Adaptive Methods within a Sequential Bayesian Approach for Structural Health Monitoring

by

Daniel W. Huff

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2013 by the
Graduate Supervisory Committee:

Antonia Papandreou-Suppappola, Chair
Narayan Kovvali
Chaitali Chakrabarti
Aditi Chattopadhyay

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

Structural integrity is an important characteristic of performance for critical components used in applications such as aeronautics, materials, construction and transportation. When appraising the structural integrity of these components, evaluation methods must be accurate. In addition to possessing capability to perform damage detection, the ability to monitor the level of damage over time can provide extremely useful information in assessing the operational worthiness of a structure and in determining whether the structure should be repaired or removed from service.

In this work, a sequential Bayesian approach with active sensing is employed for monitoring crack growth within fatigue-loaded materials. The monitoring approach is based on predicting crack damage state dynamics and modeling crack length observations. Since fatigue loading of a structural component can change while in service, an interacting multiple model technique is employed to estimate probabilities of different loading modes and incorporate this information in the crack length estimation problem. For the observation model, features are obtained from regions of high signal energy in the time-frequency plane and modeled for each crack length damage condition. Although this observation model approach exhibits high classification accuracy, the resolution characteristics can change depending upon the extent of the damage. Therefore, several different transmission waveforms and receiver sensors are considered to create multiple modes for making observations of crack damage. Resolution characteristics of the different observation modes are assessed using a predicted mean squared error criterion and observations are obtained using the predicted, optimal observation modes based on these characteristics.

Calculation of the predicted mean square error metric can be computationally intensive, especially if performed in real time, and an approximation method is proposed. With this approach, the real time computational burden is decreased significantly and

the number of possible observation modes can be increased.

Using sensor measurements from real experiments, the overall sequential Bayesian estimation approach, with the adaptive capability of varying the state dynamics and observation modes, is demonstrated for tracking crack damage.

To my parents, Margaret and Richard Huff,

and to Anastasia and our two wonderful daughters,

Marguerite and Christina,

for their constant love, support and encouragement.

ACKNOWLEDGEMENTS

I extend tremendous gratitude to my advisor, Professor Antonia Papandreou-Suppappola, for her extraordinary support, encouragement, and patience during my graduate work at Arizona State University. Her assistance and constant, positive attitude have been invaluable.

Dr. Narayan Kovvali has provided tremendous assistance. This has included numerous, after-hours discussions to review and clarify various topics related to my work, and I am very grateful for his mentoring and support.

I would also like to thank Professor Chaitali Chakrabarti for supporting me in this effort. I greatly appreciate the time, effort, and contributions she has provided.

Professor Aditi Chattopadhyay has been a strong pioneer in the development of structural health monitoring methods, as exemplified by her Adaptive, Intelligent Materials and Systems (AIMS) Center of Excellence at Arizona State University. The primary focus of this work would not be possible without her innovative efforts and expertise in this area, and I am very thankful for her support.

Although I have enjoyed the company of all fellow students while studying at Arizona State University, my associations and interactions with several former students have been particularly beneficial for this work. Within the Department of Electrical Engineering, discussions with Dr. Himanshu Shah, Dr. Debejyo Chakraborty and Dr. Wenfan Zhou were particularly helpful with respect to various approaches towards flaw detection and classification. During their studies within the Department of Mechanical and Aerospace Engineering, Dr. Subhasish Mohanty and Dr. Sunil Soni performed stress intensity analyses of mechanical test data and provided Lamb wave signal data used within this work.

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 INTRODUCTION 1

1.1 Motivation 1

1.2 Prior Work on Damage Detection 7

1.3 Proposed Thesis Work 9

1.4 Report Organization 12

2 SIGNAL PROCESSING ALGORITHMS FOR DAMAGE TRACKING . . 13

2.1 Bayesian analysis 14

2.2 Markov chains 15

2.3 Markov Chain Monte Carlo Methods 18

2.3.1 The Metropolis-Hastings Algorithm 19

2.3.2 MCMC implementation 21

2.3.3 Simulated annealing 22

2.4 Sequential Importance Sampling 23

2.4.1 Particle filters 23

2.4.2 Importance sampling 25

2.5 General sequential estimation of dynamic states 26

2.5.1 Sequential Bayesian estimation using SIS 29

2.5.1.1 Particle degeneration, resampling, and rejuvenation . . 30

2.6 Two-dimensional, time-frequency signal representations 32

2.6.1 Matching Pursuit Decomposition 35

2.6.1.1 MPD using Gaussian atoms 36

CHAPTER Page

2.7 Hidden Markov Models and Data Classification 39

2.7.1 Use of HMM for Classification 41

2.7.1.1 The Forward-Backward algorithm 43

2.7.2 Training and re-estimation of HMM parameters 44

2.7.3 Continuous observation densities 46

3 CRACK GROWTH DYNAMICS 48

3.1 Early crack growth theories 48

3.2 Paris' Law 52

3.3 Simplifications and the stochastic nature of crack growth 54

4 INTERACTIVE MULTIPLE MODEL FOR MULTIPLE LOADING MODES 57

4.1 Interacting Multiple Model Approach 57

4.2 IMM Implementation for Multiple Loading Conditions 62

5 ADAPTIVE WAVEFORM SELECTION FOR CRACK GROWTH ESTIMATION 75

5.1 Crack Length Observation Models 75

5.1.1 Observations Models Using HMMs 75

5.1.2 HMM Implementation Issues 76

5.2 Waveform Selection using Multiple Observation Modes 78

5.2.1 Use of HMM for Predicting Observations 78

5.2.2 Predicted crack lengths 79

5.2.3 Waveform Design Using Error Minimization 80

5.2.4 Test Specimen Configuration 82

5.2.5 Transmission Waveforms for Selection 83

5.2.6 Data Collection for Model Training 85

5.2.7 State Dynamics Model 87

5.2.8 MPD-HMM Observation Model For Different Waveforms . . . 91

5.2.9 Crack Growth Tracking Results with Waveform Selection . . . 95

CHAPTER Page

5.3 Implementation Issues and Alternative MSE Metric 105

5.3.1 Approximate Pre-calculated Approach For Waveform Selection 109

5.3.2 General Observation Mode Selection 111

6 CONCLUSIONS AND FUTURE WORK 121

6.1 Summary of Proposed Work 121

6.2 Possible Future Work Extensions 122

REFERENCES 126

LIST OF TABLES

Table Page

4.1 Mode transition probabilities for 2 modes. 70

4.2 Mode Transition Probabilities, 3 modes. 71

4.3 Mode transition probabilities for 4 modes. 73

5.1 Fatigue cycle numbers where crack length data was obtained. 87

5.2 Initial and MCMC-refined estimates of state dynamics model parameter m. . 88

5.3 Initial and MCMC-refined estimates of state dynamics model parameter C. . 89

5.4 Initial and MCMC-refined estimates of state dynamics model parameter K_I. 89

5.5 Confusion matrix for tone burst waveform crack length classification. . . . 94

5.6 Confusion matrix for pseudo-noise random width pulse waveform crack length classification. 95

5.7 Confusion matrix for linear chirp waveform crack length classification. . . . 95

LIST OF FIGURES

Figure Page

1.1 Overall adaptive sequential Bayesian approach. 12

2.1 Example of a Markov chain state diagram (left) and a transition matrix (right). 16

2.2 Particle filter representations. (a) Mono-modal (single peak) particle distribution. (b) Bi-modal particle distribution (two peaks). 24

2.3 Particle filter weight convergence and degeneration. 30

2.4 Resampling. 31

2.5 Extraction of signal energy from an example received burst signal using the MPD procedure. 38

2.6 Three hidden urns containing different distributions of colored balls. 41

2.7 Examples of a state transition diagram, a state transition probability matrix, an observation probability matrix, and an initial state probability matrix corresponding to the *colored ball and urn* model described in [1]. 42

3.1 Reduction in the amount of applied load producing a given deflection when crack area increases from A to $A+\Delta A$. 49

3.2 Increase in crack length with increase in number of fatigue cycles. 49

3.3 Localized stress in vicinity of crack tip. 51

3.4 Crack fracture loading modes. 51

3.5 K_I estimates and least squares cubic model. 52

3.6 Typical $\log(da/dn)$ vs. $\log(K)$ curve, with intercept value 'C' and slope 'm' as used in the Paris equation. 53

4.1 Compact tension specimen with actuator and receiver sensors. 63

4.2 Transmitted tone burst signal. 64

4.3 Received tone burst signal. 65

4.4 MPD-TFR of received tone burst signal. 65

Figure Page

4.5 Maximum and minimum loading conditions and resulting crack growth characteristics for specimen CT419a. 67

4.6 Calculated stepwise ΔK_{eff} of compact tension specimen CT419a. 67

4.7 HMM Confusion matrix for compact tension specimen CT419a. 69

4.8 IMM results assuming 2 loading modes. 70

4.9 IMM mode probabilities assuming 2 loading modes. 71

4.10 IMM results assuming 3 loading modes. 72

4.11 IMM mode probabilities assuming 3 loading modes. 72

4.12 IMM results assuming 4 loading modes. 73

4.13 IMM mode probabilities assuming 4 loading modes. 74

5.1 Biaxial mechanical test machine. 82

5.2 Closeup of tested specimen fatigue crack. 83

5.3 Cruciform specimen with sensors. 84

5.4 Different types of transmitted waveforms. (a) Tone burst, (b) linear chirp, and (c) pseudo-noise random width pulse. 86

5.5 Measured crack length vs. number of fatigue cycles, with least squares cubic model. 87

5.6 Correct crack length classification values for the three different waveforms (diagonal values of the corresponding confusion matrices). 96

5.7 Sequential Bayesian estimation of crack length with MMSE waveform selection. 97

5.8 Predicted MSE results for different crack lengths. 98

5.9 Overall crack tracking results. 99

5.10 Crack tracking with noise added to the data. The waveform selection approach, proposed within this work, was capable of following the general crack growth development, whereas, without waveform selection (e.g., [2]), the tracking procedures tend to drift from the actual crack growth. 100

Figure Page

5.11 Quadratic curves created at crack length estimates using the estimate and its two neighboring estimates. 102

5.12 Smoothed crack growth monitoring for $k = 5$ crack length estimates. 103

5.13 Smoothed crack growth monitoring for $k = 8$ crack length estimates. 104

5.14 Smoothed crack growth monitoring for $k = 14$ crack length estimates. . . . 104

5.15 Smoothed crack growth monitoring using four crack length estimates to obtain preliminary estimates of localized (da/dn) 105

5.16 Smoothed crack growth monitoring using a maximum of the 5 most recent, preliminary crack rate estimates to define the linear relationship between crack rate and $\log(\Delta K_{\text{eff}})$. 106

5.17 Predicted MSE using pre-calculated approach. 110

5.18 Approximation MSEs at fatigue cycle 110009 (actual crack length 4.2 mm) 113

5.19 Approximation MSEs at fatigue cycle 140011 (actual crack length 4.7 mm). 113

5.20 Approximation MSEs at fatigue cycle 161761 (actual crack length 5.13 mm). 114

5.21 Approximation MSEs at fatigue cycle 174238 (actual crack length 7.0 mm). 114

5.22 Approximation MSEs at fatigue cycle 184009 (actual crack length 8.92 mm). 115

5.23 Approximation MSEs at fatigue cycle 192961 (actual crack length 11.12 mm).115

5.24 Approximation MSEs at fatigue cycle 200013 (actual crack length 13.15 mm).116

5.25 Approximation MSEs at fatigue cycle 208012 (actual crack length 15.04 mm).116

5.26 Approximation MSEs at fatigue cycle 213794 (actual crack length 18.09 mm).117

5.27 Approximation MSEs at fatigue cycle 218462 (actual crack length 20.11 mm).117

5.28 Approximation MSEs at fatigue cycle 225008 (actual crack length 22.62 mm).118

5.29 Approximation MSEs at fatigue cycle 229009 (actual crack length 24.75 mm).118

5.30 Approximation MSEs at fatigue cycle 234216 (actual crack length 28.09 mm).119

5.31 Overall crack growth tracking using selection of multiple waveform and sensor combinations). 119

Chapter 1

INTRODUCTION

1.1 Motivation

When evaluating the integrity of critical civil, military, transportation, and other structures and systems, it is crucial that accurate conclusions are obtained. From a safety standpoint, failure of a critical structure can result in bodily injury and loss of life. From an economic standpoint, the amount of time, money, and other investments that must be made to address a failure incident can be enormous. The failure incident must be investigated and possible causes must be determined. Testing must be performed to recreate and simulate the failure as well as to devise and test proposed remedies to prevent similar future incidents. Assuming appropriate solutions to the problem are determined, there can be significant expenditures for implementing and monitoring them, as well as expenditures related to legal liabilities and compensation for parties affected by a failure incident. There can also be significant damage with respect to the reputations of the businesses, personnel, or other organizations deemed to be responsible for allowing an unanticipated failure to occur. Future business and revenue could be severely affected.

Faults and damage within structures exist at various levels, and in appraising the condition of a structure, it is advantageous if one can define the presence of damage, as well as the level of damage. Worden and Delieu-Barton [3] provide specific definitions of the concepts of *faults*, *damage*, and *defects* within structures. A *fault* is when the structure can no longer operate satisfactorily. If one defines the quality of a structure or system as its fitness for purpose of its ability to meet customer or user requirements, it suffices to define a *fault* as a change in the system that produces an unacceptable reduction in quality. *Damage* is when the structure is no longer operating in its ideal condition but can still function satisfactorily, i.e., in a sub-optimal manner. A *defect* is inherent in the material and statistically all materials will contain some unknown amount

of defects, this means that the structure can operate at its design condition even if the constituent materials contain defects.

Farrar and Worden [4] emphasize that the concept of flaws within structures begins at the material level, and that *all materials possess some degree of defects*. As structures are put into service and experience different types of loading scenarios, the faults or defects might grow into more significant damage states. This emphasizes the need for accurate evaluation methods for critical structures and systems, since defects exist to some extent in all structures.

During the design and development of critical structures, *destructive testing* is commonly used to determine or validate the strength and other properties of structures, components, and materials. Representative specimens are mechanically tested under various loads and conditions to evaluate strength capabilities under these situations. Extreme loading conditions are often employed, possibly up to specimen breakage, to obtain measurements within the ultimate strength range of the structure. Artificial flaws can be incorporated into specimens to determine knockdowns in strength characteristics for different levels of damage. Destructive testing often includes procedures where specimens are physically sectioned and internal regions are examined visually, microscopically or with other evaluation methods. Therefore, destructive testing is commonly used during the preliminary design and development of structural components. Even after a component is designed, developed, and put into production, destructive testing is often performed periodically for quality control evaluations using a small sampling of structures pulled from production or service. Although destructive testing information can be valuable, it generally renders the structure of interest unsuitable for its intended operation. There are obvious economic and practical limitations with respect to the extensive use of destructive testing for obtaining structural information and assessing part quality.

In addition to destructive testing, there are a variety of less damaging procedures used to monitor structural quality and obtain information with respect to the health condition of critical structures. In these methods, minimal alteration, if any, occurs to the evaluated components, so that structural integrity is not compromised. Obviously, these are the types of procedures that can be used to evaluate and monitor components currently in or intended for service.

As described by Farrar, Worden, and Dulieu-Barton [3, 4], there are several key multidisciplinary areas used for monitoring actual components and assessing damage. These include *statistical process control*, *nondestructive evaluation*, *structural health monitoring*, *condition monitoring*, and *damage prognosis*. Note that there can be significant overlap within these various areas.

Statistical process control and the related concept of *statistical quality control* pertain to the general philosophy of inherently creating quality within a product during design and manufacturing. Parameters that are particularly critical for part quality are specifically identified and weighed accordingly. This is done so that higher quality awareness is given to parameters that play key roles in determining the quality of a component. Statistical techniques are used to actively monitor and analyze the parameters and methods of evaluating these parameters [5].

Nondestructive evaluation (NDE) refers to a collection of quality assessment techniques that evaluate parts and components using methods that cause insignificant or no damage to the part or component. NDE is commonly performed as a quality control procedure during manufacturing of a part and after the part is put into service, especially when proper performance of the part is critical. A key characteristic of NDE is that it is usually performed off-line, and hence, the component typically has to be removed from manufacturing or service for evaluation. Radiography, ultrasonic inspection, magnetic particle inspection, liquid penetrant inspection, eddy current inspection, and acoustic

emission are some of the most commonly-used NDE methods.

Structural health monitoring (SHM) [6, 7] represents an alternative approach towards NDE of critical structures. In SHM, sensors are attached or otherwise incorporated into a structural component, and information from the embedded sensors is used to monitor the structural integrity of the component. Unlike NDE, a common objective of SHM is that articles can be monitored while in service, rather than having to be taken offline and inspected. Since damage detection can occur in-between periodic inspection and maintenance procedures, this can improve system quality and safety. Furthermore, since components are monitored while in service, this can also reduce the frequency in which they are removed from service for maintenance and inspection. When maintenance is required, previous information obtained from SHM could be beneficial in isolating and resolving structural problems.

Although SHM offers many advantages, it also poses certain challenges. In some cases, embedding a sensor into a structure can perturb the system, creating a structural discontinuity and small flaw within a region that does not have discontinuities to begin with. In weight critical structures, such as those used in aerospace, weight concerns inherently place limitations with respect to the level of sophistication of on-board SHM systems. Other factors that have to be considered include the methods of attaching sensors to the structure of interest, the types of transmission channels used to transmit SHM data, and the anticipated service environments that the structure and the SHM system encounter. Furthermore, the attachment and operation of the sensors must not interfere with the critical operation of the monitored components. Hence, for many potential applications, the use of SHM is still in a developmental stage.

Common approaches used in SHM include Lamb wave based and optical fiber based methods. When using Lamb wave propagation, signals are transmitted into a material of interest, and piezoelectric materials are commonly used in the transmission and

receiver sensing devices [8]. Lamb wave propagation along a material depends upon the density, elastic properties and thicknesses of the material, as well as the signal's frequency components. When Lamb waves encounter interfaces, such as those caused by defects within the material, the wave propagation velocity changes, and hence, frequency, phase and modal relationships also change. These types of changes can be used to detect and classify defects within materials. With fiber Bragg grating methods, light is sent through fibers possessing grated optical sensors within the optical fibers. When light is sent into the grated sensors, a portion of the light is reflected, and the spacing of the grating determines which wavelengths are most affected. When the material is loaded, this causes straining within the material and sensor grating, thereby shifting the wavelengths that are most affected by the grating. When changes occur to the material due to flaws or defects, the strain characteristics of the material and attached grated sensors also change. Changes in wavelength behavior can be used to determine changes in the structural condition of a material.

Note that there can be considerable overlap between approaches traditionally categorized as either NDE or SHM. Limited scope SHM could be used for real time monitoring structures for possible damage, and in the event SHM data detects possible damage, the structure could be pulled off-line for more intensive evaluations using SHM and NDE. Alternatively, SHM sensors internally embedded within a structure might have better capability of detecting internal damage, or damage in limited access areas, as compared to conventional NDE methods. Therefore, information obtained using one approach can be used to supplement information obtained with the other.

The objectives of *condition monitoring* (CM) are somewhat similar to SHM. Like SHM, CM tends to be used in an on-line manner. However, the typical SHM objective is to directly monitor the structural condition of a component, whereas the typical CM objective is to monitor some external force or event that regularly impacts the component and eventually causes predictable degradation over time. The degradation of the

component can be correlated with the magnitude and frequency of the external forces or events. CM is often used in rotating structures or structures that are subjected to significant vibration while in service. In such structures, the number of rotations or the amount of vibration can be monitored using accelerometers, seismic transducers, piezoelectric transducers, or other types of sensors.

Damage prognosis (DP) is a closely-related and somewhat overlapping area of study with respect to assessing structural integrity. Whereas the intentions of NDE and SHM often focus on measuring the current condition of a structure, damage prognosis generally includes predictive assessments related to defining the remaining service life of a damaged structure or the allowable conditions under which a damaged structure can still operate in a reliable, safe manner. Methods and models are created for predicting damage growth and future conditions of a structure. Various approaches have been devised to model fatigue crack growth during the past century. In the 1960s, Paris and fellow researchers [9, 10] proposed that crack growth state could be modeled using the stress intensity factor range associated with the loading conditions, as well as other key material and fatigue-related parameters. In recent years, more-sophisticated physics-based and *multi-scale modeling* methods [11] have been used. The objective of these methods is to model physical processes governing crack growth and overall fatigue behavior, and where the multi-scale methods have endeavored to model the physical processes and interactions from the microscopic level up through the macroscopicl structural levels. Various statistical, pattern recognition methods have been employed to incorporate past damage information for predicting future damage conditions. tMohanty, Liu, Chattopadhyay, and collaborators [12–15] employed Gaussian Process statistical pattern recognition methods for estimating residual useful life and general damage estimation within structures.

Successful implementation of damage evaluation approaches requires accurate and efficient feature extraction and data analysis techniques. The various data analysis

techniques, including filtering, feature extraction and statistical methods, can be applied to data obtained using both an off-line, NDE style inspection, or an in-service, SHM style, real-time monitoring. Regardless of the approach used to obtain the data, analysis approaches should possess high capability of detecting and accurately classifying damage state within structures.

1.2 Prior Work on Damage Detection

A number of studies involving active sensing have been performed to improve the ability to detect and locate damage within structures. Kessler, et al. [16] varied frequency, amplitude, envelope shape, and other parameters related to tone burst waveforms transmitted as Lamb waves in composite laminates, and evaluated the suitability of the resulting waveforms using a general assessment of time domain plots and wavelet transform coefficients. Sohn, et al. [17] employed a wavelet-based approach towards detection of simulated delaminations in composites, where they defined a *damage index* metric based upon a damaged/undamaged ratio of wavelet transform time marginals. Fasel, et al. [18] examined the use of sinusoids with various chaotic signal envelopes. Received signals were modeled using autoregressive time series models, and statistics of the residual error signals were compared for undamaged and damaged specimens. A genetic algorithm was used to modify the chaotic signal envelope to increase statistical differences between residues of the damaged condition specimens as compared to those associated with the undamaged, baseline condition. Lee and Staszewski [19,20] studied the propagation of Lamb waves within materials for detecting and locating damage. A local interaction simulation approach (LISA) was used, along with power spectral density, wavelet analysis, and time-space wave propagation analysis approaches. Kuttig, et al. [21], and Raghavan, et al. [22], employed chirplets to represent regions of Lamb wave dispersive mode energy for detecting and locating damage. Borkowski, et. al, [23] developed a fully-coupled, electromechanical elastodynamic model for Lamb wave propagation applicable to heterogeneous, anisotropic material systems, where mechanisms

related to Lamb wave dispersion, coupling, and interaction with material defects were addressed. In experimental studies involving SHM, there can be significant time and cost when exploring and fine tuning experimental parameters, especially when materials, geometries, and general specimen design can vary. One key advantage stated in [23] is the proposed model can be used to predict Lamb wave behavior within arbitrary specimen designs such that experimental parameters can be defined in a more intelligent, efficient manner.

In monitoring the progression of damage, parameters related to a damage growth models need to be estimated. With the massive increase in computational power during the last few decades, the use of computer simulation and sampling methods have dramatically increased the ability to perform parameter and state estimation. These techniques are particularly useful when the estimation problem includes mathematically intractable quantities or when there are other difficulties in obtaining samples of quantities of interest. In [24], the Markov chain Monte Carlo (MCMC) method [25] was used for damage estimation; it is an example of computer simulation methods commonly employed in these types of situations.

A variety of advanced signal processing and estimation methods have been proposed for SHM applications. In [26, 27], the matching pursuit decomposition (MPD) [28] was used for feature extraction in damage detection applications to create compact time-frequency representations of signals corresponding to different damage conditions. The MPD was combined with the statistical hidden Markov model (HMM) approach to model different damage conditions for use in SHM damage classification [29, 30]. Sequential Monte Carlo techniques have also been employed for monitoring of progressive damage in structures [2, 30].

1.3 Proposed Thesis Work

In this work, we build upon previous signal processing SHM efforts, and propose an overall SHM-based sequential Bayesian estimation approach with several improvements to expand the capability for adapting to different loading and damage conditions.

Within the basic sequential Bayesian methodology, we employ physics-based models such as Paris law to provide estimates of crack rate and damage propagation within a fatigue environment. We use an active sensing approach to develop our observation model. This model integrates the MPD time-frequency method and hidden Markov models (HMMs) to create and learn feature vectors from signals corresponding to different damage conditions. When new sensor measurements are obtained, the trained HMMs can be used to classify observations with respect to crack length. We show how the combination of MPD and HMMs can be used to obtain accurate crack length observations. As compared to previous work using MPD and HMMs for damage classification, in which there were typically several different levels of damage, within this work, we specifically show how the MPD-HMM approach can be to obtain accurate damage estimation in situations involving large numbers (e.g., up to 32) of possible damage levels.

Within our proposed approach, we recognize that the ability to monitor conditions can change, and we address this by including different model variations. Since dynamic systems and structural components can be subjected to multiple types of fatigue loading conditions, we propose an interacting multiple model (IMM) approach [31–33] to weigh the probabilities of different loading conditions and loading models. The IMM is an efficient filtering method that has been successfully used in other applications such as object and target tracking. Different types of movements and maneuvers of an object, as well as different fatigue loading conditions, can be modeled separately. The IMM approach can help recognize changes to the state dynamics and weigh the corre-

sponding models accordingly, thereby expanding the capability for monitoring different conditions.

Within this work, using an active sensing approach towards crack length observation, we explore the use of different transmission waveforms, as well as different receiver sensors used to capture the transmitted signal information. Although the MPD-HMM based observation model demonstrates exceptional capability for recognizing different damage conditions, the ability to resolve different levels of damage can vary with changes to the damage condition and observation-related parameters. As a result, for different estimated levels of damage, we improve damage resolution capability by selecting the optimal transmission waveform at each time step. In this selection approach, we adapt the predicted mean-squared error (MSE) of estimation as our performance metric [34]. The use of this metric relates to differences that can occur in the modeling capability of HMM models that have been trained with different waveforms given a particular crack length damage condition.

Previous work by Zhou [30] discussed the use of an MSE-based approach for sensor scheduling in SHM applications; however, it was also recognized there can be severe limitations on the number of possible sensor combinations due to increased computational burden as this number increases. Therefore, we develop an approximate MSE approach, where most of the computation burden can be performed off-line. Although this approach produces an approximation of the MSE metric, we show how this approximation approach produces results that are in good agreement with the real-time approach. Since the approximation approach reduces the real-time computational burden substantially, this allows a significant increase in the number of possible transmission waveforms, receiver sensors, and other parameters used for observing the damage condition. We demonstrate this with an example involving three different transmission waveforms and five different receiver sensors.

In performing sequential Bayesian estimation, certain quantities might exhibit levels of variation that are larger than desired, and therefore, we demonstrate several approaches for reducing variation. When extracting parameters related to the Paris Law crack propagation model, the inherent power law structure of the model makes parameter estimation particularly susceptible to extreme levels of variability. Using assumed distributions of the parameters, we demonstrate a Markov chain Monte Carlo approach to define acceptable, constrained values for the parameters. When obtaining crack length observations within extreme noise environments, the resulting crack length estimates can also exhibit significant levels of noise and crack rate inconsistency. Employing the assumed linear log-scale relationship between the crack growth rate and the corresponding stress intensity factor range, we extract noisy, localized crack rate values from our noisy crack length estimates and use the linear log-scale relationship to smooth over the noise.

Figure 1.1 shows key components within our work. Within the diagram, given we have a crack length estimate at time step $k-1$, our overall objective is to obtain an estimate of the crack state damage at fatigue time step k. In general, the state dynamics model provides an initial estimate of the distribution of the crack length at time step k. However, given there can be multiple loading conditions, multiple approaches towards modeling the propagation, or other variations in the damage propagation, the IMM approach provides flexibility for addressing the different possibilities. As the damage condition of the structure changes, or as other factors cause changes in methods used to observe the damage condition, the MSE approach allows us to assess and select the best predicted mode(s) for observing the damage. If there are more than a handful of possible modes of observing the damage, computational costs associated with calculating the MSE metric can become significant. Therefore, from a practical implementation standpoint, we demonstrate an approximation approach towards calculating an MSE estimate, such that a larger number of observation parameters (e.g, transmission wave-

forms, receiver sensors, etc.) can be evaluated for making the damage observation. For our observation model, we use an MPD-HMM approach for observing the crack length damage. As compared to previous efforts employing an MPD-HMM approach, we show how this observation model can be used to obtain accurate damage classification when there are numerous possible levels of damage.

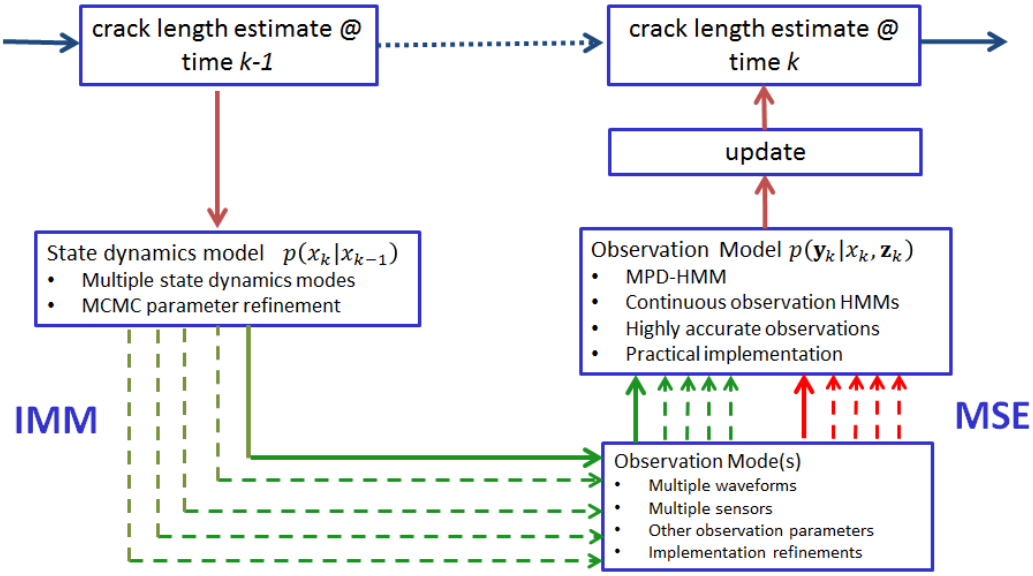


Figure 1.1: Overall adaptive sequential Bayesian approach.

1.4 Report Organization

This report is organized as follows. In Chapter 2, we give background regarding several of the key analytical and estimation tools employed within our basic overall damage monitoring approach. Chapter 3 consists of a brief historical overview of basic fracture mechanics concepts, since the state dynamics models used within our work possess a fundamental fracture mechanics basis. The use of the IMM approach to combine or select from several different state dynamics models is described in Chapter 4, as this method is used to. Chapter 5 describes a predictive mean squared error approach for combining or selecting from various modes of observing damage to improve the robustness in making these observations. Finally, in Chapter 6, concluding remarks and potential future work are discussed.

Chapter 2

SIGNAL PROCESSING ALGORITHMS FOR DAMAGE TRACKING

In this chapter, we describe the basic analytical tools utilized within our overall crack growth tracking methods. Our overall method uses a sequential Bayesian analysis approach, so we briefly discuss Bayes theorem and Bayesian analysis, including the general concepts of prior and posterior knowledge, and later in the chapter, we consider sequential Bayesian analysis. Markov chain Monte Carlo methods can be used for parameter refinement in our state dynamics model, and hidden Markov modeling is used to perform statistical pattern recognition of signal feature vectors corresponding to different damage conditions. Therefore, the basic idea of the Markov property and how it can be used to create Markov chains are discussed, and later in the chapter, Markov chain Monte Carlo and hidden Markov modeling techniques are addressed. For our crack length observations, we transmit short duration signals through a structural material of interest and receive the signals at various sensors positioned on the material. The received signal features are analyzed within the time-frequency plane, where the time-frequency signal features, in conjunction with hidden Markov modeling, are used to observe the crack length damage condition within the material. Therefore, we briefly discuss time-frequency signal processing, and then focus upon matching pursuit decomposition (MPD), which is the method employed to extract signal features within the time-frequency plane. We then discuss hidden Markov modeling (HMMs), which is a statistical, artificial learning method used to correlate the MPD, time-frequency signal features with different crack length damage conditions. The extraction of received signal features using MPD, and the crack length recognition of the MPD information by the hidden Markov models (HMMs), forms the based of our crack length observation model. Finally, we discuss the use of the MPD-HMM observation model in a reverse, emission mode, where we can use trained HMMs to generate predicted sequences of

MPD feature vectors. The predicted sequences of MPD feature vectors can then be analyzed by trained HMMs to give probabilities of predicted crack lengths. These predicted crack length probabilities are used to estimate error characteristics of the predicted observations. We employ several different options for modeling damage propagation or for observing the damage, and we use predicted error characteristics for deciding among the different available options.

2.1 Bayesian analysis

General approaches towards classification, decision, and general data analysis often rely upon Bayesian approaches. The original ideas were proposed in 1763 by Reverend Thomas Bayes [35], and the statistical concept known as Bayes theorem has been a powerful tool in statistical inference and data analysis [36–39]. The inherent structure of Bayes theorem and Bayesian data analysis is well suited for iterative estimation methods and many iterative, computer-based estimation methods employ a sequential, Bayesian approach [40].

Given a set of parameters θ and a set of related observations $\mathbf{y}$, the relationships between conditional and joint probabilities can be arranged such that it is possible to reverse the conditioning of one variable upon the other. Given $p(\mathbf{y},\theta) = p(\mathbf{y}|\theta)p(\theta)$, and $p(\mathbf{y},\theta) = p(\theta|\mathbf{y})p(\mathbf{y})$, then one conditional probability can be derived using the other:

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)\ p(\theta)}{p(\mathbf{y})} \quad \text{or} \quad p(\mathbf{y}|\theta) = \frac{p(\theta|\mathbf{y})\ p(\mathbf{y})}{p(\theta)} \tag{2.1}$$

Alternatively, one can view Bayes Theorem as a way of updating the *prior* knowledge or belief of some event given a new observation (or new *evidence*) related to the event. Focusing upon the first expression in (2.1), the term $p(\theta)$ can be thought of as the *prior* distribution of the parameters θ. After obtaining an observation $\mathbf{y}$, knowing the *likelihood* of the observation $\mathbf{y}$ given the parameters θ, and knowing the probability of the observation $p(\mathbf{y})$, then the *posterior* distribution $p(\mathbf{y}|\theta)$ can be calculated.

In an iterative, Bayesian approach to estimation of parameters θ given an observation $\mathbf{y}$, several steps are typically performed. These include predicting θ using the available prior information $p(\theta)$, obtaining an observation $\mathbf{y}$ related to θ, where the likelihood of the observation is $p(\mathbf{y}|\theta)$, and updating the prior information to obtain posterior knowledge of the parameter given the observation $p(\theta|\mathbf{y})$ using (2.1). In a sequential time step approach, if the posterior at time step k (i.e., $p(\theta_k|\mathbf{y}_k)$) can be used in creating the prior for the next time step $p(\theta_{k+1})$, a Bayesian procedure can be performed in an iterative, sequential manner, and this is generally referred to as *sequential Bayes estimation*. In many applications, the updated, posterior probability from one time step is used without modification as the prior probability for the next time step (i.e.,$p(\theta_{k+1}) = p(\theta_k|\mathbf{y}_k)$).

In the expressions shown in (2.1), it is common to express the denominator as an expectation of the observations given the parameters. For example, we can express $p(\mathbf{y})$ as:

$$p(\mathbf{y}) = \begin{cases} \int p(\mathbf{y}|\theta)\ p(\theta)\ d\theta, & \text{for } \theta \text{ continuous.} \\ \sum p(\mathbf{y}|\theta)\ p(\theta), & \text{for } \theta \text{ discrete.} \end{cases}$$

where the integral (continuous case) or summation (discrete case) is performed over the possible range of values for θ. In the discrete case, (2.1) becomes:

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)\ p(\theta)}{\sum p(\mathbf{y}|\theta)\ p(\theta)} \quad \text{or} \quad p(\mathbf{y}|\theta) = \frac{p(\theta|\mathbf{y})\ p(\mathbf{y})}{\sum p(\theta|\mathbf{y})\ p(\mathbf{y})} \tag{2.2}$$

where the denominators in (2.2) act as normalizing constants so that the posterior probability sums up to one. In many situations, it is only necessary to determine relative probabilities up to some proportionality:

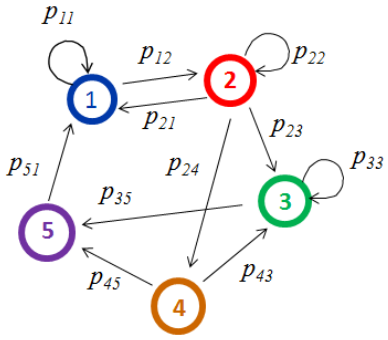
$$p(\theta|\mathbf{y}) \propto p(\mathbf{y}|\theta)\ p(\theta) \quad \text{or} \quad p(\mathbf{y}|\theta) \propto p(\theta|\mathbf{y})\ p(\mathbf{y}) \tag{2.3}$$

2.2 Markov chains

Consider a system of N states $X = \{x_1, x_2, \ldots, x_N\}$, where a sequence of these states $Q = \{q_0, q_1, q_2, \ldots\}$ is generated at time steps $k \geq 0$. Given a state x_i at time step k,

if the probability of the next state x_j at time step $k+1$ is $\Pr(x_j|x_i)$, where this probability depends only upon the current state x_i and is independent of previous states and previous transitions, then the sequence exhibits the *Markov property*. The sequence exhibiting the Markov property is called a *Markov chain*, or a *Markov process*, and $\Pr(\cdot|\cdot)$ is called the *transition kernel* of the chain [41]. If a system possesses the Markov property, this can be especially useful in a Bayesian context. For example, in Bayesian, computer-based estimations that assume the Markov property, only the most recent prior $p(\theta)$ has to be stored before making observation y and updating to obtain the posterior $p(\theta|y)$. All previous observations and priors do not have to be stored, thereby simplifying computer storage and computational requirements.

Using the notation p_{ij} to represent the transition probability from state i to state j, these probabilities $\{p_{ij}\}$ can be assembled into a state transition probability matrix $\mathbf{P}$. Figure 2.1 shows an example state transition diagram, along with its corresponding state transition probability matrix. In this example, there are several instances where a state is allowed to transition back to itself, creating a *self transition*.



i \ j	1	2	3	4	5
1	p_{11}	p_{12}	0	0	0
2	p_{21}	p_{22}	p_{23}	p_{24}	0
3	0	0	p_{33}	0	p_{35}
4	0	0	p_{43}	0	p_{45}
5	1	0	0	0	0

Figure 2.1: Example of a Markov chain state diagram (left) and a transition matrix (right).

Given N possible states, the transition probability matrix $\mathbf{P}$ is a $N \times N$ square matrix. Furthermore, each row contains the probabilities of transitioning from state i to

state j, and therefore, the sum of the elements of each row must sum up to 1:

$$\sum_{j=1}^{N} p_{ij} = 1$$

When discussing Markov chains and Markov processes, it is common to use π notation to designate the probability of being in some state, for example, using π_j to designate the probability of being in state j, or alternatively, to denote the proportion of time spent in state j during a long-run Markov chain. Since π_j represents a probability, the sum of all probabilities π_j must equal one:

$$\sum_{j=1}^{N} \pi_j = 1$$

Note that the collection of the $\{\pi_j\}$ probabilities represents a probability distribution of the states, where this distribution is commonly notated as a row vector π:

$$\pi = \begin{bmatrix} \pi_i & \pi_j & \dots \end{bmatrix} \tag{2.4}$$

In addition, since each element π_j of row vector π represents a probability of being in state i, and each element p_{ij} of the state transition matrix $\mathbf{P}$ represents the probability of transitioning from state i to state j, the sum of the product $\pi_i p_{ij}$ over all states $i = 1, \dots, N$ is equal to the probability π_j (i.e., the probability of being in state j):

$$\pi_j = \sum_{i=1}^{N} \pi_i \, p_{ij}$$

Using the vector-matrix versions of these quantities, common interests associated with long-run Markov chains pertains to determining if a *stationary distribution* π exists, where the key characteristic of the stationary distribution is that $\pi = \pi\mathbf{P}$, and if the stationary distribution exists, under what conditions does it exist.

A Markov chain is *irreducible* when there is a positive probability that every state i can transition to every state j ($j = i$ included) in a finite number of steps. An irreducible Markov chain is *aperiodic* if for some time step $t \geq 0$ and some state j,

$$\Pr(x_t = j \mid x_0 = j) \; > \; 0 \text{ and } \Pr(x_{t+1} = j \mid x_0 = j) \; > \; 0$$

. Therefore, a Markov chain is *aperiodic* when the number of steps to return to the same state is not a multiple of a number greater than one. The *detailed balance condition* occurs when

$$\pi_i \, p_{ij} = \pi_j \, p_{ji}, \quad \forall \; i \neq j. \tag{2.5}$$

With the detailed balance condition, there is equal probability within the Markov chain of a transition from state j to state i as there is of a transition from state i to state j.

2.3 Markov Chain Monte Carlo Methods

Within our work, there are situtations in which it is desired to obtain refined estimates of model parameters used within our crack damage monitoring approach. In some cases, the initial estimates of model parameters can vary greatly. Given previously-obtained or real-time crack length observations, and given various target values and defined constraints, techniques such as Markov chain Monte Carlo can be used to refine parameter values within the defined constraints. In this section, we briefly discuss Markov chain Monte Carlo techniques.

Although a common Markov chain problem relates to determining if and under what conditions the stationary distribution π exists, Markov chain Monte Carlo (MCMC) methods often assume that π is a known target distribution. A typical objective of MCMC applications is to obtain samples from this distribution, especially in situations where π is a multivariate distribution or otherwise presents difficulties when it is desire to obtain samples from π. Rather than attempting to obtain samples directly from π, in MCMC, we obtain samples from a *proposal distribution* $q(\cdot)$, and then, using MCMC procedures, samples obtained from the proposal distribution eventually converge to the desired target distribution π.

Various notational, terminology, and conceptual differences exist between general Markov chain and MCMC procedures. In the previous discussion of Markov chains, we discussed a state transition probability matrix $\mathbf{P}$ possessing elements p_{ij}, where these

elements represented the probability of transitioning from state i to state j. In MCMC, the analogous transition matrix is commonly notated as $\mathbf{Q}$ with matrix elements $q(x,y)$, where $q(x,y)$ represents a transition probability from x to y. However, in the context of MCMC, $q(x,y)$ is thought of as the probability of a proposed (and not necessarily accepted) transition – given a current value x in the Markov chain, $q(x,y)$ represents a probability of obtaining a *proposed* value y to be the next value in the Markov chain. Given the current sample x and proposed sample y, MCMC methods also include a probability $\alpha(x,y)$ of accepting y as the next value in the Markov chain. If y is rejected, the current value x is used again as the next value in the Markov chain. To be consistent with MCMC notation commonly used in literature, we are using both x and also y to represent state quantities in our MCMC Markov chain, where x represents the current state quantity in the chain, and y represents a *proposed* state quantity to be added to the chain. This is specifically mentioned because in other discussions within this document, y and variations of y are used to represent observations of a state, rather than the value of the state. With respect to other notational differences, in MCMC notation, $\pi(x)$ and $\pi(y)$ are commonly used to designate the probabilities of states x and y, respectively, whereas in our discussion of Markov chains, we used subscripts i and j to designate different states, and used the notation π_i and π_j to denote the probabilities of being in state i and state j.

One of the most common types of MCMC methods is the Metropolis-Hastings algorithm and this are discussed below.

2.3.1 The Metropolis-Hastings Algorithm

The Metropolis-Hastings (M-H) algorithm was originally developed by Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller [42] in 1953 and was subsequently generalized by Hastings [43] in 1970. However, it was not until the mid 1990s before it became popular, when increased computational capability became available and pro-

ponents of the algorithm, including Tierney [44], illustrated its usefulness. The 1995 tutorial by Chib and Greenberg also provides a good overview of M-H methods [45].

Referring to (2.5) for the detailed balance condition for Markov chains, the analogous expression for the M-H algorithm is

$$\pi(x)\, q(x,y)\, \dot{\alpha}(x,y) \;=\; \pi(y)\, q(y,x)\, \dot{\alpha}(y,x) \tag{2.6}$$

where $\dot{\alpha}(x,y)$ and $\dot{\alpha}(y,x)$ are the probabilities of accepting a proposed transition from x to y and from y to x, respectively, and where the specific values of these acceptance probabilities allow for the equality condition. If we take the ratio of the two values and redefine this quantity as $\alpha(x,y) = \dot{\alpha}(x,y)/\dot{\alpha}(y,x)$, then we can define a new, combined acceptance probability for the transition from state x to proposed state y:

$$\alpha(x,y) = \frac{\pi(y)\, q(y,x)}{\pi(x)\, q(x,y)} \tag{2.7}$$

To account for situations in which $\alpha(x,y) > 1$, the probability of accepting the proposed sample is defined as follows:

$$\Pr(\text{accept } y) \;=\; \min\, \{\alpha(x,y),\, 1\} \tag{2.8}$$

Therefore, given that we are currently at time step t in the Markov chain, and the current value is x, the M-H algorithm will generate the sample at time step $t+1$ as follows:

1. Generate a proposed sample y from the proposal distribution $q(x,y)$ and obtain the probability of the proposed sample y.
2. Calculate the probabilities $\pi(y)$ and $q(y,x)$ using the proposed sample y and current sample x.
3. Calculate the probability $\pi(x)$ (or retrieve from storage if it was calculated and stored in the previous M-H iteration).
4. Calculate $\alpha(x,y)$ using (2.7). If $\alpha(x,y) \geq 1$, accept the proposed sample y for time step $t+1$.
5. If $\alpha(x,y) < 1$, draw a random sample $u \sim U[0,1]$. If $\alpha(x,y) > u$, accept the proposed sample y for time step $t+1$. Otherwise, use the current state x for time step $t+1$.

If the proposal density is symmetric such that $q(y,x) = q(x,y)$, then (2.7) simplifies to:

$$\alpha(x,y) = \frac{\pi(y)}{\pi(x)} \tag{2.9}$$

In a version of M-H known as *independent Metropolis-Hastings*, the proposal density is only concerned with the next proposed value y and is "independent" of the current value x in the chain, so that $q(x,y)$ becomes $q(y)$ and $q(y,x)$ becomes $q(x)$, and then (2.7) becomes:

$$\alpha(x,y) = \frac{\pi(y)\ q(x)}{\pi(x)\ q(y)} \tag{2.10}$$

Note that using the term "independent" in (2.10) is not completely accurate since the acceptance of proposed value y depends upon the current value x.

2.3.2 MCMC implementation

In implementing MCMC, there are typically several issues that arise. These include selecting the proper models, selecting the proper distributions, choosing an appropriate Markov chain length, and choosing an appropriate burn-in length, which is the length of the initial portion of the chain (before convergence) that is typically ignored.

A number of studies have been performed to study convergence diagnostics of Markov chains, especially with respect to implementing the proper burn-in and overall chain lengths. One of the first studies to address MCMC convergence issues was performed by Gelman and Rubin [46]. In addition, Robert and Casella [47], as well as a collection of works edited by Gilks, Richardson and Spiegelhalter [41], provide additional information regarding MCMC convergence. A summary of MCMC convergence methods has been provided by Cowles and Carlin [48]. Typical approaches to check for convergence include the use of multiple sequences, multiple starting points and/or graphical methods to provide multiple and alternative sources of evidence to detect convergence. In many cases, researchers have used previous experience and trial and error to determine burn-in and overall lengths. If results are not satisfactory, then lengths are increased.

Note that the use of multiple sequences can also be used to improve independence of samples resulting from MCMC methods. Instead of using the last L samples (after burn-in) from a single chain to represent samples from a distribution, the last sample(s) obtained from multiple chains could be collected to create a sequence with inherently more independent samples.

2.3.3 Simulated annealing

In computer simulation and estimation methods such as MCMC, the use of *simulated annealing* can be useful to improve the ability of the estimation procedure to converge to a global optimal estimate, rather than settling on a suboptimal, local estimate [25]. The term "annealing" refers to a metallurgical procedure in which a metal is raised above its critical temperature and then is slowly cooled to room temperature. At temperature, the material softens, which allows internal stresses within the material to relieve and dislocations within the material to redistribute, such that ductility and uniformity of the material is improved. In computer-based estimation techniques involving *simulated annealing*, samples are initially drawn from a sampling distribution possessing relatively wide sampling boundaries such that global optimal regions are included. Eventually, after the estimation procedure has progressed and estimates have started to converge, the width of the sampling boundaries is slowly decreased such that sampling becomes more and more focused towards an eventual region of convergence.

Simulated annealing can be performed at more than one level. For example, in our work, we perform the M-H algorithm multiple times, and we include simulated annealing within each M-H run and also between different M-H runs. Within a M-H run, during the "burn-in" period, simulated annealing is performed by reducing the variance of the proposed samples after every M-H sampling iteration. After "burn-in," the variance is held constant, and a set of M-H samples is obtained, the mean of the post "burn-in" samples is obtained, and this mean value is used as the initial value for the

next M-H run. Prior to beginning the next M-H run, the initial variance is widened to begin the "burn-in" period, but this initial variance is reduced as compared to the initial variance of the previous M-H run. Therefore, within each M-H run, simulated annealing is performed by reducing the variance during the "burn-in" period, and between M-H runs, simulated annealing is performed by reducing the overall variance values used in the next M-H run as compared to the previous M-H run.

2.4 Sequential Importance Sampling

Sequential importance sampling (SIS) is another computer-based method that has gained significant popularity during the past several decades with the increases in computational capability. In SIS, the concept of importance sampling is performed in an iterative manner to estimate and track a dynamically changing state quantity. Whereas MCMC methods tend to focus on fixed target distributions and generating samples from such distributions, SIS methods typically focus on the estimation and tracking of parameters within dynamic states.

SIS involves several intermingling concepts, including particle filters, importance sampling, and sequential Bayesian estimation of dynamic state quantities. Therefore, we will first give a brief overview on particle filters, followed by a discussion on importance sampling and how particle filters can be used in this application. We will then examine overall concepts related to sequential estimation and tracking of dynamic states. After discussing the general, overall approach, we will then discuss the SIS approach. We will then discuss various implementation issues and methods that have been used to deal with SIS implementation.

2.4.1 Particle filters

Estimation in the SIS procedure is performed using *particle filters*. Particle filters consist of a collection of discrete support values, also known as *particles*, which are stored with a corresponding set of weight values, where each weight designates the

importance of the corresponding particle value. The collection of paired particles and weights is created in a manner so that the collection can be related to the distribution of one or more parameters of interest.

Figure 2.2 shows two example particle filter diagrams where the particle filters are depicted at given iterations within a SIS procedure. In these diagrams, the particle values are represented by the location of the circles along a value axis, and their corresponding weights are represented by the sizes of the circles. Note that a particle filter can represent a distribution related to some multi-modal quantity possessing peaks at more than just one location. Traditional mathematical estimation methods might have difficulty estimating or tracking a multi-modal quantity, whereas particle filters are less hindered for characterizing such distributions.

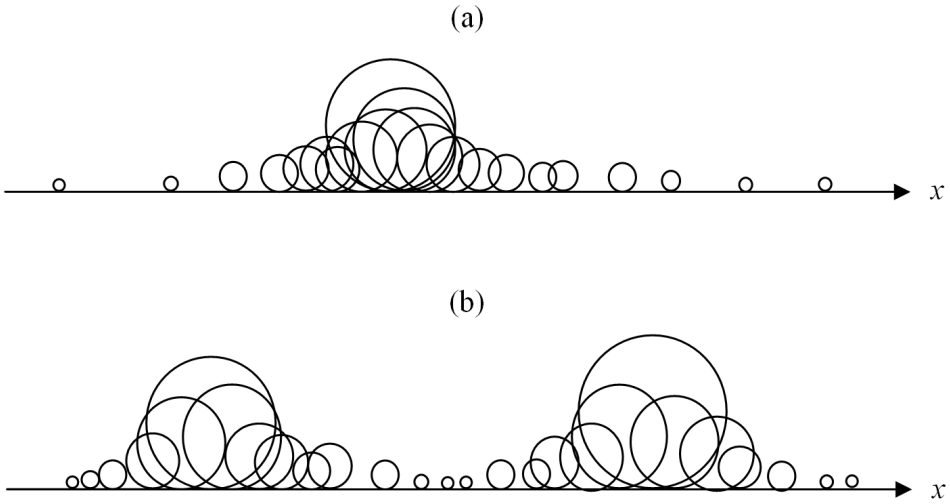


Figure 2.2: Particle filter representations. (a) Mono-modal (single peak) particle distribution. (b) Bi-modal particle distribution (two peaks).

As with other iterative Bayesian estimation methods such as the Kalman Filter [49], given a dynamically-changing state vector of interest $\mathbf{x}$, where we possess the ability to make observations $\mathbf{y}$ that are related to the state, a typical objective with SIS is to approximate the posterior distribution $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ of the state as if evolves over time, where the notation $1:k$ is used to denote iterations 1 through k. However, with

SIS, particle filter representations are employed to represent a distribution related to the state. Assume we have N_p particles $\mathbf{x}^{(i)}, i = \{1, 2, \ldots, N_\mathrm{p}\}$ that have been sampled from some appropriate proposal density and each particle possesses an associated weight $w^{(i)}$. Using the notation $\left\{\mathbf{x}_k^{(i)}, w_k^{(i)}\right\}$ to represent the {value, weight} pair of the i^{th} particle of a particle filter at iteration k, and where the particle weights have been normalized such that $\sum_i w_k^{(i)} = 1$, then the approximation of the distribution can be expressed as:

$$p\left(\mathbf{x}_k \mid \mathbf{y}_{1:k}\right) \approx \sum_i w_k^{(i)}\ \delta\left(\mathbf{x}_k - \mathbf{x}_k^{(i)}\right). \tag{2.11}$$

From the approximation of the state distribution at various iterations, we can obtain estimates related to the state distribution, such as mean and variance.

2.4.2 Importance sampling

Assume we are trying to estimate a posterior distribution of a state vector $\mathbf{x}$ at iteration k, given observations $\mathbf{y}$ from iterations $1:k$. In other words, we are trying to estimate $p\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right)$. Unfortunately, this distribution may be difficult, if not impossible, to express in closed form, and it might also be difficult, if not impossible, to obtain samples from the distribution. Therefore, as with the previous discussion of MCMC methods, we choose a proposal distribution $q\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right)$ to obtain samples, where the distribution $q(\cdot)$ can be written in closed form and it is relatively easy to obtain samples from the distribution.

More generally, given that the state vector $\mathbf{x}$ is distributed according to $p\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right)$, it is often desired to estimate characteristics of a function of the state vector $h\left(\mathbf{x}_k\right)$. For example, the expectation over the state distribution $p\left(\mathbf{x}_k\right)$ can be expressed as:

$$\mathbb{E}_p\left[h\left(\mathbf{x}_k\right)\right] = \int h\left(\mathbf{x}_k\right) p\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right) d\mathbf{x}. \tag{2.12}$$

Assuming samples will be drawn from proposal distribution $q\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right)$, and assuming we can obtain some measure of an estimate for the ratio of densities

$$\frac{p\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right)}{q\left(\mathbf{x}_k|\mathbf{y}_{1:k}\right)}$$

we can express the integral in Equation (2.12) as:

$$\int h(\mathbf{x}_k)\, p(\mathbf{x}_k|\mathbf{y}_{1:k})\, d\mathbf{x} = \int h(\mathbf{x}_k) \left(\frac{p(\mathbf{x}_k|\mathbf{y}_{1:k})}{q(\mathbf{x}_k|\mathbf{y}_{1:k})} \right) q(\mathbf{x}_k|\mathbf{y}_{1:k})\, d\mathbf{x} = \mathbb{E}_q \left[h(\mathbf{x}_k) \left(\frac{p(\mathbf{x}_k|\mathbf{y}_{1:k})}{q(\mathbf{x}_k|\mathbf{y}_{1:k})} \right) \right] . \tag{2.13}$$

For a particle filter with N_p particles $\mathbf{x}_k = \left\{ x_k^{(i)} \right\},\ i = 1 : N_\mathrm{p}$, the expectation in (2.13) can be approximated as:

$$\mathbb{E}_q \left[h(\mathbf{x}) \left(\frac{p(\mathbf{x})}{q(\mathbf{x})} \right) \right] \approx \frac{1}{N_\mathrm{p}} \sum_{i=1}^{N_\mathrm{p}} h\left(x_k^{(i)} \right) \left(\frac{p\left(x_k^{(i)}|\mathbf{y}_{1:k} \right)}{q\left(x_k^{(i)}|\mathbf{y}_{1:k} \right)} \right) = \sum_{i=1}^{N_\mathrm{p}} h\left(x_k^{(i)} \right) w_k^{(i)}$$

where

$$w_k^{(i)} = \frac{1}{N_\mathrm{p}} \left(\frac{p\left(x_k^{(i)}|\mathbf{y}_{1:k} \right)}{q\left(x_k^{(i)}|\mathbf{y}_{1:k} \right)} \right) \propto \frac{p\left(x_k^{(i)}|\mathbf{y}_{1:k} \right)}{q\left(x_k^{(i)}|\mathbf{y}_{1:k} \right)} \tag{2.14}$$

represents the weight of the i^{th} particle.

Therefore, using a particle filter and a proposal density $q(\cdot)$, an approximate measure of the expectation of a function $h(\mathbf{x})$ can be obtained.

2.5 General sequential estimation of dynamic states

Several methods are available to estimate and track dynamically changing states, and in this section, we examine general principles related to such estimation procedures.

Following the discussion in [50], given a dynamically changing state vector $\mathbf{x}$, the inherent dynamics of $\mathbf{x}$ can be represented by:

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \tag{2.15}$$

where $\mathbf{f}_k$ is some function at iteration k, possibly non-linear, which propagates the state $\mathbf{x}$ from iteration $k-1$ to iteration k, and $\mathbf{v}_{k-1}$ is the process noise, possibly non-Gaussian, associated with state $\mathbf{x}$ and the process $\mathbf{f}(\cdot)$.

Using the sequential Bayesian estimation approach, assuming we already have a prior distribution estimate of the propagated state $\mathbf{x}$ at iteration k, to obtain the posterior

estimate at iteration k, an observation $\mathbf{y}_k$ is made, typically in the presence of some type of observation noise. The observation at iteration k can be represented by:

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k) \tag{2.16}$$

where $\mathbf{h}_k$ is some observation function at iteration k, possibly non-linear, which relates the state $\mathbf{x}_k$ to the observation $\mathbf{y}_k$, and $\mathbf{n}_k$ is the observation noise, possibly non-Gaussian, associated with the observation $\mathbf{y}_k$.

Within the sequential Bayesian approach, the prior distribution can be expressed as:

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k, \mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1})\, d\mathbf{x}_{k-1}. \tag{2.17}$$

The joint density on the right side of (2.17) can be factored to produce the Chapman-Kolmogorov equation:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{1:k-1})\ p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})\ d\mathbf{x}_{k-1}$$

Assuming a Markov process, where $p(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k-1}) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$, the expression for the prior distribution becomes:

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1})\ p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})\ d\mathbf{x}_{k-1}. \tag{2.18}$$

After estimation of the prior distribution in (2.18), an observation generally becomes available, and this new information is used to update the prior distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})$ to the posterior distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$. The posterior distribution can be re-expressed

as follows:

$$\begin{aligned} p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) &= p(\mathbf{x}_k|\mathbf{y}_{1:k-1},\mathbf{y}_k) \\ &= \frac{p(\mathbf{x}_k,\mathbf{y}_{1:k-1},\mathbf{y}_k)}{p(\mathbf{y}_{1:k-1},\mathbf{y}_k)} \\ &= \frac{p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{y}_{1:k-1})\ p(\mathbf{x}_k,\mathbf{y}_{1:k-1})}{p(\mathbf{y}_{1:k-1},\mathbf{y}_k,)} \\ &= \frac{p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{y}_{1:k-1})\ p(\mathbf{x}_k|\mathbf{y}_{1:k-1})\ p(\mathbf{y}_{1:k-1})}{p(\mathbf{y}_{1:k-1},\mathbf{y}_k)\ p(\mathbf{y}_{1:k-1})} \\ &= \frac{p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{y}_{1:k-1})\ p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_{1:k-1}|\mathbf{y}_k)} \end{aligned} \tag{2.19}$$

In the numerator of (2.19), if we assume we have *memoryless* observations, i.e.,

$$p(\mathbf{y}_k|\mathbf{x}_k,\mathbf{y}_{1:k-1}) = p(\mathbf{y}_k|\mathbf{x}_k),$$

and furthermore, if we express the denominator in (2.19) as an integration over all $d\mathbf{x}_k$, then the expression for the posterior distribution becomes:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)\ p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{\int p(\mathbf{y}_k|\mathbf{x}_k)\ p(\mathbf{x}_k|\mathbf{y}_{1:k-1})\ d\mathbf{x}_k} \tag{2.20}$$

Note that the expression (2.20) for the posterior distribution depends upon the likelihood function $p(\mathbf{y}_k|\mathbf{x}_k)$, which depends upon the observation model shown in (2.16). Also note that the form of the denominator in (2.20) is consistent with the discussion concerning the denominator in (2.2), where the denominator essentially serves as a scaling constant to normalize the density.

Therefore, several key equation models are addressed during sequential Bayesian estimation. The dynamics model, shown in (2.15), represents the propagation of the state of interest as it moves from one time step to the next. The observation model (2.16) represents a relationship of our observations with respect to the state during these time steps. The expression for the prior distribution of the state, shown in (2.18), portrays our knowledge of the state prior to making an observation at the given iteration, and the expression for the posterior distribution, shown in (2.20), represents the updated knowledge of the state after making the observation.

If the dynamics and observation functions in (2.15) and (2.16) are linear and the corresponding types of noise are Gaussian, then the Kalman filter (KF) [49] can be used to obtain optimal estimates of the state vector. In addition, variations of the KF, such as the extended Kalman filter (EKF) [51], are available for situations in which deviations from linearity and Gaussian noise are not too severe. If these assumptions are not valid, expressions for the prior (2.18) and posterior (2.20) distributions are typically difficult, if not impossible, to express in convenient closed form, especially with a multidimensional state vector $\mathbf{x}$. Therefore, SIS procedures, which are discussed next, offer an alternative approach to numerical estimation methods.

2.5.1 Sequential Bayesian estimation using SIS

With SIS, importance sampling is performed in an iterative manner so that information and estimates obtained during previous iterations are not completely thrown away, but are used to develop the next estimates. As shown with general importance sampling in (2.14), calculation of particle weights is proportional to the ratio of the distribution of interest to the proposal distribution. With SIS, as shown in [50], and assuming the importance density only depends upon the latest observation and the previous state, a recursive expression relating the weights to the previous weights is as follows:

$$w_k^{(i)} \propto \frac{p\left(\mathbf{y}_k|\mathbf{x}_k^{(i)}\right) p\left(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}\right)}{q\left(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)},\mathbf{y}_k\right)} w_{k-1}^{(i)} \tag{2.21}$$

Therefore, in SIS, assuming we draw samples $\mathbf{x}_k^{(i)} \sim q\left(\mathbf{x}_{k-1}^{(i)}|\mathbf{x}_{k-1}^{(i)},\mathbf{y}_k\right)$ at time step k, we can then use (2.21) to obtain the updated weights $w_k^{(i)}$ corresponding to the particles. This will give us the collection of $\left\{\mathbf{x}_k^{(i)}, w_k^{(i)}\right\}$ particle-weight pairs to construct the particle filter at time step k, as previously discussed and depicted in Figure 2.2.

An interesting simplification occurs if we can use the prior distribution $p\left(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)}\right)$ as the proposal distribution $q\left(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)},\mathbf{y}_k\right)$. This simplified version is sometimes re-

ferred to as the *naive* proposal distribution. In this case, the weight update expression becomes:

$$w_k^{(i)} \propto p\left(\mathbf{y}_k|\mathbf{x}_k^{(i)}\right) w_{k-1}^{(i)} \tag{2.22}$$

Figure 2.3 depicts several iterations of a particle filter. Observations and estimates obtained during the iterations cause the particle filter to reweigh the particle weights in accordance with (2.21) or (2.22), where particles representing better estimates of the state are weighted heavier than particles representing poorer estimates of the state.

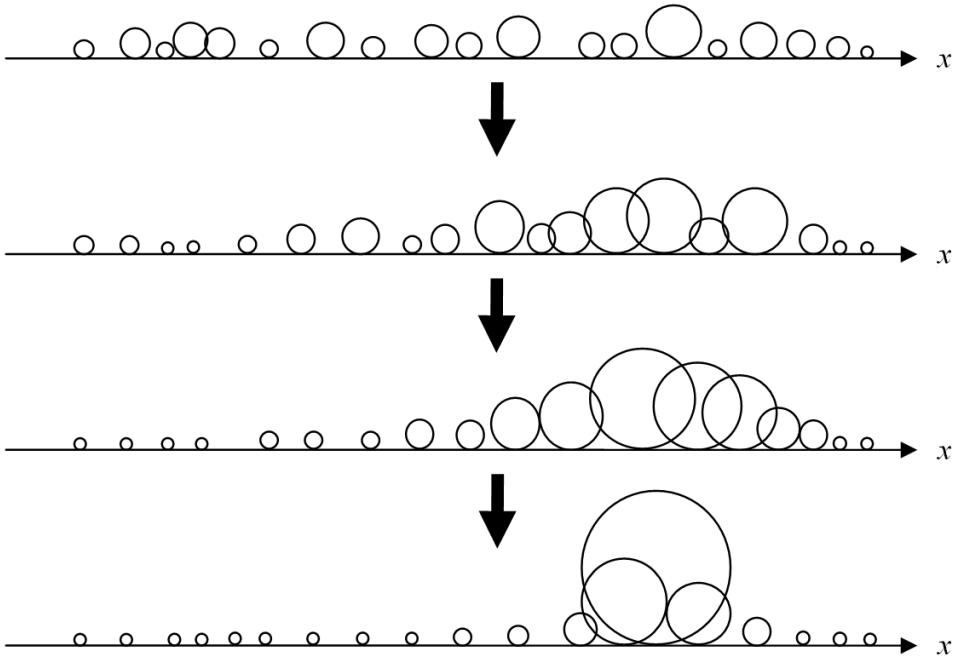


Figure 2.3: Particle filter weight convergence and degeneration.

2.5.1.1 Particle degeneration, resampling, and rejuvenation

As shown by the last frame in Figure 2.3, after only a few iterations, particle filters tend to converge to a situation where only a few particles are weighted significantly. Most other particles possess insignificant weights and contribute very little to the estimation process, and therefore, computations involving these particles are largely wasteful. This situation is known as *particle degeneration*. To improve efficiency and

reduce the number of computations involving insignificant particles, it is common to perform a procedure known as *resampling*. In this procedure, the particle filter weights are used to create a cumulative distribution function (cdf) of the particle values, and new particle values are sampled from the cdf based upon the sizes of the corresponding steps in the cdf. As shown in Figure 2.4, this produces a situation where all particles are weighted equally, but there are a greater number of particles in regions where particles of the previous iteration possessed higher weights. Note that as random process noise is added to the particles in accordance with the state dynamics model (2.15), replicate particle values will spread apart to create more of a continuous distribution in the region.

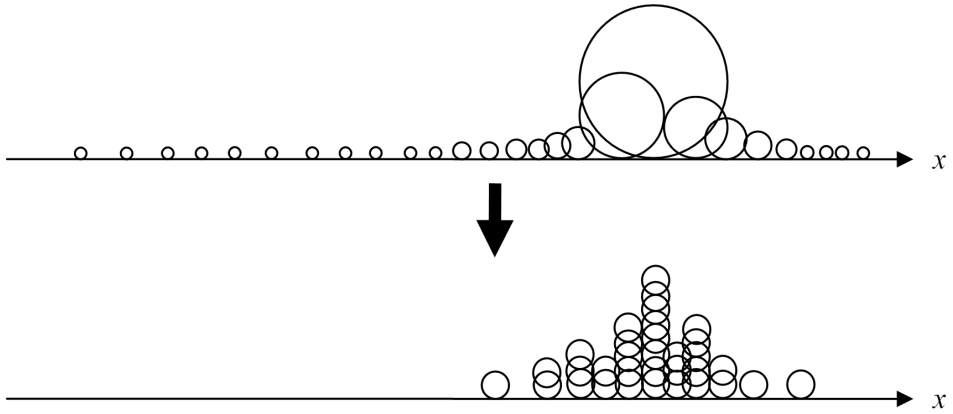


Figure 2.4: Resampling.

Resampling can be performed when the *effective* percentage of contributing particles falls below a preset threshold value. In this case, an estimate for the effective number of particles is:

$$N_{\text{eff}} = \frac{1}{\sum_i \left(w_k^{(i)}\right)^2} \tag{2.23}$$

where $w_k^{(i)}$ is the normalized weight of the i^{th} particle after iteration k. It is also common in many applications to perform resampling after each iteration to ensure that the next iteration has a complete set of contributing particles. If resampling is applied every iteration and the proposal distribution used is the *naive* distribution $p\left(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}\right)$, the resulting procedure has been deemed the *sampling importance resampling* (SIR) filter

[50]. Since resampling is performed every iteration and all weight values are equal after the resampling procedure, the weight update expression in (2.22) becomes:

$$w_k^{(i)} \propto p\left(\mathbf{y}_k|\mathbf{x}_k^{(i)}\right). \tag{2.24}$$

As an alternative to resampling, Lee and Chia [52] have proposed the use of MCMC to create a set of *rejuvenated* particles, which they claim possess a reduced amount of noise accumulated over various SIR iterations. In this procedure, an $(m+1)$-step batch observation approach is used where observations consist of the current and m-step *future* observations $(\mathbf{y}_{k:k+m})$, where the term *future* refers to iterations that occurred after the current iteration of interest, which is physically possible when the "current" iteration actually occurred previously and is now being evaluated. In this approach, a *naive* proposal distribution is used to update the weights:

$$w_k^{(i)} \propto p\left(\mathbf{y}_{k:k+m} \mid \mathbf{x}_k^{(i)}, \mathbf{y}_{1:k-1}\right) w_{k-1}^{(i)}.$$

After updating the weights, a Kullback-Leibler (KL) distance is calculated at each iteration:

$$\kappa(w_k, w_{k-1}) = \sum_{i=1}^{N} w_k^{(i)} \left(\log w_k^{(i)} - \log w_{k-1}^{(i)}\right).$$

If the KL distance is less than a designated threshold value, then the MCMC particle rejuvenation is performed using the independent M-H algorithm, where the proposal density is a normal distribution, and where the sample mean and covariance of the latest set of particles are used as the normal distribution parameters. The target density is $p(\mathbf{x}_k|\mathbf{y}_{k:k+m})$. As discussed earlier, for independent M-H, the probability of accepting the proposed sample is given by (2.8) and (2.10).

2.6 Two-dimensional, time-frequency signal representations

Within our sequential Bayesian, crack growth tracking method, after making an initial prediction of the crack length distribution at a particular time step, we then transmit a signal waveform and make an observation of the received waveform, where

this observation is used to update the predicted crack length distribution. In addition to more familiar, one-dimension signal representations, such as time-domain and frequency domain signal representations, a signal can also be described simultaneously in both dimensions using time-frequency representations. Although one-dimensional, Fourier-based analysis methods are very important in the analysis of signals and numerous other engineering problems, many real-world signals and problems are inherently non-stationary. In our application, having a signal representation in two dimensions can be useful in characterizing the time-varying nature that occurs within our received signal waveforms, where different time-frequency signal features can correspond to different structural damage conditions.

In the 1940's, Dennis Gabor recognized there were limitations using Fourier approaches to analyze many real-world signals [53]. For example, the Fourier transform is specifically defined for time = $-\infty$ to $+\infty$, whereas real-world signals are finite in length and furthermore, can possess non-stationarity. Gabor was specifically cognizant of these facts, and gave particular interest to studying signals with respect to localized characteristics within a two-dimensional, time-frequency viewpoint.

When representing a signal simultaneously with respect to both time and frequency, note that time and frequency are inversely related. Therefore, there is an inherent tradeoff between the amount of resolution that can be obtained in the time dimension with the amount of resolution that can be obtained simultaneously in the frequency dimension. Based upon Heisenberg's original concept of the Uncertainty Principle, Weyl, and later, Gabor (in his studies mentioned above) adapted the Uncertainty Principle concept towards signal processing and time-frequency relationships. In the context of time-frequency, the Uncertainty Principle is sometimes referred to as the Heisenberg-Gabor Uncertainty Principle, which includes the following time duration - frequency

bandwidth relationship:

$$T_s B_s \geq \frac{1}{4\pi}. \tag{2.25}$$

In (2.25), which is known as the *Gabor limit*, T_s is the time duration resolution (in seconds), B_s is the frequency bandwidth resolution (in Hertz), and the time-bandwidth product $T_s B_s$ cannot be less than $1/4\pi$. Using the Cauchy-Schwarz inequality, the lower limit equality condition can be realized when the time-domain signal is a Gaussian, where the Fourier transform is also a Gaussian, such that it is possible to have a linear-dependent, inverse relationship between the time-domain and frequency-domain representations. In this case, this results in a two-dimensional Gaussian in the time-frequency plane possessing a highly localized concentration of signal energy. A basic Gaussian signal can be described by:

$$g(t) = Ce^{-\kappa t^2} \tag{2.26}$$

where C controls the amplitude of the Gaussian signal and κ controls the Gaussian "spreading".

Many types of time-frequency representations (TFRs) have been and continue to be developed [54–56]. However, with the Heisenberg-Gabor Uncertainty Principle, the Gabor limit (2.25), and the inherent trade-off between time resolution and frequency resolution, the choice of a particular type of TFR, and specific parameter values associated with a selected TFR, depend heavily upon the desired time-frequency resolution characteristics. In the next section, we discuss a method that can be used to extract key features of a signal within time-frequency space, where these key signal features can be represented by localized signal structures. The use of localized signal structures has advantages with respect to time-frequency resolution. Furthermore, the localized signal structures can be utilized as features within pattern recognition techniques.

2.6.1 Matching Pursuit Decomposition

Matching pursuit decomposition (MPD) is one approach for representing signals within the time-frequency plane [28]. Considering Gabor's focus upon localized signal energy within time-frequency space, the MPD procedure exploits this concept by projecting a signal onto a dictionary of localized signal structures that have been positioned within the time-frequency plane. After determining which dictionary element corresponds to the greatest magnitude inner product (i.e., the best *matching* dictionary element), this dictionary element is weighted by the value of the inner product and then subtracted from the signal, creating a residual signal. This procedure is then performed on the residual signal, and is repeated iteratively on each subsequent residual signal until a sufficient amount of the signal energy has been extracted. The original signal can then be approximated by a summation of the weighted dictionary elements that were extracted from the signal during the MPD procedure.

One advantage of the MPD approach relates to the high resolution characteristics that can be obtained by creating TFRs of the individual extracted signal structures, and then combining the individual TFRs, resulting in a representation designated as the MPD-TFR [28]. This is in contrast to other approaches in which a TFRs are created of the entire signal. When a TFR of an entire signal is created using a linear TFR, the signal is multiplied with some external function to generate the two-dimensional signal representation. However, the dependency on the external function inherently causes a reduced resolution, spreading effect. In quadratic TFRs (QTFRs), two signal terms are multiplied together, thereby creating an intrinsically two-dimensional quantity. Since QTFRs are not dependent upon an external function (although some QTFRs employ windowing type functions), QTFRs can exhibit significantly better resolution as compared to linear TFRs. However, the basic quadratic nature of QTFRs creates cross-term components that can clutter and interfere with ability to distinguish the auto-term com-

ponents within the TFR. With the MPD-TFR approach, high resolution QTFRs can be performed on individual, localized signal structures within the time-frequency plane. Since cross-terms are only created within individual QTFRs and are not generated between individual QTFRs, overall cross-term interference can be reduced significantly as compared to performing a QTFR on the entire signal.

This general procedure has been extended into adaptive approaches suitable for decomposition and reconstruction using non-linear, as well as linear signal structures [57]. A key advantage of this approach is that the overall signal can be decomposed and concentrated into a limited number of adaptively-modified signal substructures. Since signal energy is concentrated into a limited number of individual structures that have been specifically adapted to the overall signal, the signal can be represented in the time-frequency plane with high resolution.

2.6.1.1 MPD using Gaussian atoms

In addition to matching a signal to a dictionary of specific, or adaptively-modified signal substructures, it is also possible to use a dictionary consisting of highly-localized, Gaussian atoms that have been positioned within the time-frequency plane in known regions of interest and at multiple scalings at each time-frequency location. A general construction of a basic Gaussian signal was previously discussed and is shown in Equation (2.26), and the time-shifted, frequency-shifted and scaled versions are of the form:

$$g_\eta(t) = C_\eta e^{\kappa^2(t-\tau)^2} \cos(2\pi\nu t) \tag{2.27}$$

where τ is the time shift, ν is the frequency shift, and κ is a (positive) scaling parameter controlling the amount of "spreading" exhibited by the atom. Therefore, the parameter set $\eta = \{\tau, \nu, \kappa\}$ describes each Gaussian atom $g_\eta(t)$ within the dictionary, and the normalizing constant C_η is used such that each Gaussian atom possesses unit energy.

In the MPD procedure used within this work, multiple sets of these time-shifted, frequency-shifted Gaussian atoms were employed. In each set, the Gaussian atoms were

positioned over the time-frequency plane at the same time-shift τ and frequency-shift ν locations, but each set possessed a unique scale parameter κ.

The choice of the Gaussian atoms for the dictionary offers several advantages. In addition to offering highly localized time-frequency characteristics, the Gaussian atoms also possess convenient, closed form expressions for use in mathematical transformations such as the Fourier transform and Wigner distribution. Furthermore, when calculating inner products in the MPD procedure, the inherent time-frequency structure of the dictionary atoms can significantly increase computational efficiency, since fast Fourier transform methods can be employed [28].

To initialize the MPD procedure using the Gaussian atom dictionary described above, an inner product ρ is obtained between the signal of interest $s(t)$ and each normalized Gaussian atom. Then, based upon the inner product exhibiting the maximum magnitude $|\rho|$, the *best matching* Gabor atom g_η is identified, along with its corresponding set of parameters $\eta = \{\tau, \nu, \kappa\}$. This procedure is iteratively repeated upon each resulting residue signal, and with the time-shifted, frequency-shifted and scaled Gaussian atoms as the elements of the dictionary, the original signal can be expressed as follows:

$$s(t) = \sum_{\ell=0}^{L-1} \rho_l \; g_{\eta\ell}(t) + r_L(t) \tag{2.28}$$

where $r_L(t)$ is the residue signal after L MPD iterations have been performed, where the calculation of the inner product is performed as follows:

$$\rho_\ell = \int_{-\infty}^{+\infty} r_l(t) \; g^*_{\eta,\ell}(t) \; dt, \quad \ell = 0, \ldots, L-1 \tag{2.29}$$

At each MPD iteration, the specific Gaussian atom is selected from dictionary $\mathscr{D}$ according to the maximum magnitude of the inner product:

$$\hat{g}_{\eta,\ell}(t) = \underset{g_\eta \in \mathscr{D}}{\operatorname{argmax}} \left| \int_{-\infty}^{+\infty} r_l(t) \; g^*_{\eta,\ell}(t) \; dt \right| \tag{2.30}$$

If the summation shown in (2.28) is truncated to L terms, then an approximation of the signal can be expressed as:

$$s(t) \approx s_L(t) = \sum_{\ell=0}^{L-1} \rho_\ell \, \hat{g}_{\eta,\ell}(t) \tag{2.31}$$

where L is typically chosen such that a sufficient amount of the signal energy has been extracted from the original signal. Figure 2.5 shows an example of performing MPD on one of the received, tone burst signals included in this study. The figure shows the amount of signal energy remaining after each of the 40 MPD iterations used to decompose the signal.

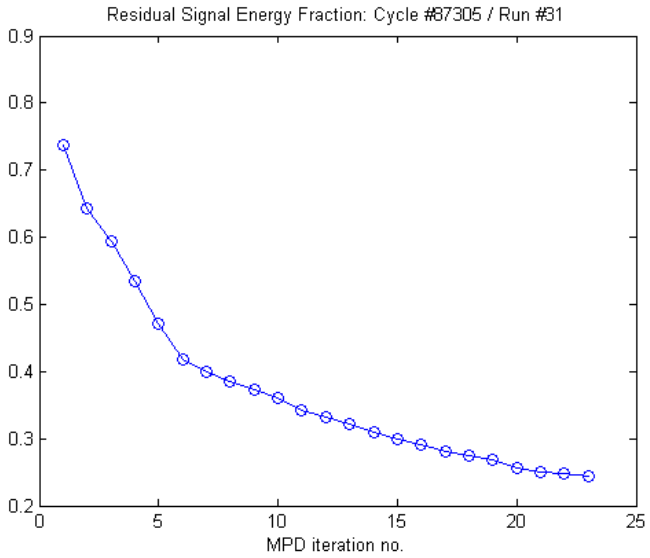


Figure 2.5: Extraction of signal energy from an example received burst signal using the MPD procedure.

Note that each MPD iteration results in a vector of four elements, where three of the elements are given by the set of parameters characterizing the Gaussian atom chosen at the lth iteration $\eta_\ell = \{\tau_\ell, \nu_\ell, \kappa_\ell\}$, and the fourth parameter is the inner product of the

atom with the residual signal, given by (2.29), such that:

$$\mathbf{y}_l = \begin{bmatrix} \tau_\ell \\ \nu_\ell \\ \kappa_\ell \\ \rho_\ell \end{bmatrix}, \quad \ell = 0,\ 1,\ \ldots,\ L-1 \tag{2.32}$$

If the collection of four-element MPD vectors $\{\mathbf{y}_l\}$ are arranged with respect to their time-shift parameter τ, this essentially creates a time-based sequence of four-element feature vectors occurring in the time-frequency plane. Arranging the vectors into a matrix, each successive column represents a 4-parameter, MPD feature vector that is *emitted* from the time-frequency plane:

$$\mathbf{Y} = [\mathbf{y}_{\tau_1}\ \mathbf{y}_{\tau_2}\ \cdots\ \mathbf{y}_{\tau_L}] \tag{2.33}$$

If this time-based sequence of emitted MPD feature vectors within the time-frequency plane can be related to some other dynamic process or overall condition of interest, it might be possible to use an artificial learning or pattern recognition method to establish the relationships. Within our crack growth monitoring approach, we use hidden Markov models (discussed next) to relate different MPD feature vector sequences to different structural damage conditions.

2.7 Hidden Markov Models and Data Classification

We briefly reconsider the Markov property and employ it for developing hidden Markov models (HMMs). Given a set of N possible states $H = \{h_1, h_2,\ \ldots\ , h_N\}$, a sequence of K state transitions $\mathbf{Q} = \{q_1, q_2,\ \ldots,\ q_k, q_{k+1},\ \ldots,\ q_K\}$, where allowed state transitions might include self-transitions from the current state back to the same state, and a transition probability matrix $\mathbf{P}$, where matrix elements p_{ij} represent the probabilities of transitioning from hidden state h_i to hidden state h_j, the fundamental idea of the Markov property and a Markov process relates to the memoryless nature of the process

in transitioning from one state to another. The probability of the transition to the next state is only dependent upon the current state and not the history of previous states.

In an HMM, the sequence of the state transitions $\mathbf{Q} = \{q_1, q_2, \ldots, q_K\}$ is commonly *hidden* and not directly observed, but observations are obtained from the emissions of the hidden states, where the probabilities of the observed symbol emissions depend upon the particular current state h_j. Given a set of M possible symbols $\mathbf{v} = \left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots, \mathbf{v}^{(M)}\right\}$ that can be present within some observation sequence, the probabilities of the symbols are organized in a matrix $\mathbf{B} = \left\{b_j\left(\mathbf{v}^{(m)}\right)\right\}$, where $1 \leq j \leq N$ represents the state index and $\mathbf{v}^{(m)}$ represents the mth observation symbol. In addition, a vector of the initial probabilities for the N possible states $\pi = \{\pi_1, \pi_2, \ldots, \pi_N\}$, where $\pi_i = \Pr(q_i = h_i)$, must also be specified. Therefore, for a simple *discrete* HMM, given N possible states, with M possible observation symbols that can be emitted, the HMM λ consists of a state transition probability matrix $\mathbf{P}$, a state-dependent, observation symbol probability matrix $\mathbf{B}$, and an initial state probability vector π. Therefore, in addition to specifying the number of states N and the number of distinct observation symbols M, the key parameters θ_λ in the *discrete* HMM λ are:

$$\theta_\lambda = \{\mathbf{P}, \mathbf{B}, \pi\} \tag{2.34}$$

The basic theory behind HMMs was first attributed to Baum and colleagues in the mid 1960s and early 1970s, at which time they were referred to as probabilistic functions of Markov chains [58, 59]. Speech recognition is one application in which HMMs have been used extensively [1, 60, 61].

A toy example was presented in [1], using a *colored ball and urn* model. In this example, there are N hidden urns, and each urn has a different distribution of M colored balls. This is illustrated in Figure 2.6 with $N = 3$ urns (*hidden states*) and $M = 5$ colored balls (*observation symbols*). Assume one of the hidden urns is selected without our knowledge of which urn was selected. A ball is randomly removed from the hidden

urn, but the balll is displayed to us. The ball is then placed back into the same hidden urn (such that the distribution in the urn remains the same). If this process is repeated multiple times, this will generate a sequence of colored ball observations. Using the HMM parameters in (2.34), the observation sequence depends on the probabilities of transitioning between urns (state transition matrix **P**), the probabilities of the color of the selected ball in a given urn (observation probability matrix **B**), and the initial urn probabilities (initial state probability vector π). An example of a state transition diagram, together with the HMM corresponding parameters, is demonstrated in Figure 2.7.

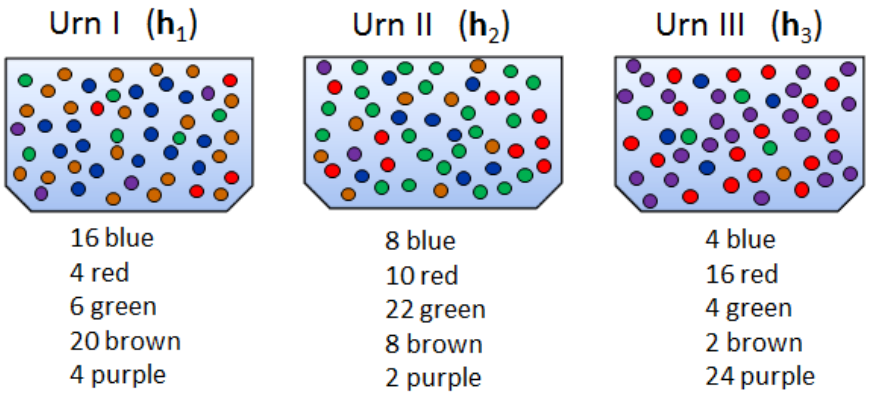
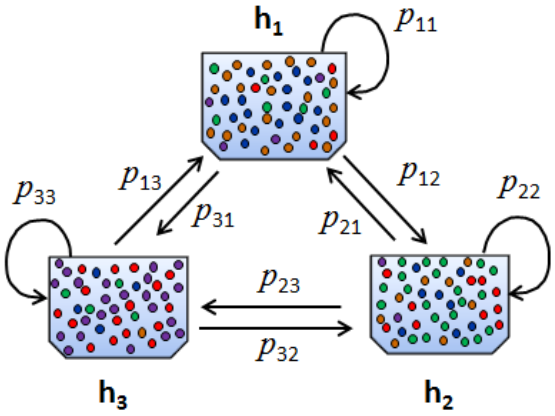


Figure 2.6: Three hidden urns containing different distributions of colored balls.

Figure 2.7 shows an example state transition diagram for the urn and colored balls example (where all possible state transitions are allowed), along with the state transition matrix **P**, the state observation probability matrix **B**, and the initial state probability vector π.

2.7.1 Use of HMM for Classification

The HMMs can be used to solve three main problems [1], and in our use of HMMs for crack length classification, we are primarily concerned about two of these three problems. Given a training observation data sequence $\mathbf{Y}^{\text{train}}$ corresponding to a damage condition d, and given a new or partially trained HMM λ with parameters



State transition probability matrix $\mathbf{P}$

q_k \ q_{k+1}	h_1	h_2	h_3
h_1	p_{11}	p_{12}	p_{13}
h_2	p_{21}	p_{22}	p_{23}
h_3	p_{31}	p_{32}	p_{33}

Observation probability matrix $\mathbf{B}$

	1)blue	2)red	3)green	4)brown	5)purple
h_1	b_1(blue)	b_1(red)	b_1(green)	b_1(brown)	b_1(purple)
h_2	b_2(blue)	b_2(red)	b_2(green)	b_2(brown)	b_2(purple)
h_3	b_3(blue)	b_3(red)	b_3(green)	b_3(brown)	b_3(purple)

Initial state probability vector $\boldsymbol{\pi}$

$$\boldsymbol{\pi} = [\pi_1 \quad \pi_2 \quad \pi_3]$$

Figure 2.7: Examples of a state transition diagram, a state transition probability matrix, an observation probability matrix, and an initial state probability matrix corresponding to the *colored ball and urn* model described in [1].

$\theta = \{\mathbf{P}, \mathbf{B}, \pi\}$, we can use the training data to adjust the model parameters θ to maximize $\Pr\left(\mathbf{Y}^{\text{train}}|\lambda\right)$. Alternatively, given we have a set of D HMMs λ_d, $d = 1, \ldots, D$, that have been trained for D different damage conditions, and then we are presented with a new, unknown test sequence $\mathbf{Y}^{\text{test}}$, we can use each of the trained HMMs λ_d to determine

the likelihood $\Pr(\mathbf{Y}^{\text{test}}|\lambda)$ of the observation sequence given λ_d, and from this information, we can infer the probabilities of the different damage conditions. (In the third common problem addressed by HMMs, given an observation sequence $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K\}$ and a HMM model λ, the problem relates to determining a state transition sequence $\mathbf{Q} = \{q_1, q_2, \ldots, q_K\}$ that is optimal in some meaningful sense such that it best "explains" the observation sequence. This issue is not a concern within this work and is not discussed further.)

2.7.1.1 The Forward-Backward algorithm

Probability measures in the HMM are systematically and iteratively calculated at each time step using a lattice structure approach. For example, within our overall damage monitoring procedure, the *forward-backward algorithm* is employed to obtain the likelihood of a set of observations given a damage condition $p(\mathbf{Y}^{\text{test}}|\lambda_d)$. In this algorithm, given K time steps and N possible states, the forward variable $\alpha_{k+1}(j)$ is the probability of the observations $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{k+1}\}$ up to time step $k+1$, given the model λ and given the system is in hidden state h_j at time step $k+1$. This probability is calculated in an iterative manner for each state at each observation time, creating a lattice structure over the states and the observation times. Calculations at time step $k+1$ are performed using the previous results obtained at time step k, along with the transition probabilities to this state from each of the possible states at the previous time step, as well as the observation probability $b_j(\mathbf{y}_{k+1})$ of the observed symbol at time step $k+1$ for hidden state h_j:

$$\alpha_{k+1}(j) = \left[\sum_{i=1}^{N} \alpha_k(i) p_{ij}\right] b_j(\mathbf{y}_{k+1}), \quad 1 \le k \le K-1 \text{ and } 1 \le j \le N \tag{2.35}$$

Initialization of the forward variable is performed for each hidden state h_i using the initial state probability π_i and the probability of the first observation given the state $b_i(\mathbf{y}_1)$:

$$\alpha_1(j) = \pi_i \, b_i(\mathbf{y}_1) \, , \; 1 \le i \le N$$

Somewhat analogous to the forward variable $\alpha_k(j)$, a backward variable $\beta_k(i)$, can be solved iteratively in the backwards direction:

$$\beta_k(i) = \sum_{j=1}^{N} p_{ij}\, b_i(\mathbf{y}_{k+1})\, \beta_{k+1}(j), \;\; k = K-1,\, K-2,\, \ldots, 1 \;\text{ and }\; 1 \leq i \leq N \tag{2.36}$$

where initialization for this backwards variable for each state h_i occurs at the last observation at final time step K:

$$\beta_K(i) = 1,\; 1 \leq i \leq N$$

Note that after spanning the entire observation sequence, a summation of the forward variables $\alpha_K(j)$ at time step $k = K$, or alternatively, a summation over the backward variables $\beta_1(i)$ at time step $k = 1$, can be used to obtain the probability of the data given an HMM. For example, given observation sequence $\mathbf{Y}$ and HMM λ, using the forward variables at the final time step K, we can obtain

$$\Pr(\mathbf{Y}|\lambda) = \sum_{j=1}^{N} \alpha_K(j) \tag{2.37}$$

2.7.2 Training and re-estimation of HMM parameters

In order to train HMM λ, given that a new training observation sequence $\mathbf{y}$ has become available, the parameters $\theta = \{\mathbf{P}, \mathbf{B}, \pi\}$ in HMM model λ can be re-estimated.

As explained in [1], focusing on the state transition from time step k to time $k+1$, a quantity $\xi_k(i,j)$ can be calculated, where $\xi_k(i,j)$ represents the probability of being in state h_i at time step k and in state h_j at time step $k+1$, given the HMM λ and observation sequence $\mathbf{Y}$:

$$\xi_k(i,j) \;=\; \Pr\left(q_k = h_i,\; q_{k+1} = h_j | \mathbf{Y}, \lambda\right) \tag{2.38}$$

This quantity is calculated as follows:

$$\begin{aligned} \xi_k(i,j) &= \frac{\alpha_k(i)\; p_{ij}\; b_j(\mathbf{y}_{k+1})\; \beta_{k+1}(j)}{\Pr(\mathbf{Y}|\lambda)} \\ &= \frac{\alpha_k(i)\; p_{ij}\; b_j(\mathbf{y}_{k+1})\; \beta_{k+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_k(i)\; p_{ij}\; b_j(\mathbf{y}_{k+1})\; \beta_{k+1}(j)} \end{aligned} \tag{2.39}$$

In Equation (2.39), the numerator represents $\Pr\left(q_k = h_i,\ q_{k+1} = h_j,\ \mathbf{Y}|\ \lambda\right)$, and then dividing by $\Pr(\mathbf{Y}|\lambda)$ gives the desired probability in Equation (2.38).

Another quantity, $\gamma_k(i)$, is defined as the probability of being in state h_i at time step k, given the observation sequence $\mathbf{y}$ and the model λ:

$$\gamma_k(i) \ = \ \Pr(q_k = h_i|\mathbf{Y},\lambda)$$

where $\gamma_k(i)$ can be obtained by performing a summation of $\xi_k(i,j)$ over all possible transition states h_j that could occur at the next time step $k+1$:

$$\gamma_k(i) = \sum_{j=1}^{N} \xi_k(i,j) \tag{2.40}$$

If a summation of $\gamma_k(i)$ is performed over all applicable time indices k, the resulting quantity $\left(\sum_k \gamma_k(i)\right)$ quantity can be thought of as the expected (over time) number of times that state h_i is visited, or the number of transitions made from state h_i (excluding the last time step K in the summation). Furthermore, if a summation of $\xi_k(i,j)$ is performed from $k=1$ to the next to the last time step $k=K-1$, the resulting quantity can be thought of as the expected number of transitions from state h_i to state h_j.

Given an existing estimate for an HMM, at each time step k, the probabilities $\gamma_k(i)$ and $\xi_k(i,j)$ can be calculated for each state or state transition pair and then used to update the components of the HMM. The elements of the initial state probability π_i can be updated as follows:

$$\hat{\pi}_i = = \ \gamma_1(i) \tag{2.41}$$

which is the expected number of times in state h_i at time step $k=1$. The p_{ij} element of the transition probability matrix $\mathbf{P}$ can be obtained as the ratio of the expected number of transitions from state h_i to state h_j, to the expected number of transitions from state h_j:

$$\hat{p}_{ij} = \frac{\sum_{k=1}^{K-1} \xi_k(i,j)}{\sum_{k=1}^{K-1} \gamma_k(i)} \tag{2.42}$$

Updates to the elements in observation probability matrix **B** can be made as the ratio of the expected number of times in state h_j and also observing symbol $\mathbf{v}^{(m)}$, to the expected number of times in state h_j:

$$\hat{b}_j\left(\mathbf{v}^{(m)}\right) = \frac{\sum\limits_{k=1}^{K} \gamma_k(i)}{\sum\limits_{k=1}^{K} \gamma_k(j)} \quad \text{s.t. } \mathbf{y}_k = v^{(m)} \tag{2.43}$$

2.7.3 Continuous observation densities

Often, the observations **y** are not specifically restricted to a limited number of discrete values. Since observations are inherently continuous in nature, we use a mixture of Gaussian distributions to model the possible (continuous) observations [29]. With this modification to the types of observations, we now define M to be the number of Gaussian mixture components that we use to model each of the continuous-valued observation quantities. Recall that in the discrete observation case, we used M to represent the number of different types of discrete observation symbols $\mathbf{v}^{(m)}$ that could be emitted. Given M four-dimensional Gaussian distributions in the mixture distribution, we create the $N \times M$ matrix **C** of mixture coefficients c_{jm}, for $j = 1,\ldots,N$ possible states, and $m = 1,\ldots,M$ weighting coefficients corresponding to the Gaussian mixture components. Using the weighted Gaussian mixture distribution, given the system is in the jth state, the observation probability can be modeled as:

$$b_j(\mathbf{y}_k) = \sum_{m=1}^{M} c_{jm}\, \mathcal{N}\left[\mu_{jm}, \Sigma_{jm}\right] \tag{2.44}$$

where $\mathcal{N}[\mathbf{y}_k; \mu, \Sigma]$ is the Gaussian probability density function of $\mathbf{y}_k$ with mean vector μ and covariance matrix Σ. Note that the weighting coefficient c_{jm} is the probability that the mth individual Gaussian distribution was responsible for emitting the observed vector $\mathbf{y}_k$, such that:

$$\begin{aligned} \sum_{m=1}^{M} c_{jm} &= 1, \quad 1 \le j \le N \\ c_{jm} &\ge 0, \quad 1 \le j \le N,\ 1 \le m \le M \end{aligned} \tag{2.45}$$

and hence, the observation probabilities integrate to 1:

$$\int_{-\infty}^{+\infty} b_j(x)dx = 1, \;\; 1 \leq j \leq N$$

With continuous observation HMMs, when training or updating the HMM, the weighting coefficient c_{jm}, the mean vector μ_{jm}, and the covariance matrix Σ_{jm} can be re-estimated as:

$$\hat{c}_{jm} = \frac{\sum_{k=1}^{K} \gamma_k(j,m)}{\sum_{k=1}^{K}\sum_{j=1}^{M} \gamma_t(j,m)}$$

$$\hat{\mu}_{jm} = \frac{\sum_{k=1}^{K} \gamma_k(j,m) \cdot \mathbf{y}_k}{\sum_{k=1}^{K} \gamma_k(j,m)}$$

$$\hat{\Sigma}_{jm} = \frac{\sum_{k=1}^{K} \gamma_k(j,m) \cdot \left(\mathbf{y}_k - \mu_{jm}\right)\left(\mathbf{y}_k - \mu_{jm}\right)^K}{\sum_{k=1}^{K} \gamma_k(j,m)}$$

where $\mathbf{v}^K$ represents the transpose of $\mathbf{v}$, and with the continuous-valued, mixture model observation, $\gamma_t(j,m)$ is now the probability of being in hidden state h_j at time step t with observation $\mathbf{y}_k$ being emitted by the mth Gaussian mixture component. This quantity can be calculated as:

$$\gamma_k(j,m) = \left[\frac{\alpha_k(j)\beta_k(j)}{\sum_{j=1}^{N} \alpha_k(j)\beta_k(j)}\right]\left[\frac{c_{jm}\,\mathcal{N}\left(\mu_{jm},\Sigma_{jm}\right)}{\sum_{\ell=1}^{L} c_{j\ell}\,\mathcal{N}\left(\mu_{j\ell},\Sigma_{j\ell}\right)}\right]$$

Chapter 3

CRACK GROWTH DYNAMICS

In a sequential Bayesian estimation approach towards estimating and monitoring the crack growth state of a structure, we use a state dynamics model that can propagate the damage state from one time step to the next; therefore, in this chapter, we describe historical concepts related to fracture mechanics, crack propagation, and basic crack growth modeling.

3.1 Early crack growth theories

As discussed in the previous chapter, *all materials possess some degree of defects*, and when a material is loaded in fatigue, a minute material defect existing within the material could provide an initiation point for a small separation to occur within the material, possibly leading to the development of a crack. As shown in Figure 3.1, given a material is subjected to a load and a crack propagates within the material, as the area of the crack grows, there is a resulting reduction in the cross sectional area of the material available to carry the load. Therefore, there will be a reduction in the amount of applied load required to produce a given amount of deflection within the material.

Figure 3.2 shows crack growth data obtained on an example fatigue test specimen, where the measured crack length has been plotted with respect to the corresponding number of fatigue cycles experienced by the specimen. A least squares cubic model has been fitted to the crack length data.

Propagation of the crack and the rate of crack growth are related to the magnitude of the stress as well as the geometry of the crack. In the 1920's, Griffith [62] studied the fatigue characteristics of glass, a brittle material, and recognized that a crack will propagate if the total energy within the system is lowered by propagation of the crack. In the 1940's, Irwin [63] extended the study of crack propagation to include duc-

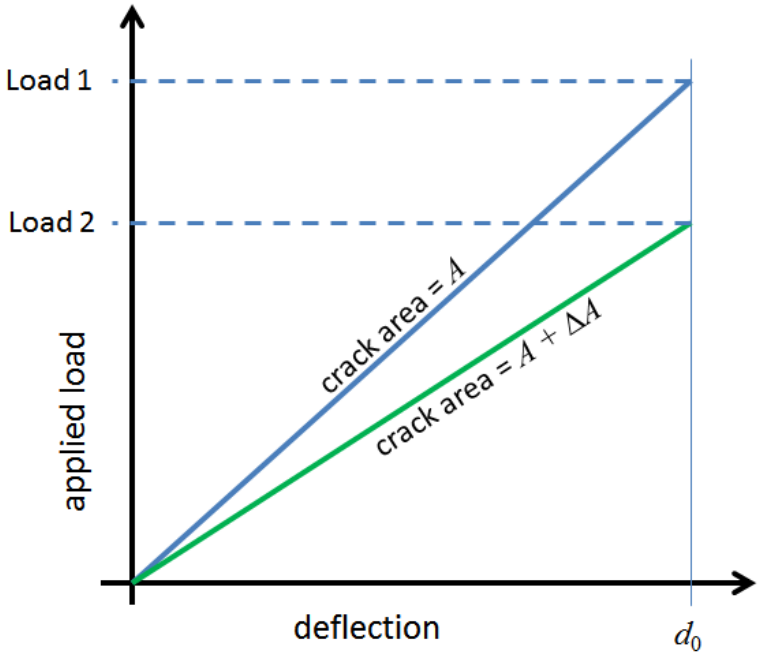


Figure 3.1: Reduction in the amount of applied load producing a given deflection when crack area increases from A to $A + \Delta A$.

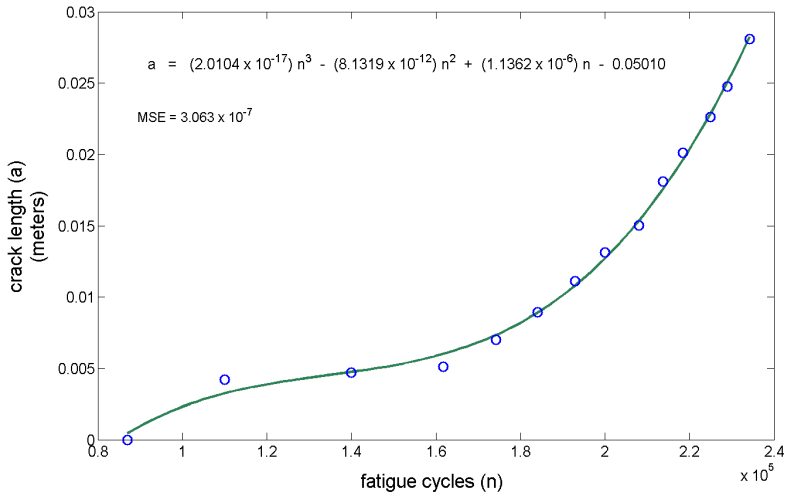


Figure 3.2: Increase in crack length with increase in number of fatigue cycles.

tile materials. The concept of strain energy release rate (G) was proposed, where this quantity represents the change in the strain energy U per change in crack area A:

$$G = \frac{dU}{dA}$$

Although the concept of strain energy was useful for studying conditions related to large, unstable crack growth, there were complications in using this approach, particularly for situations involving low rates of crack growth. In the 1950's, Irwin proposed the concept of a *stress intensity factor K* [64], where the stress intensity factor defines the magnitude of the local stresses around the crack tip. The stress intensity factor is dependent upon the loading conditions, the length of the crack, the geometry of the crack, and is of the general form [65]:

$$K = f(g)\ \sigma\sqrt{\pi a} \tag{3.1}$$

where σ represents an external, global stress that is being applied to the crack region, a is the crack length, and $f(g)$ is a function that depends upon the specimen and crack geometry. Crack propagation will occur if the value of the stress intensity factor K reaches a critical value K_c.

Figure 3.3 is an illustration of the region around a crack tip, where the local stresses in this region take the form:

$$\sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) + \cdots \tag{3.2}$$

where r and θ are the radius and angle from the crack of the tip to the local point of interest, and $f_{ij}(\theta)$ is a modification factor based upon θ.

As widely discussed in fracture mechanics textbooks (e.g., [65–68]), there are three basic loading modes leading to displacements and fracture in the cracked surface. These three modes are shown in Figure 3.4 and are as follows:

Mode I: Tensile opening mode.

Mode II: In-plane sliding mode.

Mode III: Tearing or anti-plane shear mode.

The value of the stress intensity factor K depends upon the material, the geometry of the flaw, the geometry of the specimen, loading conditions, and other influences.

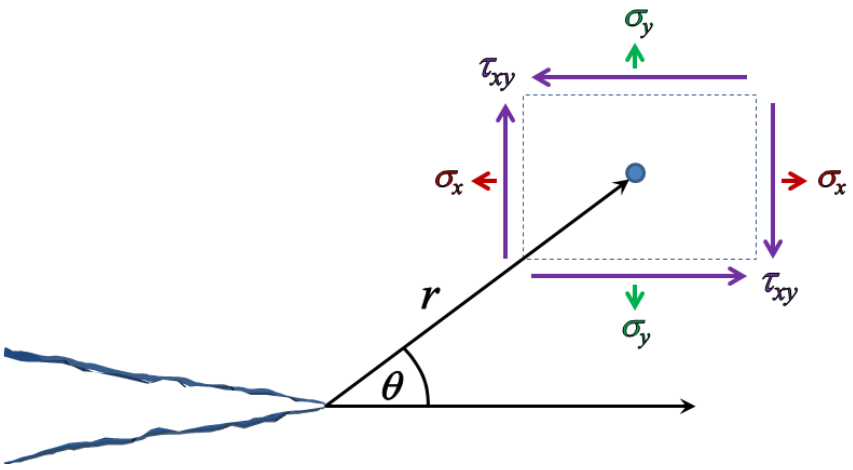


Figure 3.3: Localized stress in vicinity of crack tip.

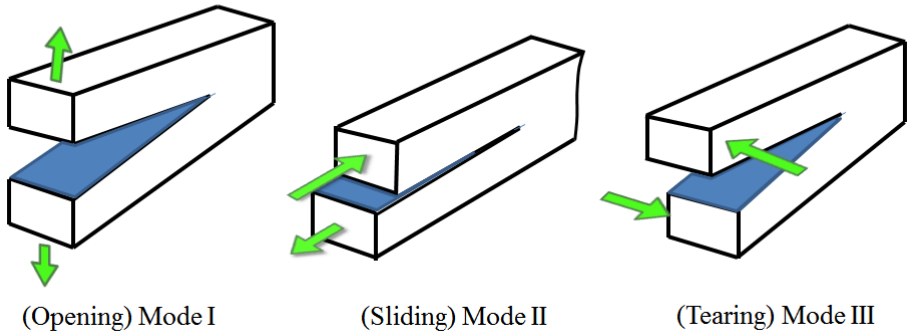


Figure 3.4: Crack fracture loading modes.

Many standard values for K are available in handbooks [69] for simple loading conditions, and the use of superposition can be used to estimate the stress intensity factor in multi-mode loading conditions. Alternatively, a value for K can be estimated using a variety of numerical methods, especially in cases where specimen geometry, crack geometry, or loading conditions are more complicated. For example, in previous work by Soni [70], estimates for K_{I} were obtained for a 2024-T3 aluminum cruciform specimen at different crack lengths using finite element methods and the displacement method described by Sanford [68]. The resulting estimates for K_{I} are shown in Figure 3.5. A cubic least squares model was created from the estimates such that a predicted estimate

for K_I could be obtained given a designated crack length value.

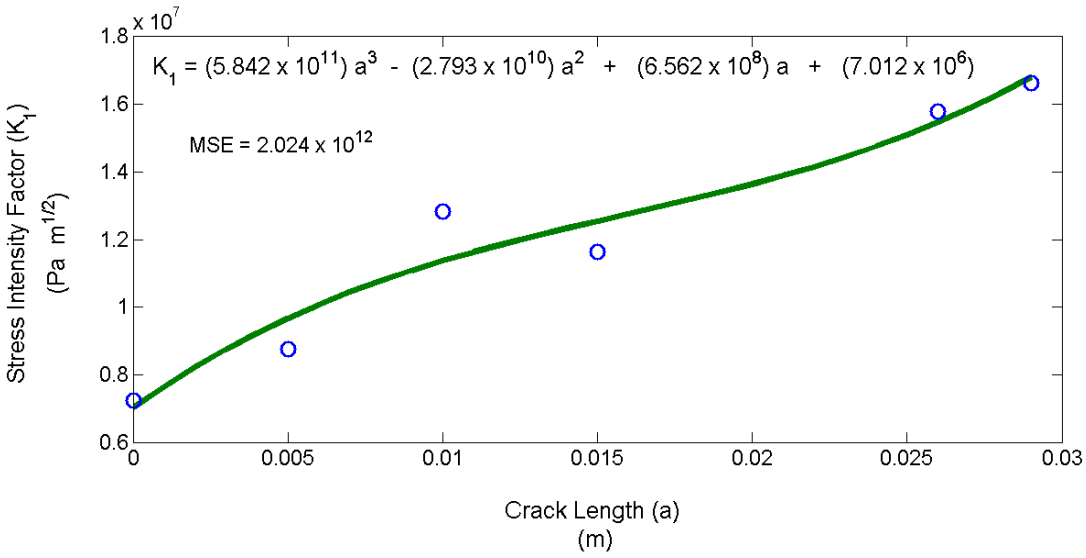


Figure 3.5: K_I estimates and least squares cubic model.

3.2 Paris' Law

As discussed in the 1960's by Paris and Erdogan [10], there were a number of early attempts to model crack propagation due to fatigue loading in materials. Some model approaches were largely theoretical, attempting to model crack growth as a function of material properties and loading characteristics. Other model approaches were primarily data driven, where mathematical models were created primarily by fitting a model to observed data. A key concept proposed by Paris and his associates related to the range of stress intensity factors, and how this range contributed to crack growth rate [9, 71]. This idea has become popularized into what is known as the Paris-Erdogan Law, or Paris' Law:

$$\frac{da}{dn} = C(\Delta K)^m \tag{3.3}$$

In this equation, the change in crack length a per loading cycle n is expressed as the difference between the stress intensity factor at the maximum and minimum fatigue loadings (ΔK), which is raised to a power m, and then multiplied by a factor C, where values for m and C are available in material handbooks. The logarithmic version of (3.3)

is as follows:

$$\log\left(\frac{da}{dn}\right) = \log C + m\log\Delta K \tag{3.4}$$

If a plot of $\log(da/dn)$ vs. $\log(\Delta K)$ is created for a specimen from crack initiation through overload and specimen failure, a sigmoidal curve is typically obtained, as shown in Figure 3.6. Note that the constant slope region of the curve can be used to obtain an estimate for m, and the intercept can be used to obtain an estimate for C.

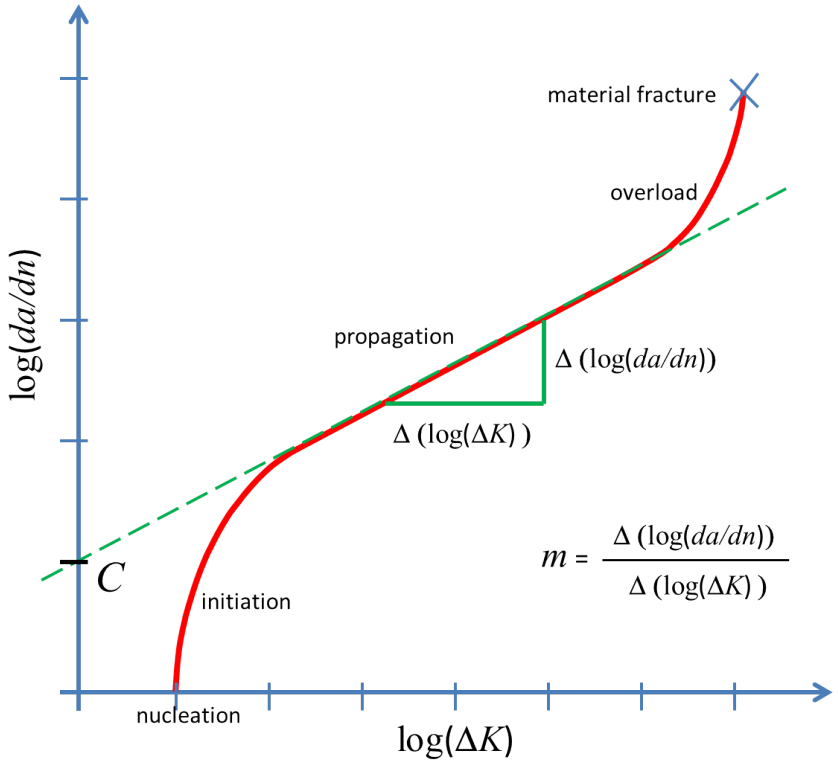


Figure 3.6: Typical $\log(da/dn)$ vs. $\log(K)$ curve, with intercept value 'C' and slope 'm' as used in the Paris equation.

From the origination of a crack until material failure, there typically are several stages related to crack development. Stages of crack development generally include *nucleation*, *initiation*, *propagation*, *overload* and *material fracture*. During crack *nucleation*, a microscopic material separation occurs at some point within the material, where this separation typically occurs along a free edge within the material, where the free edge could be located on an exterior surface of the component, within a small in-

ternal defect, or along an edge due to an intentionally-machined fastener hole or cutout. During crack *initiation*, the size of the material separation increases relatively quickly for a short period of time, but then the crack growth rate transitions to more of a stable, predictable, crack *propagation* stage, and it is in this stage that the Paris equation dynamics are applicable. If the size of the crack develops to the extent that the remaining undamaged material is unable to carry the load, the rate of crack growth accelerates, *overload* occurs, which eventually results in *material fracture*. These various stages are shown in Figure 3.6. The combination of *nucleation* and *initiation* is often referred to as *Stage I* of the crack growth process. The more predictable, crack *propagation* stage, which is shown as the linear region in Figure 3.6, is referred to as *Stage II*, and it is in this region that the Paris equation is applicable. Finally, the region containing crack *overload* and *material fracture* is commonly referred to as *Stage III*.

3.3 Simplifications and the stochastic nature of crack growth

In real-life scenarios, boundaries between different crack development stages might not be well defined and could vary greatly upon numerous geometric, loading, environmental, and other factors. If one is trying to use the Paris equation to predict crack propagation, values for m and C might not behave as well defined constants, even within the Stage II, *propagation* region shown in Figure 3.6. Simplifications to the Paris equation have been proposed, including the following model discussed by Yang, Manning, Rudd and Artley [72]:

$$\frac{da(t)}{dt} = Q\,[a(t)]^b \tag{3.5}$$

where $Q\,[a(t)]$ represents some general function of the crack length $a(t)$. Alternatively, if the number of fatigue cycles (n) is used as a measure of time, (3.5) becomes:

$$\frac{da(n)}{dn} = Q\,[a(n)]^b \tag{3.6}$$

Recognizing there can be inherent randomness within the crack growth process in a fatigue-loaded structure and that a pure deterministic approach towards modeling crack

growth might not always be optimal, Yang, Manning, Rudd and Artley [72] compared deterministic crack growth approaches (DCGA) with stochastic crack growth approaches (SCGA), and found that by including a random variable within the model, SCGA was more capable of modeling crack growth. For example, using Equation (3.5), a lognormal random variable $\chi(t))$ was included to account for variability that can occur in the crack growth rate:

$$\frac{da(t)}{dt} = \chi(t)\, Q\,[a(t)]^b, \tag{3.7}$$

Other stochastic crack growth studies have been discussed (e.g., [73–76]) using (3.7) or other generalized models. For example, a generalized function $g(*)$ has been proposed, where $g(*)$ can be based upon a number of crack growth related factors, such as the difference in stress intensities ΔK, the fatigue stress ratio R = (maximum stress)/(minimum stress), the crack length a, and other possible influences:

$$\frac{da(t)}{dt} = \chi(t)\, g(\Delta K, R, a, \ldots) \tag{3.8}$$

There has also been recognition that stochastic properties related to crack growth can change with increasing crack length (e.g., variability in the crack growth rate has been observed to decrease with crack size), and therefore, the random characteristics of crack growth have been studied as a function of the crack length:

$$\frac{da(t)}{dt} = \chi\,[(a(t)]\; g(\Delta K, R, a,\; \ldots)$$

Alternatively, crack growth has also been studied using a piecewise, fatigue-interval manner to account for changes with increasing crack length at different fatigue crack intervals j. In this case, (3.7) has been modified to account for different fatigue crack intervals j during the life of a component [73]:

$$\frac{da(t)}{dt} = \chi_j(t)\, Q_j\,[a(t)]_j^b \tag{3.9}$$

Assuming no special crack mending or repair procedures have been used to reduce the size of the crack, the actual crack growth rate cannot be negative, and the lognormal random variable in the expressions above can model the positive-valued randomness that can occur within the crack growth rate. However, when observing and tracking crack growth, random noise and variability can occur in both positive and negative directions due to observation noise. In Chapter 5, in one of our implementations of crack growth monitoring, we also included a normal random variable to allow for both negative as well as positive random fluctuations that can occur when making crack length observations.

Finally, as discussed in the previous chapter (see Section 2.3), Markov chain Monte Carlo (MCMC) techniques can be useful to refine parameters of interest. In Chapter 5, given observed data and corresponding constraints related to crack length and stress intensity factors, and given estimated target values for the Paris constants C and m, MCMC can be used to combine this information to obtain refined values of the constants that approach desired target values and are also consistent with the observed data.

Chapter 4

INTERACTIVE MULTIPLE MODEL FOR MULTIPLE LOADING MODES

Within our overall sequential Bayesian approach, the initial prediction of the crack length distribution from the previous time step to the next time step is determined by the state dynamics model. As described in the previous chapter, we employ a Paris law based model to provide this prediction of damage at the next time step. In many, if not most dynamic structures, the loading characteristics upon the structure do not remain constant but can vary at multiple levels. The interacting multiple model (IMM) is an approach that can address multiple loading conditions.

4.1 Interacting Multiple Model Approach

The IMM is often used when tracking multiple objects [31, 32] for a variety of military (e.g., radar, target tracking) and commercial (e.g., air traffic control, ground vehicle traffic) applications. In IMM, multiple dynamic models are used to represent different modes. In our work, we apply the IMM to address situations with more than one fatigue loading condition. The different loading conditions are represented by different state dynamics models. For each fatigue loading time step, the IMM is used to assess crack length observations and adjust the probabilities of each loading condition. The overall crack length state estimate is based upon a weighted summation of the individual loading condition state estimates.

The IMM consists of three main stages: interaction/mixing stage, filtering stage, and combination stage. In the interaction/mixing stage, information from the different modes is integrated together using assumed mode transition probabilities, as well as the mode densities and mode probabilities established during the previous time step. The filtering stage consists of standard stochastic filtering, except multiple filters are employed so that each of the modes are processed separately. In the combination step,

the mode probabilities are used to combine the information of the different modes to create an overall state estimate.

The IMM is discussed in [33] and used for target tracking, with emphasis on Kalman filter and extended Kalman filter for the filtering stage. We implement the IMM particle filter approach, described by Boers and Driessen in [77], as we use particle filtering in our estimation. Furthermore, our state dynamics model employs crack rates based on Paris' law, which is given by

$$\dot{a} = \frac{da}{dn} = C\,\Delta K_{\text{eff}}^{m}$$

where ΔK_{eff} represents the effective stress intensity factor range, and C and m are the Paris law constants. In our IMM implementation, the different state dynamics modes are defined by different loading conditions, which can be represented by different values of ΔK_{eff}.

Assuming M different state dynamic modes within an IMM, the probabilities of transitioning from mode ℓ to mode j are estimated and stored in a mode transition matrix with transition probability elements $p_{\ell,j}$. Furthermore, we apply M separate particle filters, one for each mode. In our implementation, the particle filters are resampled at the end of each time step and all particles are assigned equal weights. Therefore, for the ℓth mode, $\ell = 1,\ldots,M$, the resampled particle filters represent the posterior state density $p_{\ell}(x_{\ell,k-1}|y_{k-1})$, where $x_{\ell,k-1} = a_{k-1}$ is the state value at mode ℓ and time step $k-1$ and y_{k-1} is the observation at time step $k-1$. In addition, the mode probabilities $\mu_{\ell,k-1}$, which were calculated at the end of the previous time step $k-1$, are also available. For the initial time step, it is assumed the initial value of the crack length state x_0 is known, and all particles are initialized to this value with equal weighting. It is also assumed that the initial mode probabilities $\mu_{\ell,0}$ are known.

To begin the *interaction/mixing* stage, the mode mixing probabilities are calculated for all modes. The mode mixing probabilities, denoted as $\mu_{\ell|j,k-1|k-1}$, are the

probabilities of moving time step $k-1$ posterior density information of mode ℓ into mode j during this mixing step. Given the posterior state density $p_\ell(x_{\ell,k-1}|y_{k-1})$ for the ℓth mode, the mixing probabilities can be calculated as

$$m_{\ell|j,k-1|k-1} = \frac{1}{c_j}\, p_{\ell,j}\, \mu_{\ell,k-1} \tag{4.1}$$

where c_j is a probability normalization factor based upon the mode probabilities $\mu_{\ell,k-1}$ and mode transition probabilities $p_{\ell,j}$ as

$$c_j = \sum_{\ell=1}^{M} p_{\ell,j}\, \mu_{\ell,k-1}\,. \tag{4.2}$$

Then, for each mode j, using posterior density $p_j(x_{j,k-1}|y_{k-1})$, an initial mixed probability density for mode j can be calculated. This is performed as a summation of the posterior densities for each mode from the previous time step $k-1$, weighted by the corresponding mixing probabilities from (4.1):

$$\hat{m}_{j,k-1}(x_{j,k-1}|y_{k-1}) = \sum_{\ell=1}^{M} p_\ell(x_{\ell,k-1}|y_{k-1})\, m_{\ell|j,k-1|k-1}\,. \tag{4.3}$$

After the interaction/mixing stage, the *filtering stage* is performed. As the initial step in the filtering stage, the initial mixed, densities for each mode are propagated according to their respective state dynamics model. Using Paris' law corresponding to mode j, the predicted crack growth rate $\dot{a}$ can be obtained as

$$\dot{a}_k = \frac{\Delta a_k}{\Delta n_k} = \chi\, C(\Delta K_{\text{eff},j})^m \tag{4.4}$$

where Δn_k is the number of fatigue cycles occurring between time step $k-1$ and k, Δa_k is the corresponding increase in crack length within this time period, $\Delta K_{\text{eff},j}$ represents the effective stress intensity factor range for the jth mode loading condition, C and m are the Paris law constants, and χ is a lognormal random variable that is used to account for modeling errors in the dynamics model. Using $x^{(i)}_{\ell,k-1|k-1}$ to represent the ith particle of the mixed particle set of mode ℓ at time step $k-1$, the particles are propagated as

follows:

$$x^{(i)}_{\ell,k|k-1} = x^{(i)}_{\ell,k-1|k-1} + \dot{a}_k \, \Delta n_k \; + \; v_k \tag{4.5}$$

where $x^{(i)}_{\ell,k|k-1}$ represents the *prior* density at time step k for mode ℓ, $\dot{a}_k$ is from Equation (4.4), and v_k is a normal random variable that is added to account for modeling errors.

Our observation model has been previously discussed in Chapter 2. The observation is a sequence of matching pursuit decomposition (MPD) feature vectors that is used as the input into a set of trained hidden Markov models (HMMs), where each HMM is trained for a different crack length. For each of the trained HMMs, the forward-backward algorithm (Section 2.7.1.1) is used to obtain a probability measure of the feature vector. Hence, the observation step provides a probability distribution of the different crack lengths used to train the HMMs. In our implementation, we obtain multiple individual observations at each time step and the resulting distribution is based upon multiple observations. Given a distribution of discrete crack lengths, the particle weights $w^{(i)}_{j,k}$ for mode j at time step k can be obtained by rounding the value of the particle to the nearest HMM training crack length, and then using the corresponding probability of the HMM training crack length to weigh the particle.

After weighing the particles corresponding to mode j, the state estimate of mode j can be obtained as a weighted summation of N_p particles

$$\hat{x}_{j,k|k} = \sum_{i=1}^{N_p} w^{(i)}_{j,k} \, x^{(i)}_{j,k|k-1} \, . \tag{4.6}$$

For the ith particle, a predicted crack length can be calculated. The value of the crack length stored in the particle can be rounded to one of the D crack lengths used to train the HMMs. Using the corresponding trained HMM, an MPD observation can be created using the procedure to be discussed in Section 5.2.1. For each of the D trained HMMs, the forward-backward algorithm is used to create a probability measure that the MPD vector was generated by the HMM. This also provides a probability measure related to

assigning the value of the crack length used to train the HMM as the predicted output. After this is performed for all D HMMs and their corresponding training crack lengths, a distribution of predicted crack length outputs is obtained for the particle. Since the D possible crack lengths are known and D HMMs have been trained offline for these crack lengths, the distributions of predicted crack length outputs can be created off-line for each of the D training crack length values, and numerous simulations can be performed to improve the statistical accuracy. In our work, 2000 predicted MPD feature vectors were generated and analyzed off-line to generate each predicted crack length distribution. After creating the predicted output distributions, the means of these distributions can also be obtained and stored. For each of the D HMM training crack lengths, we use the mean values of the corresponding predicted crack length distributions to represent the predicted crack length output. Since the predicted crack length distributions and means of the distributions can be calculated and stored offline, this significantly reduces the computational burden during runtime.

Given mode ℓ and its corresponding set of particles, and given the predicted particle observation $y_{\ell,k}^{(i)}$ for each particle i at time step k, the predicted mean output $\bar{y}_{\ell,k}$ can be calculated for mode ℓ as a weighted summation of the particle predicted outputs:

$$\bar{y}_{\ell,k} = \sum_{i=1}^{N_p} w_{\ell,k}^{(i)} \, y_{\ell,k}^{(i)}. \tag{4.7}$$

Using the predicted output of each particle $y_{\ell,k}^{(i)}$ as well as the predicted mean output $\bar{y}_{\ell,k}$, the mode ℓ residual covariance can be calculated as

$$\hat{S}_{\ell,k} = \sum_{i=1}^{N_p} \left(y_{\ell,k}^{(i)} - \bar{y}_{\ell,k} \right)^2 \tag{4.8}$$

The difference between the actual observation y_k and the predicted particle observations $y_{\ell,k}^{(i)}$ at time step k, called particle innovations, can be obtained as

$$r_{\ell,k}^{(i)} = y_k - y_{\ell,k}^{(i)}. \tag{4.9}$$

If we assume that the particle innovations $r_{\ell,k}^{(i)}$ have zero-mean, Gaussian distributions, given by $r_{\ell,k}^{(i)} \sim \mathcal{N}(0,\ \hat{S}_{\ell,k})$, then we can calculate the likelihoods of the particle innovations. Recognizing that the probabilities can be very small numbers, and that the operations that involve these values might impinge upon computer precision limits, in our implementation, we calculate log values of the particle innovation likelihoods. Furthermore, since the log likelihoods are relative values, we perform other value checks and collective manipulations, as necessary, to ensure values are maintained within computer precision limits. For the ℓth mode, since we employ relative log likelihood values of the particle innovations, we can calculate a relative log likelihood of the mode using a summation of the particle log likelihoods. The mode likelihood $L_{\ell,k}$ can then be obtained using the exponential of the relative log likelihood and normalizing this value.

As the final step of the filtering stage, given the mode likelihoods $L_{\ell,k}$, the mode probabilities $\mu_{\ell,k}$ can be obtained as

$$\mu_{\ell,k} = \frac{1}{c_\ell}\, L_{\ell,k}\, c_\ell \tag{4.10}$$

where c_ℓ was calculated in (4.2).

In the *combination stage*, the state estimates of the individual modes are combined according to their respective mode probabilities to create an overall state estimate at time step k, given by

$$\hat{x}_k = \sum_{\ell=1}^{M} \hat{x}_{\ell,k|k}\ \mu_{\ell,k}\,, \tag{4.11}$$

where $\hat{x}_{\ell,k|k}$ was calculated in (4.6) and $\mu_{\ell,k}$ was calculated in (4.10).

4.2 IMM Implementation for Multiple Loading Conditions

As explained above, we apply the IMM methodology to show how it can be used to address situations where more than one loading condition is present. The different loading conditions are represented by different state dynamics models, and the IMM methodology is used to assess crack length observations such that the probabilities of

each loading condition are adjusted, thereby improving the ability to track changes in damage condition.

Aluminum compact tension (CT) specimens are used in our implementations of IMM towards damage state monitoring. An example CT specimen in shown in Figure 4.1. The CT specimens consist of a machined block of 2024-T3 aluminum. Referring to Figure 4.1, in the center of the specimen off one edge, there is a wide machined notch between the two loading attachment holes. The wider notch transitions into a thinner, shorter, electrochemical machined (ECM) notch. As the figure shows, after the specimen is fatigue loaded, a fatigue crack develops from the ECM notch and propagates into the specimen. Specimens are loaded in fatigue tension with several different minimum and maximum load combinations.

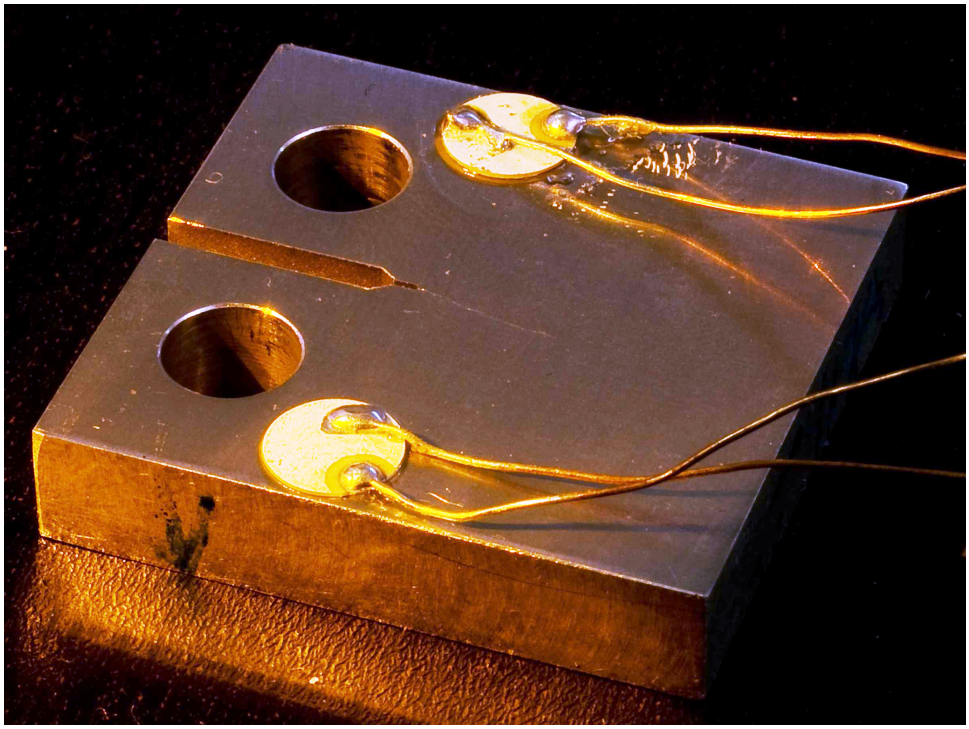


Figure 4.1: Compact tension specimen with actuator and receiver sensors.

In the active sensing approach used to observe the damage, an actuator piezoelectric sensor is positioned on one side of the notch to transmit the signal, and a receiver

sensor is positioned on the other side of the notch to receive the transmitted signal. A 135 kHz tone burst signal is used as the active sensing waveform. The tone burst waveform is an enveloped sinusoid, with amplitude modulation $A_{bu}(t)$, and can be generated using

$$s_{bu}(t) = A_{bu}(t)\cos(2\pi f_{bu}t).$$

Figure 4.2 shows an example of a transmitted tone burst signal used in this study. Figure 4.3 shows the received version of the signal, and Figure 4.4 shows the MPD time-frequency representation (TFR) of the received signal; the MPD-TFR is created as a summation of individual MPD feature vector components.

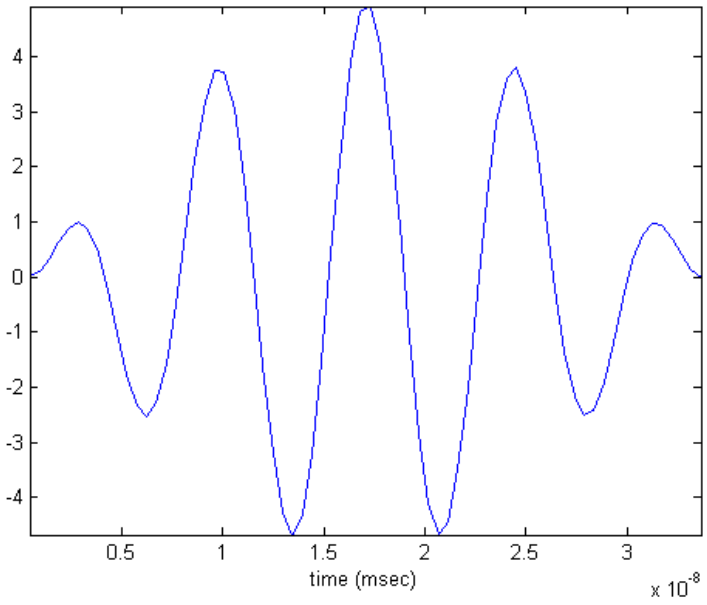


Figure 4.2: Transmitted tone burst signal.

The tone burst waveform possesses several characteristics that make it a popular choice for damage detection within structures. It possesses high localization in both the time and frequency domains. Whether using a time domain and/or a frequency domain representation of the received signal, damage detection methods typically depend upon the ability to recognize differences in the received waveform corresponding to differ-

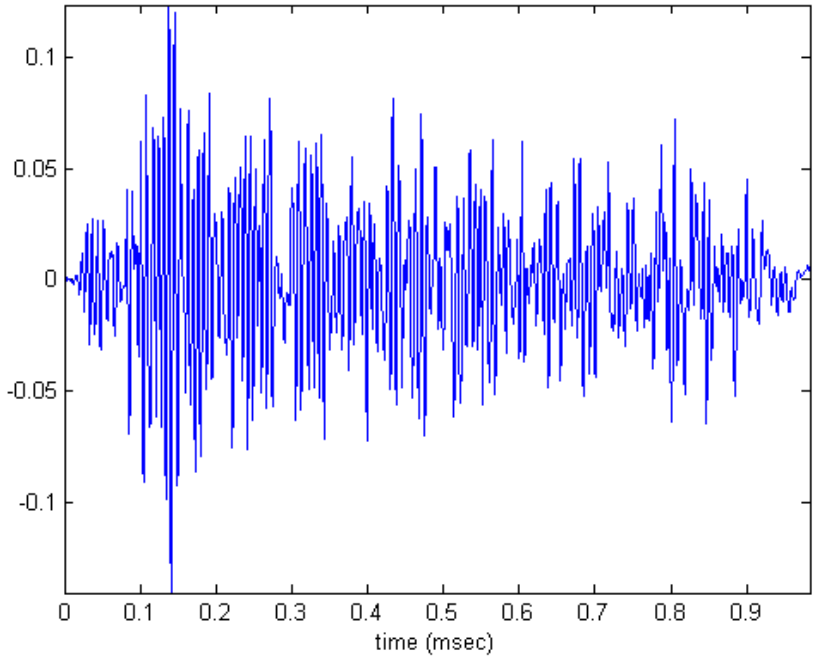


Figure 4.3: Received tone burst signal.

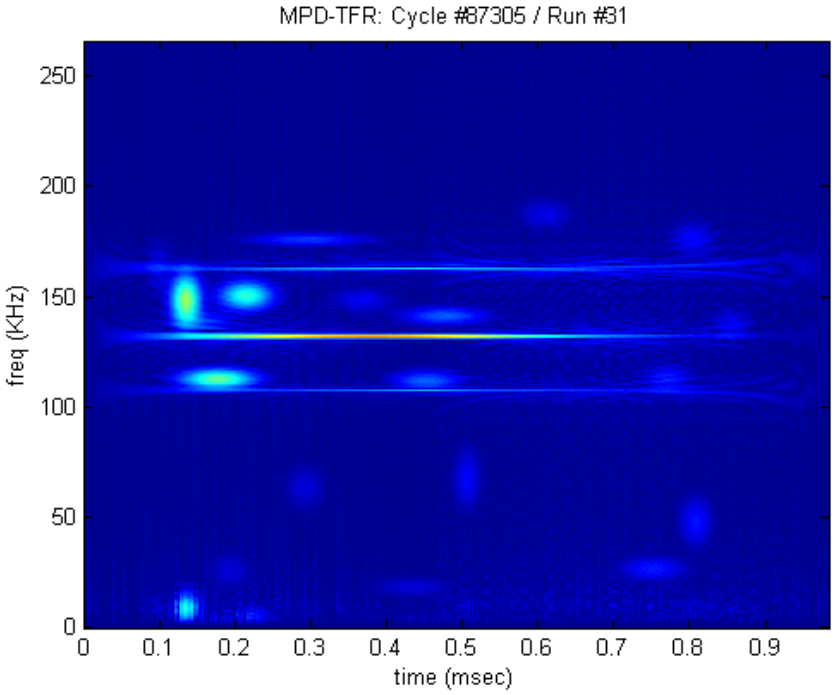


Figure 4.4: MPD-TFR of received tone burst signal.

ences in the condition of a structure. The high localization of the tone burst waveform improves the ability to recognize it in a received signal and to detect changes in the received waveform corresponding to changes in the condition of the structure.

As discussed in Chapter 2, in our work, we employ an MPD-HMM observation model. Unlike many other signal-based methods used in damage detection, our observation model does not focus upon the characteristics of only one or two visually-identifiable peaks within a time domain or frequency domain representation of the received signal. Therefore, the high localization characteristics associated with the tone burst waveform is less critical as compared to other damage detection approaches In Chapter 5, we explore the use of several different types of transmission waveforms for classification of crack lengths within an overall sequential Bayesian crack growth monitoring procedure.

In our example, demonstrations of using the IMM to assess and weigh the different state dynamic loading modes, crack length observations are obtained at approximately 30 different time steps, with approximately 30 different crack lengths, within a range of 6-17 mm crack length. Although the high localization of the tone burst waveform can be advantageous for basic binary damage detection, the limited amount of signal information within this particular waveform might constrain its capabilities with respect to crack length resolution and multiple crack length, classification accuracy.

Figure 4.5 shows a plot of the loading and crack growth characteristics of compact tension specimen CT419a, where 32 different crack length measurements were obtained during the fatigue loading of this specimen. Assuming we have training data available, where the number of fatigue cycles Δn_k and the crack increase Δa_k between time steps k and $k+1$ are known, estimates of the crack growth rate $\dot{a}_k$ can be obtained, and assuming values for the Paris constants $C = 5 \times 10^{-11}$ and $m = 4.07$, then stepwise values for ΔK_{eff} can be obtained from Paris' law. These are plotted in Figure 4.6.

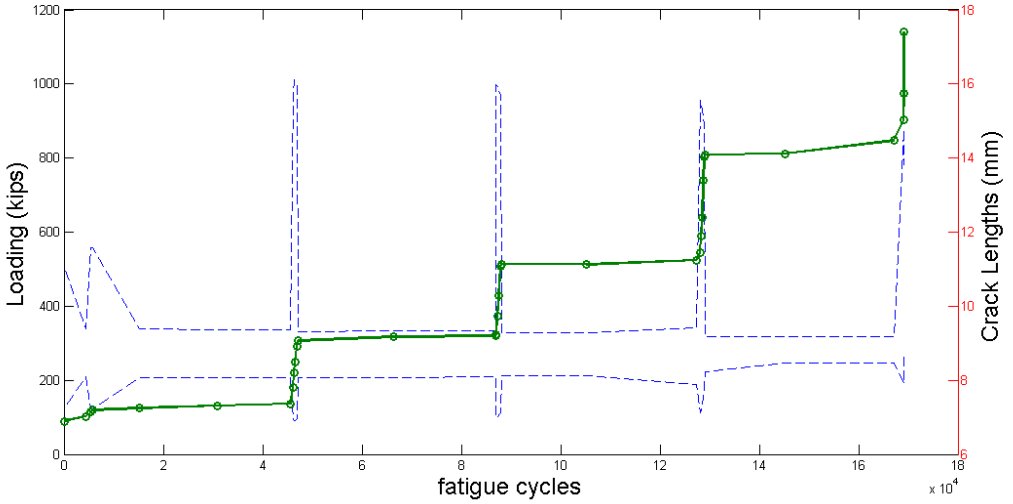


Figure 4.5: Maximum and minimum loading conditions and resulting crack growth characteristics for specimen CT419a.

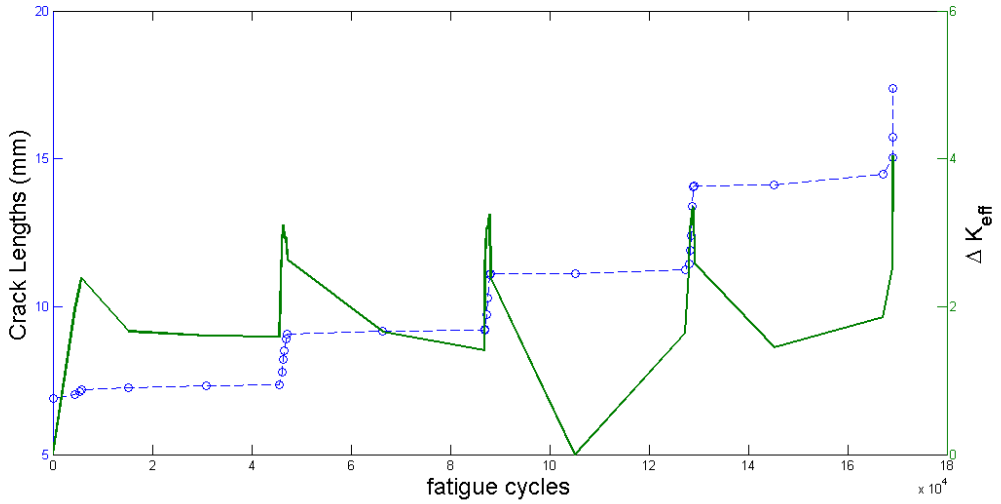


Figure 4.6: Calculated stepwise ΔK_{eff} of compact tension specimen CT419a.

As mentioned above, a 135 kHz tone burst waveform was used as the actuator signal, and the sampling frequency was 2 MHz. The received signals were downsampled by a factor of 2, and then the signals were high pass filtered to attenuate low frequency interference and low pass filtered to attenuate extraneous high frequency noise. After preprocessing, the MPD algorithm was performed to extract 4-element feature vectors. After several preliminary HMM training studies using MPD feature vectors, it

was decided that $L = 23$ was an optimal number of MPD feature vectors for training the HMMs.

The MPD feature vectors were used to train the HMMs as follows. Due to the high number of different crack lengths, a relatively high number of hidden states (about 7) was assumed to exist within the time-frequency plane, and 8 Gaussian mixture components were used to model the continuous-valued, 4-element, MPD feature vector emissions. Given one of the 32 different crack lengths, for each of the approximately 110 training signals representing that crack length, the first 23 extracted MPD feature vectors were arranged according to their time-shift parameter. The resulting time-based MPD sequence was used in training the HMM for that crack length. This was performed for all training signals corresponding to the crack length, and it was repeated 12 times for the entire collection of MPD observations. For each of these 12 training *epochs*, the order in which the different observations were presented to the HMM was randomized prior to performing the epoch.

After HMMs were trained for all 32 crack lengths, a confusion matrix was created to assess the classification accuracy of the trained HMMs. For each crack length, using the maximum likelihood classification approach, approximately 110 testing runs (corresponding to that crack length) were inputted into each of the 32 trained HMMs, and a maximum likelihood crack length classification was obtained for each of the testing runs. The decimal percentage classification results were then calculated and represented as a row of the confusion matrix corresponding to the crack length of interest. This was performed for all 32 different crack lengths. Therefore, the confusion matrix consisted of 32 rows, representing the 32 different crack lengths, and 32 columns, representing the 32 different HMMs that had been trained for each of the crack lengths. Note that for a situation in which there was high classification accuracy, the confusion matrix should exhibit a high concentration of crack length classifications along the diagonal. A two-dimensional image plot of the confusion matrix is shown in Figure 4.7. As this

plot shows, there is heavy concentration along the diagonal, indicating relatively high classification accuracy. However, with over 30 different crack lengths within a 6 to 17 mm crack length range, the image plot also reveals there is some level of error away from the diagonal.

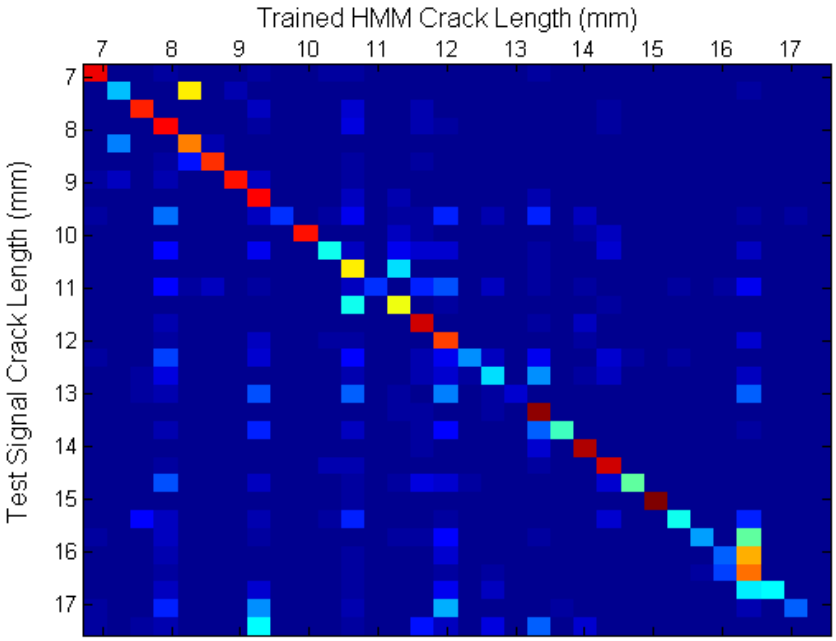


Figure 4.7: HMM Confusion matrix for compact tension specimen CT419a.

In our IMM implementation for crack growth monitoring, the different loading modes were modeled using different Paris' Law dynamics models, where the different models were created using different values for ΔK_{eff}. Referring to Figures 4.5-4.6, specific values for ΔK_{eff} and the mode transition probabilities were estimated using the values shown in these figures. Note that values for ΔK_{eff} do not necessarily exist at well-defined, discrete levels. Therefore, when selecting values for ΔK_{eff} to represent different modes of loading, averaged values were used. Furthermore, in the examples discussed below, we varied the number of assumed loading conditions to observe how this affected the damage tracking results.

Figure 4.8 shows the IMM crack tracking results when two different loading

modes were assumed. In this example, $\Delta K_{\text{eff}} = 1.3$ for low loading, $\Delta K_{\text{eff}} = 3.2$ for high loading, and the mode transition probability matrix is shown in Table 4.1.

	Mode 1	Mode 2
Mode 1	0.7368	0.2632
Mode 2	0.3333	0.6667

Table 4.1: Mode transition probabilities for 2 modes.

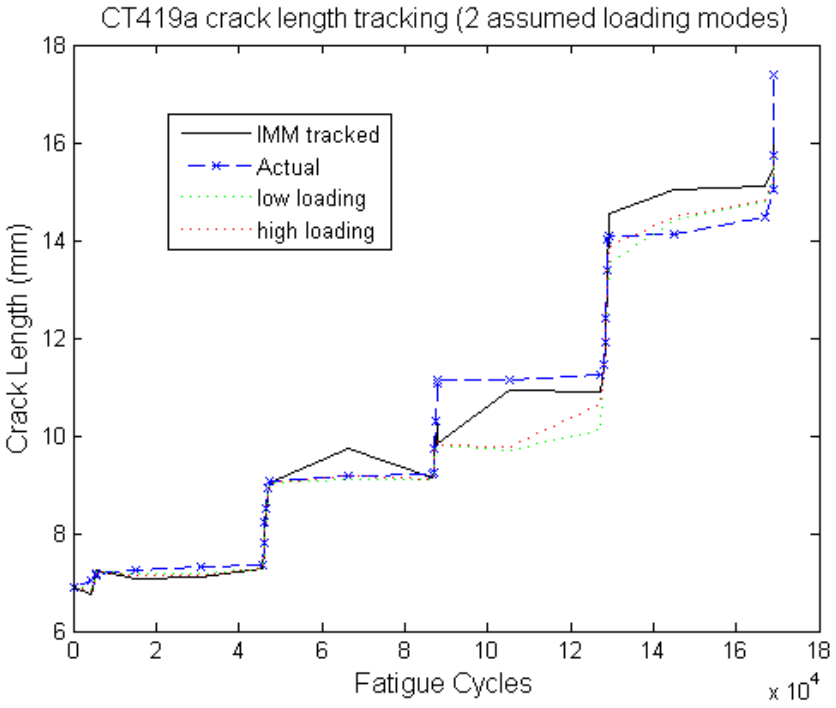


Figure 4.8: IMM results assuming 2 loading modes.

Figure 4.9 shows the mode probabilities during the fatigue cycling. Comparing this figure to Figure 4.5, the IMM procedure detects a significant increase in the probability of the high loading condition at the time or soon after the time the high loading condition has actually occurred.

Figure 4.10 shows the IMM crack tracking results when three different loading modes were assumed. In this example, $\Delta K_{\text{eff}} = 1.3$ for low loading, $\Delta K_{\text{eff}} = 2.5$ for medium to high loading, $\Delta K_{\text{eff}} = 3.5$ for high loading, and the mode transition probability matrix is shown in Table 4.2.

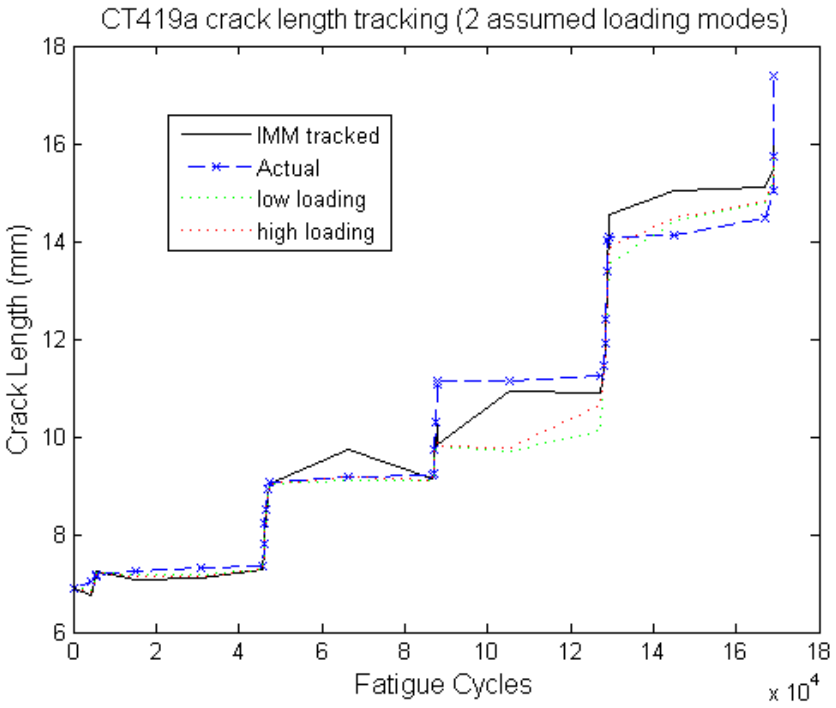


Figure 4.9: IMM mode probabilities assuming 2 loading modes.

	Mode 1	Mode 2	Mode 3
Mode 1	0.500	0.400	0.100
Mode 2	0.500	0.125	0.375
Mode 3	0.100	0.300	0.600

Table 4.2: Mode Transition Probabilities, 3 modes.

Figure 4.11 shows the mode probabilities during the fatigue cycling. Comparing this figure to Figure 4.5, the IMM procedure detects a significant increase in the probability of the high loading condition at the time or soon after the time the high loading condition has actually occurred.

Figure 4.12 shows the IMM crack tracking results when four different loading modes were assumed. In this example, $\Delta K_{\text{eff}} = 1.3$ for low loading, $\Delta K_{\text{eff}} = 1.7$ for low-medium loading, $\Delta K_{\text{eff}} = 2.5$ for medium-high loading, $\Delta K_{\text{eff}} = 3.5$ for high loading, and the mode transition probability matrix is shown in Table 4.3.

Figure 4.13 shows the mode probabilities during the fatigue cycling. Comparing

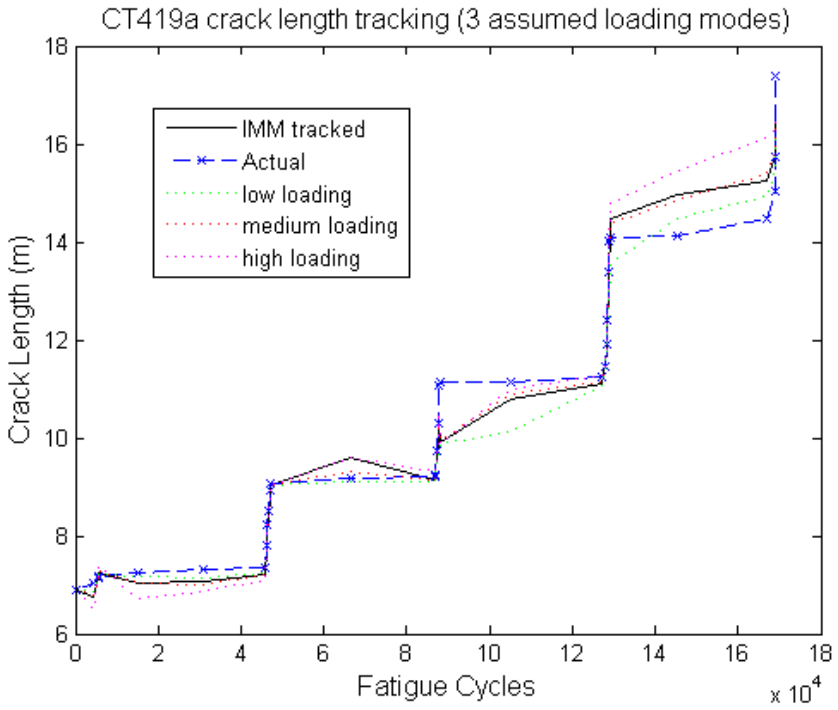


Figure 4.10: IMM results assuming 3 loading modes.

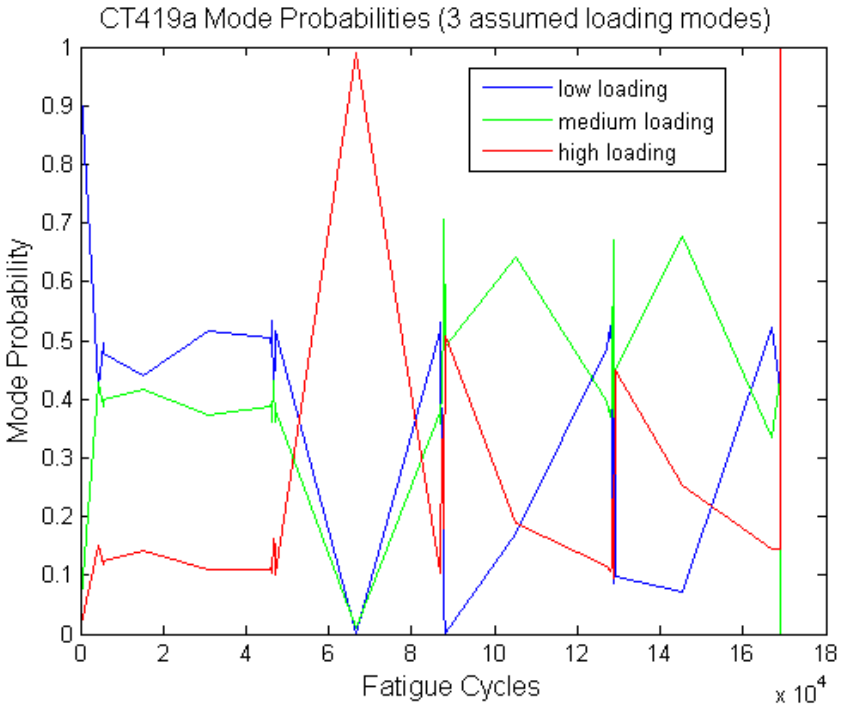


Figure 4.11: IMM mode probabilities assuming 3 loading modes.

	Mode 1	Mode 2	Mode 3	Mode 4
Mode 1	0.4444	0.2222	0.2222	0.1111
Mode 2	0.2000	0.3000	0.4000	0.1000
Mode 3	0.2500	0.2500	0.1667	0.3333
Mode 4	0.0909	0.0909	0.2727	0.5455

Table 4.3: Mode transition probabilities for 4 modes.

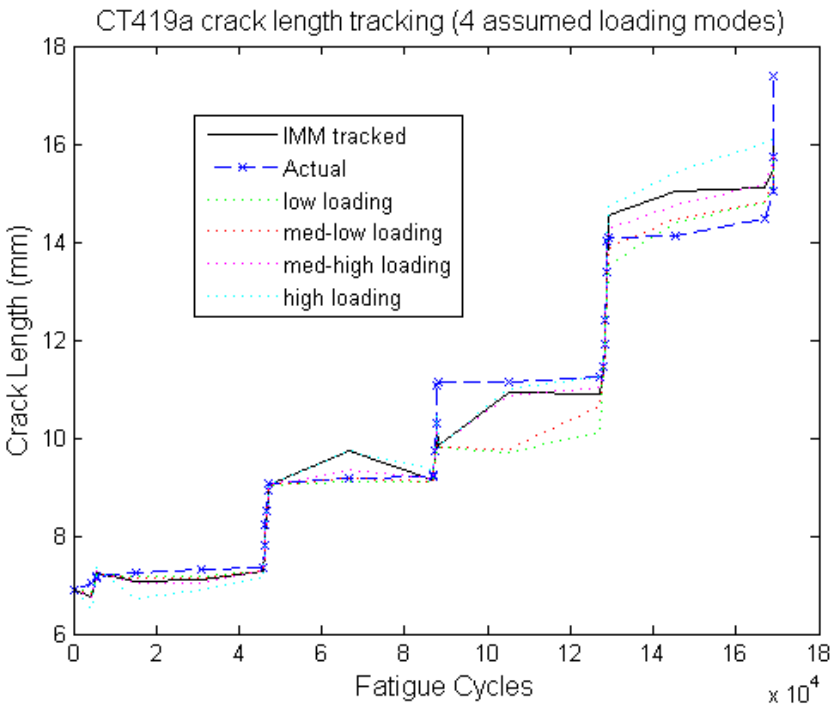


Figure 4.12: IMM results assuming 4 loading modes.

this figure to Figure 4.5, the IMM procedure detects a significant increase in the probability of the high loading condition at the time or soon after the time the high loading condition has actually occurred.

In this chapter, we explored the use of IMM for modeling structural loading situations in which more than one type of loading condition exists. As the examples show, IMM can assess changes in the crack length growth rate and adjust the probabilities of the different loading modes. These probabilities are then used to weigh the individual mode state estimates to obtain the overall crack length state estimate.

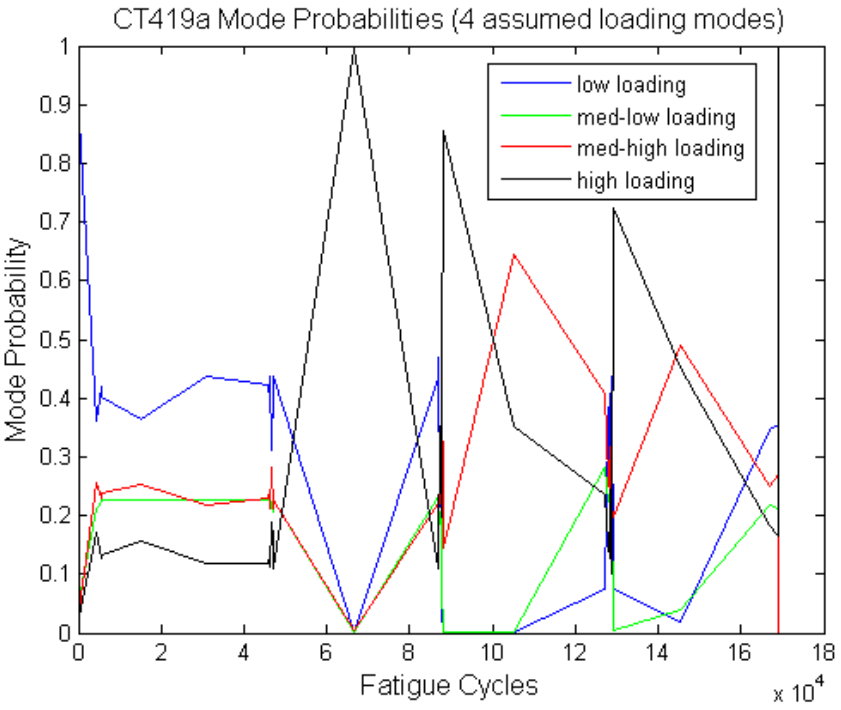


Figure 4.13: IMM mode probabilities assuming 4 loading modes.

Chapter 5

ADAPTIVE WAVEFORM SELECTION FOR CRACK GROWTH ESTIMATION

In the sequential Bayesian approach used to monitor crack growth, in addition to the possibililty of multiple modes for damage growth (discussed in the previous chapter), we can also intentionally apply multiple modes for the damage state observations. Even though the matching pursuit decomposition based hidden Markov model (MPD-HMM), that we use for the observation model, has exceptional capability for recognizing different damage conditions, the ability to resolve and observe different damage conditions can vary depending upon specific observation parameters. As a result, within our observation model, we design the transmission sensing waveform in order to optimize resolution capability given a particular estimated level of damage. Allowing for different options of transmission sensing waveforms leads to the use of multiple observation modes n the sequential Bayesian approach. With the different waveform modes, the observation mode corresponding to the optimal resolution capability for that level of damage can then be selected. In selecting the optimal transmission waveform, we employ a predicted mean-squared error (MSE) cost function [34]. The motivation for the MSE criterion relates to differences that can occur in the modeling capability of HMM models that have been trained with different waveforms, given a particular crack length damage condition.

5.1 Crack Length Observation Models

5.1.1 Observations Models Using HMMs

Within this work, we use HMMs into our crack growth monitoring method to determine the probability of the data given an HMM and the corresponding HMM parameters that maximize the probability. Given training data obtained for different dam-

age conditions, as well as different observation conditions used to obtain the data, the HMM parameters for each damage condition have to be adjusted to maximize the likelihood of the training observations. This results in training the different HMMs for the different damage and observation conditions. Our observations $\mathbf{Y}_k$ consist of L MPD feature vectors, with 4 feature elements per vector. The feature vectors are arranged according to their time shift parameter in order to represent a time-based feature vector in the time-frequency plane.

After training D different HMMs, represented by $\{\lambda_1, \ldots, \lambda_D\}$ for different conditions, and considering a known observation mode, our problem is to find the damage condition of a new test observation $\mathbf{Y}_k^{\text{test}}$ using that observation mode but obtained from a material with unknown damage condition. This damage condition corresponds to the HMM $\hat{\lambda}$ obtained using maximum likelihood estimation as

$$\hat{\lambda} = \underset{\lambda}{\text{argmax}} \; p\left(\mathbf{Y}_k^{\text{test}} | \lambda\right) \tag{5.1}$$

where $p\left(\mathbf{Y}_k^{\text{test}} | \lambda\right)$ can be calculated efficiently using the forward-backward algorithm.

5.1.2 HMM Implementation Issues

When implementing the HMM procedures, numerical values of the various probabilities can be extremely small and can often fall below the precision range of real-world, computational machines [1]. Therefore, scaling procedures and the use of logarithmic values are typically employed in order to perform computations using numbers more compatible with the precision range of currently-available, computing equipment. For example, if we are using a maximum likelihood approach to decide crack lengths, rather than use the classifier shown in (5.1), a log likelihood classifier is used as

$$\hat{\lambda} = \underset{\lambda}{\text{argmax}} \left\{ \log p\left(\mathbf{Y}_k^{\text{test}} | \lambda\right) \right\}. \tag{5.2}$$

We then choose the HMM $\hat{\lambda}$ corresponding to the maximum log likelihood and assign the value of the crack length associated with the MPD observation $\mathbf{Y}_k^{\text{test}}$.

Alternatively, we can use the log likelihood outputs from the forward-backward algorithm to create a distribution of likelihoods. This is especially useful when we have multiple observations. Using sequential importance sampling, when updating particle weights as in Equations (2.21-2.24), we can first obtain the log likelihoods of the MPD observations with respect to the different HMM training crack lengths, shift the log likelihood values to an appropriate numerical range (where their exponentials are within the precision range of typical computational machines), and then obtain the exponentials and normalize to convert to probabilities. Error-catching code can also be included to address any values that still remain less than the precision range. The likelihoods can then be used to update particle weights as in Equations (2.21-2.24).

Note that when using HMMs for the crack-growth estimation problem, we consider two different levels of state transitions and two different types of time steps which we clarify next to avoid possible confusion. In our overall crack growth modeling approach, we have different crack length damage states that occur at different fatigue cycle time steps k, and our overall objective includes monitoring crack length damage state transitions at each fatigue time step k. However, at each fatigue time step k, we make an observation of the crack length consisting of a sequence of MPD feature vectors. The MPD feature vectors are time-ordered in the time-frequency plane. Using the time-ordered MPD feature vectors, HMMs are used to correlate the feature vectors with corresponding crack length damage conditions. Therefore, in our overall crack length tracking, we make observations of crack length state transitions at time step k, and for each of these crack length observations, we obtain MPD feature vectors, in actual time, as time-frequency state transitions.

Given an observation mode using a fixed transmission waveform, and given D trained HMMs $\{\lambda_d\}$, $d = 1,\ldots,D$, corresponding to D different damage conditions, once a likelihood $\Pr(\mathbf{Y}_k|\lambda_d)$ (or a log likelihood $\Pr(\log \mathbf{Y}_k|\lambda_d)$) is obtained for each candidate HMM λ_d, the maximum likelihood classifier in (5.1) (or the maximum log

likelihood classifier in (5.2)) can be used to select the HMM $\hat{\lambda}$. Then the corresponding damage class used to train that HMM can be chosen. In a particle filter, crack length tracking implementation, the distribution of likelihoods is used in the particle weight updating procedures described in Equations (2.21), (2.22) and (2.24).

5.2 Waveform Selection using Multiple Observation Modes

In this section, we discuss how to assess the damage resolution characteristics of different observation modes given that they originated from different trained HMMs. We assess these characteristics for a given predicted damage level in order to predict which optimal observation mode, corresponding to an optimal transmission waveform, can be used to observe a crack length.

5.2.1 Use of HMM for Predicting Observations

Given a trained collection of HMMs $\{\lambda_d\}$ for different damage conditions $d = 1,\ldots,D$, and a sequence of MPD feature vectors $\mathbf{Y}_k^{\text{test}}$ representing an unknown damage condition, we could efficiently calculate $\Pr\left(\mathbf{Y}_k^{\text{test}}|\lambda_d\right)$ for each applicable damage condition and obtain a distribution of the likelihoods corresponding to different crack lengths.

In particular, given a crack length damage condition of interest, if we find the nearest HMM training crack length d, we can use the corresponding trained HMM λ_d in a reverse emission mode, where we sample from $p(\mathbf{Y}_k|\lambda_d)$ to generate predicted MPD observation vectors from λ_d. If these predicted MPD emissions are evaluated to obtain the corresponding predicted crack length values, the amount of error can be assessed between the original crack length of interest and the predicted crack length corresponding to the MPD observations sampled from $p(\mathbf{Y}_k|\lambda_d)$. If we have several different HMMs that have been trained using different modes of observation within this crack length region, this type of error assessment can provide useful information with respect to choosing the observation mode corresponding to minimum estimation error or max-

imum damage resolution characteristics. Furthermore, since resolution characteristics can change with damage conditions, the error assessment could be performed before each new observation. This can increase the probability that the optimal observation mode or transmission waveform is selected to make the observation.

Given a continuous-observation HMM λ_d that is trained using a particular damage crack length d at cycle time step k, we can sample from $p(\mathbf{Y}_k|\lambda_d)$ to generate a simulated MPD observation $\mathbf{Y}_k = \{\mathbf{y}_{k,1}, \mathbf{y}_{k,2}, \ldots, \mathbf{y}_{k,L}\}$ using the following procedure:

Initialization: Sample an initial state q_1 from $\pi = \{\pi_1, \pi_2, \ldots, \pi_N\}$.

for $\ell = 1$ **to** L**:**

1. Choose a four-dimensional Gaussian mixture component by sampling c_{jm}, $m = 1, \ldots, M$, from Gaussian mixture coefficient matrix $\mathbf{C}$; this corresponds to the row of mixture coefficients that correspond to the current state h_j.
2. From the four-dimensional Gaussian mixture component, sample an MPD observation vector $\mathbf{Y}_{k,\ell}$

$$\mathbf{Y}_{k,\ell} = \begin{bmatrix} \tau_{k,\ell} \\ \nu_{k,\ell} \\ \kappa_{k,\ell} \\ \rho_{k,\ell} \end{bmatrix}$$

3. If $\ell < L$, choose the $(\ell+1)$th state by sampling from the row of state transition probabilities in matrix $\mathbf{P}$ that correspond to the current lth state.
4. Repeat steps 1-3 until all L MPD vectors are generated.

5.2.2 Predicted crack lengths

Within our sequential Bayesian methodology, given a current crack length state estimate, the HMM trained to that crack length can be used to generate predicted observations using the procedure described in Section 5.2.1. The predicted observations can

be used to obtain predicted crack length estimates which can be compared to the current crack length state estimate. For example, using a particle filter to track crack growth at cycle time step k, the ith particle, $i = 1,\ldots,N_{\mathrm{p}}$ has a crack length state estimate $x_k^{(i)}$. Given an observation mode for each particle value $x_k^{(i)}$, we can find the HMM trained using that observation mode at the crack length closest to the value of this particle. Employing the HMM and the steps for sampling MPD vectors in Section 5.2.1, we generate $\mathscr{M}$ sets of L MPD feature vectors $\mathbf{y}_{k,l,m}^{(i)}$, $l = 1,\ldots,L$ and $m = 1,\ldots,\mathscr{M}$. For each of the $\mathscr{M}$ sequences of L MPD feature vectors, the forward-backward algorithm can be used to obtain the log likelihood of the L MPD vectors. The HMM $\hat{\lambda}_d$ corresponding to the maximum log likelihood is selected using (5.2), and the associated crack length d used to train $\hat{\lambda}_d$ can be assigned as the predicted crack length value $\hat{x}_{k,m}^{(i)}$. Another way to obtain the predicted crack length is to use a probability distribution of the different HMM training crack lengths and then use the mean estimate of the distribution as the predicted crack length value $\hat{x}_{k,m}^{(i)}$. Once a value has been assigned to the predicted crack length $\hat{x}_{k,m}^{(i)}$, the error between this predicted crack length $\hat{x}_{k,m}^{(i)}$ and the corresponding particle value $x_k^{(i)}$ can be computed. This error can then be used to assess the resolution characteristics of the particular observation mode used to train the HMMs, within the local crack length region corresponding to $x_k^{(i)}$.

5.2.3 Waveform Design Using Error Minimization

If there are several different transmission waveforms available for obtaining crack length observations, the approach described in Section 5.2.2 can be used to select the transmission waveforms that can be used to perform damage monitoring as accurately as possible. For a particular damage condition, if we have several different HMMs trained for this damage condition but using different observation modes, the errors between the current crack length estimate and the crack lengths obtained from predicted observations can be used to choose the observation mode corresponding to the minimum error. This minimum error corresponds to the maximum damage res-

olution characteristics. Furthermore, since resolution characteristics can change with different damage conditions, the error assessment can be performed before each new damage observation to ensure that the optimal observation mode is used. In our crack growth tracking methodology, we employ this waveform design approach to select an optimal observation mode when making observations of feature vectors corresponding to the current damage condition.

A mean-squared error (MSE) metric is used to select the optimal transmission waveform. Given a particle filter implementation at cycle time step k with corresponding particles and weights $(x_k^{(i)}, w_k^{(i)})$, the predicted crack lengths $\hat{x}_{k,m}^{(i)}$, $m = 1, \ldots, \mathscr{M}$, are obtained for each particle $x_k^{(i)}$. Given J available transmission waveforms, $\mathbf{z}_k^{(j)}$, $j = 1, \ldots, J$, the predicted MSE estimate for waveform $\mathbf{z}_k^{(j)}$ at cycle time step k can be obtained as

$$\mathrm{MSE}(\mathbf{z}_k^{(j)}) = \sum_{i=1}^{N_\mathrm{p}} w_k^{(i)} \left\{ \frac{1}{\mathscr{M}} \sum_{m=1}^{\mathscr{M}} \left[x_k^{(i)} - \hat{x}_{k,m}^{(i,j)} \right]^2 \right\}, \tag{5.3}$$

where $\hat{x}_{k,m}^{(i,j)}$ is the mth predicted crack length corresponding to particle $x_k^{(i)}$ and the jth waveform. Once the predicted MSE estimates are obtained for each of the J transmission waveforms, the transmission waveform $\hat{\mathbf{z}}_k$ corresponding to the minimum predicted MSE is chosen to obtain the observation at cycle time step k as

$$\hat{\mathbf{z}}_k = \underset{\mathbf{z}_k^{(j)}}{\operatorname{argmin}} \, \mathrm{MSE}(\mathbf{z}_k^{(j)}) \tag{5.4}$$

By selecting the transmission waveform corresponding to the minimum predicted MSE, we expect the best overall crack length resolution characteristics and the highest accuracy when making crack length observations. Using this waveform selection procedure, only a single waveform corresponding to the minimum MSE is selected. If more than one waveform yields the minimum MSE, we select the waveform with the shortest duration. Later in this chapter, we discuss an approach where we can select combinations of observation modes.

5.2.4 *Test Specimen Configuration*

In this demonstration of waveform selection, a 2024-T3 aluminum cruciform specimen was tested in fatigue to obtain crack growth data. A 0.25 inch diameter hole was drilled in the center of the specimen, and a 1.0 mm notch was electrochemically machined into an edge region of the hole to create a nucleation origin for a crack to develop. The cruciform specimen was positioned within a biaxial mechanical test machine (shown in Figure 5.1), and the specimen was loaded simultaneously in both the horizontal (x-axis) and vertical (y-axis) directions. As shown in Figure 5.2, a crack developed and propagated from the notch origin to the edge of the test specimen gage area. The test was discontinued prior to overload and specimen failure.

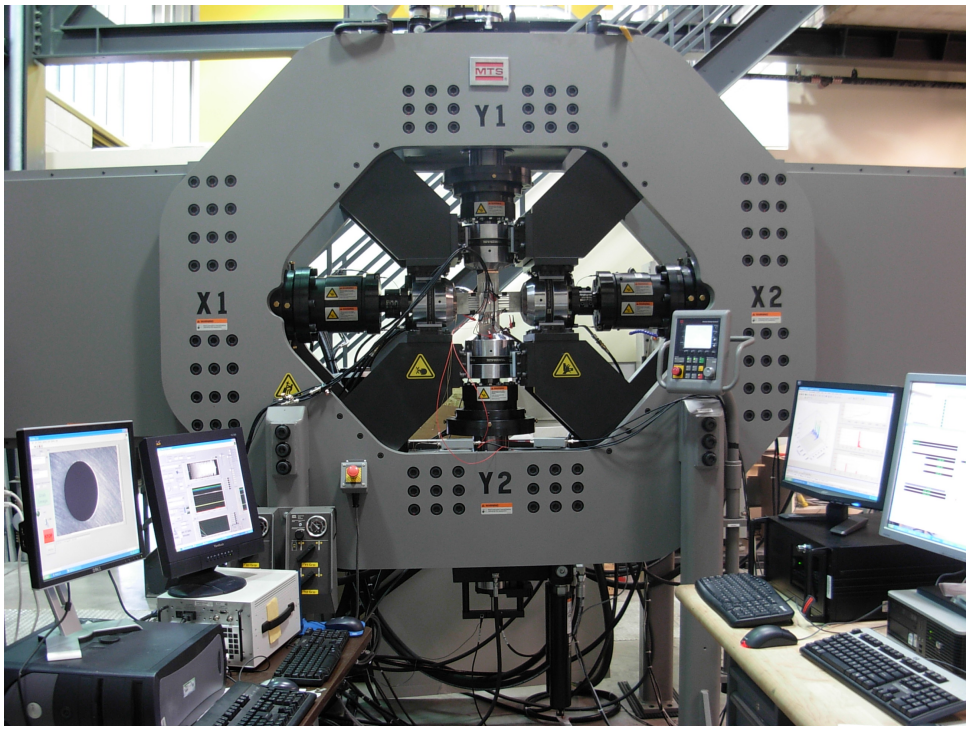


Figure 5.1: Biaxial mechanical test machine.

The fatigue testing was performed in the Integrated Mechanical Testing Laboratory, that is part of the Adaptive Intelligent Materials and Systems (AIMS) Center

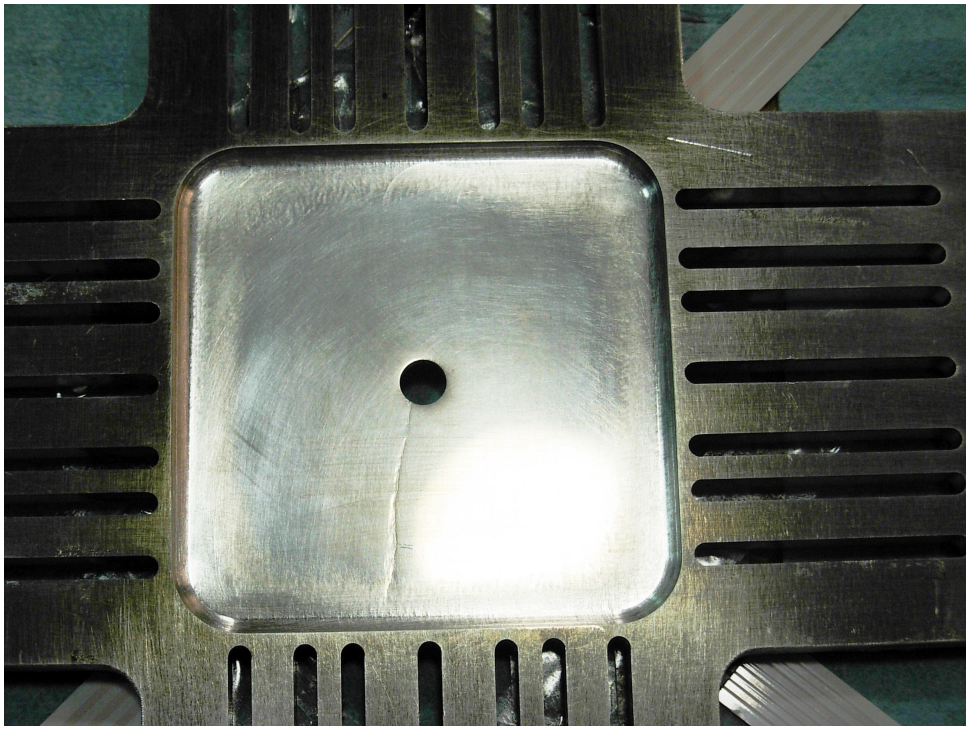


Figure 5.2: Closeup of tested specimen fatigue crack.

(`http://aims.engineering.asu.edu/`) at Arizona State University. The target and actual mean loads were 2.64 kips and 2.16 kips, respectively, with minimum and maximum loads at 0.48 and 4.8 kips, respectively. The load ratio was $R = 0.1$. Note that 1 kips corresponds to 1000 pounds load.

As shown in Figure 5.3, multiple piezoelectric transducers were placed in the specimen gage area, with one transducer used as an actuator (signal transmitter) transducer (as designated in the figure) and five other transducers used as receiver sensors. Initially, we focus our discussion upon the data that was received by Sensor 2. Later in this chapter, we consider data received by other sensors.

5.2.5 Transmission Waveforms for Selection

Three different types of waveforms were transmitted, as shown in Figure 5.4. These waveform included a tone burst waveform, a linear chirp waveform, and a signal consisting of a series of pseudo random width pulses. All waveforms were generated

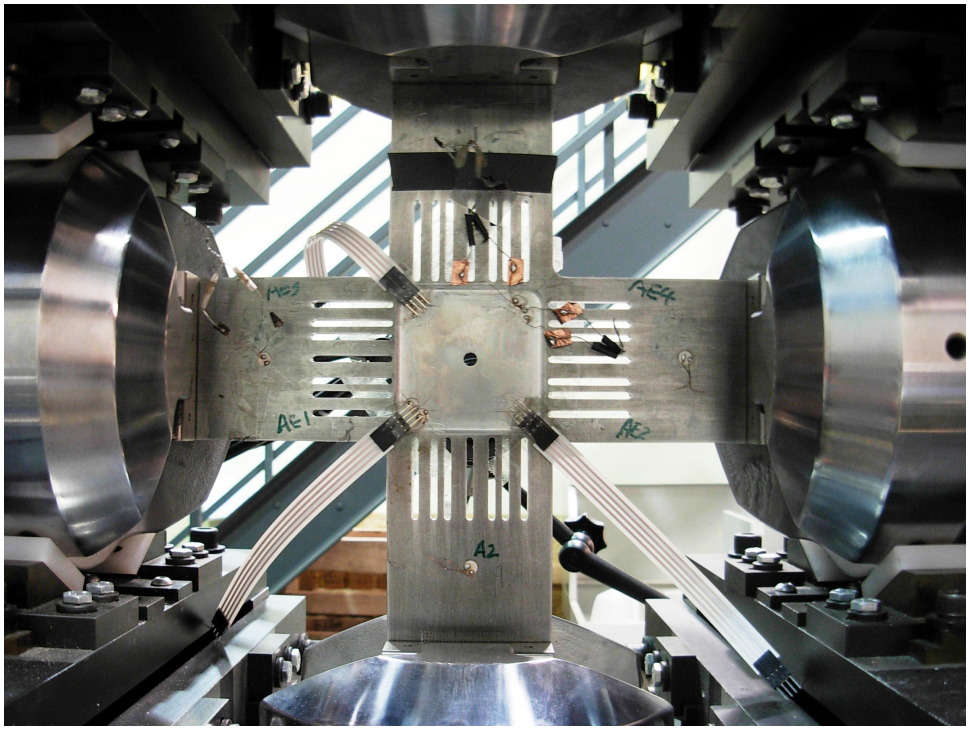


Figure 5.3: Cruciform specimen with sensors.

using a 2 MHz sampling frequency.

The tone burst waveform is an amplitude enveloped sinusoid given by

$$s_{bu}(t) = A_{bu}(t)\cos(2\pi f_{bu}t)$$

where f_{bu} is the sinusoidal frequency and $A_{bu}(t) = Ce^{-\kappa t^2}$ is a Gaussian window. The constant C adjusts the overall amplitude and κ adjusts the shape of the window. Key characteristics of this waveform are its high localization with respect to both time and frequency, with a relatively short time duration and a limited bandwidth. In our implementation, the tone burst signal had a target frequency of 230 kHz and an overall signal duration of approximately 22 μs.

In contrast to the highly-localized tone burst signal, a linear chirp waveform has significantly longer time duration wider bandwidth. In our work, the linear chirp

waveform was constructed as

$$s_{ch}(t) = A_{ch} \sin\left(2\pi f_0 t + \frac{c}{2}t^2\right)$$

where A_{ch} is a constant target amplitude, f_0 is the initial frequency at time $t = 0$, and c is the chirp rate. In our implementation, the chirp signal was generated using a linearly-increasing frequency, ranging approximately 0-950 kHz, over a duration of approximately 950 μs, such that the chirp rate of the signal was approximately 1 GHz/s.

In addition to the highly-localized tone burst signal and the long-duration, broad-bandwidth linear chirp, we also included a pseudo-noise random width pulse waveform. This waveform consisted of a sequence of random width pulses, where the periods of individual pulses and periods between pulses varied between approximately 5 to 15 μs, with a total multi-pulse signal duration of approximately 165 μs. This waveform was included to observe the general damage classification characteristics when a random, noise-like signal was used as the transmitted waveform.

5.2.6 Data Collection for Model Training

In a real-life implementation of our crack growth tracking method on a real structure, we assume that there would be previously-obtained training data available. This includes actual crack length measurements obtained at the different fatigue cycles of interest, as well as signal data that was obtained at each of the measured crack lengths. The signal data was obtained using different modes of observation by using different transmission waveforms. Therefore, in the model training portion of our overall crack monitoring procedure, the physical measurement data is used in developing the crack growth state dynamic model, and the active-sensing, signal data is used in developing the observation model.

Within this demonstration of our method, at various cycles during the fatigue testing of the cruciform specimen, the fatigue loading was temporarily stopped, the crack length was physically measured, and signal data was generated by transmitting

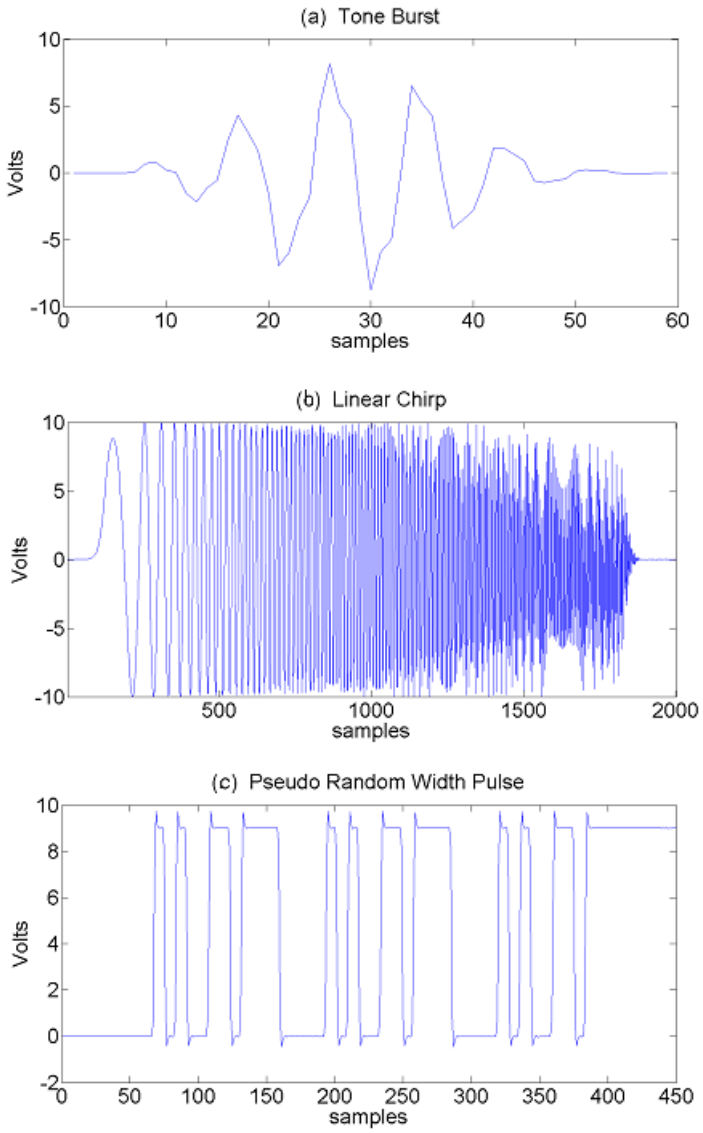


Figure 5.4: Different types of transmitted waveforms. (a) Tone burst, (b) linear chirp, and (c) pseudo-noise random width pulse.

and receiving the different waveforms using the piezoelectric transducers. For each waveform, approximately 50 different signal observations were obtained at each of the five receiver sensors at each of the designated fatigue cycles. The specific number of fatigue cycles and corresponding crack lengths are shown in Table 5.1 and also in the crack growth plot shown in Figure 5.5. In this plot, a least squares cubic model was fitted to the data.

Number of Cycles	Measured Crack Length (mm)
0	0
87037	0
110009	4.20
140011	4.70
161761	5.13
174238	7.00
184009	8.92
192961	11.12
200013	13.15
208012	15.04
213794	18.09
218462	20.11
225008	22.62
229009	24.75
234216	28.09

Table 5.1: Fatigue cycle numbers where crack length data was obtained.

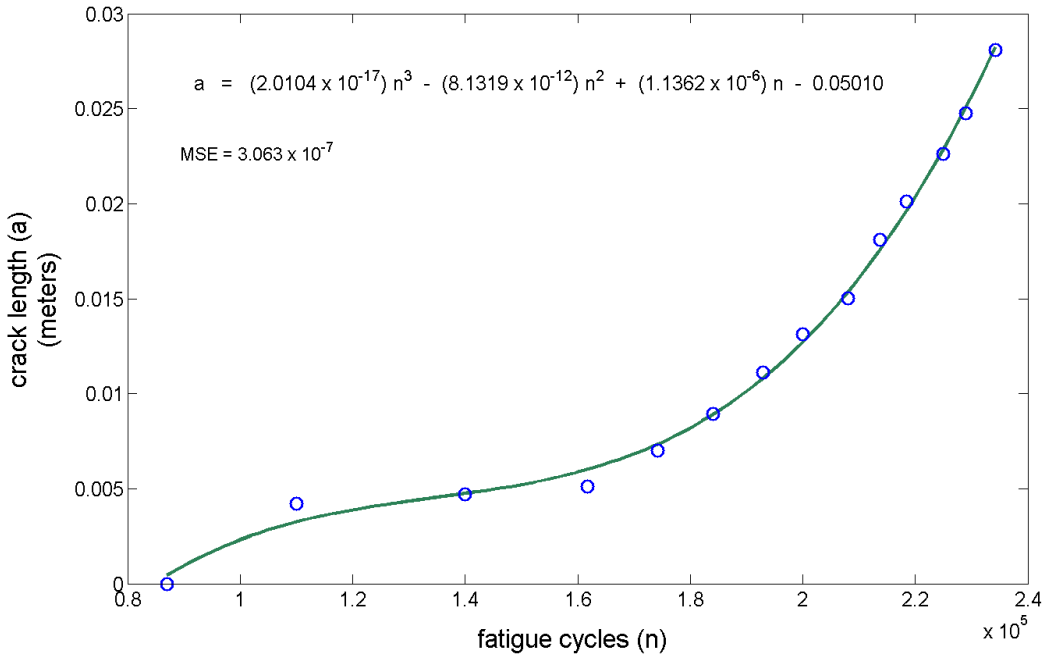


Figure 5.5: Measured crack length vs. number of fatigue cycles, with least squares cubic model.

5.2.7 State Dynamics Model

In Section 3.3, simplifications and stochastic variations of Paris' law were discussed, and in this demonstration of our crack tracking method, we employ a Paris-like

power law based upon the discussed variations of Paris' Law. Using the fitted, crack length versus fatigue cycle cubic model shown in Figure 5.5, we used the derivative of the cubic model and evaluated it at fatigue cycles at which crack measurements were obtained. In addition, using the values of the crack length measurements, we obtained initial estimates of the stress intensity factor $K_{\mathrm{I}}(a)$ using the experimentally-determined cubic model shown in Figure 3.5, in which we used this as a measure of the effective stress intensity factor range ΔK_{eff}. Using (3.4), and as shown in Figure 3.6, we could obtain initial estimates of the Paris constants C and m. However, there can be wide variability in these initial estimates. Hence, we employed Markov chain Monte Carlo (MCMC) techniques to refine the values of these parameters. Results of our example MCMC implementation are shown in Tables 5.2-5.4. As the tables show, relatively constant values for m and C could be obtained, while maintaining values for K_{I} that were still within the range of predicted values shown in Figure 3.5.

cycles	crack(mm)	m(initial est)	m(MCMC)
110009	4.20	-9.0888	0.4172
140011	4.70	8.0926	0.3982
161761	5.13	12.2138	0.4236
174238	7.00	8.2708	0.4542
184009	8.92	6.3777	0.4504
192961	11.12	5.3171	0.4636
200013	13.15	4.7060	0.4890
208012	15.04	3.9934	0.4788
213794	18.09	3.3387	0.4990
218462	20.11	2.7281	0.4942
225008	22.62	1.8850	0.4697
229009	24.75	1.4531	0.5023
234216	28.09	1.0237	0.4874

Table 5.2: Initial and MCMC-refined estimates of state dynamics model parameter m.

After refining the values of these parameters, and using a discrete time step k notation, where k represents the indices of the fatigue cycles where measurements were

cycles	crack(mm)	m(initial est)	m(MCMC)
110009	4.20	1.0807×10^{56}	2.1545×10^{-8}
140011	4.70	1.3369×10^{-64}	2.3782×10^{-8}
161761	5.13	2.4383×10^{56}	2.3545×10^{-8}
174238	7.00	1.0847×10^{-65}	2.4727×10^{-8}
184009	8.92	2.232×10^{-52}	2.4976×10^{-8}
192961	11.12	6.7391×10^{-45}	2.6814×10^{-8}
200013	13.15	1.4119×10^{-40}	2.3844×10^{-8}
208012	15.04	1.5961×10^{-35}	2.4590×10^{-8}
213794	18.09	7.200×10^{-31}	2.3414×10^{-8}
218462	20.11	1.6128×10^{-26}	2.5109×10^{-8}
225008	22.62	1.7019×10^{-20}	2.4496×10^{-8}
229009	24.75	2.1239×10^{-17}	2.5837×10^{-8}
234216	28.09	2.6097×10^{-14}	2.5917×10^{-8}

Table 5.3: Initial and MCMC-refined estimates of state dynamics model parameter C.

cycle(k)	cycle(k+1)	crack(mm)(k)	crack(mm)(k+1)	KI (k)	KI (k+1)
87037	110009	0.00	4.20	8165454	8722830
110009	140011	4.20	4.70	9429409	8560679
140011	161761	4.70	5.13	9631181	8794103
161761	174238	5.13	7.00	10079931	8037457
174238	184009	7.00	8.92	10747788	11940649
184009	192961	8.92	11.12	11358623	9895599
192961	200013	11.12	13.15	11899865	9536098
200013	208012	13.15	15.04	12346258	10431927
208012	213794	15.04	18.09	12877235	11020609
213794	218462	18.09	20.11	13434196	12665571
218462	225008	20.11	22.62	13997026	13738065
225008	229009	22.62	24.75	14665976	1372468
229009	234216	24.75	28.09	15681117	15749785

Table 5.4: Initial and MCMC-refined estimates of state dynamics model parameter K_I.

taken (reference Table 5.1), we then estimated a crack growth rate $\dot{a}_k$ at time step k:

$$\dot{a}_k \approx C_k \left[K_{\mathrm{I}(k,k+1)}\right]^{m_k} . \tag{5.5}$$

In (5.5), C_k and m_k are discrete time step versions of Paris parameters C and m, in which we assume that these parameters are not true constants but have some variability. $K_{\mathrm{I}(k,k+1)}$ is the average value of stress intensity factor $K_{\mathrm{I}}(a)$ at time steps k and $k+1$,

which we are using as a measure of the effective stress intensity factor range ΔK_{eff}. To obtain this average value, crack lengths were obtained from Table 5.1 corresponding to time step indices k and $k+1$. Then the crack lengths were used as input to the experimentally-determined cubic model shown in Figure 3.5 and the average value was used.

As discussed in Section 3.3, a lognormal random variable is often used to model variability in crack growth rate in theoretical stochastic crack growth studies. However, when tracking actual crack growth, there can be both negative and positive random variability, where negative variability could possibly occur due to random noise within crack length observations. The combination of a lognormal random variable (to model variability within the state dynamics, crack growth rate) with a normal random variable (to model variability within crack length observations) represents one approach towards modeling the combined variability that could occur when tracking crack growth. Alternatively, a single, normal random variable might be sufficient to model the randomness in both positive and negative directions, and that is the approach we used within the effort discussed in this chapter.

Using a single, normal random variable to model the variability in the crack growth, and using the discrete time step, crack growth rate expression shown in (5.5), we formed the following state dynamics model

$$a_{k+1} = a_k + \dot{a}_k \Delta n_k + v_k \tag{5.6}$$

where Δn_k is the number of fatigue cycles between time steps k to $k+1$ and v_k is a zero-mean Gaussian random variable with variance σ_a^2, included to account for randomness during crack growth tracking. An estimate of σ_a^2 is obtained using the MSE of the crack growth cubic model shown in Figure 3.2.

5.2.8 MPD-HMM Observation Model For Different Waveforms

In accordance with procedures described in Section 2.7.2, HMMs were trained for each combination of waveform and crack length, using sequences of MPD feature vectors as the observable quantities emitted from hidden states in the time-frequency plane. Within this particular study, all signals were received by a single sensor, and we implemented different observation modes by using different transmission waveforms. After training the HMMs, given a new MPD vector sequence corresponding to an unknown crack length, and given the particular transmission waveform and the collection of HMMs trained with that waveform for different crack lengths, the specific HMM and its associated training crack length could be selected based upon the procedures described in Section 5.1.2.

In extracting MPD feature vectors from the received signals, we performed $L = 40$ MPD iterations, such that our observations consisted of $L = 40$ length sequences of four-element MPD feature vectors. This relatively large number of MPD iterations was necessary due to characteristics of the pseudo random width pulse waveform, since more MPD iterations were required to extract a sufficient amount of signal energy from this noise-like waveform as compared to the other two waveforms. Even though the tone burst and linear chirp waveforms required fewer MPD iterations to extract comparable levels of signal energy, to maintain consistency in our comparison of the three different transmission waveforms, we intentionally chose to use the same number ($L = 40$) of MPD iterations for all three waveforms when performing the MPD procedure. In the MPD algorithm, most of the signal energy is extracted during the early iterations, so that after 40 iterations, the amount of signal energy extracted from all three waveforms was approximately the same. After obtaining the $L = 40$ feature vectors using the MPD procedure, the vectors were sorted according to their time shift (τ) parameter, such that they were arranged according to their time-based, sequential order of emission within

the time-frequency plane.

In creating and training the HMMs, we assumed there were $N = 8$ hidden states that existed within time-frequency plane. For data generated with the tone burst and linear chirp transmission waveforms, we were able to train the HMMs adequately with a somewhat lower assumed number of hidden states; however, training the pseudo random width pulse waveform required a higher assumed number ($N = 8$) of hidden states in order to obtain reliable crack length classification results. As with our selection of the number of MPD iterations, in order to maintain consistency when comparing the observation characteristics of the three different transmission waveforms, we intentionally chose to use the same assumed number ($N = 8$) of hidden states for all three waveforms when creating and training the HMMs.

We chose to use $M = 2N = 16$ Gaussian mixture components in our HMMs. We assumed this number of mixture components would be sufficient for various feature vector distributions that might occur, including mono-modal distributions, multi-modal distributions, and situations where different damage classes possessed common mixture components. As with the number of states N, we intentionally chose to keep the same number of Gaussian mixture components ($M = 16$) for all waveforms.

The HMMs were trained using 24 training observations for each waveform and crack length combination, where observations consisted of a time-shift ordered sequence of $L = 40$ MPD feature vectors. Training was performed using procedures described in Section 2.7.2. Multiple training epochs were used for each set of observations. In each training epoch, the specific order in which the MPD observations were presented to the HMM was randomized. Although we tried to be as consistent as possible when processing the different waveforms in other procedures, we had to used different numbers of training epochs for the different waveforms when training the HMMs. In supervised learning algorithms, there is a possibility of *overtraining* or *overfitting* the

data, where the learning algorithm not only learns key features associated with a particular class of training data, but also can learn random noise and other nuances within the training data unrelated to the particular data class. This leads to reduced generalization when classifying new data, since random characteristics that occur within new data will generally be different from those that occurred in the training data. To improve generalization, a sufficient number of training epochs should be used such that key features associated with a particular data class are learned, but after the key features have been learned, *early stopping* should be performed to reduce the probability of overtraining. For the tone burst waveform, 12 training epochs were sufficient to learn the key features of the training data. Since the linear chirp waveform had longer duration and wider bandwidth, we used slightly more epochs (24) when training HMMs with this waveform. The pseudo random width pulse waveform presented an interesting example with respect to generalization and overtraining. This particular waveform was intentionally created as a random, noise-like waveform, and consequently, the training process with this waveform was particularly susceptible to overtraining. If we used more than a limited number of epochs to train the HMMs with this waveform, we saw a dramatic drop-off in the ability to classify new data. We finally chose to use only 4 epochs to train with this waveform. This appeared to be a sufficient number to allow the HMMs to learn overall, macro envelope features related to different damage classes. If we used more training epochs, the HMMs apparently began to learn individual pulse width characteristics within the training data. Since these characteristics varied from run to run and would likely be different in new data, it was particularly important to limit the number of HMM training epochs with the pseudo random width pulse waveform such that the trained HMMs did not become overtrained to these random pulse width characteristics.

Tables 5.5-5.7 show confusion matrix, crack length classification results for the three different waveforms, where test signals were received using Sensor 2, and then processed using our MPD-HMM observation model to determine the corresponding

crack length. The test signals were generated on the same specimen (see Figure 5.3) as the training signals, as well as at the same crack lengths experienced during the fatigue loading of the specimen. There were 24 test signals generated for each transmission waveform and crack length combination. Test signals of a particular known transmission waveform were processed by the collection of HMMs trained with that waveform at the different crack lengths. Using the procedures described in Section 5.1.2, crack lengths were determined for the different test signals. As the confusion matrices show, we obtained excellent crack length classification for all three transmission waveforms. This is impressive, since there was a relatively large number (13) of different damage classes, and the results show that our MPD-HMM observation model was capable of distinguishing between signals obtained from the different crack length classes. In the few cases where notable misclassifications occurred, these misclassifications primarily occurred at crack levels immediately adjacent to the correct crack level. Therefore, even in these cases, the error in the crack length classification was relatively small.

crack(mm)	4.20	4.70	5.13	7.00	8.92	11.12	13.15	15.04	18.09	20.11	22.62	24.75	28.09
4.20	**0.7083**	0.2917	0	0	0	0	0	0	0	0	0	0	0
4.70	0	**1**	0	0	0	0	0	0	0	0	0	0	0
5.13	0	0.7083	**0.2917**	0	0	0	0	0	0	0	0	0	0
7.00	0	0.0417	0	**0.9583**	0	0	0	0	0	0	0	0	0
8.92	0	0	0	0	**1**	0	0	0	0	0	0	0	0
11.12	0	0	0	0	0	**0.9583**	0	0	0	0	0	0.0417	0
13.15	0	0	0	0	0	0	**0.9583**	0	0.0417	0	0	0	0
15.04	0	0	0	0	0	0	0	**0.9583**	0.0417	0	0	0	0
18.09	0	0	0	0	0	0	0	0	**0.9583**	0.0417	0	0	0
20.11	0	0	0	0	0	0	0	0	0	**0.9583**	0.0417	0	0
22.62	0	0	0	0	0	0	0	0	0	0	**0.9167**	0	0.0833
24.75	0	0	0	0	0	0	0	0	0	0	0	**0.8750**	0.1250
28.09	0	0	0	0	0	0	0	0	0	0	0	0	**1**

Table 5.5: Confusion matrix for tone burst waveform crack length classification.

The diagonal values of the three confusion matrices, that provide correct classification, were combined and plotted in Figure 5.6. As the figure shows, when notable misclassification occurred for one of the waveforms at a particular crack length damage level, there was accurate classification at that same crack length for at least one of the other waveforms, if not both of the other waveforms. Therefore, there can be an

crack(mm)	4.20	4.70	5.13	7.00	8.92	11.12	13.15	15.04	18.09	20.11	22.62	24.75	28.09
4.20	**1**	0	0	0	0	0	0	0	0	0	0	0	0
4.70	0	**1**	0	0	0	0	0	0	0	0	0	0	0
5.13	0	0	**1**	0	0	0	0	0	0	0	0	0	0
7.00	0	0	0	**1**	0	0	0	0	0	0	0	0	0
8.92	0	0	0	0	**1**	0	0	0	0	0	0	0	0
11.12	0	0	0	0	0	**1**	0	0	0	0	0	0	0
13.15	0	0	0	0	0	0.2083	**0.7917**	0	0	0	0	0	0
15.04	0	0	0	0	0	0	0	**1**	0	0	0	0	0
18.09	0	0	0	0	0	0	0	0	**0.7083**	0.2917	0	0	0
20.11	0	0	0	0	0	0	0	0	0	**1**	0	0	0
22.62	0	0	0	0	0	0	0	0	0	0.9167	**0.0833**	0	0
24.75	0	0	0	0	0	0	0	0	0	0.1250	0	**0.8750**	0
28.09	0	0	0	0	0	0	0	0	0	0	0	0.0417	**0.9583**

Table 5.6: Confusion matrix for pseudo-noise random width pulse waveform crack length classification.

crack(mm)	4.20	4.70	5.13	7.00	8.92	11.12	13.15	15.04	18.09	20.11	22.62	24.75	28.09
4.20	**1**	0	0	0	0	0	0	0	0	0	0	0	0
4.70	0	**1**	0	0	0	0	0	0	0	0	0	0	0
5.13	0	0	**1**	0	0	0	0	0	0	0	0	0	0
7.00	0	0	0	**1**	0	0	0	0	0	0	0	0	0
8.92	0	0	0	0	**1**	0	0	0	0	0	0	0	0
11.12	0	0	0	0	0	**1**	0	0	0	0	0	0	0
13.15	0	0	0	0	0	0	**0.8750**	0	0.1250	0	0	0	0
15.04	0	0	0	0	0	0	0	**1**	0	0	0	0	0
18.09	0	0	0	0	0	0	0	0.1667	**0.7917**	0	0	0.0417	0
20.11	0	0	0	0	0	0.0417	0	0.0417	0.2500	**0.6667**	0	0	0
22.62	0	0	0	0	0	0	0	0	0.0833	0	**0.9167**	0	0
24.75	0	0	0	0	0	0	0	0	0.1667	0	0	**0.8333**	0
28.09	0	0	0	0	0	0	0	0	0	0	0	0	**1**

Table 5.7: Confusion matrix for linear chirp waveform crack length classification.

advantage of having different waveforms available for making the observations.

5.2.9 Crack Growth Tracking Results with Waveform Selection

In this section, we discuss the results of our overall sequential Bayesian crack growth tracking procedure, where we incorporated the waveform selection procedure described in Section 5.2 to select the specific transmission waveforms to be used in the observation model. Our overall crack growth tracking method is shown in Figure 5.7 summarizes the procedures used to transition from the posterior state estimate $p\left(x_{k-1}|\hat{\mathbf{Y}}_{k-1}\right)$, to the posterior state estimate $p\left(x_k|\hat{\mathbf{Y}}_k\right)$. These procedures include the state dynamics model, the minimum MSE waveform selection and the observation model.

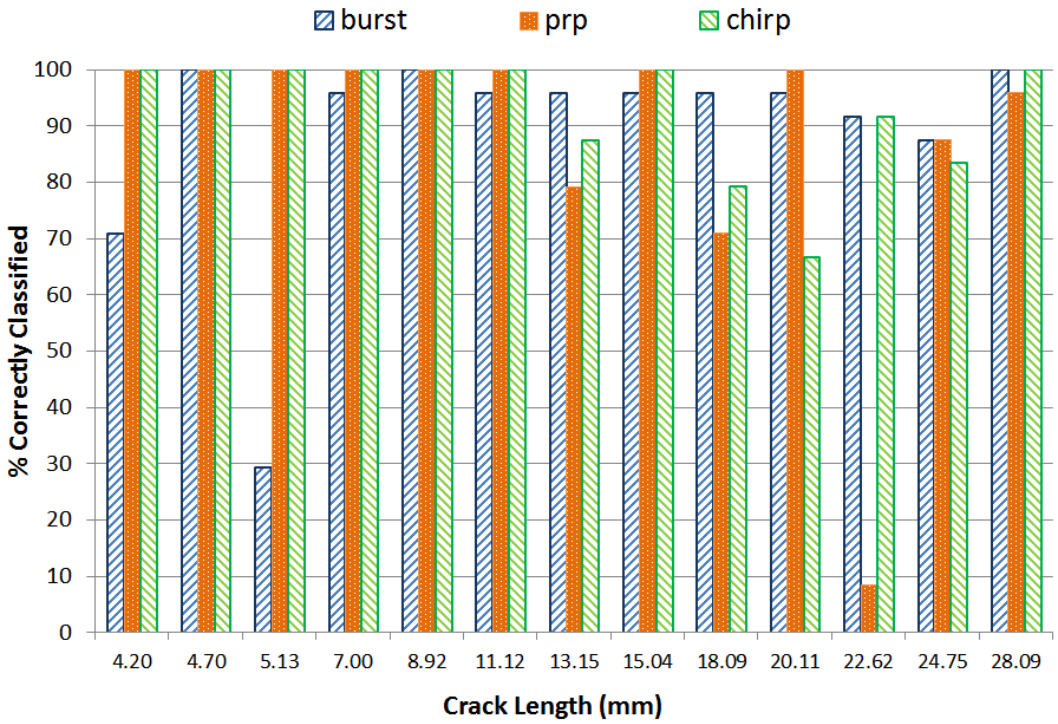


Figure 5.6: Correct crack length classification values for the three different waveforms (diagonal values of the corresponding confusion matrices).

With respect to the waveform selection procedure, Figure 5.8 shows averaged results of several example runs in which the predicted MSE, waveform selection procedure was performed. In this procedure, the MSE value was deemed to be an inverse indicator of crack length resolution, and was used to choose the transmission waveform for obtaining the crack length observation. Based upon MSE values, the pseudo random width pulse waveform, which is inherently noisy, possessed high resolution characteristics at relatively low levels of crack damage, as evidence by low predicted MSE values. It is assumed at low crack lengths, the MPD-HMM observation method was capable of recognizing overall, envelope characteristics of the waveform. However, at longer crack lengths, the inherent randomness within the waveform possibly affected the ability to distinguish between different crack lengths, as evidenced by higher MSE values. The linear chirp waveform, with its relatively long duration and wide bandwidth, exhibited

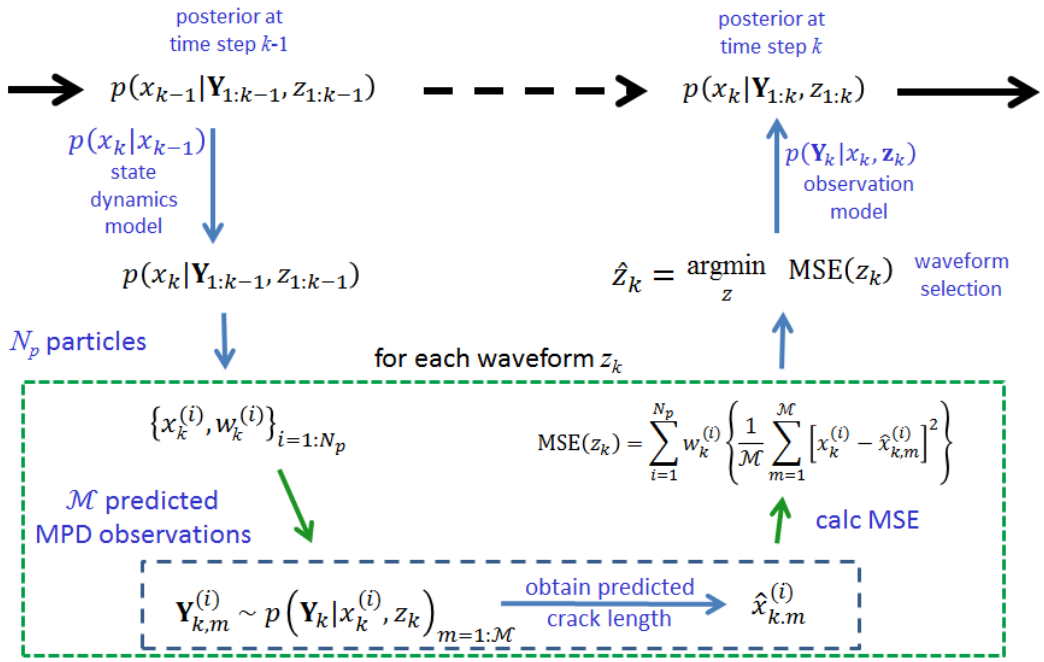


Figure 5.7: Sequential Bayesian estimation of crack length with MMSE waveform selection.

good crack length resolution (low predicted MSE) over much of the crack growth range, although there was a less resolution within crack lengths located in the lower and mid crack length regions. At these lower resolution regions, it is possible the longer duration and higher bandwidth qualities of the linear chirp waveform overwhelmed key time-frequency features related to crack length recognition, thereby reducing resolution characteristics in these regions. In contrast to the linear chirp waveform, the highly localized, tone burst waveform possessed highest resolution within the center portion of the crack growth range. The particular tone burst frequency (230 kHz) used in this study has been found to provide good results for detecting damage in the particular type of cruciform specimen used in this study. The predicted MSE results shown in Figure 5.8 confirm how the 230 kHz tone burst waveform is well-*tuned* to the center region of possible crack lengths for this specimen. As the figure shows, the damage resolution characteristics can be dependent upon both the damage conditions and the chosen transmitted waveform, and hence, to improve robustness in a damage classification or crack

growth monitoring procedure, selection of the optimal waveform could be beneficial.

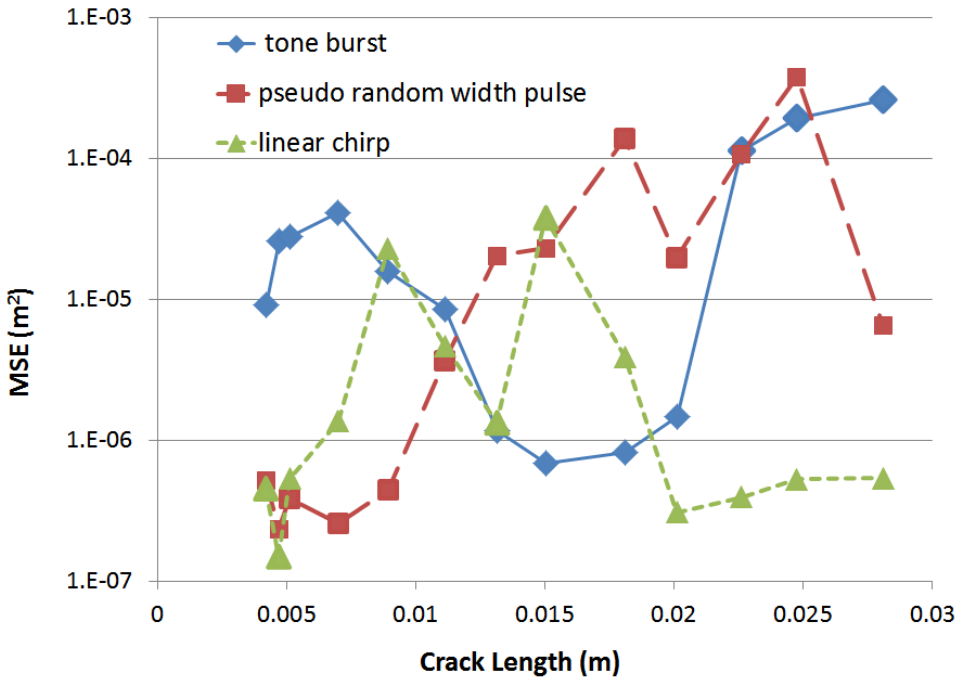


Figure 5.8: Predicted MSE results for different crack lengths.

Figure 5.9 shows the overall crack length monitoring results using our method. At each fatigue cycle time step, 24 test signals were obtained, the corresponding crack lengths were obtained using the implementation approach discussed in Section 5.1.2. In making the observations, the optimal, highest resolution (i.e., minimum MSE) waveform was determined as shown in Figure 5.8 and used to make the observation. As the figure shows, our method exhibited excellent ability to track crack growth accurately.

When waveform scheduling was not used and only a single waveform was used in the crack growth monitoring procedure, there could be occasional errors in the observations, since the single waveform might not be the optimal choice in making observations at all crack lengths, causing estimates to deviate slightly from the true crack growth curve. However, even with this situation, the overall state-space, sequential Bayesian approach was helpful in *smoothing* over occasional observation errors and maintaining

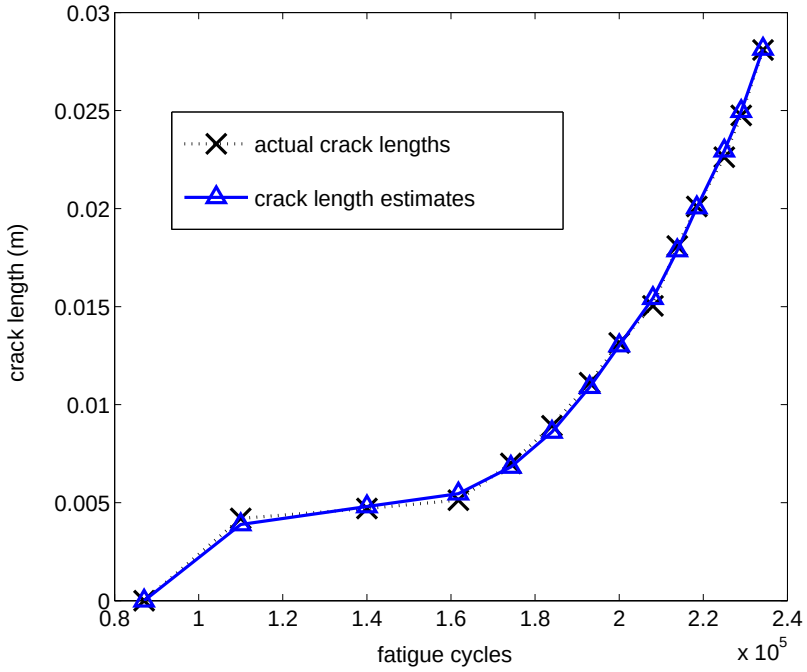


Figure 5.9: Overall crack tracking results.

general accuracy in the crack growth monitoring procedure.

In the run shown in Figure 5.10, we added Gaussian noise to the original (noisy) Sensor #2 observation data, where the (original) signal to (additional) noise ratio was 3 dB. As shown in the figure, the overall tracking procedure with the optimal waveform scheduling was able to maintain general tracking of the crack growth, whereas, without the waveform switching, the crack growth monitoring procedure was less capable of tracking crack growth accurately. In addition to the robustness achieved by including both a predictive state dynamics model with an accurate observation model, we enhanced the accuracy of the observation model, as well as the robustness of our overall procedure, by having several waveforms available, and then evaluating and selecting the optimal waveform prior to each observation.

Although the waveform selection procedure improved the ability to track the overall crack length within an elevated noise environment, the resulting crack length es-

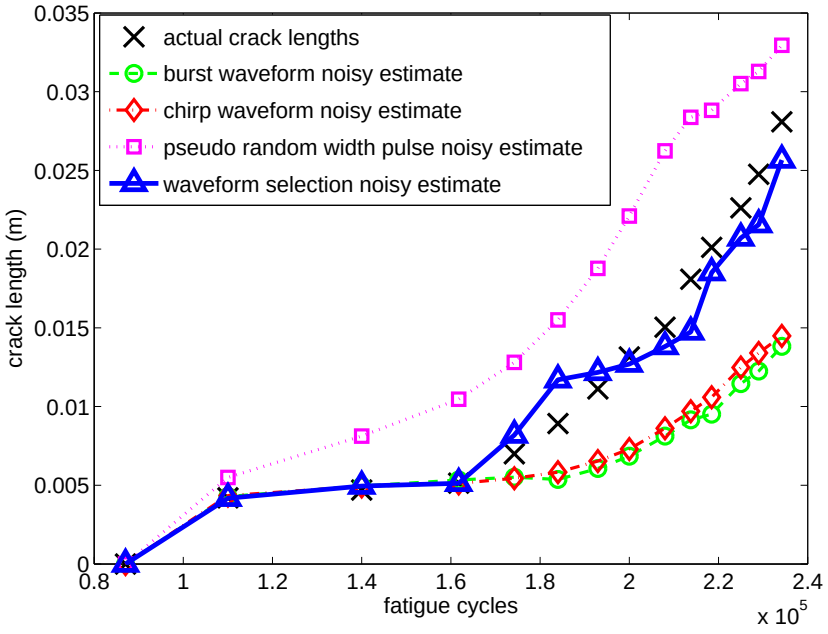


Figure 5.10: Crack tracking with noise added to the data. The waveform selection approach, proposed within this work, was capable of following the general crack growth development, whereas, without waveform selection (e.g., [2]), the tracking procedures tend to drift from the actual crack growth.

timates possessed elevated levels of noise, and this resulted in widely-varying, localized estimates of the crack growth rate. Since the rate of crack growth is a vital component in monitoring the health of a structure, the wide variation of the localized crack growth rates could create uncertainty in the assessment of the structure.

One way in which the noisy crack rate estimates can be addressed is by employing the assumed, linear relationship between the log crack rate and $\log(\Delta K_{\text{eff}})$, as shown in the linear Stage II of Figure 3.6. Neerukatti, et al. [78] have proposed the use of this linear relationship for fitting crack data. In our approach, given k crack length estimates in the monitoring procedure, there are $k-2$ crack length estimates that are preceded and followed by other crack length estimates. At each crack length estimate, the collection of the crack estimate with its neighboring crack estimates can be used to generate a polynomial curve, and the tangent to the curve can be used as a measure of the local

crack rate. If this is performed for several crack length values within the Stage II crack propagation region, the localized, noisy crack rate estimates can be used to construct a smoothed, linear relationship between the log crack rate and $\log(\Delta K_{\text{eff}})$. This linear relationship can then be used to obtain smoothed estimates of the localized crack rate.

As an example, a noisy crack length estimate can be used with its two surrounding, noisy, crack length estimates to create a quadratic curve of the form:

$$a_{est} = \beta_0 + \beta_1 n + \beta_2^2 \tag{5.7}$$

where a_{est} is an estimate of the crack length based upon the quadratic curve, n is the number of fatigue cycles, and β_0, β_1, and β_2 are constant, linear, and quadratic coefficients. Given the crack length estimate at fatigue cycle $\tilde{n}$, the localized crack growth rate $\dot{a}_est$ at this fatigue cycle can be obtained as the tangent to the curve at fatigue cycle $\tilde{n}$:

$$\dot{a}_est = \beta_1 + 2\beta_2\tilde{n}\,. \tag{5.8}$$

In Figure 5.11, given $k = 8$ crack length estimates have been obtained, each of the $(k-2) = 6$ internal crack length estimates are shown with a portion of the associated quadratic curve created using the estimate and its two neighboring estimates. For each of these estimates, the tangent to the curve represents the crack growth rate at that estimate, where this tangent crack growth rate can be evaluated per (5.8).

Within this smoothing approach, we assumed the initial crack length and its associated crack rate were known. Note that the crack rate behavior within the crack initiation stage does not generally possess a linear log crack rate versus $\log(\Delta K_{\text{eff}})$ relationship. Furthermore, at low fatigue cycles, there might be considerable amount of existing data to provide an adequate initial crack rate estimate. We also expect that as the crack monitoring procedure obtains more data at higher fatigue cycles, error in the initial crack rate estimate will have less effect upon overall estimates.

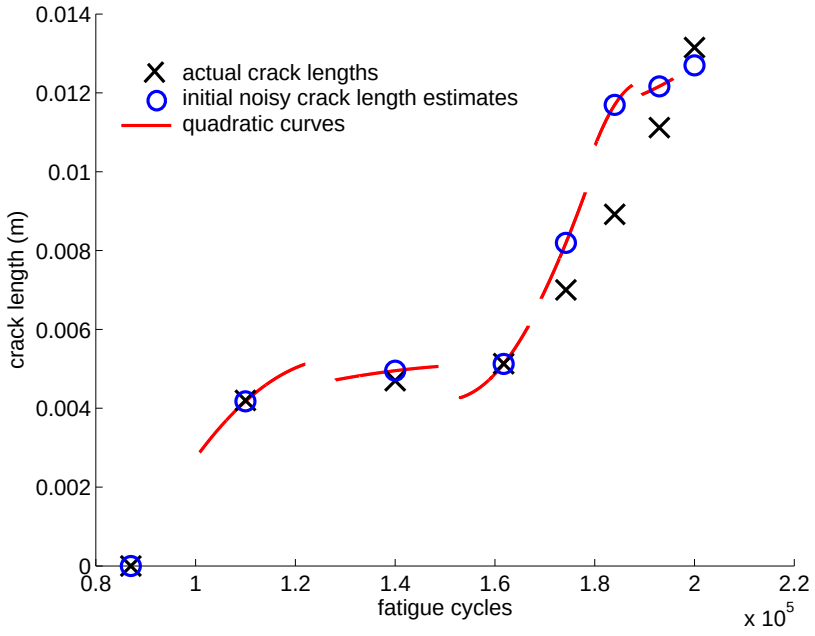


Figure 5.11: Quadratic curves created at crack length estimates using the estimate and its two neighboring estimates.

After obtaining localized estimates for crack rate at each crack length, the cubic model shown in Figure 3.5 was used to obtain estimates for ΔK_{eff}, and a least squares approach was used to construct a linear, crack rate versus $\log(\Delta K_{\text{eff}})$ equation of the form shown in (3.4). After defining the this linear relationship, smoothed values for the log crack rate and crack rate can be obtained. Using the smoothed crack rate values, a smoothed version of the crack length versus fatigue cycles curve was constructed.

Figures 5.12-5.14 show smoothed crack growth curves for different levels of crack growth. Given there have been k crack length estimates, and excluding the first crack estimate, crack rate estimates were obtained at the previous $k-2$ crack lengths using the quadratic curve, tangent approach of (5.7) and (5.8). A linear relationship of the form shown in (3.4) was created using all $k-2$ crack rates. This relationship was used to obtained smoothed values of the crack rate for 2 to $k-1$ crack lengths. Given these smoothed values at each crack length, the smoothed crack growth curves shown

in Figures 5.12-5.14 were created.

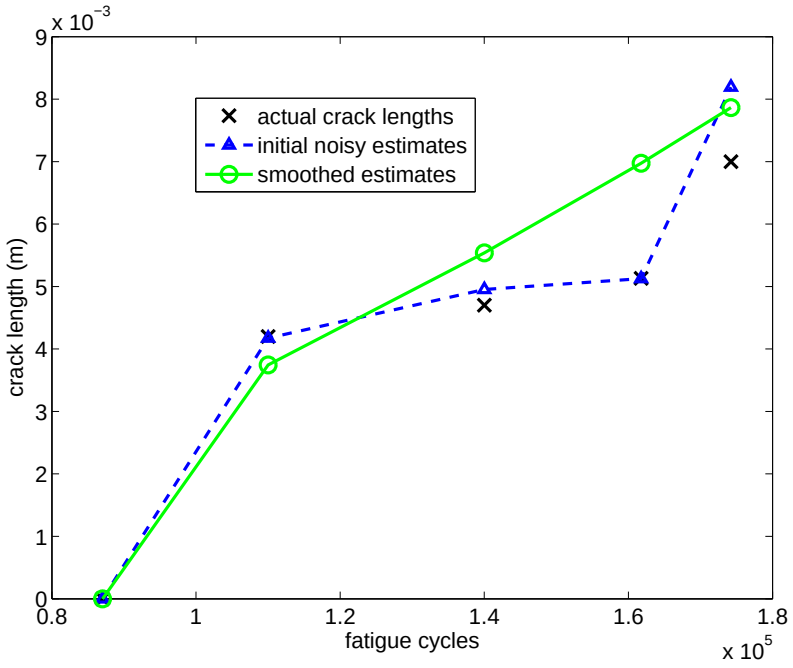


Figure 5.12: Smoothed crack growth monitoring for $k = 5$ crack length estimates.

Variations of this smoothing approach are possible. For example, the procedure for creating a base polynomial curve at each crack length could be modified. In Figure 5.15, four crack length estimates were used to create the quadratic curves of (5.7). These consisted of the crack length of interest, the preceding crack estimate, and the following two crack estimates. (For the $k-1$ crack, since there was only one following crack length (the kth crack), the kth crack was included twice in creating the quadratic curve.) Since the crack growth rate tends to increase at higher fatigue cycles, the inclusion of more future crack length data adjusted the tangent crack rate estimates accordingly.

Although the crack rate versus $\log(\Delta K_{\text{eff}})$ relationship is generally assumed to be linear within the Stage II region of crack propagation, this linear relationship can

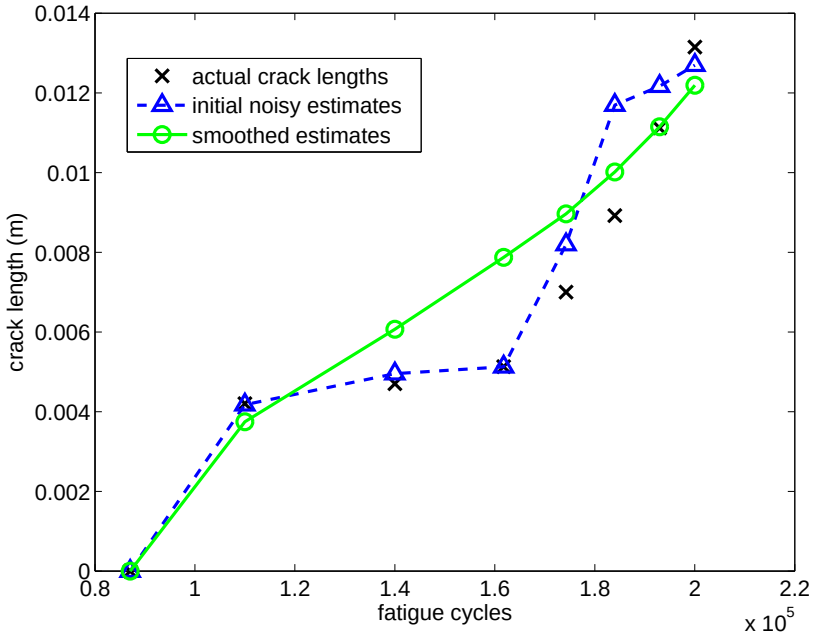


Figure 5.13: Smoothed crack growth monitoring for $k = 8$ crack length estimates.

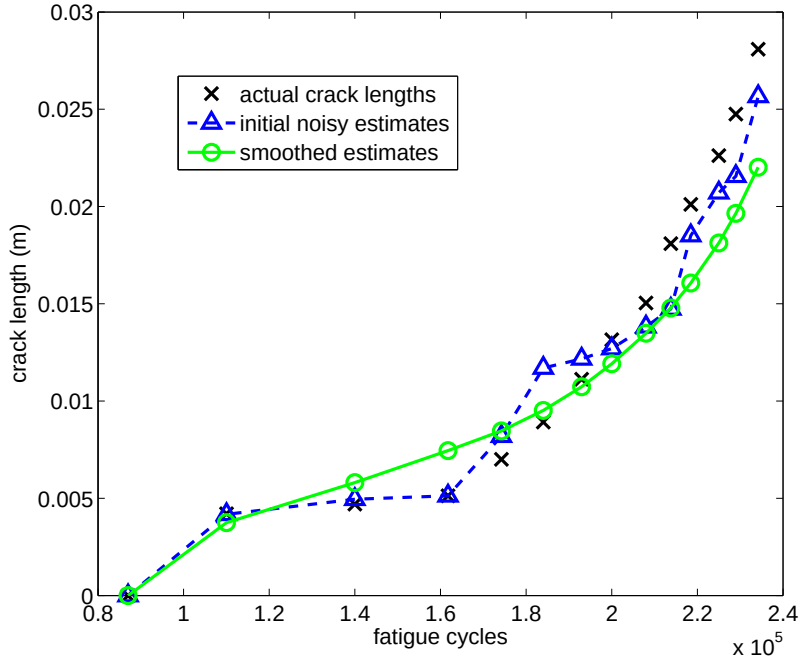


Figure 5.14: Smoothed crack growth monitoring for $k = 14$ crack length estimates.

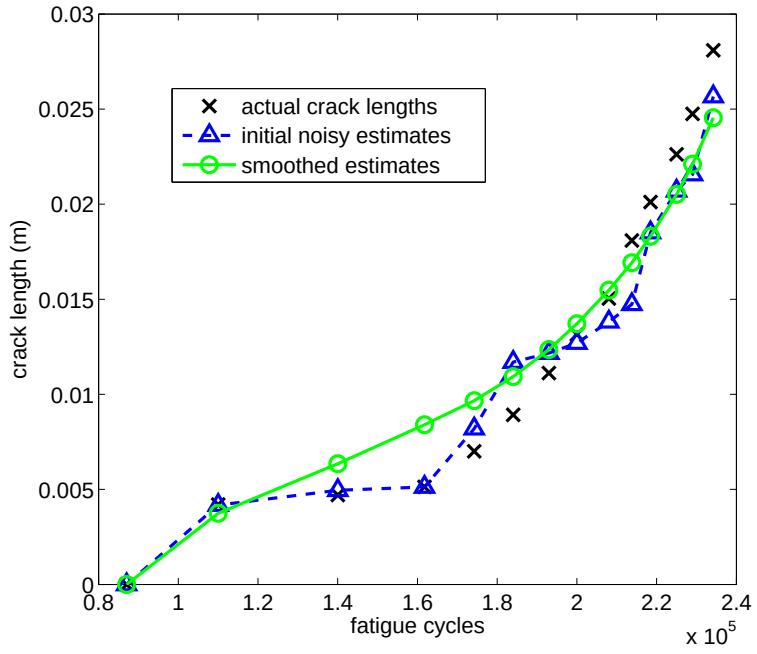


Figure 5.15: Smoothed crack growth monitoring using four crack length estimates to obtain preliminary estimates of localized (da/dn) .

shift as the crack length increases and changes occur within the loading geometry of a structure. Rather than use the entire collection of crack rate tangent estimates obtained at all previous crack lengths to define the linear relationship, it might be advantageous to use only a limited number of the most recent crack rate tangent estimates to define this relationship. In Figure 5.16, the linear relationship was defined using a maximum of the 5 most recent crack rate tangent estimates. Since crack rate smoothing only incorporated a limited number of the most recent crack length information, the smoothed curve reflected more of the localized changes within the crack growth.

5.3 Implementation Issues and Alternative MSE Metric

Although the predicted MSE estimate shown in (5.3) provides an excellent metric for selecting an optimal resolution observation mode (e.g, an optimal transmission waveform), implementing this approach can be computationally expensive; therefore, for our future work, we propose an alternative, approximation approach which signifi-

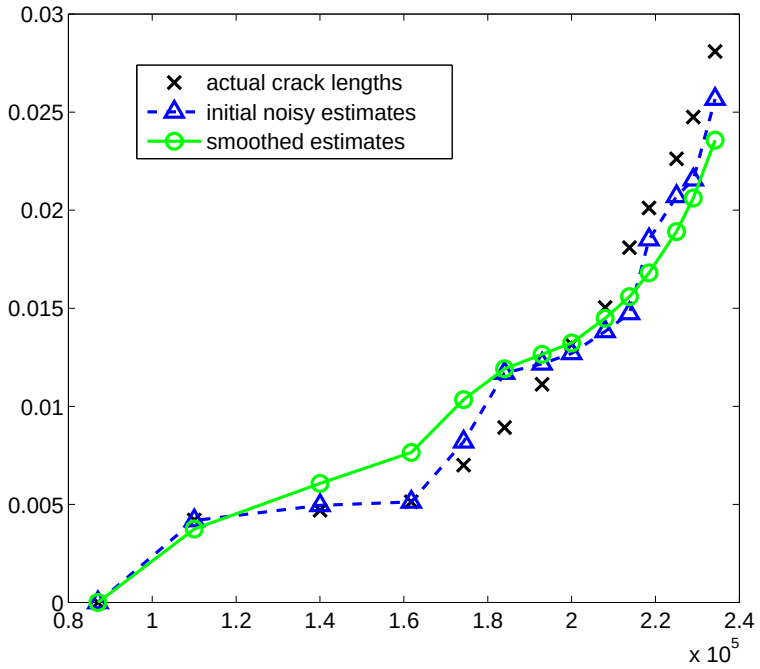


Figure 5.16: Smoothed crack growth monitoring using a maximum of the 5 most recent, preliminary crack rate estimates to define the linear relationship between crack rate and $\log(\Delta K_{\text{eff}})$.

cantly reduces the computational complexity.

Note that within (5.3), there are N_{p} particles and there are $\mathscr{M}$ predicted observations generated for each particle. For each of the $\mathscr{M}$ predicted observations, there are L MPD iterations that have to be performed to generate a single observation, where each observation consists of a L-length MPD sequence of 4-dimensional MPD vectors. After generating the L-length vector sequence observation, the lot likelihood of this observation has to be obtained for each of D HMMs, where each HMM corresponds to one of the D training crack lengths. The log likelihoods of the observations, given the D training crack lengths, are used to determine a value of the predicted crack length value in (5.3). Therefore, for *each possible observation mode*, the generation of a L-length predicted feature vector sequence and analysis of the vector sequence has to be performed $N_{\text{p}} \times \mathscr{M}$ times.

This high level of computational complexity has significant effects when implementing our crack growth tracking method. Considering we wish to choose an observation mode that possesses maximum crack length resolution, and since each of the possible observation modes has to be analyzed, there are inherent restrictions upon the number of different observation modes that can be used to obtain data. When we evaluate the resolution of different observation modes, we generate hypothetical, predicted MPD observations for each particle in our crack length estimation particle filter, there are practical limits on the number of particles that we can use in the particle filter, as well as practical limits upon the number of predicted MPD observations that we can generate for each particle. These limits could affect statistical accuracy in estimating crack lengths and in evaluating damage resolution characteristics of the different observation modes. Furthermore, since the predicted MPD feature vector sequences have to be processed through a HMM corresponding to each of D damage conditions, and where each MPD feature vector sequence contains L feature vectors, this could put practical restrictions on the number of different damage conditions D that can be evaluated and the number of MPD iterations L that can be used to decompose a signal into feature vectors.

Alternatively, we propose using an approximation to (5.3) that significantly reduces the computational burden and accomplishes the primary objective of generating a real-time, MSE-based metric for selecting the observation mode. Rather than use the particle value $x_k^{(i,j)}$ in (5.3), we round $x_k^{(i,j)}$ to the nearest of the D crack lengths used to train the D HMMs with the jth waveform, where we designate this discretized, approximation of the particle value as $\tilde{x}_{k,\lambda_d}^{(i,j)}$. With this approximated version of the particles, (5.3) becomes

$$\mathrm{MSE}(z_k^j)) \approx \sum_i^{N_\mathrm{p}} w_k^{(i)} \left\{ \frac{1}{\mathscr{M}'} \sum_{m=1}^{\mathscr{M}'} \left[\tilde{x}_{k,\lambda_d}^{(i,j)} - \hat{x}_{k,m}^{(i,j)} \right]^2 \right\} \tag{5.9}$$

With this approach, the quantity in brackets that multiples the weight $w_k^{(i,j)}$ in (5.9)

represents an approximate, predicted MSE value for a single, discretized crack length damage condition d, given the jth transmission waveform was used to make the observations. This quantity can be created offline for all the HMMs, where each HMM represents a different combination of crack length damage condition d and observation mode j. Since these calculations are performed offline and do not have a relationship with a fatigue cycle time step index k or a particle index (i), we can remove the references to these indices and rewrite the approximate predicted MSE as

$$\frac{1}{\mathscr{M}'}\sum_{m=1}^{\mathscr{M}'}\left[\tilde{x}_{\lambda_d}^{(j)}-\hat{x}_{k,m}^{(i,j)}\right]^2. \tag{5.10}$$

Since we can perform the approximated MSE calculations offline for each HMM λ_d and its associated crack length, this can significantly reduce the real-time computational burden. Performing calculations offline, a very large number $\mathscr{M}'$ of predicted MPD feature vectors can be generated using the procedure described in Section 5.2.1. The predicted MPD vectors are analyzed as described in Section 5.1.2, resulting in the predicted crack lengths $\hat{x}_{k,m}^{(i,j)}$. Then, using the crack length $\tilde{x}_{\lambda_d}^{(j)}$ corresponding to the HMM training crack length λ_d, the approximated MSE in (5.10) can be calculated for each individual HMM λ_d. These individual MSE values for each HMM can be stored in a lookup table according to the crack length damage index d and transmission waveform j.

When implementing our crack growth monitoring procedure, given we are at fatigue cycle time step k and determining a predicted (approximated) MSE value for waveform j, for each particle index (i), we obtain the value of the particle and find the HMM λ_d corresponding to the nearest HMM training crack length value $\tilde{x}_{\lambda_d}^{(j)}$. We then use the look-up table to find the corresponding, pre-calculated MSE value for λ_d. With this approach, calculation of the approximate MSE metric per (5.9) becomes a relatively simple, weighted summation of individual, pre-calculated MSE values from the look-up table that are weighted by the corresponding particle weight.

Although the calculation of the predicted MSE metric using (5.9) is an approximation to the MSE value obtained using (5.3), our primary interest is to obtain a relative metric that can be obtained efficiently for selecting an optimal observation mode, rather than to determine the most accurate value for the predicted MSE. Furthermore, there are numerous realtime computational advantages obtained when using the approximate MSE method, where most of the computational complexity issues discussed above for the original MSE calculation (5.3) are reduced considerably within the realtime implementation of the approximation MSE approach. With the approximation approach, it is significantly more practical to have larger numbers of available observation modes, to have significantly higher number of particles to perform damage estimation, as well as to maintain or increase the number of possible damage states and number of MPD iterations used to decompose the received signals.

5.3.1 Approximate Pre-calculated Approach For Waveform Selection

Figure 5.17 shows the MSE results using the pre-calculated, approximation method of (5.9), and these results can be compared with the real-time approach of (5.3), which were shown previously in Figure 5.8. Note that the pre-calculated, approximation approach produced MSE values that were consistently lower than those obtained using the realtime approach, although the waveform selection criteria is based upon relative MSE values between the waveforms, rather than the actual MSE values. The pre-calculated MSE approach rounds the particle value to the nearest HMM training crack length, and hence, excludes the error between the actual particle value and its nearest HMM training crack length. However, if these errors are assumed to be zero mean distributed, the general MSE relationships between the waveforms will be similar whether the realtime or pre-calculated approximation approach is used. This is evident when comparing the two figures. In both figures, starting at low crack lengths and moving to higher levels of damage, the linear chirp and then the pseudo random width pulse waveforms exhibited the lowest MSE values. In the middle of the crack length range,

the 230 kHz tone burst waveform exhibited high resolution characteristics with respect to possessing the lowest MSE characteristics. At higher crack lengths, the linear chirp exhibited the lowest MSE characteristics.

In the pre-calculated, approximation MSE approach, since the error between particle value and HMM crack length is excluded, only a finite number of possible crack lengths (i.e., the HMM training crack lengths) are employed when using this calculation approach. Since the crack lengths are constrained to a finite number of possibilities, the values of the associated errors are also constrained, and there is a possibility several predicted MSEs corresponding to different waveforms will have equal MSE values, or as Figure 5.17 shows, there is a possibility some waveforms will exhibit a zero-valued MSE. Note that log-spacing was used in this figure, and any zero-valued MSEs were located at a bottom horizontal axis (labeled ZERO).

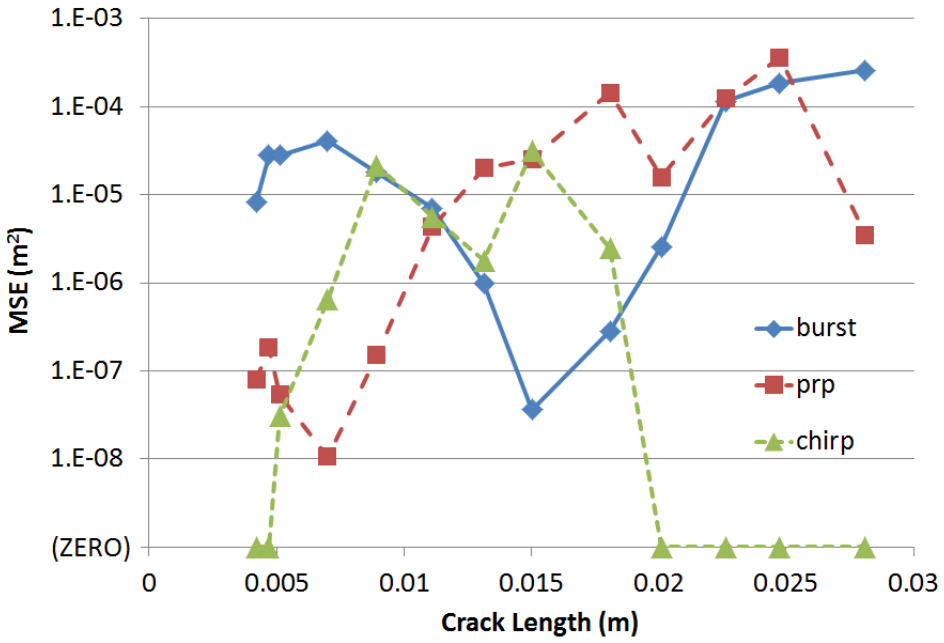


Figure 5.17: Predicted MSE using pre-calculated approach.

The real-time efficiency of the pre-calculated, approximation method of (5.9) allows greater number of samples to be used in creating estimates. As discussed above,

with the pre-calculated, approximate MSE approach, the MSE for an individual particle can easily be obtained as a pre-calculated, stored MSE value corresponding to the nearest HMM training crack length. This is in contrast to the real-time MSE approach, where the MSE is based upon multiple predicted observations, and evaluations of the predicted observations, that have to be performed real-time for each particle. (Technically, in the real-time MSE approach, it could be possible to have more particles in the actual particle filter but then use only a sampling of the particles to obtain an MSE estimate. In this discussion, we are assuming that all particles in the particle filter are used to calculate the overall MSE value for the particle set.) In the pre-calculated, approximation MSE example, $N_{\mathrm{p}} = 5000$ particles were used in the particle filter tracking. However, in the real-time approach, it was not practical to use that many particles, and only $N_{\mathrm{p}} = 100$ particles were used. In addition, even though the pre-calculated, approximation approach generates an approximation for individual particle MSEs, since the calculations are performed offline, significantly more samples can be used to generate the approximate MSE values. In the real time approach shown in Figure 5.8, the procedure described in 5.2.1 had to be performed $\mathscr{M}$ times, in real time, for each particle, and only $\mathscr{M} = 50$ samples were generated for each particle. With the pre-calculated, approximation approach, since calculations were performed offline, and since there was only a defined, finite number of HMM training crack lengths, $\mathscr{M}' = 2,000$ samples were generated for each of the HMM training crack lengths.

5.3.2 General Observation Mode Selection

Although the observation mode selection procedure can be employed to evaluate and select optimal transmission waveforms to improve the robustness of damage observations, other observation parameters could also be included within a more general, observation mode selection procedure. For example, in addition to three different transmission waveforms, the cruciform specimen shown in Figure 5.1 and Figure 5.3 had five different receiver sensors. Therefore, in addition to selecting one or more transmission

waveforms to obtain an observation, the particular receiver sensor(s) used to obtained the information could also be included as part of the observation mode. With multiple available modes of making the crack length observation, this could further improve the robustness of the overall crack length monitoring method.

With three different transmission waveforms and five different receiver sensors, fifteen possible waveform/sensor combinations were available. Furthermore, as discussed above, with a finite number of possible crack lengths, there would be a good possibility that some of the combinations would exhibit equal predicted MSE values. This could include situations where several of the waveform/sensor combinations exhibit the lowest, most optimal MSE value. Furthermore, since the MSE represents a measure of crack length resolution, a threshold MSE value could be set using a defined acceptable level of resolution. Observations could then be made using the miminum MSE waveform/sensor combinations, but also using waveform/sensor combinations possessing MSE values below the defined threshold value.

In a demonstration of this concept was performed, in which an acceptable MSE threshold value of 10^{-10} was used, such that all waveform/sensor combinations exhibiting the lowest predicted MSE value, as well as any waveform/sensor combinations exhibiting a MSE value below the threshold value, were used to obtain the crack length observations. Figures 5.18-5.30 show plots of the MSEs for all five sensors, with all three transmission waveforms, obtained at each of the thirteen different particle filter time step iterations. The resulting crack length tracking results are shown in Figure 5.31. As this figure shows, the overall crack tracking method exhibited excellent capability to track the actual crack length.

An important concept that can be seen by the collection of figures is that no single sensor, single waveform, or single waveform/sensor combination produced the optimal MSE for all levels of fatigue crack damage. Since the predicted MSE values

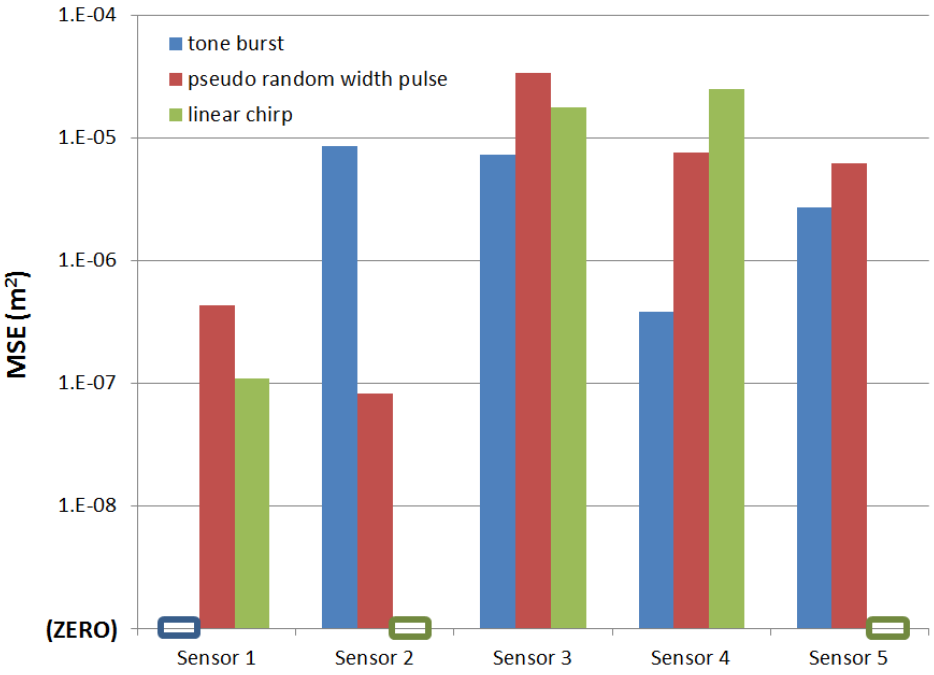


Figure 5.18: Approximation MSEs at fatigue cycle 110009 (actual crack length 4.2 mm)

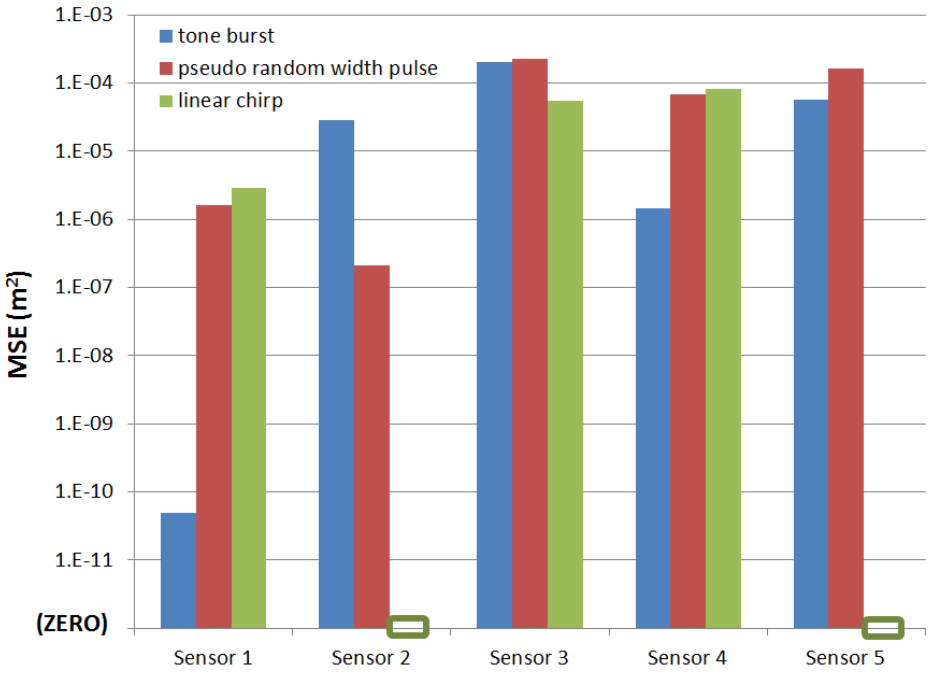


Figure 5.19: Approximation MSEs at fatigue cycle 140011 (actual crack length 4.7 mm).

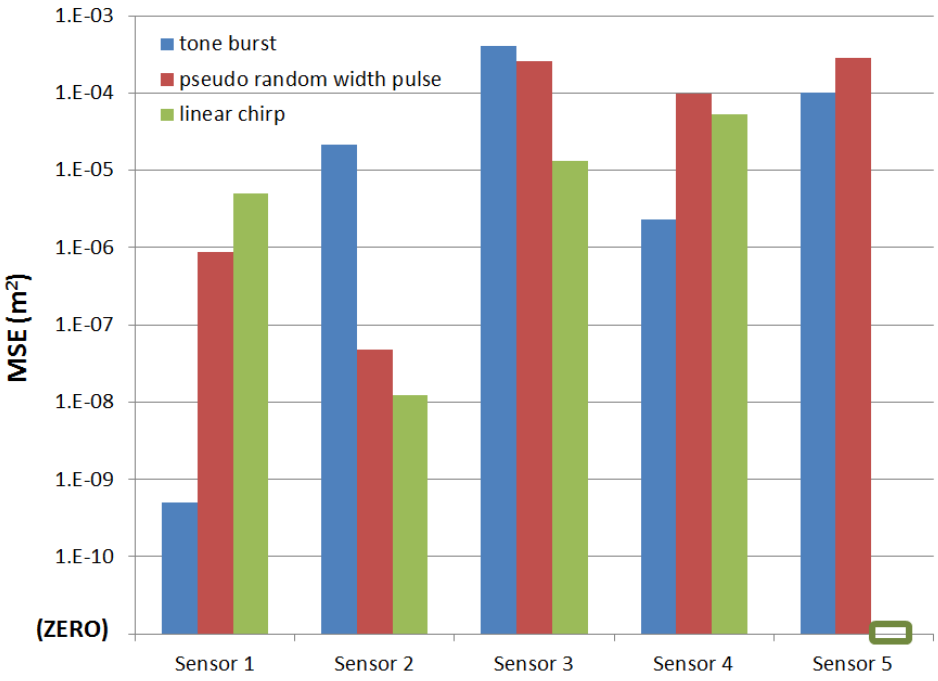


Figure 5.20: Approximation MSEs at fatigue cycle 161761 (actual crack length 5.13 mm).

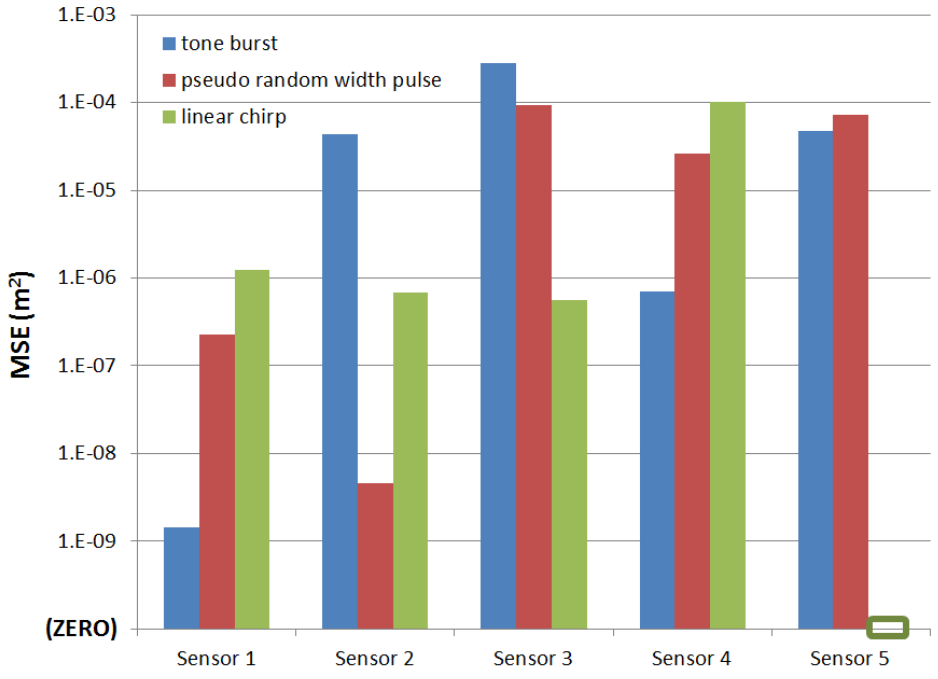


Figure 5.21: Approximation MSEs at fatigue cycle 174238 (actual crack length 7.0 mm).

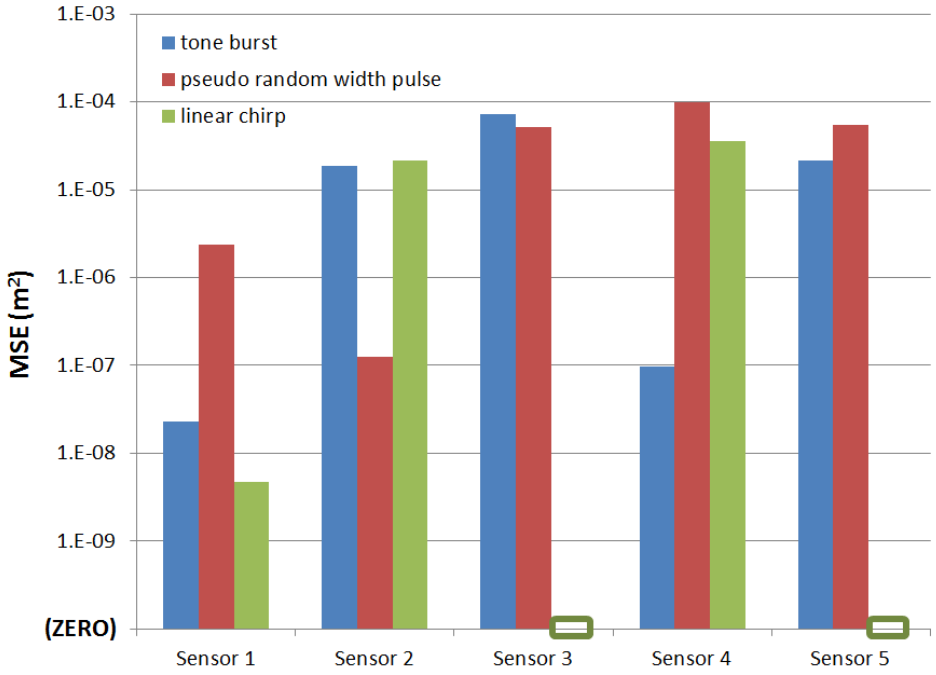


Figure 5.22: Approximation MSEs at fatigue cycle 184009 (actual crack length 8.92 mm).

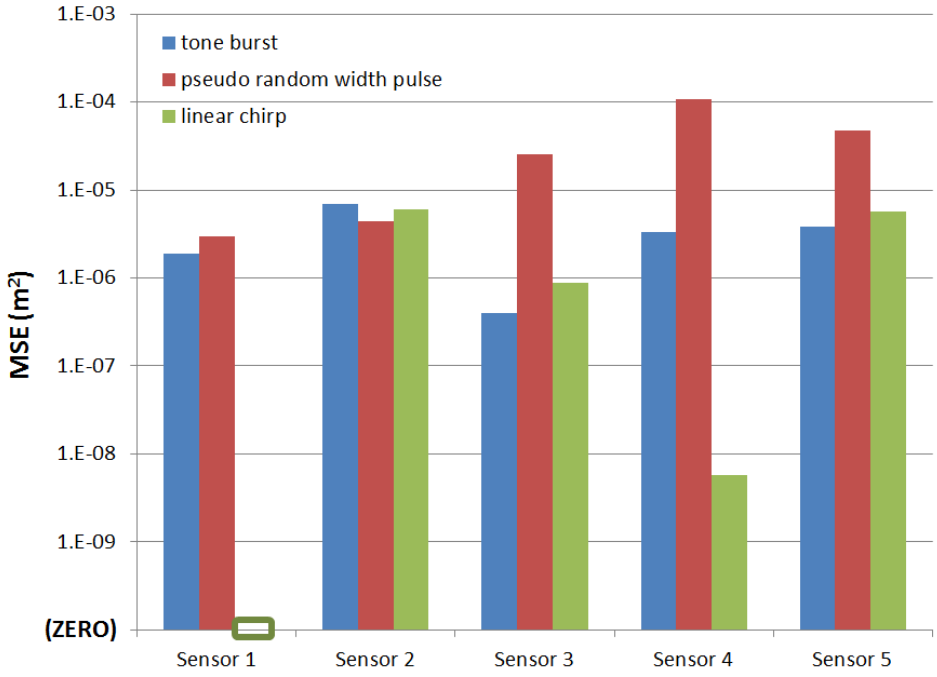


Figure 5.23: Approximation MSEs at fatigue cycle 192961 (actual crack length 11.12 mm).

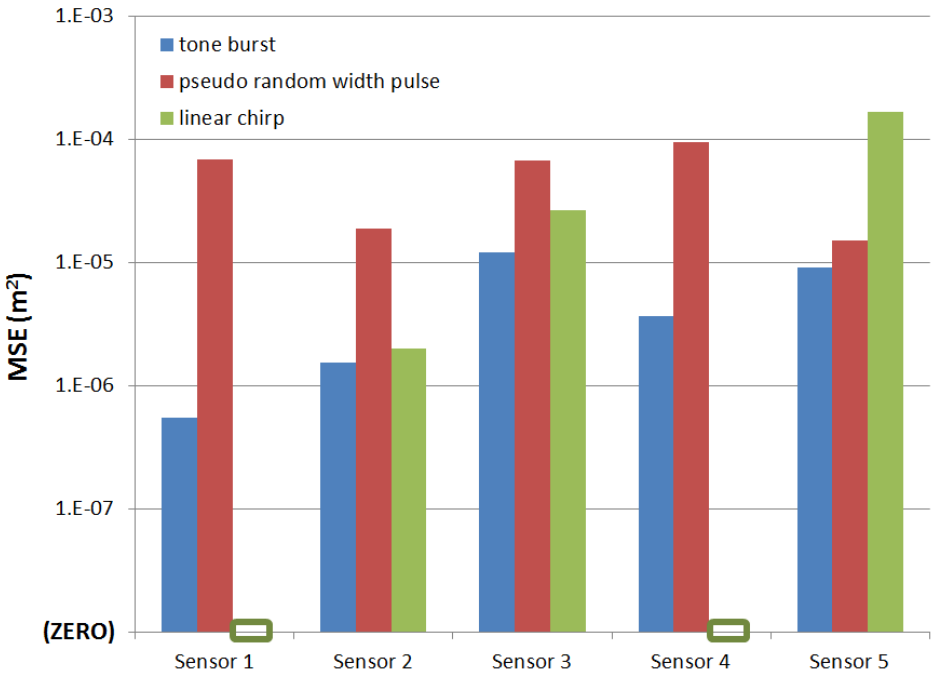


Figure 5.24: Approximation MSEs at fatigue cycle 200013 (actual crack length 13.15 mm).

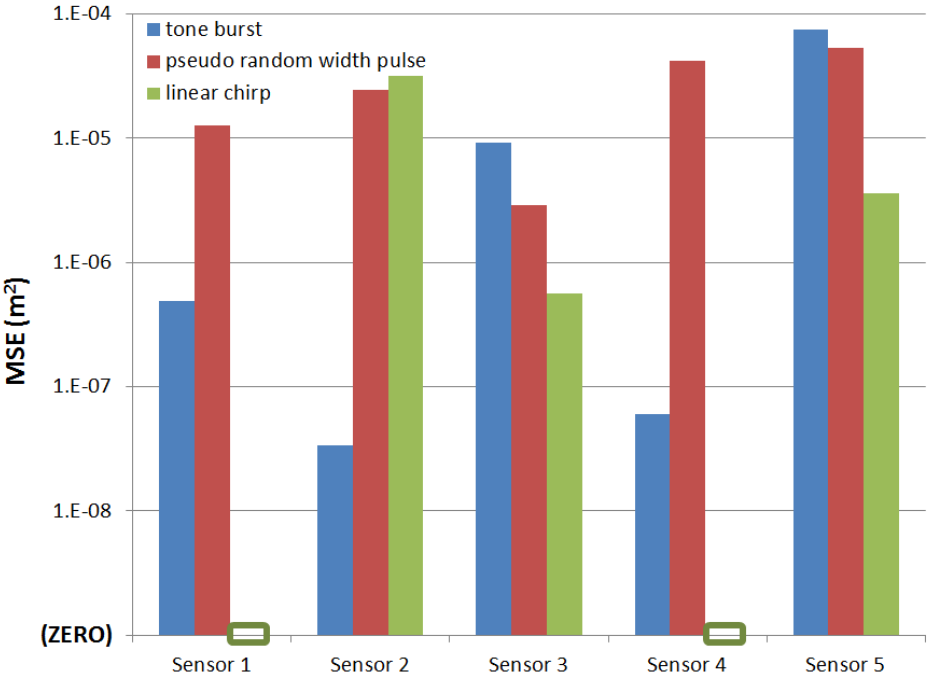


Figure 5.25: Approximation MSEs at fatigue cycle 208012 (actual crack length 15.04 mm).

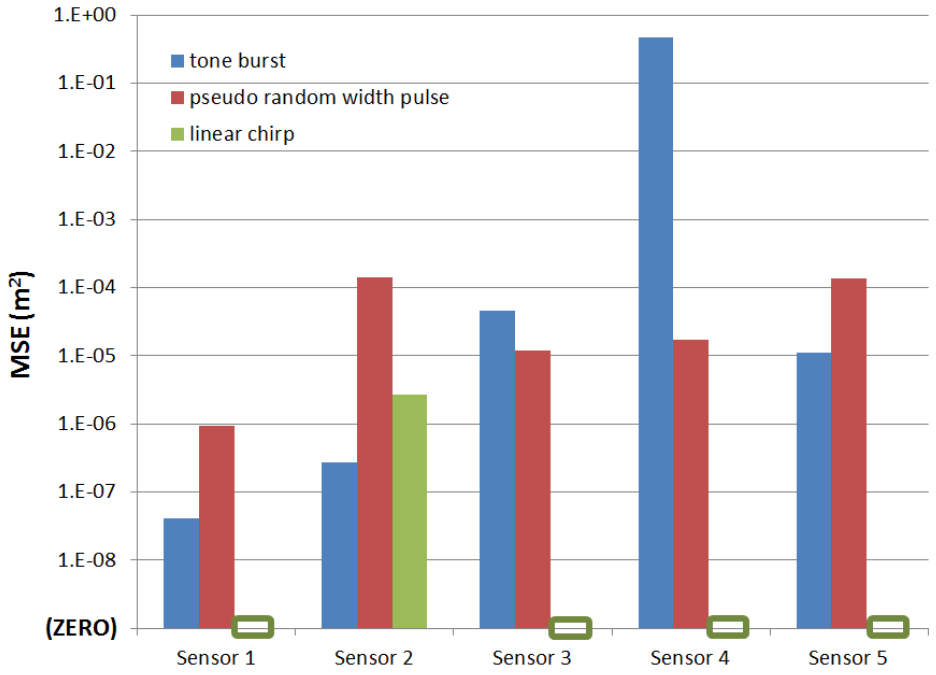


Figure 5.26: Approximation MSEs at fatigue cycle 213794 (actual crack length 18.09 mm).

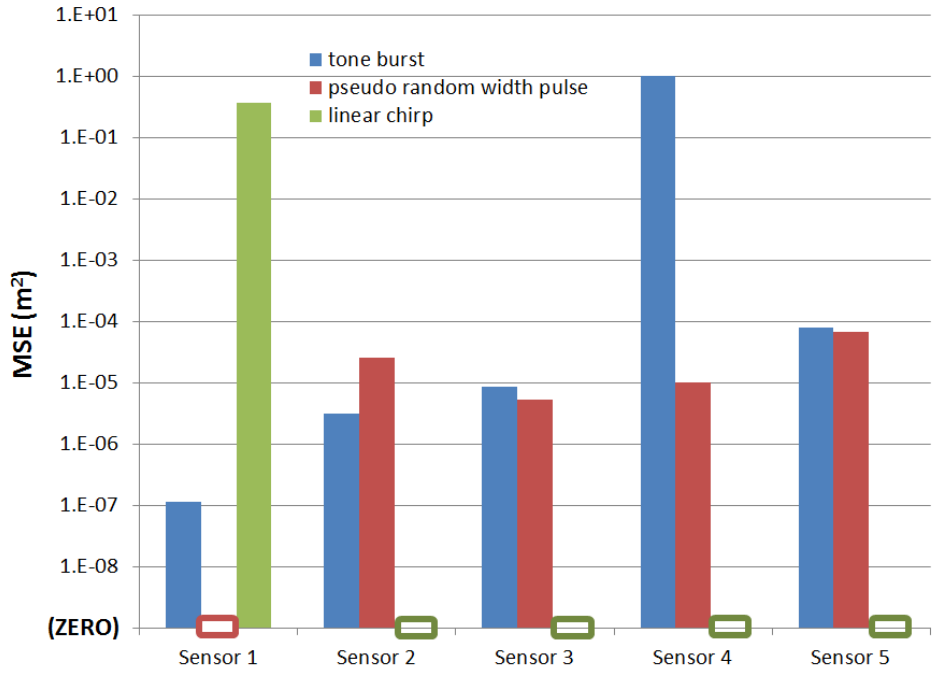


Figure 5.27: Approximation MSEs at fatigue cycle 218462 (actual crack length 20.11 mm).

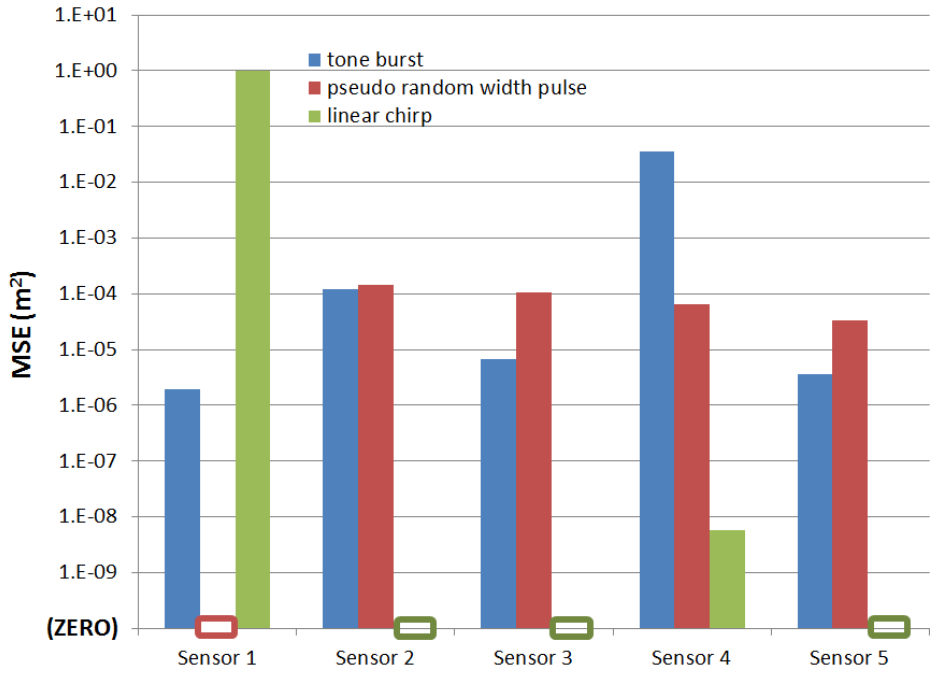


Figure 5.28: Approximation MSEs at fatigue cycle 225008 (actual crack length 22.62 mm).

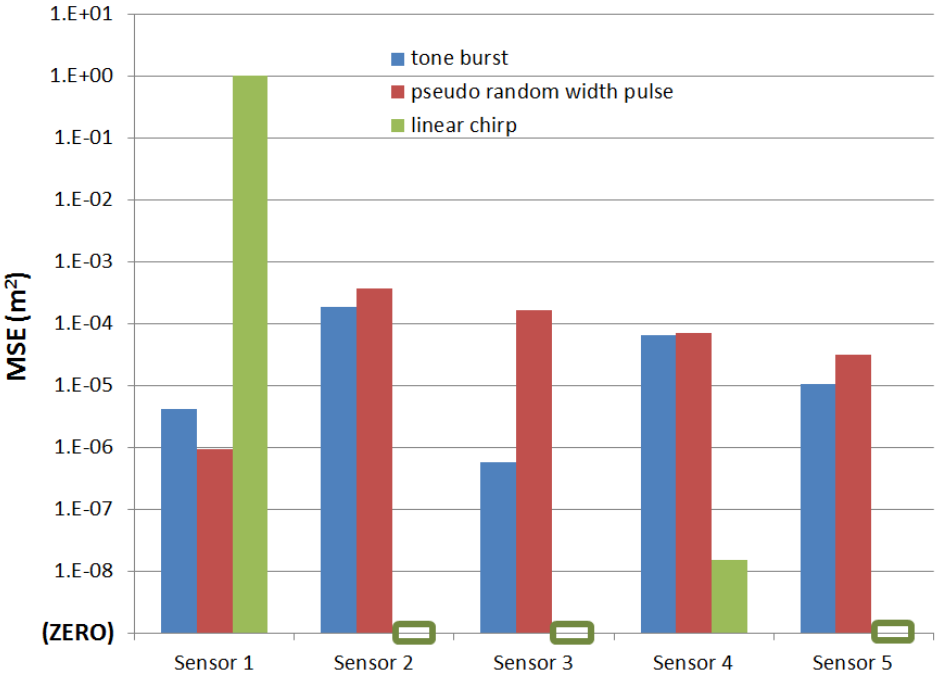


Figure 5.29: Approximation MSEs at fatigue cycle 229009 (actual crack length 24.75 mm).

Figure 5.30: Approximation MSEs at fatigue cycle 234216 (actual crack length 28.09 mm).

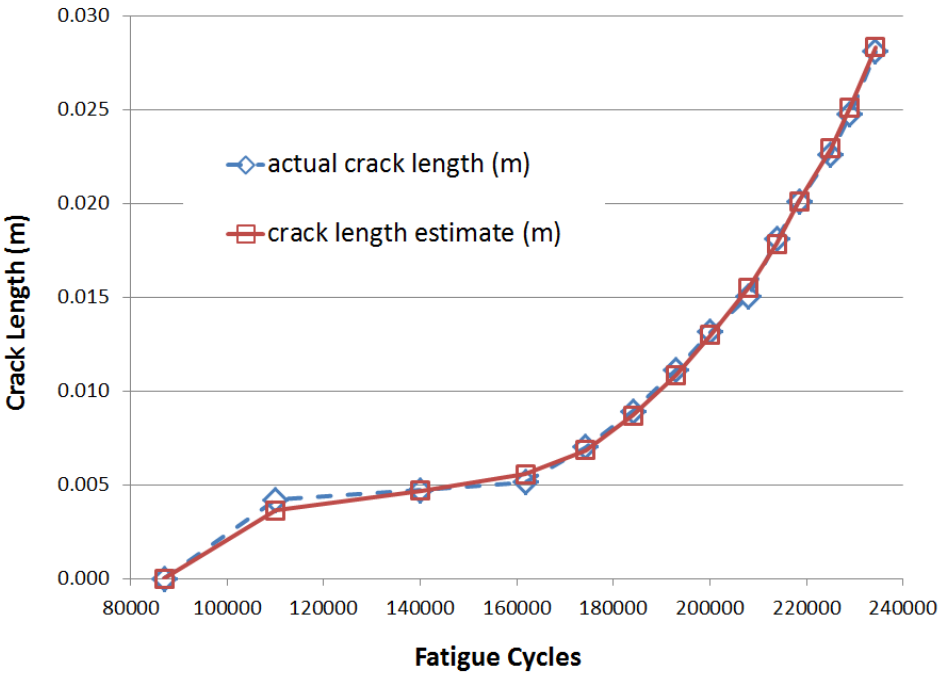


Figure 5.31: Overall crack growth tracking using selection of multiple waveform and sensor combinations).

are an indication of the crack length resolution capabilities of the different observation modes, the ability to assess and then choose the optimal observation parameter combinations can increase the robustness of the overall crack growth monitoring method.

Chapter 6

CONCLUSIONS AND FUTURE WORK

6.1 Summary of Proposed Work

We develop a sequential Bayesian, state-space approach, integrated with two main contributions, that exhibits high accuracy in tracking damage crack growth. Given a fatigue loaded structure with a posterior distribution of the crack length at an existing fatigue cycle, our first contribution is a predictive, state dynamics model that estimates a prior distribution of the crack length at the next fatigue cycle step. Based upon values within the prior distribution, our second contribution is the selection of an optimal transmission waveform to be used in obtaining a crack length observation. After the observation is obtained, it is used to update the prior distribution to obtain the posterior distribution at the next designated fatigue cycle step. At each step, the crack length estimate is obtained as the mean of the posterior distribution.

Within the predictive, state dynamics model, we can incorporate an interacting multiple model (IMM) for situations in which multiple loading conditions are possible. The IMM offers at least two key benefits. First, IMM improves the ability to track the crack growth damage, since state dynamics models are adjusted appropriately to changes in loading conditions. Second, IMM specifically tracks changes in loading conditions, and this basic loading information can be useful from operational, maintenance, and other standpoints.

We also demonstrate the use of Markov chain Monte Carlo (MCMC) methods for constraining state dynamics parameters, given additional knowledge, desired values, or general constraints regarding the parameters. In our demonstration, this was performed off-line on training data such that parameters of state dynamics models could be adjusted prior to real time, crack growth monitoring. However, we also note that MCMC could be used in a real time mode to create constraints in temporary situations

involving significant observation noise and extreme outlier observation results.

As the crack length classification results of Tables 5.5-5.7 show, we were able to extract key signal features within joint time-frequency space using the matching pursuit decomposition (MPD) and use continuous-observation hidden Markov models (HMMs) to learn and identify features corresponding to different damage conditions.

As the minimum mean-squared error (MSE) estimation results of Figure 5.8 show, we can assess the waveform resolution capability at different levels of crack damage by optimally selecting the transmission waveform to impart robustness when obtaining observations at different damage levels. We integrate these various concepts within an overall particle filtering methodology, and as Figure 5.9 shows, the particle filter performs exceptionally well in tracking the crack damage growth. Furthermore, the waveform selection procedure improved the basic ability to track crack growth within enhanced noisy environments. Even though the resulting crack length estimates exhibited significant noise, using the assumption of $log(da/dn)$ vs. $log(\Delta K_{\text{eff}})$ liinearity for Stage II crack propagation, we demonstrate how the da/dn crack growth rates associated with the noisy estimates can be smoothed using various post-processing techniques.

Using a particle value approximation approach to reduce real time computational burden in calculating MSEs, we extend the waveform selection concept to a more general, observation mode selection procedure. Since most of the computational burden is performed off-line, we demonstrate how multiple waveforms, multiple sensors, and possibly other observation parameters, can be included in the observation mode selection method.

6.2 Possible Future Work Extensions

For future work, we intend to explore several more approaches towards damage monitoring, such as expanding the crack growth monitoring procedure described within this report, as well investigation of non-parametric Bayesian approaches.

A key contribution within this work was the waveform selection procedure, as discussed in Chapter 5. During our Chapter 4 state dynamics studies with IMM, although our primariy focus related to multiple loading conditions and multiple state dynamics models, we only had a single 135 KHz tone burst waveform available in these studies. Considering the successful result of our Chapter 5 waveform selection studies, it became apparent it would have been useful if we had more than one waveform available in the IMM studies. Although the MPD-HMM observation model produced acceptable results, as shown by the confusion matrix plot in Figure 4.7, in the IMM studies, the crack length data corresponded to approximately 30 different crack lengths and varying levels of observation error were observed. Furthermore, in our Chapter 5 waveform selection studies, as shown by the tone burst MSE characteristics in Figures 5.8 and 5.17, a single, constant-frequency, tone-burst waveform can exhibit different crack length resolution characteristics.

In the waveform selection studies described in Chapter 5, we focused upon different *types* of transmission waveforms, since the objective was to study crack tracking in terms of different transmission waveform *types*. The tone burst represented a simple, well-defined waveform possessing short duration and small bandwidth. The linear chirp waveform also represented a well-defined waveform, but with somewhat longer duration and wider bandwidth. The pseudo random width pulse represented an ill-defined, noisy, medium duration waveform. Although this collection of different types of waveforms produced interesting results, for future work, the overall waveform selection procedure could be simplified.

For example, rather than using different types of waveforms, it might be easier, and possibly of greater value, to use multiple versions of the same waveform type, where one or possibly several waveform (or other observation mode) parameters were systematically varied at several different levels. After obtaining training data for different crack lengths using the different waveform variations, the MSE characteristics for

each waveform variation and crack length could be obtained. With this information, it might be possible to create a model for the predicted MSE as a function of the observation mode parameters. If this was achievable, then given a predicted crack length, the model cold be used to predict which waveform (or other observation mode) parameter values should be chosen to obtain maximum predicted resolution. With this capability, the overall waveform *selection* procedure becomes somewhat of a simple, data-driven, waveform *design* approach.

We also intend to explore non-parametric Bayesian approaches. In the supervised learning approaches discussed in this report, many of the quantities, structures, and procedures were predetermined within the damage monitoring method. Although the various parameters and groupings were defined by a pre-examination of the data, this approach inherently forces all data to conform to predetermined parameters and groupings. Real data does not always conform within a limited number of predetermined data structures or parameters, especially in situations where new types of data might be encountered, where the new data types might not readily conform to existing classes and parameters. In non-parametric Bayesian approaches, data is allowed to define groupings and other structures within the data.

The Dirichlet process (DP) and the beta process (BP) represent two examples of non-parametric Bayesian approaches. With the DP, the DP mixture model provides a stochastic description for data with an unbounded number of mixture components or latent clusters. In this non-parametric Bayesian framework, there is no restrictive assumption that the appropriate number of clusters are finite or known *a priori*. In particular, the number of clusters is learned from the data and can change as new data is available. However, in many problems the data are not always well-clustered. Depending on the mixture distributions used, the clusters can sometimes overlap and a single data point may belong to multiple clusters at the same time. As an alternative to the DP based latent clustering model, the BP provides a non-parametric Bayesian framework

where each measurement can be assigned not to only one cluster (class) but rather to a subset of clusters (attributes or features). Unlike the DP, which induces a prior distribution over partitions of the data (latent clusters), the BP imposes a prior over infinite binary matrices whose entries indicate collections of attributes for the data (latent features). Since the number of possible clusters representable using m features is 2^m, this also addresses the issue of data overfitting – fewer parameters are needed as compared to the DP. The DP and BP based analyses of data can be performed efficiently using MCMC and variational methods.

REFERENCES

[1] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, 1989.

[2] W. Zhou, N. Kovvali, A. Papandreou-Suppappola, P. Peralta, and A. Chattopadhyay, "Progressive damage estimation using sequential Monte Carlo techniques," *7th International Workshop on Structural Health Monitoring*, 2009.

[3] K. Worden and J. M. Dulieu-Barton, "An overview of intelligent fault detection in systems and structures," *Structural Health Monitoring*, vol. 3, pp. 85–98, 2004.

[4] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philosophical Transactions of the Royal Society Series A*, vol. 365, pp. 303–315, 2007.

[5] D. C. Montgomery, *Statistical Quality Control*, 7th ed. Hoboken, NJ: Wiley, 2012.

[6] W. J. Staszewski, C. Boller, and G. R. Tomlinson, *Health Monitoring of Aerospace Structures*. Chichester, West Sussex, England, 2004.

[7] C. R. Farrar and K. Worden, *Structural Health Monitoring - A Machine Learning Approach*. Chichester, West Sussex, UK: John Wiley and Sons Ltd, 2013.

[8] V. Giurgiutiu, *Structural Health Monitoring with Piezoelectric wafer Active Sensors*. Burlington, MA: Elsevier Academic Press.

[9] P. C. Paris, M. P. Gomez, and W. E. Anderson, "A rational analytic theory of fatigue," *The Trend in Engineering (University of Washington)*, vol. 13, pp. 9–14, 1961.

[10] P. Paris and F. Erdogan, "A critical analysis of crack propagation laws," *Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers*, vol. 12, pp. 528–534, 1963.

[11] C. Luo, J. Wei, M. Parra-Garcia, A. Chattopadhyay, and P. Peralta, "Fatigue damage prediction in metallic materials based on multiscale modeling," *AIAA Journal*, vol. 47, pp. 2567–2576, 2009.

[12] S. Mohanty, A. Chattopadhyay, and P. Peralta, "Adaptive residual useful life estimation of a structural hotspot," *Journal of Intelligent Material Systems and Structures*, vol. 21, pp. 321–335, 2010.

[13] Y. Liu, S. Mohanty, and A. Chattopadhyay, "Condition based structural health monitoring and prognosis of composite structures under uniaxial and biaxial loading," *Journal of Nondestructive Evaluation*, vol. 29, pp. 181–188, 2010.

[14] S. Mohanty, A. Chattopadhyay, P. Peralta, and D. Quech, "Fatigue damage prognosis of a cruciform structure under biaxial random and flight profile loading," *Proceedings of SPIE, Nondestructive Characterization for Composite Materials, Aerospace Engineering, Civil Infrastructure, and Homeland Security*, vol. 7649, 2010.

[15] S. Mohanty, A. Chattopadhyay, and J. N. Rajadas, "Dynamic strain mapping and real-time damage-state estimation under random fatigue loading," *AIAA Journal*, vol. 50, pp. 769–777, 2012.

[16] S. S. Kessler, S. M. Spearing, and C. Soutis, "Damage detection in composite materials using Lamb wave methods," *Smart Materials and Structures*, vol. 11, pp. 269–278, 2002.

[17] H. Sohn, G. Park, J. R. Wait, N. P. Limback, and C. R. Farrar, "Wavelet-based active sensing for delamination detection in composite structures," *Smart Materials and Structures*, vol. 13, pp. 153–160, 2004.

[18] T. R. Fasel, C. C. Olson, and M. D. Todd, "Optimized guided wave excitation for health monitoring of a bolted joint," *Proceedings of SPIE*, vol. 6935, 2008.

[19] B. C. Lee and W. J. Staszewski, "Modeling of Lamb waves for damage detection in metallic structures, part i: Wave propagation; part ii: Wave interactions with damage," *Smart Materials and Structures*, vol. 12, pp. 804–814, 2003.

[20] ——, "Lamb waves for damage detection in metallic structures, part i: Two dimensional analysis; part ii: Damage monitoring strategy," *Smart Materials and Structures*, vol. 16, pp. 249–274, 2007.

[21] H. Kuttig, M. Niethammer, S. Hurlebaus, and L. J. Jacobs, "Model-based analysis of dispersion curves using chirplets," *Journal of the Acoustic Society of America*, vol. 11, pp. 269–278, 2002.

[22] A. Raghavan and C. E. S. Cesnick, "Guided-wave signal processing using chirplet matching pursuits and mode correlation for structural health monitoring," *Smart Materials and Structures*, vol. 16, pp. 355–366, 2007.

[23] L. Borkowski, K. Liu, and A. Chattophadyay, "Fully coupled electromechanical elastodynamic model for guided wave propagation analysis," *Journal of Intelligent Material Systems and Structures*, 2013.

[24] E. Figueiredo, L. Radu, R. Westgate, J. Brownjohn, E. Cross, K. Worden, and C. Farrar, "Applicability of of a Markov chain Monte Carlo method for damage detection on data from the z-24 and Tamar suspension bridges," *European Workshop on Structural Health Monitoring*, 2012.

[25] S. M. Ross, *Simulation*, 3rd ed. Academic Press, San Diego, CA, 2002.

[26] S. Das, A. Papandreou-Suppappola, X. Zhou, and A. Chattopadhyay, "On the use of the Matching Pursuit Decomposition signal processing technique for structural health monitoring," vol. 5764, pp. 583–594, 2005.

[27] D. Chakraborty, N. Kovvali, J. Wei, A. Papandreou-Suppappola, D. Cochran, and A. Chattopadhyay, "Damage classification structural health monitoring in bolted structures using time-frequency techniques," *Journal of Intelligent Material Systems and Structures*, vol. 20, pp. 1289–1305, 2009.

[28] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, 1993.

[29] W. Zhou, D. Chakraborty, N. Kovvali, A. Papandreou-Suppappola, D. Cochran, and A. Chattopadhyay, "Damage classification for structural health monitoring using time-frequency feature extraction and continuous hidden Markov models," *Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, pp. 848–852.

[30] W. Zhou, N. Kovvali, A. Papandreou-Suppappola, and A. Chattopadhyay, "Sensor optimization for progressive damage diagnosis in complex structures," *Proceedings of SPIE*, vol. 7650, 2010.

[31] H. A. P. Blom, "An efficient filter for abruptly changing systems," *Proceedings of IEEE 23rd Conference on Decision and Control*, pp. 656–658, 1984.

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

[33] E. Mazor, A. Averbuch, Y. Bar-Shalom, , and J. Dayan, "Interacting multiple model methods in target tracking: A survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, pp. 103–123, 1998.

[34] S. P. Sira, A. Papandreou-Suppappola, and D. Morrell, *Advances in Waveform-Agile Sensing for Tracking*. Morgan and Claypool, 2009.

[35] T. Bayes and R. Price, "An essay towards solving a problem in the doctrine of chances," *Philosophical Transactions of the Royal Society of London*, vol. 53, pp. 370–418, 1763.

[36] G. E. P. Box and G. C. Tao, *Bayesian Inference in Statistical Analysis*. Addison-Wesley, 1973.

[37] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubino, *Bayesian Data Analysis second edition*. Chapman and Hall CRC, 2004.

[38] C. P. Robert, *The Bayesian Choice, 2nd ed.* Springer-Verlag, 2001.

[39] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*, 2nd ed. Springer-Verlag, 1997.

[40] M. Chen, Q. Shao, and J. G. Ibrahim, *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag, 2000.

[41] W. Gilks, S. Richardson, and D. Spiegelhalter, Eds., *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1996.

[42] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, pp. 1087–1092, 1953.

[43] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[44] L. Tierney, "Markov chains for exploring posterior distributions (with discussions)," *The Annals of Statistics*, vol. 22, pp. 1701–1702, 1994.

[45] S. Chib and E. Greenberg, "Understanding the metropolis-hastings algorithm," *The American Statistician*, vol. 49, pp. 327–335, 1995.

[46] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences," *Statistical Science*, vol. 7, pp. 457–511, 1992.

[47] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 3rd ed. Springer-Verlag, 2004.

[48] M. K. Cowles and B. P. Carlin, "Markov chain Monte Carlo convergence diagnostics: A comparative review," *Journal of the American Statistical Association*, vol. 91, pp. 883–904, 1996.

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

[50] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.

[51] L. A. McGee and S. F. Schmidt, "Discovery of the kalman filter as a practical tool for aerospace and industry," *National Aeronautics and Space Administration, Technical Memorandum 86847*, 1985.

[52] D. S. Lee and N. K. K. Chia, "A particle algorithm for sequential parameter estimation and model selection," *IEEE Transactions on Signal Processing*, vol. 50, pp. 326–336, 2002.

[53] D. Gabor, "Theory of communication," *Journal of IEE (London)*, vol. 93, pp. 429–457, 1946.

[54] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, 1992.

[55] L. Cohen, *Time-Frequency Analysis*. Upper Saddle River, New Jersey: Prentice-Hall.

[56] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. Academic Press, 1999.

[57] A. Papandreou-Suppappola and S. B. Suppappola, "Analysis and classification of time-varying signals with multiple time-frequency structures," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 92–95, 2002.

[58] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state markov chains," *The Annals of Mathematical Statistics*, vol. 37, pp. 1554–1563, 1966.

[59] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *The Annals of Mathematical Statistics*, vol. 41, pp. 164–171, 1970.

[60] L. Rabiner and B. Juang, *Fundamentals of Speech Recognition*. Prentice Hall, Englewood Cliffs, NJ, 1993.

[61] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. Prentice Hall, Englewood Cliffs, NJ, 1978.

[62] A. A. Griffith, "The phenomena of rupture and flow in solids," *Philosophical Transactions of the Royal Society, Series A*, vol. 221, pp. 163–198, 1921.

[63] G. Irwin, "Fracture mechanics," in *Fracturing of Metals*. Cleveland, OH: American Society for Metals, 1948, pp. 163–198.

[64] G. R. Irwin, "Analysis of stresses and strains near the end of a crack traversing a plate," *Journal of Applied Mechanics*, vol. 79, pp. 361–364, 1957.

[65] J. A. Bannantine, J. J. Comer, and J. L. Handrock, *Fundamentals of Metal Fatigue Analysis*. Prentice Hall, Pearson Education, Upper Saddle River, NJ, 1990.

[66] S. Suresh, *Fatigue of Materials*, 2nd ed. New York, NY: Cambridge University Press, 1998.

[67] M. Janssen, J. Zuidema, and R. J. H. Wanhill, *Fracture Mechanics*, 2nd ed. The Netherlands: VSSD, 2002.

[68] R. J. Sanford, *Principles of Fracture Mechanics*. Upper Saddle River, NJ: Prentice Hall Pearson Education Inc., 2003.

[69] G. C. Sih, *Handbook of Stress Intensity Factors*. Bethlehem, PA: Institute of Fracture and Solid Mechanics, Lehigh University, 1973.

[70] S. Soni, "Damage detection, localization, and quantification in lug joints, MURI annual review meeting," *MURI SHM and Prognosis, DOD MURI, Grant FA95550-06-1*, May 2010.

[71] F. Erdogan, "Stress intensity factors," *Journal of Applied Mechanics Transactions of the American Society of Mechanical Engineers*, vol. 50, pp. 992–1002, December 1983.

[72] J. N. Yang, S. D. Manning, J. L. Rudd, and M. E. Artley, "Probabilistic durability analysis methods for metallic airframes," *Probabilistic Engineering Mechanics*, vol. 2, pp. 9–15, 1987.

[73] J. N. Yang and S. D. Manning, "Stochastic crack growth analysis methodologies for metallic structures," *Engineering Fracture Mechanics*, vol. 37, pp. 1105–1124, 1990.

[74] ——, "A simple second order approximation for stochastic crack growth analysis," *Engineering Fracture Mechanics*, vol. 53, pp. 677–686, 1996.

[75] W. F. Wu and C. C. Ni, "A study of stochastic fatigue crack growth modeling through experimental data," *Probabilistic Engineering Mechanics*, vol. 18, pp. 107–118, 2003.

[76] C. Willhauck, S. Mohanty, A. Chattopadhyay, and P. Peralta, "Stochastic crack growth under variable loading for health monitoring and prognosis."

[77] Y. Boers and J. N. Driessen, "Interacting multiple model particle filter," *IEE Proceedings on Radar, Sonar and Navigation*, vol. 150, pp. 344–349, 2003.

[78] R. K. Neerukatti, K. Liu, N. Kovvali, and A. Chattopadhyay, "Fatigue life prediction using hybrid prognosis for structural health monitoring," *AIAA Journal of Aerospace Information Systems*.