Machine Learning Based Prediction and Classification of Properties of Engineered Composite Materials

By

Mason Alan Hickman

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Civil Engineering

May 8, 2020

Nashville, Tennessee

Approved:
Prodyot Basu, D.Sc.
Çağlar Oskay, Ph.D.
Sankaran Mahadevan, Ph.D.
Douglas Adams, Ph.D.
D. Mitchell Wilkes, Ph.D

Copyright © 2020 by Mason A. Hickman

All Rights Reserved

ACKNOWLEDGMENTS

First, I would like to thank Vanderbilt University, where I have had the pleasure of studying for the past decade. It has been inspiring to work in an environment surrounded by such intelligent, creative colleagues from diverse backgrounds. I am grateful to the university for providing funding for my graduate education as well as my employment as a teaching and research assistant. I am also thankful to the Department of Civil and Environmental Engineering faculty and staff.

I am especially grateful to my advisor, Dr. Basu, for his guidance and mentorship throughout my graduate studies. I would also like to thank Drs. Oskay, Mahadevan, Adams, and Wilkes for serving on my Ph.D. committee. The skills I learned from their courses and counseling were crucial in my development as an engineer and researcher.

I would like to thank my fellow graduate students, notably Niko Nordendale, William Heard, and Matthew Pike, with whom I had the pleasure of collaborating on research over the years.

I am grateful to Rich Teising for his assistance constructing the experimental guitar rig detailed in this dissertation. I am also thankful to Giuseppe Del Giudice for his advice on designing and controlling the test rig, and to Scott Marquart for providing samples for the experiments.

Last but certainly not least, I am grateful to my family and friends for their support throughout my graduate career. The pursuit of this degree was a mentally and emotionally challenging process and I could not have succeeded without their encouragement and understanding.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

1 Introduction 1

  1.1 Motivation 1

    1.1.1 Short-Fiber Reinforced Composites 2

    1.1.2 Composite Musical Instrument Strings 3

  1.2 Literature Review 5

    1.2.1 Short-Fiber Reinforced Composites 6

    1.2.2 Composite Musical Instrument Strings 9

  1.3 Objectives 13

  1.4 Dissertation Organization 14

2 Machine Learning of Structure-Property Relationships 16

  2.1 The Descriptor-Based Machine Learning Approach 17

    2.1.1 Descriptor Selection and Sampling 20

    2.1.2 Random Process Evaluation/Data Acquisition 24

    2.1.3 The Learning Machine 25

3 Prediction of Random Short-Fiber Reinforced Composite Properties 28

  3.1 Characteristic Response of FRC 31

  3.2 Multiscale Modeling of SFRC 34

    3.2.1 Macroscale Model 36

    3.2.2 Microscale Model 39

3.3 DBML Implementation 47
  3.3.1 Descriptor Selection and Sampling 47
  3.3.2 Random Process Evaluation/Data Acquisition 49
  3.3.3 The Learning Machine 53
3.4 Conclusion 57

4 Vibrational Response of Composite Musical Instrument Strings 59
4.1 Governing Equations 62
  4.1.1 String Vibration 62
  4.1.2 Signal Analysis 68
4.2 Musical Problem Background 72
  4.2.1 Tuning and Timbre 72
  4.2.2 The Electric Guitar System 74
4.3 DBML Implementation 80
  4.3.1 Descriptor Selection and Sampling 80
  4.3.2 Random Process Evaluation/Data Acquisition 81
  4.3.3 The Learning Machine 92
4.4 Conclusion 103

5 Summary, Conclusions, and Future Work 105

Appendices 110

A SFRC Descriptor Space Sample Points 110

B Guitar String Damping Coefficients 114

REFERENCES 118

LIST OF TABLES

Table Page

2.1 2-phase random short-fiber reinforced composite descriptors and response metrics. 22

2.2 Wound musical instrument string descriptors and response metrics. 22

3.1 SFRC descriptor variable ranges for numerical models. 47

3.2 Phase material property constants for numerical models. 51

3.3 Composite response parameters measured in each simulation. 53

3.4 Comparison of error magnitudes obtained from GP regression of each output variable. 57

4.1 Standard guitar tuning and string configurations. 76

4.2 Description of gauge 46 roundwound sample classes used for experiment. . . 81

4.3 Ranking of string classes by stiffness. 89

4.4 SVM training data organization. 97

4.5 Accuracy (%) of SVM on four classification problems using different feature sets for training, $n = 12$. 100

A.1 SFRC sample points, $\mathbf{x}_i, i = 1...20$. 110

A.2 SFRC sample points, $\mathbf{x}_i, i = 21...40$. 111

A.3 SFRC model results, $\mathbf{x}_i, i = 1...20$. 112

A.4 SFRC model results, $\mathbf{x}_i, i = 21...40$. 113

B.1 Mean A_n for each class. 114

B.2 Coefficient of variation of A_n for each class. 115

B.3 Mean α_n for each class. 116

B.4 Coefficient of variation of α_n for each class. 117

LIST OF FIGURES

Figure Page

2.1 Example 2-phase microstructure optimization problem. 17

2.2 Input/output space view of the descriptor approach with respect to (a) the SFRC problem, and (b) the composite guitar string problem. 19

2.3 System diagram of supervised learning framework. 20

3.1 (a) Isotropic aligned, (b) anisotropic aligned, and (c) random microstructures with interparticle spacing s. 30

3.2 Random SFRC microstructure realizations with $\lambda = 100$ mm, $\phi_f = 0.1\%$, $l_f = 3.27$ mm. 30

3.3 ACK model composite response phases. 32

3.4 Two types of fiber surface deformities: crimping (top) and twisting (bottom). 33

3.5 Example pullout test data for a ductile fiber in a cementitious matrix. 33

3.6 Local variation of fiber volume content in a random SFRC domain. 35

3.7 Proposed sequential multiscale modeling framework for SFRC. 36

3.8 CDP model constitutive law. 39

3.9 Visualization of (a) strain discontinuity enrichment function and (b) debonding enrichment function for an arbitrarily placed fiber in a 5x5 mm element domain. 43

3.10 2-dimensional feature space visualization of LHS for SFRC. 48

3.11 RSA algorithm overview. 49

3.12 Three random realizations of sample points (a) $\mathbf{x}_{20}$ and (b) $\mathbf{x}_{32}$ obtained using RSA algorithm. 50

3.13 XFEM problem domain and boundary conditions. 51

3.14 Stress-strain curve with elastic and post-peak region defined by parameters ε_i, ε_f, and σ_i. 53

3.15 Sample tensile stress-strain curves for 20 realizations of 2 random points $\mathbf{x}_{20}$ and $\mathbf{x}_{32}$. 54

4.1 Comparison of geometries of roundwound strings with circular and hexagonal cores. 60

4.2 Surface profiles of gauge 46 roundwound strings with (a) iron-cobalt alloy wrap and (b) pure nickel wrap. 61

4.3 String vibration problem formulation. 63

4.4 String vibration boundary conditions. 64

4.5 Harmonic content for three different plucking locations. 65

4.6 Theoretical inharmonicity due to string stiffness. 67

4.7 Musical note frequencies and intervals. 73

4.8 2-second time signal of note C4, f_1 =261.6 Hz, plucked on (a) guitar and (b) cello. Frequency spectrum obtained from FFT for (c) guitar and (d) cello. 74

4.9 Typical electric guitar configuration. 76

4.10 Typical guitar string boundary conditions at the (a) nut and (b) bridge. . . . 76

4.11 Comparison of vibrational response measured from (a) experiment and (b) finite element model of a plain round ASTM A228 wire. 80

4.12 (a) Experimental system diagram and (b) laboratory setup. 84

4.13 Spectrograms of a sample from each string class measured from S1. Plots (a)-(g) correspond to Classes A-G. 86

4.14 Spectrograms of a sample from each string class measured from S2. Plots (a)-(g) correspond to Classes A-G. 87

4.15 Average measured inharmonicity from S1. Plots (a)-(g) correspond to Classes A-G. 90

4.16 Average measured inharmonicity from S2. Plots (a)-(g) correspond to Classes A-G. 91

4.17 Time-dependent amplitude of first five partials measured from S1 for each string class. Plots (a)-(g) correspond to Classes A-G. 93

4.18 Time-dependent amplitude of first five partials measured from S2 for each string class. Plots (a)-(g) correspond to Classes A-G. 94

4.19 Support vector machines for (a) binary classification and (b) multiple classes. 96

4.20 Experimental data for seven classes plotted in two arbitrary dimensions of feature space, A_2 and A_8. 98

4.21 SVM prediction error using all features, $\mathbf{x}_i^{(3)}$, with respect to number of partials used for four problems, (a) String Class, (b) Core Geometry, (c) Core Material, (d) Wrap Material. 101

4.22 SVM confusion matrices using all features for four problems: (a)-(b) String Class, (c)-(d) Core Geometry, (e)-(f) Core Material, (g)-(h) Wrap Material. The left column corresponds to S1 data and the right column corresponds to S2 data. 102

Chapter 1

Introduction

1.1 Motivation

Well-designed composites can effectively utilize the desirable attributes of their constituent phases, resulting in a material greater than the sum of its parts. Structural composites may be optimized to realize performance which meets some demand threshold for the structure. The ideal composite material for a structure may be selected based on characteristics like desired stiffness, strength-weight ratio, thermal or acoustical conductivity, fatigue life, or some combination of these. The influence of the characteristics of the constituents and the interfaces at small scales on the continuum scale response of the composite is well understood. The material properties of the composite treated as a continuum are controlled by the behavior of individual phases and interactions thereof. All physical systems, natural or engineered, are prone to some degree of inherent uncertainty. Uncertainty is present at all scales, from the atomic interactions of a material microstructure to the demand imposed on the system at the macroscale. Design of composites with engineered microstructures requires characterization of uncertainties to assess the reliability of material performance in expected scenarios, so that the design life of the structure is predictable within a desired level of accuracy.

The main purpose of this research is to formalize and implement a computational framework for the analysis and design of arbitrary composite material microstructures to optimize the desired structural performance of the composite. This will enable intelligent tailoring of a mix of material phases of different types, shapes, and sizes satisfying the needs of an application in terms of performance, cost, and ease of manufacture. Key to the approach is the use of machine learning concepts for identifying patterns in higher-dimensional design spaces based on data from numerical simulations and experiments.

The framework presented in this research is implemented for two problems from apparently widely different scenarios, described in Sections 1.1.1 and 1.1.2. The first scenario is concerned with the prediction of mechanical properties of short-fiber reinforced composites with randomly distributed fibers, widely used in many engineering applications. The second problem focuses on the vibration of metallic composite wires used in stringed musical instruments. In each of these two problems, the objective is to identify descriptors that can be used to characterize a composite morphology, and determine the correlation between descriptors and the overall performance of the composite. Once a machine learning model has been successfully trained using data from experiments and/or simulations, it can be used to efficiently simulate the response of composites with microstructures that have not been previously tested, allowing for data-based rapid prototyping of composites in a variety of applications.

1.1.1 Short-Fiber Reinforced Composites

Fiber reinforced composites (FRC) generally consists of short fibers, long fibers, or some combination of the two embedded in a lower strength matrix material. Both short fibers and long fibers show advantages for different applications. Fiber inclusions enable crack bridging, show enhanced ductility, and absorb energy without significantly increasing the weight of the composite. For fibers continuous and straight throughout the material, it is possible to derive simple phenomenological mechanical property relationships for the matrix-fiber combination in the form of constitutive equations. It is also possible to estimate the failure strength of adequately reinforced aligned long fiber composites using a mechanics-based approach. A unidirectional FRC lamina is orthotropic in nature. The toughness is maximum in the direction of the fibers, but relatively weak in a direction transverse to the fibers. For this reason, unidirectional continuous fibers are useful when the direction of the load with respect to the material orientation is known with some degree of certainty. For example, some aircraft components use laminates with layers of uniformly oriented fibers designed to specifically handle the unique load profiles imposed under operating conditions.

In some applications, randomly dispersed short fibers may be preferred due to reduced directional dependence of strength because irrespective of the load direction some of the fibers will always share a portion of the load. For example, in blast-resistant structural design, the location and magnitude of the design load is often unknown. Consequently, moment reversal may occur in a supporting beam. The toughness of a brittle-matrix composite can be significantly improved by adding a small volume fraction ($<2\%$) of randomly dispersed higher-strength ductile or brittle short fibers to the matrix. In addition to improved mechanical properties, random short-fiber reinforced composites (SFRC) provide an advantage in ease of fabrication, as the fibers can simply be added to the matrix as it is mixed. In the case of SFRC, apart from the strength of the matrix material, its degree of ductility can be a function of volumetric ratio, physical properties of fiber like diameter, strength, length, distribution, elastic modulus, and the interfacial mechanical and chemical bond characteristics of the two. Random SFRC exhibits variation in morphology throughout the problem domain that must be accounted for in the design. In practical applications, the designer must select fiber shapes and sizes to achieve optimal performance. This dissertation addresses the problem of predicting the response of a random SFRC microstructures with varying fiber configurations based on state-of-the art methods for material modeling and machine learning.

1.1.2 Composite Musical Instrument Strings

Stringed instruments produce a signal when a transverse excitation force is applied to a tensioned string or wire by a plucking, bowing, or striking mechanism. In acoustic instruments, sound is produced when the wire vibrations coupled with the instrument body create a pressure wave in the surrounding air. Rather than creating pressure waves in the air, the vibrations of the metal string on an electric guitar are converted into electrical impulses using a magnetic pickup. The signal is transmitted to an amplifier, which is capable of producing a signal significantly more powerful than that of an acoustic instrument. The pitch that the instrument produces is related to the length, density, and stiffness of the string material, as well as the magnitude of the tension in the string. The effect of each of these factors must be considered when selecting materials to construct an instrument string.

Composite electric guitar strings consist of a tensioned core wire helically wrapped with an additional wire that is used to control the fundamental frequency of vibration by manipulating the overall density and stiffness of the composite. Most string producers use either round or hexagonal core wires wrapped with another round, semicircular, or rectangular metallic wire. A variety of materials are used for core and wrap wires. Limited data exists to directly compare composite strings of different materials and geometry, since information such as ideal core and wrap wire sizes and precise alloy composition are studied by corporations in-house and treated as proprietary information. The strings are marketed to users using subjective, relative terms such as “maximum toughness” and “improved brightness”. Occasionally a manufacturer will provide information such as a frequency response curve indicating the expected sound of a given set of strings, but limited information is given regarding the experimental basis of the results. There is a lack of published literature on the vibrational response of strings with different geometry and material property combinations. One contribution of the current research is to analyze various strings using the same series of experiments, to mitigate industrial bias. The development of a method to optimize string geometries to obtain a desired frequency response and design life is of interest to manufacturers seeking to modernize their production technology. This research will compare the performance of commercially available instrument strings with varying geometries and material properties in an unprecedented manner.

1.2 Literature Review

Generally, composites may be classified as fibrous composites, particulate composites, and laminated composites, or some combination of the three [1]. For example, a laminate may consist of some layup with particulate inclusions and another layup with fiber inclusions. Historically composite materials have been characterized on the basis of strength tests on representative specimens, resulting in a number of empirical relationships that have been used by engineers [2]. Such empirical characterization methods may be a viable option for relatively inexpensive bulk materials subjected to quasi-static loads, but the design of new-age composites for a broad range of applications can be time consuming and expensive tests may not be an acceptable option. Early phenomenological attempts to model the behavior of heterogeneous materials used principles of continuum mechanics of materials to analytically represent the properties [3, 4, 5]. Recent efforts to solve the problem more realistically are based on micro-mechanical models coupled with numerical homogenization [6, 7, 8], which is preferred for its computational efficiency. The representative volume element (RVE), defined as a volume of heterogeneous material that is sufficiently large to be statistically representative of the composite [9], is used extensively in numerical homogenization techniques. The RVE approach generally assumes periodicity in the material composition at the small scale, which can be coupled with an appropriate method for linking the scales of the problem.

The current emphasis on rational performance driven design approaches demands a more critical consideration of the ways to define the relationship between the constituents of a material and structural performance [10]. Previous work has investigated the link between micro- and macroscale behavior for systems of polycrystal grain boundaries [11, 12], randomly dispersed spheres and spherocylinders [13, 14, 15, 16, 17], randomly dispersed circular inclusions [18, 19, 20], elliptical inclusions [18], arbitrarily shaped particle inclusions [21, 22, 23, 24], randomly dispersed uniaxial fibers [25, 26, 27], randomly dispersed voids [28, 19], and long fibers with ellipsoidal cross-sections [29].

Hierarchical design methods have been proposed [30, 31, 22] in which optimal microstructure processing techniques and associated characteristics are identified to achieve a target structural response for a range of composite materials. Fundamental to the approach is identification of a quantitative relationship between microstructural composition and material response [23]. This relationship is viewed as a random function which maps the input space to the output space. Machine learning is an attractive tool for identifying such mappings due to its flexibility and robustness for a wide variety of engineering problems [32]. There has been significant new research on machine learning for material design in recent years, resulting in the publication of several state-of-the-art reviews of the subject [33, 34]. Machine learning using neural networks was applied to the problem of brittle-matrix FRC as early as 1996 [35], when the effect of inclusion volume fraction on the stress-strain relationship of the material was analyzed. Other previous studies have used a descriptor-based approach to characterize the behavior of heterogeneous cementitious materials containing voids in the microscale [22]. Some research has focused on accounting for large numbers of random variables when characterizing random heterogeneous media [36]. Work relevant to the musical instrument string problem includes the use of a clustering approach for identifying different modes of damage in composite bridge wires using acoustic emissions data [37].

1.2.1 Short-Fiber Reinforced Composites

Early tests on fibrous composites showed that for a given matrix material, some fiber materials provide an increase in stiffness with higher volume fraction, whereas other fiber materials lead to a decrease in composite stiffness [38] . In a long fiber composite, the fibers have sufficient bond strength to remain embedded until the fiber fractures. In a short-fiber composite, however, failure at the interface occurs long before the fibers reach the stress required to break [39]. Actual interface structures are highly specific to the fiber and matrix combination. Pullout tests are commonly used to directly characterize interface resistance between the two phases in such material [40, 41]. The pullout test measures the resisting force as the fiber is pulled out of the matrix, which can be used to experimentally characterize the traction-separation relationship at the interface. Cohesive properties can be manipulated in a design by coating fibers with sprayed layers to enhance bond strength and/or introducing deformities to provide mechanical anchorage [42]. The effect of fiber crimping on the bond behavior for a polymeric fiber embedded in a cementitious matrix was studied by Bentur [43].

The Aveston Cooper Kelly (ACK) theory was developed in an early attempt to describe the tensile stress-strain behavior of FRC with a brittle matrix [44], and serves as the basis for later attempts to stochastically describe the cracking behavior of such composites [45, 46]. A number of approaches exist for modeling the homogenized behavior of FRC. The Concrete Damage Plasticity model (CDP) is one constitutive model capable of accurately predicting the nonlinear, multiaxial behavior of concrete using plasticity theory [47]. The CDP model is an attractive macroscale model for FRC due to its independent definition of tensile and compressive stiffness degradation. For a given state of strain in the composite, the resulting tensile stress is interpolated from data obtained from uniaxial tension and compression experiments. The work of Jankowiak [48] provides a detailed description of the parameter calibration process for CDP using experimental results. Nordendale recently demonstrated the efficacy of the CDP model for simulating the performance of SFRC under blast and impact loads [49].

While macromechanics is the study of the composite behavior when the material is treated as homogeneous with averaged apparent properties, micromechanics focuses on the interaction of the inherent heterogeneities. The scale at which heterogeneities interact is termed the microscale (or, more appropriately the mesoscale). As a result of advances in computational tools over the last few decades, numerical multiscale modeling methods have become a popular tool for analyzing and predicting the response of heterogeneous materials [50]. In general, multiscale methods are classified as sequential or concurrent.

Concurrent models solve systems at different scales in parallel, exchanging information between the models used in each scale [51, 52, 53]. In the process, the macroscale response gets influenced by behavior at the microscale, and vice-versa. In a sequential model [54, 55, 56] simulations at each scale are performed separately. Unknown coarse-scale parameters are determined using the solution of the microscale problem. A sequential multiscale approach is particularly useful when a model exists to capture the coarse scale behavior of the material, but one or more macroscale model parameters are dependent on the microstructure response. In a sequential multiscale framework, the governing equations at each scale are uncoupled, so complete scale separation is assumed. The Heterogeneous Multiscale Method (HMM) [57] was developed as a framework for either sequential or concurrent multiscale techniques where a different physical or numerical model is needed at each scale, in contrast to traditional homogeneous multiscale methods.

Previous work has assessed the effect of fiber-matrix interfacial debonding and fiber orientation on the elastic response of a composite using the finite element method (FEM) to model the RVE [58, 59] . Unique challenges arise in identifying the RVE of a material with randomly distributed inclusions, requiring special treatment from a multiscale modeling perspective. In the work of Xu [21], random RVEs containing elliptical inclusions were decomposed into smaller statistical volume elements (SVE) whose response was analyzed and quantified for upscaling. The approach is similar to Voronoi cell FEM, used by Ghosh [27], where a statistically equivalent RVE (seRVE) is used to model FRC at the microscale within the context of a concurrent multiscale model. Greene [31] identifies scenarios where uncertainties significantly affect macroscale behavior using a statistical description of the microstructure for three benchmark problems. Recent work by Clement [29] focused on the development of a database of elementary cells of fiber reinforced polymers containing arbitrary heterogeneities within the stochastic dimension. One useful aspect of the present effort is that different microstructures may continue to be modeled and added to the database, which can be used within a hierarchical multiscale framework.

As the nonlinear response of a composite is dependent upon the arrangement of fibers at the fine scale, phenomenological parameters for the coarse scale must be determined experimentally, or through a number of microstructural simulations capable of capturing the local interaction effect of fibers. While FEM can be used to model fiber-matrix debonding at the fine scale using cohesive elements at the interfacial phase boundary, conventional FEM is ill-suited to handle complex microstructure geometries due to the need for extremely fine meshing of the domain. The extended finite element method (XFEM) has emerged as a computationally efficient alternative to cohesive zone modeling, and its usefulness has been demonstrated for woven fibers [60] and random short-fibers [61, 62, 63].

1.2.2 Composite Musical Instrument Strings

The work of Abbott provides a thorough history of stringed instruments [64, 65]. The earliest musical strings were made of various animal intestines, colloquially known as *catgut*. Catgut strings suffered from a lack of homogeneity and as a result their tone and durability were rather inconsistent. Evidence has shown that metallic strings can be traced back as early as the 12th century, but technological limitations on the wire drawing process kept them from becoming commonplace at the time. Early metal strings were made of brass, copper, or silver since these materials are easy to draw and resist corrosion.

The earliest strings consisted of a single wire with a uniform circular cross section. Today, such strings are referred to as plain strings. Smaller diameters are used for strings tuned to higher pitches and larger diameters are used for low frequencies so that the tension in each string is similar. Instrument makers later realized that they could expand the frequency range of instruments by wrapping a plain string with an additional wire. In this configuration, the core wire carries the tensile force, while the wrap wire provides an increase in the effective mass for vibration control. The bond between the core and the wrap wire relies on the friction between the two materials. Some sources suggest that wrapped strings were made in Germany as early as the 14th or 15th century, although they did not become the norm until several centuries later [66]. These types of strings are referred to in literature by various names today, including overwrapped strings [67], wound strings [68, 69, 70], overwound strings [71, 72], and overspun strings [73]. Wound strings may be further classified by the cross section geometry of the wrap wire. Roundwound strings use a circular wrap wire, halfwound strings use a semicircular wrap wire, and flatwound strings use a rectangular wrap wire.

The analysis of vibrating strings dates back to the work of Pythagoras, who noted that the length of a string is correlated to its fundamental frequency. In 1749, D'Alembert derived the solution to the 1D wave equation, which describes the motion of a plucked string. The work of Fourier later provided the basis for spectral analysis. Early work on the theoretical behavior of strings relied on assumptions of an ideal string, one that is infinitely long and flexible with rigid supports. Lord Rayleigh derived equations for the vibration of homogeneous strings with finite stiffness more than a century ago [74]. In 1964, Fletcher expanded upon this work to account for the effect of material stiffness and wrap wire on the performance of piano strings [75]. His later work provided a review of equations relating the physical properties of strings to their nonlinear response, inharmonicity, and damping [73]. Three sources of damping are present in a vibrating instrument string: energy loss through the supports, viscous air damping from the surrounding media, and internal material damping. Internal damping is generally negligible for solid homogeneous strings, but very significant in wrapped strings. In 1982, Davis developed a graphical design guide for flexible strings [69]. The design guide related string diameter, frequency, and stress using several connected plots with common axes. A few years later Firth developed an updated design guide that incorporated Fletcher's equations for a stiff string with inharmonicity [67, 76]. Pickering later studied the influence of core tension during the manufacturing process on the performance of a wrapped violin string [71] .

The computer revolution and corresponding improvement in instrumentation allowed for researchers to experimentally study the response of real strings with increasing fidelity.

The effect of aging on the frequency response of guitar strings was studied by Allen [77]. It was found that new strings have more upper harmonics present that decay less rapidly than aged strings. Houtsma attempted to measure the change in these frequency dependent damping parameters as electric guitar strings aged [68]. His experimental work used a modified Gibson SG guitar body with brass wound steel strings. However, brass is not commonly used for electric guitar string wrap wire since it is not magnetic and electric guitars such as the SG generally use magnetic pickups. No detail was provided on how the signals were measured in the study. Hanson designed an experimental apparatus to study the response of an isolated wound nylon string (as found on a classical guitar) and an isolated wound steel string (as found on a folk guitar) in an attempt to measure results without the influence of body coupling [70]. Hancock took a similar approach in his studies on piano and cello strings [78, 72]. Several anomalies were identified between the theoretical and real response of the strings. He suggested that flexibility, the indeterminate location of endpoints, tension gradient among the string length, and coupling between transverse and longitudinal modes were the cause of deviation. Penttinen showed that it is possible to identify the location of the pluck based on the harmonics present in the signal [79]. Kemp studied the effect of two different wrap materials on pitch deviation using an electric guitar with a tremolo arm that is used by some guitarists to stretch strings while playing [80].

More recent work has investigated characteristic damage in steel instrument strings using microscopy. Bulbul studied damage modes in steel strings used on a bağlama [81, 82]. Three primary damage mechanisms were identified: abrasion in the neck region from pressing on the strings, erosion in the body region from plucking, and environmental wear due to exposure to moisture, oxygen, oil, and dirt. Olver studied a number of fractured steel strings and identified that the primary cause of failure was the onset of transverse fatigue crack initiation in the cross section, followed by ductile failure [83]. The majority of failures occurred in the region where the strings were plucked.

It is important to note that the design and construction of instruments and their strings has a significant link to the field of psychoacoustics. Human perception of sound is highly dependent on social and cultural factors. Based on the way that humans are conditioned, certain frequencies, intervals, or rhythms may be perceived as pleasing or irritating. For this reason, there is no consensus among musicians or string manufacturers on what makes an "optimal" string. However, psychoacoustic studies have identified correlations between characteristics of a measured sound and perceived quality. One such study was performed by Dayan, who studied the characteristics of the sound produced by four different types of string for acoustic guitars [84]. The study attempted to correlate the magnitude of certain harmonics with subjective quality according to listeners. However, the tests did not specify the material or geometry of the strings used. Other authors studied the effect of acoustic guitar body style and material on perceived quality [85]. In this study a professional guitarist performed a variety of musical pieces on 15 different types of guitars from various makers. From the results, a correlation was determined between the measured properties in the signal and the perceived quality of the tone. A follow-up study [86] further investigated the link between perceived quality and constructional details. In these studies the researchers held constant, or attempted to do so, all the variables except for the different guitar bodies used. Because the electric guitar transmits signal using electromagnetic pickups rather than acoustic wave propagation, there is less influence of the electric guitar body geometry and material than in an acoustic guitar [87].

Recently there has been an increased interest in producing instrument sounds using digital synthesis. Woodhouse developed a synthesis model for plucked guitar string transients [88] and calibrated the synthesis model using experimental results from a nylon core string with metal a wrap wire [89]. His later work used this model to generate samples for a psychoacoustic study measuring the sensitivity of listeners to changes in frequency dependent damping and other parameters [90]. Synthesis of string vibrations from three instruments, the Chin, acoustic guitar, and the Pipa, was undertaken by Liang [91]. The three instruments used nylon-wrapped silk-wound steel, steel wound, and nylon wound strings, respectively. Karjalainen developed a digital signal processing technique to modify the sounds produced by an electric guitar so that they resemble the sound of an acoustic guitar [92]. These synthesis models provide useful insight into methods for parameterizing an experimentally measured signal.

Most of the studies discussed previously focused on only a few types of strings. For example, Woodhouse examined the difference between D'Addario composite nylon strings and Martin steel strings [90]. Hancock analyzed the response of two types of cello strings and piano strings [72]. Some of the studies presented experimental results but did not provide a thorough discussion of the test method. Houtsma, for example, does not provide any discussion of how the strings were plucked in his study on electric guitar strings [68]. In some cases, the signal recording rig is detailed, but the strings were plucked by the researcher, so the results are not reproducible.

1.3 Objectives

The objective of this research is to demonstrate the efficacy of a generalized descriptor-based machine learning framework to identify structure-property relationships in composite materials. The research synthesizes concepts from multiple fields related to computational mechanics, including material characterization and modeling, numerical methods, homogenization, digital signal processing, and machine learning. For the first problem, a recently developed XFEM approach is implemented to efficiently explore SFRC design space and study the effect of fiber randomness on composite behavior. In the second problem an experiment is designed, and a data-driven approach is used to analyze the previously unexplored relationship between wound electric guitar string geometry, material, and vibrational response. In each problem, the complex nonlinear interaction between the two material phases informs the overall behavior of the composite. Using data from numerical models and experiments, it is demonstrated that an appropriately trained learning machine can predict and classify the behavior of the composite using only parameters describing the phase properties and geometry. Once trained, the machine can be used to efficiently and accurately predict the response of untested composites, foregoing the need for expensive experiments and simulations.

Objectives for SFRC Problem:

1. Identify descriptor variables that influence the response of random SFRC microstructures and generate SVEs.
2. Numerically model SVEs using XFEM to generate database of composite performance.
3. Implement machine learning model to predict composite tensile response using microstructural descriptors.

Objectives for Composite Music String Problem:

1. Identify descriptor variables that influence the performance of composite electric guitar strings.
2. Design an experiment for capturing the time and frequency dependent response of a vibrating string in a controlled setting.
3. Implement machine learning approach to classify strings based on features of the signals they produce.

1.4 Dissertation Organization

This dissertation is organized as follows. Chapter 2 formulates the composite material design problem in the generic sense. Here, a framework for designing a multiphase material using concepts of stochastic processes and machine learning is introduced. Methods for characterizing a multiphase material are discussed in the context of a descriptor variable approach. The remainder of the document focuses on the two problems discussed above that can be approached using the framework described in Chapter 2. Chapter 3 addresses the SFRC problem. Chapter 4 applies the framework to identifying structure-property relationships in electric guitar strings. Chapter 5 summarizes the findings of the research to date and gives an outline of potential future work.

Chapter 2

Machine Learning of Structure-Property Relationships

In the composite material design problem, let Ω_M denote the coarse scale domain of a simple periodic linear elastic two-phase material with uniformly distributed inclusions and displacement and traction boundary conditions prescribed on $\partial\Omega_M^u$ and $\partial\Omega_M^t$, as shown in Figure 2.1. The elastic constants of the matrix and fibers are given as E_m, E_f, ν_m, and ν_f, where E is Young's Modulus, ν is Poisson's ratio, and the subscripts m and f denote matrix and fiber, respectively. Assuming the phases to be perfectly bonded, meaning no slip or delamination can occur at the phase interfaces, the elastic properties E_c and ν_c, of the composite in the domain Ω_M will be dependent on the elastic constants of the individual phases, as well as the geometry, or morphology, of the microstructure, which in this case can be characterized by the inclusion diameter d_f and the spacing, s. The matrix and inclusions interact in the fine scale domain Ω_m. Depending on the application of interest, the designer may wish to determine values of s and d_f such that the composite stiffness E_c is maximized, or select an inclusion material with an optimal stiffness for a fixed particle size and/or diameter. This example design problem is formalized as a constrained optimization problem in Equation 2.1:

$$\begin{aligned} \text{maximize} \quad & E_c = F(E_m, E_f, \nu_m, \nu_f, s, d_f) \\ \text{subject to} \quad & s \in S \\ & d_f \in D \\ & E_m, E_f \in E \\ & \nu_m, \nu_f \in \mathscr{V} \end{aligned} \tag{2.1}$$

Where S, D, E and $\mathscr{V}$ are sets of feasible design values for each parameter. If the material microstructure is random in nature it is necessary to analyze the structural reliability, which

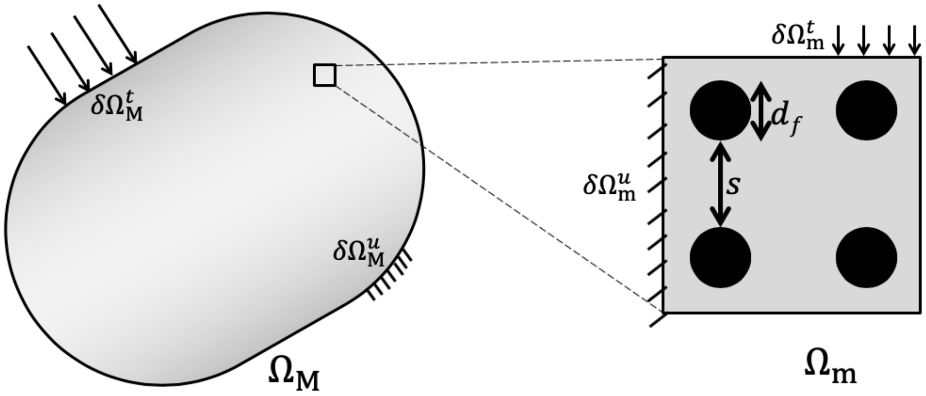


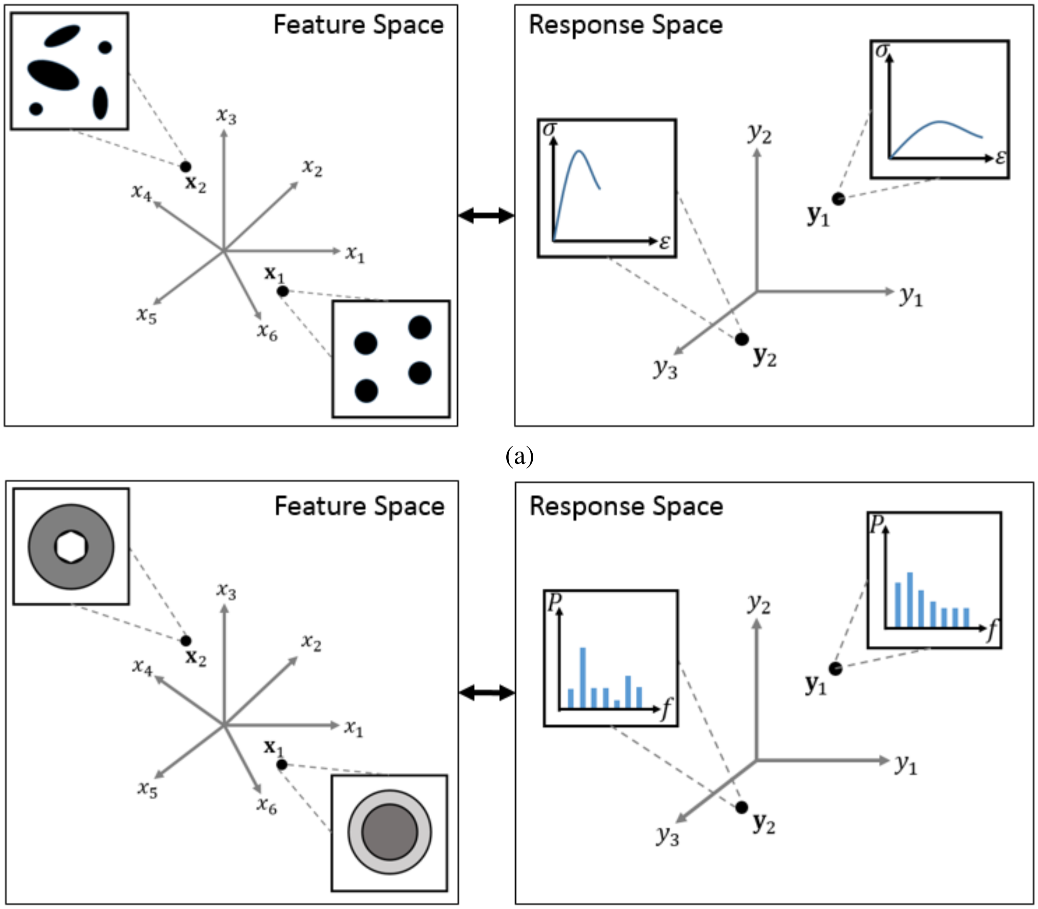
Figure 2.1: Example 2-phase microstructure optimization problem.

requires consideration of uncertainty in the material as well as external effects like loading. Random materials are characterized by higher variability in resistance, which cannot be neglected in design. Such variability cannot be represented by deterministic approaches and requires data-driven algorithms based on statistical concepts. Machine learning has emerged as a powerful tool for analyzing data in high-dimensional spaces obtained from experiments where randomness is present in the material or loading conditions. In the material science and computational mechanics community the descriptor-based machine learning approach has been implemented for prediction of the behavior of composite materials for a range of design applications. The following section will formalize the composite material analysis and design problem in this context.

2.1 The Descriptor-Based Machine Learning Approach

In a descriptor-based approach, a correlation is assumed between variables that describe the material on the fine scale, called descriptors, and measurable material properties. Descriptors may be thought of as design variables for an engineered composite optimization problem, or sources of uncertainty in an analysis problem. The descriptors may be discrete or continuous, depending on the nature of the parameter being constrained. For example, consider the design problem outlined in Figure 2.1. If a matrix and fiber material have already been selected, the engineer is tasked with determining optimal values of s and d_f so that E_c is maximized. If these are the only two design variables being considered, the descriptor space is 2-dimensional. In practical applications the feasible region D of values for fiber diameter is dependent on the range of diameters available from a manufacturer, which in turn depends upon the precision of the machine used for extrusion. Thus, d_f is a discrete random variable with a finite number of possible values defined by the manufacturer. The spacing s is a continuous random variable with a lower bound set by aggregate size of the matrix material, and an upper bound dictated by the dimensions of the matrix. To determine the optimal values of d_f and s, each dimension should be uniformly sampled. Random realizations of the composite are constructed using various values of d_f and s and tested using experiments or numerical models. If the matrix and fiber materials exhibit a relatively high degree of homogeneity, in the context of modeling, Young's modulus and Poisson's ratio may be considered deterministic. However, if the matrix and inclusion materials exhibit high variability, they must also be treated as descriptors with associated probability distributions. The response metric E_c is measured for each realization. By analyzing several microstructures with differing values of s and d_f, a correlation between design parameters and composite stiffness can be developed. For simple linear elastic systems such as that depicted in Figure 2.1, the composite properties can be explicitly solved for using mechanics of materials principles. In many applications, however, it is necessary to account for nonlinear behavior and damage, which requires an alternative modeling approach.

This process of identifying a function that maps descriptor space to response space can be achieved by machine learning. Using machine learning algorithms, it is possible to identify complex nonlinear patterns in large data sets obtained from experiments or simulations, enabling researchers to study the effect of input variables in high-dimensional design spaces in an unprecedented manner. Machine learning can be used to identify a surrogate model for a random process that can be used to robustly predict the response of an unknown structure using a database of previously tested samples. Machine learning algorithms can be paired with optimization tools such as genetic algorithms and simulated annealing to



(a)

(b)

Figure 2.2: Input/output space view of the descriptor approach with respect to (a) the SFRC problem, and (b) the composite guitar string problem.

identify optimal material structures. Figure 2.2 depicts the input-output space relationship for the problems discussed in Chapters 3 and 4. For the SFRC problem, random microstructures are mapped to composite tensile response using machine learning. For the vibrating string problem, a machine learning approach is implemented to map between a string's material phase, cross-sectional geometry, and the frequency response of the composite.

Machine learning methods generally fall into one of two classes: unsupervised learning or supervised learning. In unsupervised learning methods, algorithms search for structures within input data without regard for outputs. In supervised learning approaches, an optimal model $F^*(\mathbf{x})$ is identified that most accurately generates output data $\mathbf{y}$ from input data $\mathbf{x}$.

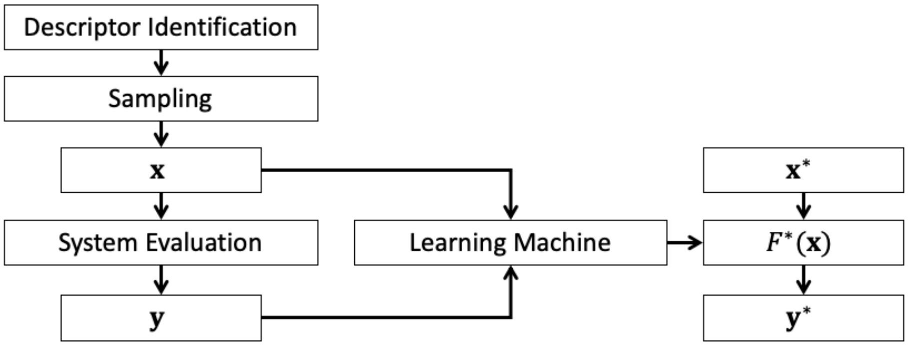


Figure 2.3: System diagram of supervised learning framework.

Figure 2.3 illustrates the supervised learning framework. The field of supervised learning can be further divided into classification and regression problems. In classification problems, each output falls into some category, and the model seeks to predict the category of the output from the input. In regression analysis, the output is a real-valued continuous variable. Both classification and regression methods are implemented in this research. Regression of SFRC properties based on material morphology is undertaken in Chapter 3. In Chapter 4, a classification method is used to identify composite string properties from their signals. The generalized learning framework consists of four components that are described in the following subsections:

1. Identification of descriptors
2. A sampling method to generate random input vectors $\mathbf{x}$
3. A system that produces outputs $\mathbf{y}$ for a given $\mathbf{x}$
4. A learning machine that estimates unknown values $\mathbf{y}^*$ at test points $\mathbf{x}^*$ using training data $\mathbf{x}$, $\mathbf{y}$

2.1.1 Descriptor Selection and Sampling

Identification of descriptors concerns the challenge of reducing the design space into a finite number of dimensions [93]. Use of too many dimensions creates a complex learning problem; whereas, too few dimensions may not be able to accurately model the random process. Principal component analysis can be used to analyze the sensitivity of a machine learning model to changes in descriptor values, with the objective of identifying irrelevant descriptors in order to reduce model complexity. Let a point $\mathbf{x}$ in d-dimensional descriptor space be defined by the row vector $\mathbf{x} = [x_1, x_2, ..., x_{d-1}, x_d]$. Each component of $\mathbf{x}$ represents the value of the descriptor $x_j,\ j = 1...d$. Each point in the descriptor space corresponds to a microstructure with a unique set of descriptors. It is assumed that the composite response $\mathbf{y}$ is dependent upon the location of $\mathbf{x}$ in the descriptor space. $\mathbf{y}$ is a p-dimensional vector of response quantities measured by experiments or simulations. p is related to the number of objective functions in the optimization problem. In the problem shown in Figure 2.1, where maximizing E_c is the only objective, $p = 1$ and $\mathbf{y} = E_c$. If one wished to maximize E_c in addition to maximizing or minimizing other parameters, then $p > 1$. The values of d and p are selected by the user based on the problem being considered. In the context of the descriptor approach, the constrained optimization problem in Figure 2.1 is generalized as:

$$\begin{aligned} \text{maximize} \quad & \mathbf{y} = F(\mathbf{x}) & \\ \text{subject to} \quad & x_j \geq a_j & j = 1...d \\ & x_j \leq b_j & j = 1...d \end{aligned} \tag{2.2}$$

Here $[a_j, b_j]$ is the range of feasible design values for descriptor x_j. The feasible region is defined independently in each dimension of the descriptor space. In the case of aligned or periodic material microstructures, descriptors such as volume fraction, geometry, and spacing of different material phases are sufficient to describe a unique composite material structure. This will be the case for the guitar string problem discussed in Chapter 4. For a composite guitar string, the diameter of the string, the core and wrap geometries, and the phase material properties sufficiently describe the location of the string in design space. However, in the case of random heterogeneous materials, such as the SFRC analyzed in Chapter 3, an infinite number of unique fiber distributions can be obtained using the same

Table 2.1: 2-phase random short-fiber reinforced composite descriptors and response metrics.

Morphological Descriptors	Composite Response Features
Fiber Volume Fraction	Peak Strength
Fiber Length	Elongation
Fiber Diameter	Strain Energy Capacity
Bond Characteristics	

Table 2.2: Wound musical instrument string descriptors and response metrics.

Morphological Descriptors	Composite Response Features
Core Wire Geometry	Internal Damping/Sustain Time
Core Wire Material	Partial Amplitudes
Wrap Wire Geometry	Inharmonicity
Wrap Wire Material	

fiber dimensions and volume fraction as inputs to a random microstructure generator. To address this, several statistically representative realizations of each point in descriptor space must be tested. Tables 2.1 and 2.2 show the descriptors and output variables considered in the problems presented in Chapters 3 and 4 respectively. For the SFRC problem, the fiber dimensions, volume fraction, and bond characteristics are considered as descriptors, while the matrix and fiber mechanical properties are treated as deterministic. Parameters describing the tensile stress-strain curve of the material are the response metrics to be predicted by the learning machine. For the guitar string problem, core and wrap wire material and geometries are used as descriptors, and signal metrics including modal frequencies and damping parameters are used as the response.

For material optimization, one must sufficiently explore the feasible region of design space using some random sampling technique. For the purpose of sampling, a probability distribution type must be identified for each input descriptor. Descriptors may be continuous or discrete, bounded or unbounded based on the physical property they represent. Different descriptors may be selected for the structure in a design or analysis problem. For a ceramic matrix composite, for example, the matrix material of interest may have some cracking strength or modulus that can be expressed as a normally distributed random variable. It may be of interest to determine the ideal diameter of fibers for reinforcement, so a uniform distribution may be used to explore the dimension of the design space corresponding to fiber diameter in an unbiased manner. Let $f_{x_j}(x)$ representing the probability density function (PDF) for the assumed distribution of descriptor x_j. For a general continuous random variable x_j with lower bound a_j and upper bound b_j

$$f_{x_j}(x) = P(x_j = x) \quad \forall \quad x \in [a_j, b_j] \tag{2.3}$$

$P(x_j = x)$ represents the probability that x_j is equal to x. The corresponding cumulative distribution function (CDF) is then defined as

$$F_{x_j}(x) = \int_{a_j}^{x} f_{x_j}(x)dx \quad \forall \quad x \in [a_j, b_j] \tag{2.4}$$

The CDF is non-decreasing and varies between 0 and 1 over the domain $x \in [a_j, b_j]$. To generate random descriptor values for sample $\mathbf{x}_i$, a vector of independently generated random numbers, $\mathbf{r}_i$ is constructed in the form of Equation 2.5.

$$\mathbf{r}_i = [r_1, r_2, ..., r_{d-1}, r_d], \quad r_j \in (0,1), \quad j = 1...d \tag{2.5}$$

Using the inverse CDF of each x_j, entries of $\mathbf{r}_i$ are converted into components of the random sample point $\mathbf{x}_i$ by Equation 2.6.

$$\mathbf{x}_i = [F_{x_1}^{-1}(r_1), F_{x_2}^{-1}(r_2), ...F_{x_{d-1}}^{-1}(r_{d-1}), F_{x_d}^{-1}(r_d)] \tag{2.6}$$

Let N denote the number of samples desired. In order to determine a meaningful relationship between $\mathbf{x}$ and $\mathbf{y}$, it is important that the feasible region is appropriately sampled in each dimension of the descriptor space. If components of $\mathbf{r}_i$ are obtained independently for each of the N samples, it is possible for multiple random samples to contain similar values in some dimensions. As a result, some regions of design space may be insufficiently sampled. To avoid this issue, stratified sampling methods are commonly used.

2.1.2 Random Process Evaluation/Data Acquisition

Experiments or numerical simulations are used to determine the response $\mathbf{y}_i$ of each sample. Some post-processing is generally required to obtain the output quantities $\mathbf{y}_i$. For the SFRC problem, the raw data from each simulation is in the form of stress-strain curve. Calculation of output quantities such as stiffness, strength, and ductility requires digitally searching the curve for maxima and minima and performing integration. For the musical instrument string problem, raw data is in the form of an AC signal in the time domain. Features such as signal rise and fall time, power, and harmonic content must be extracted from the raw data for learning purposes. Feature extraction for these problems will be discussed more thoroughly in Chapters 3 and 4.

Machine learning goes hand-in-hand with fields like data mining and knowledge discovery from databases (KDD) [94]. The development of a database is especially useful for material design and discovery because machine learning models can be updated to take advantage of new data as it becomes available [29]. It is important that the training data is representative of the overall system being considered. The data extraction process may require the user to clean noisy data or remove incomplete or corrupt data. When available, experimental data is ideal for training the stochastic model. However, due to the expensive and time-consuming nature of some material tests, data from computational models can be used to make informed predictions about the behavior of theoretical materials. Previous work has demonstrated the efficacy of using a combination of experimental data and numerical simulations to train stochastic models. Numerical models can be used to supplement datasets where some information is obscured or corrupt [93]. When insufficient experimental data exists, artificially generated data may be used to complete the set using the technique of bootstrapping [95]. In Chapter 3, computational models are used to simulate the response of random SFRC microstructures. In Chapter 4, string vibration data is obtained experimentally.

2.1.3 The Learning Machine

If the system being modeled is deterministic, it can be expressed simply as a function $\mathbf{y} = F(\mathbf{x})$. However, many real-world systems are not deterministic, or contain missing data that contributes uncertainty to the model. Rather than fitting an explicit function $F(\mathbf{x})$, the goal of machine learning is to determine a hypothesis $F^*(\mathbf{x})$ that approximates $F(\mathbf{x})$ most accurately using sample data $\{\mathbf{x}_i, \mathbf{y}_i\},\ i = 1...N$. This is achieved by testing a number of functions $F^*(\mathbf{x}, \omega),\ \omega \in W$, where ω are parameters in generalized parameter space W. The validity of a hypothesis is measured using a loss function $L(\mathbf{y}, F^*(\mathbf{x}, \omega))$. For a 2-class classification problem, where the machine is tasked with differentiating between two symbolic values, a common loss function is:

$$L(\mathbf{y}, F^*(\mathbf{x}, \omega)) = \begin{cases} 0, & \mathbf{y} = F^*(\mathbf{x}, \omega) \\ 1, & \mathbf{y} \neq F^*(\mathbf{x}, \omega) \end{cases} \tag{2.7}$$

For regression problems, squared error is a common metric loss function:

$$L(\mathbf{y}, F^*(\mathbf{x}, \omega)) = (\mathbf{y} - F^*(\mathbf{x}, \omega))^2 \tag{2.8}$$

The expected value of the loss function is given by the risk functional:

$$R(\omega) = \iint L(\mathbf{y}, F^*(\mathbf{x}, \omega)) P(\mathbf{x}, \mathbf{y})\, dx\, dy \tag{2.9}$$

The risk functional can be interpreted differently based on the class of problem. For classification, the risk functional returns a value on the interval [0,1] that indicates the probability of misclassification. For the regression problem, the risk functional is nonnegative but unbounded, with a value close to zero indicating high accuracy. Training the learning machine is thus a process of determining optimal parameters ω that minimize $R(\omega)$.

The final step of the machine learning process is model validation and verification. Selection of an appropriate function class for a random process is of crucial importance. There is no one-size-fits-all algorithm for machine learning; generally, the model building process consists of some trial and error to determine which type of model yields the best performance for the problem of interest. Several machine learning algorithms are tested during this phase and reliability analyses are later used to identify the most appropriate model. In Chapter 3, a Gaussian Process (GP) is used for regression of SFRC properties. In Chapter 4, a Support Vector Machine (SVM) is used to classify strings based on their signals.

One issue that must be addressed during validation is the finite nature of the data. For an asymptotically large number of samples, the measured error will be the true error of the model. However, for a smaller number of samples it is necessary to estimate the true error rate of the model given data from only a small portion of the overall population. This issue can be addressed by properly selecting training and test points from the data. If too few training points are used, the model may not be general enough to predict points outside the training set. If too many training points are used, there is a risk of overfitting the model. An overfitted model captures the noise in the response more than the true response. Resampling methods are used to address the issues of over and under-fitting, at the expense of some additional computational cost. The simplest resampling method is re-substitution, in which all samples are used for training and testing. While straightforward, this approach generally results in an under-approximation of model error. Another form of resampling is the holdout method, in which the available data is partitioned into a training set and a test set. The response of the trained model at the test points is then compared to the data to provide a measure of accuracy. Since partitioning the data differently will affect the performance of the model, an alternative approach is k-fold cross validation. Here, the data is divided into k equally sized disjoint subsets. One of the subsets is used for testing and the remaining $k-1$ subsets are used for training. The process is repeated k times and the error is taken as the average over all folds. In this approach, all data is used for training and validation.

Chapter 3

Prediction of Random Short-Fiber Reinforced Composite Properties

Random heterogeneous materials are common in nature and also in a number of engineering applications. Accurate prediction of material properties based on microstructural information requires an approach that combines statistical mechanics, homogenization theory, and stochasticity of geometrical parameters [96]. Materials with microstructures requiring statistical description are defined as random heterogeneous materials. Microstructure characterization is the process of identifying the stochastic geometry of a microscale domain, with the eventual goal of determining the correlation between microstructural geometry and macroscopic material performance [23]. A number of approaches have been tried to characterize random heterogeneous materials. Sources of heterogeneity can manifest as random distributions of pores of different shape and size, inclusions of different shape and size, or even as different phases of the same material, as in polycrystals. It is therefore necessary to consider a variety of microstructure descriptors, depending on the relevant properties of interest on the macroscopic scale.

The response of a composite is influenced by the volume fraction of the constituents, as well as their distribution, geometry, size, and interfacial bond properties. It is thus necessary to define a set of effective properties, K_e (Equation 3.1), of the composite, accounting for the presence of local heterogeneities so that the composite can be treated as a continuum at the structural scale.

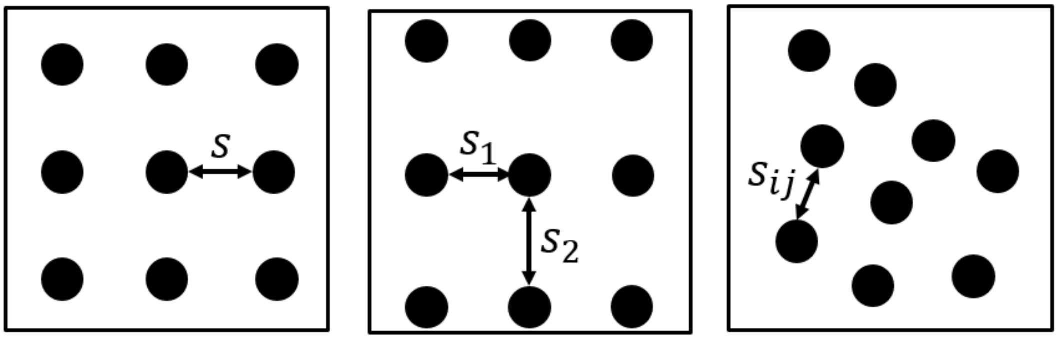
$$K_e = f(K_1, K_2, ...K_M; \phi_1, \phi_2, ...\phi_M; \Omega) \tag{3.1}$$

In Equation 3.1, M is the number of constituent phases in the composite. K_i and ϕ_i are the phase properties and volume fractions, respectively. Ω is a set of higher-order microstructural information such as interfacial bond descriptors. The rule of mixtures (Equation 3.2) is often used to define the effective properties of a composite at the macroscale as the volume average of phase properties, to predict the response of the composite with uniformly distributed heterogeneities and perfectly bonded phases.

$$K_e = \phi_1 K_1 + \phi_2 K_2 + ... + \phi_M K_M \tag{3.2}$$

While volume fraction and size of inclusion may influence the effective property of the composite, such descriptors do not completely account for the microstructural geometry of a random material. The bounds proposed by Hashin and Shtrikman [3] assume a quasi-isotropic material, signifying that the volume fraction is uniform throughout the composite domain. Figure 3.1 shows three 2D particle composite microstructures with equal volume fraction and size of inclusion. In Figure 3.1a the inclusions are uniformly spaced in each direction, so the geometrical descriptor may simply be defined as the interparticle spacing s. Figure 3.1b displays an anisotropic microstructure, where two descriptors (s_1 and s_2) are needed to adequately describe the geometry. Figure 3.1c depicts a random microstructure devoid of any uniformity. In this case, a statistical distribution may be used to describe the locations of the particles in terms of inter-particular distances, such as s_{ij} between particles i and j. For normally distributed particles, the mean and variance of s_{ij} may be used to characterize the stochastic geometry. Figure 3.2 depicts three possible randomly generated SFRC microstructures in 2D, based on the same input parameters for fiber length and volume fraction. It can be clearly observed that significant variation in the fiber concentration exists despite the use of equivalent values of design variables. In addition to local volume fraction fluctuation, the distance between the fiber tips and orientation of neighboring fibers also vary, which can influence damage accumulation and the propagation of matrix cracks.

In this chapter, a 2-phase composite is considered, in which the matrix material is a brittle ceramic and the inclusions are short ductile fibers to improve the tensile performance. In addition to discussing the influence of design factors on the response of such composites,



(a) (b) (c)

Figure 3.1: (a) Isotropic aligned, (b) anisotropic aligned, and (c) random microstructures with interparticle spacing s.

(a) (b) (c)

Figure 3.2: Random SFRC microstructure realizations with $\lambda = 100$ mm, $\phi_f = 0.1\%$, $l_f = 3.27$ mm.

the mechanical characterization of a brittle matrix SFRC is used as a test case for the descriptor-based machine learning (DBML) approach. The organization of this chapter is as follows: Section 3.1 discusses the behavior of FRC materials subjected to loading. Section 3.2 gives an overview of multiscale methods for modeling the effect of inclusions on the composite response. Section 3.3 formalizes the problem in the context of the DBML method and discusses the results. The conclusions from the study are presented in Section 3.4.

3.1 Characteristic Response of FRC

The ACK theory [44] suggests that the stress-strain curve of FRC may be described in three phases: pre-cracking, multiple-cracking, and post-cracking. It assumes that fibers only provide load-bearing capacity along their longitudinal axis only and that the bond strength between matrix and fiber is weaker than the material strength. Figure 3.3 illustrates the phases of ACK theory with respect to a typical stress-strain curve for an elastic fiber, a nonlinear matrix, and their combined composite response. During the pre-cracking phase (1), the fiber and matrix exhibit elastic behavior, and the elastic response of the composite is some weighted average of the constituent responses. In this stage, the stress is transferred between matrix and fibers by cohesive bonds. In the multiple-cracking phase (2), the matrix begins to crack when the cracking strain ε_{mc} is exceeded. The point of crack initiation is generally at the fiber tips, where stress concentration in the matrix is maximum. The fibers play the crack-bridging role in the now-damaged matrix. The fibers continue to debond until the cohesive bond strength is fully exhausted, after which a constant frictional stress is mobilized between matrix and fiber. The composite exhibits nonlinear response as cracking spreads throughout the matrix. In the post-cracking phase (3), the matrix is completely damaged and all loads are carried by the fibers, which eventually fail by fracture when the local strain exceeds ε_{fr}. In some applications, an elastic structure may be desired so that the nonlinear behavior of the composite can be ignored. Often, however, it is of interest to

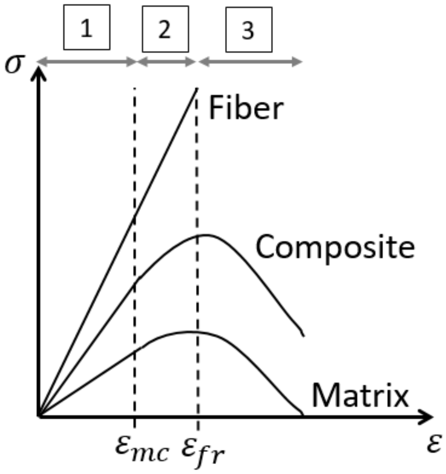


Figure 3.3: ACK model composite response phases.

study the post-peak behavior of the material in order to assess its total energy absorption capacity.

Pullout tests are commonly used to directly characterize the interfacial resistance between two phases in a material [40, 41]. This test measures the variation of the resisting force as the fiber is pulled out of the matrix, which can be used to experimentally characterize the traction-separation relationship at the interface. Cohesive properties can be manipulated in design by coating the fibers with sprayed layers to enhance the bond strength and/or introducing deformities to the fiber, say, by crimping or twisting, to provide active mechanical anchorage [42], as shown in Figure 3.4. Figure 3.5 illustrates the effect of fiber crimping on the bond behavior for a polymeric fiber embedded in a cementitious matrix [43]. The domain δ corresponds to the measured slip and the range T is the measured traction force. In Figure 3.5, a bilinear traction-separation model is superimposed on the experimental data. Such models are commonly fitted to experimental data for use in cohesive zone models.

The improvement in tensile response provided by fiber reinforcement is well documented. Fiber compressive strength influences composite compressive strength, but the effects cannot be easily deciphered because of the occurrence of micro-buckling in fibers

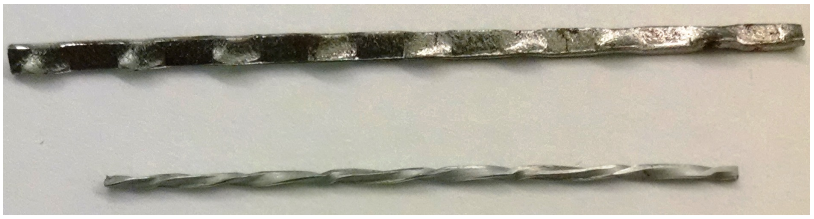


Figure 3.4: Two types of fiber surface deformities: crimping (top) and twisting (bottom).

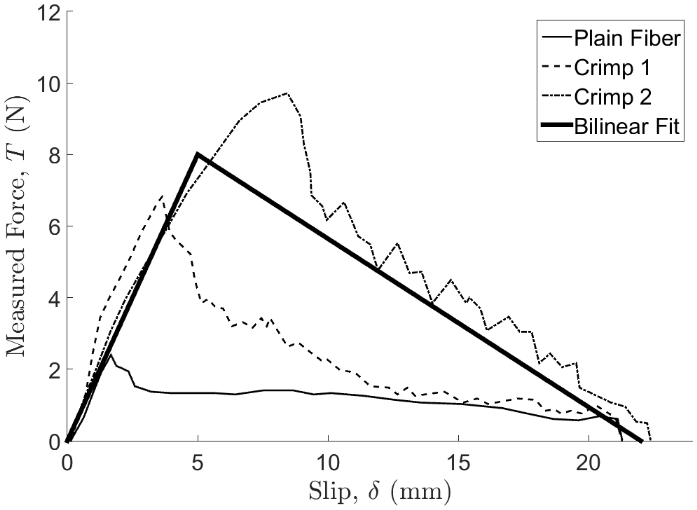


Figure 3.5: Example pullout test data for a ductile fiber in a cementitious matrix.

and imperfections in fiber orientation [39]. The tension and compression behavior of FRC is identical in the initial elastic stages of loading. Under higher tension, the stiffness degrades as multiple cracking and fiber pullout occurs.

In a long-fiber composite, the ratio of fiber length to fiber diameter is large ($l_f/d_f \rightarrow \infty$); whereas, short fibers are typically characterized by a finite aspect ratio in the order of 10 to 10^2. For full effectiveness, the length of shorter fibers should exceed a critical value given by $l_c = \sigma_{yf} d_f / 2\tau_b$, where σ_{yf} is fiber yield strength and τ_b is fiber-matrix bond strength. If fiber length is greater than $15l_c$, it can be treated as a continuous fiber. In many situations, cross-ply and multiaxial laminates are used. The concept of volumetric weighted average of the individual phases can then be used to determine the load share of each lamina in a given orientation.

3.2 Multiscale Modeling of SFRC

An engineered FRC at the structural or macroscale (10^{-1}-10^{0} m) may contain thousands of short fibers ($\sim 10^{-3}$ m in length). As it is impractical to explicitly discretize every single fiber in the problem domain, homogenized strength properties are often used, under the assumption that fibers are uniformly distributed. However, over a large problem domain, local heterogeneity may be present, introducing uncertainty in the performance of the material. Discontinuously reinforced composites often contain regions where the volume fraction varies significantly from that of the entire composite domain [3]. Figure 3.6 shows a domain containing randomly distributed fibers with a prescribed volume fraction of 0.1%. A division of this domain into four subdomains shows that the volume fraction observed in each subdomain appears to differ significantly from the average for the global domain with a minimum value of 0.047% and a maximum of 0.141%. Therefore, when an engineer prescribes a design value for the inclusion volume fraction, it cannot be assumed that the structure behaves in a homogeneous manner.

In the present study, a heterogeneous sequential multiscale modeling approach is de-

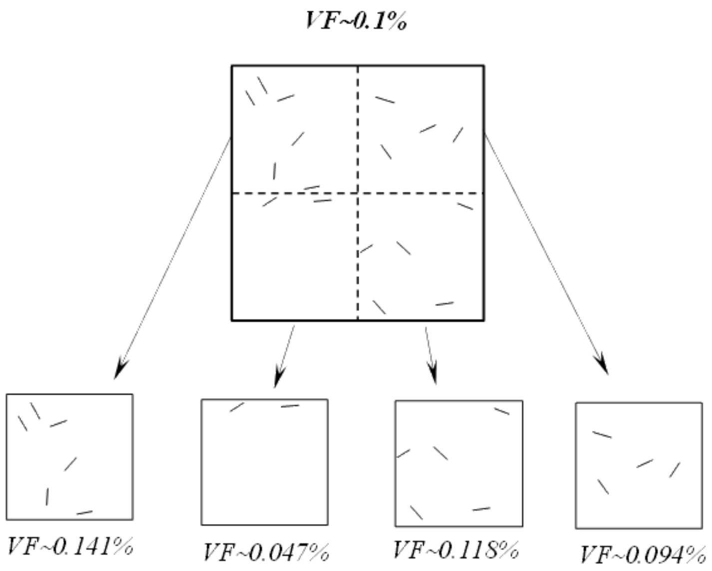


Figure 3.6: Local variation of fiber volume content in a random SFRC domain.

veloped for random SFRC. A database relating the microstructure composition to coarse scale parameters is developed using numerical methods. The observed composite properties can then be passed on to the macroscale, where a homogenized nonlinear model can be used with stochastic finite elements to assess the large-scale performance of the random composite in an efficient and convenient manner. A schematic of the approach used is shown in Figure 3.7. Points in the descriptor space are sampled and used to generate random microstructures over a range of design values. The selection of descriptors and sampling technique are discussed in Section 3.3. The random samples are analyzed using XFEM. Gaussian process regression is used to quantify the multivariate nonlinear relationship between descriptors and measured effective properties. The GP is also used to predict properties of subdomains in a macroscale model that correspond to unknown points in the descriptor space. At the macroscale, the material is modeled as homogeneous with each subdomain described by unique material parameters with stochastic distribution. Fiber randomness effects are accounted for by locally prescribed variations in macroscale parameters rather than explicitly modeling the fiber-matrix interaction effects. This enables assessment of the reliability of a macro structure, accounting for the local randomness of

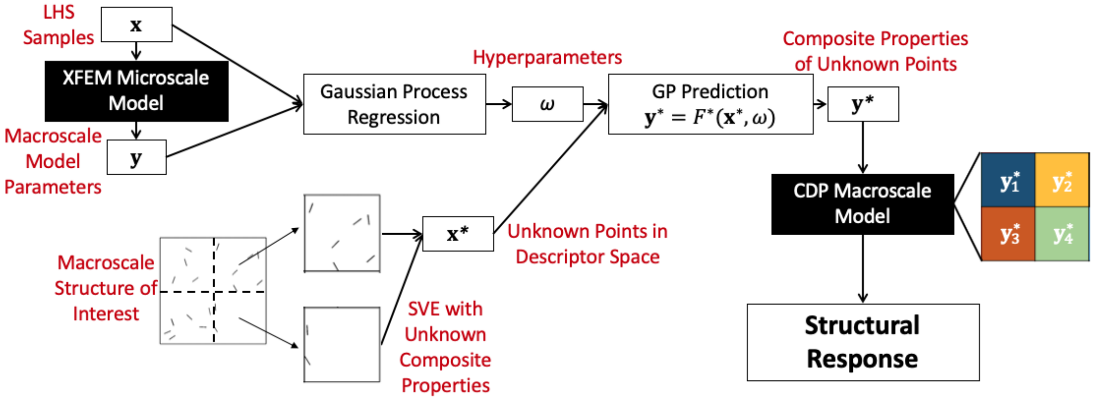


Figure 3.7: Proposed sequential multiscale modeling framework for SFRC.

design variables.

3.2.1 Macroscale Model

Several approaches exist for modeling the homogenized behavior of FRC. In 1989, Lubliner presented a constitutive model capable of accurately predicting the nonlinear, multiaxial behavior of concrete using plasticity theory [47]. The model is commonly known today as the Concrete Damage Plasticity (CDP) model and can be easily and efficiently implemented using commercial finite element analysis software such as Abaqus. The model consists of a yield criterion, a hardening rule, and a flow rule that can be fully defined by four parameters calibrated to uniaxial tension, uniaxial compression, biaxial compression, and triaxial compression test data. The CDP model is an attractive macroscale model for homogenized FRC due to the way it handles stiffness degradation. A stress-strain curve for the composite material, obtained by experiments or simulations, can be provided by the user during the model definition. For a given state of strain, the stress in the material is interpolated from the data. This allows for modeling the effect of fiber reinforcement without explicitly modeling the fibers.

Jankowiak provided a detailed description of the parameter calibration process using experimental results [48]. A brief overview of the CDP model is given here for the sake of completeness. The CDP constitutive equation is given by Equation 3.3.

$$\bar{\boldsymbol{\sigma}} = \mathbf{D}_0^{el}(\varepsilon - \varepsilon^{pl}) \in \{\bar{\boldsymbol{\sigma}} | F(\bar{\boldsymbol{\sigma}}, \tilde{\varepsilon}^{pl}, f, K_c) \leq 0\} \tag{3.3}$$

where $\bar{\boldsymbol{\sigma}}$ is the effective stress, defined in terms of total strain ε, plastic strain ε^{pl}, and initial elasticity matrix $\mathbf{D}_0^{el}$. F is the yield surface with parameters f, the ratio of the biaxial compressive strength to the uniaxial compressive strength, and K_c, the ratio of the second stress invariant on the tensile meridian to the second stress invariant on the compressive meridian, which shapes the deviatoric load surface. The effective stress is related to the Cauchy stress tensor $\boldsymbol{\sigma}$ by Equation 3.4.

$$\boldsymbol{\sigma} = [1 - d(\tilde{\varepsilon}^{pl})]\bar{\boldsymbol{\sigma}} \tag{3.4}$$

where $d(\tilde{\varepsilon}^{pl})$ is the scalar damage variable. $\tilde{\varepsilon}^{pl} = [\tilde{\varepsilon}_c^{pl} \quad \tilde{\varepsilon}_t^{pl}]^{\mathrm{T}}$ is the hardening variable, with components corresponding to compression and tension. The value of the hardening variable is non-decreasing and taken as $\tilde{\varepsilon}^{pl} = \int_0^t \dot{\tilde{\varepsilon}}^{pl} dt$. The evolution of the hardening variable is given by Equation 3.5.

$$\dot{\tilde{\varepsilon}}^{pl} = \begin{bmatrix} \dot{\tilde{\varepsilon}}_t^{pl} \\ \dot{\tilde{\varepsilon}}_c^{pl} \end{bmatrix} = \begin{bmatrix} r(\hat{\bar{\boldsymbol{\sigma}}})\hat{\dot{\varepsilon}}_{max}^{pl} \\ -[1 - r(\hat{\bar{\boldsymbol{\sigma}}})]\hat{\dot{\varepsilon}}_{min}^{pl} \end{bmatrix} \tag{3.5}$$

here $r(\hat{\bar{\boldsymbol{\sigma}}})$ is a stress weight factor for the multiaxial case that is dependent upon the principal effective stresses, $\hat{\bar{\boldsymbol{\sigma}}}$. $\hat{\dot{\varepsilon}}_{max}^{pl}$ and $\hat{\dot{\varepsilon}}_{min}^{pl}$ are the maximum and minimum eigenvalues of the plastic strain rate tensor. Equation 3.6 defines the stress weight factor.

$$r(\hat{\bar{\boldsymbol{\sigma}}}) = \frac{\sum_{i=1}^{3} \langle \hat{\bar{\sigma}}_i \rangle}{\sum_{i=1}^{3} |\hat{\bar{\sigma}}_i|} \tag{3.6}$$

$\langle \cdot \rangle = ((\cdot) + |\cdot|)/2$ represents the Macaulay brackets. The stress weight factor is equal to 1 for pure tension and 0 for pure compression. The scalar damage variable $d(\tilde{\varepsilon}^{pl})$ is a function of the user-defined uniaxial damage variables $d_t(\tilde{\varepsilon}_t^{pl})$ and $d_c(\tilde{\varepsilon}_c^{pl})$. For multiaxial stress it is assumed that the damage variable can be calculated using Equation 3.7.

$$d = 1 - (1 - s_t d_c)(1 - s_c d_t) \tag{3.7}$$

Computed parameters s_t and s_c are introduced to incorporate stiffness recovery effects into the model. Equation 3.8 gives the expression for the stiffness recovery parameters.

$$\begin{bmatrix} s_t \\ s_c \end{bmatrix} = \begin{bmatrix} 1 - w_t r(\hat{\bar{\boldsymbol{\sigma}}}) \\ 1 - w_c[1 - r(\hat{\bar{\boldsymbol{\sigma}}})] \end{bmatrix} \tag{3.8}$$

The parameters w_t and w_c dictate whether stiffness is recovered when loading changes from tension to compression. In general, quasi-brittle materials show compressive stiffness recovery when cracks open in tension and close in compression, but no stiffness is recovered in tension following initiation of cracking. This effect is captured by setting $w_t = 0$ and $w_c = 1$. The influence of the damage variables and recovery parameters on material stiffness are summarized in Figure 3.8. The non-associative flow rule is given by Equation 3.9.

$$\dot{\varepsilon}^{pl} = \dot{\lambda} \frac{\partial G(\bar{\boldsymbol{\sigma}}, \psi, \varepsilon)}{\partial \bar{\boldsymbol{\sigma}}} \tag{3.9}$$

Here $\dot{\lambda}$ is the non-negative plastic multiplier and ψ and ε are parameters that shape the flow surface, referred to as the dilation angle and eccentricity, respectively. G is the scalar-valued flow potential function, given by the Drucker-Prager hyperbolic function in Equation 3.10.

$$G(\bar{\boldsymbol{\sigma}}, \psi, \varepsilon) = \sqrt{(\varepsilon \tan\psi)^2 + \bar{q}^2} - \bar{p} \tan\psi \tag{3.10}$$

$\bar{p} = -\frac{1}{3} I_1$ is the effective hydrostatic pressure with I_1 representing the first invariant of the effective stress tensor and $\bar{q} = \sqrt{\frac{3}{2} \bar{\mathbf{S}} : \bar{\mathbf{S}}}$ is the Mises equivalent effective stress, defined in terms of the deviatoric effective stress $\bar{\mathbf{S}} = \bar{p}\mathbf{I} + \bar{\boldsymbol{\sigma}}$. Since $\dot{\lambda} = 0$ when $F < 0$ and $\dot{\lambda} > 0$

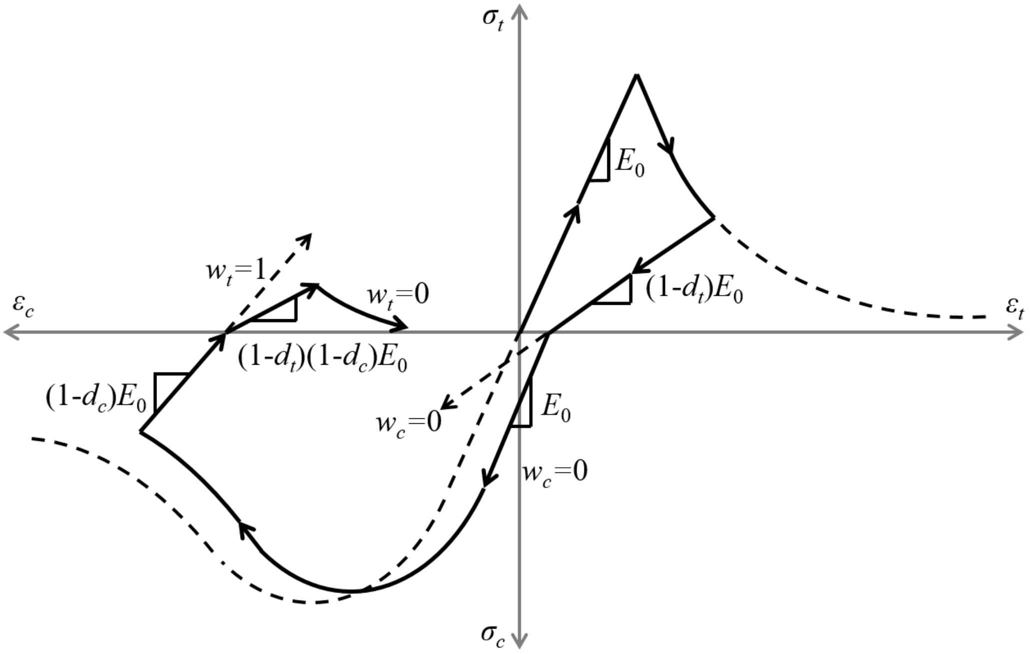


Figure 3.8: CDP model constitutive law.

when $F = 0$, the Kuhn-Tucker condition is given by $\dot{\lambda}F = 0$. The consistency condition $\dot{F} = 0$ follows from the assumption that when plastic slip is occurring the state of stress remains on the yield surface.

3.2.2 Microscale Model

In the previous section, a model for the homogenized, coarse-scale response of SFRC was discussed. In such a model, a characteristic stress-strain curve for the material must be provided by the user from experiments or simulations. Parameters describing the characteristic stress-strain curve are sensitive to the local arrangement of fibers and bond properties between the matrix and fiber. Since there are several design variables that influence these properties, it is necessary to identify a method for efficiently exploring the design space. Since fabrication of such materials is expensive and time consuming, a numerical modeling approach is preferred. This section details an extended finite element method (XFEM) for modeling random short fibers.

XFEM has been tailored for a variety of problems in material analysis and design. The approach can be used to model fracture, dislocations, grain boundaries, and phase interfaces [97]. The method has been extensively used to model cracks and inclusions within multiphase materials. Recently, XFEM has been employed to capture the inelastic response of textile-reinforced polymers [60], and to stochastically analyze the effect of elliptical inclusion aspect ratio on elastic behavior [24]. A method for handling multiple cracks and crack junctions within an element was proposed by Daux et al. [98]. Later work demonstrated that the junction of two cracks can be handled more easily by combining two step enrichments rather than using a special enrichment function for the junction [99]. Modeling multiple fibers in an element requires superposition of enrichment functions, and characterization of interaction effects. Recent work proposed a displacement approximation for elements containing multiple inclusions [18].

XFEM uses two-part approximation with a standard part and enrichment part. In the standard part, regular shape functions are used to approximate the displacement field. The enrichment part captures the rapid local variation of the displacement field due to the presence of discontinuities at dissimilar material interfaces resulting from inclusions, or at cracks. The local enrichment functions are selected *a priori* according to the nature of discontinuity and satisfy the partition of unity, so that the local enrichment does not affect the global solution. If a n-dimensional $\Omega \in \Re^n$ domain is discretized by a finite element mesh with a set of nodes m, out of which m^* nodes ($m^* \subset m$) are subjected to enrichment with the function ψ_j, the general form of the XFEM approximation field can be expressed as:

$$\mathbf{u}(\mathbf{x}) = \sum_{i \in m} N_i(\mathbf{x})\mathbf{u}_i + \sum_{j \in m^*} N_j^*(\mathbf{x})\psi_j(\mathbf{x})\mathbf{a}_j \tag{3.11}$$

The first term of this expression is the standard part and the second one is the enrichment part. Here, $\mathbf{u}$ is the displacement; N_i is the standard shape function for node i satisfying the Kronecker-delta property; N_j^* is the partition of unity shape function, normally set equal to N_i; ψ_j is the enrichment function for the inclusion at node j in set m^*; $\mathbf{u}_i$ are the displacement of the standard nodes; and $\mathbf{a}_j$ are coefficients of the enrichment shape functions. The level-set method, proposed by Osher and Sethian [100] to track moving interfaces, is based upon the idea of representing an interface as a level-set curve of a higher dimensional function, to account for, say, the presence of an arbitrarily oriented inclusion within a finite element. Weakly discontinuous enrichment functions can capture the strain discontinuity in an element, whereas a strongly discontinuous function can capture discontinuity in the displacement field. The implementation of XFEM requires addressing the following issues:

- Tracking the interface of different phases by using line segments or using the level-set method implicitly [101].
- Selecting the node set m^* for the enrichment region of interest.
- Identifying the enrichment function(s) to reflect the physics of the problem.
- Integrating the non-smooth enrichment functions.

An approach for modeling the behavior of an elastic matrix containing stretchable short-fiber inclusions with no delamination using XFEM was outlined by Pike and Oskay [61]. The enrichment functions were derived to model the strain discontinuity in an element resulting from an arbitrarily placed high aspect ratio fiber. More recently, the XFEM model was improved to include the effect of cohesive debonding between the matrix and inclusions. An additional enrichment function was derived to model the displacement discontinuity in an element as a result of debonding [62]. A nonlocal damage model was used to capture the stiffness degradation of the matrix material. It was demonstrated that with sufficient mesh refinement XFEM is capable of accurately reproducing the local mechanical response observed in a reference FEM model. The primary advantage of using XFEM for random SFRC modeling is that the mesh does not need to conform with the morphology of the problem domain. Depending upon the arrangement of the inclusions, the standard FEM mesh may prove to be geometrically complex and require an unreasonably large number of elements to accurately solve the problem. In recent work by the author [63] it was shown that multiple fibers in the same element domain can be modeled using XFEM by superimposing the enrichment functions of each fiber. The XFEM formulation is discussed briefly in this section, but a complete derivation of the model considered herein is reported in previous work [61, 62, 63].

The approximation of the displacement field within the composite domain is given by Equation 3.12.

$$\mathbf{u}(\mathbf{x}) = \sum_{i \in m} N_i(\mathbf{x})\mathbf{u}_i + \sum_{i \in n_f} \left[\sum_{j \in m_f^i} N_{ji}(\mathbf{x}) \left(\psi_i(\mathbf{x})\mathbf{a}_{ji} + \Upsilon_i(\mathbf{x})\mathbf{b}_{ji} \right) \right] \tag{3.12}$$

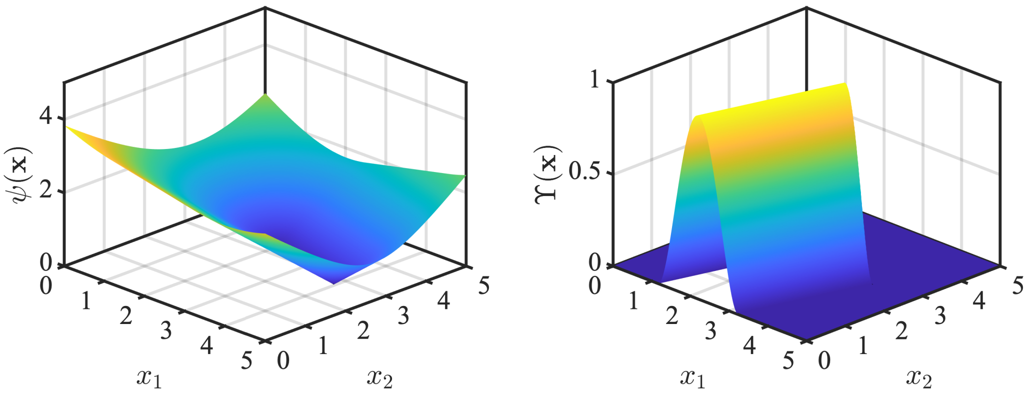
The formulation is similar to Equation 3.11 but contains an additional enrichment function $\Upsilon_i(\mathbf{x})$ to capture debonding along the fiber-matrix interface, while $\psi_i(\mathbf{x})$ captures the discontinuity in the strain field. The enrichment functions depend on the location of fiber i within the domain. n_f is the number of fibers in the domain, and m_f^i is the number of nodes enriched by fiber i. ji is the index set of enriched nodes for fiber i. $\mathbf{u}_i$, $\mathbf{a}_{ji}$, and $\mathbf{b}_{ji}$ are the nodal coefficients of the standard nodes, fiber motion enrichment, and debonding enrichment shape functions.

The fibers are assumed to lie entirely within the domain of the composite. The high aspect ratio fiber is idealized as a line segment. Let the domain $\mathbf{x}$ of a fiber be defined as:

$$\mathbf{x} = \mathbf{x}_c + \frac{\mathbf{x}_2 - \mathbf{x}_1}{2}s; \ -1 \leq s \leq 1; \ \mathbf{x} \in \Gamma \tag{3.13}$$

Where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the fiber tips, $\mathbf{x}_c$ is the fiber midpoint, and s is a normalized length parameter. The position of a fiber within the domain is defined using level-set functions. The level-set function along the fiber length, ϕ_c, is given by Equation 3.14.

$$\phi_c(\mathbf{x}) = \|\mathbf{x} - \mathscr{P}(\mathbf{x})\| \tag{3.14}$$



(a) (b)

Figure 3.9: Visualization of (a) strain discontinuity enrichment function and (b) debonding enrichment function for an arbitrarily placed fiber in a 5x5 mm element domain.

$\mathscr{P}(\mathbf{x})$ is the projection of $\mathbf{x}$ onto the fiber, given by:

$$\mathscr{P}(\mathbf{x}) = \mathbf{x}_1 + [(\mathbf{x}-\mathbf{x}_1)\cdot\mathbf{t}_2]\mathbf{t_2} = \mathbf{x}_2 + [(\mathbf{x}-\mathbf{x}_2)\cdot\mathbf{t}_1]\mathbf{t_1} \tag{3.15}$$

where $\mathbf{t}_1$ and $\mathbf{t}_2$ denote the tangent at each of the fiber tips. ϕ_c is equal to 0 along the fiber line segment and takes positive values on either side of the fiber. The level-set function for each of the fiber tips is given by Equation 3.16.

$$\phi_\lambda(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_\lambda)\cdot\mathbf{t}_\lambda; \quad \lambda = 1,2 \tag{3.16}$$

The enrichment function for the strain discontinuity across the fiber is expressed in terms of the level-set functions by Equation 3.17. The function is smooth and nonzero everywhere in the domain except along the fiber, as illustrated in Figure 3.9a.

$$\psi_i(\mathbf{x}) = \left[\prod_{\lambda=1}^{2} H(-\phi_\lambda)\right]\phi_c(\mathbf{x}) + \sum_{\lambda=1}^{2} H(\phi_\lambda)d_\lambda(\mathbf{x}) \tag{3.17}$$

H is the Heaviside function and $d_\lambda(\mathbf{x}) = \|\mathbf{x}-\mathbf{x}_\lambda\|$ is the distance between $\mathbf{x}$ and the fiber tip.

The debonding enrichment function was derived using the same level-set functions but introduces a displacement discontinuity rather than a strain discontinuity along the fiber. The debonding enrichment function is given by:

$$\Upsilon_i(\mathbf{x}) = \phi_p(\mathbf{x}) H(r(\phi_c(\mathbf{x}))) \left(\prod_{\lambda=1}^{2} H(-\phi_\lambda(\mathbf{x})) \right) \tag{3.18}$$

$r = \pm\phi_c$ is the signed distance function, defined as positive on one side of the fiber level-set ϕ_c and negative on the opposite side. The shape of the debonding enrichment function is controlled by the discontinuity function ϕ_p. In previous work [62] an expression was derived for ϕ_p as a fourth order polynomial subject to constraints that the maximum debonding along the fiber occurs at the center of the fiber and the ends of the fiber remain embedded in the matrix:

$$\phi_p(\mathbf{x}) = 1 + \frac{\tan\theta_d}{2} s(\mathbf{x})^2(1 - s(\mathbf{x})^2) - s(\mathbf{x})^2(2 - s(\mathbf{x})^2) \tag{3.19}$$

where θ_d is a parameter describing the slope of the discontinuity at the fiber tips. The resulting enrichment function Υ_i is illustrated in Figure 3.9b.

The governing equation for equilibrium in the model domain Ω is given by Equation 3.20:

$$\nabla \cdot \boldsymbol{\sigma}(\mathbf{x}) = 0; \quad \mathbf{x} \in \Omega \tag{3.20}$$

Using a continuum damage mechanics approach, the constitutive law for the matrix is given by Equation 3.21:

$$\boldsymbol{\sigma} = (1 - w(\mathbf{x},t))\mathbf{L} : \varepsilon(\mathbf{x}) \tag{3.21}$$

In the above equations $\boldsymbol{\sigma}$ is the stress tensor and $\nabla(\cdot)$ is the divergence operator. The stress is related to strain ($\varepsilon = \nabla^s\mathbf{u}$) by the elastic moduli tensor $\mathbf{L}$. $w(\mathbf{x},t)$ is a history-dependent damage scalar that ranges from 0 (undamaged) to 1 (fully damaged). The displacement and traction boundary conditions are defined by Equations 3.22 and 3.23, respectively.

$$\mathbf{u}(\mathbf{x}) = \tilde{\mathbf{u}}(\mathbf{x})\,; \quad \mathbf{x} \in \Gamma_u \tag{3.22}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \tilde{\mathbf{t}}(\mathbf{x})\,; \quad \mathbf{x} \in \Gamma_t \tag{3.23}$$

The boundaries Γ_u and Γ_t denote non-overlapping regions of the domain on which $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{t}}$ are prescribed, with $\partial\Omega = \Gamma_u \cup \Gamma_t$. The stress in a fiber i is assumed to be axial and proportional to the axial strain by the relation $\sigma_f^i = E_f \varepsilon_f^i$, where E_f is the elastic modulus of the fiber. For a fiber of length l_f and thickness t_f, as $t_f/l_f \to 0$, the weak form of Equations 3.20-3.23 can be expressed as:

$$\int_\Omega \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon}\, d\Omega + \sum_{i \in n_f} t_f E_f \int_{\Omega_i} \varepsilon_f^i \delta\varepsilon_f^i\, d\Omega + \sum_{i \in n_f} \int_{\Gamma_i} \boldsymbol{T} \cdot \delta[\![\mathbf{u}]\!]\, d\Gamma - \int_{\Gamma_t} \tilde{\mathbf{t}} \cdot \delta\mathbf{u}\, d\Gamma = 0 \tag{3.24}$$

Ω_i denotes the domain of fiber i and Ω represents the domain of the matrix. Γ_i is the interface between the matrix and fiber i. $\delta\mathbf{u}$ and $\delta\boldsymbol{\varepsilon}$ represent the test function and the gradient of the test function, respectively. $\boldsymbol{T}$ is the traction resulting at the interface due to jump discontinuity of displacement across the fiber, $[\![\mathbf{u}]\!]$. The traction separation relationship is included in the model using an intrinsic cohesive law with uncoupled normal and shear components. In this work a bilinear cohesive law is used to characterize the interfacial bond. Three parameters define the idealized curve shown in Figure 3.5: peak cohesive strength (T_n), characteristic displacement at the peak cohesive strength (δ_n) and displacement at bond failure (δ_{max}). The subscript n denotes the normal component of the traction T and the displacement δ. For the sake of simplicity, the shear parameters are assumed to be proportional to the normal components in the initial study. A range of values are used for the cohesive parameters in the numerical studies detailed in Section 3.3, in order to assess the influence of the bond characteristics on the response of the composite material.

A nonlocal continuum damage model is used to capture the progressive degradation of matrix stiffness based on the weighted average of the principal strains within a radial basis.

As the damage parameter increases, the stiffness of the material progressively decreases until failure. An alternative continuum damage law can easily be used if deemed suitable for the matrix material of interest. At an arbitrary material point $\hat{\mathbf{x}}$ the evolution of the damage variable follows the arctangent function of Equation 3.25.

$$w(\hat{\mathbf{x}},t) = \frac{\arctan(a\,k(\hat{\mathbf{x}},t) - b) + \arctan(b)}{\frac{\pi}{2} + \arctan(b)} \tag{3.25}$$

In Equation 3.25, a and b are parameters fit to data from tests on the unreinforced matrix material. $k(\hat{\mathbf{x}},t)$ is an increasing non-negative history dependent variable defined by Equation 3.26.

$$k(\hat{\mathbf{x}},t) = \max_{\tau\in[0,t]} \left(\langle \hat{v}(\hat{\mathbf{x}},\tau) - v_{\text{ini}}\rangle\right) \tag{3.26}$$

$\hat{v}$ is the nonlocal equivalent strain obtained from weighting the local equivalent strains v within a radial basis defined by $\hat{\lambda}(\mathbf{x},\hat{\mathbf{x}})$. v_{ini} is a threshold value of equivalent strain below which damage is assumed not to progress. Equation 3.27 defines the local equivalent strain in two dimensions as the norm of the positive principle strains $\hat{\varepsilon}_i$, which is intended to simulate stiffness degradation as cracks form in the matrix under tensile loads.

$$v(\hat{\mathbf{x}},t) = \sqrt{\sum_{i=1}^{2} \langle \hat{\varepsilon}_i(\hat{\mathbf{x}},t)\rangle^2} \tag{3.27}$$

The nonlocal equivalent strains are then computed using Equation 3.28.

$$\hat{v}(\hat{\mathbf{x}}) = \frac{\int_\Omega \hat{\lambda}(\mathbf{x},\hat{\mathbf{x}})v(\mathbf{x},t)dx}{\int_\Omega \hat{\lambda}(\mathbf{x},\hat{\mathbf{x}})dx} \tag{3.28}$$

The weights are determined using the Wendland Radial Basis Function [102] given by

Table 3.1: SFRC descriptor variable ranges for numerical models.

Descriptor x_j	Lower Bound (a_j)	Upper Bound (b_j)
d_f, Fiber diameter (x_1)	0.005 mm	0.1 mm
l_f, Fiber length (x_2)	5 mm	15 mm
VF, Fiber volume fraction (x_3)	.01%	1.5%
T_n, Peak normal cohesive strength (x_4)	5 MPa	15 MPa
δ_n, Characteristic normal cohesive separation (x_5)	0.000005 mm	0.0001 mm
δ_{max}, Maximum cohesive separation (x_6)	0.0002 mm	0.0005 mm

Equation 3.29, where l_c defines the span of the radial basis.

$$\hat{\lambda}(\mathbf{x},\hat{\mathbf{x}}) = \left(1 - \frac{\|\mathbf{x}-\hat{\mathbf{x}}\|}{l_c}\right)^4 \left(4\frac{\|\mathbf{x}-\hat{\mathbf{x}}\|}{l_c} + 1\right) \tag{3.29}$$

3.3 DBML Implementation

3.3.1 Descriptor Selection and Sampling

Given some matrix and fiber material, it is necessary to identify the ideal combination of fiber length, diameter, volume content, and interfacial cohesive properties to achieve desirable inelastic response. Here, $d = 6$ descriptors are selected, as listed in Table 3.1.

In this work a uniform distribution is used for each descriptor so that the entire range of the feasible region for a descriptor can be considered with equal bias. A continuous distribution is used for each variable in this problem. In reality, possible values of fiber diameter and cohesive parameters would be discrete variables, with values dictated by what is available from a manufacturer. Using a uniform distribution, components of $\mathbf{x}_i$ in Equation 2.6 are calculated as

$$F_{x_j}^{-1}(r_j) = a_j + r_j(b_j - a_j) \tag{3.30}$$

In this study, no correlation between parameters is used in sampling the descriptor space, i.e. each descriptor is treated as an independent design variable. Latin Hypercube Sampling (LHS) was used to generate $N = 40$ samples in the 6-dimensional descriptor space, the

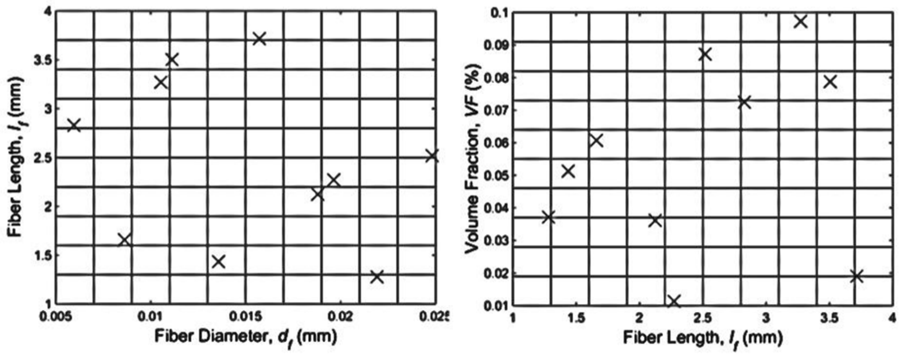
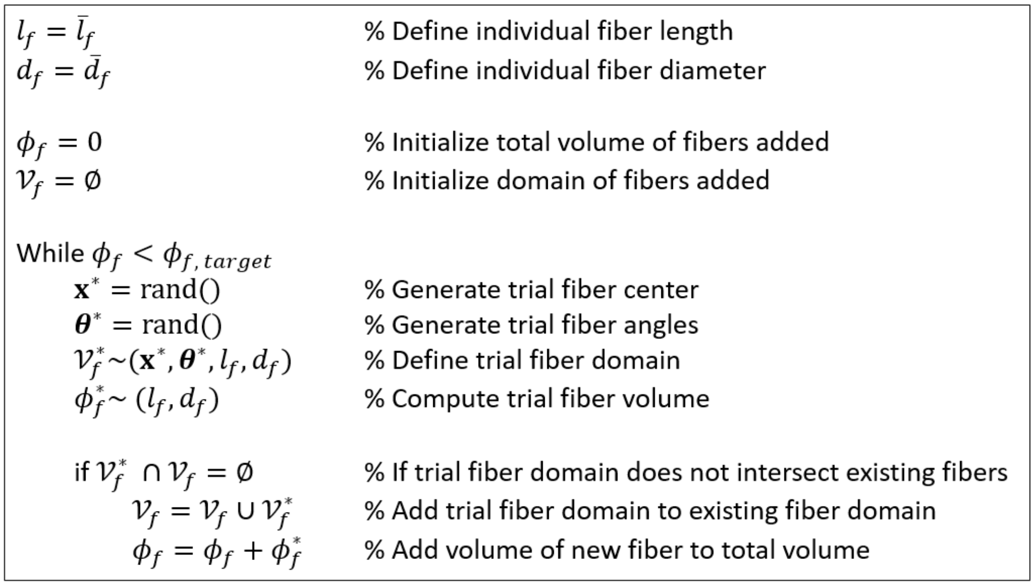


Figure 3.10: 2-dimensional feature space visualization of LHS for SFRC.

span of which is defined in Table 3.1. A list of the 40 resulting sample points is given in Appendix Tables A.1-A.2. LHS partitions each dimension of descriptor space into N equal intervals. Samples are generated such that only one sample will be drawn from each interval in each dimension. Figure 3.10 illustrates the LHS method. In this illustration, $N = 10$ samples were drawn from 3-dimensional descriptor space, with dimensions corresponding to fiber length, diameter, and volume fraction. Plotting the samples in 2 dimensions shows that each dimension of space has been partitioned into 10 subregions, each containing 1 sample.

A random sequential addition (RSA) algorithm was used to generate the random short-fiber microstructures corresponding to each sample point. The RSA process was formalized in the context of hard spheres being randomly packed into a fixed volume [96]. The center of a sphere is placed at a randomly selected point within the microstructure domain. From the remaining unoccupied volume a new random point is selected and the next sphere is placed. The process is repeated until the desired inclusion volume fraction is reached. A similar approach is used herein, but with fibers instead of spheres. Figure 3.11 summarizes the manner in which RSA is implemented algorithmically for the case of short-fiber reinforcement. The descriptors listed in Table 3.1 were used as inputs for the RSA algorithm. The user must define the length l_f and diameter d_f of the fibers to be placed in the matrix, as well as the target volume fraction $\phi_{f,target}$. The volume ϕ_f of an individual fiber is



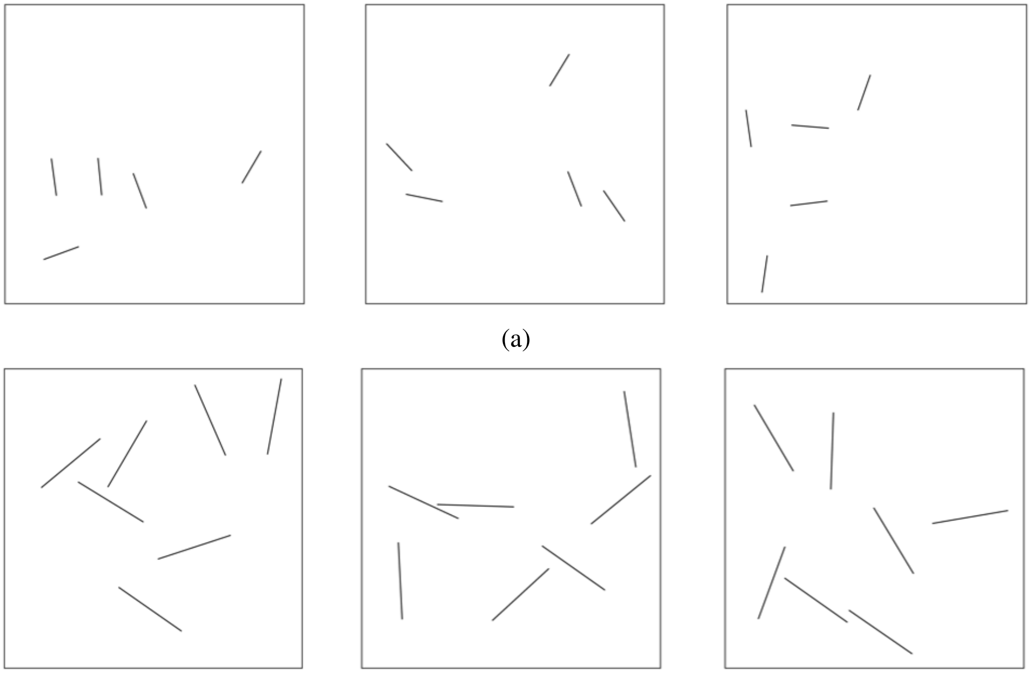
```
$l_f = \bar{l}_f$                      % Define individual fiber length
$d_f = \bar{d}_f$                      % Define individual fiber diameter

$\phi_f = 0$                           % Initialize total volume of fibers added
$\mathcal{V}_f = \emptyset$            % Initialize domain of fibers added

While $\phi_f < \phi_{f,target}$
    $\mathbf{x}^* = \mathrm{rand}()$                        % Generate trial fiber center
    $\boldsymbol{\theta}^* = \mathrm{rand}()$               % Generate trial fiber angles
    $\mathcal{V}_f^* \sim (\mathbf{x}^*, \boldsymbol{\theta}^*, l_f, d_f)$   % Define trial fiber domain
    $\phi_f^* \sim (l_f, d_f)$                              % Compute trial fiber volume

    if $\mathcal{V}_f^* \cap \mathcal{V}_f = \emptyset$     % If trial fiber domain does not intersect existing fibers
        $\mathcal{V}_f = \mathcal{V}_f \cup \mathcal{V}_f^*$    % Add trial fiber domain to existing fiber domain
        $\phi_f = \phi_f + \phi_f^*$                        % Add volume of new fiber to total volume
```

Figure 3.11: RSA algorithm overview.

computed as $(d_f \cdot l_f)$ in 2D or $(\pi d_f^2 l_f/4)$ in 3D. A trial fiber center $\mathbf{x}^*$ and orientation angle $\boldsymbol{\theta}^*$ are randomly generated. In 2D, $\mathbf{x}^*$ is 2-dimensional and $\boldsymbol{\theta}^*$ is 1-dimensional. In 3D, $\mathbf{x}^*$ is 3-dimensional and $\boldsymbol{\theta}^*$ is 2-dimensional. The domain $\mathcal{V}_f^*$ of the trial fiber is related to $\mathbf{x}^*$, $\boldsymbol{\theta}^*$, and user-defined fiber dimensions. A check is performed to see if the trial fiber intersects any of the previously placed fibers. If the trial fiber is found to occupy a unique region of the microstructural domain, the volume of placed fibers is increased and the existing fiber domain is updated. In practice, some convergence criteria should be added to the algorithm to prevent the occurrence of an infinite loop when the maximum packing limit is reached. Figure 3.12a shows three samples of point $\mathbf{x}_{20}$ (with $d_f = .027$ mm, $l_f = 6.269$ mm, $VF = 0.127\%$) using the RSA algorithm and Figure 3.12b illustrates 3 samples for point $\mathbf{x}_{32}$ (with $d_f = 0.081$ mm, $l_f = 12.872$ mm, $VF = 1.042\%$).

3.3.2 Random Process Evaluation/Data Acquisition

The XFEM model described in Section 3.2.2 deals with a 2D square matrix reinforced by 1D fibers subject to uniaxial tension. An SVE domain of 50x50 mm was considered (so, $\lambda = 50$ mm). Based on square elements of 0.5 mm size a 10,000 element mesh was



(a)

(b)

Figure 3.12: Three random realizations of sample points (a) $\mathbf{x}_{20}$ and (b) $\mathbf{x}_{32}$ obtained using RSA algorithm.

created. Figure 3.13 depicts the problem domain for a random sample. The left and bottom boundaries of the matrix are constrained against displacement in their normal directions. The right boundary of the domain is assigned a displacement of magnitude $u_0 = 0.1$ mm. The problem was solved using 250 fixed strain increments. Stress vs. strain data was recorded at each increment of the analysis. Table 3.2 lists the properties of the matrix and fiber used in the models. While the fiber length, diameter, and cohesive properties were sampled from descriptor space, the phase material properties were assumed to be constant. The numerical values used are typical for a cementitious matrix material and steel fiber.

Figure 3.15 illustrates the curves obtained for 20 SVEs at two locations, $\mathbf{x}_{20}$ and $\mathbf{x}_{32}$. Damage in the matrix material initiated at the fiber tips and propagated through the material to the edges of the SVE domain. As a result, higher concentrations of fibers lead to additional potential damage paths, reducing the peak strength of the composite and producing more uncertainty in the post-peak response. At first glance, it can be clearly seen

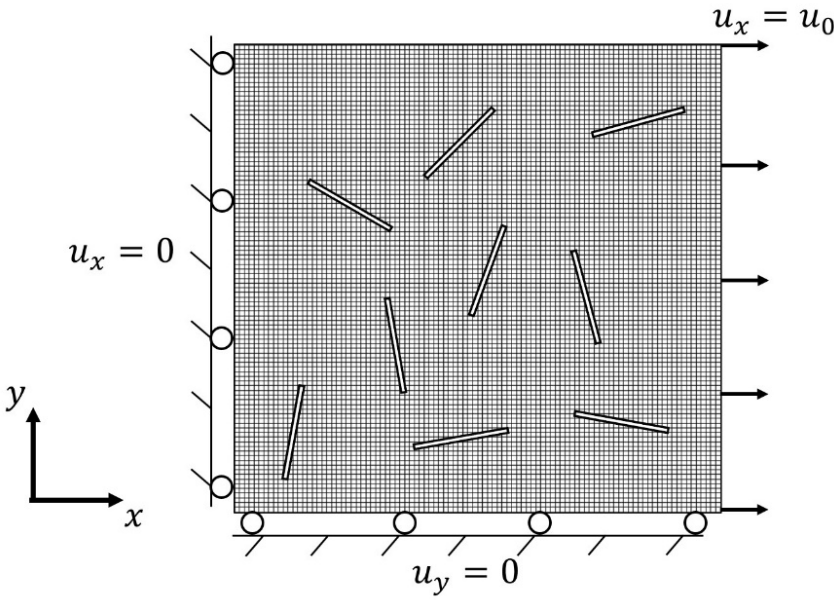


Figure 3.13: XFEM problem domain and boundary conditions.

Table 3.2: Phase material property constants for numerical models.

Property	Matrix	Fiber
Young's Modulus, E	14,000 MPa	207,000 MPa
Poisson's Ratio, ν	0.3	0.3
Nonlocal Damage Radius, l_c	0.75 mm	-
Nonlocal Damage Parameter, a	49,000	-
Nonlocal Damage Parameter, b	19.5	-

that $\mathbf{x}_{32}$ exhibits lower peak tensile strength and a smaller failure strain than $\mathbf{x}_{20}$. The response of each realization appears nearly identical in the elastic range where the response is controlled by the undamaged matrix material. Increased variation appears in the post-peak response due to the effect of the random fiber placement. It is observed that the stress-strain curves follow the general pre- and post-peak behavior described by the ACK model, as discussed earlier in this chapter. The following piecewise equation is selected to parametrize the curve:

$$\sigma(\varepsilon) = \begin{cases} (\frac{\sigma_i}{\varepsilon_i})\varepsilon & 0 \leq \varepsilon < \varepsilon_i \\ \sigma_i \mathrm{e}^{-\alpha(\varepsilon - \varepsilon_i)} & \varepsilon_i \leq \varepsilon \leq \varepsilon_f \end{cases} \tag{3.31}$$

The equation consists of an initial elastic region followed by an exponentially decaying post-peak softening region. Here σ_i is the peak stress at which the concrete matrix cracks and the fibers are engaged and ε_i is the strain corresponding to σ_i. ε_f is the strain at which the composite fails. α is an exponential decay parameter that controls the shape of the post-peak softening region. ε and σ are the strain and stress data to which the equation must be fit. The failure strain of the composite, ε_f, is defined as the strain at which the tensile stress has degraded to 10% of the peak value. The parameters are shown on a sample stress-strain curve in Figure 3.14. The area under the stress-strain curve represents the total strain energy capacity of the material in tension, known as the modulus of toughness, G^t. The area under the linear elastic region of the curve is known as the modulus of resilience, denoted by G^r and defined as the amount of energy that can be absorbed by the material prior to inelastic deformation. The modulus of resilience is calculated using Equation 3.32.

$$G^r = \int_0^{\varepsilon_i} (\frac{\sigma_i}{\varepsilon_i})\varepsilon d\varepsilon = \frac{1}{2}\sigma_i\varepsilon_i \tag{3.32}$$

The modulus of toughness is obtained by summing the modulus of resilience with the area

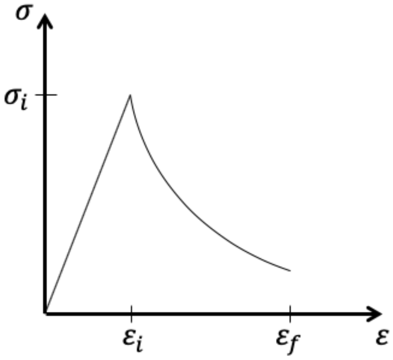


Figure 3.14: Stress-strain curve with elastic and post-peak region defined by parameters ε_i, ε_f, and σ_i.

Table 3.3: Composite response parameters measured in each simulation.

Identifier (y_q)	Material Property	Description
y_1	σ_i	Peak tensile stress
y_2	ε_i	Tensile strain at peak stress
y_3	ε_f	Strain at failure of composite
y_4	G^r	Modulus of resilience
y_5	G^t	Modulus of toughness

under the post-peak region of the stress-strain curve as given by Equation 3.33.

$$G^t = G^r + \int_{\varepsilon_i}^{\varepsilon_f} \sigma_i \mathrm{e}^{-\alpha(\varepsilon-\varepsilon_i)} d\varepsilon \tag{3.33}$$

The parameters that characterize the FRC stress-strain curve, denoted by $y_q,\ q = 1...5$, are listed in Table 3.3. The mean and standard deviation of the computed values for each sample point are listed in Appendix Tables A.3-A.4.

3.3.3 The Learning Machine

While clear differences may be observed in the curves obtained for $\mathbf{x}_{20}$ and $\mathbf{x}_{32}$ (Figure 3.15), the complex relationship between descriptor variables and composite properties cannot be obtained by directly comparing output quantities, because each sample is characterized by a point in the 6-dimensional space. Each of the six descriptors is assumed

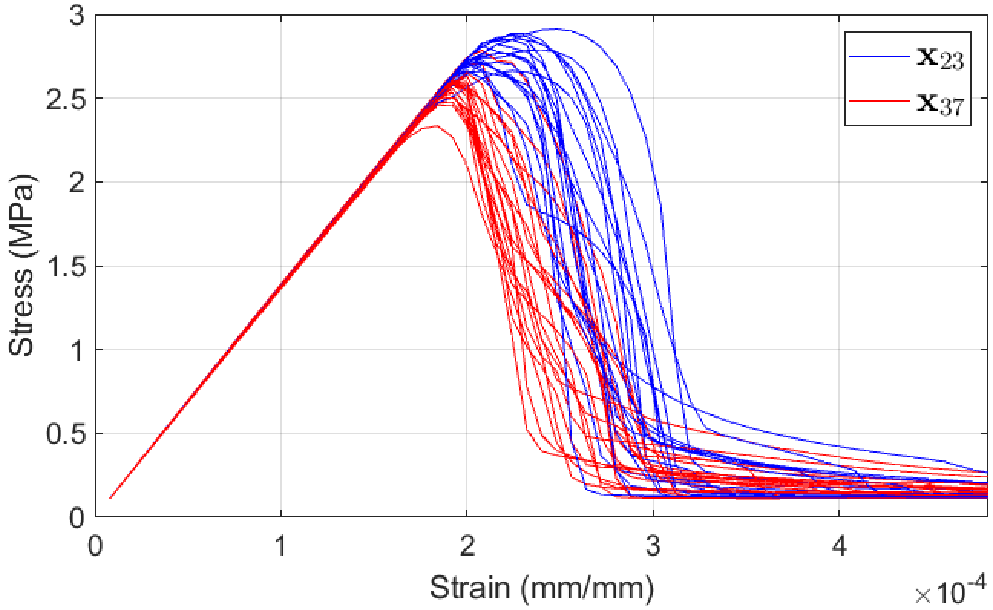


Figure 3.15: Sample tensile stress-strain curves for 20 realizations of 2 random points $\mathbf{x}_{20}$ and $\mathbf{x}_{32}$.

to have some influence on the measured material properties. To enable mapping from descriptor space to composite response space, Gaussian process regression is implemented to model the function $\mathbf{y} = F^{*}(\mathbf{x})$ (presented in Chapter 2). The work of Rasmussen provides a thorough state-of-the-art discussion of GP methods [103]. This section provides a brief overview of GP, focusing primarily on how the concepts can be applied to SFRC characterization and property prediction.

A GP is a collection of random variables with joint Gaussian distributions defined by its mean and covariance functions, similar to a multivariate Gaussian distribution. The expression for the random function $F^{*}(\mathbf{x})$ assumed to relate descriptor variables to output quantities is given by Equation 3.34.

$$F^{*}(\mathbf{x}) \sim \mathrm{GP}\Big(m(\mathbf{x}, \boldsymbol{\omega}_m), k(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\omega}_k)\Big) \tag{3.34}$$

Here, $m(\mathbf{x}, \boldsymbol{\omega}_m)$ and $k(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\omega}_k)$ represent the prior mean and covariance functions with hyperparameters $\boldsymbol{\omega}_m$ and $\boldsymbol{\omega}_k$, respectively. The covariance function governs the manner in which the value of $F^*(\mathbf{x}, \omega)$ changes throughout descriptor space. The squared exponential covariance function is commonly used in GP regression for its simplicity and flexibility. This function takes the form:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\Big(-\frac{1}{2l^2}(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)\Big) \tag{3.35}$$

where σ_f is a scaling parameter and l is the characteristic length scale parameter. The squared exponential covariance function provides that for two points $\mathbf{x}_i$ and $\mathbf{x}_j$ in close proximity to each other there is a high likelihood that the function will return similar output values at each point, with the likelihood decreasing exponentially as the distance between the points increases. The characteristic length influences the rate of the exponential decay. However, when the descriptor variables differ significantly in magnitude, it is not reasonable to assume a common characteristic length scale for all dimensions. In this work, the following anisotropic equivalent of the squared exponential covariance function is used:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\Big(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{\Lambda}^{-2}(\mathbf{x}_i - \mathbf{x}_j)\Big) \tag{3.36}$$

where $\mathbf{\Lambda}$ is a dxd diagonal matrix of characteristic correlation length values in each dimension of the descriptor space. For a given dataset the covariance matrix is constructed as:

$$\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) \tag{3.37}$$

The experimental data is partitioned into a training set T and prediction set P. A covariance matrix $\mathbf{K}_{TT}$ is constructed using the training points and $\mathbf{K}_{PP}$ is constructed using the predicted points. $\mathbf{K}_{TP}$ denote the covariance matrix between training and prediction points. The hyperparameters of the covariance function are determined such that the log marginal likelihood (Equation 3.38) is maximized, signifying that the probability of obtaining data

$\mathbf{y}$ from input $\mathbf{x}$ using model $F^*(\mathbf{x},\boldsymbol{\omega})$ is optimal.

$$\log p(\mathbf{y}_T|\mathbf{x}_T,\omega) = -\frac{1}{2}\mathbf{y}_T^{\mathrm{T}}(\mathbf{K}_{TT}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y}_T - \frac{1}{2}\log|\mathbf{K}_{TT}+\sigma_n^2\mathbf{I}| + \frac{N_T}{2}\log(2\pi) \tag{3.38}$$

In this equation N_T is the number of training points used, σ_n^2 is the variance of the noise in the experimental data, and $\mathbf{I}$ is an identity matrix with size equal to that of $\mathbf{K}_{TT}$. Once the parameters have been determined, the predictive means and variances are calculated using the following equations:

$$\mathrm{E}[\mathbf{y}_P] = \mathbf{K}_{TP}^{\mathrm{T}}(\mathbf{K}_{TT}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y}_T \tag{3.39}$$

$$\mathrm{Var}[\mathbf{y}_P] = \mathbf{K}_{PP} - \mathbf{K}_{TP}^{\mathrm{T}}(\mathbf{K}_{TT}+\sigma_n^2\mathbf{I})^{-1}\mathbf{K}_{TP} \tag{3.40}$$

To assess the accuracy of the GP in predicting each response metric y_q from the simulation data, k-fold cross validation was used, with $k = 5$. The 40 points in descriptor space were divided into five randomly selected subsets of 8 points each. In the first step, four of the sets were used as training points, with the remaining set used as test points. The process was repeated five times, using each of the subsets as test points in successive iterations. The root mean squared error (RMSE) was calculated using the following equation:

$$\mathrm{RMSE}[y] = \sqrt{\sum_{i=1}^{N}\frac{(y_i^* - y_i)^2}{N}} \tag{3.41}$$

where $N = 40$, y_i^* is the GP predicted value at point $\mathbf{x}_i$, and y_i is the value at point $\mathbf{x}_i$ obtained by XFEM. The RMSE provides a measure of the standard deviation of the model error. By comparing the RMSE to the magnitude of the mean of the predicted value, the reliability of the model can be assessed. The measured RMSE is in the same units as the output variable being predicted. The coefficient of variation (CV) provides a dimensionless measure of model uncertainty by normalizing the RMSE by the expected value of the output quantity.

Table 3.4: Comparison of error magnitudes obtained from GP regression of each output variable.

y	E[y]	RMSE[y]	CV[y]
σ_i (MPa)	2.58	0.097	3.76%
ε_i (mm/mm)	$1.98\text{x}10^{-4}$	$8.74\text{x}10^{-6}$	4.41%
ε_f (mm/mm)	$3.24\text{x}10^{-4}$	$3.55\text{x}10^{-5}$	10.96%
G^r (mJ/mm^3)	$2.57\text{x}10^{-4}$	$2.53\text{x}10^{-5}$	9.20%
G^t (mJ/mm^3)	$3.92\text{x}10^{-4}$	$3.10\text{x}10^{-5}$	7.91%

The CV is computed using Equation 3.42:

$$\mathrm{CV}[y] = \frac{\mathrm{RMSE}[y]}{\mathrm{E}[y]} \tag{3.42}$$

The RMSE calculated between the GP predictions and the model outputs is given in Table 3.4. The results indicate that the GP is very accurate in predicting σ_i and ε_i, where the CV is $< 5\%$. The model shows slightly more variability in predicting the modulus of resilience and modulus of toughness, but still show $< 10\%$ variation. The variability in the total strain to failure is the highest, with $\mathrm{CV}[y] = 10.96\%$.

3.4 Conclusion

In this chapter a method for multiscale modeling of random SFRC using a descriptor-based machine learning approach was demonstrated. A coarse scale model for the homogenized behavior of SFRC subjected to tension was presented. Key descriptors influencing the coarse scale model parameters were identified, reducing the design space to a finite number of dimensions. Samples were drawn from the descriptor space using stratified sampling techniques. Statistically representative microstructures were generated for each sample point and their tensile response was analyzed using an extended finite element method tailored specifically to the case of random SFRC with nonlinear damage. A Gaussian process was trained from the data and used to predict the coarse scale model parameters for several microstructures using only descriptors. The following conclusions can be drawn from the study:

1. A descriptor-based machine learning approach can be used to predict the elastic and inelastic response of a random SFRC with a high-degree of accuracy.
2. XFEM is a useful tool for efficiently analyzing the microstructural response of random SFRC due to its ability to model complex geometries with a fixed mesh.
3. The GP makes an effective surrogate model for the microstructural domain within the context of a sequential multiscale modeling approach. It foregoes the need for expensive explicit modeling of the fine-scale domain by using a microstructural database.
4. The GP was most accurate in predicting the peak stress and the corresponding strain in the composite, but was less accurate in predicting the failure strain of the composite, which is more strongly influenced by fiber randomness. This may be the result of non-inclusion of fiber fracture in the XFEM model, which was treated by defining failure strain as the point corresponding to a 90% reduction in strength.

Since the machine learning model is data-driven and non-parametric it is most useful in predicting the response in and around the sampled subspace, but can be retrained as more data becomes available. A discussion of potential future work on this topic is provided in Chapter 5.

Chapter 4

Vibrational Response of Composite Musical Instrument Strings

In this chapter a similar approach to characterization and prediction of properties using descriptors is followed, as was done in the case of random SFRC, to evaluate the response of pretensioned vibrating electric guitar strings. As opposed to acoustic guitars in which the sound box acts as an amplifier to produce audible sound, in an electric guitar the vibrations of plucked strings are captured up by means of electrical pickups (electromagnetic, piezoelectric, or optical) and amplified to produce audible sound. As a guitar string is plucked, it vibrates with a certain frequency depending upon its unsupported length, density, tension, and stiffness. Depending upon these parameters the string has a number of frequencies at which it will naturally vibrate producing a low intensity sound requiring amplification. These natural frequencies are known as the harmonics of the guitar string. The durability and damping characteristics of guitar strings are dependent upon several random variables, particularly the string material properties and cross section geometry. Due to the availability of several new wire materials for musical instrument strings, it is of interest to develop a computational model capable of linking the vibrational response of a composite string to its mechanical properties and geometry.

Depending upon the geometric configuration, modern electric guitar strings can be classified as either plain or wound. A plain string is characterized by a solid homogeneous circular cross section. The diameter of the string is commonly referred to as a gauge, measured in thousandths of an inch. For example, a gauge 10 string has a diameter of 0.010”. In general, strings less than 0.018” in diameter are of plain type. Larger diameter strings have composite cross sections formed by winding a wrap wire around a plain core wire. When a composite string is subject to tension, the force is primarily carried by the core wire. The purpose of the wrap wire is to increase the effective mass of the vibrating string so that

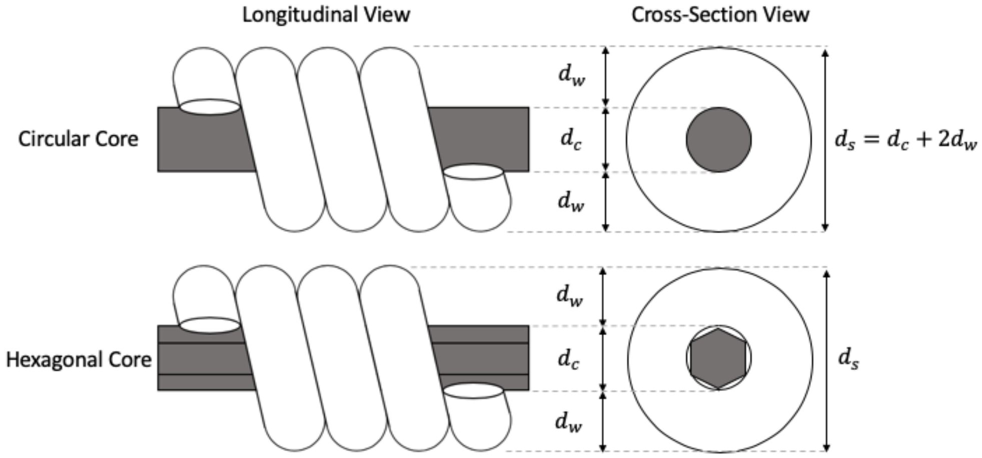
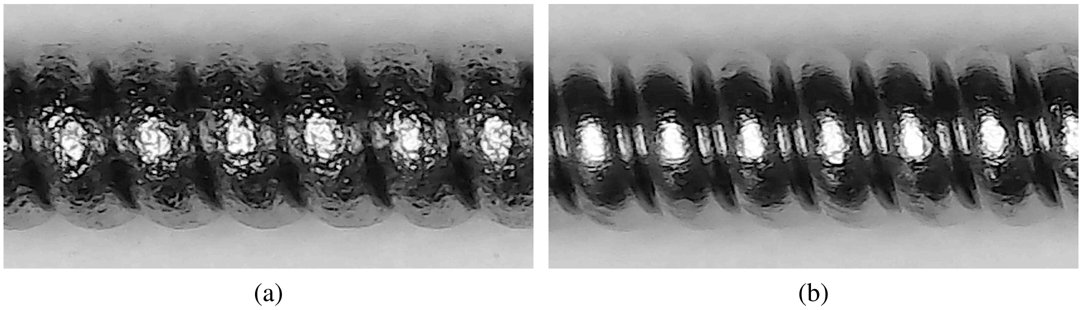


Figure 4.1: Comparison of geometries of roundwound strings with circular and hexagonal cores.

lower frequencies can be realized without significantly increasing the tension in the string. The bond between the core and the wrap wire relies on the friction between the two materials. Hexagonal core wires are typically used to achieve optimal anchorage between the core and the wrap because the sharp edges of the core tend to restrict slippage between the two. However, many manufacturers also use round core wires to create a different sound. Figure 4.1 illustrates the geometry of composite strings with circular and hexagonal cores. The diameter of the composite string, d_s, is the sum of the core diameter d_c and twice the wrap diameter d_w. The composite geometry can be described by the core-to-wrap ratio (d_c/d_w), the diameter of the core relative to the wrap. Different string manufacturers use different core-to-wrap ratios for a given gauge of string. Wound strings can be further classified based on the cross section shape of the wrap wire used. Roundwound strings consisting of a wrap wire with a circular cross section are most commonly used. Use of a semicircular or rectangular wrap wire results in halfwound and flatwound strings, respectively.

The most common material used for plain strings and core wire is music wire. Per ASTM A228 [104], music wire is a cold-drawn high-carbon steel consisting of 0.7-1.0% C, 0.2-0.7% Mn, 0.025% P, 0.03% S, and 0.10-0.30% Si. The material is intended especially for use in high stress or fatigue prone environments. Some manufacturers use other



(a) (b)

Figure 4.2: Surface profiles of gauge 46 roundwound strings with (a) iron-cobalt alloy wrap and (b) pure nickel wrap.

variations of high-carbon steel (HCS) with unspecified chemical composition. One new core material is maraging steel, an iron alloy with 6-16% Co, 15-20% Ni, 2-6% Mo, and <2% Ti [105]. Each core wire material is characterized by high strength, toughness, and malleability.

The most prevalent wrap wire material is nickel-plated steel (NPS). NPS is produced by electroplating a low carbon steel wire with 2-8% Ni by weight. Other available wrap materials include pure nickel, stainless steel (<0.1% C, 10-19% Cr, <4.5% Cu,, 0.2-3.0% Mn, <6% Mo, 4-10% Ni, <0.05% P, <0.1% S, <1.5% Si) [106], and iron-cobalt alloy (8-35% Co, 0.25-7% Cr, 0.25-1% Mn, 0.25-7% Ni, 1-5% V) [107]. Figure 4.2 shows the surface profiles of iron-cobalt and pure nickel wrap wires studied in this work obtained using a digital microscope. There are significantly more surface defects present in the iron-cobalt alloy wrap, which will influence the vibration of the string.

The organization of this chapter is as follows: Section 4.1 provides an overview of the physics of string vibration and methods for modeling the response of idealized and real strings. Methods for capturing and parameterizing an audio signal are also discussed. Section 4.2 discusses essential concepts of music theory and describes the electric guitar as an electromechanical system. Section 4.3 formalizes the problem in the context of the descriptor-based machine learning approach outlined in Chapter 2 and details an experimental method for measuring the response of strings. Conclusions drawn from the study are discussed in Section 4.4.

4.1 Governing Equations

4.1.1 String Vibration

An ideal string has no bending stiffness based on the assumption that the transverse displacement is small relative to the length of the string. Figure 4.3 depicts the deformed shape of a string with length l and linear density μ at instant t in time. The string is subject to a tensile force T, which is assumed to be constant when the displacement $y(x,t)$ is small. Dynamic equilibrium of net y-component of the tensile force and the inertial force in the differential element dx at any instant of time yields the equation of motion in the y-direction:

$$T\left(\theta+\frac{\partial\theta}{\partial x}dx\right)-T\theta=\mu dx\frac{\partial^2 y}{\partial t^2} \tag{4.1}$$

As $\theta=\frac{\partial y}{\partial x}$, Equation 4.1 can be rewritten as:

$$\frac{\partial^2 y}{\partial x^2}=\frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} \tag{4.2}$$

where the wave speed c of the vibrating string is related to its tension and density by Equation 4.3.

$$c=\sqrt{\frac{T}{\mu}} \tag{4.3}$$

A true end support on a stringed instrument is nearly fixed but may yield slightly [87]. Assuming that the displacement at the ends of the strings is ideally fixed, the solution $y(x,t)$ to Equation 4.2 must satisfy the elliptic boundary conditions $y(0,t)=0$ and $y(l,t)=0$. Substitution shows that waves of the form given by Equation 4.4 satisfy the boundary conditions.

$$y=A\sin(Kx\pm\omega t) \tag{4.4}$$

$\omega=2\pi v$ is the angular frequency and K is the wave number. Let $y_1=A\sin(Kx-\omega t)$ denote

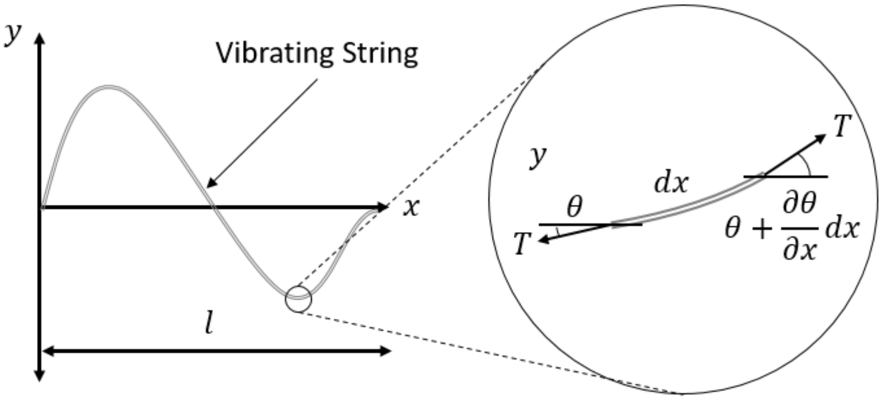


Figure 4.3: String vibration problem formulation.

a continuous sine wave traveling in the positive x-direction and $y_2 = A\sin(Kx + \omega t)$ is a similar wave traveling in the opposite direction. If the waves of equal frequency, amplitude, and speed are reflected perfectly off of the ends, superposition of the waves yields:

$$y = y_1 + y_2 = A[\sin(Kx - \omega t) + \sin(Kx + \omega t)] = 2A\sin(Kx)\cos(\omega t) \tag{4.5}$$

The trigonometric function in Equation 4.5 satisfies the boundary condition $y(l,t) = 0$ when $K = n\pi/l$ only, where n is a positive integer. Equation 4.5 is expressed in terms of the vibration mode n in Equation 4.6.

$$y_n = 2A_n \sin(K_n x)\cos(\omega_n t) = 2A_n \sin\left(\frac{n\pi}{l}x\right)\cos\left(\frac{n\pi c}{l}t\right) \tag{4.6}$$

When a string with fixed ends is plucked, the response can be expressed as a sum of many waves in the form of a generalized Fourier series:

$$y(x,t) = \sum_{n=1}^{\infty} [A_n \cos(\omega_n t) + B_n \sin(\omega_n t)] \sin\left(\frac{n\pi}{l}x\right) \tag{4.7}$$

The initial magnitude of each mode of vibration is defined by Fourier coefficients A_n and

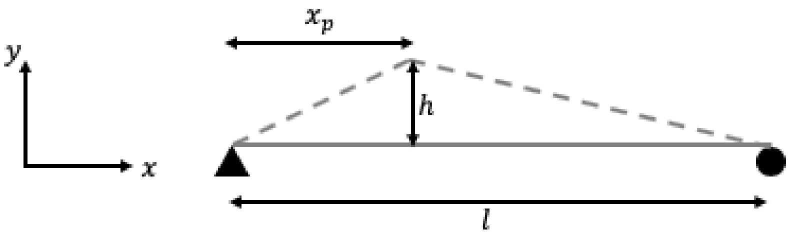


Figure 4.4: String vibration boundary conditions.

B_n, which are calculated by the following equations:

$$A_n = \frac{2}{\omega_n l}\int_0^l \dot{y}(x,0)\sin\frac{n\pi x}{l}dx \tag{4.8}$$

$$B_n = \frac{2}{l}\int_0^l y(x,0)\sin\frac{n\pi x}{l}dx \tag{4.9}$$

The initial conditions $y(x,0)$ and $\dot{y}(x,0)$ can be used to determine the values of A_n and B_n. For a string initially at rest, $\dot{y}(x,0) = 0$ and $A_n = 0$ for all values of n. The boundary condition $y(x,0)$ is defined according to the location and amplitude of the pluck, as shown in Figure 4.4. The initial configuration of the string between the supports is shown by the solid gray line. The configuration at the time when the pluck is applied is shown as a dashed line. The location of the pluck is x_p and the plucked height is given as h. The initial displacement boundary condition at $t = 0$ can then be written as:

$$y(x,0) = \begin{cases} \left(\frac{h}{x_p}\right)x & 0 \le x \le x_p \\ \frac{h}{l-x_p}\left(1-\frac{x}{l}\right) & x_p \le x \le l \end{cases} \tag{4.10}$$

In Figure 4.5, B_n is plotted for three different plucking point locations along a 25.5” string. β denotes the normalized location of the pluck along the string, x_p/l. The amplitude of harmonic n changes significantly with a change in x_p.

The fundamental frequency of the string is related to its tension and length by the fol-

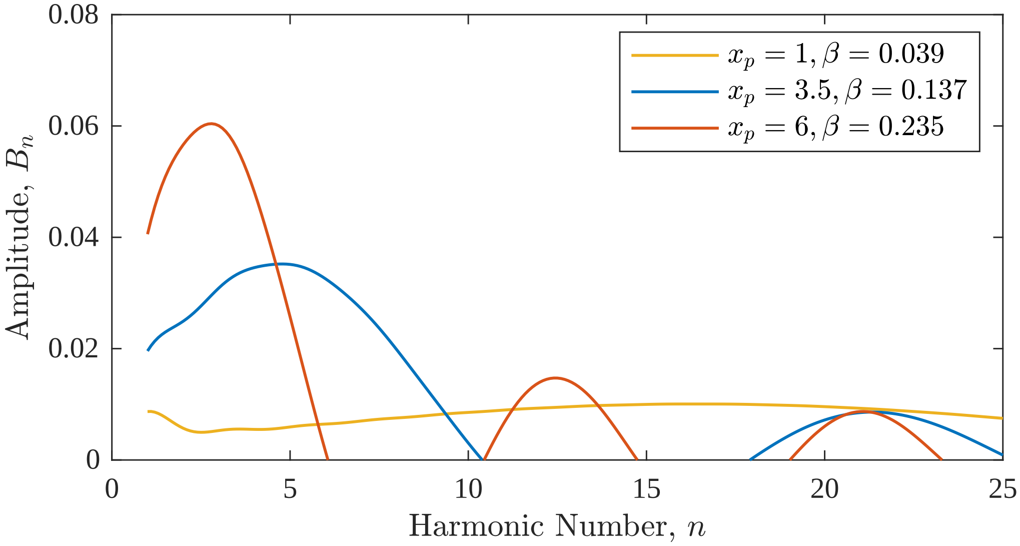


Figure 4.5: Harmonic content for three different plucking locations.

lowing equation:

$$f_1 = \frac{1}{2l}\sqrt{\frac{T}{\mu}} \tag{4.11}$$

An ideal string, which is infinitely flexible, has resonant frequencies which are integer multiples of the fundamental, that is, harmonics f_n will exist at integer multiples of the fundamental frequency nf_1 for $n = 2, 3, 4 \ldots \infty$. However, in real strings, there is some inharmonicity due to the stiffness of the material and the presence of geometric nonlinearities. The stiffness influences particularly the higher vibration modes. Equation 4.2, which governs the motion of an ideal string, can be modified to account for the stiffness of the string as:

$$\mu \frac{\partial^2 y}{\partial t^2} = T \frac{\partial^2 y}{\partial x^2} - EI \frac{\partial^4 y}{\partial x^4} \tag{4.12}$$

where E is Young's Modulus of the wire and I is the moment of inertia of the cross section. Equations 4.13 and 4.14 can be used to calculate the inharmonic frequencies f_n^* of a stiff string assuming pinned ends and clamped ends, respectively [75].

$$f_n^* = nf_1\left(1 + Bn^2\right)^{1/2} \tag{4.13}$$

$$f_n^* \approx n f_1 \left(1+Bn^2\right)^{1/2} \left[1+(2/\pi)B^{1/2}+(4/\pi^2)B\right] \tag{4.14}$$

where the unitless stiffness parameter B is defined by the relationship:

$$B = \frac{\pi^2 EI}{Tl^2} \tag{4.15}$$

Calculation of B is straightforward for a plain wire with a uniform cross section. Figure 4.6 shows the theoretical effect of string stiffness on inharmonicity for the first 20 partials. The results were calculated using Equations 4.13-4.15 for a string of plain geometry. The length of the scale is 25.5" and the diameter of the circular cross section is 0.017". Young's modulus is 30,000 ksi and the density of the material is 0.284 pci, similar to music wire. The string is tensioned to 16.6 lbs, which results in a fundamental frequency of 196 Hz. The y-axis of the plot has been normalized by the partial number for readability. For small values of n, the deviation between the ideal string and the real string is minimal, but the difference becomes apparent as n increases. The clamped end approximation yields larger deviations from the theoretical response than the pinned case. For a wound string, the composite Young's modulus is influenced by the bond between core and wrap wire, which is difficult to characterize. Explicit calculation of the moment of inertia for the composite string is also a complicated problem since the spiral pattern of the wrap wire creates a nonuniform cross section. For this reason, previous work [88, 89, 90] on the inharmonic response of wrapped strings has focused on the inverse problem of fitting Equations 4.13-4.14 using experimental data.

Musical instruments produce sound using a combination of the vibration of several system components, which are coupled to some extent. In the case of the electric guitar, several oscillating strings are attached to the body at the bridge and nut. Each string has potential energy due to its stiffness and inertial energy due to its mass. The solid body of the electric guitar is intended not to color the sound significantly, and not to absorb vibrational energy from the strings [108], however some damping due to body coupling is

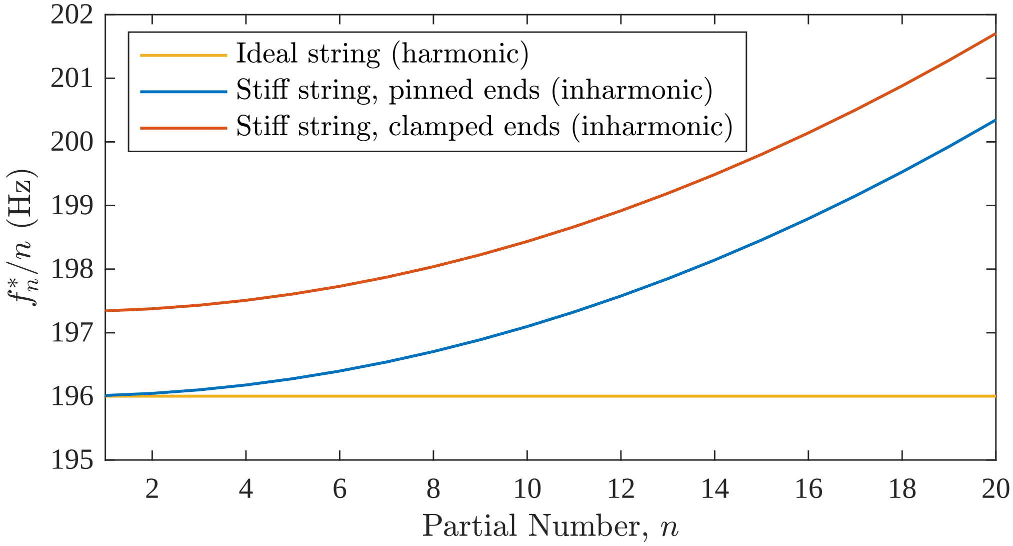


Figure 4.6: Theoretical inharmonicity due to string stiffness.

inevitable. Damping is present in a string due to a combination of air damping, internal material damping, and energy loss at the supports [73]. Frequency-dependent damping characteristics of strings have been studied in previous work [68, 89, 90, 109]. Viscous drag on a string lowers the modal frequencies slightly and produces exponential decay of amplitude. For the size of strings commonly used in musical instruments, the retarding force due to air damping is expressed as:

$$F_r = 2\pi^2 \rho_a f v r^2 l \left(\frac{\sqrt{2}}{M} + \frac{1}{2M^2} \right) \tag{4.16}$$

where r is the radius of the string, ρ_a is the density of air, f is the frequency of vibration, and v is the velocity. $M = (r/2)\sqrt{2\pi f/\eta_a}$. η_a is the kinematic viscosity of air. For oscillation at a single frequency f, the amplitude decays exponentially with time constant τ_1:

$$\tau_1 = \frac{\rho}{2\pi\rho_a f} \left(\frac{2M^2}{2\sqrt{2}M + 1} \right) \tag{4.17}$$

Where ρ is the density of the string. In the equation of motion for a string (4.12), the string is treated as an elastic isotropic material with Young's Modulus E. In reality, the string behavior is viscoelastic; when a stress is applied, there is a slight delay before the corresponding strain is fully realized. To capture this effect, Young's modulus may be treated as a complex number, with $E = E_1 + jE_2$. Substituting this relationship into Equation 4.12 yields the decay time τ_2 due to internal damping:

$$\tau_2 = \frac{1}{\pi f}\frac{E_1}{E_2} \tag{4.18}$$

This type of damping is usually negligible for solid metal strings, but there is energy loss due to internal friction between components in wrapped strings. The decay time due to energy loss through the supports is given by the following equation:

$$\tau_3 = \frac{1}{8\mu l f^2 G} \tag{4.19}$$

G is the conductance, defined as the real part of the mechanical admittance. The decay time τ in the presence of all three types of damping is calculated by the following equation:

$$1/\tau = 1/\tau_1 + 1/\tau_2 + 1/\tau_3 \tag{4.20}$$

For thin metal wires, as in the case of guitar strings, air damping contributes most significantly to the total damping force.

4.1.2 Signal Analysis

The signal produced by a vibrating string can be analyzed in the time domain or frequency domain. Within the time domain, it may be of interest to study the decay time of the vibrations following excitation of the system, or the shape of the decay envelope. In the frequency domain, it is often useful to study the concentration of power over certain frequency bands. Digital signal processing concepts relevant to the instrument string problem are discussed in this section, including analog to digital conversion, noise in signals, and methods for representing a signal in time and frequency domains.

In the previous section the signal produced by a vibrating string was defined as $y(x,t)$. Let $y(t),\ t \in [0,T]$ denote the displacement of the string with respect to time at an arbitrary fixed location x, where T is the duration of the signal in seconds. The signal $y(t)$ is continuous in time and amplitude domains. In order to record and analyze the signal, $y(t)$ must be discretized in each domain. This is achieved using an analog-to-digital converter (ADC). An ADC captures samples of the audio at some sample rate, F_s, with the data size of each sample referred to as bit-depth. Sampling discretizes the signal in the time domain, while the bit-depth of the samples determines the quantization resolution in the amplitude domain. Let the discretized signal $y[n]$ be defined as a vector of equally spaced sampled values from the signal $y(t)$. If the signal is sampled at a rate of F_s (samples/second), the length of the sample vector $y[n]$ is given by $N = F_sT$. In general, parentheses are used to indicate functions of continuous variables and brackets are used for the case of discrete variables.

The work of Nyquist and Shannon [110, 111] showed that in order to accurately sample a wave of frequency f, a sampling frequency of at least $2*f$ Hz is required to detect the peaks and the troughs. For the musical instrument string problem, the frequency range of interest is the range of human hearing, which is approximately 20 Hz - 22,500 Hz for the average person. The average dynamic range of human hearing is 0 dB sound pressure level (SPL) to 140 dB SPL. According to the Nyquist sampling theorem, a sample rate of at least 44.1 kHz is required for audio applications to capture the entire spectrum of human hearing. As the bit depth of each sample is decreased, noise in the signal becomes more apparent. The equation for signal-to-noise ratio (SNR) as a function of bit-depth Q is:

$$\mathrm{SNR} = 20\log_{10}(2^Q)\ \mathrm{dB} \tag{4.21}$$

Using Equation 4.21 with a bit depth of 24 results in a theoretical SNR of 144.5 decibels, which is approximately equal to the range of human hearing. Modern analog-to-digital conversion technology is capable of capturing 24 bit-samples at a sample rate of 192 kHz. At this rate, frequencies of up to 96 kHz can be captured with a theoretical dynamic range of 144 dB SPL. Increasing the sample rate results in a digital signal that will more accurately represent the analog signal. Modern analog-to-digital audio converters commonly use sampling frequencies up to 192 kHz, which results in more than 8 samples per oscillation of a 22,500 Hz wave.

Fourier Transform (FT) is one of the most commonly used tools for frequency domain decomposition of a signal. The FT, denoted $Y(\omega)$, of a continuous signal $y(t)$ is defined in Equation 4.22.

$$Y(\omega) = \int_{-\infty}^{\infty} y(t)e^{-j\omega t}dt \tag{4.22}$$

$Y(\omega)$ is a complex-valued scalar, defined in the frequency domain, $\omega \in (-\infty, \infty)$. The FT maps a function of a single variable (time) into a function of another single variable (frequency). For a certain frequency ω the scalar part of $Y(\omega)$ indicates the prominence of ω in the signal, and the imaginary part indicates the phase shift. The components $Y(\omega)$ are averaged over the time domain, so they do not provide precise information on how the frequency components decay over the duration of the signal.

The FT is derived for a continuous periodic signal of infinite length. For signals that have been sampled in time, the Discrete Time Fourier Transform (DTFT) applies:

$$Y[\omega] = \sum_{n=-\infty}^{\infty} y[n]e^{-j\omega n} \tag{4.23}$$

In this case, $\omega \in [-\pi, \pi]$. For a finite length signal sampled in time, the Discrete Fourier Transform (DFT) is used:

$$Y[k] = \sum_{n=0}^{N-1} y[n]e^{-j\omega_k n} \tag{4.24}$$

Where, k is an index corresponding to the sampled frequency $\omega_k = \frac{2\pi k}{N}$, $k = 0, 1, ..., N-1$.

N denotes the length of signal $y[n]$ in samples.

When there is a discontinuity between the last sample of one period and the first sample of the next period (which is the case for a finite length signal), artifacts will be present in the DFT, which is a phenomenon called spectral leakage. Windowing methods were introduced in 1946 as an attempt to address the limitations of FT in analyzing finite-length signals [112]. In this approach, known as the short-time Discrete Fourier transform (STDFT), the signal $y[n]$ is multiplied by a smooth symmetric function $w[n]$ centered at sample index m. The spacing and amplitude of artifacts in the frequency spectrum of a non-periodic signal can be controlled by applying a window function to the signal of interest. The STDFT is given by Equation 4.25.

$$Y[k,m] = \sum_{n=0}^{N-1} y[n]w[n-m]e^{-j\omega_k n} \tag{4.25}$$

The window function is designed to assume nonzero values on the domain $n \in [m-N/2, m+N/2]$, where N is the width of the window, and zero elsewhere.

Many types of windows have been introduced in literature for different applications. Generalized cosine windows are popular due to their smoothness and ease of implementation. Equation 4.26 defines the raised cosine window function.

$$w[n] = \alpha - \beta \cos\left(\frac{2\pi n}{N-1}\right) \tag{4.26}$$

When $\alpha = \beta = 0.5$ the function is known as the Hanning window function. Another popular variation is the Hamming window, where $\alpha = 0.54$ and $\beta = 1 - \alpha = 0.46$. The STDFT provides a trade-off between frequency resolution and time resolution. The precision of signal characterization in the time-domain is dependent on the window size. If a large window is used, high frequency resolution is obtained, but the frequencies are smeared over a larger region of the time domain. If a short window is used to obtain greater time localization, high-frequency components of the signal are lost. Computation of the DFT requires

N^2 floating-point operations and is inefficient for large amounts of data. For this reason, the fast Fourier transform (FFT) is preferable in many applications. The FFT is a more efficient algorithm capable of computing the DFT in only $N \log N$ operations. Analyzing a time-based signal using multiple overlapping windows of constant size can provide information on how the frequency components of the signal vary over time. A plot that depicts this change in frequency content with respect to time is known as a spectrogram.

All real signals contain noise from a combination of sources. Noise can be inherent in the system or in the environment in which it operates, or can be introduced by the tools used to induce, capture, and transmit the signal. Techniques exist to identify the noise part and remove it from the measured signal to recover the signal of interest. Such techniques generally includes signal decomposition, thresholding, and reconstruction [113]. The simplest method for noise elimination is filtering, in which a threshold amplitude is defined and frequencies below the threshold are neglected. However, in practice, the engineer must be conscious of the frequency range of interest and choose thresholding values with care, which can make the process tedious.

4.2 Musical Problem Background

4.2.1 Tuning and Timbre

A musical note is classified by its fundamental frequency. As per international standard tuning [114], A4 is the name given to a note with fundamental frequency f_1 =440 Hz. An octave is defined as the interval between a note and another note with double or half of its frequency. Thus, A3 corresponds to f_1 =220 Hz and A5 corresponds to f_1 =880 Hz. Western music uses a twelve-tone equal temperament system for pitch classification, where an octave is divided into 12 notes, with the frequency ratio between subsequent notes of pitch f_1 and f_2 given by the equation $f_2/f_1 = 2^{1/12}$. The interval between subsequent pitches on a twelve-tone equal-tempered scale is referred to as a semi-tone. Each semitone is further divided into 100 cents. For reference, Figure 4.7 depicts a standard 88-key piano,

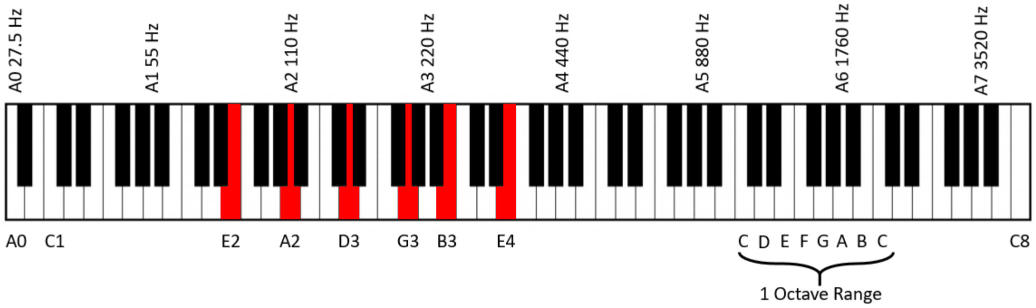
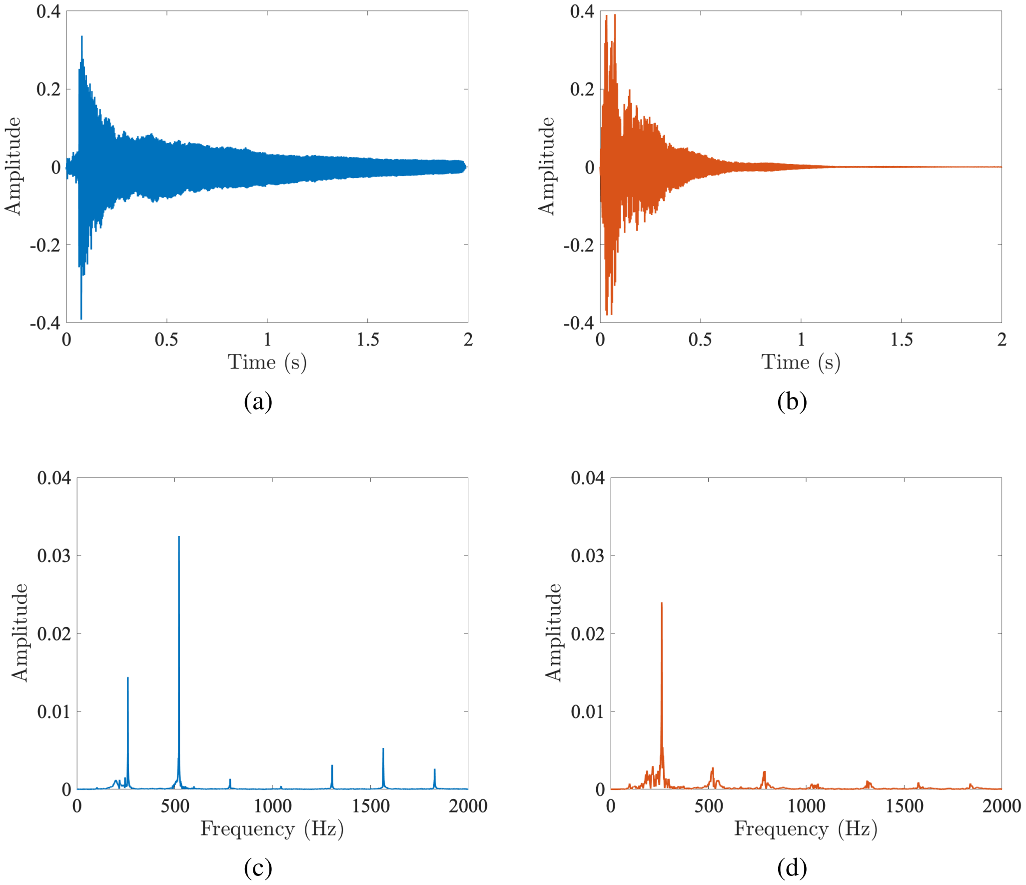


Figure 4.7: Musical note frequencies and intervals.

with each octave of A labeled with its fundamental frequency. The fundamental frequencies of the 88 keys range from A0 (27.5 Hz) to C8 (4186.0 Hz). The range of fundamental frequencies for a 6-string electric guitar with 24 frets is E2-E6 (82.4-1318.5 Hz) in standard tuning. The notes highlighted in red in Figure 4.7 correspond to the tuning of the 6 strings of the electric guitar.

A string's timbre is defined by the unique time-dependent combination of overtones it produces when the musician plays a note, and can be classified quantitatively using spectral analysis. The timbre of a string can be fully defined by measuring the magnitude of the partials in the signal, the inharmonicity of the partials, and the frequency-dependent damping parameters. In the case of the wound electric guitar string, altering the core-to-wrap ratio, the geometry of the core wire, and the material properties of the wires will influence the timbre of the sound it produces. Timbre is also influenced by the material properties of the instrument body, the force mechanism by which sound is produced, and in the case of electronic instruments, the circuit components in the signal path.

Figure 4.8 compares the frequency spectrum obtained from a guitar and a cello plucking C4 (f_1 =261.6 Hz) averaged over 2 seconds following the pluck using FFT. The spectrum is plotted on the interval $f \in [0, 2000]$ Hz so that the fundamental and following six partials can be observed. In both cases, the lowest significant frequency component in the signal is at f_1. In the case of the guitar, however, the second harmonic at 523.2 Hz is largest in



(a) (b) (c) (d)

Figure 4.8: 2-second time signal of note C4, f_1 =261.6 Hz, plucked on (a) guitar and (b) cello. Frequency spectrum obtained from FFT for (c) guitar and (d) cello.

magnitude. The guitar and cello are compared here to illustrate a rather extreme difference in timbre that can be heard as well as visually identified from the plot. The difference between timbre of different types of electric guitar strings is more subtle, but nonetheless present, and is examined in Section 4.3.2.

4.2.2 The Electric Guitar System

The development of the electric guitar began in the 1920s. Early models used piezoelectric sensors or microphones to capture the vibrations of traditional acoustic guitars with hollow bodies. These models had problems with signal distortion and feedback, as the signal was influenced more by the body response than the string response. A solution was introduced in the form of the electromagnetic pickup, which detects vibrations directly from the strings rather than the body, and hence requires the use of magnetic strings. Electromagnetic pickups are the standard for modern electric guitars. The solid body electric guitar was developed in the 1940s and is the most common and popular type of electric guitar today. By using a solid slab of wood with increased mass, strings are able to vibrate for a longer period of time, with less energy transferred to the body at the bridge. In 1948, Fender released the solid body electric guitar that would later become known as the Telecaster, and in 1952 Gibson released the Les Paul. These models remain popular today and are used around the world. Six-string guitars are most prevalent, but less common models utilize anywhere between 4 and 8 strings.

Figure 4.9 shows a typical electric guitar configuration. The electric guitar consists of two main parts - the body and the neck. The length of the span of the vibrating string is the distance between the bridge on the body and the nut on the neck. These end supports are illustrated in Figure 4.10. The distance between the bridge and the nut is termed the scale length. Common values for scale length are 25.5” for Fender guitars and 24.75” for Gibson guitars. Other makers are known to use different scale lengths that may vary between 22” and 28”. When a new string is placed on a guitar it is led through a hole in the bridge where the ball end catches, providing support at one end. The plain end of the string is typically wound around a post called a tuning peg behind the nut. Rotating the post changes the tension in the string, which is how the string is tuned to a reference pitch. The neck has several intermediary supports called fret wires. When the guitarist’s finger presses the string against a fret wire, the span of the string is shortened to the distance between the bridge and the fret wire, resulting in a change of pitch. Table 4.1 lists the pitch and diameter of typical strings for a 6-string guitar with a 25.5” scale in standard tuning. Smaller diameters are used for strings tuned to higher pitches and large diameters are used for low frequencies so that the tension in each string is similar and the longitudinal torque applied to the neck of the guitar is minimized.

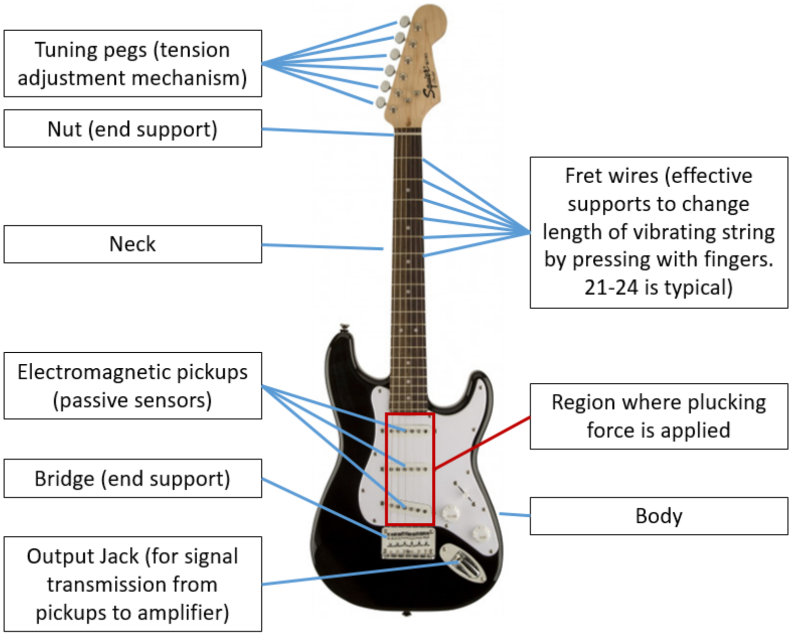
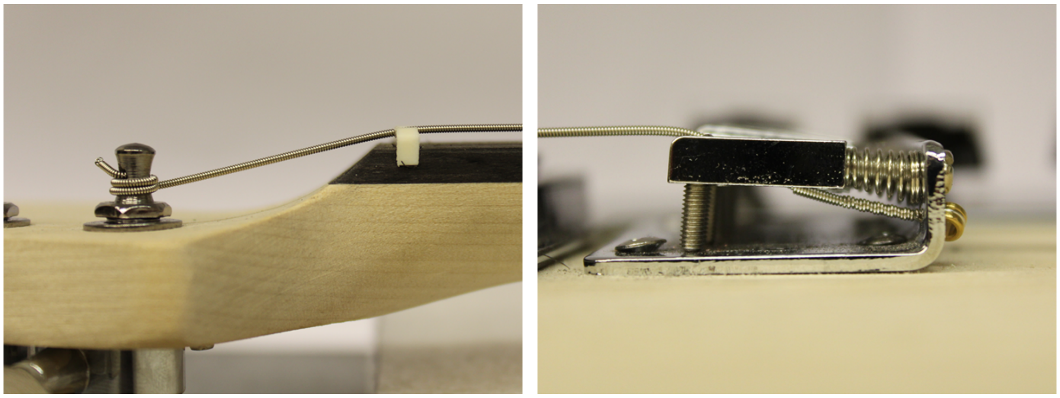


Figure 4.9: Typical electric guitar configuration.



(a)

(b)

Figure 4.10: Typical guitar string boundary conditions at the (a) nut and (b) bridge.

Table 4.1: Standard guitar tuning and string configurations.

String Number	Note	f_1 (Hz)	Typical Diameter (in)	Geometry
1	E4	329.6	0.009-0.011	Plain
2	B3	246.9	0.011-0.014	Plain
3	G3	196.0	0.016-0.018	Plain or Wound
4	D3	146.8	0.024-0.028	Wound
5	A2	110.0	0.032-0.038	Wound
6	E2	82.4	0.042-0.048	Wound

A magnetic pickup for a standard six-string guitar consists of six poles (one for each string) wrapped with either one or two multi-turn coils of copper wire. When an oscillating string moves closer to the pole, the magnetic field relaxes, and when the string moves away, the magnetic field contracts. This induces a voltage in the coil which is transmitted to an amplifier or signal processing device via 1/4” phone jack output. The voltage output from the pickup is proportional to the velocity of the string. The two most common types of passive pickups are single-coil and double-coil, the latter of which are also known as humbuckers. Single-coil pickups are susceptible to noise from AC power sources. In the United States, the AC noise has a frequency of 60 Hz. Humbucking pickups are designed to suppress the presence of noise in the signal. A humbucker has a second coil wound in the opposite direction, which removes the noise by way of phase cancellation. Humbucking pickups have lower resonant frequencies and more high-frequency cutoff than their counterparts. Most pickups are passive and do not require electrical power to transmit signal, however active pickups that require a power source are sometimes preferred.

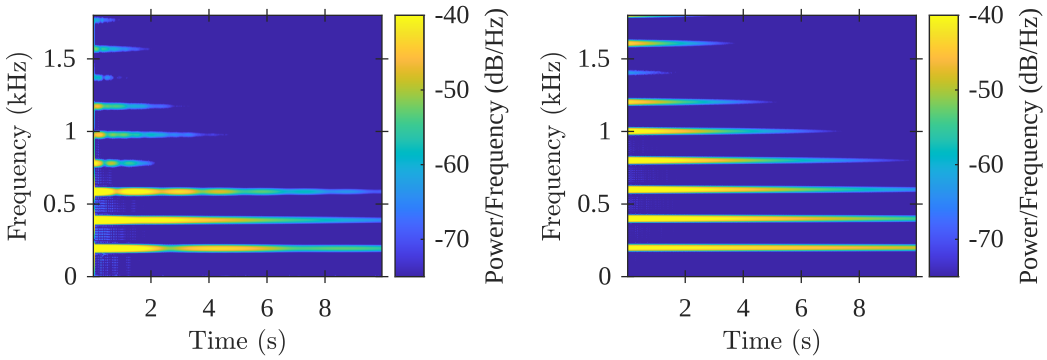
The location of the pickup along the string influences the magnitude and frequency of partials present in the signal [115]. The magnetic field between the string and the pickup creates a stiffness effect that causes an increase in inharmonicity. Since the magnitude of the vibrating string’s displacement varies along its length, changing the sensor location will affect measured inharmonicity. If the pickup is placed at a normalized distance β along the length of the string, there will be no output of the partial $1/\beta$ or any multiple of that frequency because the partials would have nodes at the pickup location. If the pickup is placed close to the bridge ($\beta \to 0$), some of the lower and higher partials will be attenuated, but the full range of harmonics will be present in the signal. If the pickup is placed closer to the fingerboard, there will be fewer harmonics in the signal but the strength of the lower partials will be more significant. Electric guitars typically have 1, 2, or 3 pickups, which can be alternated between by using a switching system. Using the switch, a user can blend the signal captured by pickups in different locations.

Guitarists either use a plectrum called a pick to excite the strings, or their own fingers. The type of plectrum used will also influence the signal produced by the instrument. The string is typically in contact with the fingers or plectrum for about 100 ms. While the plectrum is in contact with the string, there are local torsional and transverse forces acting on the string. Using a heavier plectrum results in a longer contact period, whereas using a stiffer plectrum reduces the contact time [108].

Instrument strings undergo damage from mechanical and environmental wear during their design life. Damage begins to accumulate the moment a new string is placed on an instrument. When a new string is tuned, the mean force is set in the range of 10-20 lbs depending on the diameter of the string and the desired fundamental frequency. When a wound string is initially tensioned, there is some slippage between the core and wrap wire. When the string is plucked, the force oscillates within a few pounds of the mean stress, with the magnitude and frequency of the short-scale oscillations dependent upon the musician and the style of music being played. After some time, the player notices that the strings have stretched and the mean stress is re-adjusted accordingly to maintain the pitch. As the strings plastically deform over their lifecycle, the mean stress required to achieve the desired fundamental frequency may increase. Over time, the cross section of the strings in the neck region is reduced in area from being repeatedly pressed against the fret wires. Mechanical wear is also observed in used strings from erosion due to impact where the plectrum strikes near the pickups and bridge. If the instrument is stored in a particularly cold or hot environment between uses, the mean stress will increase or decrease as the strings experience thermal elongation or contraction, requiring the user to retune before next use. Environmental wear is also present in used strings from exposure to moisture and oxygen, as well as oils and dirt introduced from the fingers of the musician. For this reason most manufacturers package their strings in corrosion-inhibiting envelopes for storage and transport. It is possible for a string to be used well past the onset of damage. While abrasion, erosion, and corrosion may weaken the string and change its harmonic signature, the string is usable as long as it has not fractured.

Finite element models have been used effectively to model the time-dependent response of a plain undamaged string [116]. To demonstrate this a preliminary experiment was conducted to capture the signal produced by a round ASTM A228 gauge 17 wire tuned to G3 (196 Hz). The length of the string was 25.5" and it was plucked at a normalized distance $\beta = 0.147$. A spectrogram of the resulting signal is shown in Figure 4.11a. A finite element model of the experiment was then constructed in Abaqus. An elastic isotropic material model was used for the wire with density $\rho = 7850$ kg/m^3, Young's modulus $E =$ 210.29 GPa, and Poisson's ratio $\nu = 0.3$. A modal dynamics analysis was used to measure the response of the string for 10 seconds after it was plucked. In the absence of damping, the wire will oscillate indefinitely with no decrease in magnitude. Composite damping was used to model the time-dependent decay observed in the experiment. The damping parameter was calibrated to minimize the difference between the model and experiment. A plot of the frequency spectrum with respect to time obtained from the model is shown in Figure 4.11b. The fundamental and first two harmonics show good agreement between the two plots. More variation is observed in the upper harmonics. The model accuracy is limited somewhat by computational efficiency. For the FE analysis, 4,000 steps per second were used to implicitly solve the problem, which is effectively the sample rate for the measured signal. As a result, the model is only capable of measuring harmonic content up to 2 kHz. Attempts to use a higher number of steps to solve the problem caused Abaqus to crash on a computer with a 3.00 GHz 6-core processor and 12 GB RAM.

While the implementation of FEM for the plain string is relatively straightforward, modeling of a wound composite string is challenging due to the problem of meshing the composite geometry, similar to the SFRC problem discussed in Chapter 3. The interaction between the core and wrap wire cannot be captured using a 1D model, and the length to diameter ratio of the string is so large that an excessive number of elements are required for 3D discretization. Modeling the effect of the bond between the core wire and wrap



(a)

(b)

Figure 4.11: Comparison of vibrational response measured from (a) experiment and (b) finite element model of a plain round ASTM A228 wire.

wire further increases computational cost. The composite string cross section cannot be treated as axisymmetric because of the angle of the helically wrapped wire, which rules out a 2D model of the cross section. Due to these computational modeling constraints, a data-driven machine learning approach is implemented in this work to determine a link between composite string descriptors and response using experimental data.

4.3 DBML Implementation

To implement the DBML approach, a database will first be constructed using recordings from several samples of new strings and used to train a machine learning model. When a new audio sample is provided from a randomly chosen string, the model will be used to deduce what type of material and geometry produced the signal.

4.3.1 Descriptor Selection and Sampling

In Chapter 3, a computational mechanics approach was taken to evaluate samples of SFRC microstructures, so the descriptors used as model inputs were sampled from a continuous range of numeric values. In this chapter, a physical experiment is used to measure the response of strings, so samples are picked from a discrete set of categorical descriptors

Table 4.2: Description of gauge 46 roundwound sample classes used for experiment.

Class	Core Material	Core Geometry	Wrap Material	Unit Mass, μ (g m^{-1})
A	ASTM A228	Hexagonal	NPS	7.21
B	ASTM A228	Round	NPS	7.24
C	High-Carbon Steel	Hexagonal	NPS	7.02
D	Maraging Steel	Hexagonal	Iron-Cobalt Alloy	7.11
E	ASTM A228	Round	Pure Nickel	7.50
F	ASTM A228	Hexagonal	Pure Nickel	7.98
G	ASTM A228	Hexagonal	Stainless Steel	7.09

based on what is available from manufacturers. As discussed previously, factors influencing the response of a composite guitar string include the core wire geometry and material properties, and the wrap wire geometry and material properties.

This study focuses on roundwound strings as they are most widely used. Table 4.2 lists seven classes of strings available from manufacturers that will be used for the study. These types were selected because they use a variety of core and wrap material and geometry combinations. Gauge 46 strings ($d_s = 0.046''$), tuned to E2 ($f_1 = 82.4$ Hz), were considered in this work. Other gauges and tunings are for future study.

4.3.2 Random Process Evaluation/Data Acquisition

As discussed earlier in this chapter, there are many factors aside from material and geometry that contribute to the signal generated from a vibrating string, including but not limited to the force with which it is plucked, its vibrating length, and the damping caused by air and the guitar body to which it is attached. Therefore in order to study the effect of geometry and material, it was necessary to design an experiment where all other variables are controlled. Experiments have been designed previously to test the response of a plucked ideal string [91, 117]. While different authors have taken different approaches to excite the strings and measure the response, the common theme is to isolate the response of the string from the supporting structure. This is achieved by using supports attached to a body with sufficient mass and stiffness so that the string is lightweight and flexible compared to the structure. Guitar string manufacturers use their own test methods to analyze frequency response of their strings for design. However, they do not publish the test methods used to obtain such curves, and therefore it is not possible to directly compare different types of strings from such information.

A guitar test rig with a Telecaster-style body was constructed in the lab to be used for all audio tests, shown in Figure 4.12. The scale length of the guitar is 25.5” and its overall mass is 3.21 kg. The body is solid basswood and the neck is constructed of maple with a rosewood fretboard. Two humbucking pickups were installed in the body, labeled as sensors “S1” and “S2”. S1 is located at a distance of 5.875” along the string ($\beta = 5.875$”/25.5” = 0.230) and S2 at 1.625” ($\beta = 0.064$). Rather than summing the two pickups to a single output using a switch as seen in a typical guitar, the pickups were wired separately to two output jacks so that the signal could be measured at each location independently.

The string is plucked at a distance halfway between the two pickups ($\beta = 0.147$) by a medium Fender 351 Standard Celluloid pick attached to the shaft of a NEMA 23 bipolar stepper motor. The mass of the pick is 0.48 g. The stepper motor is powered using an 18V DC supply and controlled using an Arduino Uno and TB6600 4A motor driver . A script was written to control the motor speed and position with respect to time. During the test, the output from each pickup was recorded using a PreSonus Audiobox USB ADC and saved as an uncompressed audio file with 24-bit resolution and 44.1 kHz sampling frequency. The test procedure is as follows:

1. Place new string on guitar.

2. Tune to the target fundamental frequency using a digital tuner ($f_1 = 82.4$ Hz).

3. Begin recording audio signals from ADC.

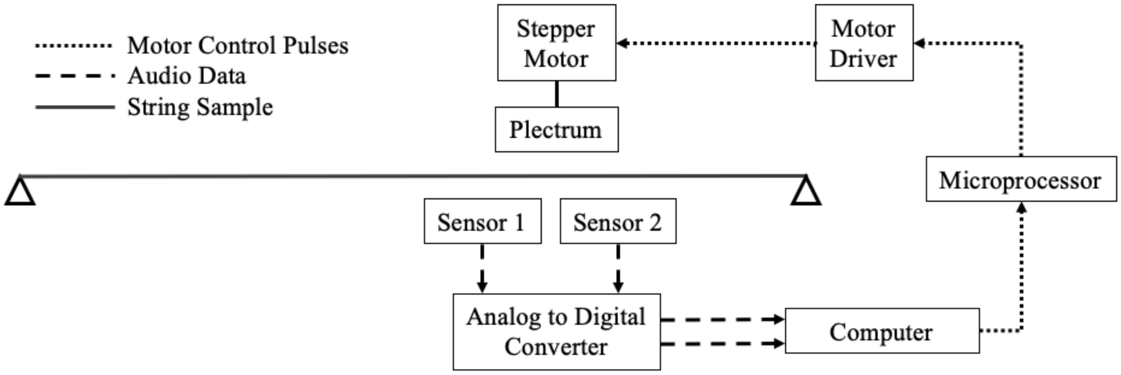
4. Pluck the string once forward (counterclockwise motion of plectrum), wait 30 sec-

onds.

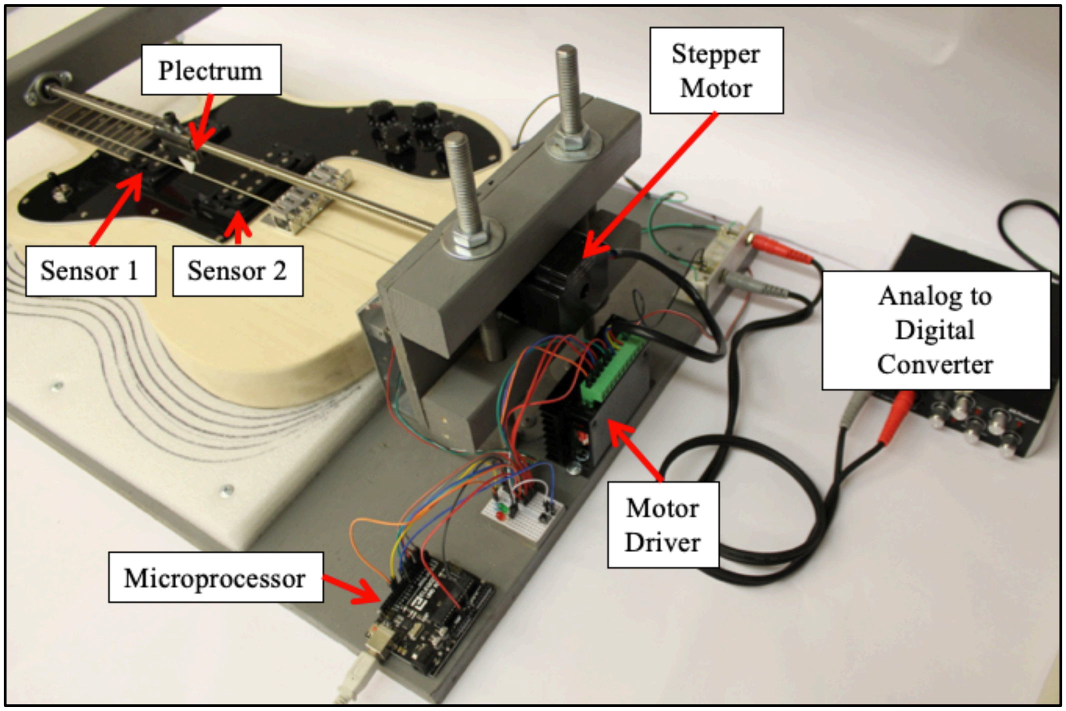
5. Pluck the string once backward (clockwise motion of plectrum), wait 30 seconds.
6. Pluck the string in alternating directions at a rate of 2 Hz until it has been plucked 200 times.
7. Repeat steps 4 to 5.
8. Stop recording audio.
9. Repeat steps 2 to 8 five times.
10. Remove string.

The purpose of the 30 second pause in steps 4 and 5 is to observe the entire decay envelope of the string. Preliminary tests indicated that 30 seconds is more than enough time for vibrations to be fully damped out. The purpose of repeating steps 2-8 is to allow the state of stress in the string to stabilize. It is common for new strings to fall out of tune easily due to material creep and slippage. It was found from preliminary tests that each string typically needed to be tuned 2 to 3 times before the fundamental frequency stabilized.

The test procedure was repeated for five samples of each of the seven string classes listed in Table 4.2. The spectrogram of each recording was obtained by taking the Short-Time Fourier Transform (STFT) of overlapping windows of 0.1 seconds, or 4410 samples in length, in order to observe the way the frequency content in the signal changes over time. Each window of the signal was multiplied by a Hamming window of equal length before taking the STFT in order to reduce spectral leakage in the finite length signal. The number of sample points for the discrete Fourier Transform was $10F_s$, giving a frequency domain resolution of 0.1 Hz. The spectrograms are shown in Figures 4.13 and 4.14 for the neck pickup (S1) and the bridge pickup (S2), respectively. One plot from each string class is shown. The spectrogram is plotted over the 0-1 kHz range so that the fundamental and the next ten partials can be observed. The first 15 seconds following the pluck are shown.



(a)



(b)

Figure 4.12: (a) Experimental system diagram and (b) laboratory setup.

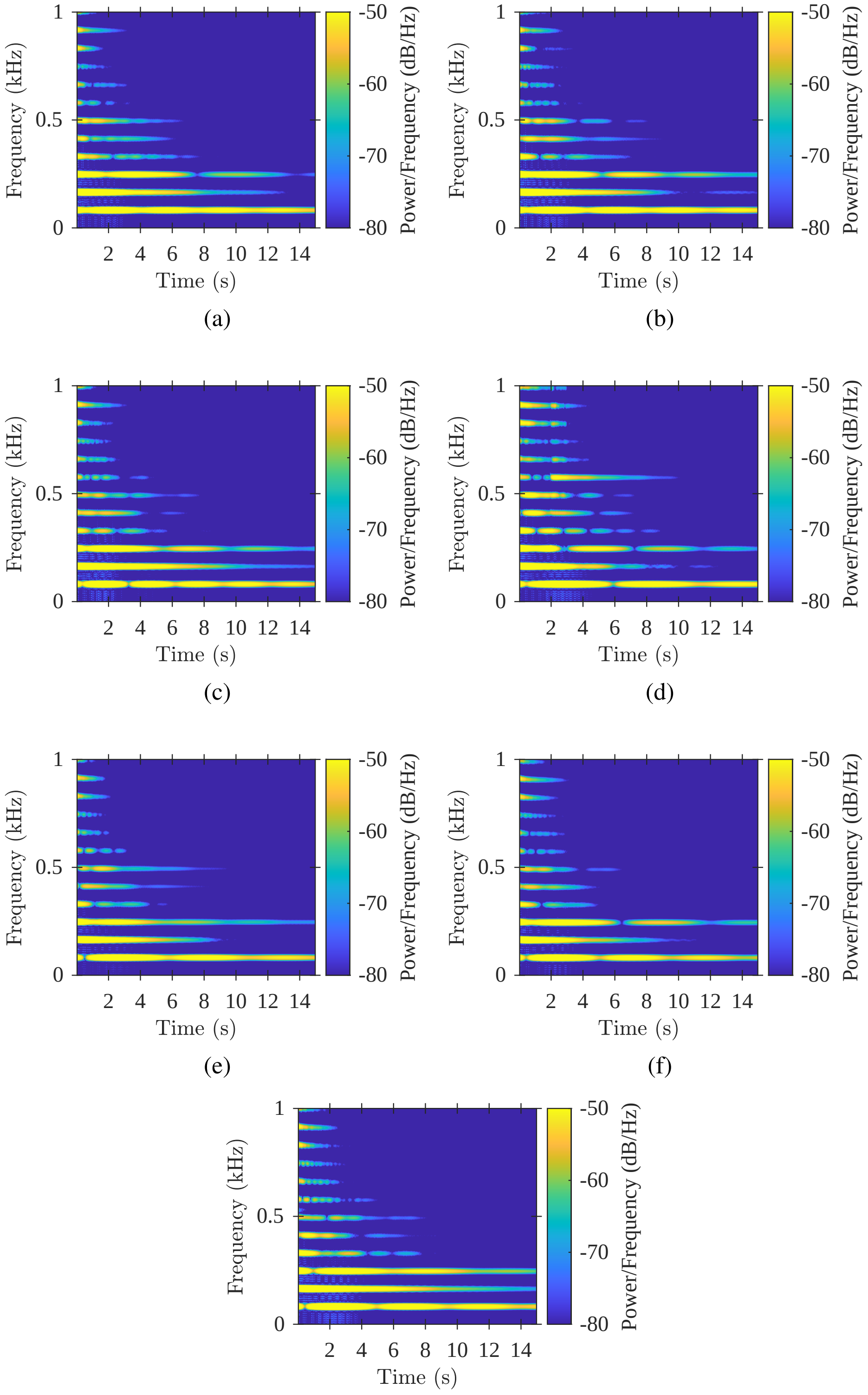
From visual observation, the partials lie at approximately integer multiples of the fundamental. The higher partials decay more quickly than the fundamental and lower partials as a result of the frequency-dependent damping behavior of the materials. The spectrogram was computed using the same parameters for each of the audio files recorded during the experiment.

Using the data from the spectrograms, the inharmonicity of each string class was quantified by measuring the difference in the observed frequency of the *n*th partial and the ideal harmonic frequency nf_1. The stiffness factor B for each string was calculated by fitting Equations 4.13 and 4.14 to the experimental data using the method of least squares. The resulting fitted curves for each class of string are shown in Figure 4.15 for the data obtained from S1 and in Figure 4.16 for the data obtained from S2. The blue trend lines correspond to the pinned case and the orange trend lines correspond to the clamped case. The calculated value of B for each class is indicated in the corresponding plot legend. Analysis of the spectrogram data indicated that the maximum amplitude of partials $n > 12$ were less than 10% of the maximum fundamental amplitude for each string sample, so the first 12 partials were used to calculate B. The data was averaged across the five strings tested from each class, and the error bars on the plots represent 2 standard deviations above and below the mean for each partial. For the sake of illustration, the vertical axis was transformed from frequency to inharmonicity using the following relationship:

$$\Delta f_n = 1200 \cdot \log_2 \left(\frac{f_n^*}{n f_1} \right) \tag{4.27}$$

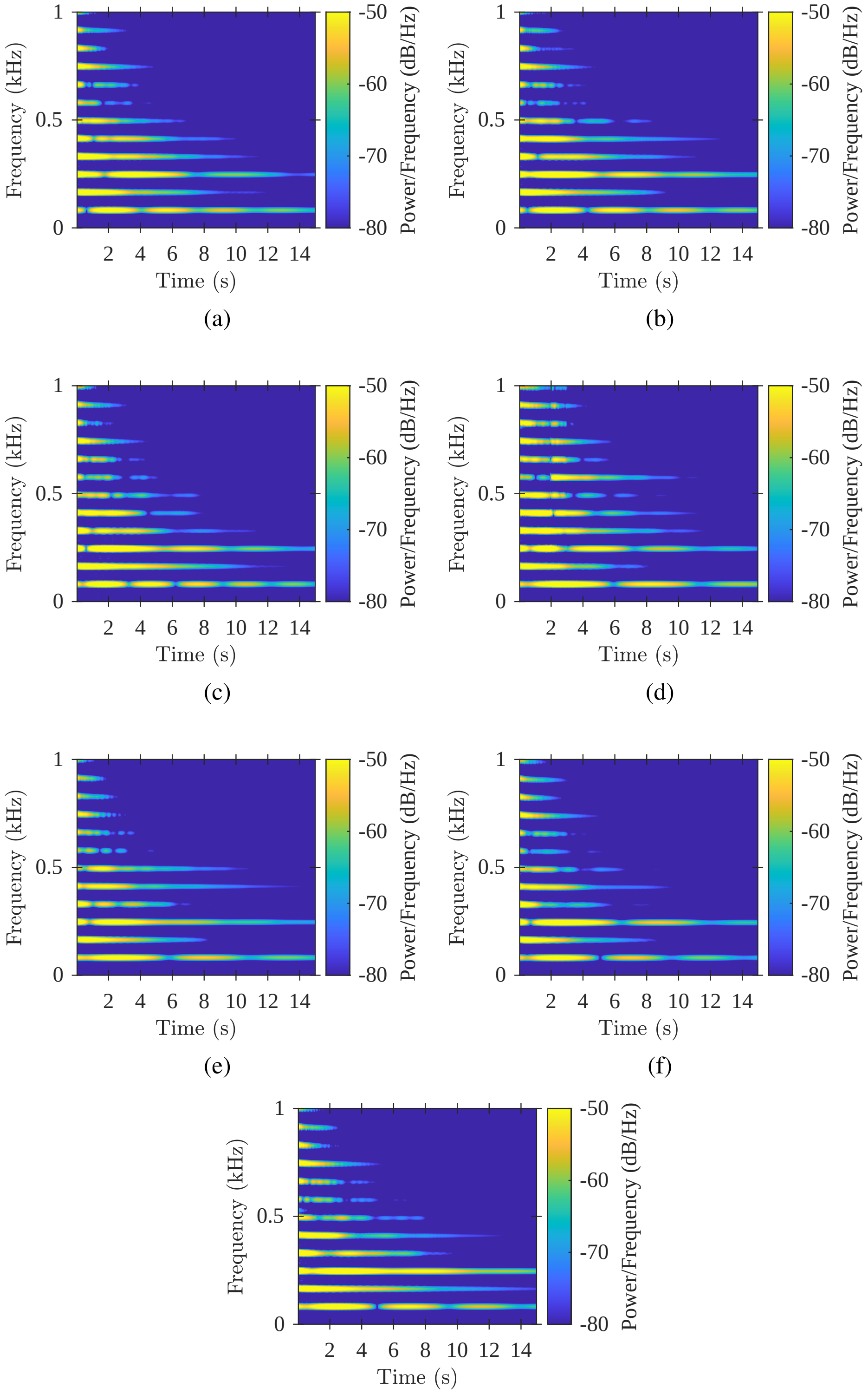
where f_n^* is the experimentally measured value of the frequency of the *n*th partial. Δf_n is the inharmonicity in cents. It should be noted that the equations for B are in terms of frequency in Hz and were fit prior to this nonlinear transformation, which is why the plotted curves do not appear to be an exact least squares fit of the data upon visual inspection. Since each strings is tuned to f_1, $\Delta f_1 = 0$ for all samples.

Figures 4.15-4.16 indicate that the data agrees well with Equation 4.13, corresponding



(a) (b) (c) (d) (e) (f) (g)

Figure 4.13: Spectrograms of a sample from each string class measured from S1. Plots (a)-(g) correspond to Classes A-G.



(a) (b) (c) (d) (e) (f) (g)

Figure 4.14: Spectrograms of a sample from each string class measured from S2. Plots (a)-(g) correspond to Classes A-G.

to the pinned end solution, while the form of Equation 4.14 tends to overestimate the inharmonicity of the lower partials and underestimate the inharmonicity of the higher partials. The results indicate that in reality the behavior of the supports on the guitar is somewhere between pinned and clamped. The second partial ($n = 2$) was at a slightly higher than expected frequency for each of the string classes tested, but the data otherwise follows a smooth increase in inharmonicity with respect to n. The only exception is Class G, where there are a few decreases observed in Δf_n with respect to n. The error bars provide insight into the variance in response observed across the five samples of each class. Classes D, E, and F showed the most consistent performance among samples, with standard deviations of only a few cents at each partial, while Classes A, B, and C showed slightly more variation. As discussed in Section 4.2, the location of the sensor along the string influences the measured inharmonicity. For each of the string classes, a larger value of B was calculated from the S2 data compared to S1, as seen in Figures 4.15-4.16.

Using the average calculated values of B for the pinned case and the measured unit masses reported in Table 4.2, Equations 4.11 and 4.15 were used to approximate EI for each of the string classes. The results are presented in Table 4.3. By ranking the classes in order of their stiffness, a few trends become apparent. First, the stiffness of the string is highly correlated to the wrap material used. Class G was the stiffest string tested, and was the only class with stainless steel wrap wire. The next highest values of EI were observed for Classes A, B, and C, which each utilize NPS wrap wire, followed by Classes F and E, which utilize pure nickel wrap wire. Class D was the most flexible string, with its unique iron-cobalt alloy wrap and maraging steel core. The data also indicates that hexagonal core wire provides an increase in stiffness compared to round core wire. Classes A and B use identical core and wrap material, but the hexagonal core of Class A results in a slightly higher value of EI than the round core of Class B. The same can be said of Classes E and F. Each has a music wire core and a pure nickel wrap, but the hexagonal core of Class F results in a stiffer string than the round core of Class E. The percent difference in stiffness

Table 4.3: Ranking of string classes by stiffness.

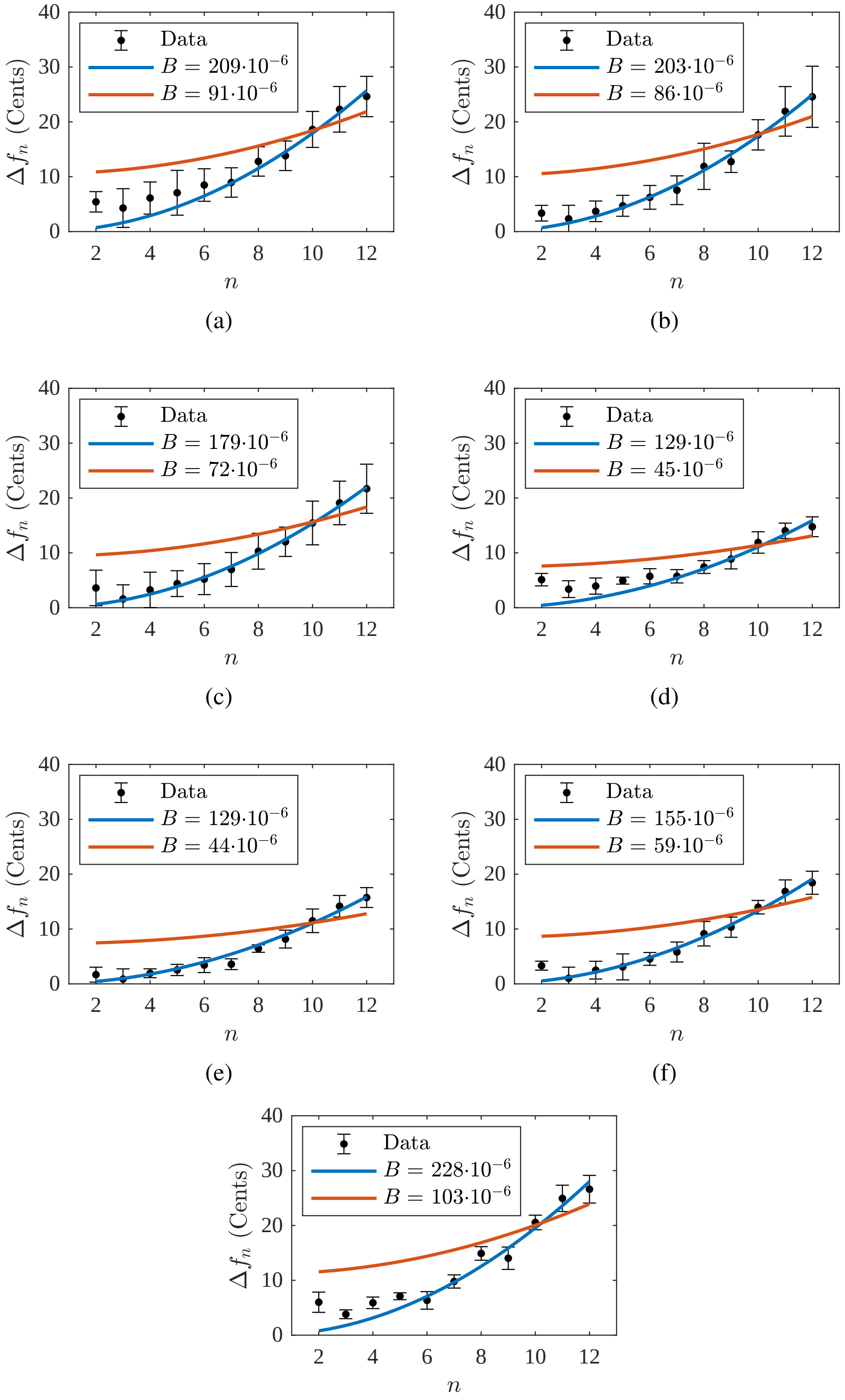
Rank	Class	Core Material	Core Geometry	Wrap Material	S1 EI $(10^{-3}\ \text{lb}\cdot\text{in}^2)$	S2 EI $(10^{-3}\ \text{lb}\cdot\text{in}^2)$	Difference (%)
1	G	ASTM	Hex	SS	272.6	285.8	4.8
2	A	ASTM	Hex	NPS	254.2	270.0	6.2
3	B	ASTM	Round	NPS	247.9	267.4	7.9
4	C	HCS	Hex	NPS	211.9	234.4	10.6
5	F	ASTM	Hex	Ni	208.6	230.2	10.3
6	E	ASTM	Round	Ni	163.2	173.3	6.2
7	D	MS	Hex	FeCo	154.7	171.5	10.9

measured between S1 and S2 is reported in the final column of the table.

After identifying the inharmonic frequencies of each string, the decay of each partial with respect to time was analyzed. Let $|S(f_n^*,t)|$ denote magnitude of the spectrogram bin corresponding to measured inharmonic frequency f_n^* calculated over the 0.1 second STFT window centered at t. Figures 4.17 and 4.18 show the decay of the first five partials over 10 seconds following a pluck using data measured from S1 and S2, respectively. Some low frequency oscillation can be seen in the data. This phenomenon has been observed in previous studies on string vibration [68, 91, 89], and can be attributed to coupling between the string and guitar body, which is most significant for small values of n. The following exponential decay model was used to parameterize the data for the nth partial for the first five seconds after the pluck:

$$|S(f_n^*,t)| = A_n \exp(\alpha_n t) \tag{4.28}$$

where parameter A_n is equal to the initial amplitude of the nth partial and α_n represents the time constant for the exponential decay curve. A larger value of α_n indicates that the frequency is damped more quickly, whereas smaller values correspond to longer decay times. A_n and α_n were calculated for each partial of each string class tested by fitting the data to Equation 4.28. The resulting mean values for A_n and α_n are presented in Appendix Tables



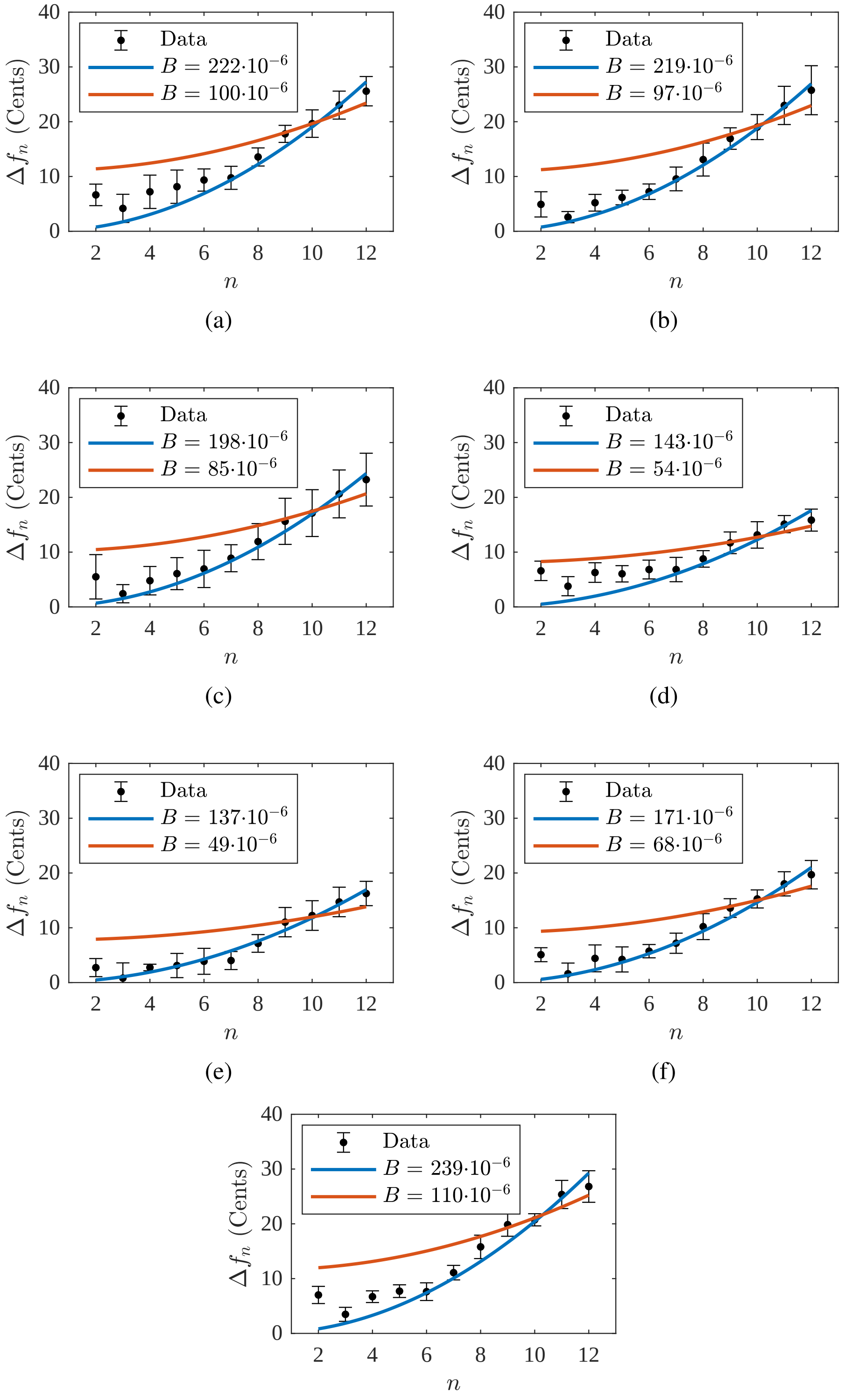
(a) (b)

(c) (d)

(e) (f)

(g)

Figure 4.15: Average measured inharmonicity from S1. Plots (a)-(g) correspond to Classes A-G.



(a) (b) (c) (d) (e) (f) (g)

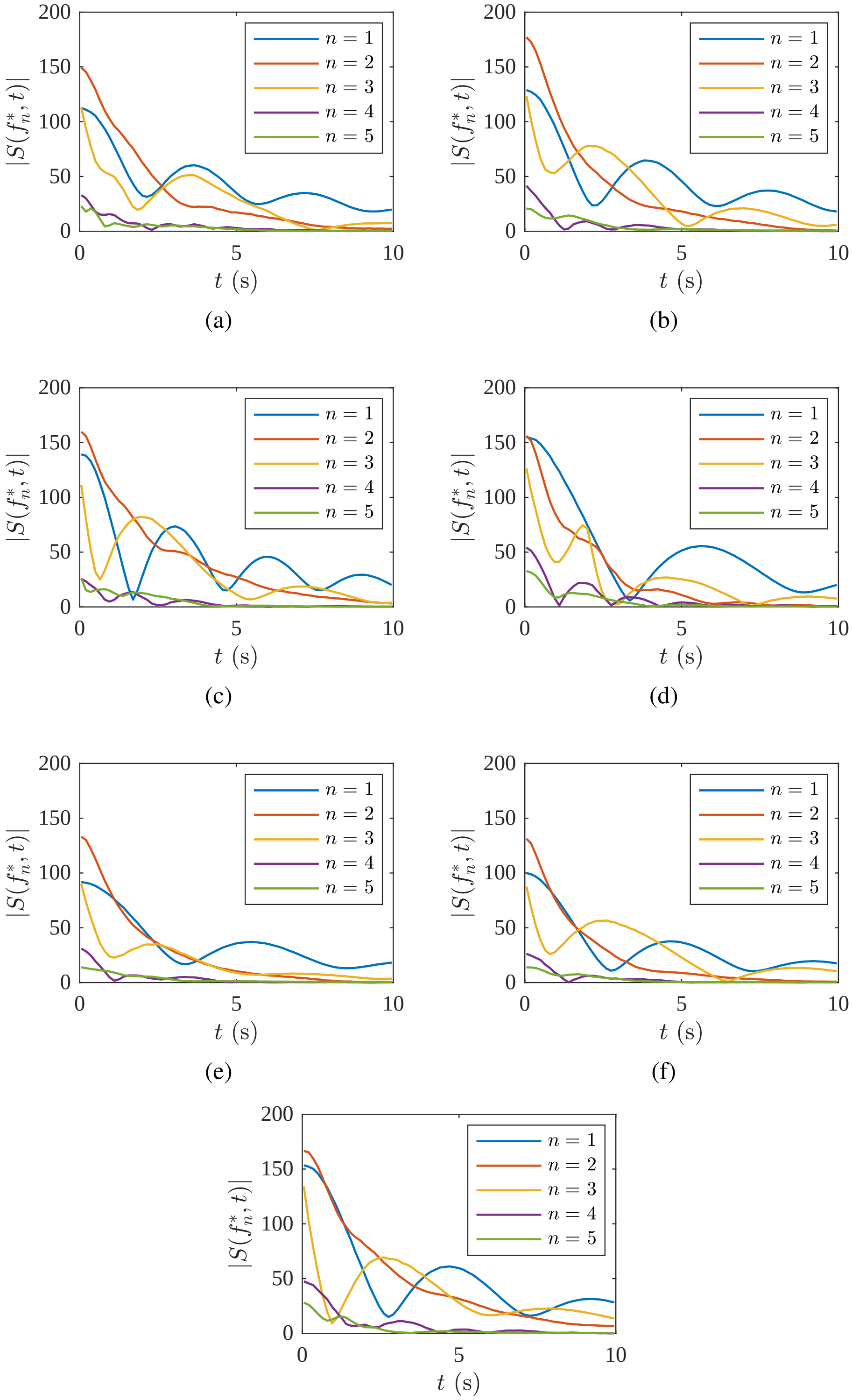
Figure 4.16: Average measured inharmonicity from S2. Plots (a)-(g) correspond to Classes A-G.

B.1 and B.2, respectively. The coefficients of variation for each parameter are presented in Appendix Tables B.3-B.4.

Typically, pickups located closer to the bridge, such as S2, measure more high frequency content than a pickup located near the neck. This trend is observed in the results of Tables B.1-B.2. The values of A_n indicate that pickup S1 measured significantly more of partials 1-3 than S2. The average peak amplitude of the second partial A_2 measured by S1 is largest for each of the string classes. The subsequent partials tend to decrease in magnitude as n increases. The data from S2 indicates that partials 3-5 are prominent for each of the string classes. Class G shows the largest average amplitude for most of the partials presented in the table. This seems to indicate that the stainless steel wrap wire of Class G exhibits a stronger magnetic attraction to the pickups than the other wrap materials tested, resulting in an overall louder signal. Incidentally, Class G also exhibited the most inharmonicity of any class. However, the values of A_n computed for the rest of the classes do not directly correlate to their stiffness ranking. The damping parameters α_n tend to increase with n, indicating that the higher frequencies are damped out more quickly, but the tabulated values do not follow a purely monotonic trend. This phenomenon can be observed in the spectrograms depicted in Figures 4.13-4.14.

4.3.3 The Learning Machine

To determine whether a clear distinction between the string classes can be established from the data, a support vector machine was trained to classify the type of string that produced a randomly selected audio signal. In the SVM approach, a signal produced by a sample from some class is represented by a point in feature space. The features that constitute the dimensions of feature space are parameters that can be measured from a signal, such as B, A_n, and α_n. An algorithmic approach is used to identify hyperplanes in feature space that optimally separate data from different classes. Nonlinear boundaries can be determined using a kernel method. The foundation of the SVM dates back to the work



(a)

(b)

(c)

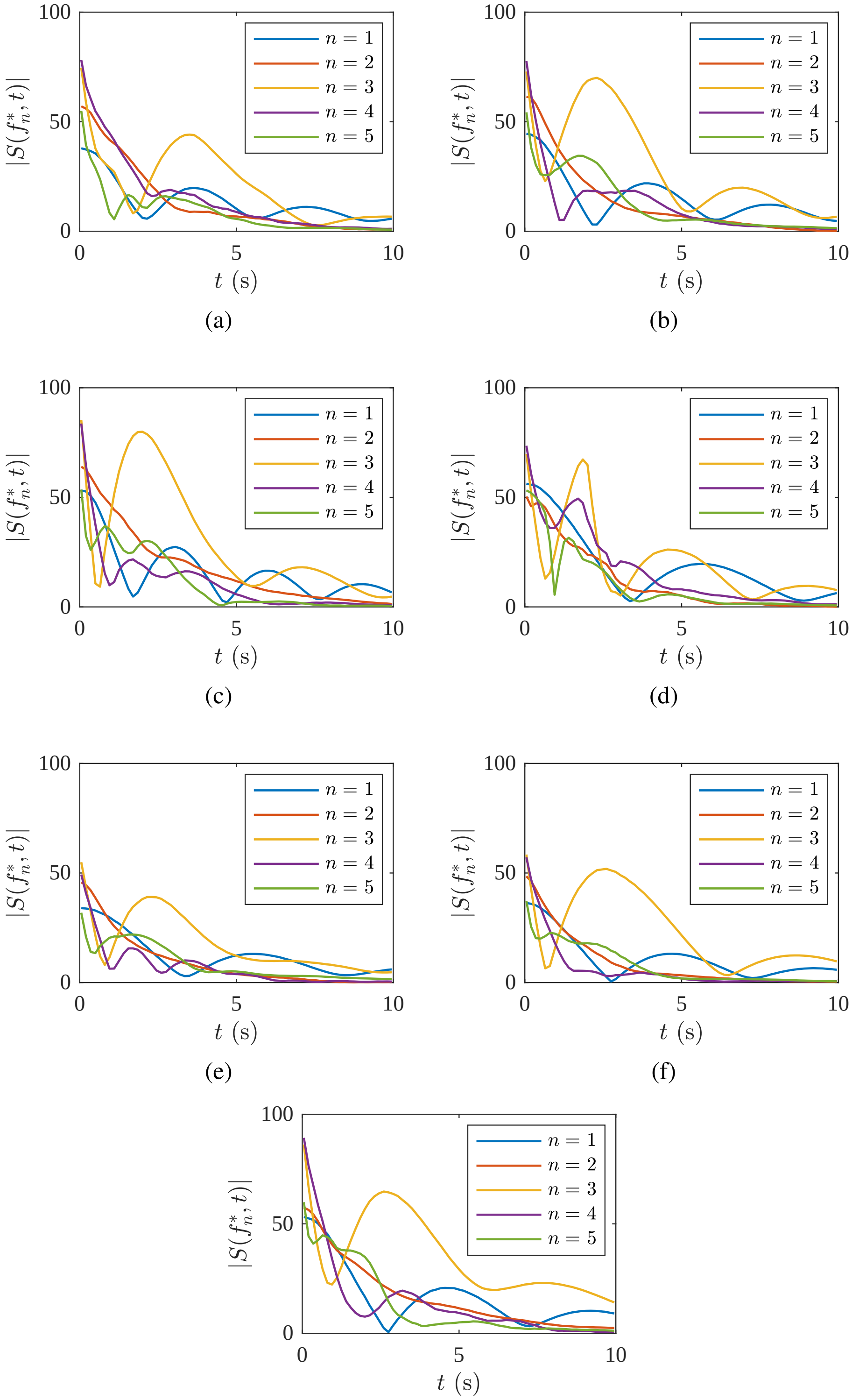
(d)

(e)

(f)

(g)

Figure 4.17: Time-dependent amplitude of first five partials measured from S1 for each string class. Plots (a)-(g) correspond to Classes A-G.



(a)

(b)

(c)

(d)

(e)

(f)

(g)

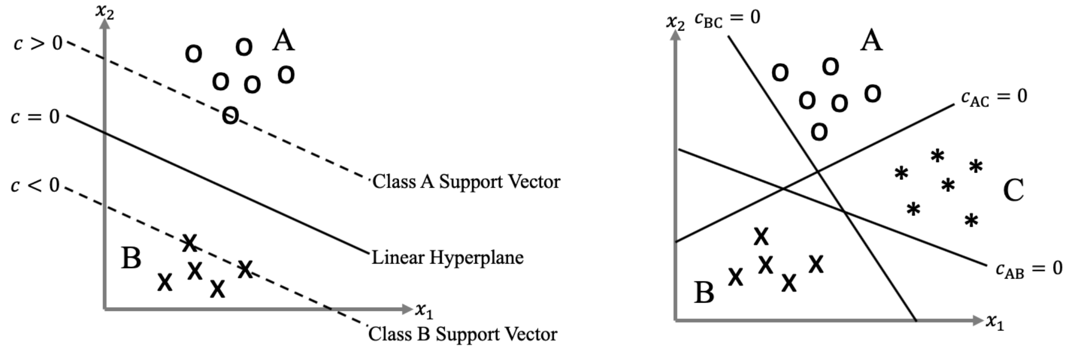
Figure 4.18: Time-dependent amplitude of first five partials measured from S2 for each string class. Plots (a)-(g) correspond to Classes A-G.

of Vapnik [118], who in 1963 introduced a linear classifier for identifying images . The modern definition of the SVM was formalized in his later collaborations with Boser [119] and Cortes [120]. The SVM has since been implemented in speech recognition algorithms [121], an audio classification problem similar to the guitar string problem considered here. The theory of SVMs and the algorithms used to train them have been discussed in detail in the literature, but a brief overview is provided here. A more exhaustive explanation is available in Vert [122].

Let the dataset used to train the SVM be denoted $\{\mathbf{x}_i, y_i\}$, $i = 1...N$, where $\mathbf{x}_i \in \Re^d$, with d being the number of dimensions of feature space and N being the number of data points. By this definition, the signal obtained from testing string sample i can be represented as a point in the feature space denoted by $\mathbf{x}_i = [x_1, x_2, ..., x_d]$, where each component of $\mathbf{x}_i$ corresponds to a random variable measured from the signal. The output y_i represents the class of sample i and can take a value of 1 or -1, with each value corresponding to one of two classes. Let a classifying hyperplane $c(\mathbf{x})$ be defined by the following linear equation:

$$c(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0 \tag{4.29}$$

where $\mathbf{w} \in \Re^d$ is a vector of weights and b is a constant. The model is trained to minimize $||\mathbf{w}||$ subject to the constraint $y_i c(\mathbf{x}_i) \geq 1 \, \forall \, i$. The boundary points in each class, where $y_i c(\mathbf{x}_i) = 1$, are the support vectors. Once trained, the model outputs $c(\mathbf{x}_i) > 0$ for points where $y_i = 1$ and $c(\mathbf{x}_i) < 0$ for points where $y_i = -1$. Figure 4.19a illustrates the SVM approach to classification for two classes of data, A and B, in a 2-dimensional feature space with dimensions x_1 and x_2. SVMs are inherently binary classifiers by the definition $y_i \in \{+1, -1\}$ but may be extended to account for more classes using $m * (m-1)/2$ hyperplanes, where m is the number of classes, as illustrated in Figure 4.19b for the case of three classes. In this example, one hyperplane is used to distinguish between Class A and Class B, another between B and C, and a third between A and C. This is called a one-



(a) (b)

Figure 4.19: Support vector machines for (a) binary classification and (b) multiple classes.

vs-one approach. An alternate method is the one-vs-all approach, where a hyperplane is identified to separate one class from all other classes, resulting in m hyperplanes. As n increases, the number of hyperplanes needed using the one-vs-one approach grows much larger, but the size of the linear system for each hyperplane stays reasonably small because data is being compared from only 2 classes. Conversely, the one-vs-all approach requires fewer hyperplanes, but each hyperplane is determined using all of the data, resulting in a large linear system with a trade-off in computational efficiency. For the size of the dataset presented here, using a one-vs-all approach does not result in a significant improvement in computational speed. Training the SVM takes a matter of seconds.

Algorithms for optimizing the weights of the SVM require computation of the dot product $\mathbf{x}_i \cdot \mathbf{x}_j$ for all $i, j \in \{1, 2, ..., N\}$. If the points in feature space are not separable by a linear hyperplane, a nonlinear boundary may be determined by replacing the dot product with a kernel function, $k(\mathbf{x}_i, \mathbf{x}_j)$. A nonlinear kernel is particularly useful for separating data in high-dimensional spaces. Many kernel functions can be used within the SVM framework. The training process generally involves testing the performance of different kernel functions to identify the kernel that most accurately separates the classes. In this work, three kernels were initially considered: a linear kernel (Equation 4.30), a quadratic kernel (Equa-

Table 4.4: SVM training data organization.

i	B (x_1)	Δf_2 (x_2)	$\cdots$	Δf_n	A_1	$\cdots$	A_n	α_1	$\cdots$	α_n (x_d)	Class (y_i)
1	$219 \cdot 10^{-6}$	5.26		21.12	98.23		26.38	0.22		1.76	A
2	$204 \cdot 10^{-6}$	4.20		20.13	111.73		23.49	0.21		1.15	A
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
N	$254 \cdot 10^{-6}$	8.39		27.51	177.34		38.83	0.28		1.55	G

tion 4.31), and a Gaussian kernel (Equation 4.32).

$$k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j \tag{4.30}$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i \cdot \mathbf{x}_j)^2 \tag{4.31}$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \tag{4.32}$$

In Equation 4.32, σ is a scale parameter that is optimized during the training process. Preliminary results using each of the three functions for training indicated that the quadratic kernel provides the most accurate predictions for the experimental data, so it was selected for use.

Table 4.4 shows the manner in which the experimental data is organized for training the SVM. Each row of the table corresponds to a sample i and each column of the table is a feature, with the exception of the last column which indicates the known class of the sample and serves as the output to be predicted. The damping and inharmonicity parameters presented earlier in this section are used as features for the SVM. The stiffness parameter B from the pinned case (Equation 4.13) was used since it provided the most accurate fit of the data. Five recorded plucks were used from each of the five specimen tested from the seven string classes, resulting in $N = 175$ samples. Figure 4.20 shows the data obtained from the seven considered classes plotted in two arbitrary dimensions of the feature space.

Four classification problems were attempted. In the first problem, the SVM was trained

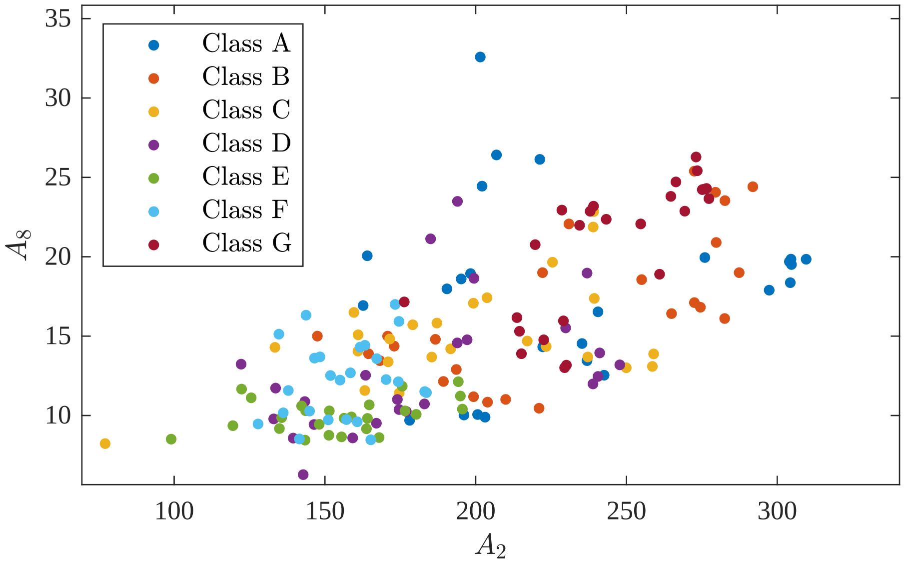


Figure 4.20: Experimental data for seven classes plotted in two arbitrary dimensions of feature space, A_2 and A_8.

to identify which of the seven different sample classes produced a signal of interest, so each sample was associated with a class $y_i \in \{\text{A, B, C, D, E, F, G}\}$. In the second problem, the algorithm was trained to identify whether a signal was produced by a string with a round or hexagonal core wire, with classes $y_i \in \{\text{Hex, Round}\}$. The third problem was predicting the core material, with $y_i \in \{\text{ASTM A228, High-Carbon Steel, Maraging Steel}\}$. The fourth problem was concerned with identifying the wrap materials, with $y_i \in \{\text{NPS, Iron-Cobalt Alloy, Pure Nickel, Stainless Steel}\}$.

For each problem, the SVM was trained three times using different combinations of feature vectors to determine which parameters are most useful for differentiation among the classes. The SVM was first trained using only inharmonic features, such that $\mathbf{x}_i^{(1)} = \{B, \Delta f_2, \ \dots \ , \Delta f_n\}_i$, and the number of dimensions of feature space is $d = n$. Δf_1 is omitted because there is no inharmonicity at the fundamental frequency. The SVM was then trained using only damping features, with $\mathbf{x}_i^{(2)} = \{A_1, \ \dots \ , A_n, \alpha_1, \ \dots \ , \alpha_n\}_i$, where $d = 2n$. Finally, the SVM was trained using all of the parameters for each sample i, so that $\mathbf{x}_i^{(3)} = \mathbf{x}_i^{(1)} \cup \mathbf{x}_i^{(2)} = \{B, \Delta f_2, \ \dots \ , \Delta f_n, A_1, \ \dots \ , A_n, \alpha_1, \ \dots \ , \alpha_n\}_i$, and $d = 3n$.

The performance of the SVM classifier is assessed by partitioning the available data into a training set and a test set. The training set is used to determine the weights for the SVM. The trained SVM is then used to predict the class of each of the points in the test set. Comparing the predicted class to the known class of each test point indicates whether or not the SVM is accurate. The performance of the trained SVM is dependent on which training points are selected. To address this dependency, 5-fold cross validation was used. The 175 data points in descriptor space were divided into five randomly selected disjoint subsets of 25 points each. In the first iteration four of the sets were used as training points, with the remaining set used as test points. The process was repeated four more times, using each of the subsets as test points in successive iterations. The accuracy of the SVM for each problem was calculated as the average over all five iterations.

The prediction accuracy of the SVM for each of the problems and feature variable sets is summarized in Table 4.5. The learning machine exhibited similar accuracy using data from each of the two sensors. The machine performed better using the partial amplitude and time constant features, $\mathbf{x}^{(2)}$, than using the inharmonic features, $\mathbf{x}^{(1)}$, for each of the problems considered. However, the best performance was obtained by using all features, $\mathbf{x}^{(3)}$, for training. The accuracy of the SVM was greater than 93% for each of the classification problems using $\mathbf{x}^{(3)}$. The SVM was most accurate at differentiating between strings with different wrap materials (>97%). The SVM demonstrated similar performance between the core geometry and core material identification problems (94-95%). The seven-class problem of identifying the particular string that produced a signal was most challenging, but the classifier still performed with 93.7% accuracy for each sensor.

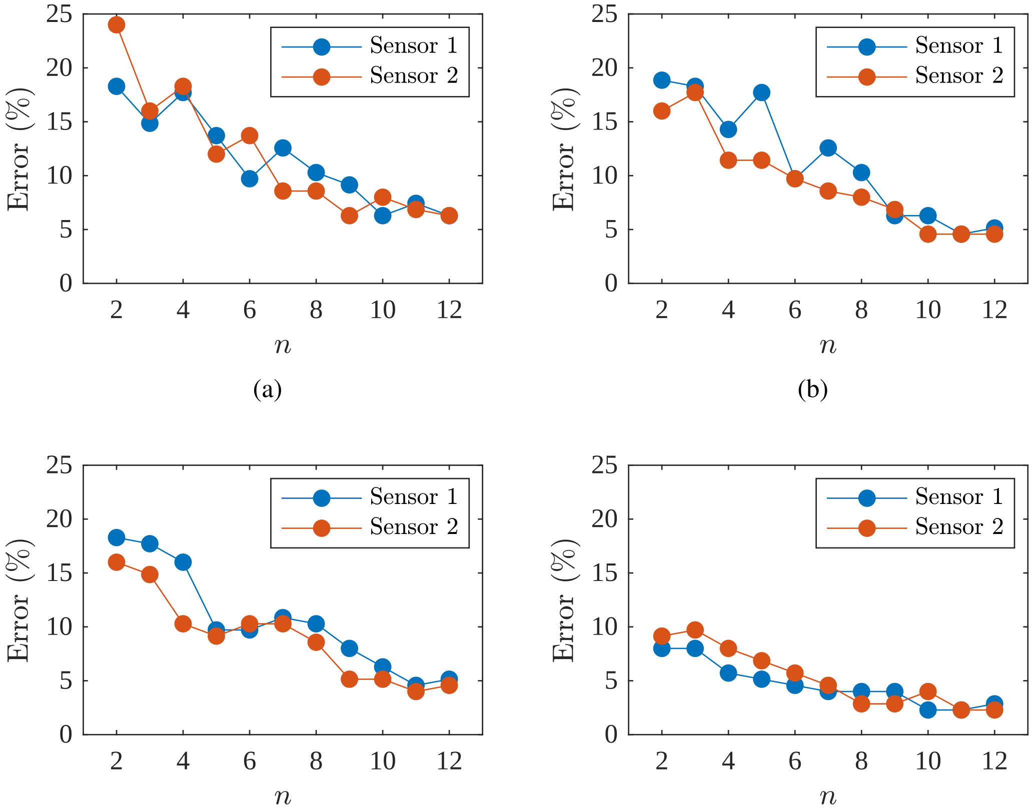
The results of Table 4.5 were obtained using $n = 12$. As previously mentioned, the amplitude of the partials corresponding to $n \geq 13$ were less than 10% of the fundamental amplitude, and in many cases were indistinguishable from the noise floor of the signal. To show the effect of the number of partials considered on the accuracy of the classifier, the SVM was trained using the entire feature set, $\mathbf{x}^{(3)}$, with values of n ranging from 2 to

Table 4.5: Accuracy (%) of SVM on four classification problems using different feature sets for training, $n = 12$.

Sensor	Problem	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$
1	String Type	90.9	93.1	93.7
	Core Geometry	90.3	92.0	94.9
	Core Material	89.1	90.3	94.9
	Wrap Material	93.7	96.0	97.1
2	String Type	90.3	90.3	93.7
	Core Geometry	90.3	91.4	95.4
	Core Material	91.4	93.1	95.4
	Wrap Material	93.7	94.9	97.7

12. The results for each of the four classification problems are shown in Figure 4.21. The plots indicate that the performance of the SVM improves as more partials are considered for each of the problems. Figures 4.21a-4.21c exhibit similar trends, with prediction error >15% when only a few partials are considered. With $n \geq 10$, the error is reduced to about 5%. Figure 4.21d indicates that the SVM exhibits >90% accuracy in identifying the wrap material of a string even when only two partials are considered.

Figure 4.22 shows the confusion matrices obtained for each of the four test problems using $\mathbf{x}^{(3)}$ to train the quadratic SVM with $n = 12$. Plots a, c, e, and g correspond to S1 and plots b, d, f, and h correspond to S2. In each plots, the rows correspond to the true class of a test point and the columns correspond to the predicted class of a test point. The green shaded boxes on the diagonal show the number of correct predictions for each class while the red shaded boxes show the incorrectly predicted classes. For example, Figure 4.22a indicates that of the 25 samples from Class A, the SVM correctly identified 23 samples, while 2 samples were misidentified as belonging to Class C.



(a) (b)

(c) (d)

Figure 4.21: SVM prediction error using all features, $\mathbf{x}_i^{(3)}$, with respect to number of partials used for four problems, (a) String Class, (b) Core Geometry, (c) Core Material, (d) Wrap Material.

(a)

	A	B	C	D	E	F	G
A	23		2				
B		25					
C	2		22		1		
D				25			
E					23	2	
F						25	
G		2	1				22

(b)

	A	B	C	D	E	F	G
A	25						
B		23	2				
C		2	19		1	3	
D				25			
E					24	1	
F						25	
G		2	2				21

(c)

	Hex	Round
Hex	121	4
Round	5	45

(d)

	Hex	Round
Hex	119	6
Round	7	43

(e)

	ASTM	HCS	MS
ASTM	123	2	
HCS	7	18	
MS			25

(f)

	ASTM	HCS	MS
ASTM	124	1	
HCS	6	19	
MS			25

(g)

	FeCo	NPS	Ni	SS
FeCo	24		1	
NPS		74	1	
Ni			50	
SS		2		23

(h)

	FeCo	NPS	Ni	SS
FeCo	25			
NPS		74	1	
Ni			50	
SS		3		22

Figure 4.22: SVM confusion matrices using all features for four problems: (a)-(b) String Class, (c)-(d) Core Geometry, (e)-(f) Core Material, (g)-(h) Wrap Material. The left column corresponds to S1 data and the right column corresponds to S2 data.

4.4 Conclusion

A study was performed to analyze the effects of composite guitar string geometry and material properties on the vibrational signal produced by a plucked string. An experiment was designed to isolate the response of a string within the context of a typical electric guitar system. The guitar body material and geometry, sensor type and location, and plectrum material and force were used as control variables. Tests were performed on a variety of string specimen and response was measured via electromagnetic sensors at two locations along the vibrating length of the string. The measured inharmonicity, stiffness, and frequency-dependent damping properties of each string class were reported. A support vector machine was trained to learn the vibrational characteristics of a string based on its morphological descriptors. The following conclusions can be drawn from this study:

1. The stiffness of the wound string is highly correlated to its wrap material. The strings with a stainless steel wrap were the stiffest tested, while the iron-cobalt alloy wrap resulted in the most flexible string.
2. Strings utilizing a hexagonal core wire were stiffer than strings made of the same materials with round core wires, resulting in more inharmonicity.
3. Oscillations in the time-dependent response of each partial indicate a coupling effect between the string and the instrument body. Characterization of this coupling will require future study.
4. The power and frequency of partials in the measured signals are strongly influenced by the string's composite morphology. A support vector machine is capable of accurately differentiating between signals produced by different types of strings (>90% success rate), with a quadratic kernel producing the best performance.
5. The classifier performs most successfully when using features that describe the time-dependent response of partials in addition to inharmonicity. The accuracy of the algorithm improves as the number of partials used increases.

While previous work has shown that modeling the vibration of a homogeneous wire is straightforward, no analytical method exists to account for wound strings where slippage between core and wrap wire is present. The results of this study are particularly useful for a data-driven analysis of this phenomenon. Potential avenues for future work are discussed in the following chapter.

Chapter 5

Summary, Conclusions, and Future Work

In this dissertation a generalized framework for predicting the properties of composites with stochastic geometries using machine learning concepts was proposed and verified using a combination of numerical models and experimental data. The approach was applied to solve the problem of a 2-phase random short-fiber composite subject to nonlinear damage under static tension, as presented in Chapter 3. Suitable computational models for the fine and coarse scale response of the material were identified. A recently developed XFEM model was used to account for the effect of randomly located fibers with cohesive bonds in the matrix. A Gaussian process was used to map from an input space defined by the fiber size, volume fraction, and bond properties to a characteristic composite response curve. In Chapter 4, the approach was extended to the case of vibrating composite electric guitar strings with a variety of phase geometries. An experiment was designed to capture the response of various strings in a controlled environment. The results constitute the most extensive study to date on the effect of string material properties and geometry on the quality of sound produced when plucked. It was shown that using the data a trained support vector machine classifier can identify the string configuration that produces an arbitrary signal. In each of the two test cases drawn from distinctly different fields, a link was established between material morphology on the small scale and the performance of the global system.

The purpose of this dissertation was to demonstrate how a modern data-driven approach can be used to solve nonlinear problems in high-dimensional spaces that cannot be treated as deterministic. In the case of SFRC a numerical microscale model was tailored specifically for the treatment of random in-plane fibers for a parametric study. The XFEM model that was used allowed for the treatment of fibers in close proximity using a uniform mesh, making the method more computationally efficient than traditional discretization methods. The GP surrogate model allowed for prediction of properties using only microstructural descriptors. In the context of a multiscale modeling approach, this will further expedite the analysis process since effective properties at random material points can be inferred from the trained GP without resorting to explicit simulation. There is ample room for future extension of the methods discussed in this work. In the case of SFRC, the following issues may merit further consideration:

- The XFEM model used in this work does not account for fiber fracture, which dictates the true failure strain and stress of the composite material. In this work a 90% reduction in strength was defined as failure, but modeling of fiber fracture would provide a more accurate picture of the point at which the composite fails.

- The physical properties of fiber and matrix were treated as deterministic variables, while the fiber volume fraction, size, and bond properties were treated as probabilistic. Future work treating the matrix and fiber materials as random may lead to deeper understanding of the effect of such uncertainties.

- The concept of statistical descriptors for fiber distribution in the matrix was introduced in this work but was not implemented in the machine learning process. In addition to using volume fraction as a descriptor, future work may consider parameters of an n-point correlation function, lineal path function, misorientation function, or some other function to better characterize the spatial distribution of fibers. While the 6-descriptor model used here agreed well with the test cases, it would be interesting to see if an additional higher-order descriptor improves the performance of the learning machine.

- The work presented here focuses on characterizing random SFRC microstructures, modeling the response, and predicting the properties using machine learning. The CDP model was presented as an effective approach to macroscale modeling of SFRC.

Within a sequential multiscale modeling framework, the results obtained from XFEM and GP should be passed to the CDP model as stochastic material parameters. The final step to be addressed is implementation of SFEM on a macroscale structure of interest using the predictions of the learning machine.

- Scalability of the machine learning model must be addressed as more data becomes available. A Gaussian process was used in this work, but the GP represents one of many options for regression using machine learning. While it performed well on this dataset, its efficiency requires evaluation as more data becomes available.

For the electric guitar string problem an experiment was custom-designed to measure the response of a string in a realistic operational environment. While previous work has focused on modeling the response of an electric guitar string, to the knowledge of this author there is no published work comparing the response of many classes of strings in a clearly defined, repeatable manner. The ability to correlate composite string descriptors such as core geometry and wrap material to the sound that they produce when used on a guitar is of interest to manufacturers seeking to improve their products, as well as musicians that wish to experiment with alternate timbres. The results of this study can be used to make informed decisions on material and topology selection for the manufacture of new-age strings. The following topics have been identified for future study:

- This study considered a single gauge of roundwound strings. Additional gauges should be analyzed in future work. While less common, some guitarists prefer flatwound strings, which are constructed by polishing the outer layer of a roundwound string to create a smoother surface profile. It is also of interest to study flatwound strings, and the effect of different polishing techniques on their performance.

- By building upon the results presented here, it is possible to identify optimal string configurations. The optimization problem is dependent upon mechanical performance as well as psychoacoustic response. The optimal string should not only have a maximal fatigue life, but also must be acoustically pleasing. A psychological study in which several types of strings are subjectively ranked by guitarists and listeners could provide insight into what signal metrics are most desirable.

- Future studies should assess the manner in which string response changes as damage progresses. As discussed in Chapter 4, there are three common damage modes for the instrument string. Each of the modes results in local variation of material properties that influence the response, such as a reduction in cross-sectional area, variation in density, or a loss of stiffness. As it is challenging to artificially induce realistic damage in strings, an SFEM approach is envisioned for modeling this phenomenon computationally. Using SFEM, “damaged” elements could be randomly placed throughout the length of the string and their influence on its response measured.

- No method currently exists for damage identification and prognosis in instrument strings, so the replacement of used strings often occurs on an arbitrary timetable. The long term goal of this research is to develop a tool capable of accurately predicting impending failure of a string by monitoring the manner in which the signal changes over time. The envisioned approach would utilize the concepts of acoustic emissions based structural health monitoring. In general, acoustic emissions techniques in structural applications associate acoustic events with sources of damage in the structure. Detection of a peak in the signal at some time or frequency may indicate a crack in the material. For the musical string application, the problem is not the explicit identification of cracks associated with signal peaks, but rather characterizing the modification of signal characteristics as damage occurs.

- To help realize the points listed above, it is of interest to eventually acquire data from various guitarists who use different guitars and pickups and pluck strings at different intensities. These random variables could be treated as additional descriptors in the context of the machine learning approach. The envisioned network of sensors would feed into a database, allowing the learning machine to update as new data becomes available. Guitarists commonly use a digital tuner to analyze the fundamental frequency of a string before playing. These tuners do not make use of higher frequency content in the signal, and they operate in real time, with no memory. The development of a new type of tuner capable of tracking signal metrics and storing the information to memory or transmitting it over a network would allow for an unprecedented in-depth study of the behavior of guitar strings.

Appendix A

SFRC Descriptor Space Sample Points

Table A.1: SFRC sample points, $\mathbf{x}_i, i = 1...20$.

i	d_f (mm)	l_f (mm)	ϕ_f (%)	T_n (N)	δ_n (10^{-5} mm)	δ_{max} (10^{-4} mm)
1	0.079	7.167	0.97	9.40	7.19	2.44
2	0.043	7.667	1.28	7.19	2.08	4.96
3	0.094	5.421	0.70	14.81	7.91	3.00
4	0.065	11.697	0.84	8.80	1.48	3.26
5	0.050	8.459	0.08	8.26	1.84	4.81
6	0.074	10.969	0.35	6.04	6.43	3.88
7	0.086	12.209	1.07	8.49	3.91	4.66
8	0.096	10.292	0.66	5.22	6.89	3.63
9	0.084	9.793	0.27	12.87	5.87	2.08
10	0.054	6.438	1.46	11.37	8.25	4.35
11	0.099	10.050	0.94	6.96	3.11	3.78
12	0.039	13.376	1.13	6.56	7.97	3.83
13	0.072	8.335	0.55	14.58	9.50	2.15
14	0.025	8.091	0.77	10.28	0.65	3.48
15	0.036	11.120	0.87	7.41	4.08	4.72
16	0.090	11.917	1.20	14.13	4.33	4.44
17	0.044	9.564	1.25	9.12	7.38	2.49
18	0.068	14.796	0.39	12.61	9.41	2.27
19	0.052	6.015	0.74	5.06	4.27	2.02
20	0.027	6.269	0.13	5.40	2.76	3.68

Table A.2: SFRC sample points, $\mathbf{x}_i, i = 21...40$.

i	d_f (mm)	l_f (mm)	ϕ_f (%)	T_n (N)	δ_n (10^{-5} mm)	δ_{max} (10^{-4} mm)
21	0.063	6.813	0.91	11.45	0.79	4.32
22	0.058	11.416	0.34	10.07	2.42	3.94
23	0.061	11.241	0.47	12.15	4.66	2.36
24	0.066	5.685	1.40	11.94	5.81	4.28
25	0.088	14.112	1.09	13.46	2.86	2.58
26	0.095	12.548	0.89	7.89	6.62	2.68
27	0.071	12.650	1.42	11.72	5.38	2.34
28	0.049	5.954	0.25	10.68	2.37	2.77
29	0.092	13.787	1.17	9.25	8.99	2.86
30	0.046	14.247	1.15	5.88	1.22	3.04
31	0.015	13.457	0.40	8.88	3.18	3.52
32	0.081	12.872	1.04	11.10	5.51	3.20
33	0.033	10.516	0.59	12.58	1.05	3.24
34	0.077	7.522	0.29	7.26	4.80	4.85
35	0.023	9.237	0.50	7.60	1.44	4.51
36	0.040	6.737	1.31	10.53	1.73	4.00
37	0.056	8.830	0.06	13.82	6.24	2.20
38	0.028	8.699	0.79	12.31	9.72	4.93
39	0.020	12.111	0.44	6.77	8.60	3.40
40	0.017	5.188	0.19	13.74	5.07	2.82

Table A.3: SFRC model results, $\mathbf{x}_i, i = 1 \ldots 20$.

i	σ_i (MPa)	ε_i (10^{-4} mm/mm)	ε_f (10^{-4} mm/mm)	G^r (10^{-4} mJ/mm^3)	G^t (10^{-4} mJ/mm^3)
1	2.55	1.97	3.28	2.52	3.73
2	2.32	1.79	3.39	2.08	3.37
3	2.75	2.16	3.23	2.97	4.35
4	2.65	2.00	3.47	2.65	4.06
5	2.88	2.21	3.04	3.21	4.67
6	2.60	1.97	4.03	2.57	4.61
7	2.51	1.87	3.73	2.35	3.98
8	2.65	2.00	3.20	2.65	4.14
9	2.67	2.08	2.93	2.78	3.83
10	2.38	1.84	3.04	2.19	3.22
11	2.72	2.08	2.99	2.83	4.33
12	2.45	1.89	3.12	2.32	3.33
13	2.57	2.16	3.71	2.79	4.36
14	2.38	1.84	2.99	2.19	3.31
15	2.43	1.89	3.17	2.30	3.36
16	2.61	1.95	2.96	2.55	3.98
17	2.46	1.92	2.67	2.36	3.24
18	2.65	1.97	2.72	2.61	3.70
19	2.51	1.95	3.25	2.45	3.92
20	2.75	2.16	2.96	2.98	4.32

Table A.4: SFRC model results, $\mathbf{x}_i, i = 21...40$.

i	σ_i (MPa)	ε_i (10^{-4} mm/mm)	ε_f (10^{-4} mm/mm)	G^r (10^{-4} mJ/mm^3)	G^t (10^{-4} mJ/mm^3)
21	2.57	1.97	2.93	2.53	3.79
22	2.64	1.97	3.25	2.61	4.24
23	2.59	1.95	3.81	2.52	4.24
24	2.49	1.92	3.52	2.39	3.95
25	2.55	1.92	2.72	2.45	3.69
26	2.62	2.00	3.52	2.62	4.34
27	2.53	1.89	3.04	2.39	3.86
28	2.77	2.08	3.20	2.88	4.38
29	2.63	1.97	2.91	2.60	3.90
30	2.53	1.92	2.91	2.43	3.49
31	2.44	1.89	3.25	2.31	3.39
32	2.52	1.92	3.55	2.43	3.88
33	2.53	1.92	3.73	2.43	3.80
34	2.94	2.27	3.92	3.37	5.26
35	2.44	1.89	3.28	2.32	3.59
36	2.27	1.81	3.01	2.06	3.14
37	3.09	2.45	3.23	3.88	5.36
38	2.40	1.87	3.36	2.25	3.53
39	2.43	1.92	3.23	2.34	3.46
40	2.55	2.03	3.31	2.59	3.85

Appendix B

Guitar String Damping Coefficients

Table B.1: Mean A_n for each class.

Sensor	n	A	B	C	D	E	F	G
1	1	150.1	146.1	145.9	156.6	110.7	122.9	167.0
	2	231.9	229.0	196.3	182.6	154.6	156.5	244.1
	3	102.6	126.1	106.4	123.9	78.6	80.6	111.2
	4	62.1	59.8	49.3	58.6	36.5	32.8	57.5
	5	40.0	36.6	28.9	39.3	24.2	24.3	33.7
	6	36.8	29.5	28.0	44.3	27.2	26.5	42.2
	7	20.0	17.8	17.2	20.8	11.5	10.3	26.1
	8	17.9	16.7	15.1	12.9	10.0	12.3	20.6
	9	8.4	9.0	7.6	9.0	7.0	6.2	9.3
	10	29.9	30.8	26.1	32.9	22.2	22.8	31.5
	11	39.2	35.3	31.0	39.4	27.2	29.1	36.3
	12	22.7	19.8	17.9	18.8	13.2	13.6	17.3
2	1	57.0	53.1	55.3	63.0	43.0	48.2	65.0
	2	82.2	78.5	69.2	63.7	52.5	56.1	87.7
	3	62.9	86.6	89.7	89.2	51.6	71.1	105.4
	4	94.0	73.1	82.7	96.9	60.6	67.7	96.5
	5	71.8	63.3	59.5	76.2	43.8	51.3	60.7
	6	39.9	36.6	36.9	55.6	34.2	33.3	48.4
	7	17.8	17.4	17.2	24.6	14.0	12.3	23.9
	8	32.7	29.2	26.1	28.6	15.6	22.1	40.7
	9	43.7	47.5	41.4	40.0	31.7	36.3	51.5
	10	34.8	34.2	28.4	38.9	21.9	24.8	29.7
	11	30.3	25.6	23.9	32.9	19.4	23.2	27.2
	12	20.9	17.8	16.7	15.8	10.7	12.2	15.0

Table B.2: Coefficient of variation of A_n for each class.

Sensor	n	A	B	C	D	E	F	G
1	1	0.06	0.10	0.12	0.13	0.04	0.07	0.03
	2	0.18	0.19	0.17	0.14	0.12	0.05	0.08
	3	0.41	0.16	0.16	0.14	0.25	0.27	0.12
	4	0.36	0.39	0.35	0.26	0.34	0.08	0.10
	5	0.50	0.30	0.39	0.19	0.34	0.20	0.18
	6	0.31	0.21	0.33	0.17	0.22	0.18	0.31
	7	0.34	0.12	0.33	0.24	0.32	0.15	0.25
	8	0.19	0.26	0.14	0.11	0.05	0.14	0.15
	9	0.27	0.09	0.18	0.25	0.08	0.16	0.08
	10	0.11	0.20	0.29	0.10	0.08	0.17	0.16
	11	0.19	0.18	0.12	0.09	0.10	0.13	0.10
	12	0.16	0.17	0.23	0.21	0.09	0.13	0.13
2	1	0.13	0.08	0.14	0.08	0.04	0.08	0.04
	2	0.11	0.13	0.14	0.12	0.13	0.03	0.09
	3	0.49	0.25	0.31	0.30	0.32	0.29	0.13
	4	0.30	0.20	0.35	0.27	0.13	0.09	0.28
	5	0.38	0.13	0.40	0.18	0.28	0.25	0.06
	6	0.35	0.19	0.28	0.20	0.19	0.22	0.33
	7	0.34	0.11	0.31	0.27	0.18	0.11	0.28
	8	0.21	0.21	0.19	0.21	0.07	0.21	0.22
	9	0.08	0.08	0.17	0.08	0.05	0.10	0.06
	10	0.11	0.24	0.23	0.11	0.08	0.19	0.30
	11	0.19	0.20	0.18	0.10	0.13	0.14	0.17
	12	0.11	0.17	0.25	0.13	0.07	0.12	0.15

Table B.3: Mean α_n for each class.

Sensor	n	A	B	C	D	E	F	G
1	1	0.35	0.32	0.35	0.36	0.35	0.33	0.34
	2	0.58	0.56	0.51	0.72	0.60	0.76	0.53
	3	0.33	0.39	0.29	0.51	0.32	0.31	0.22
	4	0.80	0.76	0.76	0.92	0.66	0.66	0.54
	5	0.75	0.71	0.74	1.03	0.79	0.71	0.68
	6	0.60	0.61	0.59	0.65	0.70	0.64	0.92
	7	0.78	0.91	0.86	0.63	0.67	0.59	0.80
	8	1.60	1.14	1.15	0.78	1.08	1.09	0.94
	9	1.15	1.18	1.07	0.87	1.18	1.09	0.91
	10	2.50	2.51	2.34	0.97	1.51	1.48	2.84
	11	1.98	1.85	1.38	1.24	1.58	1.56	1.53
	12	2.71	2.42	2.44	1.52	2.21	1.83	1.64
2	1	0.36	0.34	0.38	0.37	0.41	0.37	0.34
	2	0.59	0.53	0.48	0.69	0.56	0.71	0.48
	3	0.13	0.33	0.28	0.32	0.27	0.36	0.41
	4	0.59	0.52	0.55	0.87	0.60	0.71	0.63
	5	0.55	0.49	0.60	0.73	0.60	0.56	0.54
	6	0.47	0.52	0.72	0.76	0.69	0.61	0.71
	7	0.81	0.75	0.80	0.55	0.71	0.70	0.70
	8	1.33	0.93	0.96	1.23	0.88	0.97	1.22
	9	1.29	1.28	1.10	0.85	1.17	1.03	1.14
	10	2.24	2.11	1.83	1.01	1.26	1.19	1.86
	11	2.07	1.69	1.38	1.32	1.36	1.43	1.50
	12	2.79	2.45	2.56	1.27	2.03	1.58	1.45

Table B.4: Coefficient of variation of α_n for each class.

Sensor	n	A	B	C	D	E	F	G
1	1	0.18	0.12	0.12	0.19	0.08	0.19	0.13
	2	0.13	0.22	0.19	0.25	0.15	0.07	0.10
	3	0.74	0.47	0.37	0.11	0.49	0.55	0.22
	4	0.18	0.20	0.22	0.25	0.41	0.27	0.12
	5	0.45	0.51	0.61	0.24	0.43	0.37	0.18
	6	0.12	0.18	0.23	0.08	0.14	0.38	0.31
	7	0.30	0.47	0.34	0.47	0.39	0.38	0.24
	8	0.14	0.34	0.14	0.24	0.10	0.23	0.12
	9	0.16	0.20	0.09	0.26	0.08	0.18	0.10
	10	0.49	0.27	0.54	0.16	0.07	0.26	0.29
	11	0.28	0.28	0.13	0.23	0.05	0.22	0.07
	12	0.21	0.32	0.36	0.61	0.14	0.41	0.15
2	1	0.14	0.13	0.10	0.16	0.06	0.22	0.12
	2	0.20	0.22	0.19	0.23	0.10	0.06	0.09
	3	1.27	0.18	0.32	0.31	0.75	0.61	0.20
	4	0.25	0.43	0.29	0.38	0.40	0.33	0.41
	5	0.52	0.34	0.63	0.21	0.55	0.45	0.13
	6	0.15	0.12	0.48	0.20	0.10	0.18	0.27
	7	0.16	0.30	0.26	0.23	0.13	0.31	0.20
	8	0.14	0.31	0.14	0.23	0.14	0.29	0.31
	9	0.22	0.21	0.13	0.33	0.14	0.17	0.14
	10	0.27	0.24	0.47	0.19	0.13	0.17	0.26
	11	0.25	0.40	0.20	0.25	0.07	0.14	0.17
	12	0.19	0.34	0.45	0.44	0.19	0.34	0.17

REFERENCES

[1] Robert M. Jones. *Mechanics of Composite Materials*. Taylor & Francis, 2 edition, 1999.

[2] G.P. Sendeckyj. A brief survey of empirical multiaxial strength criteria for composites. In *Composite Materials: Testing and Design (2nd Conference)*, 1972.

[3] Z. Hashin and S. Shtrikman. A variational approach to the theory of the elastic behaviour of multiphase materials. *J. Mech. Phys. Solids*, 11:127–140, 1963.

[4] R. Hill. Elastic properties of reinforced solids: Some theoretical principles. *J. Mech. Phys. Solids*, 11:357–372, 1963.

[5] T. Mori and K. Tanaka. Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metallurgica*, 21(5):571–574, 1973.

[6] K. Terada and N. Kikuchi. A class of general algorithms for multi-scale analyses of heterogeneous media. *Computer methods in applied mechanics and engineering*, 190(40):5427–5464, 2001.

[7] Z. Yuan and J. Fish. Toward realization of computational homogenization in practice. *International Journal for Numerical Methods in Engineering*, 73(3):361–380, 2008.

[8] J.E. Bishop, J.M. Emery, R.V. Field, C.R. Weinberger, and D.J. Littlewood. Direct numerical simulation in solid mechanics for understanding the the macroscale effects of microscale material variability. *Computer methods in applied mechanics and engineering*, 287:262–289, 2015.

[9] T. Kanit, S. Forest, I. Galliet, V. Mounory, and D. Jeulin. Determination of the size of

the representative volume element for random composites: statistical and numerical approach. *Int. J. of Solids and Structures*, 40:3647–3679, 2003.

[10] V.C. Li. Postcrack scaling relations for fiber reinforced cementitious composites. *J. Materials in Civil Engineering*, 4(1):41–57, 1992.

[11] H. Hilgenkamp and J. Mannhart. Grain boundaries in high-T_c superconductors. *Reviews of Modern Physics*, 74(2):485–549, 2002.

[12] P.S. Lee, H.R. Piehler, A.D. Rollett, and B.L. Adams. Texture clustering and long-range disorientation representation methods: application to 6022 aluminum sheet. *Metallurgical and Materials Transactions A*, 33(12):3709–3718, 2002.

[13] P. Louis and A.M. Gokhale. Application of image analysis for characterization of spatial arrangments of features in microstructure. *Metallurgical and Materials Transactions*, 26:1449–1456, 1995.

[14] B. L. Lu and S. Torquato. Lineal-path function for random heterogeneous materials. *Phys. Rev. A*, 45:922–929, 1992.

[15] B. L. Lu and S. Torquato. Nearest-surface distribution functions for polydispersed particle systems. *Phys. Rev. A*, 45:5530–5544, 1992.

[16] M. Stroeven, H. Askes, and L.J. Sluys. Numerical determination of representative volumes for granular materials. *Comput. Methods Appl. Mech. Engrg*, 193:3221–3238, 2004.

[17] S. Torquato. Nearest-neighbor statistics for packings of hard spheres and disks. *Physical Review E*, 51(4):3170–3182, 1995.

[18] B. Hiriyur, H. Waisman, and G. Deodatis. Uncertainty quantification in homogenization of heterogeneous microstructures modelled by XFEM. *Int. J. Numer. Meth. Engng.*, 88:257–278, 2011.

[19] R. Bostanabad, A.T. Bui, W. Xie, D.W. Apley, and W. Chen. Stochastic microstructure characterization and reconstruction via supervised learning. *Acta Materialia*, 103:89–102, 2016.

[20] D. Savvas, G. Stefanou, and M. Papadrakis. Determination of RVE size for random composites with local volume fraction variation. *Computer Methods in Applied Mechanics and Engineering*, 305:340–358, 2016.

[21] H. Xu, M.S. Greene, H. Deng, D. Dikin, C. Brinson, W.K. Liu, C. Burkhart, G. Papakonstantopoulos, M. Poldneff, and W. Chen. Stochastic reassembly strategy for managing information complexity in heterogeneous materials analysis and design. *Journal of Mechanical Design*, 135:1–12, 2013.

[22] H. Xu, Y. Li, C. Brinson, and W. Chen. A descriptor-based design methodology for developing heterogeneous microstructural materials system. *Journal of Mechanical Design*, 136:1–12, 2014.

[23] S. Torquato. Statistical description of microstructures. *Annual Review of Materials Research*, 32:77–111, 2002.

[24] D. Savvas, G. Stefanou, M. Papadrakakis, and G. Deodatis. Homogenization of random heterogeneous media with inclusions of arbitrary shape modeled by XFEM. *Comput. Mech.*, 54:1221–1235, 2014.

[25] S. Swaminathan, S. Ghosh, and N.J. Pagano. Statistically equivalent representative volume element for unidirectional composite microstructures: Part I - without damage. *Journal of Composite Materials*, 40(7):583–604, 2006.

[26] S. Swaminathan and S. Ghosh. Statistically equivalent representative volume element for unidirectional composite microstructures: Part II - with interfacial debonding. *Journal of Composite Materials*, 40(7):605–621, 2006.

[27] S. Ghosh, J. Bai, and P. Raghavan. Concurrent multi-level model for damage evolution in microstructurally debonding composites. *Mechanics of Materials*, 39:241–266, 2007.

[28] H.L. Frisch and F.H. Stillinger. Contribution to the statistical geometric basis of radiation scattering. *The Journal of Chemical Physics*, 38(9):2200–2207, 1963.

[29] A. Clement, C. Soize, and J. Yvonnet. Computational nonlinear stochastic homogenization using nonconcurrent multiscale aproach for hyperelastic heterogeneous microstructures analysis. *Int. J. Numer. Meth. Engng.*, 91:799–824, 2012.

[30] V.C. Li and G. Fischer. Reinforced ECC - an evolution from materials to structures. In *Proceedings of the first FIB congress*, pages 105–122, 2002.

[31] M.S. Greene, H. Xu, S. Tang, W. Chen, and W.K. Liu. A generalized uncertainty propagation criterion from benchmark studies of microstructured material systems. *Comput. Methods Appl. Mech. Engrg.*, 254:271–291, 2013.

[32] Ian H. Witten, Eibe Frank, Mark A. Hall, and Christopher J. Pal. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 4 edition, 2017.

[33] T. Mueller, A.G. Kusne, and R. Ramprasad. Machine learning in materials science: recent progress and emerging applications. *Reviews in Computational Chemistry*, 29:186–273, 2016.

[34] Y. Liu, T. Zhao, W. Ju, and S. Shi. Materials discovery and design using machine learning. *Journal of Materiomics*, 3(3):159–177, 2017.

[35] H.S. Rao and A. Mukherjee. Artificial neural networks for predicting the macromechanical behavior of ceramic-matrix composites. *Computational Materials Science*, 5:307–322, 1996.

[36] A. Clement, C. Soize, and J. Yvonnet. Uncertainty quantification in computational stochastic multiscale analysis of nonlinear elastic materials. *Comput. Methods Appl. Mech. Engrg.*, 254:61–82, 2013.

[37] D. Li, W. Yang, and W. Zhang. Cluster analysis of stress corrosion mechanisms for steel wires used in bridge cables through acoustic emission particle swarm optimization. *Ultrasonics*, 77:22–31, 2017.

[38] N.F. Dow and B.W. Rosen. Evaluations of filament-reinforced composites for aerospace structural applications. Technical report, National Aeronautics and Space Administration, 1965.

[39] M. Piggott. *Load Bearing Fibre Composites*. Kluwer Academic Publishers, 2 edition, 2002.

[40] A. Naaman and H. Najm. Bond-slip mechanisms of steel fibers in concrete. *ACI Materials Journal*, 88(2):135–145, March 1991.

[41] S. Hamoush, T. Abu-Lebdeh, T. Cummins, and B. Zornig. Pullout characterizations of various steel fibers embedded in very high-strength concrete. *Journal of Engineering and Applied Sciences*, 3(2):418–426, 2010.

[42] D. Kalman. Use of steel fiber reinforced concrete for blast resistant design. Master's thesis, Kansas State University, 2010.

[43] A. Bentur and S. Mindess. *Fibre Reinforced Cementitious Composites*. Taylor & Francis, 1990.

[44] J. Aveston and A. Kelly. Theory of multiple fracture of fibrous composites. *Journal of Materials Science*, 8(3):352–362, March 1973.

[45] W.A. Curtin. Stochastic damage evolution and failure in fiber-reinforced compostes. *Advances in Applied Mechanics*, 36:163–253, 1998.

[46] H. Cuypers and J. Wastiels. Stochastic matrix-cracking model for textile reinforced cementitious composites under tensile loading. *Materials and Structures*, 39:777–786, 2006.

[47] J. Lubliner, J. Oliver, S. Oller, and E. Onate. A plastic-damage model for concrete. *Int. J. Solids and Structures*, 25:299–326, 1989.

[48] T. Jankowiak and T. Lodygowski. Identification of parameters of concrete damage plasticity model. Technical report, Poznan University of Technology, 2005.

[49] N.A. Nordendale. *Modeling and Simulation of Brittle Armors Under Impact and Blast Effects*. PhD thesis, Vanderbilt University, 2013.

[50] S. Ghosh. Foundational aspects of a multi-scale modeling framework for composite materials. *Integrating Materials and Manufacturing Innovation*, 4(9):1–28, 2015.

[51] J.T. Oden and T.I. Zohdi. Analysis and adaptive modeling of highly heterogeneous elastic structures. *Comput. Methods Appl. Mech. Engrg*, 148:367–391, 1997.

[52] P. Raghavan and S. Ghosh. Concurrent multi-scale analysis of elastic composites by a multi-level computational model. *Comput. Methods Appl. Mech. Engrg*, 193:497–538, 2004.

[53] S. Ghosh. Adaptive hierarchical-concurrent multiscale modeling of ductile failure in heterogeneous metallic materials. *The Journal of The Minerals, Metals, & Materials Society (TMS)*, 67:129–142, 2015.

[54] J.M. Guedes and N. Kikuchi. Preprocessing and postprocessing for materials based on the homogenization method with adaptive finite element methods. *Comput. Methods Appl. Mech. Engrg*, 83:143–198, 1990.

[55] S. Ghosh, K. Lee, and S. Moorthy. Multiple scale analysis of heterogeneous elastic

structures using homogenization theory and Voronoi cell finite element method. *Int. J. Solids and Structures*, 32(1):27–62, 1995.

[56] J. Fish and K. Shek. Multiscale analysis of composite materials and structures. *Composites Science and Technology*, 60:2547–2556, 2000.

[57] W. E and B. Engquist. The heterogeneous multiscale methods. *Comm. Math. Sci.*, 1(1):87–132, 2003.

[58] F.G. Yuan, N.J. Pagano, and X. Cai. Elastic moduli of brittle matrix composites with interfacial debonding. *Int. J. Solids and Structures*, 34:177–201, 1997.

[59] P. Kabele. Multiscale framework for modeling of fracture in high performance fiber reinforced cementitious composites. *Engineering Fracture Mechanics*, 74:194–209, 2007.

[60] M. Kastner, S. Muller, and V. Ulbricht. XFEM modelling of inelastic material behaviour and interface failure in textile-reinforced composites. *Procedia Materials Science*, 2:43–51, 2013.

[61] M.G. Pike and C. Oskay. Modeling random short nanofiber- and microfiber-reinforced composites using the extended finite-element method. *J. Nanomech. Micromech.*, 5:1–11, 2015.

[62] M.G. Pike and C. Oskay. XFEM modeling of short microfiber reinforced composites with cohesive interfaces. *Finite Elements in Analysis and Design*, 106:16–31, 2015.

[63] M. Pike, M.A. Hickman, and C. Oskay. Interactions between multiple enrichments in extended finite element analysis of short fiber reinforced composites. *International Journal for Multiscale Computational Engineering*, 13(6):507–531, 2015.

[64] D. Abbott and E. Segerman. Historical background to the strings used by catgut-scrapers. *Journal of the Catgut Acoustical Society*, (25):24–26, 1976.

[65] D. Abbott and E. Segerman. More on the history of strings and instruments. *Journal of the Catgut Acoustical Society*, (29):29–33, 1978.

[66] N.C. Pickering. Strings and metallurgy. *Journal of the Catgut Acoustical Society*, 3(4):24–29, 1997.

[67] I. Firth, R. Bain, and A. Gallaher. String design equations. *Journal of the Catgut Acoustical Society*, (46):3–6, 1986.

[68] A.J.M. Houtsma. Inharmonicity of wound guitar strings. *Journal of Guitar Acoustics*, (6):60–64, 1982.

[69] E.B. Davis. A design guide for steel core, bronze-wound musical instrument strings. *Journal of the Catgut Acoustical Society*, (38):40, 1982.

[70] R.J. Hanson. Analysis of ”live” and ”dead” guitar strings. *Journal of the Catgut Acoustical Society*, (48):10–16, 1987.

[71] N.C. Pickering. Nonlinear behavior in overwound violin strings. *Journal of the Catgut Acoustical Society*, 1(3):46–50, 1989.

[72] M. Hancock. The dynamics of musical strings. *Journal of the Catgut Acoustical Society*, 1(8):23–25, 1991.

[73] N.H. Fletcher. Plucked strings - a review. *Journal of the Catgut Acoustical Society*, (26):13–17, 1976.

[74] Lord Rayleigh. *The Theory of Sound*, volume 1. MacMillan and Co. Ltd., 1894.

[75] Harvey Fletcher. Normal vibration frequencies of a stiff piano string. *The Journal of the Acoustical Society of America*, 36(1):203–209, 1964.

[76] I. Firth. Overwrapped strings: Design guide incorporating acoustical limitations. *Journal of the Catgut Acoustical Society*, (45):7–9, 1986.

[77] J.B. Allen. On the aging of steel guitar strings. *Catgut Acoust. Soc. Newsletter*, 1(26):27–29, 1976.

[78] M. Hancock. The dynamics of isolated musical strings. *Journal of the Catgut Acoustical Society*, (38):24–29, 1982.

[79] H. Penttinen and V. Valimaki. A time-domain approach to estimating the plucking point of guitar tones obtained with an under-saddle pickup. *Applied Acoustics*, 65(12):1207–1220, 2004.

[80] J.A. Kemp. The physics of unwound and wound strings on the electric guitar applied to the pitch intervals produced by tremolo/vibrato arm systems. *PLoS ONE*, 12(9), 2017.

[81] F. Bulbul. Microscopic examination of play-worn steel music strings. *Friction and Wear Research*, 3:1–14, 2015.

[82] F. Bulbul and T. Karacali. Investigation of play-wear damage on steel music strings using atomic force microscope and low level resistance method. *Savart Journal*, 1(5):1–10, 2015.

[83] Andrew V. Olver, Daniel Wilson, and P. Shaun J. Crofton. Investigation of service failures of steel music wire. *Engineering Failure Analysis*, 14:1224–1232, 2007.

[84] H.G. de Dayan and A. Behar. The quality of strings for guitars: An experimental study. *Journal of Sound and Vibration*, 64(3):421–431, 1979.

[85] Jurgen Meyer. Quality aspects of the guitar tone. In Erik V. Jansson, editor, *Function, Construction and Quality of the Guitar*, number 38, pages 51–76. Royal Swedish Academy of Music, 1983.

[86] Jurgen Meyer. The function of the guitar body and its dependence on construction

details. In Erik V. Jansson, editor, *Function, Construction and Quality of the Guitar*, number 38, pages 77–100. Royal Swedish Academy of Music, 1983.

[87] Neville H. Fletcher and Thomas D. Rossing. *The Physics of Musical Instruments*. Springer-Verlag New York, Inc., 2nd edition, 1998.

[88] J. Woodhouse. On the synthesis of guitar plucks. *Acta Acustica United With Acustica*, 90:928–944, 2004.

[89] J. Woodhouse. Plucked guitar transients: Comparison of measurements and synthesis. *Acta Acustica United With Acustica*, 90:945–965, 2004.

[90] J. Woodhouse, E.K.Y. Manuel, L.A. Smith, A.J.C. Wheble, and C. Fritz. Perceptual thresholds for acoustical guitar models. *Acta Acustica*, 98:475–486, 2012.

[91] S. Liang and Alvin W.Y. Su. Modeling and analysis of acoustic musical strings using Kelly-Lochbaum lattice networks. *Journal of Information Science and Engineering*, 20:1161–1182, 2004.

[92] M. Karjalainen, H. Penttinen, and V. Valimaki. Acoustic sound from the electric guitar using DSP techniques. In *IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings*, pages 773–776, 2000.

[93] L.M. Ghiringhelli, J. Vybiral, S.V. Levchenko, C. Draxl, and M. Scheffler. Big data of materials science: Critical role of the descriptor. *Physical Review Letters*, 114(10):105503, 2015.

[94] Margaret H. Dunham. *Data Mining: Introductory and Advanced Topics*. Prentice Hall, 2003.

[95] Mehmed Kantardzic. *Data Mining: Concepts, Models, Methods, and Algorithms*. John Wiley & Sons, Inc., 2003.

[96] S. Torquato. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*. Springer, 2002.

[97] T. Belytschko and R. Gracie. A review of extended/generalized finite element methods for material modeling. *Modelling Simul. Mater. Sci. Eng.*, 17:1–24, 2009.

[98] C. Daux, N. Moes, J. Dolbow, N. Sukumar, and T. Belytschko. Arbitrarily branched and intersecting cracks with the extended finite element method. *International Journal for Numerical Methods in Engineering*, 48:1741–1760, 2000.

[99] G. Zi, J. Song, E. Budyn, S. Lee, and T. Belytschko. A method for growing multiple cracks without remeshing and its application to fatigue crack growth. *Modelling Simul. Mater. Sci. Eng.*, 12:901–915, 2004.

[100] S. Osher and J.A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.

[101] J.A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision and Material Science*. Cambridge University Press, 1999.

[102] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Comp. Math*, 4:389–396, 1995.

[103] C. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[104] ASTM International. ASTM A228/A228M-18 Standard Specification for Steel Wire, Music String Quality. Standard, ASTM International, West Conshohocken, PA, 2018.

[105] Brian N. Ball and James C. Harrington. Ultra-high tensile strength maraging steel music instrument string, February 4, 2014. US Patent 8642861B2.

[106] Sina Vosough and Goran Berglund. Music string, August 17, 2010. US Patent 7777108B2.

[107] Brian N. Ball. Musical instrument string having cobalt alloy wrap wire, July 17, 2012. US Patent 8222504B1.

[108] T.D. Rossing, editor. *The Science of String Instruments*. Springer Science and Business Media, LLC, 2010.

[109] C. Valette. *The mechanics of vibrating strings*. Springer-Verlag, 1995.

[110] H. Nyquist. Certain factors affecting telegraph speed. *Bell Labs Technical Journal*, 3(2):324–346, 1924.

[111] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(4):623–666, 1948.

[112] D. and Gabor. Theory of communcation. Part 1: The analysis of information. *Journal of the institution of electrical engineers*, 93(26):429–441, 1946.

[113] O. Yapar. *AE Based Health Monitoring and Bonded FRP Patch Repair in Bridge Management*. PhD thesis, Vanderbilt University, May 2015.

[114] International Organization for Standardization. ISO 16:1975 acoustics - standard tuning frequency (standard musical pitch). In *ICS 17.140.01 Acoustic measurements and noise abatement in general*. International Organization for Standardization, Geneva, Switzerland, 1975.

[115] Masahiro Harazono. Electromagnetic pickup response excited by a string vibration. *Journal of the Acoustical Society of Japan (E)*, 10:23–29, 1989.

[116] Michael Collins. Finite element analysis of a vibrating guitar string and project management. Master's thesis, University of Limerick, 2011.

[117] Richard Mark French. *Engineering the Guitar: Theory and Practice*. Springer, 2009.

[118] V. Vapnik and A. Lerner. Pattern recognition using generalized portrait method. *Automation and Remote Control*, 24:774–780, 1963.

[119] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152. ACM Press, 1992.

[120] C. Cortes and V Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[121] S. E. Kruger, M. Schaffoner, M. Katz, E. Andelic, and A. Wendemuth. Speech recognition with support vector machines in a hybrid system. In *Interspeech*, pages 993–996, 2005.

[122] Jean Philippe Vert, Koji Tsuda, and Bernhard Schölkopf. A primer on kernel methods. In Bernhard Schölkopf, Koji Tsuda, and Jean Philippe Vert, editors, *Kernel Methods in Computational Biology*. MIT Press, 2004.