 5.11 can be rewritten as follows:

$$\mathbf{U}_{N_p}[k+1] = (\bar{\mathbf{A}}_u + \bar{\mathbf{B}}_{uc})\mathbf{U}_{N_p}[k] + \bar{\mathbf{B}}_u \bar{\mathbf{f}}[k] \tag{5.13}$$

where

$$\bar{\mathbf{f}}[k] \equiv [\mathbf{w}[k]^T \quad \mathbf{n}[k]^T]^T \tag{5.14}$$

and $\bar{\mathbf{A}}_u$, $\bar{\mathbf{B}}_u$ and $\bar{\mathbf{B}}_{uc}$ are given by

$$\bar{\mathbf{A}}_u \equiv \begin{bmatrix} \bar{\mathbf{A}} & \mathbf{0}_{(N_{wf}+2N_d+N_{yf})\times N_p N_o} \\ \mathbf{L}_o & \mathbf{0}_{N_o \times N_p N_o} \\ \mathbf{0}_{(N_p-1)N_o \times (N_{wf}+2N_d+N_{yf})} & \mathbf{I}_{(N_p-1)N_o} \quad \mathbf{0}_{(N_p-1)N_o \times N_o} \end{bmatrix} \tag{5.15}$$

$$\bar{\mathbf{B}}_{uc} \equiv \begin{bmatrix} \bar{\mathbf{B}}_c \mathbf{G}_0 \mathbf{L}_o & \bar{\mathbf{B}}_c \mathbf{G}_1 & \cdots & \bar{\mathbf{B}}_c \mathbf{G}_{N_p} \\ \mathbf{0}_{N_p N_o \times (N_{wf}+2N_d+N_{yf}+N_p N_o)} \end{bmatrix} \tag{5.16}$$

$$\bar{\mathbf{B}}_u \equiv \begin{bmatrix} \bar{\mathbf{B}} & \bar{\mathbf{B}}_c \mathbf{G}_0 \\ \mathbf{0}_{N_o \times N_w} & \mathbf{I}_{N_o} \\ \mathbf{0}_{(N_p-1)N_o \times N_w} & \mathbf{0}_{(N_p-1)N_o \times N_o} \end{bmatrix} \tag{5.17}$$

Therefore, the covariance matrix $\mathbf{\Sigma}_u \equiv E[\mathbf{U}_{N_p}[k]\mathbf{U}_{N_p}[k]^T]$ of the augmented vector

$\mathbf{U}_{N_p}$ is readily obtained:

$$\begin{aligned}\boldsymbol{\Sigma}_u &= (\bar{\mathbf{A}}_u + \bar{\mathbf{B}}_{uc})\boldsymbol{\Sigma}_u(\bar{\mathbf{A}}_u + \bar{\mathbf{B}}_{uc})^T + \bar{\mathbf{B}}_u\boldsymbol{\Sigma}_{\bar{f}}\bar{\mathbf{B}}_u^T \\ \boldsymbol{\Sigma}_{\bar{f}} &= \begin{bmatrix} \boldsymbol{\Sigma}_w & \boldsymbol{\Sigma}_{wn} \\ \boldsymbol{\Sigma}_{wn}^T & \boldsymbol{\Sigma}_n \end{bmatrix}\end{aligned} \tag{5.18}$$

where $\boldsymbol{\Sigma}_{\bar{f}}$ denotes the covariance matrix of the vector $\bar{\mathbf{f}}$ in Eqn. 5.14. Note that Eqn. 5.18 is a standard stationary Lyapunov covariance equation in discrete form.

In summary, the original continuous-time excitation, structural and output equations are transformed to a linear discrete-time state-space equation for an augmented vector $\mathbf{U}_{N_p}$. The system response is a stationary Gaussian process with zero mean and covariance matrix that can be readily calculated using Eqn. 5.18. These properties are used to design the optimal robust controller for the structure.

5.3 Optimal Controller Design

The optimal robust controller is defined here as the one which maximizes the robust reliability (Papadimitriou et al. 2001) with respect to the feedback gain matrices in Eqn. 5.10, that is, the one which minimizes the robust failure probability for a structural model with uncertain parameters representing the real structural system. Failure is defined as the situation in which at least one of the performance quantities (structural response or control force) exceeds a given threshold level. This is the classic 'first passage problem', which has no closed form solution (Lin 1976). Therefore, the proposed method utilizes an approximate solution bases on Rice's 'out-crossing' theory (Lin 1976).

5.3.1 Conditional Failure Probability

Use $\mathbf{q}[k] \in \mathbb{R}^{N_q}$ to denote the control performance vector of the system at time $k\Delta t$. Its components may be structural interstory drifts, floor accelerations, control

force, etc. The system performance is given by

$$\mathbf{q}[k] = \mathbf{P}_0 \mathbf{U}_{N_p}[k] + \mathbf{m}[k] \tag{5.19}$$

where $\mathbf{P}_0 \in \mathbb{R}^{N_q \times (N_{wf}+2N_d+N_{yf}+N_p N_o)}$ is a performance matrix which multiplies the augmented vector $\mathbf{U}_{N_p}$ from Eqn. 5.12 to give the corresponding performance vector of the model. In order to account for the unmodeled dynamics, the uncertain prediction error $\mathbf{m} \in \mathbb{R}^{N_q}$ in Eqn. 5.19 is introduced because the goal is to control the system performance, not the model performance; it is modeled as discrete white noise with zero mean and covariance matrix $\mathbf{\Sigma}_m$.

For a given failure event $F_i = \{|q_i(t)| > \beta_i \text{ for some } t \in [0,T]\}$, the conditional failure probability $P(F_i|\boldsymbol{\theta})$ for the performance quantity q_i based on the structural model and excitation model specified by $\boldsymbol{\theta}$ can be estimated using Rice's formula if $\nu_{\beta_i} << 1$ (Lin 1976):

$$P(F_i|\boldsymbol{\theta}) \approx 1 - \exp[-\nu_{\beta_i}(\boldsymbol{\theta})T] \tag{5.20}$$

where $\nu_{\beta_i}(\boldsymbol{\theta})$ is the mean out-crossing rate for the threshold level β_i and is given by

$$\nu_{\beta_i}(\boldsymbol{\theta}) = \frac{\sigma_{\dot{q}_i}}{\pi \sigma_{q_i}} \exp(-\frac{{\beta_i}^2}{2\sigma_{q_i}^2}) \tag{5.21}$$

where σ_{q_i} and $\sigma_{\dot{q}_i}$ are the standard deviation for the performance quantity q_i and its derivative $\dot{q}_i$, respectively. In implementation, $\dot{q}_i$ must be included in $\mathbf{y}_f$ in Eqn. 5.4 if it is not already part of $\mathbf{y}$.

Now consider the failure event $F = \cup_{i=1}^{N_q} F_i$, that is, the system fails if any $|q_i|$ exceeds its threshold β_i. Since the mean out-crossing rate of the system can be approximated by: $\nu = \sum_{i=1}^{N_q} \nu_{\beta_i}$ (Veneziano et al. 1977), the probability of failure $P(F|\boldsymbol{\theta})$ of the controlled structural system is given approximately by

$$P(F|\boldsymbol{\theta}) \approx 1 - \exp[-\sum_{i=1}^{N_q} \nu_{\beta_i}(\boldsymbol{\theta})T] \tag{5.22}$$

where N_q denotes the number of performance quantities considered.

5.3.2 Robust Failure Probability

No matter what technique, e.g., finite-element method or system identification, is used to develop a model for a structural system, the structural parameters are always uncertain to some extent. Furthermore, the excitation model is uncertain as well. Therefore, a probabilistic description is used to describe the uncertainty in the model parameters $\boldsymbol{\theta}$ defined earlier. Such probability distributions can be specified using engineering judgement or they can be obtained using system identification techniques. This leads to the concept of the robust failure probability given by the theorem of total probability (Papadimitriou et al. 2001):

$$P(F|\boldsymbol{\Theta}) = \int_{\boldsymbol{\Theta}} P(F|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\boldsymbol{\Theta}) d\boldsymbol{\theta} \tag{5.23}$$

which accounts for modeling uncertainties in deriving the failure probability. This robust failure probability is conditional on the probabilistic description of the parameters which is specified over the set of possible models $\boldsymbol{\Theta}$. Note that this high dimensional integral is difficult to evaluate numerically, so an asymptotic expansion is utilized (Papadimitriou et al. 1997a). Denote the integral of interest by

$$I = \int_{\boldsymbol{\Theta}} e^{l(\boldsymbol{\theta})} d\boldsymbol{\theta} \tag{5.24}$$

where $l(\boldsymbol{\theta})$ is given by

$$l(\boldsymbol{\theta}) = \ln[P(F|\boldsymbol{\theta})] + \ln[p(\boldsymbol{\theta}|\boldsymbol{\Theta})] \tag{5.25}$$

The basic idea here is to fit a Gaussian density centered at the 'design point' at which $e^{l(\boldsymbol{\theta})}$, or $l(\boldsymbol{\theta})$, is maximized. It is assumed here that there is a unique design point;

see (Au et al. 1999) for a more general case. Then, this integral is approximated by

$$I \equiv P(F|\Theta) \approx (2\pi)^{\frac{N_\theta}{2}} \frac{P(F|\boldsymbol{\theta}^\star)p(\boldsymbol{\theta}^\star|\Theta)}{\sqrt{\det \mathbf{L}(\boldsymbol{\theta}^\star)}} \tag{5.26}$$

where $\boldsymbol{\theta}^\star$ is the design point at which $l(\boldsymbol{\theta})$ has a maximum value and $\mathbf{L}(\boldsymbol{\theta}^\star)$ is the Hessian of $-l(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}^\star$. The optimization of $l(\boldsymbol{\theta})$ to find $\boldsymbol{\theta}^\star$ can be performed, for example, by using MATLAB subroutine 'fmins'.

The proposed control design can be summarized as follows: By solving Eqn. 5.18, the covariance matrix of the structural response can be obtained. Then, the robust failure probability can be calculated using the asymptotic expansion formula in Eqn. 5.26 along with Eqn. 5.20 - 5.22. The optimal robust controller is obtained by minimizing the robust failure probability over all possible controllers parameterized by their gain matrices, which again can be performed, for example, using MATLAB subroutine 'fmins'.

The optimal controller can be readily updated when dynamic data $\mathcal{D}$ is available from the systems (Beck and Katafygiotis 1998; Papadimitriou et al. 2001). In this case, Bayes' Theorem is used to get an updated PDF $p(\boldsymbol{\theta}|\mathcal{D}, \Theta)$ that replaces $p(\boldsymbol{\theta}|\Theta)$ in Eqn. 5.23 and hence the updated robust failure probability $p(F|\mathcal{D}, \Theta)$ (Papadimitriou et al. 2001) is minimized to obtain the optimal control gains.

5.4 Illustrative Examples

5.4.1 Example 5-1: Four-story Building under Seismic Excitation

The first example refers to a four-story building under seismic excitation with an active mass driver and a sensor on each floor above the ground level. In this example, the stochastic ground motion model is fixed during the controller design but the shear-building model of the structure (Fig. 5.1) is uncertain. The nominal model of the structure has a floor mass and interstory stiffness uniformly distributed over its height. The stiffness-to-mass ratios $\frac{k_i}{M_i}, i = 1, \ldots, 4$ is $1309.3\,\mathrm{sec}^{-2}$, where M_i is the mass of floor i. The nominal damping-to-mass ratios $\frac{c_i}{M_i}, i = 1, \ldots, 4$ are all chosen to be equal

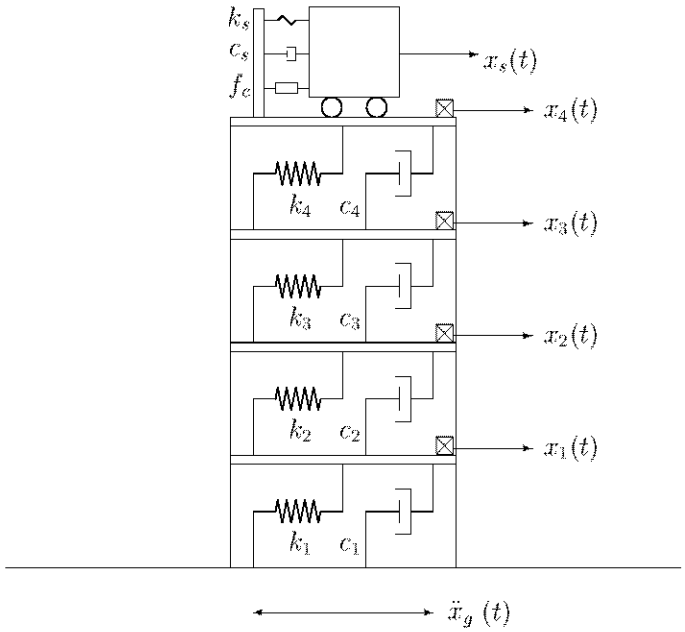

Figure 5.1: Four-story shear building with active mass driver on the roof (Example 5-1).

to $2.0\,\mathrm{sec}^{-1}$. As a result, the nominal modal frequencies of the uncontrolled structure are 2.00 Hz, 5.76 Hz, 8.82 Hz and 10.82 Hz and the nominal damping ratio of the first mode is 1.00%. In order to take into account the uncertainty in the structural model parameters, all the stiffness and damping parameters are assumed to be Gaussian distributed, truncated for positive values of the stiffness and damping, with mean at their nominal values and coefficients of variation 5% (stiffness) and 20% (damping), respectively. To provide more realism, the structure to be controlled is defined by model parameters sampled from the aforementioned probability distributions rather than being equal to the nominal structural model. This gave stiffness-to-mass ratios of $1253\,\mathrm{sec}^{-2}$, $1177\,\mathrm{sec}^{-2}$, $1304\,\mathrm{sec}^{-2}$ and $1344\,\mathrm{sec}^{-2}$ for the 1^{st} to 4^{th} floor, respectively. The corresponding damping-to-mass ratios are $2.50\,\mathrm{sec}^{-1}$, $2.16\,\mathrm{sec}^{-1}$, $1.68\,\mathrm{sec}^{-1}$ and $2.22\,\mathrm{sec}^{-1}$.

The ratio μ of the actuator mass M_s to the total structure mass $M_o = \sum_{i=1}^{4} M_i$ is chosen to be 1%. The natural frequency ω_s and the damping ratio ζ_s of the actuator may be chosen according to the following expressions which give the optimal passive control system for the first mode of the nominal structure under white-noise excitation

(Warburton and Ayorinde 1980):

$$\omega_s = \omega_1 \sqrt{\frac{2-\mu}{2(\mu+1)^2}}$$
$$\zeta_s = \sqrt{\frac{\mu(3\mu+4)}{2(\mu+1)(\mu+2)}} \tag{5.27}$$

where ω_1 is the fundamental frequency of the nominal uncontrolled structure. Then, the stiffness-to-mass ratio $\frac{k_s}{M_s}$ and the damping-to-mass ratio $\frac{c_s}{M_s}$ of the actuator are given by: $\frac{k_s}{M_s} = \frac{k_s}{\mu M_o} = \omega_s^2$ and $\frac{c_s}{M_s} = \frac{c_s}{\mu M_o} = 2\zeta_s\omega_s$. In this example, $\frac{k_s}{M_s} = 1.540 \times 10^2\,\mathrm{sec}^{-2}$ and $\frac{c_s}{M_s} = 2.473\,\mathrm{sec}^{-1}$ are the optimal parameters based on Eqn. 5.27. However, they are assumed to be $\frac{k_s}{M_s} = 1.60 \times 10^2\,\mathrm{sec}^{-2}$ and $\frac{c_s}{M_s} = 2.0\,\mathrm{sec}^{-1}$ in the following since it might not be possible to build a controller with the optimal values of $\frac{k_s}{M_s}$ and $\frac{c_s}{M_s}$ in reality; these parameters are assumed to be known during the controller design.

The controller design is based on maximizing the robust reliability or, equivalently, minimizing the robust failure probability, calculated for the structure with uncertain parameters subject to an uncertain white-noise ground excitation with spectral intensity of 0.01 $\mathrm{m}^2\mathrm{sec}^{-3}$ for a 20 sec interval. The threshold level for the interstory drifts, actuator stroke and the control force $f_{nc} = f_c/M_s$ (normalized by the actuator active mass) are chosen to be 2.0 cm, 2.0 m and 10 g, respectively. The failure event F of interest is the exceedence of any one of these threshold levels. For simplicity, it is assumed that displacements are measured at specified floors using a sampling interval $\Delta t = 0.01\,\mathrm{sec}$. In the next example, acceleration measurements will be assumed.

Four robust controllers are designed using the proposed methodology, each using different control feedback:

Controller 1: Displacement measurements at every floor at the current time step.

Controller 2: Displacement measurements at the 4^{th} floor at the current time step.

Controller 3: Displacement measurements at the 4^{th} floor at the current and previous two time steps.

Controller 4: Displacement measurements at the 4^{th} floor at the previous two time

steps.

Gain	Controller 1	Controller 2	Controller 3	Controller 4
$G_o(1)$	14.08	—	—	—
$G_o(2)$	11.87	—	—	—
$G_o(3)$	49.46	—	—	—
$G_o(4)$	32.66	86.15	134.58	—
$G_1(4)$	—	—	-26.63	237.45
$G_2(4)$	—	—	-20.98	-150.72

Table 5.1: Gain coefficients of the optimal controllers (Example 5-1).

	Passive	Controller 1	Controller 2	Controller 3	Controller 4
$P(F\|\boldsymbol{\Theta})$	0.56	0.0013	0.0014	0.0008	0.0009

Table 5.2: Robust failure probability (Example 5-1).

Table 5.1 shows the optimal gain parameters $G_p(i)$ for Controllers 1 - 4 where index p and index i correspond to the number of time-delay steps and the floor number, respectively. Table 5.2 shows the robust failure probability of the interstory drifts and the stroke for passive control (all gain coefficients are fixed at zero) and for Controllers 1 - 4. The active controllers give a much better design performance objective than the passive mass damper. All controllers give similar design performance objectives but Controller 3 is the best, followed by Controllers 4 and 1, and then 2. Although the number of measured degrees of freedom is different in Controllers 1 and 2, the performance of the controlled structure is almost the same. This is because the motion of the structure is dominated by the first mode in the case of ground shaking. Therefore, the measurements at one DOF contain almost all of the information regarding the motion of the structure. However, Controller 3 gives a better performance objective than Controller 1 even though Controller 3 uses only one sensor because measuring displacements at consecutive time steps gives more information, which corresponds to the structural velocities in this case.

Fig. 5.2 - 5.5 show the time histories of the interstory drifts using Controllers 1 - 4, respectively. The dashed and solid lines show the response of the uncontrolled and controlled structure, respectively, during simulated operation under the same ground motion sampled from the stochastic ground motion model. It can be seen that the interstory drifts are significantly reduced by the controllers. Furthermore, Table 5.3 shows the statistical properties (standard deviations and maximum) of the performance quantities (interstory drifts, strokes and controller accelerations) for the uncontrolled structure, passive control and Controllers 1 - 4. By comparing Controllers 1 and 2 in Table 5.2, one observes that the robust failure probabilities are very similar. Furthermore, Table 5.3 shows that the statistical properties of the responses in these two cases are almost the same. This implies that the performance of using feedback from one or four (all) degrees of freedom are virtually the same. As mentioned before, this is because the motion of the structure is dominated by the first mode in the case of ground shaking and so using the measurements at one degree of freedom is sufficient to characterize the motion of the structure. Note that although Controller 3 gives the smallest probability of failure in Table 5.2, the performance quantities in Table 5.3 are almost the same for all optimal controllers.

Controller 4 is the case in which the controller feeds back the measurements at past time steps only. Although its robust failure probability is slightly larger than Controller 3 in Table 5.2, the performance quantities in Table 5.3 are virtually the same as Controller 3. Moreover, this controller does not suffer from time-delay induced stability problems if the time-delay of the controller Δt_d is less than Δt. If Δt_d is larger than Δt, one can choose $N_p > \frac{\Delta t_d}{\Delta t}$ and fix all the matrices $\mathbf{G}_o, \ldots, \mathbf{G}_{\mathrm{INT}(\frac{\Delta t_d}{\Delta t})}$ at zero. Here, INT denotes the integer part of a number. The controller feeds back the measurements far back enough that the control system has enough time to compute and apply the control force to the structure. Fig. 5.6 and 5.7 shows the similar control force (normalized by the actuator mass) and stroke time histories respectively for Controllers 1 - 4.

In order to test the robustness of the proposed controller to the excitation, the structural responses are calculated for the uncontrolled structure and the controlled

Performance quantity	Threshold	Uncontrolled	Passive	Controller 1	Controller 2	Controller 3	Controller 4
σ_{x_1} (m)	.....	0.0143	0.0075	0.0042	0.0042	0.0040	0.0040
$\sigma_{x_1-x_2}$ (m)	.....	0.0138	0.0072	0.0039	0.0039	0.0037	0.0037
$\sigma_{x_2-x_3}$ (m)	.....	0.0095	0.0050	0.0029	0.0028	0.0027	0.0028
$\sigma_{x_3-x_4}$ (m)	.....	0.0053	0.0029	0.0021	0.0020	0.0019	0.0020
$max\|x_1\|$ (m)	0.02	0.0373	0.0213	0.0120	0.0122	0.0114	0.0115
$max\|x_1 - x_2\|$ (m)	0.02	0.0374	0.0197	0.0117	0.0116	0.0113	0.0114
$max\|x_2 - x_3\|$ (m)	0.02	0.0257	0.0143	0.0088	0.0088	0.0086	0.0087
$max\|x_3 - x_4\|$ (m)	0.02	0.0134	0.0085	0.0059	0.0058	0.0059	0.0060
σ_{x_s} (m)	.....	.....	0.1019	0.4056	0.3934	0.4101	0.4071
$\sigma_{f_{cn}}$ (g)	.....	.....	.....	2.7764	2.6656	2.7785	2.7905
$max\|x_s\|$ (m)	2.0	.....	0.2984	1.0756	1.0374	1.0897	1.0895
$max\|f_{cn}\|$ (g)	10.0	.....	.....	8.0942	7.9589	8.3669	8.509

Table 5.3: Statistical properties of the performance quantities (Example 5-1).

structure (using Controller 3) subjected to the 1940 El Centro earthquake record. In Fig. 5.8, the dashed line and the solid line show the first story drifts for the uncontrolled structure and the controlled structure, respectively. It can be seen that the structural response is significantly reduced by using the proposed controller. In this case, the peak control force normalized by the actuator mass is 7.1g and the peak actuator stroke is 1.08m.

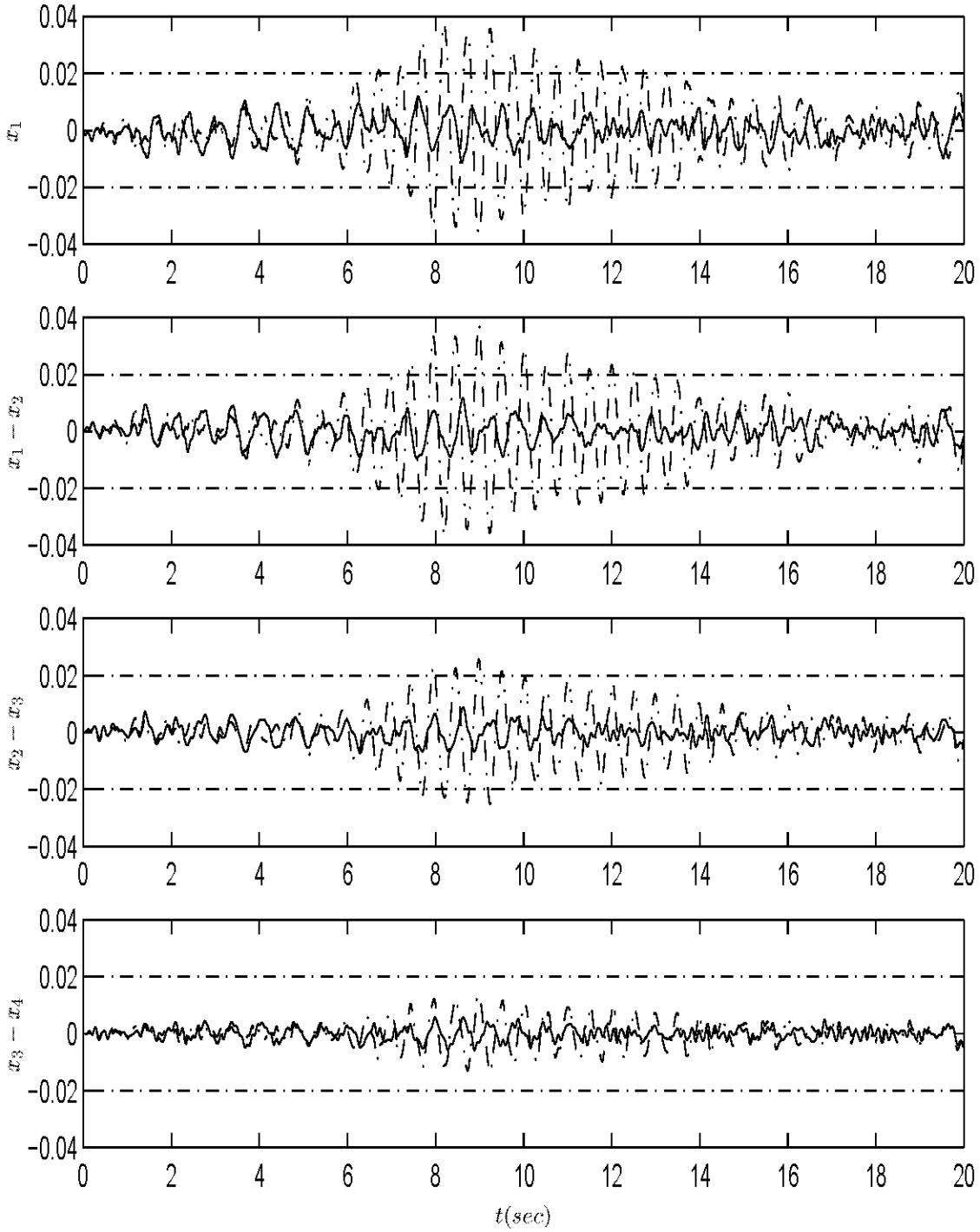

Figure 5.2: Simulated interstory drifts for the uncontrolled (dashed) and controlled structure using Controller 1 (solid) (Example 5-1).

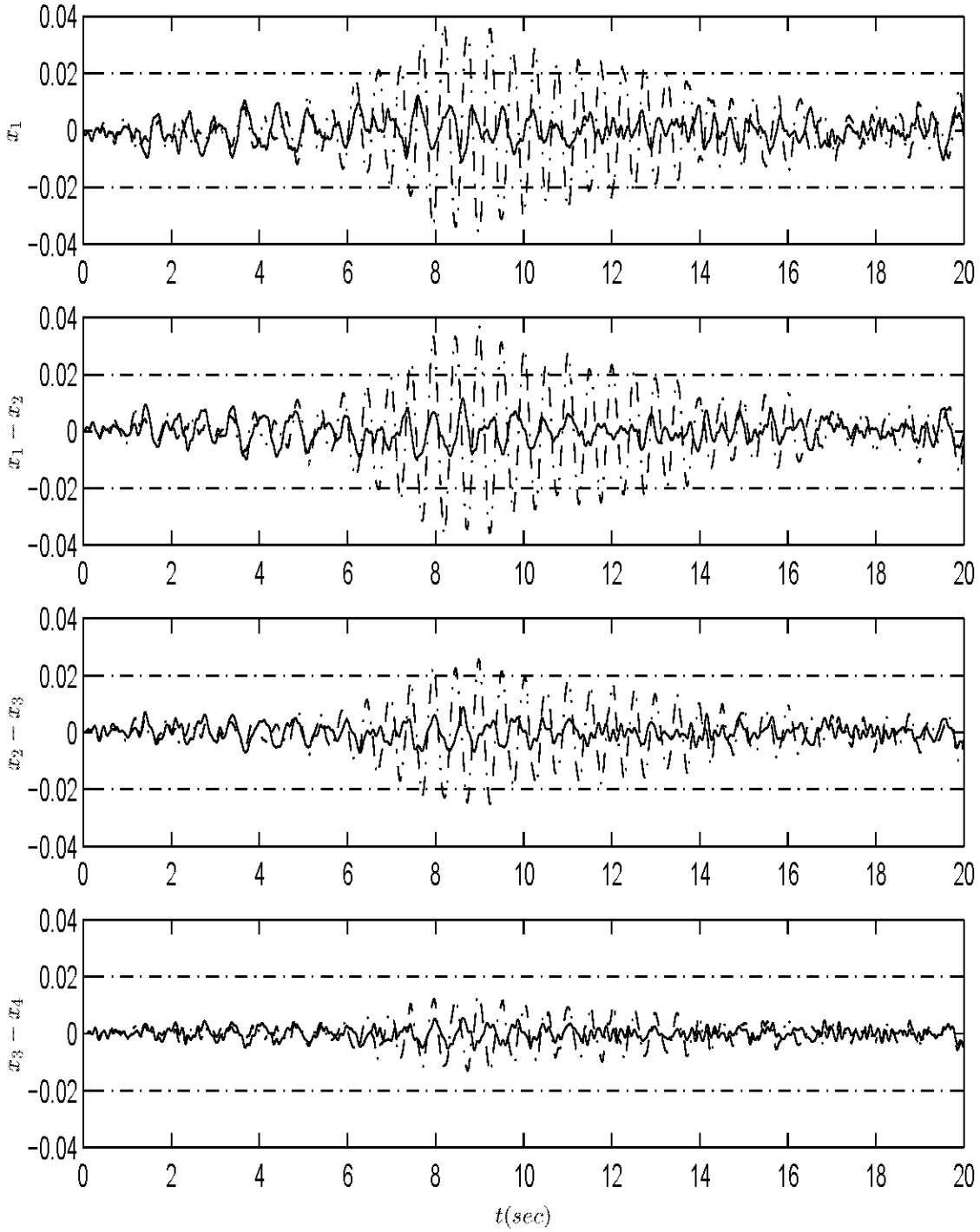

Figure 5.3: Simulated interstory drifts for the uncontrolled (dashed) and controlled structure using Controller 2 (solid) (Example 5-1).

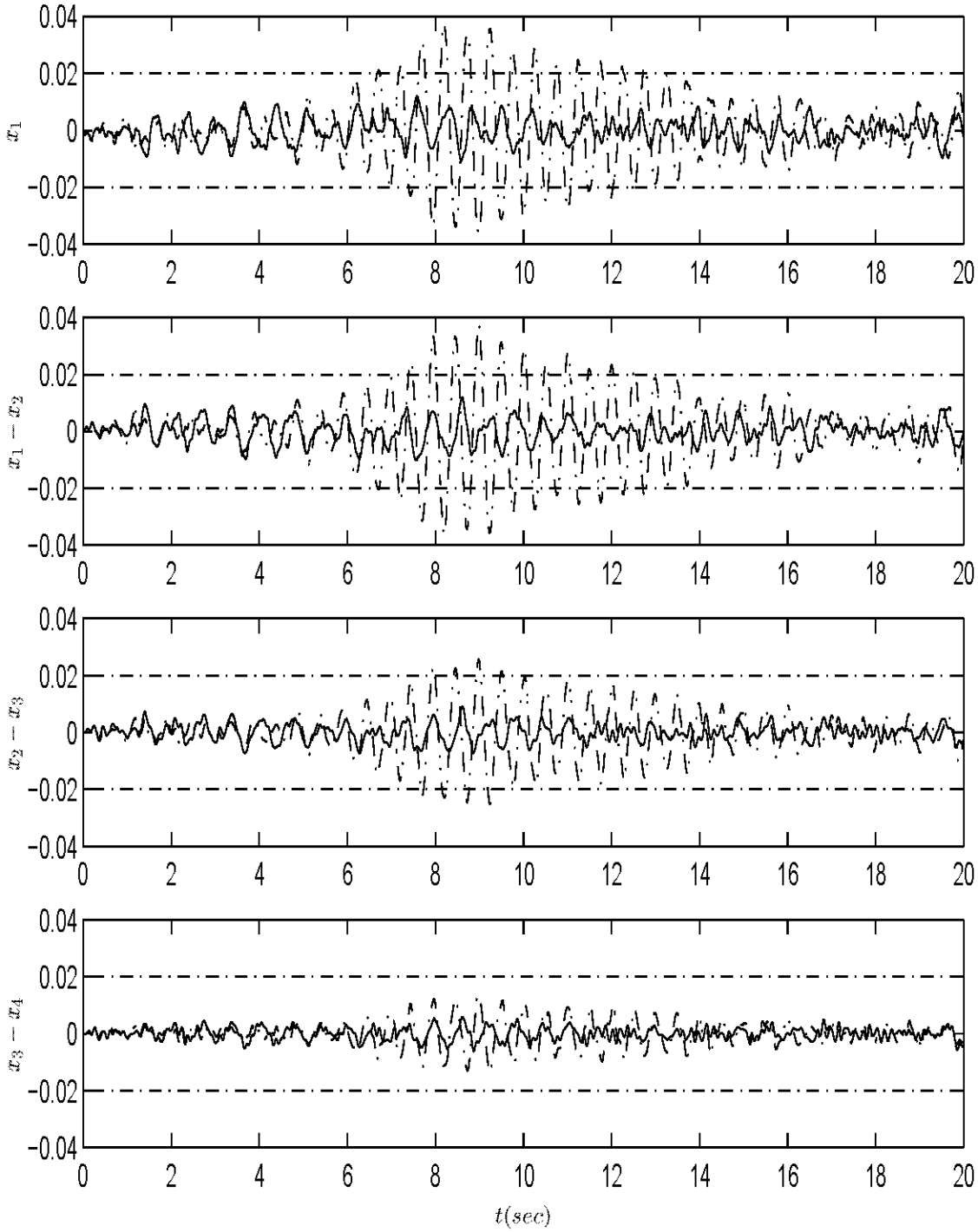

Figure 5.4: Simulated interstory drifts for the uncontrolled (dashed) and controlled structure using Controller 3 (solid) (Example 5-1).

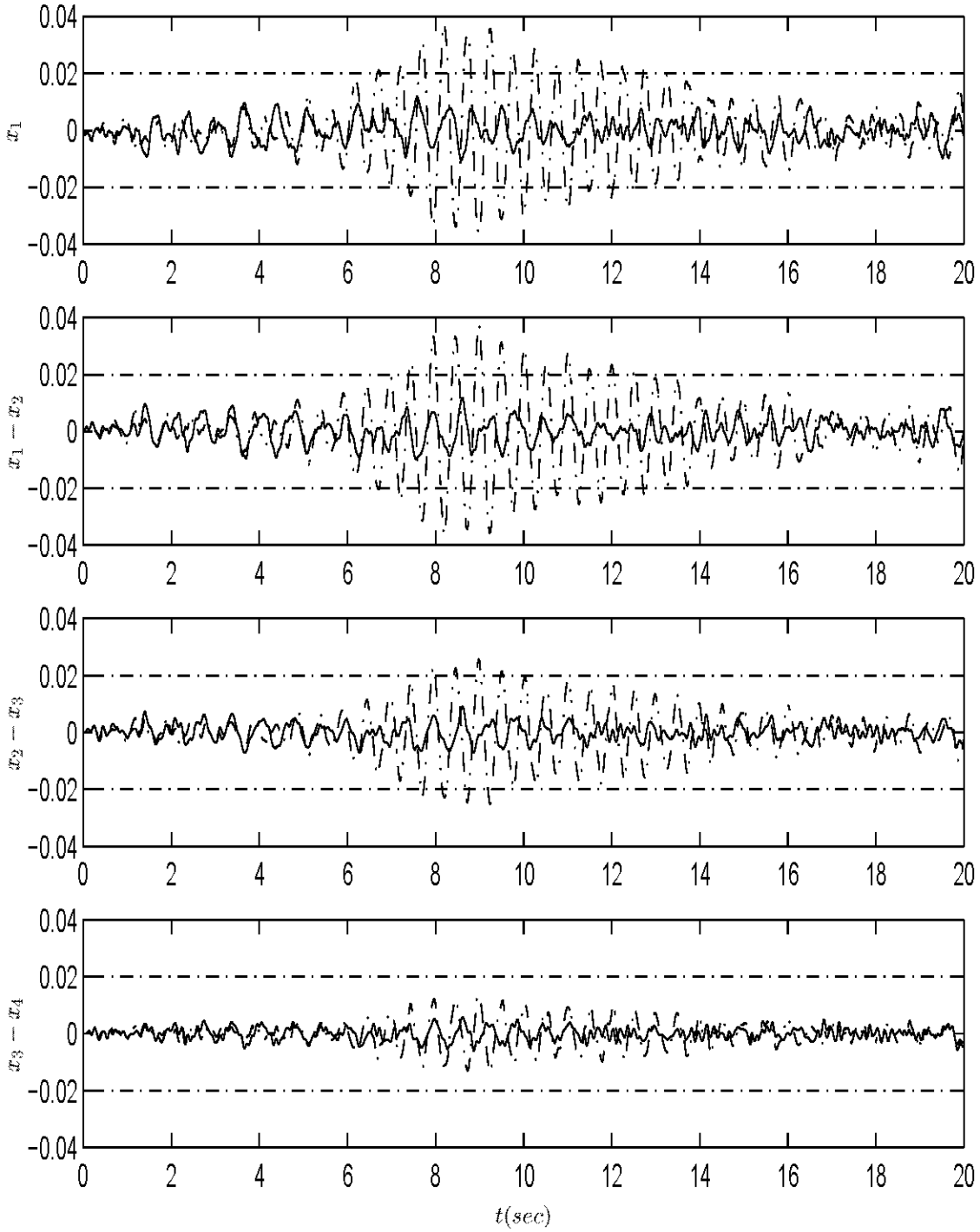

Figure 5.5: Simulated interstory drifts for the uncontrolled (dashed) and controlled structure using Controller 4 (solid) (Example 5-1).

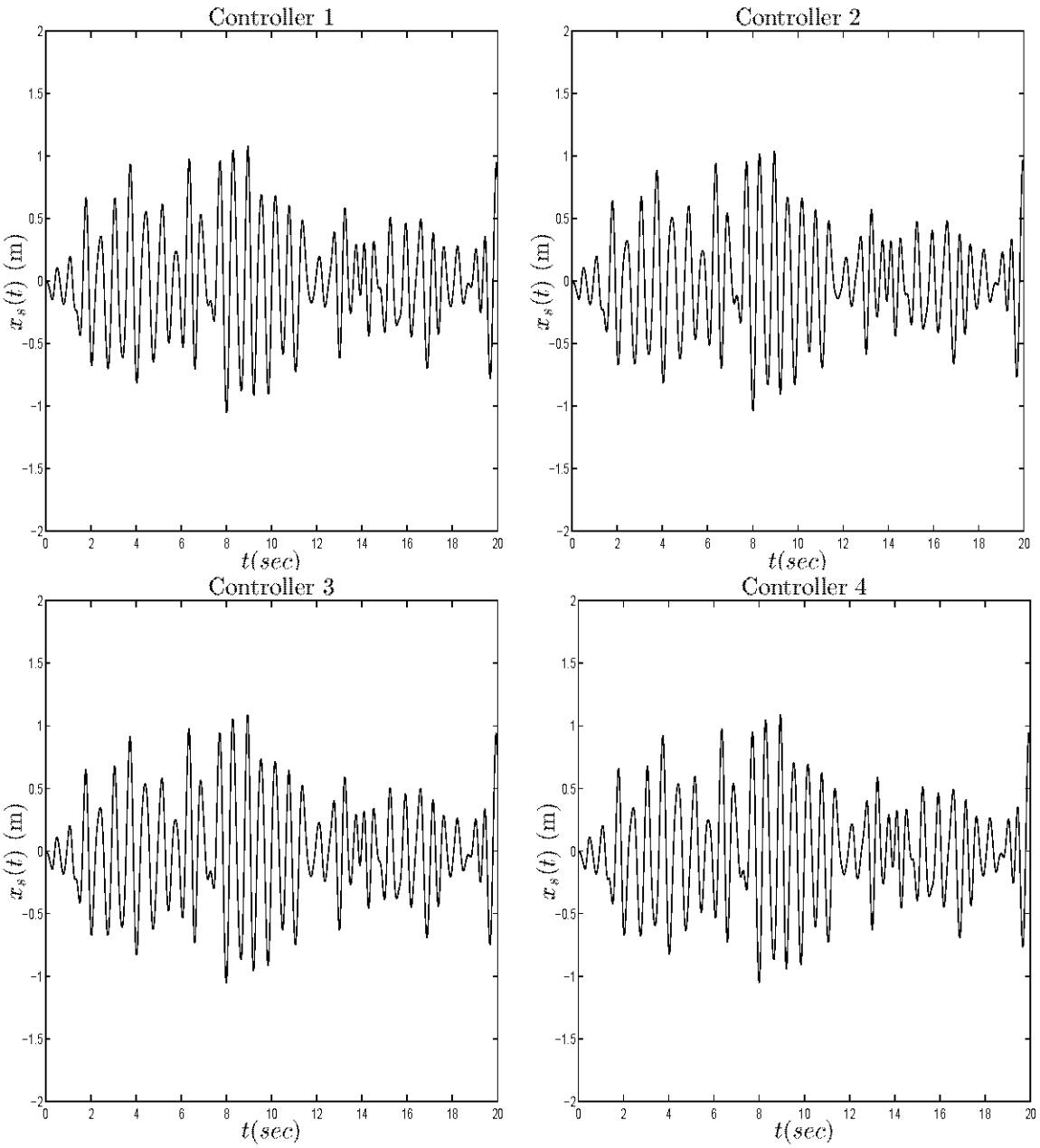

Figure 5.6: Controller stroke time histories using Controllers 1 - 4 (Example 5-1).

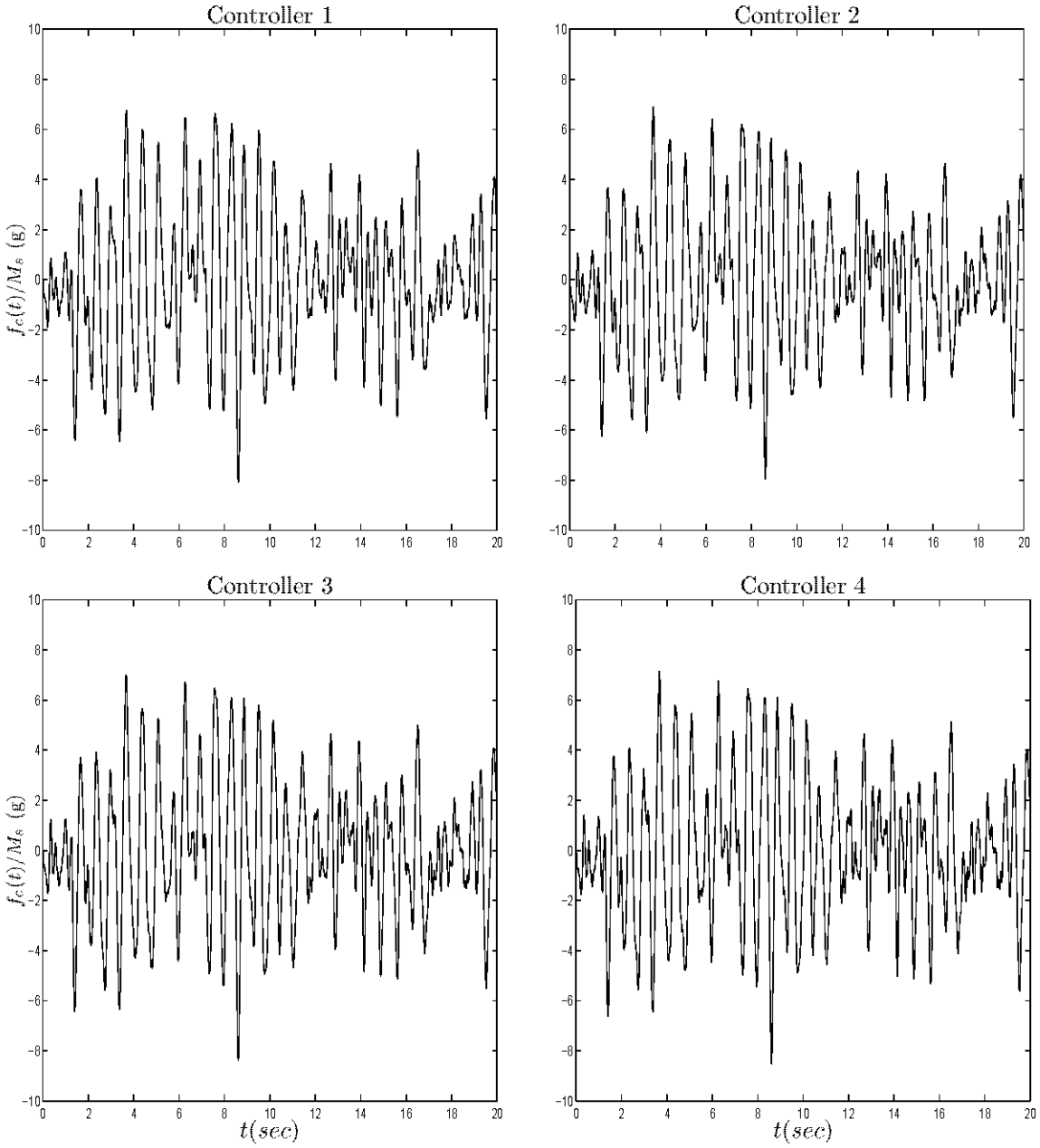

Figure 5.7: Controller force (normalized by the actuator mass) time histories using Controllers 1 - 4 (Example 5-1).

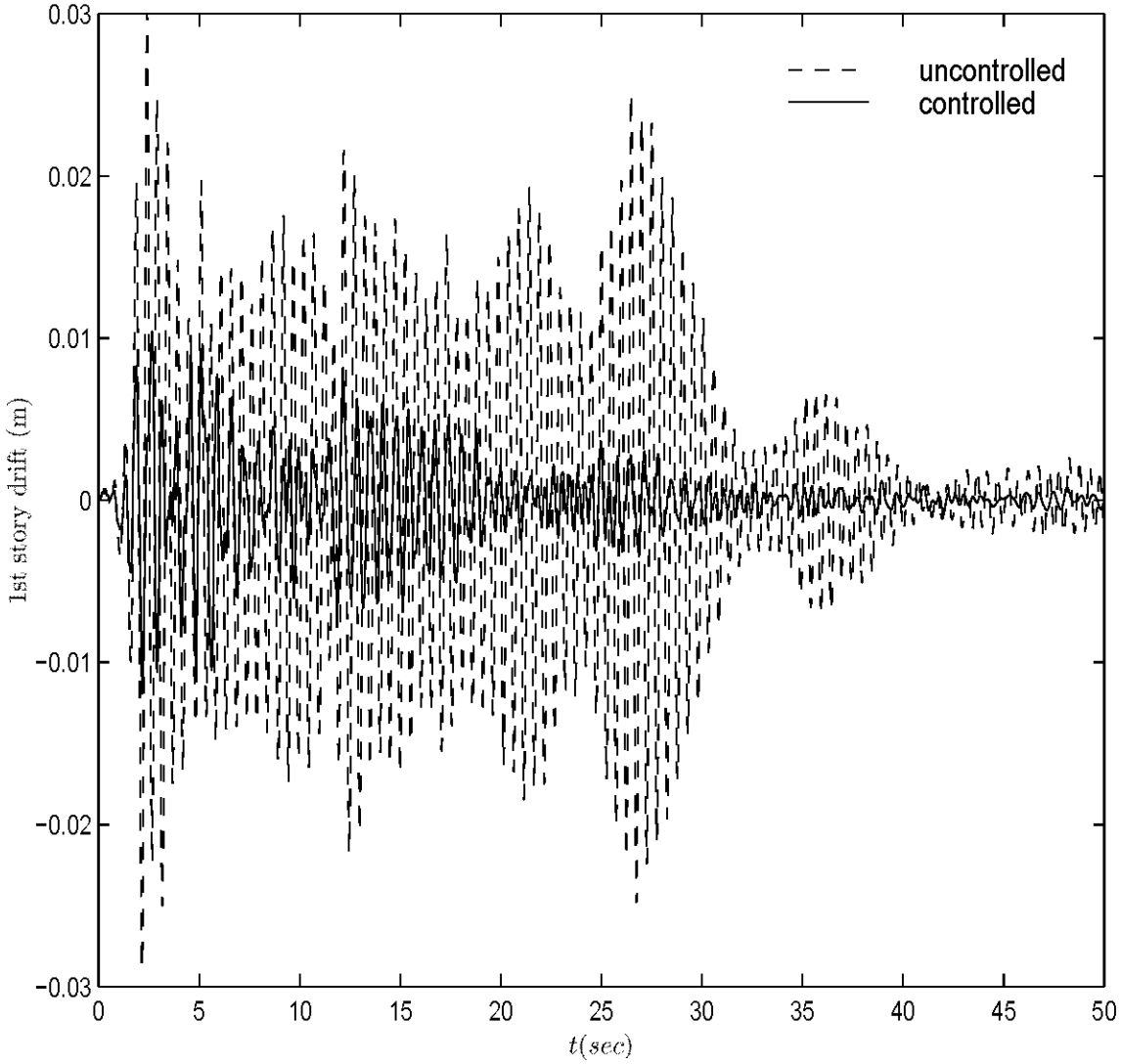

Figure 5.8: Structural response of the uncontrolled (dashed) and controlled structure using Controller 3 (solid) to the El Centro earthquake record (Example 5-1).

5.4.2 Example 5-2: Control Benchmark Problem

The proposed control strategy is applied to the well-known control benchmark problem with an active mass driver (Spencer et al. 1998). The benchmark problem is based on a three-story, single-bay laboratory test structure (Dyke et al. 1996). It is a steel frame with 158cm height. The natural frequencies of the first three modes are 5.81Hz, 17.68Hz and 28.53Hz, respectively. The associated damping ratios are 0.33%, 0.23% and 0.30%. In this example, the structural system is assumed known (an accurate dynamic model is given in the benchmark, but the stochastic excitation model is treated as uncertain). The controllers are designed and tested under the excitation of a Kanai-Tajimi filtered white noise, and further tested using a scaled 1940 El Centro earthquake record and a scaled 1968 Hachinohe earthquake record. The sampling time intervals is $\Delta t = 0.001$ sec, as specified by the benchmark. The threshold levels for the interstory drifts, actuator displacements and actuator accelerations are 1.5cm, 9.0cm and 6.0g, respectively. As the delay time of the control force is $\Delta t_d = 0.0002$ sec, the controllers in this study are chosen to feedback only the response measurements from one and two time steps back, that is, $\mathbf{G}_0$ is fixed to be zero and $\mathbf{G}_i, i = 1, 2$ are the design parameters. Two feedback cases were investigated as follows:

Controller 1: Feedback of acceleration from all floors at the previous two time steps, i.e., $\mathbf{G}_i, i = 1, 2$ are the design parameters.

Controller 2: Acceleration measurements from all floors are passed through the same second order filter, with transfer function $\omega_c^2/(-\omega^2 + 2i\zeta_c\omega_c\omega + \omega_c^2)$. Then, the controller feeds back the filtered measurements at the previous two time steps. Here, ζ_c is chosen to be $1/\sqrt{2}$ and ω_c is included in the design parameter set. This case has been previously studied using only output of the filter at the current time (May and Beck 1998).

Following the benchmark guidelines (Spencer et al. 1998), the controllers are used to control a high-fidelity linear time-invariant state-space representation of the structure which has 28 states. Quantization, saturation and time delay of the control

force are considered in this model. In order to test the robustness of the controllers with respect to modeling errors, a reduced 10-state model is used in the design process, which is provided by the official benchmark web site at http://www.nd.edu/~quake/. Furthermore, the excitation is assumed to be a stationary zero-mean Gaussian process with a spectral density defined by an uncertain Kanai-Tajimi spectrum:

$$S_{\ddot{x}_g\ddot{x}_g} = S_0 \frac{4\zeta_g^2\omega_g^2\omega^2 + \omega_g^4}{(\omega^2 - \omega_g^2)^2 + 4\zeta_g^2\omega_g^2\omega^2} \tag{5.28}$$

where ω_g, ζ_g are assumed to be log-normally distributed with mean 50 rad/sec and 0.5, respectively. Furthermore, their logarithm standard deviations are assumed to be $\sigma_{\log\omega_g} = 0.2$ and $\sigma_{\log\zeta_g} = 0.2$. The spectral intensity parameter S_0 is given by

$$S_0 = \frac{0.03\zeta_g}{\pi\omega_g(4\zeta_g^2 + 1)} \mathrm{g}^2\,\mathrm{sec} \tag{5.29}$$

such that $\sigma_{\ddot{x}_g} = 0.12\mathrm{g}$ regardless of the values of ω_g and ζ_g.

Gain\Controller	1	2
$G_1(1)$	0.0062	0.0930
$G_1(2)$	0.0014	0.0959
$G_1(3)$	0.0228	0.0931
$G_2(1)$	0.0319	0.1268
$G_2(2)$	0.0494	0.1056
$G_2(3)$	0.0838	0.1047
ω_c (rad/sec)	—	44.993

Table 5.4: Design parameters of the optimal controllers (Example 5-2).

Table 5.4 shows the optimal gains and the optimal filter parameter for Controllers 1 and 2. One can see that the control gains increase significantly when using the filter. Table 5.5 shows the performance quantities J_1 to J_{10} defined in Spencer et al. (1998) for Controllers 1 and 2, for the controller obtained by May and Beck (1998) and also for the sample controller provided in Spencer et al. (1998). All the controllers provide satisfactory performance. Note that the controller obtained by May and Beck

Excitation	Performance quantity	Controller 1	Controller 2	May and Beck	Spencer et al.
Filtered white noise	J_1	0.183	0.205	0.207	0.283
	J_2	0.301	0.310	0.345	0.440
	J_3	0.366	0.736	0.851	0.510
	J_4	0.363	0.738	0.832	0.513
	J_5	0.606	0.676	0.683	0.628
Maximum response of Hachinohe 1968 and El Centro 1940	J_6	0.492	0.380	0.380	0.456
	J_7	0.811	0.694	0.684	0.681
	J_8	0.812	1.39	1.64	0.669
	J_9	0.847	1.35	1.56	0.771
	J_{10}	1.64	1.16	0.936	1.28

Table 5.5: Performance quantities for the benchmark problem (Example 5-2).

is similar to Controller 2 except that they only feed back the response measurements at the current state. Their optimal gains are $G_0(1) = 0.431$, $G_0(2) = 0.291$, and $G_0(3) = 0.235$ and the optimal filter parameter is $\omega_c = 33.1$rad/sec. J_1 to J_5 correspond to the case of uncertain excitation for 300 sec. J_1 and J_2 correspond to the standard deviations of the maximum RMS drifts and the maximum RMS absolute acceleration of the controlled structure over all of the floors, normalized by the corresponding values for the uncontrolled structure. J_3, J_4 and J_5 correspond to the RMS actuator displacement relative to the third story, the RMS relative actuator velocity and the RMS absolute actuator acceleration. Again, they are normalized by their corresponding values for the uncontrolled structure. J_6 to J_{10} represent the peak values of the same response quantities for the deterministic response of the controlled structure to the two scaled earthquake ground motions, the north-south component of the 1940 El Centro earthquake record and the north-south component of the 1968 Hachinohe earthquake record. Again, these quantities are normalized by the peak response quantities of the uncontrolled structure for each earthquake.

May and Beck (1998) showed that directly feeding back the accelerations at the current time without a compensator leads to an unstable controlled system due to the delay-time imposed in the model of the system to be controlled (Spencer et al. 1998). However, Controller 1 provides satisfactory performance using direct feedback

of delayed accelerations because the delay-time is explicitly taken into consideration in the formulation, as described in Section 5.2. In May and Beck (1998), a filter was used in the feedback loop to produce stability. When a filter is used here (Controller 2), the control system is not as efficient as in Controller 1 when subjected to random excitation because certain information, especially the high frequency content, is filtered out. However, Table 5.5 shows it provides better performance for the El Centro and the Hachinohe earthquake records, which do not follow the Kanai-Tajimi spectrum closely.

Fig. 5.9 shows the 1st story drift for both earthquakes using Controller 2 which has the filter (solid lines). For comparison purposes, the dashed lines show the corresponding 1st story drifts of the uncontrolled structure. It can be seen that the 1st story drifts are significantly reduced by using the proposed control methodology. Fig. 5.10 shows the actuator displacements for both earthquakes. It can be seen that they are much smaller than the threshold values.

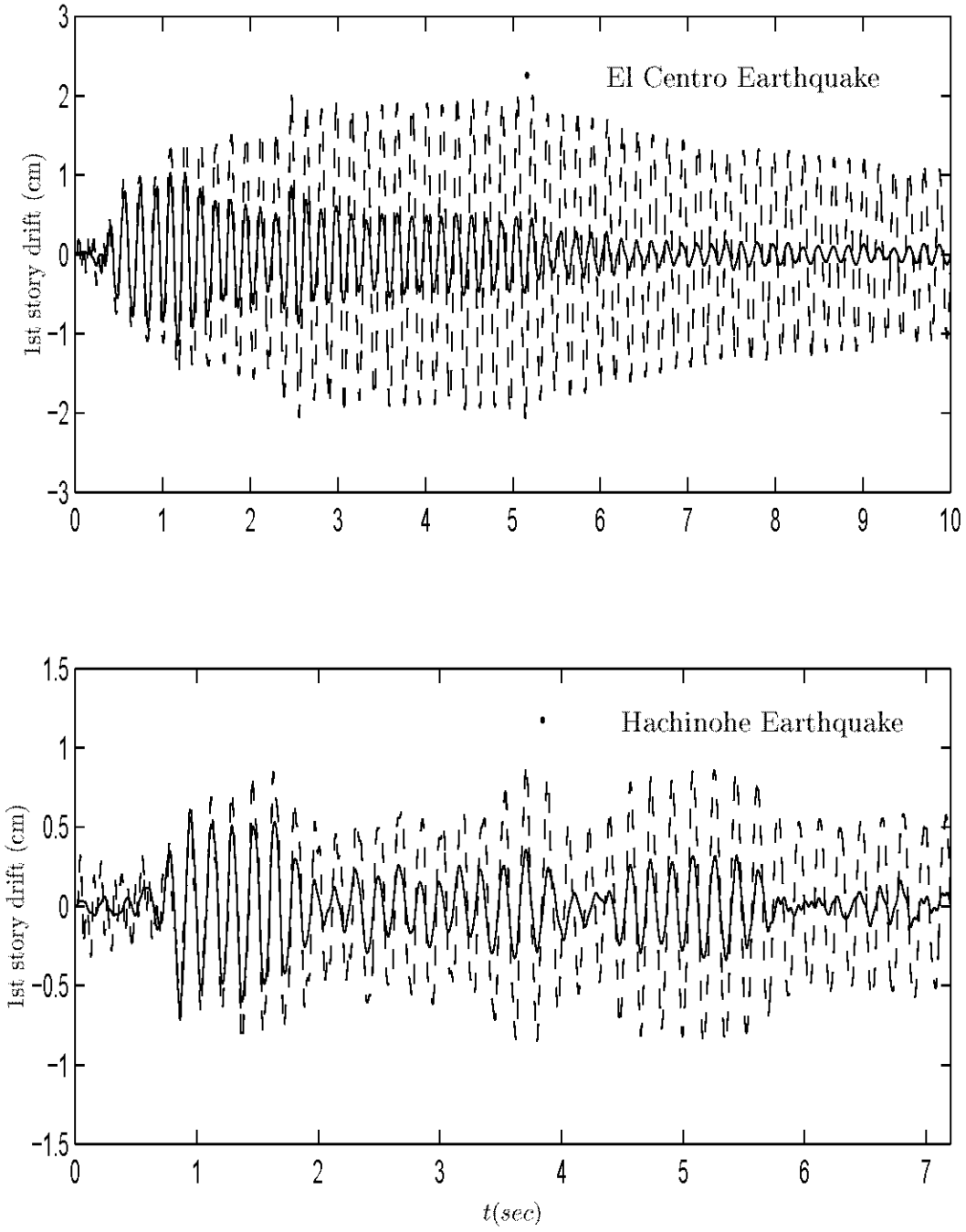

Figure 5.9: Structural response of the uncontrolled (dashed) and controlled structure using Controller 2 (solid) to the El Centro and Hachinohe earthquake records (Example 5-2).

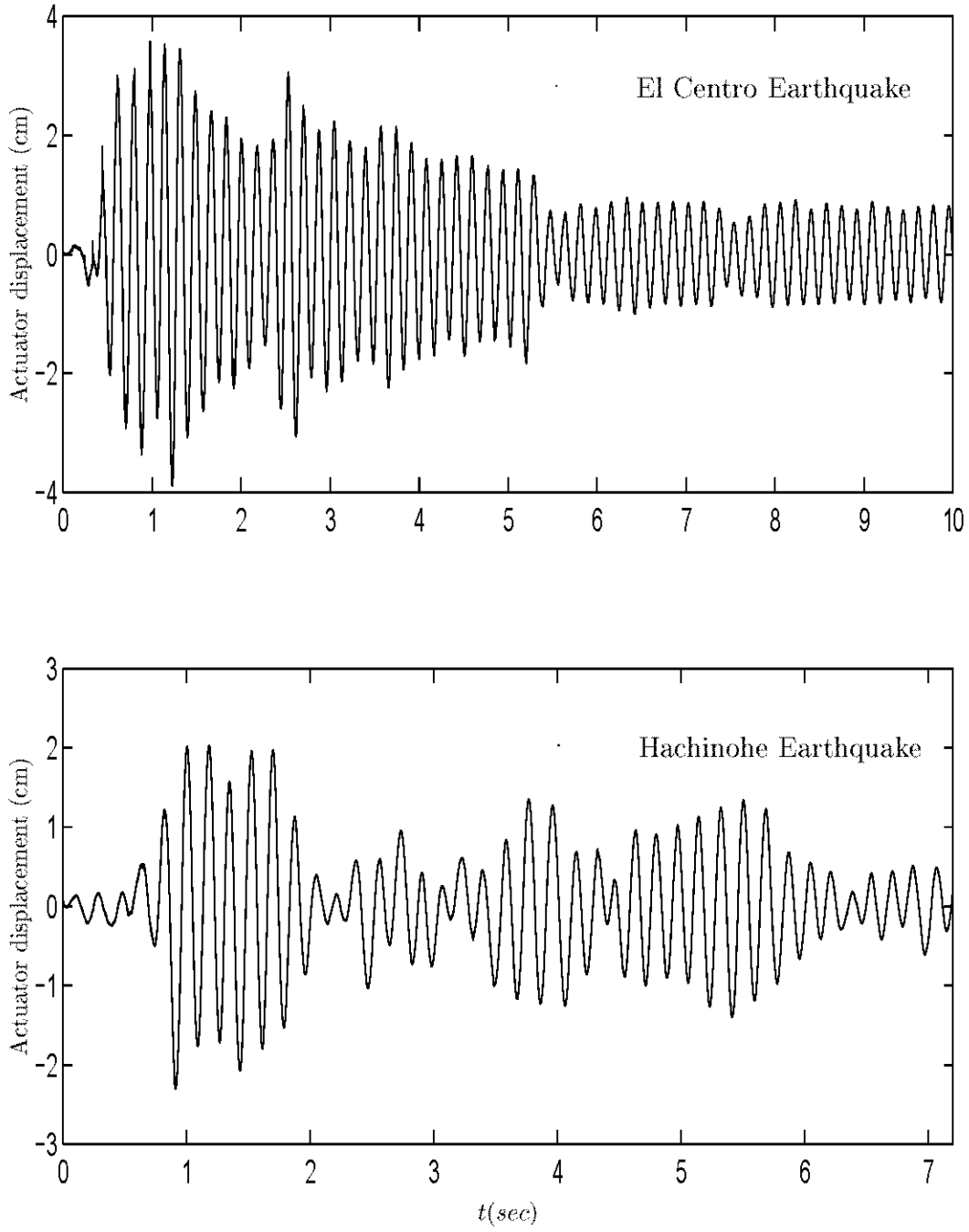

Figure 5.10: Actuator displacement using Controller 2 to the El Centro and Hachinohe earthquake records (Example 5-2).

5.5 Conclusion

A reliability-based robust feedback control approach was presented for dynamical systems adequately represented by linear state space models. The response covariance matrix is first obtained from the discrete Lyapunov equation using an augmented vector for the system. The optimal controller is then chosen from a set of possible controllers so that the robust reliability of the controlled system is maximized or, equivalently, the robust failure probability is minimized. An asymptotic approximation is utilized to evaluate high dimensional integrals for the robust failure probability. The feedback of the past output provides additional information about the system dynamics to the controller. It can also be used to avoid stability problems due to time-delay effects. The proposed approach does not require full state measurements or a Kalman filter to estimate the full state. The robust failure probability criterion provides robustness of the control for both uncertain excitation models and uncertain system models. Furthermore, it can give different weighting to the different possible values of the model parameters by using a probability description of these parameters based on engineering judgement or obtained from system identification techniques. This is in contrast to most current robust control methods which split the values for the system parameters into only two groups (possible or impossible). Although the proposed approach was presented here for linear models of dynamical systems, it can be extended to nonlinear models. The only difference is that the second order moments can not be obtained by solving the Lyapunov equation, but this can be replaced by approximate numerical techniques; for example, simulations can be utilized to obtain the response covariance matrix. We are currently investigating this extension to robust control of nonlinear structural behavior.

Chapter 6 Illustrative Example of Robust Controller Design and Updating

6.1 Problem Description

In this chapter, a 20-DOF three-bay four-story structural frame (Fig. 6.1) is used to demonstrate the probabilistic procedure of robust controller design and updating.

The stiffness-to-mass ratio is taken as $EI_{c1}/m_1 = 1.25 \times 10^5 \mathrm{m}^4\,\mathrm{sec}^{-2}$, where m_1 is the mass of the first floor. Furthermore, $m_1 = m_2 = m_3 = 1.2m_4$ and $[EI_{c2}, EI_{c3}, EI_{c4}] = [0.9, 0.8, 0.7]EI_{c1}$. The rigidity of the beams is taken to be $EI_b/m_1 = 2.00 \times 10^6 \mathrm{m}^4\,\mathrm{sec}^{-2}$. The first four natural frequencies of the structure are 4.108Hz, 11.338Hz, 17.26Hz and 21.50Hz. Rayleigh damping is assumed, so the damping matrix $\mathbf{C}$ is given by $\mathbf{C} = \alpha_m\mathbf{M} + \alpha_k\mathbf{K}$, where $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices of the system; and $\alpha_m = 0.376\,\mathrm{sec}^{-1}$, $\alpha_k = 2.07 \times 10^{-4}\,\mathrm{sec}$, which gives 1.0% damping for the first two modes.

6.2 Model Selection and Identification

Two candidate classes of models, with the same height and width of the system, are considered as shown in Fig. 6.2:

Class A: Eight-DOF structural frame models, with four rigidity parameters. The nominal rigidity of the beam is assumed to be its exact values but the nominal rigidity of the columns are taken to be $2EI_{c1}$, twice the exact values of the columns at the first story. The rigidity of the beam and columns at the j^{th} story is equal to the product of the rigidity parameter θ_j and its corresponding nominal value.

Class B: Four-DOF shear building models, with four rigidity parameters. The nominal values of the rigidity of the columns are taken to be $2EI_{c1}$, $2EI_{c2}$, $2EI_{c3}$

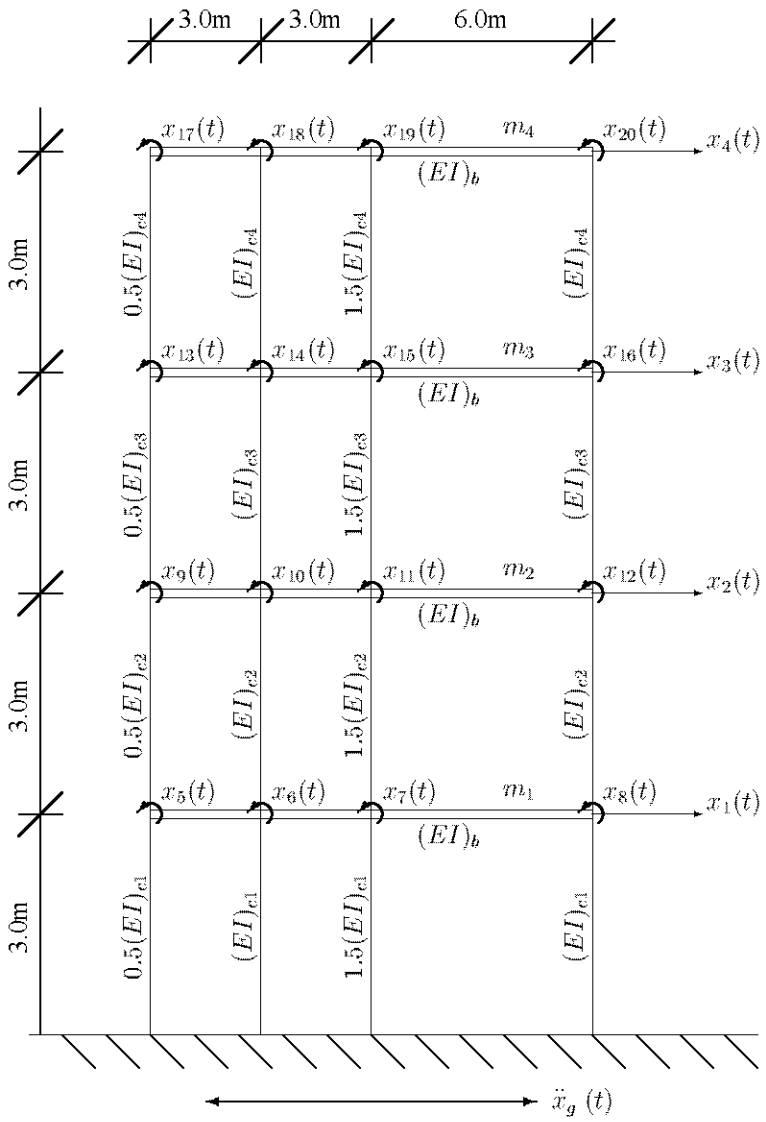

Figure 6.1: Four-story structural frame (Example 6-1).

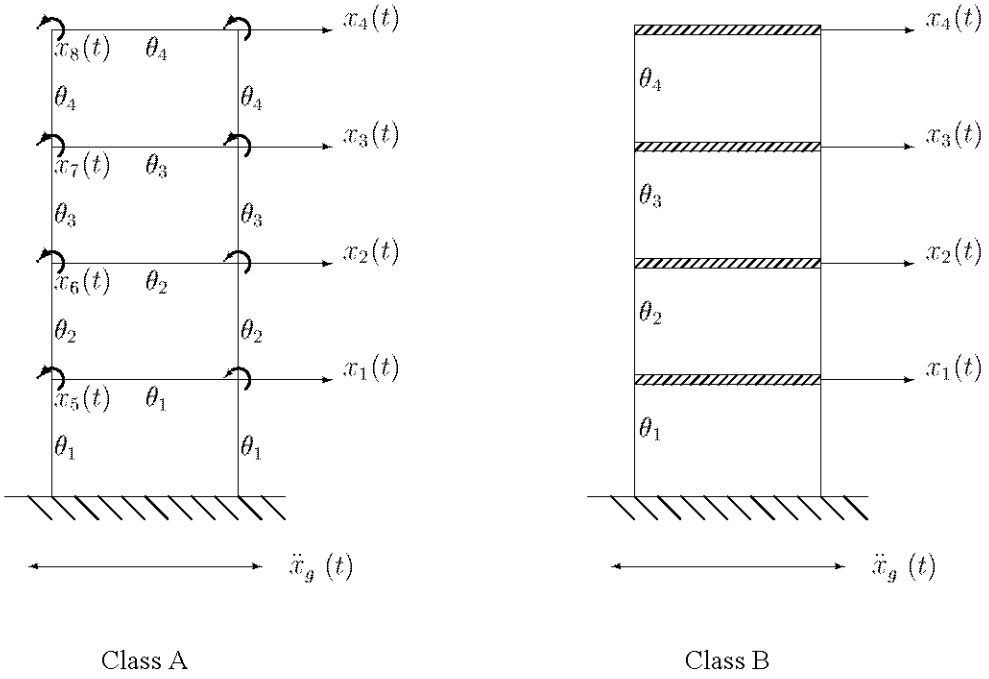

Figure 6.2: Candidate model classes (Example 6-1).

and $2EI_{c4}$ for the 1^{st} to 4^{th} floor, respectively. Note that the nominal values in this case do not affect the identification results because they are only scaling, but not constraints, of the parameters.

Assume that measurements of the absolute accelerations are available at the 2^{nd} and 4^{th} DOFs for 30 sec with a sampling frequency 200 Hz. These data are simulated using the actual model with 10% rms noise added. Furthermore, assume that the system is subjected to a white noise ground motion with spectral intensity $S_{f0} = 1.0 \times 10^{-3}\mathrm{m}^2\,\mathrm{sec}^{-3}$.

Note that although the model selection and identification approaches were presented in separate chapters, they have to work together. Here, the identification approach presented in Chapter 3 is utilized to update the rigidities and damping ratios of the structure. Although the approach in Chapter 3 is presented for modal updating for linear systems, it can be applied directly for updating the model parameters of a linear structure. In order to have better scaling, the rigidities are parameterized as follows: $EI_j = \theta_j \widetilde{EI}_j, j = 1,2,3,4$, where EI_j denotes the j^{th} story rigidity and $\widetilde{EI}_j$ is its nominal value. The rigidity parameters θ_j are considered unknown and are

determined by identification.

Table 6.1 shows the exact values and the identified rigidity parameters and damping ratios for the two classes of models. Note that the damping ratios are presented in percentages. Table 6.2 shows the corresponding frequencies for the actual and the optimal models. One can see that the optimal model in Model Class A fails to fit the natural frequency of the third mode. Furthermore, the identified fourth mode of this model is not close to an actual mode of the system. Therefore, the damping ratio of this mode is much larger than others.

Parameter	θ_1	θ_2	θ_3	θ_4	ζ_1	ζ_2	ζ_3	ζ_4
Exact	—	—	—	—	1.000	1.000	1.296	1.537
Model Class A	1.061	2.542	0.551	0.480	1.286	0.395	1.436	9.964
Model Class B	0.908	0.987	0.881	1.040	1.128	0.563	2.342	0.623

Table 6.1: Optimal (most probable) structural parameters in each model class representing the structural frame (Example 6-1).

Mode	f_1	f_2	f_3	f_4
Actual	4.108	11.34	17.26	21.50
Model Class A	4.064	11.31	22.71	33.14
Model Class B	4.158	11.35	17.40	21.02

Table 6.2: Natural frequencies (in Hz) of the optimal model in each class (Example 6-1).

Fig. 6.3 and 6.4 show the updated PDFs (solid) for the rigidity parameters $\theta_j, j = 1, 2, 3, 4$ for Model Class A and Model Class B, respectively. The crosses correspond to the Gaussian approximation. It can be seen that the Gaussian approximation is very accurate.

Here, $P(\mathcal{M}_A|\mathcal{U}) = P(\mathcal{M}_B|\mathcal{U}) = 0.5$ is assumed, implying that there is no prior preference between these two classes of models. By using the model selection approach presented in Chapter 2, it is found that $P(\mathcal{M}_A|\mathcal{D},\mathcal{U}) = 5.8 \times 10^{-25}$ and $P(\mathcal{M}_B|\mathcal{D},\mathcal{U}) = 1.0$. Therefore, it is suggested that Model Class B is much better

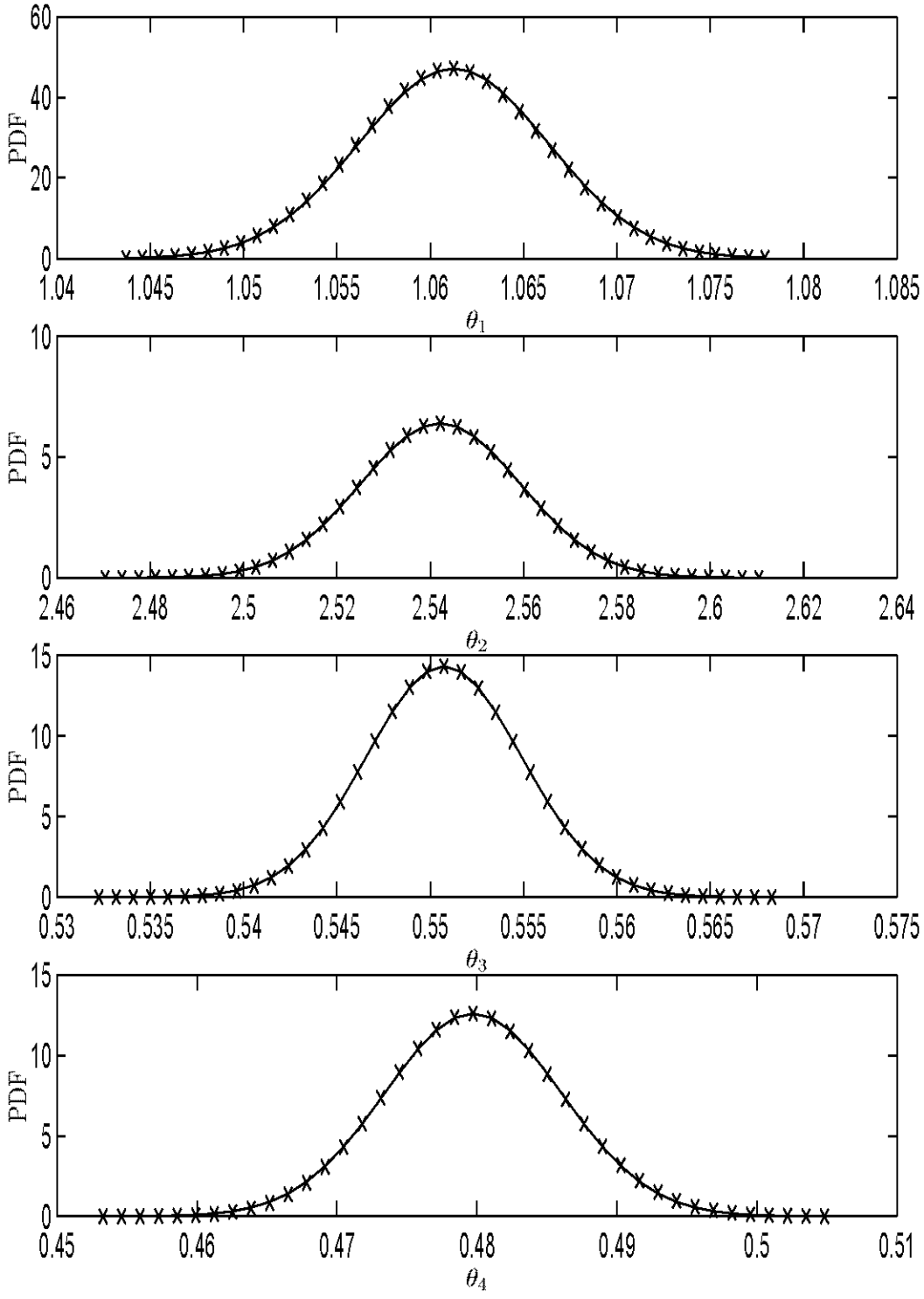

Figure 6.3: Updated PDFs of the rigidity parameters for Model Class A obtained from: i) Eqn. 3.15 - **cross**; and ii) Gaussian approximations - **solid** (Example 6-1).

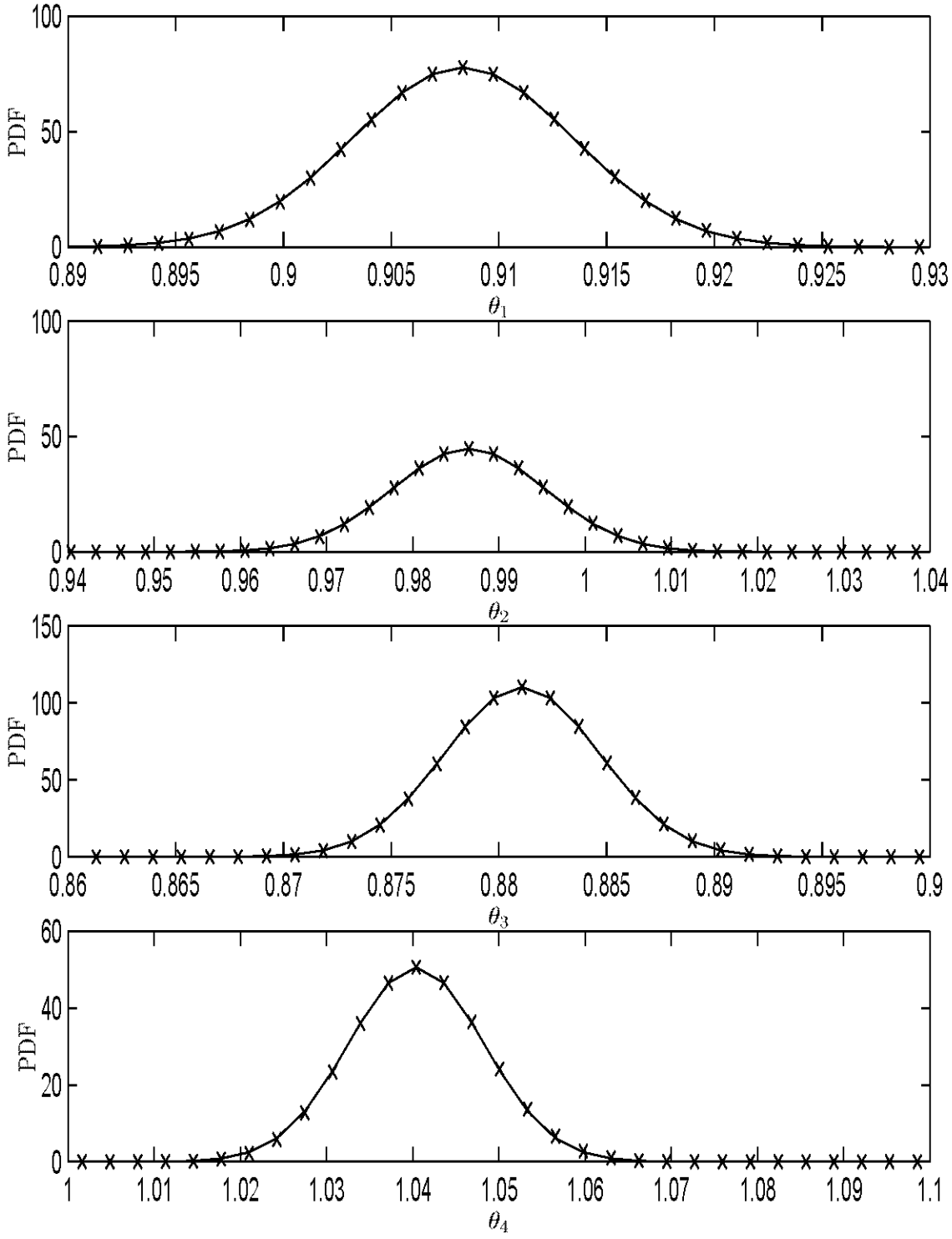

Figure 6.4: Updated PDFs of the rigidity parameters for Model Class B obtained from: i) Eqn. 3.15 - **cross**; and ii) Gaussian approximations - **solid** (Example 6-1).

than Model Class A. Therefore, Model Class B is used for the controller design. Furthermore, the updated PDF of the rigidity parameters is used for calculating the robust failure probability in the controller design.

6.3 Controller Design

The updated PDF of the rigidity parameters and the damping ratios of Model Class B (Fig. 6.4) is used for calculating the robust failure probability in the controller design. First, the fundamental mode of the identified model (Model Class B in Table 6.1) is used to design the stiffness and damping of the AMD (active mass damper), i.e., $\tilde{f}_1 = 4.158$Hz and $\zeta_1 = 1.128\%$. The AMD mass M_s is chosen to be 1% of the mass of the building. By using Eqn. 5.27, the controller stiffness and damping parameters are given by $k_s/M_s = 663.7\,\mathrm{sec}^{-2}$ and $c_s/M_s = 5.13\,\mathrm{sec}^{-1}$. However, these are rounded so that $k_s/M_s = 670\,\mathrm{sec}^{-2}$ and $c_s/M_s = 6.0\,\mathrm{sec}^{-1}$ to give a natural frequency and damping ratio approximately equal to that of the identified fundamental mode of the structure.

In Dyke et al. (1995), hydraulic actuators are modeled as follows:

$$\dot{f}_c = A_f f_c + B_f \dot{x}_a + B_{fu} u \tag{6.1}$$

where f_c is the control force applied by the actuator; $\dot{x}_a$ is the actuator velocities; u is the signal given to the actuator; and A_f, B_f and B_{fu} are given by

$$A_f = -\frac{2\beta k_a}{V}, B_f = -\frac{2\beta A^2}{V}, B_{fu} = \frac{2\beta A k_q}{V} \tag{6.2}$$

where β is the bulk modulus of the fluid; k_a and k_q are the controller constants; V is the characteristic hydraulic fluid volume of the actuator; and A is the cross-sectional area of the actuator. Schematically, the structure-actuator is shown in Fig. 6.5.

The output vector $\mathbf{y}_f$ in Eqn. 5.4 is comprised of: $\mathbf{y}_f = [f_c, \tilde{f}_c, \dot{\tilde{f}}_c]^T$, where $\tilde{f}_c$ and $\dot{\tilde{f}}_c$ are the state vectors for a low-pass filter with input f_c that approximates differentiation of f_c. Note that $\dot{\tilde{f}}_c$ is used to estimate the out-crossing rate of the

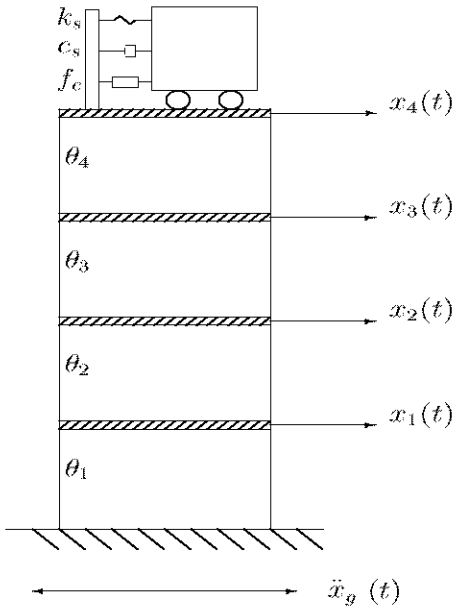

Figure 6.5: Structure-actuator model (Example 6-1).

control force and it is not used in the control system feedback.

The full state vector equation is given by

$$\dot{\mathbf{v}} = \mathbf{A}\mathbf{v} + \mathbf{B}w + \mathbf{B}_c u \tag{6.3}$$

where $\mathbf{v} = [\mathbf{x}^T, \dot{\mathbf{x}}^T, \mathbf{y}_f^T]^T$; w is the ground motion; and u is the signal given to the actuator and it will be discussed in more detail later. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{B}_c$ are given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{5\times5} & \mathbf{I}_5 & \mathbf{0}_{5\times3} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} & \mathbf{M}^{-1}\mathbf{A}_{12} \\ \mathbf{A}_{21} & & \mathbf{A}_{22} \end{bmatrix} \tag{6.4}$$

$$\mathbf{B} = -[0,0,0,0,0,1,1,1,1,1,0,0,0]^T$$

$$\mathbf{B}_c = [\mathbf{0}_{1\times10}, B_{fu}, 0, 0]^T$$

where $\mathbf{A}_{12}$ and $\mathbf{A}_{21}$ and $\mathbf{A}_{22}$ are given by

$$\mathbf{A}_{12} = \begin{bmatrix} & \mathbf{0}_{3\times3} & \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \mathbf{A}_{21} = \begin{bmatrix} \mathbf{0}_{1\times8} & -B_f & B_f \\ & \mathbf{0}_{2\times10} & \end{bmatrix}, \mathbf{A}_{22} = \begin{bmatrix} A_f & 0 & 0 \\ 0 & 0 & 1 \\ \omega_c^2 & -\omega_c^2 & -\sqrt{2}\omega_c \end{bmatrix} \tag{6.5}$$

where $\omega_c = 10.0$Hz is used.

Using an analogy of Eqn. 5.6, Eqn. 6.3 can be transformed to the following discrete-time augmented state equation:

$$\mathbf{v}[k+1] = \bar{\mathbf{A}}\mathbf{v}[k] + \bar{\mathbf{B}}w[k] + \bar{\mathbf{B}}_c u[k] \tag{6.6}$$

where $\bar{\mathbf{A}} \equiv e^{\mathbf{A}\Delta t}$, $\bar{\mathbf{B}} \equiv \mathbf{A}^{-1}(\bar{\mathbf{A}} - \mathbf{I}_{13})\mathbf{B}$ and $\bar{\mathbf{B}}_c \equiv \mathbf{A}^{-1}(\bar{\mathbf{A}} - \mathbf{I}_{13})\mathbf{B}_c$.

Absolute accelerations measurements are available at the 2^{nd} and 4^{th} DOF, which is given by

$$\mathbf{z}[k] = -\mathbf{L}_o\mathbf{M}^{-1}\mathbf{K}\mathbf{x}[k] - \mathbf{L}_o\mathbf{M}^{-1}\mathbf{C}\dot{\mathbf{x}}[k] + \mathbf{n}[k] \tag{6.7}$$

where $\mathbf{L}_o$ is an observation matrix which is given by

$$\mathbf{L}_o = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{6.8}$$

and $\mathbf{n}[k]$ is a discrete white noise, with zero mean and standard deviations 0.001g, which models the prediction error, i.e., measurement noise, the differentiator errors and modeling error.

The signal given to the AMD actuator is given by

$$u[k] = \mathbf{G}_0\mathbf{z}[k] + \mathbf{G}_1\mathbf{z}[k-1] \tag{6.9}$$

so the controller feeds back only the current and the previous time step. Here $\mathbf{G}_p \in$

$\mathbb{R}^{1\times 2}, p = 0, 1$ are design parameters. Substituting Eqn. 6.7 and 6.9 into Eqn. 6.6, one can obtain the following augmented vector equation:

$$\begin{bmatrix} \mathbf{v}[k+1] \\ \mathbf{z}[k] \end{bmatrix} = \bar{\mathbf{A}}_u \begin{bmatrix} \mathbf{v}[k] \\ \mathbf{z}[k-1] \end{bmatrix} + \bar{\mathbf{B}}_u \bar{\mathbf{f}}[k] \tag{6.10}$$

where $\bar{\mathbf{f}} = [w[k], \mathbf{n}[k]^T]^T$; and $\bar{\mathbf{A}}_u$ and $\bar{\mathbf{B}}_u$ are given by

$$\begin{aligned} \bar{\mathbf{A}}_u &= \begin{bmatrix} \bar{\mathbf{A}} - \bar{\mathbf{B}}_c \mathbf{G}_0 \mathbf{L}_o [\mathbf{M}^{-1}\mathbf{K} & \mathbf{M}^{-1}\mathbf{C} & \mathbf{0}_{13\times 3}] & \bar{\mathbf{B}}_c \mathbf{G}_1 \\ -\mathbf{L}_o \mathbf{M}^{-1}\mathbf{K} & -\mathbf{L}_o \mathbf{M}^{-1}\mathbf{C} & \mathbf{0}_{2\times 3} & \mathbf{0}_{2\times 2} \end{bmatrix} \\ \bar{\mathbf{B}}_u &= \begin{bmatrix} \bar{\mathbf{B}} & \bar{\mathbf{B}}_c \mathbf{G}_0 \\ \mathbf{0}_{2\times 1} & \mathbf{I}_2 \end{bmatrix} \end{aligned} \tag{6.11}$$

Then, the covariance matrix $\mathbf{\Sigma}_u$ of the augmented vector is the solution of the following Lyapunov's equation in discrete form

$$\mathbf{\Sigma}_u = \bar{\mathbf{A}}_u \mathbf{\Sigma}_u \bar{\mathbf{A}}_u^T + \bar{\mathbf{B}}_u \mathbf{\Sigma}_{\bar{f}} \bar{\mathbf{B}}_u^T \tag{6.12}$$

where $\mathbf{\Sigma}_{\bar{f}}$ is the covariance matrix of $\bar{\mathbf{f}}$.

The threshold levels for the performance reliability of the interstory drifts, actuator stroke and the actuator acceleration are 1.0cm, 1.0m and 10g, respectively. Two controllers are designed. A pre-test controller is found by using the following pre-test prior distribution of the structure: Gaussian distribution, truncated for positive values only, with mean 1.1 and 1.0% and standard deviation 0.2 and 0.5% for the stiffness parameters and damping ratios, respectively. Also, a post-test controller is found by using the updated PDF of the parameters obtained in Section 6.2.

Tables 6.3 and 6.4 show the performance quantities of standard LQG, with the mass matrix and the identified stiffness matrix as the weighting matrices, and the two aforementioned controllers, including the interstory drifts, AMD actuator stroke (x_s) and control force (f_{cn}), for the cases of a random excitation sample and twice the 1940 El Centro earthquake record, respectively. In these tables, σ denotes the rms value of

Performance quantity	Threshold	Uncontrolled	LQG	Pre-test (10g)	Post-test (10g)
σ_{x_1} (m)	—	0.0072	0.0024	0.0023	0.0019
$\sigma_{x_1-x_2}$ (m)	—	0.0074	0.0024	0.0023	0.0019
$\sigma_{x_2-x_3}$ (m)	—	0.0062	0.0022	0.0020	0.0017
$\sigma_{x_3-x_4}$ (m)	—	0.0037	0.0016	0.0015	0.0013
$max\|x_1\|$ (m)	0.01	0.0222	0.0107	0.0085	0.0072
$max\|x_1 - x_2\|$ (m)	0.01	0.0217	0.0094	0.0082	0.0070
$max\|x_2 - x_3\|$ (m)	0.01	0.0212	0.0084	0.0079	0.0068
$max\|x_3 - x_4\|$ (m)	0.01	0.0133	0.0067	0.0066	0.0057
σ_{x_s} (m)	—	—	0.1533	0.0482	0.0501
$\sigma_{f_{cn}}$ (g)	—	—	6.7633	0.7863	1.3642
$max\|x_s\|$ (m)	1.0	—	0.6545	0.1656	0.1739
$max\|f_{cn}\|$ (g)	10.0	—	28.059	3.1912	5.5509

Table 6.3: Statistical properties of the performance quantities under random excitation (Example 6-1).

a quantity. It can be seen that the interstory drifts are significantly reduced in both cases when the AMD is installed. The LQG controller gives comparable structural performance to the robust reliability controllers but it requires much larger control forces. Furthermore, the post-test controller gives better performance than the pre-test controller because it incorporates the updated PDF of the structural parameters. Fig. 6.6 and 6.7 show the interstory drifts for the uncontrolled and controlled structure using the post-test controller under twice the El Centro earthquake. Furthermore, the corresponding stroke and normalized control force are shown in Fig. 6.8 and 6.9, respectively.

A parametric study is performed to investigate the effect of the control force limit. Post-test controllers are designed using the following threshold levels for the control force: 2g, 5g, 20g and also for unlimited control force (but retaining the constraint of 1.0m for the actuator stroke). The controller gains for these controllers as well as the aforementioned ones are shown in Tables 6.5. Furthermore, the performance quantities for these control systems are shown in Tables 6.6 under random excitation. It is intuitive that a controller with a higher level of control force performs better but the improvement saturates when this level is large enough, so there is a kind of law of diminshing return as larger actuators are provided.

Performance quantity	Threshold	Uncontrolled	LQG	Pre-test (10g)	Post-test (10g)
σ_{x_1} (m)	------	0.0030	0.0011	0.0010	0.0009
$\sigma_{x_1-x_2}$ (m)	------	0.0030	0.0011	0.0010	0.0009
$\sigma_{x_2-x_3}$ (m)	------	0.0025	0.0009	0.0008	0.0008
$\sigma_{x_3-x_4}$ (m)	------	0.0014	0.0006	0.0005	0.0005
$max\|x_1\|$ (m)	0.01	0.0135	0.0069	0.0069	0.0068
$max\|x_1 - x_2\|$ (m)	0.01	0.0145	0.0070	0.0069	0.0069
$max\|x_2 - x_3\|$ (m)	0.01	0.0127	0.0055	0.0058	0.0055
$max\|x_3 - x_4\|$ (m)	0.01	0.0078	0.0045	0.0037	0.0046
$\sigma_{\ddot{x}_s}$ (m)	------	------	0.0683	0.0203	0.0211
$\sigma_{f_{cn}}$ (g)	------	------	1.9145	0.4871	0.5494
$max\|\ddot{x}_s\|$ (m)	1.0	------	0.5099	0.1319	0.1388
$max\|f_{cn}\|$ (g)	10.0	------	15.661	3.8865	4.4003

Table 6.4: Statistical properties of the performance quantities under twice the 1940 El Centro earthquake (Example 6-1).

Gain	LQG	Pre-test(10g)	Post-test(2g)	Post-test(5g)	Post-test(10g)	Post-test(20g)	Post-test(no limit)
$G_0(1)$	-4.2332	-0.6823	-0.1998	-0.9030	-0.6478	-0.0102	-3.31
$G_0(2)$	-17.251	-2.2305	-0.1710	-0.8906	-2.3554	-6.2100	-38.53
$G_1(1)$	------	-0.0708	0.0224	0.0157	-0.0391	0.0489	-14.19
$G_1(2)$	------	-1.1744	0.3995	0.8092	1.1090	1.8500	15.65

Table 6.5: Gain coefficients of the optimal controllers (Example 6-1).

Performance quantity	Threshold	Post-test (2g)	Post-test (5g)	Post-test (20g)	Post-test (no limit)
σ_{x_1} (m)	------	0.0025	0.0022	0.0017	0.0015
$\sigma_{x_1-x_2}$ (m)	------	0.0025	0.0022	0.0017	0.0016
$\sigma_{x_2-x_3}$ (m)	------	0.0022	0.0019	0.0015	0.0013
$\sigma_{x_3-x_4}$ (m)	------	0.0015	0.0014	0.0011	0.0010
$max\|x_1\|$ (m)	0.01	0.0095	0.0082	0.0068	0.0068
$max\|x_1 - x_2\|$ (m)	0.01	0.0086	0.0075	0.0067	0.0066
$max\|x_2 - x_3\|$ (m)	0.01	0.0082	0.0077	0.0065	0.0064
$max\|x_3 - x_4\|$ (m)	0.01	0.0060	0.0056	0.0045	0.0040
$\sigma_{\ddot{x}_s}$ (m)	------	0.0273	0.0354	0.0587	0.1268
$\sigma_{f_{cn}}$ (g)	------	0.3304	0.7889	3.7039	13.317
$max\|\ddot{x}_s\|$ (m)	1.0	0.0863	0.1149	0.2209	0.5998
$max\|f_{cn}\|$ (g)	As labeled	1.1426	2.9848	17.109	57.548

Table 6.6: Statistical properties of the performance quantities under random excitation with different control force constraints (Example 6-1).

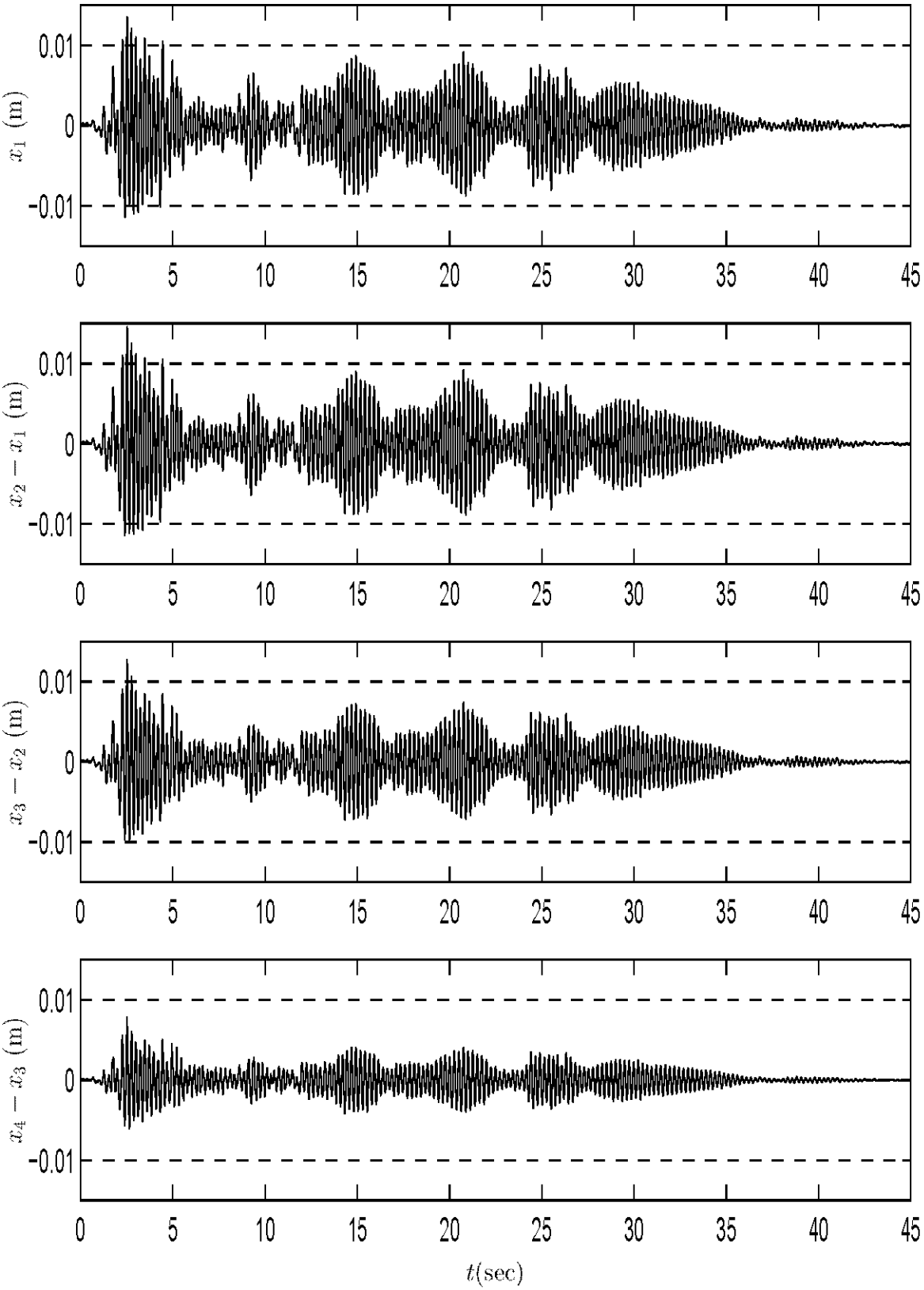

Figure 6.6: Interstory drift time histories of the uncontrolled structure under twice the 1940 El Centro earthquake record (Example 6-1).

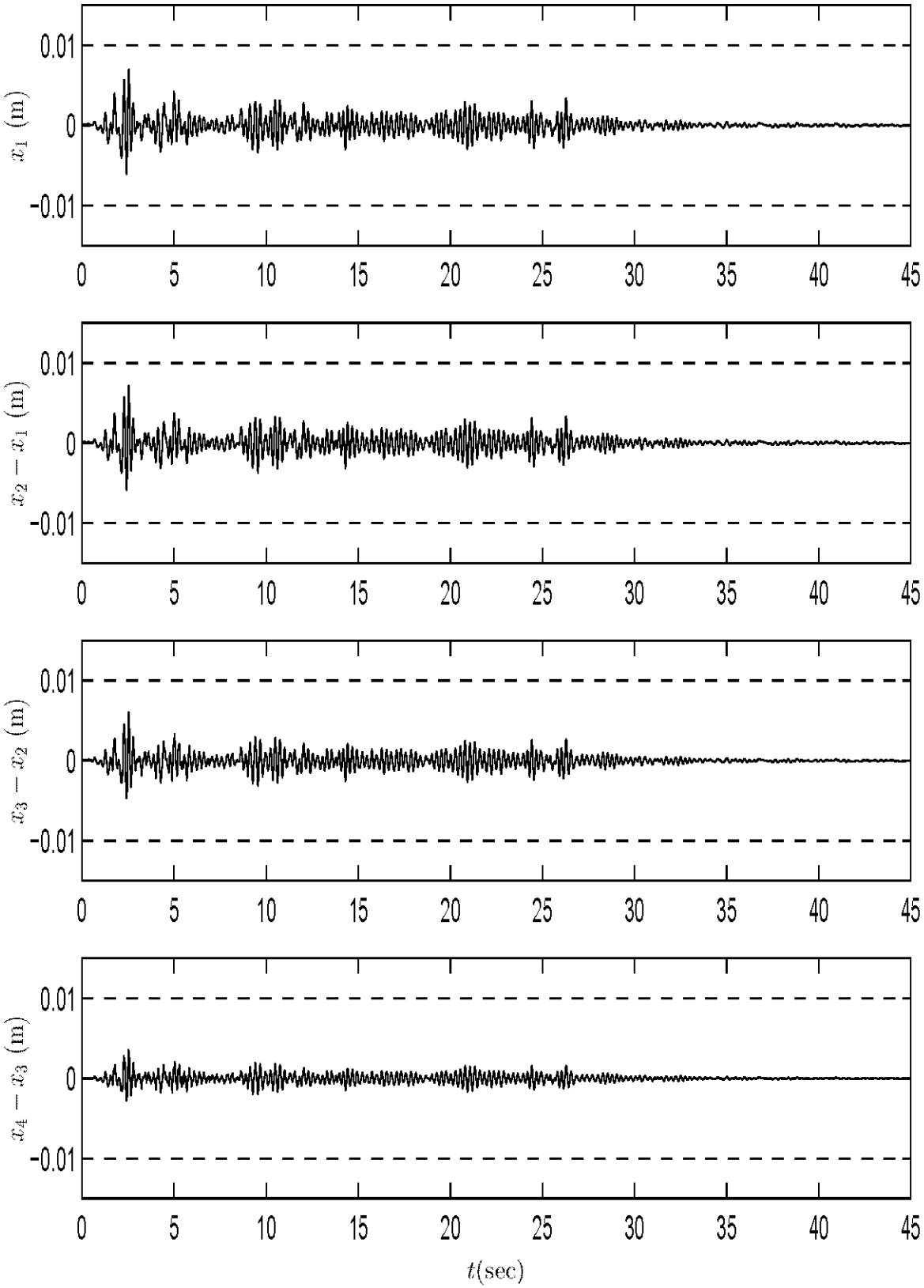

Figure 6.7: Interstory drift time histories of the controlled structure under twice the 1940 El Centro earthquake record (Example 6-1: post-test controller).

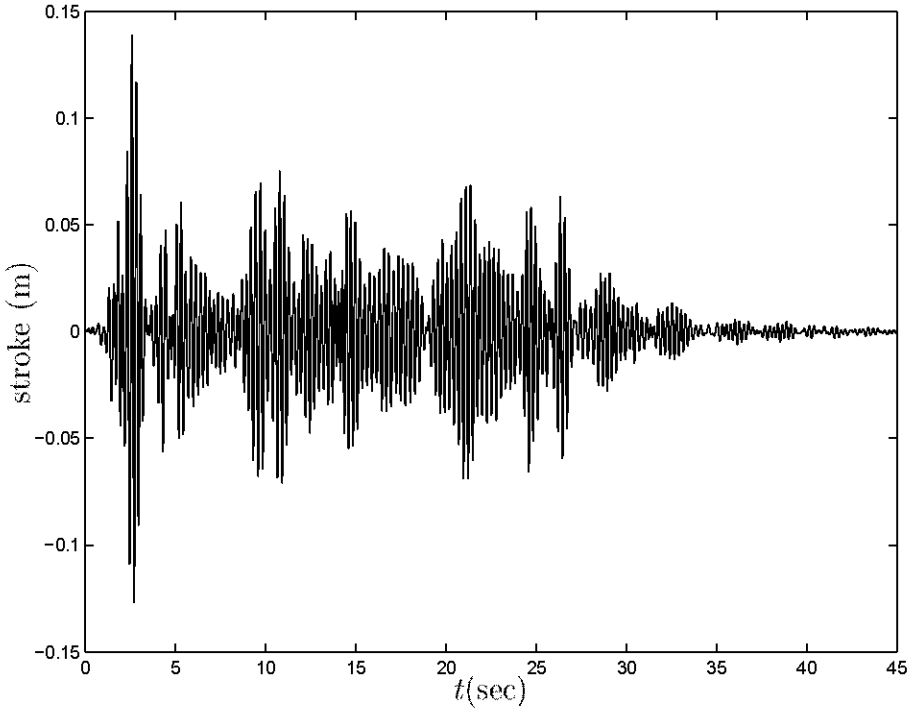

Figure 6.8: Controller stroke time histories under twice the 1940 El Centro earthquake record (Example 6-1: post-test controller).

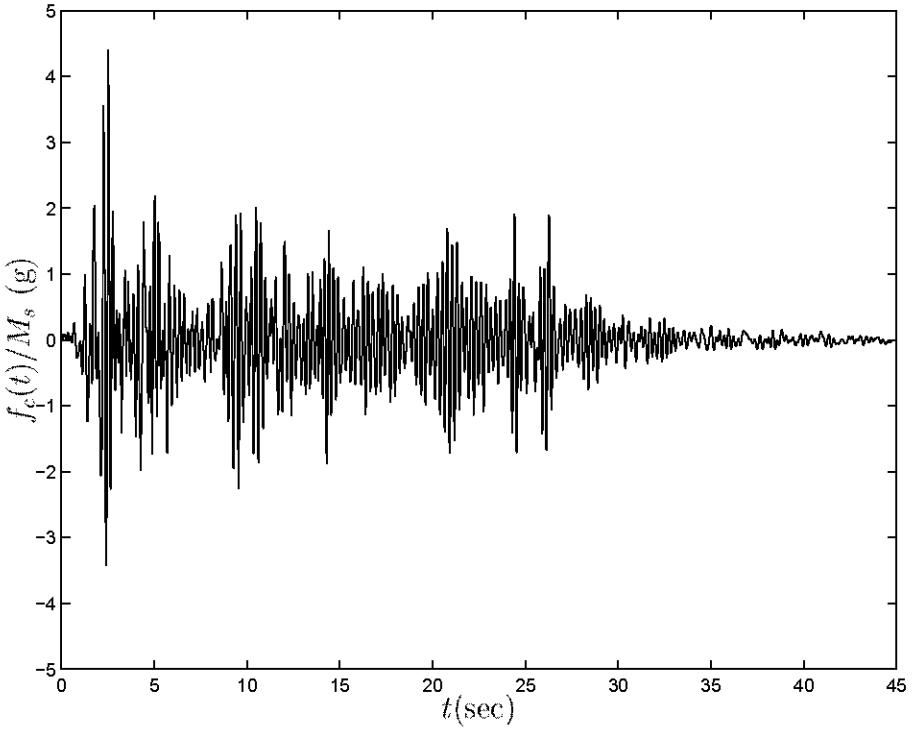

Figure 6.9: Controller force (normalized by the actuator mass) time histories under twice the 1940 El Centro earthquake record (Example 6-1: post-test controller).

Chapter 7 Conclusion and Future Work

7.1 Conclusion

Chapter 1 introduces the general idea of identification and robust control for smart structures and the organization of this thesis.

Chapter 2 introduces a probabilistic approach for choosing the most plausible class of models representing a physical structure based on its response measurements. The most plausible class of models has to compromise between its accuracy and complexity. For example, a more complex model might be able to capture the dynamics of the system more precisely but it introduces more parameters, which might over-fit the data. Therefore, the most plausible class of models should have enough, but not redundant, complexity in order to optimize this tradeoff. Examples using linear and nonlinear systems are used for demonstration.

Chapter 3 describes a Bayesian time-domain approach for modal updating using nonstationary incomplete noisy measurements. This time-domain approach is based on an approximate expansion of the updated probability density function. The proposed approach allows for direct calculation of the associated uncertainty of the identified modal parameters. Numerical examples verify the accuracy of the identified modal parameters and their associated uncertainty by simulations. The importance of treating nonstationary response is also addressed.

Chapter 4 introduces a Bayesian spectral density approach for updating nonlinear systems using incomplete noisy measurements. This frequency-domain approach is based on the statistical properties, shown in Appendix A, of the spectral density estimator. The proposed probabilistic framework is very well-suited for solving such a nonunique problem. Again, the proposed approach allows for direct calculation of the associated uncertainty of the identified model parameters using response measurements only. Numerical examples verify the accuracy of the identified model

parameters and their associated uncertainty using simulated response time histories.

Chapter 5 introduces a stochastic robust control method. The proposed method provides exact treatment for the delay time (including the buffer time for data collection, computation, and signal delivery and the internal dynamics of the actuator), so the controller can avoid the instability problems induced by time delays. Numerical examples are used to provide some insights into the proposed method. The proposed approach is also applied to a control benchmark problem with satisfactory results.

Chapter 6 uses a 20-DOF building to demonstrate the procedures for identification and robust control for smart structures. First, the model selection and identification approaches presented in Chapters 2 and 3 are used for optimally selecting the model class to be used for the controller design. Two model class candidates are assumed, which are a class of four-story structural frames and a class of four-story shear buildings. It turns out that the class of four-story shear buildings is more plausible based on the data. Furthermore, the optimal parameters and the updated PDF for the parameters are obtained. By using this information, the robust control approach presented in Chapter 5 can be performed to obtain the optimal controller. The control system (the 20-DOF building with the actuator) is tested under random excitation and twice the 1940 El Centro earthquake record. The structural response was significantly reduced, and the actuator stroke and control force did not exceed their threshold levels.

7.2 Future Work

This thesis introduces a complete framework for the identification and robust control for smart structures. However, it is desirable to extend the framework to allow for a *real-time adoptable* controller in order to capture changes in structural behavior. This is very important since the structural properties might be changing during large earthquakes or strong wind excitation.

The proposed robust control methodology was demonstrated using examples with an active mass driver. Further work can be done using semi-active devices (Kobori

et al. 1993; Iwan and Wang 1996; Dyke et al. 1996; Spencer and Sain 1997; Johnson et al. 1999; Zhang and Iwan 2002), which have very low power requirements and have been proved to be very efficient. Also, the proposed methodology can be extended to nonlinear dynamical systems. In this case, the failure probability might be obtained by simulations (Au and Beck 2001).

Cost-benefit analysis (Irfanoglu and Beck 2001) can be used with the proposed control framework for selecting the most cost-effective strategy of response reduction (e.g., choosing from passive control device or active mass driver).

Finally, the proposed system identification techniques based on output-only data can be applied for damage detection. They can also applied to past seismic response, especially in the case where the base motion time histories are unavailable to serve as input, as occurred in some damaged steel-frame buildings in the 1994 Northridge Earthquake in Los Angeles (Carlson 1999).

Appendix A

Asymptotic Independence of the Spectral Density Estimator

In this appendix it is shown that the spectral density estimators are independent at any two different frequencies ω and ω' as $N \to \infty$.

First, define the following scaled real and imaginary parts of the spectral density estimator

$$\begin{aligned} \mathcal{Y}_{N,R}^{(\alpha)}(\omega_k) &= \mathrm{Re}[\mathcal{Y}_N^{(\alpha)}(\omega_k)] \\ \mathcal{Y}_{N,I}^{(\alpha)}(\omega_k) &= \mathrm{Im}[\mathcal{Y}_N^{(\alpha)}(\omega_k)] \end{aligned} \tag{A.1}$$

where $\boldsymbol{\mathcal{Y}}_N(\omega_k)$ is given by Eqn. 4.20 and $f^{(\alpha)}$ denotes the j^{th} element of a vector $\mathbf{f}$. Note that $E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega_k)] = 0$ and $E[\mathcal{I}_N^{(\alpha)}(\omega_k)] = 0$.

A.1 $\mathcal{Y}_{N,R}^{(\alpha)}(\omega)$ with $\mathcal{Y}_{N,R}^{(\beta)}(\omega')$

Taking mathematical expectation on both sides of Eqn. A.1, one can easily obtain the following:

$$\begin{aligned} & E[(\mathcal{Y}_{N,R}^{(\alpha)}(\omega) - E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,R}^{(\beta)}(\omega') - E[\mathcal{Y}_{N,R}^{(\beta)}(\omega')])] \\ &= E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)\mathcal{Y}_{N,R}^{(\beta)}(\omega')] - E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)]E[\mathcal{Y}_{N,R}^{(\beta)}(\omega')] \\ &= \frac{\Delta t}{2\pi N} \sum_{j=0}^{N-1} \sum_{l=0}^{N-1} (E[y^{(\alpha)}(j\Delta t) y^{(\beta)}(l\Delta t)] - \mu^{(\alpha)}\mu^{(\beta)}) \cos(j\omega\Delta t) \cos(l\omega'\Delta t) \end{aligned} \tag{A.2}$$

Let $S_y^{(\alpha,\beta)}(\Omega)$, which is assumed to be finite $\forall \Omega \in \mathbb{R}$, be the cross spectral density between $y^{(\alpha)}$ and $y^{(\beta)}$ at frequency Ω. By using the fact that the cross covariance

function $\phi^{(\alpha,\beta)}(\tau) = \int_{-\infty}^{\infty} S_y^{(\alpha,\beta)}(\Omega) e^{i\Omega\tau} d\Omega$ and $\cos(z) = \frac{e^{iz}+e^{-iz}}{2}$,

$$
\begin{aligned}
& E[(\mathcal{Y}_{N,R}^{(\alpha)}(\omega) - E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,R}^{(\beta)}(\omega') - E[\mathcal{Y}_{N,R}^{(\beta)}(\omega')])] \\
&= \int_{-\infty}^{\infty} \frac{\Delta t}{8\pi N} \sum_{j=0}^{N-1} \sum_{l=0}^{N-1} e^{i\Omega(j-l)\Delta t} S_y^{(\alpha,\beta)}(\Omega)[e^{ij\omega\Delta t} + e^{-ij\omega\Delta t}][e^{il\omega'\Delta t} + e^{-il\omega'\Delta t}] d\Omega \\
&= \int_{-\infty}^{\infty} \frac{\Delta t}{8\pi N} \sum_{j=0}^{N-1} \sum_{l=0}^{N-1} [e^{ij(\Omega+\omega)\Delta t + il(\omega'-\Omega)\Delta t} + e^{ij(\Omega+\omega)\Delta t + il(-\omega'-\Omega)\Delta t} + e^{ij(\Omega-\omega)\Delta t + il(\omega'-\Omega)\Delta t} \\
&\quad + e^{ij(\Omega-\omega)\Delta t + il(-\omega'-\Omega)\Delta t}] S_y^{(\alpha,\beta)}(\Omega) d\Omega \\
&= \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') d\Omega + \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega') d\Omega + \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega') d\Omega \\
&\quad + \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega') d\Omega
\end{aligned}
\tag{A.3}
$$

where $H_N(\Omega;\omega,\omega')$ is defined as

$$
H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') \equiv \frac{S_y^{(\alpha,\beta)}(\Omega)\Delta t}{8\pi N} \sum_{j=0}^{N-1} e^{ij(\Omega+\omega)\Delta t} \sum_{l=0}^{N-1} e^{il(\omega'-\Omega)\Delta t} \tag{A.4}
$$

If $|\Omega| \neq \omega$ and $|\Omega| \neq \omega'$,

$$
H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') = \frac{[1 - e^{iN\Delta t(\Omega+\omega)}][1 - e^{iN\Delta t(\omega'-\Omega)}]}{8\pi N[1 - e^{i\Delta t(\Omega+\omega)}][1 - e^{i\Delta t(\omega'-\Omega)}]} S_y^{(\alpha,\beta)}(\Omega)\Delta t \tag{A.5}
$$

By using $\sin(z) = \frac{e^{iz}-e^{-iz}}{2i}$,

$$
H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') = \frac{e^{i(N-1)\Delta t)(\omega+\omega')/2} \sin[\frac{(\Omega+\omega)N\Delta t}{2}] \sin[\frac{(\omega'-\Omega)N\Delta t}{2}]}{8\pi N \sin[\frac{(\Omega+\omega)\Delta t}{2}] \sin[\frac{(\omega'-\Omega)\Delta t}{2}]} S_y^{(\alpha,\beta)}(\Omega)\Delta t \tag{A.6}
$$

Then, the following inequality can be obtained since $|e^{ir}| = 1$ and $|\sin(r)| \leq 1$ ($r \in \mathbb{R}$):

$$
\left| H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') \right| \leq \left| \frac{S_y^{(\alpha,\beta)}(\Omega)\Delta t}{8\pi N \sin[\frac{(\Omega+\omega)\Delta t}{2}] \sin[\frac{(\omega'-\Omega)\Delta t}{2}]} \right| \tag{A.7}
$$

By taking the limit as $N \to \infty$,

$$\lim_{N\to\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') = 0 \qquad \text{if} \quad |\Omega| \neq \omega \quad \text{and} \quad |\Omega| \neq \omega' \tag{A.8}$$

Similarly, $H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')$, $H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')$ and $H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')$ also tend to zero as $N \to \infty$, if $|\Omega| \neq \omega$ and $|\Omega| \neq \omega'$.

Now, we consider $H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')$ at $|\Omega| = \omega$ or $|\Omega| = \omega'$. First, at $\Omega = -\omega$,

$$\begin{aligned} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') &= \frac{S_y^{(\alpha,\beta)}(\Omega)\Delta t}{8\pi} \sum_{l=0}^{N-1} e^{il(\omega'-\Omega)\Delta t} \\ &= \frac{S_{y,N}^{(\alpha,\beta)}(\Omega)\Delta t[1 - e^{iN\Delta t(\omega'-\Omega)}]}{8\pi[1 - e^{i\Delta t(\omega'-\Omega)}]} \end{aligned} \tag{A.9}$$

which is finite. Similarly, it can be shown that $H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')$ is finite at $\Omega = \omega'$. Next, consider $\Omega = \omega$ or $\Omega = -\omega'$,

$$H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') = \frac{[1 - e^{iN\Delta t(\Omega+\omega)}][1 - e^{iN\Delta t(\omega'-\Omega)}]}{8\pi N[1 - e^{i\Delta t(\Omega+\omega)}][1 - e^{i\Delta t(\omega'-\Omega)}]} S_y^{(\alpha,\beta)}(\Omega)\Delta t \tag{A.10}$$

As $N \to \infty$ and $\Omega = \omega$ or $-\omega'$,

$$\lim_{N\to\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega') = 0 \tag{A.11}$$

Therefore, $H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')$ is finite at $\Omega = \pm\omega$ and $\Omega = \pm\omega'$. Similarly, it can be shown that $H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')$, $H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')$ and $H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')$ are finite at $\Omega = \pm\omega$ and $\Omega = \pm\omega'$.

Therefore, $H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')$, $H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')$, $H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')$ and $H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')$ tend to zero as $N \to \infty$ if $|\Omega| \neq \omega$ and $|\Omega| \neq \omega'$ and they are finite as $N \to \infty$ if $|\Omega| = \omega$ or $|\Omega| = \omega'$. It can be concluded that $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')d\Omega$, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')d\Omega$, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')d\Omega$ and $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')d\Omega$ tend to zero as $N \to \infty$ and

so from Eqn. A.3:

$$\lim_{N\to\infty} E[(\mathcal{Y}_{N,R}^{(\alpha)}(\omega) - E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,R}^{(\beta)}(\omega') - E[\mathcal{Y}_{N,R}^{(\beta)}(\omega')])] = 0, \qquad \text{if} \quad \omega \neq \omega' \tag{A.12}$$

A.2 $\mathcal{Y}_{N,I}^{(\alpha)}(\omega)$ with $\mathcal{Y}_{N,I}^{(\beta)}(\omega')$

Similarly, it can be proved that $\lim_{N\to\infty} E[(\mathcal{Y}_{N,I}^{(\alpha)}(\omega) - E[\mathcal{Y}_{N,I}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,I}^{(\beta)}(\omega') - E[\mathcal{Y}_{N,I}^{(\beta)}(\omega')])] = 0$ as follows:

$$\begin{aligned}
&E[(\mathcal{Y}_{N,I}^{(\alpha)}(\omega) - E[\mathcal{Y}_{N,I}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,I}^{(\beta)}(\omega') - E[\mathcal{Y}_{N,I}^{(\beta)}(\omega')])] \\
&= \frac{\Delta t}{2\pi N} \sum_{j=0}^{N-1}\sum_{l=0}^{N-1} (E[y^{(\alpha)}(j\Delta t)y^{(\beta)}(l\Delta t)] - \mu^{(\alpha)}\mu^{(\beta)}) \sin(j\omega\Delta t)\sin(l\omega'\Delta t) \\
&= -\frac{\Delta t}{8\pi N} \int_{-\infty}^{\infty} \sum_{j=0}^{N-1}\sum_{l=0}^{N-1} e^{i\Omega(j-l)\Delta t} S_y^{(\alpha,\beta)}(\Omega)[e^{ij\omega\Delta t} - e^{-ij\omega\Delta t}][e^{il\omega'\Delta t} - e^{-il\omega'\Delta t}]d\Omega \\
&= -\int_{-\infty}^{\infty} \frac{\Delta t}{8\pi N} \sum_{j=0}^{N-1}\sum_{l=0}^{N-1} [e^{ij(\Omega+\omega)\Delta t + il(\omega'-\Omega)\Delta t} - e^{ij(\Omega+\omega)\Delta t + il(-\omega'-\Omega)\Delta t} - e^{ij(\Omega-\omega)\Delta t + il(\omega'-\Omega)\Delta t} \\
&\quad + e^{ij(\Omega-\omega)\Delta t + il(-\omega'-\Omega)\Delta t}] S_y^{(\alpha,\beta)}(\Omega)d\Omega \\
&= -\int_{-\infty}^{\infty} \frac{\Delta t}{8\pi N} \left[\sum_{j=0}^{N-1} e^{ij(\Omega+\omega)\Delta t} \sum_{l=0}^{N-1} e^{il(\omega'-\Omega)\Delta t} - \sum_{j=0}^{N-1} e^{ij(\Omega+\omega)\Delta t} \sum_{l=0}^{N-1} e^{il(-\omega'-\Omega)\Delta t} \right. \\
&\quad \left. - \sum_{j=0}^{N-1} e^{ij(\Omega-\omega)\Delta t} \sum_{l=0}^{N-1} e^{il(\omega'-\Omega)\Delta t} + \sum_{j=0}^{N-1} e^{ij(\Omega-\omega)\Delta t} \sum_{l=0}^{N-1} e^{il(-\omega'-\Omega)\Delta t} \right] S_y^{(\alpha,\beta)}(\Omega)d\Omega \\
&= -\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')d\Omega + \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')d\Omega + \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')d\Omega \\
&\quad - \int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')d\Omega
\end{aligned} \tag{A.13}$$

where $H_N^{(\alpha,\beta)}$ is given by Eqn. A.4.

As shown in Section A.1, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')d\Omega$, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')d\Omega$,

$\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')d\Omega$ and $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')d\Omega$ tend to zero as $N\to\infty$, so

$$\lim_{N\to\infty} E[(\mathcal{Y}_{N,I}^{(\alpha)}(\omega)-E[\mathcal{Y}_{N,I}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,I}^{(\beta)}(\omega')-E[\mathcal{Y}_{N,I}^{(\beta)}(\omega')])]=0, \qquad \text{if} \quad \omega\neq\omega' \quad \text{(A.14)}$$

A.3 $\mathcal{Y}_{N,R}^{(\alpha)}(\omega)$ with $\mathcal{Y}_{N,I}^{(\beta)}(\omega')$

Finally, we prove $\lim_{N\to\infty} E[(\mathcal{Y}_{N,R}^{(\alpha)}(\omega)-E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,I}^{(\beta)}(\omega')-E[\mathcal{Y}_{N,I}^{(\beta)}(\omega')])]=0.$

$$\begin{aligned}
&E[(\mathcal{Y}_{N,R}^{(\alpha)}(\omega)-E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,I}^{(\beta)}(\omega')-E[\mathcal{Y}_{N,I}^{(\beta)}(\omega')])]\\
&=\frac{\Delta t}{2\pi N}\sum_{j=0}^{N-1}\sum_{l=0}^{N-1}(E[y^{(\alpha)}(j\Delta t)y^{(\beta)}(l\Delta t)]-\mu^{(\alpha)}\mu^{(\beta)})\cos(j\omega\Delta t)\sin(l\omega'\Delta t)\\
&=\int_{-\infty}^{\infty}\frac{\Delta t}{8\pi N i}\sum_{j=0}^{N-1}\sum_{l=0}^{N-1}e^{i\Omega(j-l)\Delta t}S_y^{(\alpha,\beta)}(\Omega)[e^{ij\omega\Delta t}+e^{-ij\omega\Delta t}][e^{il\omega'\Delta t}-e^{-il\omega'\Delta t}]d\Omega\\
&=\int_{-\infty}^{\infty}\frac{\Delta t}{8\pi N i}\sum_{j=0}^{N-1}\sum_{l=0}^{N-1}[e^{ij(\Omega+\omega)\Delta t+il(\omega'-\Omega)\Delta t}-e^{ij(\Omega+\omega)\Delta t+il(-\omega'-\Omega)\Delta t}+e^{ij(\Omega-\omega)\Delta t+il(\omega'-\Omega)\Delta t}\\
&\quad-e^{ij(\Omega-\omega)\Delta t+il(-\omega'-\Omega)\Delta t}]S_y^{(\alpha,\beta)}(\Omega)d\Omega\\
&=\int_{-\infty}^{\infty}\frac{\Delta t}{8\pi N i}\Bigg[\sum_{j=0}^{N-1}e^{ij(\Omega+\omega)\Delta t}\sum_{l=0}^{N-1}e^{il(\omega'-\Omega)\Delta t}-\sum_{j=0}^{N-1}e^{ij(\Omega+\omega)\Delta t}\sum_{l=0}^{N-1}e^{il(-\omega'-\Omega)\Delta t}\\
&\quad+\sum_{j=0}^{N-1}e^{ij(\Omega-\omega)\Delta t}\sum_{l=0}^{N-1}e^{il(\omega'-\Omega)\Delta t}-\sum_{j=0}^{N-1}e^{ij(\Omega-\omega)\Delta t}\sum_{l=0}^{N-1}e^{il(-\omega'-\Omega)\Delta t}\Bigg]S_y^{(\alpha,\beta)}(\Omega)d\Omega\\
&=-i\int_{-\infty}^{\infty}H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')d\Omega+i\int_{-\infty}^{\infty}H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')d\Omega-i\int_{-\infty}^{\infty}H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')d\Omega\\
&\quad+i\int_{-\infty}^{\infty}H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')d\Omega
\end{aligned} \qquad \text{(A.15)}$$

where $H_N^{(\alpha,\beta)}$ is given by Eqn. A.4.

Again, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,\omega')d\Omega$, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;\omega,-\omega')d\Omega$, $\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,\omega')d\Omega$ and
$\int_{-\infty}^{\infty} H_N^{(\alpha,\beta)}(\Omega;-\omega,-\omega')d\Omega$ tend to zero as $N\to\infty$, so

$$\lim_{N\to\infty} E[(\mathcal{Y}_{N,R}^{(\alpha)}(\omega)-E[\mathcal{Y}_{N,R}^{(\alpha)}(\omega)])(\mathcal{Y}_{N,I}^{(\beta)}(\omega')-E[\mathcal{Y}_{N,I}^{(\beta)}(\omega')])]=0, \qquad \text{if} \quad \omega\neq\omega' \quad \text{(A.16)}$$

A.4 $S_{y,N}^{(\alpha,\beta)}(\omega)$ with $S_{y,N}^{(\gamma,\delta)}(\omega')$

We conclude that any element in the set $\{\mathcal{Y}_{N,R}^{(\alpha)}(\omega), \mathcal{Y}_{N,I}^{(\alpha)}(\omega), \mathcal{Y}_{N,R}^{(\beta)}(\omega), \mathcal{Y}_{N,I}^{(\beta)}(\omega)\}$ with any element in the set $\{\mathcal{Y}_{N,R}^{(\gamma)}(\omega'), \mathcal{Y}_{N,I}^{(\gamma)}(\omega'), \mathcal{Y}_{N,R}^{(\delta)}(\omega'), \mathcal{Y}_{N,I}^{(\delta)}(\omega')\}$ gives an uncorrelated pair, where $\alpha, \beta, \gamma, \delta = 1, 2, \ldots, N_s$ and $\omega \neq \omega'$. Furthermore, $\mathcal{Y}_{N,R}^{(\alpha)}(\Omega)$ and $\mathcal{Y}_{N,I}^{(\alpha)}(\Omega)$ are Gaussian distributed $\forall \Omega \in \mathbb{R}$ and $\alpha = 1, 2, \ldots, N_s$ as $N \to \infty$ even if the stochastic process $\mathbf{y}$ is not Gaussian. Since uncorrelated Gaussian random variables are independent, as $N \to \infty$ each element in the set $\{\mathcal{Y}_{N,R}^{(\alpha)}(\omega), \mathcal{Y}_{N,I}^{(\alpha)}(\omega), \mathcal{Y}_{N,R}^{(\beta)}(\omega), \mathcal{Y}_{N,I}^{(\beta)}(\omega)\}$ is statistically independent of each element in the set $\{\mathcal{Y}_{N,R}^{(\gamma)}(\omega'), \mathcal{Y}_{N,I}^{(\gamma)}(\omega'), \mathcal{Y}_{N,R}^{(\delta)}(\omega'), \mathcal{Y}_{N,I}^{(\delta)}(\omega')\}$, where $\alpha, \beta, \gamma, \delta = 1, 2, \ldots, N_s$ and $\omega \neq \omega'$. By using Eqn. 4.19, the following can be obtained:

$$\begin{aligned} S_{y,N}^{(\alpha,\beta)}(\omega) &= \mathcal{Y}_N^{(\alpha)}(\omega)\mathcal{Y}_N^{(\beta)^*}(\omega) \\ &= [\mathcal{Y}_{N,R}^{(\alpha)}(\omega)\mathcal{Y}_{N,R}^{(\beta)}(\omega) + \mathcal{Y}_{N,I}^{(\alpha)}(\omega)\mathcal{Y}_{N,I}^{(\beta)}(\omega)] + i[\mathcal{Y}_{N,I}^{(\alpha)}(\omega)\mathcal{Y}_{N,R}^{(\beta)}(\omega) - \mathcal{Y}_{N,R}^{(\alpha)}(\omega)\mathcal{Y}_{N,I}^{(\beta)}(\omega)] \end{aligned} \tag{A.17}$$

Therefore, $S_{y,N}^{(\alpha,\beta)}(\omega)$ and $S_{y,N}^{(\gamma,\delta)}(\omega')$ are statistically independent if $\alpha, \beta, \gamma, \delta = 1, 2, \ldots, N_s$, $\omega \neq \omega'$ and $0 < \omega, \omega' < \frac{\pi}{\Delta t}$, the Nyquist frequency.

Bibliography

Akaike, H. (1974). A new look at the statistical identification model. *IEEE Transactions on Automatic Control 19*, 716–723.

Akaike, H. (1976). On entropy maximization principle. *In Applications of Statistics, P.R. Krishnaioh (Ed.), North Holland, Amsterdam*, 27–41.

Andersen, P. and P. H. Kirkergaard (1998). Statistical damage detection of civil engineering structures using ARMAV models. In *Proceedings of 16th IMAC*, Santa Barbara, CA, pp. 356–362.

Asmussen, J. C., S. R. Ibrahim, and R. Brincker (1997). Application of vector triggering random decrement. In *Proceedings of 15th IMAC*, Volume 2, Orlando, Florida, pp. 1165–1171.

Au, S. K. and J. L. Beck (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics 16(4)*, 263–277.

Au, S. K., C. Papadimitriou, and J. L. Beck (1999). Reliability of uncertain dynamical systems with multiple design points. *Structural Safety 21*, 113–133.

Beck, J. L. (1978). *Determining Models of Structures from Earthquake Records.* Technical Report EERL 78-01, California Institute of Technology, Earthquake Engineering Research Laboratory, Pasadena, California.

Beck, J. L. (1996). System identification methods applied to measured seismic response. In *Proceedings of Eleventh World Conference on Earthquake Engineering.* Elsevier, New York.

Beck, J. L. and S. K. Au (2002). Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation. *Journal of Engineering Mechanics 128(4)*, 380–391.

Beck, J. L. and L. S. Katafygiotis (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechanics 124(4)*, 455–461.

Beck, J. L., B. S. May, and D. C. Polidori (1994). Determination of modal parameters from ambient vibration data for structural health monitoring. In *Proceedings of First World Conference on Structural Control*, Pasadena, California, pp. TA3:3–TA3:12.

Box, G. E. P. and G. M. Jenkins (1970). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

Box, G. E. P. and G. C. Tiao (1973). *Bayesian Inference in Statistical Analysis.* Reading, Mass., Addison-Wesley.

Brockwell, P. J. and R. Davis (1991). *Time Series: Theory and Methods.* New York: Springer-Verlag.

Carlson, A. E. (1999). *Three-dimensional Nonlinear Inelastic Analysis of Steel Moment-frame Buildings damaged by Earthquake Excitations* . Technical Report EERL 99-02, California Institute of Technology, Earthquake Engineering Research Laboratory, Pasadena, California.

Caughey, T. K. and M. E. J. O'Kelly (1965). Classical normal modes in damped linear dynamic systems. *Journal of Applied Mechanics 123(12)*, 583–588.

Caughey (Ed.), T. K. (1998). Special issue on benchmark problems. *Earthquake Engineering and Structural Dynamics 27(11).*

Cox, R. T. (1961). *The Algebra of Probable Inference.* Baltimore: Johns Hopkins Press.

Doyle, J. C., B. A. Francis, and A. R. Tannenbaum (1992). *Feedback Control Theory.* Macmillan Publishing Company.

Doyle, J. C., K. Glover, P. P. Khargonekar, and B. A. Francis (1989). State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. *IEEE Transactions on Automatic Control 34(8)*, 831–847.

Dyke, S. J., B. F. Spencer, P. Quast, D. C. Kaspari, and M. K. Sain (1996). Implementation of an active mass driver using acceleration feedback control. *Microcomputers in Civil Engineering 11*, 305–323.

Dyke, S. J., B. F. Spencer, P. Quast, and M. K. Sain (1995). Role of control-structure interaction in protective system design. *Journal of Engineering Mechanics 121(2)*, 322–338.

Dyke, S. J., B. F. Spencer, M. K. Sain, and J. D. Carlson (1996). A new semi-active control device for seismic response reduction. In *Proceedings of 11th ASCE Engineering Mechanics Special Conference, Ft. Lauderdale, Florida.*

Field, R. V., W. B. Hall, and L. A. Bergman (1994). A matlab-based approach to the computation of probabilistic stability measures for controlled systems. In *Proceedings of First World Conference on Structural Control, International Association for Structural Control*, Pasadena, pp. TP4–13–TP4–22.

Field, R. V., P. G. Voulgaris, and L. A. Bergman (1996). Probabilistic stability robustness of structural systems. *Journal of Engineering Mechanics 122(10)*, 1012–1021.

Gersch, W. and D. A. Foutch (1974). Least squares estimates of structural system parameters using covariance function data. *Institute of Electrical and Electronics Engineers Transactions on Automatic Control AC-19(6)*, 898–903.

Gersch, W., G. T. Taoka, and R. Liu (1976). Structural system parameter estimation by two-stage least squares method. *Journal of Engineering Mechanics 102(5)*, 883–899.

Ghanem, R. and S. Sture (Eds.) (2000). Special issue on structural health monitoring. *Journal of Engineering Mechanics 126(7).*

Gull, S. F. (1988). Bayesian inductive inference and maximum entropy. *Maximum Entropy and Bayesian Methods (Ed. J. Skilling), Kluwer Academic Publisher, Boston*, 53–74.

Hoshiya, M. and E. Saito (1984). Structural identification by extended Kalman filter. *Journal of Engineering Mechanics 110(12)*, 1757–1770.

Housner, G. W., L. A. Bergman, T. K. Caughey, A. G. Chassiakos, R. O. Claus, S. F. Masri, R. E. Skelton, T. T. Soong, B. F. Spencer, and J. T. P. Yao (1997). Special issue on structural control: past, present, and future. *Journal of Engineering Mechanics 123(9)*.

Irfanoglu, A. and J. L. Beck (2001). Optimal structural design under seismic risk using engineering and economic performance objectives. In *Proceedings of ICOSSAR'01*, Newport Beach, California, USA.

Ivers, D. E. and L. R. Miller (1991). *Semi-Active Suspension Technology: An evolutionary view.* Advanced Automotive Technologies, ASME Book No. H00719.

Iwan, W. D. and L. D. Lutes (1968). Response of the bilinear hysteretic system to stationary random excitation. *Journal of the Acoustical Society of America 43*, 545–552.

Iwan, W. D. and L. J. Wang (1996). New developments in active interaction control. In *Proceedings of Second International Workshop on Structural Control, Hong Kong, China.*

Jaynes, E. T. (1983). *Papers on probability statistics and statistical physics (Ed. R. Rosenkrantz).* Reidel, Dordrecht.

Jeffreys, H. (1961). *Theory of Probability (3rd edition).* Oxford Clarendon Press.

Johnson, E. A., B. F. Spencer, and Y. Fujino (1999, March). Semi-active control of cable vibration. In *Proceedings JSME Dynamics and Design Conference*, Chiba, Japan, pp. 153–156.

Johnson, E. A., P. G. Voulgaris, and L. A. Bergman (1998). Multiobjective optimal structural control of the notre dame building model benchmark. *Earthquake Engineering and Structural Dynamics 27(11)*, 1165–1187.

Katafygiotis, L. S. and J. L. Beck (1998). Updating models and their uncertainties. II: Model identifiability. *Journal of Engineering Mechanics 124(4)*, 463–467.

Katafygiotis, L. S., C. Papadimitriou, and H. F. Lam (1998). A probabilistic approach to structural model updating. *Soil Dynamics & Earthquake Engineering 17(7-8)*, 495–507.

Kobori, T., M. Takahashi, T. Nasu, N. Niwa, and K. Ogasawara (1993). Seismic response controlled structure with active variable stiffness system. *Earthquake Engineering and Structural Dynamics 22(11)*, 92–941.

Krishnaiah, P. R. (1976). Some recent developments on complex multivariate distributions. *Journal of multivariate analysis 6*, 1–30.

Kullback, S. (1968). *Information Theory and Statistics.* Dover, Publications Inc., Mineola, N.Y.

Lin, Y. K. (1976). *Probabilistic Theory of Structural Dynamics.* Robert E. Krieger Publishing Company, Malabar, FL.

Loh, C.-H. and S.-T. Chung (1993). A three-stage identification approach for hysteretic systems. *Earthquake Engineering and Structural Dynamics 22(2)*, 129–150.

Loh, C.-H. and Y.-H. Tsaur (1988). Time domain estimation of structural parameters. *Engineering Structures 10(2)*, 95–105.

Lutes, L. D. and S. Sarkani (1997). *Stochastic Analysis of Structural and Mechanical Vibrations.* New Jersey: Prentice Hall.

Mackay, D. J. C. (1992). Bayesian interpolation. *Neural computation 4(3)*, 415–447.

Marrison, C. I. and R. F. Stengel (1995). Stochastic robustness synthesis applied to a benchmark problem. *International Journal of Robust Nonlinear Control 5*, 13–31.

May, B. S. and J. L. Beck (1998). Probabilistic control for the active mass driver benchmark structural model. *Earthquake Engineering and Structural Dynamics 27(11)*, 1331–1346.

Natke, H. G. and J. T. P. Yao (1988). *Proceedings of the Workshop on Structural Safety Evaluation Based on System Identification Approaches.* Wiesbaden:

Vieweg and Sons.

Paganini, F. (1996). *Sets and Constrains in the Analysis of Uncertain Systems.* Technical report, Ph.D. thesis, California Institute of Technology.

Papadimitriou, C., J. L. Beck, and L. S. Katafygiotis (1997a). Asymptotic expansions for reliability and moments of uncertain systems. *Journal of Engineering Mechanics 123(12)*, 1219–1229.

Papadimitriou, C., J. L. Beck, and L. S. Katafygiotis (2001). Updating robust reliability using structural test data. *Probabilistic Engineering Mechanics 16(2)*, 103–113.

Papadimitriou, C., L. S. Katafygiotis, and S. K. Au (1997b). Effects of structural uncertainties on TMD design: A reliability-based approach. *Journal of Structural Control 4(1)*, 65–88.

Peng, C. Y. and W. D. Iwan (1992). An identification methodology for a class of hysteretic structures. *Earthquake Engineering and Structural Dynamics 21(8)*, 695–712.

Pi, Y. L. and N. C. Mickleborough (1989). Modal identification of vibrating structures using ARMA models. *Journal of Engineering Mechanics 115(10)*, 2232–2250.

Quek, S. T., W. P. Wang, and C. G. Koh (1999). System identification of linear MDOF structures under ambient excitation. *Earthquake Engineering and Structural Dynamics 28*, 61–77.

Roberts, J. B., J. F. Dunne, and A. Debonos (1995). A spectral method for estimation for nonlinear system parameters from measured response. *Probabilistic Engineering Mechanics 10(4)*, 199–207.

Roberts, J. B. and P. D. Spanos (1990). *Random Vibration and Statistical Linearization.* Wiley, New York.

Safak, E. (1989). Adaptive modeling, identification, and control of dynamical structural systems I: Theory. *Journal of Engineering Mechanics 115(11)*, 2386–2405.

Sato, T. and K. Takei (1997). Real time robust identification algorithm for structural systems with time-varying dynamic characteristics. In *Proceedings SPIE 4th Annual Symposium on Smart Structures and Materials*, Bellingham, pp. 393–404.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of statistics 6(2)*, 461–464.

Shi, T., N. P. Jones, and J. H. Ellis (2000). Simultaneous estimation of system and input parameters from output measurements. *Journal of Engineering Mechanics 126(7)*, 746–753.

Sivia, D. S. (1996). *Data Analysis: A Bayesian Tutorial.* Oxford Science Publications.

Soong, T. T. (1990). *Active Structural Control: Theory and Practice.* Wiley, New York.

Spencer, B. F., S. J. Dyke, and H. S. Deoskar (1998). Benchmark problems in structural control: Part I-Active mass driver system. *Earthquake Engineering and Structural Dynamics 27(11)*, 1127–1139.

Spencer, B. F. and D. C. Kaspari (1994). Structural control design: a reliability-based approach. In *Proceedings of American Control Conference*, Baltimore, MD, pp. 1062–1066.

Spencer, B. F., D. C. Kaspari, and M. K. Sain (1994). Reliability-based optimal structural control. In *Proceedings of Fifth U.S. National Conference on Earthquake Engineering*, EERI, Oakland, California, pp. 703–712.

Spencer, B. F. and M. K. Sain (1997). Controlling buildings: A new frontier in feeback. *IEEE Control Systems Magazine on Emerging Technology 17(6)*, 19–35.

Stengel, R. F. and L. R. Ray (1991). Stochastic robustness of linear time-invariant control systems. *IEEE Transactions on Automatic Control 36(1)*, 82–87.

Vanik, M. W., J. L. Beck, and S. K. Au (2000). A Bayesian probabilistic approach to structural health monitoring. *Journal of Engineering Mechanics 126(7)*, 738–745.

Veneziano, D., M. Grigoriu, and C. A. Cornell (1977). Vector-process models for system reliability. *Journal of Engineering Mechanics 103(EM3)*, 441–460.

Warburton, G. B. and E. O. Ayorinde (1980). Optimal absorber parameters for simple systems. *Earthquake Engineering and Structural Dynamics 8*, 179–217.

Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions*. Springer: Prentice Hall.

Yuen, K.-V. (1999). Structural modal identification using ambient dynamic data, MPhil thesis. Technical report, Hong Kong University of Science and Technology, Hong Kong.

Zeldin, B. A. and P. D. Spanos (1998). Spectral identification of nonlinear structural systems. *Journal of Engineering Mechanics 124(7)*, 728–733.

Zhang, Y. F. and W. D. Iwan (2002). Active interaction control of tall buildings subjected to near-field ground motions. *ASCE Journal of Structural Engineering 128(1)*, 69–79.

Zhou, K. K. and J. C. Doyle (1996). *Robust and Optimal Control*. Prentice Hall, New Jersey.