Title

Trustworthy Machine Learning for Experimental Characterization of Ceramic Matrix Composites

Author

Muir, Caelin

Publication Date

2023

University of California
Santa Barbara

Trustworthy Machine Learning for Experimental Characterization of Ceramic Matrix Composites

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Materials

by

Caelin Muir

Committee in charge:

Professor Tresa M. Pollock, Committee Co-Chair
Professor Samantha H. Daly, Committee Co-Chair
Professor Matthew Begley
Professor Ram Seshadri
Professor Enoch Yeung

December 2023

The Dissertation of Caelin Muir is approved.

Professor Matthew Begley

Professor Ram Seshadri

Professor Enoch Yeung

Professor Samantha H. Daly, Committee Co-Chair

Professor Tresa M. Pollock, Committee Co-Chair

November 2023

Trustworthy Machine Learning for Experimental Characterization of Ceramic Matrix Composites

by

Caelin Muir

Acknowledgements

Although acknowledgements are traditionally limited to one page, I firmly believe the work contained in here would not have been possible without the personal and professional support of a large number of people. Because of this, limiting my acknowledgements to one page seems disingenuous. So, in the spirit of Lesley Gore, ”It’s my dissertation and I can acknowledge as many people as I want to”.

First are my advisors at UCSB: Professor Samantha Daly, Professor Tresa Pollock. Sam, your careful attention to detail, endless curiosity, and optimism in the face of setbacks is inspiring. I cannot overstate how important your mentorship is, and how grateful I am to have your advice to guide me through my scientific and personal challenges. Tresa, working with you and your group has been instrumental in maintaining broader perspectives and becoming a well-rounded investigator. Every time I talk with you I learn something valuable.

Next is the NASA Space Technology Research Fellowship program, for funding my studies and providing me access to research opportunities that would have otherwise been inaccessible. Moreover, I want to thank my NASA and Michigan advising team: Dr. Craig Smith, Doug Kiser, Dr. Amjad Almansour, and Dr. Kathy Sevener. Your weekly guidance and insight with respect to AE and CMCs has been invaluable for conducting my research. I always look forward to our weekly meetings and trivia questions.

To my colleagues both past and present, your friendship and scientific support has shaped me for the better. In no particular order Bhavana Swaminathan, Jeff Rosin, Neal Brodnik, Jayden Plumb, Andrew Christison, Nick Tulshibagwale, Andrew Furst, and Abed Musaffar. Talking about science with you all is so much fun and always seems to lead to new ideas. Long days in the office, conferences, trivia nights, and weekends at campus point have been some of the best experiences in grad school because of you.

To my dear friends both in and out of graduate school: John Garcia, Phillip Griffith, Austin Hardy, Aaron Engel, Mary Franitza, Michael Chin, Matthew Clawson, and everyone else un-named who if written down would truly make this section egregiously long, you all have been a constant source of joy for me. Between board games and raid nights, oceanside and mountainside adventures, I am better for having you in my life.

Finally to my parents, family, and the cat that won't read this. Mom and Dad, none of this would be possible without your love and support. I don't think I'll ever be able to truly express how grateful I am for the opportunities you gave me. Christine and Chris, Alex and Sarah, I couldn't ask for better siblings or role models. And Jenny, you are such a bright spot in my life. Your endless patience, kindness, and willingness to try my sometimes too-hot-to-eat cooking never ceases to amaze me.

Thank you all.

Curriculum Vitæ

Caelin Muir

Education

2023	Ph.D. in Materials, University of California, Santa Barbara.
2019	B.S. in Materials Science & Engineering, University of Illinois at Urbana-Champaign

Publications

1 **C. Muir**, B. Swaminathan, K. Fields, A. S. Almansour, K. Sevener, C. Smith, M. Presby, J. D. Kiser, T. M. Pollock, and S. Daly. A machine learning framework for damage mechanism identification from acoustic emissions in unidirectional SiC/SiC composites *NPJ Computational Materials*, 7. (2021).

2 **C. Muir**, B. Swaminathan, A. S. Almansour, K. Sevener, C. Smith, M. Presby, J. D. Kiser, T. M. Pollock, and S. Daly. Damage mechanism identification in composites via machine learning and acoustic emission *NPJ Computational Materials*, 7. (2021).

3 **C. Muir**, N. Tulshibagwale, A. Furst, B. Swaminathan, A. S. Almansour, K. Sevener, M. Presby, J. D. Kiser, T. M. Pollock, S. Daly, and C. Smith. Quantitative benchmarking of acoustic emission machine learning frameworks for damage mechanism identification *Integrating Materials and Manufacturing Innovation.*

4 **C. Muir**, B. Swaminathan, A. K. Musaffar, N. R. McCarthy, A. S. Almansour, T. M. Pollock, J. D. Kiser, C. Smith, S. Daly, and K. Sevener. In situ crack opening displacement growth rates of SiC/SiC ceramic matrix minicomposites *Journal of the European Ceramic Society*

5 N. R. Brodnik, **C. Muir**, N. Tulshibagwale, J. Rossin, M.P. Echlin, C.M. Hamel, S.L.B. Kramer, T.M. Pollock, J.D. Kiser, C. Smith, and S.H. Daly, Perspective: Machine Learning in Experimental Solid Mechanics *Journal of Mechanics and Physics of Solids*

6 N. R. Brodnik, S. Carton, **C. Muir**, S Ghosh, D. Downey, M.P. Echlin, T.M. Pollock, and S.H. Daly, Perspective: Large Language Models in Applied Mechanics *Journal of Applied Mechanics*

In Preparation

1 **C. Muir**, N. Tulshibagwale, A. S. Almansour, K. Sevener, M. Presby, J. D. Kiser, T. M. Pollock, S. Daly, and C. Smith. Autoencoders enable supervised real-time damage mechanism identification in simplified SiC/SiC composite structures from acoustic emissions *Proceedings of the National Academy of Science.*

2 N. Tulshibagwale, **C. Muir**, N. Brodnik, S. Daly, and C. Smith. Acoustic Emission in Composites *Applied Mechanics Reviews.*

Abstract

Trustworthy Machine Learning for Experimental Characterization of Ceramic Matrix Composites

by

Caelin Muir

SiC/SiC composites are refractory, damage tolerant, high-strength materials that are ideal for use in turbine engines. Although these properties allow engineers to design aircraft with better fuel efficiency, the safety-critical nature of their application space necessitates a detailed understanding of how damage initiates and progresses. Recent advances in machine learning (ML) and related statistical tools have created novel characterization pathways for understanding damage accumulation, however their widespread adoption is limited due to a lack of interpretability. This work lays the foundation for trustworthy ML for interpreting acoustic emissions (AE) that are produced when SiC/SiC composites sustain damage. A modified signal representation scheme was combined with unsupervised clustering to form the spectral framework. This represents the first AE-ML approach that can distinguish between fiber break and matrix crack signals in SiC/SiC minicomposites. The spectral framework was then benchmarked against 4 state-of-the-art AE-ML frameworks and shown to achieve superior performance. Community guidelines for standardized benchmarking were proposed to promote greater transparency. Finally, an *in situ* x-ray computed tomography experiment demonstrated that matrix crack signals overlap with early fiber breaks, and prevents the use of unsupervised clustering for realistic composite geometries. An autoencoder framework was created to overcome this limitation and used to demonstrate that the frequency distribution of fiber break signals is compact. These findings establish a pathway for real-time health monitoring.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Ceramic Matrix Composites

Fiber-reinforced ceramic matrix composites (CMCs) are a class of lightweight, damage-tolerant, refractive structural materials for extreme environments. Their architected design consists of an array of ceramic fibers connected to a ceramic matrix via a weak interphase (Fig. 1.1). Such design allows for substantial damage tolerance in the form of crack deflection and arrest, which allows engineers to leverage the desirable properties of monolithic ceramics while preventing catastrophic failure [15]. As a result, CMCs have been incorporated into the hot section of turbine engines as high pressure temperature vanes, shrouds, and combustor liners, and are being considered for fuel cladding in generation IV nuclear reactors [16–18]. However, despite their desirable properties, the safety critical nature of their application spaces necessitates a detailed understanding of how CMCs respond to thermo-mechanical and environmental loads.

As a result, micromechanical modeling of damage chronology in CMCs has been the subject of extensive research and has provided valuable insights into how constituent properties influence crack formation and growth. For example, Aveston, Cooper, and

Kelly (ACK) and Marshall have shown that there exists a finite range of damage tolerant CMC microstructures [10, 19]. Fiber volume fractions must be sufficiently high such that fibers remain intact in the wake of matrix cracking, and interfacial shear stresses between the fibers and matrix must be large enough to resist crack opening, but not so large as to completely prevent fiber debonding and sliding. Hutchinson and Jensen extended the ACK theoretical framework to include the effects of energy absorption from interfacial debonding, compressive thermal mismatch stresses, and non-uniform friction between the fiber and the matrix [11]. They used this framework to analyze steady-state cracking for two spatial distributions of fibers, a hexagonal and square array, and an arbitrary combination of elastic parameters. Their analysis quantified microstructural effects on crack opening displacements (CODs), and provided guidance on the optimal interfacial parameters and critical processing temperatures needed to obtain a damage tolerant microstructure.

More recent numerical simulations of composite damage by Chateau et al. led to the hypothesis that heterogeneities in fiber distributions and interfacial properties need to be considered to explain CMC mechanical behavior [20]. They found that the experimentally measured stress-strain response of their CMC could only be modeled when effects from asymmetrical interfacial shear stresses and cascading fiber breaks were included. These hypotheses arise from heterogeneities which lead to local load sharing and were supported by *in-situ* x-ray computed tomography observations. In turn, they provide evidence that models must include effects from microstructural heterogeneities. Models that do not, such as that of Hutchinson and Jensen, are expected to not be sufficiently complex to describe damage progression or severity in as-processed composite structures.

These and other foundational studies provide important insights, but the complex and interacting factors within real microstructures restrict predictive models to order-of-magnitude COD estimates [21–27]. Validation of models, for the purpose of obtaining accurate COD calculations, is critical for the characterization and mitigation of environmental effects [28, 29]. Despite its key importance, there are a limited number of *in-situ* studies that can be used for this purpose [30–33]. As a result, it is unknown which modeling choices are most appropriate and how inhomogeneities present in real CMC microstructures influence crack growth.

1.2 Acoustic Emission Based Damage Mechanism Identification

When a structure sustains damage, accumulated strain energy is rapidly released into the surrounding medium in the form of elastic waves. These waves, as well as the nondestructive evaluation (NDE) technique used to record these waves, are called acoustic emission (AE) [34]. AE is uniquely powerful because it allows researchers to triangulate damage source locations [34], identify areas of concern in the material bulk [35], and evaluate the severity of incurred damage [36] in real time using a relatively inexpensive (<\$10,000) array of piezoelectric transducers affixed to the structure. Moreover, AE allows researchers to identify damaged areas in the material bulk down to the order of nanometers in size [37], with a time resolution on the order of 100s of μs, allowing for rapid real-time damage detection of events below the optical limit of methods such as x-ray computed tomography [12, 30].

Because of its unique capabilities, AE research efforts have been on-going since the 1980s [38–40]. However, initial studies were limited by both the experimental hardware and data storage limitations of the time. The narrow frequency response of resonant transducers forced waves to be recorded as dampened sine waves, losing detailed frequency (and therefore time domain) resolution. Moreover, storage and memory constraints mandated that the full waveform could not be preserved. Instead, waves were stored as a list of time-domain parameters such as first peak amplitude, energy, and duration (Table 1.1), which resulted in further information loss [34]. The development of broadband transducers, combined with advances in computer hardware during the 1990s, enabled researchers to record the full frequency spectrum of a signal and digitally store the waveform. This method, known as modal AE (MAE), is now standard practice [13, 34, 41].

The advent of MAE spurred renewed interest in AE with a specific focus on evaluating the chronology of damage progression as a function of specimen geometries [41–43], architectural and processing choices [44], and loading conditions [45]. These efforts laid the foundation for the hypothesis that damage mechanisms in multi-phase structural materials, such as composites, can be identified directly from their acoustic emission (AE) [46–49]. Developing this capability has wide-reaching ramifications for lifetime prediction investigations and *in operando* monitoring of advanced structural materials. It would allow researchers to augment damage triangulation [34, 35], lifetime prediction [33], and high-resolution optical studies [12, 30] with complementary mechanism-informed data streams. Moreover, it would facilitate investigations on how processing paths and microstructural landscapes are tied to the failure response of advanced structural materials, and would significantly broaden the potential application space of CMCs [16, 18, 50–52]. As such, the high-fidelity damage mode identification of AE signals has become a core objective of modern AE investigations.

Yet, despite the growing availability and abundance of statistical ML tools [6, 53], there has been sparse analysis of their strengths, limitations, or trustworthiness when applied to AE data. As a result, it is unclear if the approaches adopted in the literature are suited for the task of damage mechanism identification in CMC systems, where the elastic, and therefore acoustic, properties of the constituents are similar [54–57]. In order to understand the limitations of existing damage mechanism identification frameworks in composite systems, it is necessary to understand how AE are influenced by their local environment and how each step in the damage mechanism identification pipeline impacts the ability to do such identification. To this end, I will briefly discuss wave propagation in solids and identification of relevant waveform characteristics as they relate to the original damage mechanism (Chapter 1.2.1). I will then outline the creation of a damage mechanism identification pipeline, survey the most commonly used featurization schemes (Chapter 1.3.1), clustering algorithms (Chapter 1.3.3), and data-visualization methods (Chapter 1.3.5). Strengths and limitations of each step in a damage mechanism identification pipeline will then be used to contextualize prior investigations, and to highlight existing challenges which this dissertation addresses (Chapter 1.4).

1.2.1 Wave Propagation in Solid Materials

Historically, researchers have used Lamb's homogeneous equations to understand AE phenomena [13, 58], where it is assumed that acoustic waves are confined to a thin plate of thickness h. This model has been useful for engineering applications as many modern manufactured composites typically follow this constraint [59]. When waves are modeled this way, the predicted out-of-plane flexural (w) and in-plane extensional (u^0, v^0) displacements are governed by:

$$D\nabla^4 w + \rho\frac{\partial^2 w}{\partial t^2} = 0 \tag{1.1}$$

$$\frac{\partial^2 u^0}{\partial x^2} + \left(\frac{1-\nu}{2}\right)\frac{\partial^2 u^0}{\partial y^2} + (1+\nu)\frac{\partial^2 v^0}{\partial x \partial y} = \frac{\rho}{A}\frac{\partial^2 u^0}{\partial t^2} \tag{1.2}$$

$$\frac{\partial^2 v^0}{\partial y^2} + \left(\frac{1-\nu}{2}\right)\frac{\partial^2 v^0}{\partial x^2} + (1+\nu)\frac{\partial^2 u^0}{\partial x \partial y} = \frac{\rho}{A}\frac{\partial^2 v^0}{\partial t^2} \tag{1.3}$$

with constants A and D (the bending stiffness):

$$A = \frac{Eh}{1-\nu^2} \tag{1.4}$$

$$D = \frac{Eh^3}{12(1-\nu^2)} \tag{1.5}$$

where ρ is the density, E is the elastic modulus, and ν is the Poisson ratio.

While this model is a powerful tool for predicting the basic properties of elastic waves in solids, its assumptions of elastic isotropy and material homogeneity are often violated. It fails to describe the effects pores, grain boundaries, and architectural/processing heterogeneities influence wave propagation [41, 42, 60–62].

To account for these effects, current modeling efforts simulate AE in real microstructures by solving Lamb's homogeneous equations via finite element method (FEM) calculations, a process known as forward modeling [63–69]. This allows researchers to computationally explore the effects of microstructural heterogeneities, complex geometries, and determine how waveform characteristics are impacted by these factors [70–73]. Recent studies suggest that frequency and time-frequency representations encode information related to the emitting damage mechanism, and that time domain features are impacted by factors unrelated to the damage mechanism. Sause and Horn demonstrated differences between damage modes in the time-frequency domain with an FEM model of a carbon fiber reinforced composite [67]. Their model produced frequency spectra in good agreement with experiment and supported the claim that changes to power carried by a frequency band are minimal in the near field [67, 74]. Gall et al. showed that in the absence of defect structures, rise time, amplitude, and energy scatter are dependent on source-to-sensor distance, while partial power is not [71]. Similarly, Aggelis et al. used numerical simulations to show that rise angle values (Table 1.1) increase with increasing source-to-sensor distance [72] and provide evidence that the density of scattering interfaces distorts the waveform in the time-domain.

Forward modeling studies guide us to two complementary conclusions: (i) if there is information encoded in AE which enables damage mode discrimination, it likely exists within the frequency domain; and (ii) the high degree of perturbation imposed on time-domain features by experimental factors likely renders them inadequate for differentiating source mechanisms. These deductions highlight the necessity for determining salient features for machine learning applications, where the accuracy of the output is directly dependent on the quality of feature choice.

1.3 Machine Learning

Machine learning (ML) is the set of statistical tools for building a model to approximate a function; it can be thought of as a set of non-linear curve fitting methods [75]. Subsequently, the set of tasks which an ML model can accomplish is broad and can range from clustering, where similar objects are grouped together, to the prediction of future events. Regardless of the exact details of the task, any ML algorithm can be classified as an unsupervised, semi-supervised, or supervised algorithm depending on what data is available at the time of model fitting [4, 5, 76–79]. For the case of AE damage mechanism identification, direct correlations between waveforms and damage mechanisms are unavailable, and ML models are typically restricted to the class of unsupervised clustering algorithms whose workflow follows [78, 80]:

1. **Experimentation**: An experiment is conducted and n waveforms are collected

(Figure 1.2a)

2. **Feature Extraction**: Waveforms are vectorized by extracting d pertinent features, i.e. each waveform is represented by a d-dimensional vector (Figure 1.2b)
3. **Pre-processing**: A metric or similarity function is chosen, and the set of feature vectors undergoes an *optional* transformation, such as re-scaling (Figure 1.2c)
4. **Clustering**: A clustering algorithm is selected and applied to group waveforms by their similarities, as defined by the previously chosen similarity function (Figure 1.2d)
5. **Error Analysis**: Post-clustering analysis is performed to assess the validity of results (Figure 1.2d)

The following sections present an overview of the most common feature extraction techniques, pre-processing methods, and clustering algorithms. The strengths and weaknesses of each choice are discussed, followed by a critical review of existing damage mechanism identification studies.

1.3.1 Feature Extraction

Importance of Salient Features

The accuracy of an AE-ML framework is interdependent on each step described above, however it is foremost contingent on feature choice; waveforms must be represented by features which are salient in order to find meaningful partitions. To illustrate this, consider a researcher who aims to differentiate a set of ducks and cats (Figure 1.3a). The researcher could represent each animal by color, resulting in the partition in Figure 1.3b, or by their ability to fly (Figure 1.3c). While both partitions are valid, only Figure 1.3c is a useful result if the researcher aimed to differentiate species. This example illustrates the more general principle that when objects are described by improper features, it is possible to obtain compact and well separated clusters not representative of desired categories, and a substantial loss in discriminating power will occur [78, 81].

Traditional Features

Historically, AE waveforms have been represented by parameters in the time domain (Figure 1.4a), the frequency domain, or by composite values comprised of two or more base values [13, 46, 49, 82–92]. Table 1.1 lists common time domain, frequency, and composite features, where $x(t)$ is the measured signal and $F[*]$ is the Fourier transform operator. It is worthwhile to note that the predominance of time domain features appears to be due to the ease with which they are extracted, either via commercial AE software or in-house programs. However, the efficacy of these features is currently unclear, as are the impacts when used in an AE-ML framework.

Continuous Wavelet Transform

Because the impacts of time domain features is unclear, investigators additionally include frequency domain features to represent acoustic waveforms. While this practice is supported by forward modeling studies which show frequency features contain salient information, the optimal way to extract meaningful frequency information remains unclear. This is because AE waveforms are non-stationary, meaning that their frequency content changes in time. When this is the case, the use of the usual fast Fourier transform (FFT) to extract frequency information may not be the optimal solution, as it systematically misrepresents frequency content for non-stationary waves [93]. Because of this, researchers have espoused time-frequency representations such as the continuous wavelet transform, wavelet packet transform, and Hilbert-Huang transform, which are described in detail below [1, 64, 94–97].

The continuous wavelet transform (CWT) is a popular time-frequency representation for inspecting non-stationary signals, whose frequency content changes in time (Figure 1.4b). It can be used for either qualitative AE analyses [64, 94, 98] or quantitative characterization [1, 99]. The CWT is described by:

$$CWT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\phi^{*}\left(\frac{t-b}{a}\right) dt \tag{1.6}$$

where ϕ^* is the complex conjugate of the mother wavelet ϕ (a function satisfying certain criteria [1]), a is the scale parameter corresponding to a pseudo-frequency, and b is the translation parameter. Common choices for ϕ include the Debauchies wavelet and the Morlet wavelet [1, 2].

Eqn. 1.6 is a convolution calculation, however when a and b are considered fixed, Eqn. 1.6 is the inner product of the time series with the mother wavelet. This leads to an intuitive interpretation; high values of the CWT at time b indicate a large frequency content corresponding to a. The author is aware of only one method to extract features with the CWT: calculate all desired values of $CWT(a, b)$ and represent them as a column vector [1, 100]. For example, if an AE signal has 1000 time domain samples, $CWT(a, b)$ can be calculated for N values of frequency, resulting in a feature vector whose length is $1000 \times N$.

Wavelet Packet Transform

An alternate method for interrogating non-stationary signals is the wavelet packet transform (WPT)[2, 101–103]. The WPT is a set of decompositions of the original signal, $x_{(0,0)}$. At each decomposition iteration, two down-sampled signals containing the originals low frequency (approximation) and high frequency (detail) components are retrieved. The decomposition process is applied all sub-signals until a specified decomposition level is reached. This is shown schematically in Figure 1.4c.

A signal decomposed to level j will produce 2^j waveforms, $x_{(j,i)}, i \in [1, 2, ..., 2^j]$ with non-overlapping frequencies in each sub-signal [102, 104]. The relative energies (the portion of the original signals energy contained within each down-sampled signal at the final decomposition level) are then extracted and used as features [2, 102, 104].

Hilbert-Huang Transform

The Hilbert-Huang transform (HHT) is a data-adaptive method for extracting instantaneous frequencies in the time-frequency domain that is well-suited for non-stationary signals [3, 93, 95]. An HHT analysis consists of two steps: (i) empirical mode decomposition (EMD) of a signal into a set of intrinsic mode functions (IMFs), and (ii) a Hilbert transformation of the IMFs.

In a well-behaved time series $x(t)$ (symmetric about 0 with no superimposed waves), the Hilbert transform $y(t)$ is defined as [93, 95]:

$$y(t) = \frac{1}{\pi} PV \left[\int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \right] \tag{1.7}$$

where PV is the Cauchy principal value of the improper integral. From this, the analytic function $z(t)$ is defined:

$$z(t) = x(t) + iy(y) \tag{1.8}$$

This has a natural phasor representation:

$$z(t) = a(t)e^{i\theta(t)} \tag{1.9}$$

where $a(t)$ is the instantaneous amplitude of $x(t)$, and $\theta(t)$ is the instantaneous phase of $x(t)$.

$$a(t) = |z(t)| = (x^2(t) + y^2(t))^{1/2} \tag{1.10}$$

$$\theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right) \tag{1.11}$$

The instantaneous frequency $\omega(t)$ is then calculated:

$$\omega(t) = \frac{d\theta}{dt} \tag{1.12}$$

In practice, $x(t)$ is unlikely to be well behaved, however EMD ensures that each IMF, c_i, is. By EMD, the frequency content of the first IMF is the highest and decreases monotonically with index i. For AE data, the lowest order IMFs isolate frequency information from the signal, thereby expelling statistical noise. An analogous approach is ensemble empirical mode decomposition (EEMD), which provides a superior level of robustness to EMD and should be used when computationally affordable [93]. The readers are referred to Huang et al. [93] for details of executing E/EMD; additionally, there are available MATLAB[104] and Python[105, 106] packages that readily perform both E/EMD and the Hilbert transform.

After EMD, the Hilbert transform is applied to each IMF. For a given IMF, we can define:

$$H_i(t, \omega_i(t)) = a_i(t) \tag{1.13}$$

which assigns a time value and its related instantaneous frequency to the instantaneous amplitude of c_i. The sum of all H_i is the Hilbert spectrum, which represents the extracted instantaneous frequency and amplitude from all c_i [95].

An alternative visualization of the Hilbert spectrum is the marginal Hilbert spectrum, defined as:

$$h(\omega) = \int_0^T H(t, \omega) dt \tag{1.14}$$

where T is the length of signal $x(t)$ (Figure 1.4d). The marginal Hilbert spectrum is the probability of finding an instantaneous frequency at any given point in the signal, which Huang et al. differentiated from the FFT, which represents the total energy persisting from a frequency throughout the signal [95].

While the HHT has the potential to reveal hidden insights in the data, its applicability for feature extraction is uncertain [3, 107–109]. Though previous work by Hamdi et al. suggested that HHT features could be used for unsupervised ML, their findings are inconclusive[3]. Their input data included labeled data from other studies [7, 110] that employed ML techniques unverified with experiments, as a basis for labeling their own data sets. In general, HHT properties are not as well characterized as the FFT, CWT, or WPT. For this reason, it has not as been as widely adopted.

1.3.2 Metrics and Transformations

Once features are extracted, a metric, or similarity function, $d(*, *)$, is used to define the degree of similarity between observations. In general, d must satisfy:

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \Leftrightarrow x = y$
3. $d(x, y) = d(y, x)$

4. $d(x,y) \leq d(x,z) + d(z,y)$

however some clustering algorithms, such as those using linkage-based distances or those working on local neighborhoods, relax requirement 4 [111–113]. It should be noted that the efficacy of a given metric is tied to the geometry of the input data, which manifests itself in the quality of the resulting partition [78]. For example, neighborhood-based distances are better suited for grouping nested circles than Euclidean distance [111].

While there is flexibility in choosing a metric, an alternative strategy is to pre-process feature vectors by a transformation [4, 78, 79]. In general, many transformations are available, but the business of choosing an appropriate one is tricky because they alter the geometry of the input data. While this is desired, a complete understanding of the transformation's effect on the input data is contingent on knowledge of the starting geometry. This is problematic, as data living in a high dimensional space and can never be viewed in its natural setting. As a result, the choice of transformation is typically empirical, with the final decision being justified via a (semi-)exhaustive searches over common methods.

Principal component analysis (PCA) among the most commonly employed transformations, whereby a linear and orthogonal change of basis is carried out such that axes of the new coordinate system are oriented along the directions of maximum variance. PCA transforms the original ordered basis $(\vec{e_1}, \vec{e_2}, ..., \vec{e_d})$ to a new basis, $(\vec{e_1'}, \vec{e_2'}, ..., \vec{e_d'})$, with the property that $\vec{e_1'}$ is oriented along the direction of greatest variance, $\vec{e_2'}$ is oriented along the direction of second greatest variance, and so on, with a diagonal covariance matrix (i.e. no correlation).

The usefulness of the PCA transformation is two-fold. First, by dropping the last $d - q$ basis vectors, data can be represented in a lower dimensional space with minimal reconstruction error. It is worth noting this projection could obscure clusters and is not recommended, as the accuracy of k-means is not strongly tied to the input dimension [5]. Second, PCA acts to decorrelate and rescale data [1, 56, 114, 115]. Since features with large correlation and variance tend to dominate clustering results, decorrelation and rescaling, also called PCA whitening or sphering, often increases performance of clustering methods [1, 78, 116]. Major programming languages offer packages to execute both PCA and whitening [104, 106, 117]. The reader is referred to MacGregor et al.[118] for further details.

1.3.3 Clustering Algorithms

After pre-processing, a clustering algorithm must be chosen that should complement the new geometry of the processed data. It is stressed that if an informed decision is not made at this juncture, an accurate partition cannot be achieved, even if the initial data is properly represented.

K-means and Variants

K-means is among the earliest clustering techniques and has positioned itself as the default clustering method for AE applications [80]. In order to leverage the strengths and characterize the efficacy of AE frameworks commonly used, it is necessary to understand how clustering algorithms partition waveforms. To this end, this section will detail k-means and Gaussian mixture models.

The k-means algorithm seeks to partition a set of observations into k clusters by (Figure 1.5a):

1. Initializing k centroids $\mu_j \in \{1, k\}$

2. Attaching each feature vector, $\vec{x}_i$, to the closest centroid, μ_j

3. Moving each centroid, μ_j to the center of mass of the data points to which it is attached
4. Repeating steps 2 and 3 until a user-defined stopping criterion is reached

Stepping through this algorithm is equivalent to minimizing the L_2 loss function:

$$L_2 = \sum_{j=1}^{k} \sum_{i} |\vec{\mu}_j - \vec{x}_i|^2 \tag{1.15}$$

via gradient descent. For this reason, k-means tends to fall into local minima which are dependent on how cluster centroids are initialized. [119].

In practice, this means that for any given run of k-means there is likely a better partition to be found (Figure 1.5b). To avoid falling into such local minima, variants of k-means modify the initialization scheme [7, 84, 120]. Among those, the maxmin and k-means++ variants are the most robust [5, 106, 119], as both initialize well dispersed seeds. Maxmin selects a point for the first cluster center, and at each subsequent step, the following centroid is calculated as the farthest distance from the nearest centroid. K-means++ selects subsequent seeds from a statistical distribution with farther distances being preferred [119]. Sibil et al. developed an approach using a genetic algorithm, where multiple runs of k-means are initialized, and the best results are cross-bred according to a set of user-defined rules [120]. While this genetic algorithm allows k-means to find an optimal solution with fewer restarts, appreciable gains in runtime or accuracy are not achieved by its use.

An alternative to modifying the initialization scheme is to alter how centroids move after initialization. One such approach is fuzzy c-means (FCM), in which a feature vector (x_j) is assigned a membership grade (u_i) to cluster i such that $u_i(x_j) \in [0, 1]$. FCM updates centroids according to weighted (rather than absolute) membership, which reduces the tendency to fall into local minima [78]. Despite these benefits, it is unclear if FCM is superior to k-means [110, 121–125] due to difficulties in verifying cluster results. Ultimately, the user can opt for another strategy altogether: restarting k-means multiple times and selecting the best run. When the number of clusters is low and data is poorly separated, as is often the case with AE data, this strategy yields acceptable results provided that the number of restarts is sufficiently large (>1000) [5]. Given that clustering via k-means tends to have short runtimes, this is a feasible strategy.

However, none of the aforementioned techniques address the issue of data geometry. The k-means algorithm finds spherical clusters and can therefore only cluster convex data [4]. However, AE data does not necessarily abide by this constraint and thus k-means should be the default choice. To illustrate this, two commonly used features are extracted from waveforms gathered in Chapter 4 (Figure 1.6). When plotted as a function of frequency centroid and partial power, these waveforms adopt a circular, non-convex geometry and as a result, distance-based clustering methods, such as k-means, will fail to accurately partition these waveforms by the angle of incidence. Therefore, alternative algorithms must be considered that are suited for larger ranges of data geometries at similar computational costs.

Gaussian Mixture Model

One such viable alternative to k-means is the Gaussian mixture model (GMMs)[92]. In contrast to k-means, GMMs find Gaussian distributions that model the input data and thus are well-suited to clustering data sampled from Gaussian distributions. Moreover, they provide greater flexibility for clustering anisotropic data, or data with different scales (Figure 1.5c).

GMMs model input data as belonging to a set of d-dimensional Gaussian distributions and search for the set of related parameters [77, 126]. The probability of finding feature vector $\vec{x}_j \in X$ where X is the set of feature vectors sampled is:

$$p(\vec{x}_j|\lambda) = \sum_{i=1}^{M} w_i g_i(\vec{x}_j|\vec{\mu}_i, \Sigma_i) \tag{1.16}$$

where M is the number of component Gaussians, and w_i is the weight associated with the multivariate Gaussian distribution:

$$g_i(\vec{x}|\vec{\mu}_i, \Sigma_i) = \frac{1}{(2\pi^{d/2})|\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x}_j - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x}_j - \vec{\mu}_i)\right) \tag{1.17}$$

parameterized by covariance matrix Σ_i and mean $\vec{\mu}_i$. The weights are normalized such that their sum is unity. The Gaussian mixture is then completely described by the set of parameters:

$$\lambda = \{w_i, \vec{\mu}_i, \Sigma_i\}, \quad i \in \{1, 2, ..., M\} \tag{1.18}$$

The GMM searches for λ which maximizes the likelihood of sampling X by way of the Expectation-Maximization (EM) algorithm [127]. Once the EM reaches a set stopping criterion, feature vectors are labeled according to the Gaussian that provides the highest likelihood of sampling it. Final GMM partitions are also sensitive to initialization, and therefore multiple restarts are recommended. For an in-depth discussion of GMMs, the reader is referred to Reynolds et al. [77].

1.3.4 Selection Clusters Number

Once a suitable clustering algorithm has been chosen, the number of clusters that are being searched for must be specified. In the context of damage mechanism identification, this is analogous to asking, ”how many unique damage modes are discernible?”. This is an open question which is the subject of Chapter 3.

To understand this question, it is worthwhile to revisit two hypotheses developed by the AE community: (i) an AE waveform has distinct features corresponding to the damage mode that emitted that waveform; and (ii) an individual AE waveform is generated by only one damage mode. If both (i) and (ii) are true, then follows that the number of clusters should correspond with the number of detectable damage modes. However, experimental variance can mask the true damage mechanism (further discussed in Chapter 1.4), and cause distinct clusters to overlap. Moreover, it is unclear if hypothesis (ii) holds. It has been experimentally shown *in-situ* that more than one damage mechanism can be active within the timeframe of a single AE event [36]. For example, in the case of a fiber-bridged matrix crack, the damage mechanisms of transverse matrix fracture, fiber debonding, and interfacial sliding all occur simultaneously and are captured by a single AE signal.

Ultimately, although the number of clusters will be between 1 and the number of unique damage modes, the true number of clusters is unknown. In order to estimate the correct number of identifiable damage mechanisms, researchers employ heuristic functions. These measure intercluster/intracluster spread, and operate on the premise that the optimal number of clusters yields the highest intercluster separation and lowest intracluster variance [128–134]. The true number of clusters is estimated where the function is extremized.

The Silhouette Value (SV) is a popular heuristic function[130]. It is a staple in unsupervised clustering because of its ease of use and natural graphical representation. For a feature vector $\vec{x}_k$, the SV is defined as:

$$s(\vec{x}_k) = \frac{b(\vec{x}_k) - a(\vec{x}_k)}{\max\{a(\vec{x}_k), b(\vec{x}_k)\}} \tag{1.19}$$

where $a(\vec{x}_k)$ is the average distance of $\vec{x}_k$ to all objects in its cluster, and $b(\vec{x}_k)$ is the minimum of the average distance from $\vec{x}_k$ to all objects in a different cluster. While defined for a single feature vector, the SV of a partition is the average SV over all feature vectors; this is the value used to estimate the number of clusters. The SV is bounded such that $SV \in [0, 1]$, with 1 corresponding to perfectly compact clusters. The SV is applicable to evaluate clusters independent of their geometry, but it is best suited for convex and spherical data.

When non-convex data is clustered, other heuristic functions should be considered. In GMMs, both the Akaike information criterion (AIC) and Bayesian information criterion (BIC) are established model selection techniques [106]. It has been shown that AIC generally overestimates the true number of mixture components, whereas there is considerable evidence that BIC consistently selects the correct number[135].

The author stresses that heuristic functions are estimates. When possible multiple heuristic functions should be used, and their results compared. Moreover, researchers cannot *only* rely on heuristic functions to evaluate clustering success; it is possible to obtain compact and well separated clusters whose membership does not reflect damage modes. To ensure that this is the case, investigators should compare their results to micromechanical expectations and justify their choices with secondary experimental evidence, as discussed in Chapter 1.4.

1.3.5 Manifold Learning Techniques

In addition to quantitative assessments of partition quality, such as those described in Chapter 1.3.4, it is often useful to visualize data with manifold learning techniques. These are a set of methods with act to map a high-dimensional object, e.g. waveform, to a lower dimensional space so that it may be visualized.

Self-Organizing Maps

The self-organizing map (SOM) is a type of neural network used for dimension reduction and data visualization [136]. SOMs maintain their topology, meaning that nearby points in the reduced representation correspond to similar feature vectors. Unlike PCA, SOMs are non-linear and thereby better suited to visualize a wider variety of data geometries.

The SOM consists of a grid G of $l \times l$ nodes, h_{ij}. The grid architecture describes node connections and dictates which nodes move when the SOM is learning the structure of the input data. Finally, each node has an associated weight vector, $\vec{w_{ij}}$. This can be considered the node's position within the high-dimensional feature space. The SOM learns the structure of the input data as follows:

1. Weights $\vec{w}_{ij}$ are randomly initialized
2. A random observation, $\vec{x}_k$ is selected and its distance from the high-dimensional embedding of node ij is calculated as:

$$d_{ij} = |\vec{x}_k - \vec{w}_{ij}|^2 \tag{1.20}$$

3. Step 2 is repeated for all nodes, and the closest node is identified by:

$$h^*_{ij} = \min_{(i,j)\in G} d_{ij} \tag{1.21}$$

4. The weight of node h^*_{ij} is updated according to:

$$\vec{w}^*_{ij}(t+1) = \vec{w}^*_{ij}(t) + \eta(t)\left[\vec{x_k}(t) - \vec{w}^*_{ij}(t)\right] \tag{1.22}$$

where t is the time step and η is the learning gain factor which starts at 1 and monotonically decreases with t

5. The weights of nodes in the neighborhood of h^* are updated according to Step 4, and t is incremented

6. Steps 2-5 are repeated until all data points have been selected at least once

Once completed, the high-dimensional data can be visualized by assigning a grey value to the 2D grid of the SOM according to the maximum distance of a node from its neighbors (Figure 1.7a)[137]. Thus, the SOM effectively has 2 representations: a high-dimensional representation and a 3D representation (2D grid + 1D grey values). The strength of the SOM is rooted in the fact that the high dimensional representation adapts to the structure of the feature vectors, while maintaining a meaningful 3D representation [136].

Previously, AE studies have used SOMs to pre-process data by associating input data to the closest $\vec{w}_{ij}$ and clustering these weight vectors with k-means [7, 138, 139]. However, approximating points by their closest node can only reduce the amount of available information, thereby degrading the fidelity of k-means. For this reason, the author recommends that data be clustered and subsequently visualized with a SOM, rather than clustering SOM nodes directly.

T-Distributed Stochastic Neighbor Embedding

Another method for visualizing high-dimensional data is t-distributed stochastic neighbor embedding (t-SNE), which is a manifold learning algorithm. T-SNE is powerful due to its ability to outperform other manifold learning techniques such as the SOM, Sammon mapping, Isomap, and local linear embedding [140–142] (Figure 1.7b), At the highest level, t-SNE maintains all pairwise distances between the high dimensional and low dimensional representation of feature vectors, x_i and y_i, respectively [9]. The fundamental assumption in t-SNE is that for given a feature vector, x_i, all other points follow a Gaussian distribution with standard deviation σ_i centered at x_i. The conditional probability of finding another feature vector x_j is then:

$$p_{j|i} = \frac{\exp(-|x_i - x_j|^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-|x_i - x_j|^2/2\sigma_i^2)} \tag{1.23}$$

There does not exist an optimal value for σ_i that describes all data points. Instead, this is estimated with a hyperparameter termed "perplexity," which is a measure of how much of the local structure is retained in the final low-dimensional map. As perplexity increases, local structure information is exchanged for global structure information [9]. Pairwise probabilities in the high dimension, $p_{j|i}$, are translated to pairwise probabilities in the low-dimension, $q_{j|i}$. If the low dimensional representation has correctly maintained the same pairwise distances as the high dimensional representation, then $p_{j|i} = q_{j|i}$ for all pairs i, j. In contrast to the high dimensional representation, where similarities are calculated according to a Gaussian distribution, pairwise similarities in low dimension are calculated according to a student t-distribution with a single degree of freedom. This alleviates the crowding problem in traditional SNE while permitting for tractable runtimes. T-SNE is described in greater detail in van der Maaten and Hinton [9], and a practical guide is provided by Wattenberg et al. [143]. It should be noted that despite its benefits over other manifold learning techniques, t-SNE suffers from long runtimes when large ($n > 20,000$) datasets are considered. When this is the case, the reader is referred to the recently created technique, UMAP, which has been shown to outperform t-SNE in runtime with competitive visualization performance [144].

1.4 Towards Improved AE-ML Frameworks

Many investigators have leveraged ML tools to infer damage mechanisms from AE. However, there is little overlap between experimental systems and analysis methodologies. This introduces difficulties for isolating how experimental factors, such as the effect of sensor coupling, impact the ability to fingerprint AE. Despite this lack of standardization, there is consensus on the following: (i) damage modes are distinguishable by the frequencies of their AE signatures, and (ii) robust methods for verifying results are needed. This section presents a survey of previous studies focusing on these two points. Advantages and limitations are discussed, issues that must be addressed moving forward are highlighted.

Frequency-based Fingerprinting

First, there is overwhelming agreement that damage modes can be fingerprinted from frequency-based features. There is significant evidence from both experiment [44, 45, 47, 55, 145–147] and modeling [67, 68, 71, 114, 148] indicating that damage mechanisms exhibit distinguishable differences in frequencies. Subsequently, the most suitable representations of waveforms will use methods that encode frequency information such as partial powers, CWT coefficients, or HHT spectra. The optimal representation scheme is still in dispute, as it is not known how these features are impacted by experimental configuration [72, 149, 150]. As an example, changes to the sample geometry affect waveform features, even when the dominant damage mode has not changed [68, 90, 151, 151]. Ospitia et al. demonstrated that AE generated from pencil lead breaks, and material damage, exhibited lower frequencies in plate geometries than to beam geometries [151]. This difference was attributed to the unrestricted crack extension in the beam geometry compared with incremental extensions in the plate geometry.

Moreover, variations in propagation pathways distort acoustic waveforms, resulting from a single event recorded at different sensors [65, 68, 90, 151]. Hamstad showed, via FEA calculations, that excitations will produce lower frequencies corresponding to the flexural wave mode when the excitation is farther away from the midplane of the plate [65]; this has been experimentally verified [149]. Maillet et al. experimentally demonstrated that path lengths shift frequency centroids to lower values for larger source-to-sensor distances [90]. Thus, an event not emitted from the center of a specimen will have different frequency content depending on which sensor is used to analyze the wave. Finally, the sensor contact is also known to influence the frequency spectra of a signal[152–154]. Theobald et al. showed that the acoustic impedance matching between a couplant and the sensor significantly influences the recorded frequency spectra [153]. Future AE studies aiming to study point-by-point AE information across specimens must standardize methods for ensuring experimentally consistent sensor contact, specimen geometry, and loading conditions.

Robust Verification

The verification of ML results is a closed loop problem. The optimal feature representation and ML algorithm are not known *a priori*, motivating the need for manual verification which in turn requires *a priori* knowledge of the optimal feature representation. Breaking this loop necessitates strategies that bring outside information, such as existing domain knowledge or secondary corroborating experimentation. The simplest of these approaches is the use of existing domain knowledge to correlate clusters to damage modes by inspecting average frequency values of a cluster, or the cluster activity [1, 3, 57, 85, 86, 97, 110, 125, 155].

When the average frequency characteristics of a cluster are used to verify cluster membership, conclusions should be treated with caution as there is little accounting for the large experimental variance of such characteristics. For example, singleton frequency values (e.g. peak, average, or centroid) are influenced by factors such as source-to-sensor distance and the state of damage accumulation; both affect waveform and frequency attenuation in unknown ways, and therefore reduce the robustness of verification [72, 115, 149, 150]. When cluster activity is considered, there is often considerable overlap in ways that are not physically expected. For example, Lyu et al. concluded that all modes of damage in 0/90 woven C_f/SiC CMCs, including fiber bundle failure, are active from the beginning of a creep test [156]. It is unclear if this is a real phenomenon, or if fewer damage modes are active and are instead being incorrectly distributed across all clusters. Few studies show distinct differences in chronological cluster activation [100, 155, 157, 158]. Similarly, more complex fiber architectures often show simultaneous cluster activation, however, these findings have not been verified experimentally [56, 85, 159, 160]. When corroborating experimentation is used to break the loop, high-resolution postmortem techniques (namely, microscopy) are popular for justifying cluster labels [84, 88, 139, 161–164]. However, these postmortem methods are insufficient for informing on damage chronology and can only provide information on what labels are available to be assigned.

More complex solutions involve the creation of AE libraries to serve as ground truth data sets [54, 87, 165, 166]. This is a particularly powerful approach because once ground truth sets have been established, there is no difficulty in the interpretation of final results and refinement of the learning procedure becomes possible. One method of library creation is to isolate 1-2 damage modes by using specialized sample geometries, and then leverage signal characteristics to label damage modes in more relevant geometries. For example, Godin et al. used AE from pure epoxy resin and single fiber microcomposites to label the unsupervised results from unidirectional composites [54]. Another method is to test full-scale samples in ways that promote one damage mechanism. Gutkin et al.

tested full-scale composite laminates under monotonic tension, compact tension, compact compression, double cantilever beam, and four-point end notch flexure tests [87].

The most promising library creation method involves testing specimens with known damage modes, and using secondary high-resolution *in-situ* methods to assign labels to individual signals. While such studies should elucidate the nature of the AE-mechanism relationship, they are highly non-trivial to conduct due to time and resource constraints. As a result, existing studies which attempt this lack one-to-one correlation between events and damage modes, limiting the usefulness of the library. For example, the library created by Tat et al. [166] is composed of a single tomographic scan in the period of interest, which does not meaningfully increase temporal resolution compared to a post-mortem image and is geometry-specific. Moreover, [166] used an open hole dogbone specimen geometry, which is known to alter waveform characteristics in a way such that it cannot be used to classify waveforms from other geometries [167].

The inability to use signals gathered from one experimental configuration to classify signals gathered from another experimental configuration points to a more general problem: there is no overlap between material systems, experimental configurations, and no framework which currently exists to address these differences [54, 100, 103, 120, 165, 168]. This prevents meaningful comparison of the efficacy of different representation schemes and choice of ML algorithms, and is the subject of the study presented in Chapter 4.

1.5 Objectives and Outline

Expanding the space in which ceramic matrix composites can be used hinges on the ability to accurately predict their damage progression. However, due to their heterogeneous microstructure, quantitative descriptions of damage are non-trivial to obtain. To this end, the primary focus of this dissertation is the development and utilization of novel trustworthy ML tools which can be used to obtain such quantitative descriptions of damage progression. Herein, the first framework for fingerprinting AE signals to identify their source damage mechanism in elastically similar composites is proposed and characterized. Guidelines for the development of trustworthy ML in the physical sciences are established, and perspective on the future of ML-based solutions in mechanics is offered.

In Chapter 2, I will first present an *in-situ* study of crack growth rates in SiC/SiC minicomposites, and show that current canonical models do not accurately predict crack growth rates or crack opening displacements. Results are contextualized in terms of microstructural heterogeneities, which are found to be the primary cause of discrepancy between model and measurement. Such heterogenities are non-trivial to capture in analytical and numerical models and motivate the need for high-fidelity damage mechanism identification from AE.

Chapter 3 presents a novel ML-based method for damage mechanism identification from AE in elastically similar composites via a spectral framework. AE waveforms were collected from SiC/SiC ceramic matrix minicomposites loaded under uniaxial tension. The signals were then encoded with a modified partial power scheme and subsequently clustered via spectral clustering. Matrix cracking and fiber failure were successfully identified based on the frequency information contained in the AE event they produced, despite the similar constituent elastic properties of the matrix and fiber. Importantly, the resultant identification of AE events closely followed CMC damage chronology, wherein early matrix cracking is later followed by fiber breaks, even though the approach is fully domain-knowledge agnostic. The presented approach is promising for CMCs and other composite systems with elastically similar constituents.

Following the development of the spectral framework, a broad need for characterization of AE-ML frameworks was identified. However, the lack of ground truth datasets and the limited overlap between experimental configurations has thus far precluded any direct, quantitative benchmarking of their accuracy. In Chapter 4, a ground truth dataset is generated and used to benchmark our new ML framework and currently existing frameworks for AE identification. This dataset is comprised of pencil lead breaks with known angles of incidence, θ, where each angle generates a unique frequency spectrum that changes continuously with θ; this can be considered analogous to attributes of AE signals generated from failure processes, such as those that occur in CMCs. Five frameworks from the literature, all using various clustering approaches, are benchmarked according to their ability to discriminate between two sets of signals with a fixed $\Delta\theta$. Their performance as related to choice of features is discussed, a set of best practices for feature selection is proposed, and guidelines to facilitate widespread use of AE-ML frameworks are provided.

While the work in Chapter 4 represents a first step in creating ground truth libraries of acoustic signals, it is likely that signals from the PLB library do not share characteristics which can be used for damage mechanism identification in CMCs. To address this shortcoming, Chapter 5 presents and discusses preliminary work targeted at creating a ground truth library of signals from CMCs. Here, SiC/SiC minicomposites are subjected to uniaxial tension and imaged via x-ray computed tomography. Broad damage mechanism trends are identified, allowing a training library primarily composed of matrix crack signals to be curated. An autoencoder architecture is then leveraged to perform anomaly detection to identify fiber break signals. It is found that fiber break signals have unusually low reconstruction errors, indicating a high degree of homogeneity within these signals. This behavior is contextualized by the composite microstructure, lends support to hypotheses developed in Chapter 3, and motivates choices for future AE-ML frameworks.

In Chapter 6, I discuss the implications of the research findings presented in this dissertation, provide perspective on how ML tools can be effectively leveraged to investigate the mechanical behavior of materials, and propose recommendations for future scientific investigations into the analysis of acoustic emission characterization of composites using machine learning.

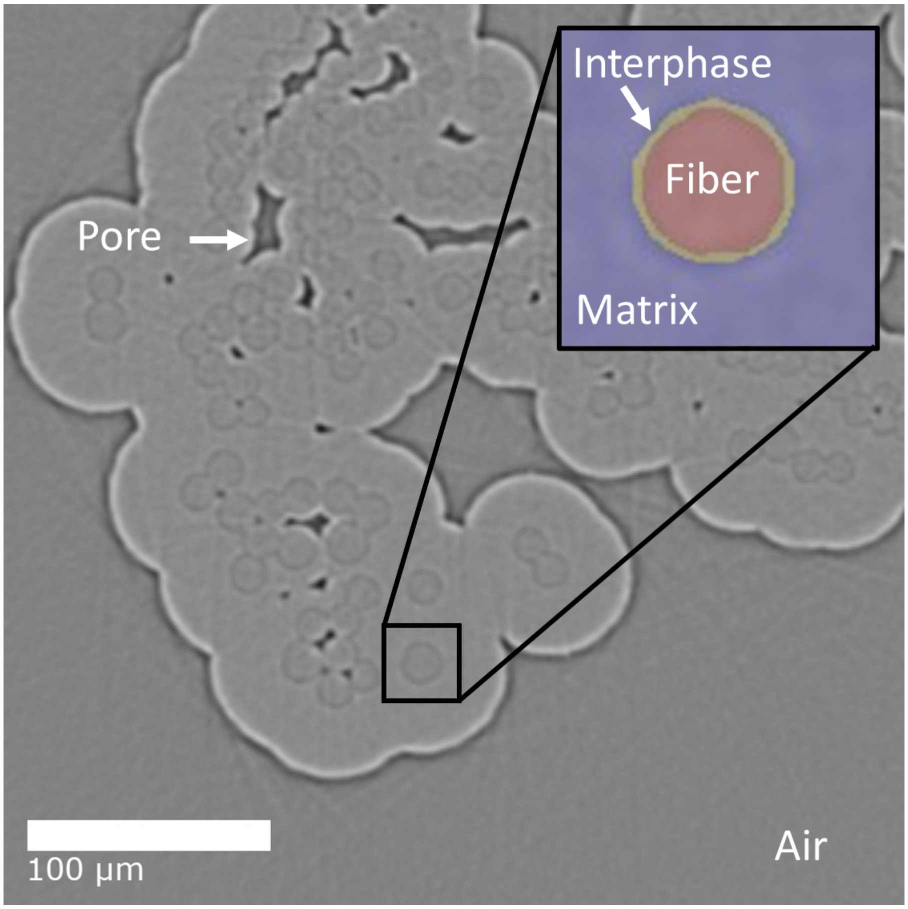

Figure 1.1: **Cross-section of a SiC/BN/SiC minicomposite.** SiC fibers are coated in a 300 ± 80 nm BN interphase are surrounded by a SiC matrix. The inset has been artificially colored to highlight the fiber, interphase, and matrix constituents.

Table 1.1: **Common AE time-domain, frequency-domain, and composite features**

ID	**Feature**	**Units**	**Description**
1	Duration[85, 120, 155]	μs	Time difference between first and last threshold crossing
2	Amplitude[85, 120]	dB, V	Maximum signal amplitude
3	Energy[85, 120]	attoJ	Area under signal envelope or $\int x(t)^2dt$, depending on software
4	Zero crossing rate [103, 159]	%	-
5	Zero crossings [169]	-	-
6	Rise time[85, 120]	μs	Time from start of signal to maximum amplitude
7	Decay time [159]	μs	Time from maximum amplitude to end of signal
8	Temporal centroid	μs	$\int tx(t)dt/\int tdt$
9	Entropy [159]	-	-
10	Counts[85, 155]	-	Number of signal peaks above a threshold
11	Counts to peak[85, 120]	-	Counts from signal onset to maximum amplitude
12	Rise angle (RA) [155, 169]	-	$\tan^{-1}$(Amplitude/Rise time)
13	Decay angle (DA) [155, 169]	-	$\tan^{-1}$(Amplitude/Decay time)
14	RA value [91, 92]	V/μs	Rise Time/Amplitude
15	Rise time/Duration [159]	-	-
16	Duration/Amplitude [159]	μs	-
17	Energy/Amplitude [159]	attoJ/V	-
18	Non-dimensional amplitude [169]	-	Maximum amplitude/mean amplitude
19	Counts/duration[85]	-	-
20	Partial Power [49, 169]	%	$\int_a^b F[x(t)](k)dk/\int_0^{f_N} F[x(t)](k)dk, \quad a, b \in [0, f_N]$.
21	Average frequency [92, 155]	kHz	$\int F[x](k)dk/\int kdk$. This is the mean frequency.
22	Frequency centroid, $f_{centroid}$ [85]	kHz	$\int kF[x](k)dk/\int kdk$. The center of gravity of the frequency spectrum.
23	Peak frequency, f_{peak}[85]	kHz	Frequency with maximum FFT power
24	Spectral spread[159]	kHz	Standard deviation of $F[x(t)]$
25	Spectral skewness [159]	-	-
26	Spectral kurtosis [159]	-	-
27	Roll-on frequency [159]	kHz	Frequency at which 10% of total FFT power has been accumulated
28	Roll-off frequency[159]	kHz	Frequency at which 90% of total FFT power has been accumulated
29	Initiation Frequency [49, 120]	kHz	Average frequency from signal onset to maximum amplitude
30	Reverberation frequency [120, 169]	kHz	Average frequency from maximum amplitude to end of signal
31	Weighted peak frequency [49]	kHz	$\sqrt{f_{centroid}f_{peak}}$

$x(t)$ is the signal, $F[*]$ is the FFT, and f_N is the Nyquist frequency.

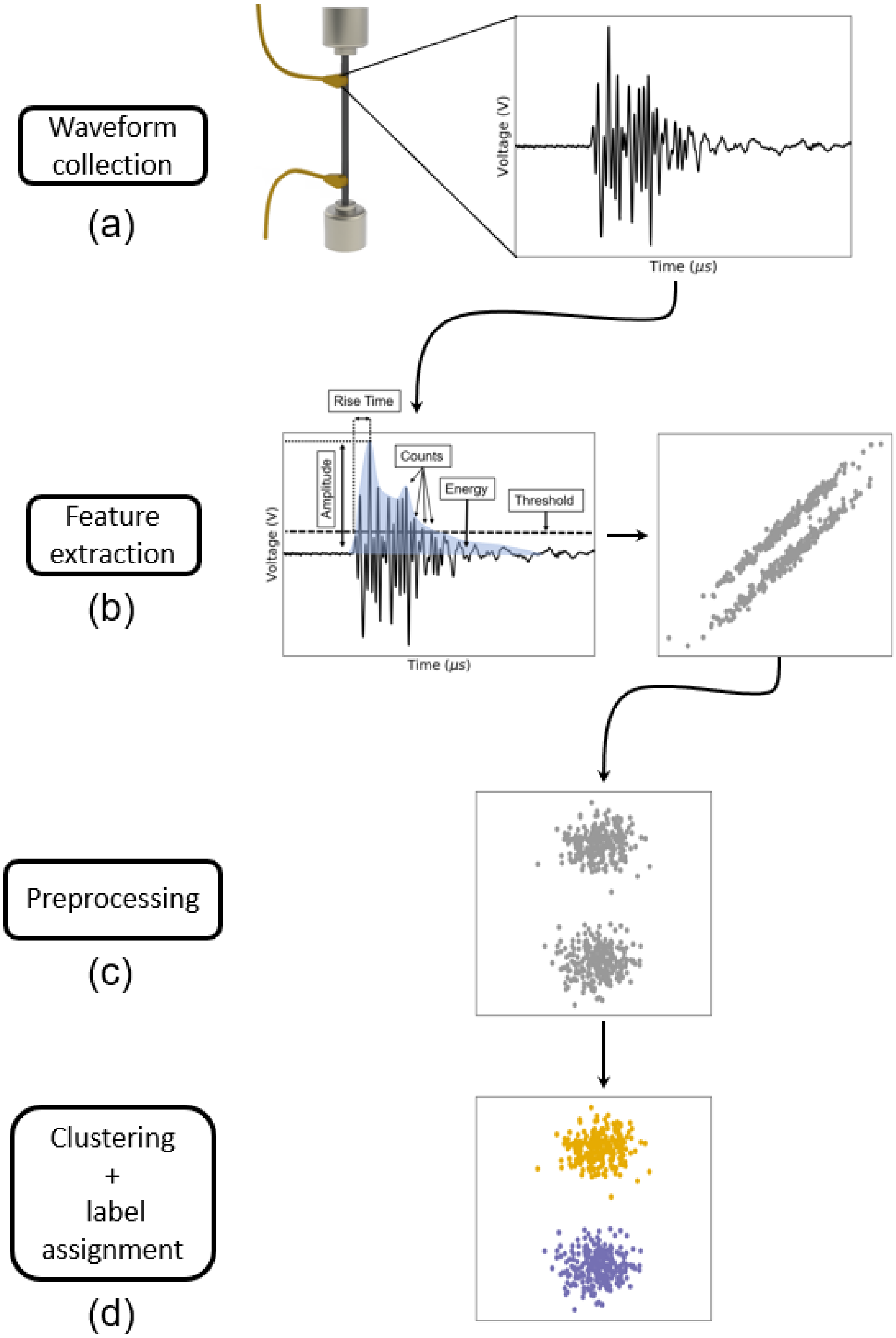

Figure 1.2: **Workflow diagram of an AE-ML framework.** (a) Waveforms are collected and (b) pertinent features are extracted from the waveforms, which are then represented as vectors in feature space. (c) Feature vectors can then be re-scaled and/or re-mapped before (d) the clustering algorithm is applied and feature vectors are labeled. Every AE-ML framework follows this procedure.

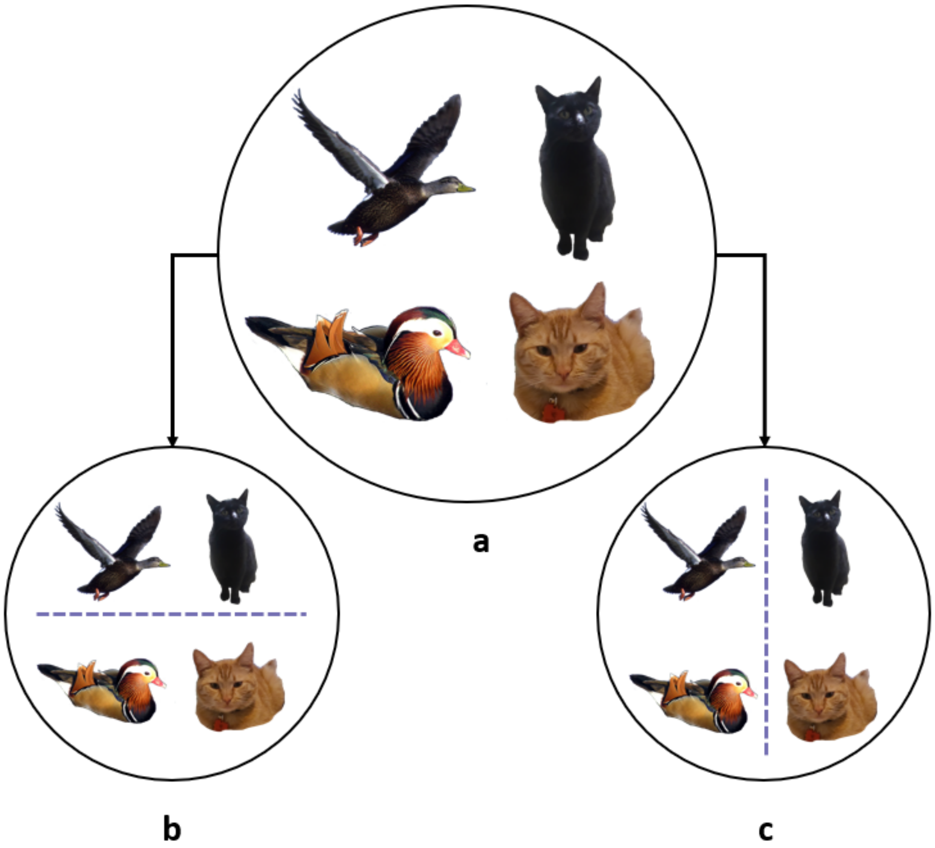

Figure 1.3: **Feature selection determines the quality of clusters.** (a) The initial group of objects is partitioned in (b) based on color and in (c) based on flight capability. The researcher must tailor the feature set to achieve the desired partition.

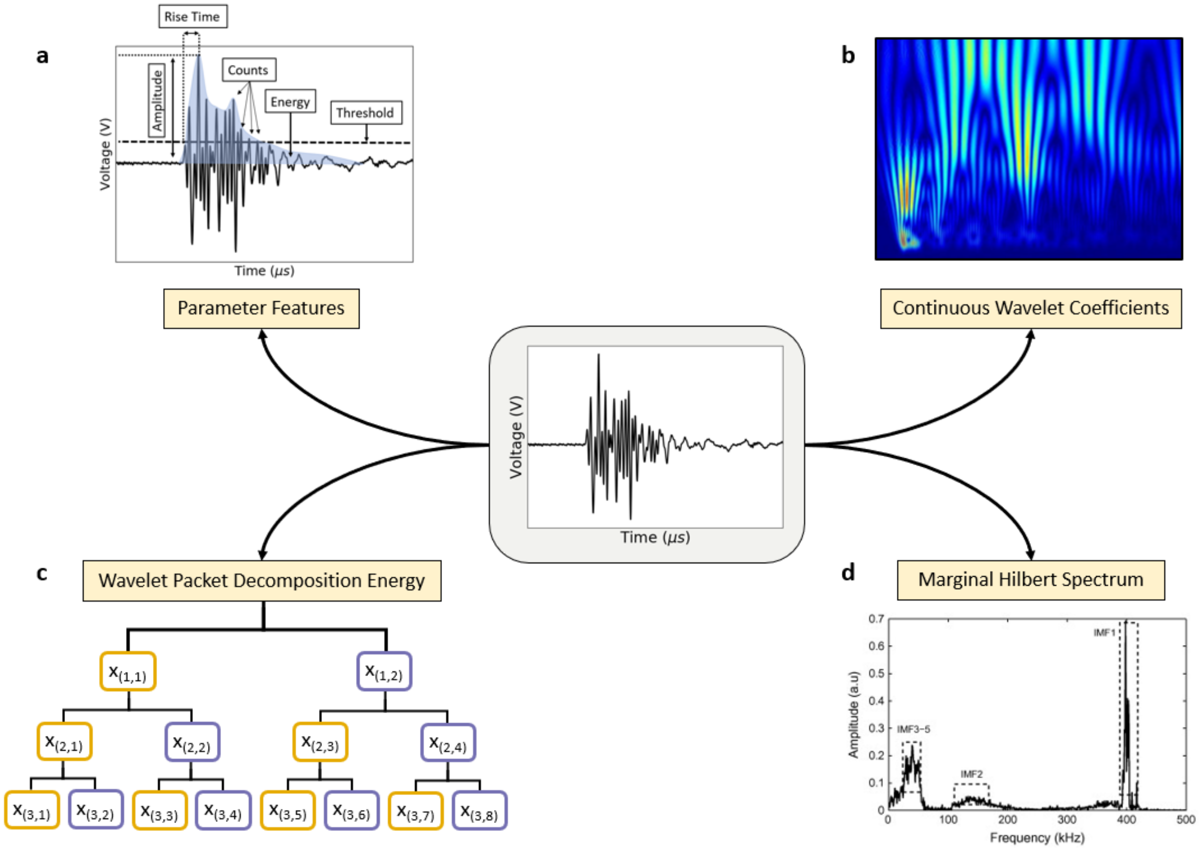

Figure 1.4: **Common feature extraction methods.** (a) Parameter features are readily extracted by means of commercial or in-house software.(b) The continuous wavelet transform is applied to the signal using Eqn. 1.6. CWT coefficients are then used as feature vector entries. (c) Wavelet packet decomposition down-samples the original signal to 2^j sub-signals containing non-overlapping frequency content. Energy contained in each of these subsignals is used as feature vector entries. (d) The marginal Hilbert spectrum is calculated and the frequency centroid of each IMF is used as feature vector entries. Panel **b** is adapted from Ref. [1] with permission from Elsevier. Panel **c** is adapted from Ref. [2] with permission from Elsevier. Panel **d** is adapted from Ref. [3] with permission from Elsevier.

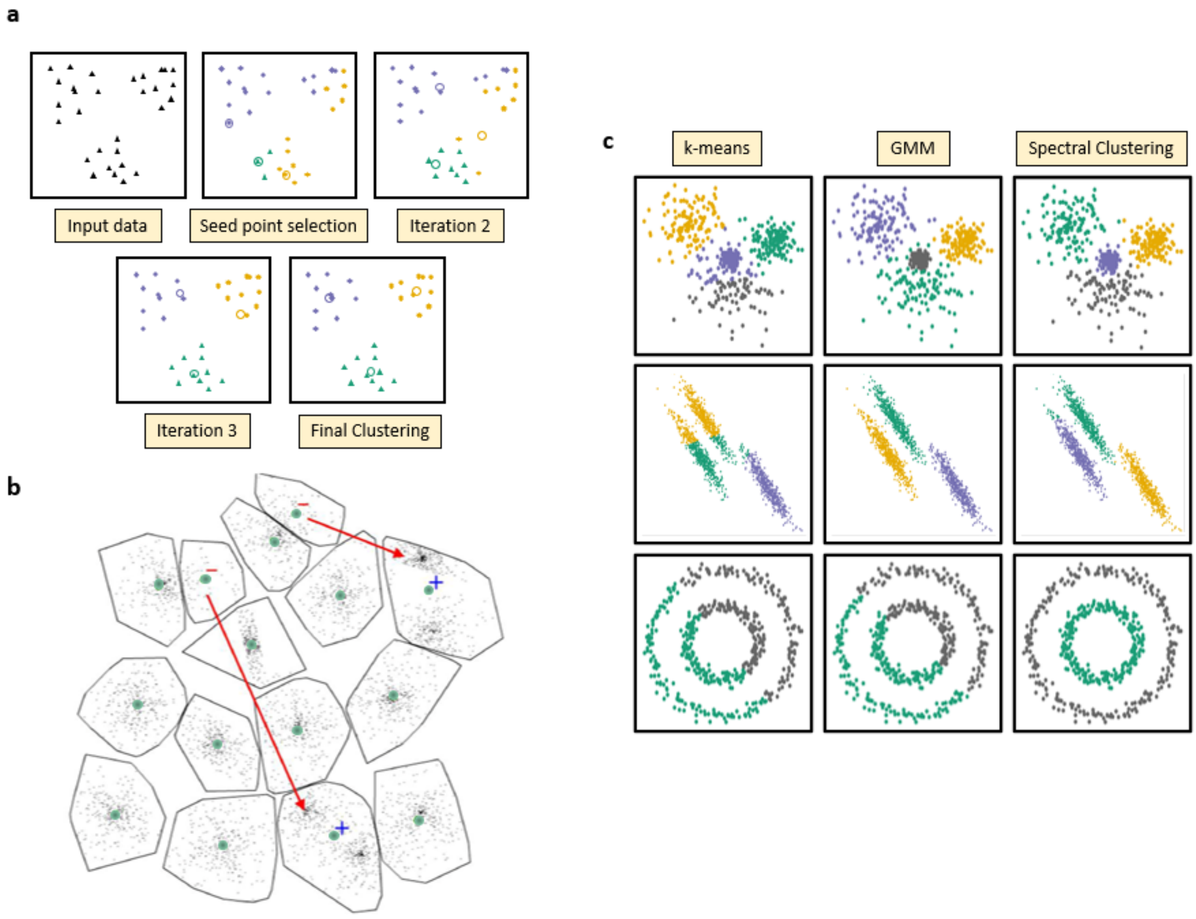

Figure 1.5: **A visual guide to k-means and other ML algorthms.** (a) When clustering data with k-means, seed points are selected to represent the initial data, which are then labeled according to the closest seed point. The seeds are then moved to the centroid of the labeled data, and points are re-labeled until a user-defined stopping criterion is reached. (b) K-means often falls into local minima, resulting in sub-optimal partitions. An initialization with 15 centroids is shown, where two centroids have fallen into a local minima (denoted with minuses) and are unable to move to their optimal position (denoted with pluses). (c) The performance of k-means, GMM, and spectral clustering is demonstrated on a Gaussian dataset (row 1), an anisotropic dataset (row 2), and a concentric circle dataset (row 3). The performance of an ML algorithm is dependent on the data scale and geometry. Panel (a) is adapted from Ref. [4] with permission from Elsevier. Panel (b) is adapted from Ref. [5] under creative commons license. Panel (c) is generated from the sci-kit learn toolbox [6].

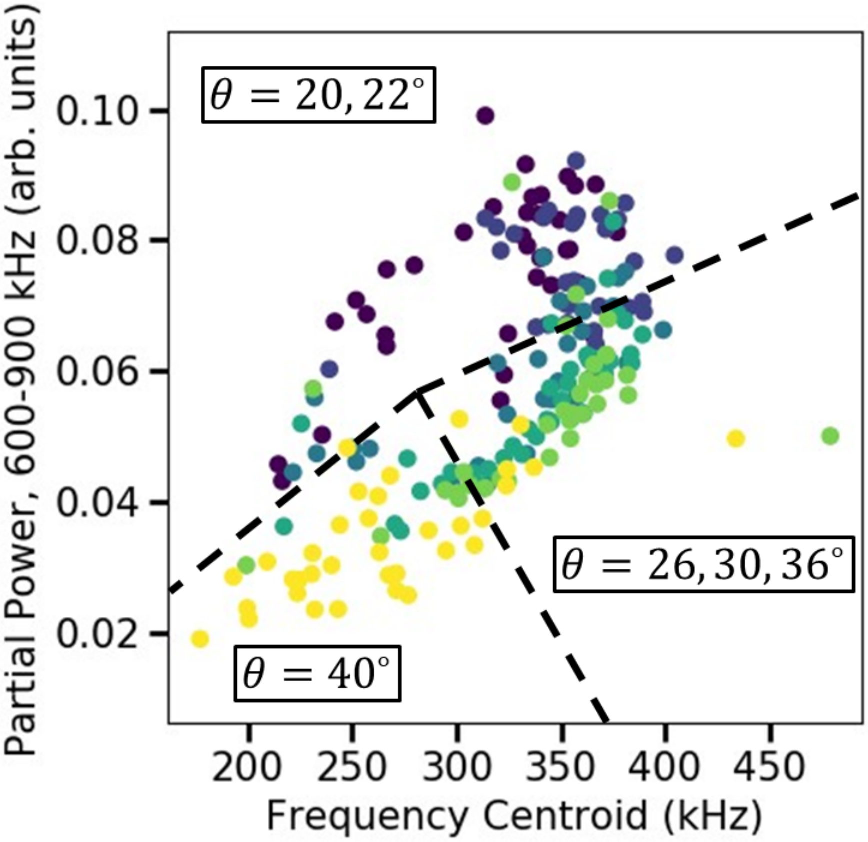

Figure 1.6: **AE waveforms generated in Chapter 4 adopt a non-convex geometry when plotted in feature space.** Waveforms are plotted as a function of frequency centroid, partial power, and colored by the angle of incidence. In this feature space waveforms adopt a circular, non-convex geometry. Unique angles of incidence occupy unique arcs on this circle, denoted by dashed lines.

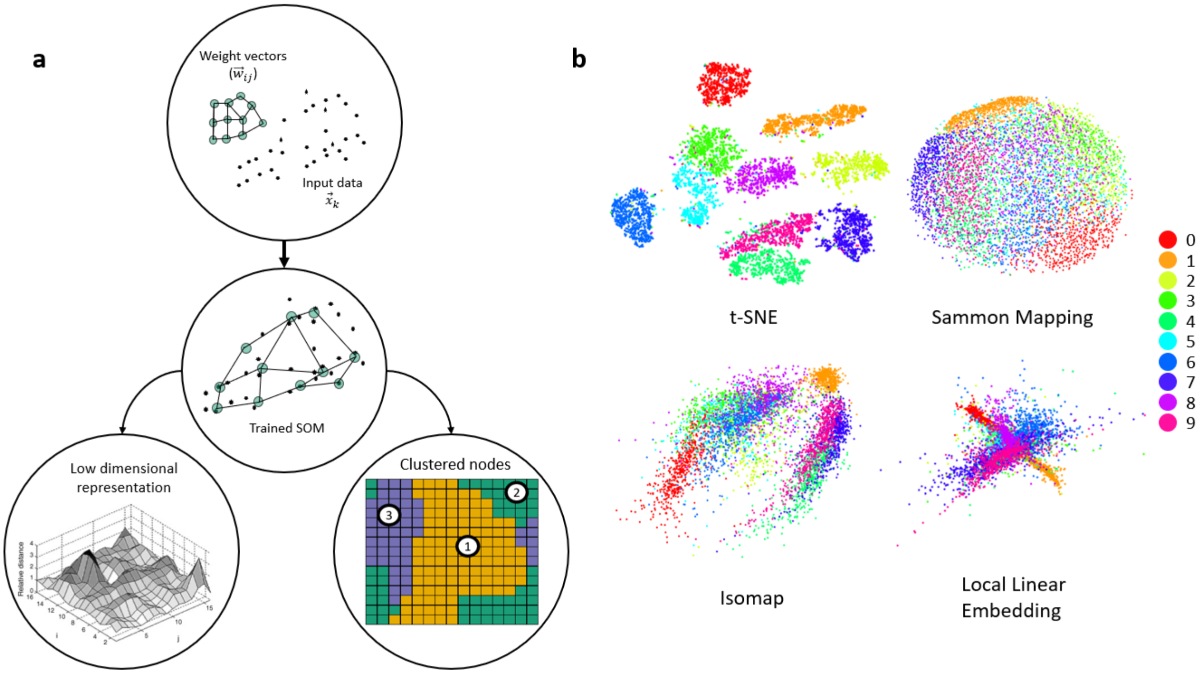

Figure 1.7: **Roadmap and comparison of data visualization techniques.** (a) The weight vectors of the SOM are initialized and then adapt to the structure of the input data in the high-dimensional feature space. The weight vectors are then represented by a grid in 2D, where each point on the grid is either assigned a grey value or a label. (b) The same set of 6000 handwritten digits is visualized using 4 common manifold learning techniques: t-SNE, Sammon mapping, Isomap, and local linear embedding. Only t-SNE maintains distinct clusters in the plane, thereby outperforming the other data visualization techniques. Panel (a) is adapted from Ref. [7] with permission from Elsevier and from Ref.[8] with permission from the Journal of Acoustic Emission. Panel (b) is adapted from Ref. [9] with permission from the authors.

Chapter 2

In situ Crack Opening Displacement Growth Rates of SiC/SiC Ceramic Matrix Minicomposites

2.1 Scope

In this chapter, I will present an *in-situ* study of crack growth rates in SiC/SiC minicomposites. In CMCs, environmental degradation rates are heavily influenced by crack opening displacements (CODs) and as a result, strategies for extending CMC lifetimes are driven by understanding COD evolution. While many models describe COD growth, few experimental studies exist to validate their accuracy. The goals of this chapter are three-fold: 1) compare experimentally observed crack-growth to two canonical COD growth models, 2) contextualize deviations from these models in terms of microstructural heterogeneity, and 3) motivate the need for ML-based characterization tools.

To this end, four SiC/SiC minicomposites fabricated by two different sets of processing conditions were imaged as they were loaded under uniaxial tension in a scanning electron microscope (SEM). The COD of each crack was tracked as a function of stress, and compared to predictions from the Marshall model [10] and the Hutchinson and Jensen model [11]. It is experimentally demonstrated that $> 45\%$ of CODs can deviate from the classical σ^2 dependence typically proposed in these models. Evidence is provided for the hypothesis that such deviation is a result of crack geometry, proximity to neighboring cracks, and intact subsurface matrix content.

2.2 Materials, Methods, and Modeling

2.2.1 In-SEM Tensile Testing and COD Measurements

Two SiC/SiC minicomposite systems, fabricated by Rolls-Royce HTC, were incrementally loaded to failure in-SEM (MAIA3, Tescan, Kohoutovice, Czech Republic) using the method described in [33]. Images of the specimen surface were captured at each load increment and stitched to form gauge-length ($\approx$ 25 mm span) micrographs of surface damage accumulation. The minicomposite test specimens consisted of a Hi-Nicalon$^{\text{TM}}$ Type S (HNS) (NGS Advanced Fibers Co., Ltd., Toyama, Japan) 500 filament SiC fiber tow, a boron nitride (BN) interphase, and a chemical vapor infiltrated SiC matrix. Interphase thicknesses and matrix volume fractions varied between systems. As such, one system is referred to as low fiber content (LFC) due to its relatively low fiber and high matrix content compared to the other system, whereas the other system is referred to as high fiber content (HFC) (Figure 2.1).

For the LFC and HFC systems, the fiber volume fractions were nominally $22.8 \pm 1.0\%$ (LFC) and $34.1 \pm 1.3\%$ (HFC), and the interphase volume fractions were $\approx 2.5\%$ (LFC) and $\approx 25\%$ (HFC). In both cases, the interphase volume fraction was consolidated into the volume fraction of the matrix. This was done because when a rule-of-mixtures was applied between the interphase and the matrix, the COD predictions changed by less than 250 nm, an amount less than the resolution limit of the micrographs captured. Therefore in both systems, the matrix volume fraction V_m was defined in terms of the fiber volume fraction V_f as $V_m = 1 - V_f$. The interfacial shear stress τ and interfacial debond energy Γ_c were previously determined by fiber push-in tests [33]. In the LFC system, $\Gamma_c = 1.2 \pm 0.5$ Jm^{-2} and $\tau = 18.1 \pm 4.8$ MPa. In the HFC system $\Gamma_c = 5.5 \pm 3.9$ Jm^{-2} and $\tau = 34.5 \pm 13.0$ MPa.

Crack opening displacements (CODs) were measured with ImageJ for each visually identifiable through-thickness crack at each load state (Figure 2.2) [170]. For each crack, 20 measurements were taken at approximately equal spacing across the entire profile of the crack, at each load state. If the crack had bifurcated, measurements were not taken in the bifurcated region. The average value of those measurements is plotted as a function of stress for each system in Figure 2.3.

2.2.2 Modeling

In minicomposites, damage chronology typically begins with the formation of a matrix crack at a surface flaw. This crack grows until it reaches a fiber and the crack is deflected orthogonal to the initial growth direction via an interface debonding mechanism. This debonding dissipates energy that would otherwise contribute to continued crack growth, thereby slowing, and in some cases completely arresting [32], the crack. Upon further loading the crack grows, the composite reaches a state of matrix crack saturation, and the traction forces which oppose COD growth reach a maximum [22, 23, 171]. After matrix crack saturation, further loading cannot increase stresses in the matrix. Instead, fibers begin to bear further increases to the load, fail and reduce the structure's ability to withstand further mechanical stresses. Within each of the two seminal models that describe this process, the microstructural properties dictating COD are the volume fraction of fibers and the interfacial parameters τ and Γ_c. The primary differences only arise when considering the state of pre-existing damage within the minicomposite.

The Marshall Model

The first model, developed by Marshall et al., assumes a pre-existing steady-state crack with an infinitely long debond region. In this model, the only toughening phenomenon is a result of fibers resisting crack growth via friction at the interface. CODs are determined by [10]:

$$\delta = 2\frac{r_f\sigma^2}{4\tau V_f^2 E_f\left(1+\frac{E_f V_f}{E_m(1-V_f)}\right)} \tag{2.1}$$

where r_f is the fiber radius, τ is the interfacial shear stress, and σ is the far-field stress. For the LFC and HFC system respectively, τ was 18.1 ± 4.8 MPa and 34.5 ± 13.0 MPa [33].

The Hutchinson and Jensen Model

The second modeling approach, from Hutchinson and Jensen, makes similar assumptions but allows for a finite, pre-existing debond region. In this model, the energy needed to continue debonding along the fiber-matrix interface and frictional stresses act to reduce the COD [11]. To best match the minicomposite geometry tested, their solution for a hexagonal fiber array subject to constant interfacial shear stress was used. For this case, Hutchinson and Jensen derived an estimate for the COD as [11]:

$$\delta = 2b_2\left[\frac{1-V_f}{V_f c_1 c_3}\left(\frac{\Gamma_c}{E_m r_f}\right)^{1/2} l + \frac{\tau l^2}{E_m r_f}\right] \tag{2.2}$$

where l is the debond length and Γ_c is the interfacial debond energy:

$$l = \frac{r_f E_m (1 - V_f)(\sigma - \sigma_i)}{2 V_f \tau E_c} \tag{2.3}$$

The debond stress σ_i is given by:

$$\sigma_i = E_m \epsilon^T \left[\frac{1}{c_1} \left(\frac{\Gamma_c}{E_m r_f \epsilon^{T2}} \right)^{1/2} - \frac{c_2}{c_1} \right] \tag{2.4}$$

where ϵ^T is the thermal residual mismatch strain of the fiber and matrix. Coefficients b_i, c_i are detailed in [11].

Note that δ has been defined as the full opening of the crack (hence the factor of two included in each model) to correspond with what has been measured. This is in contrast to the original models, which defined δ to be the half-opening of the crack. Finally, in each model the effects of microstructural properties on COD growth, as compared to experiment, are discussed in Section 2.3.3.

2.3 Results and Discussion

2.3.1 Sources of Error

Agreement between experimental data and model predictions depends on how well a model has captured the underlying physics, and also on the fidelity with which measurements can be made. In this study, it was found that discrepancies could be explained by both measurement limitations and model limitations. To deconvolute these two discrepancies, it is necessary to identify sources of measurement error that persist through the study and limit their effect on the analysis of measured CODs.

Inspection of Figure 2.4 suggests that each model significantly underpredicts CODs at the earliest loads (<700 MPa) and only matches the lowest measured CODs at higher loads. While only shown for HFC in Figure 2.4, it was also present in the LFC system. This underprediction has been determined to be the result of image resolution limitations. At loads below 700 MPa, HFC CODs are less than 2 μm, and the spatial resolution of the micrographs was 0.3-0.4 μm/pixel. This has two implications. First, although it is possible to identify the existence of a crack whose COD is less than 0.3 μm, it is not possible to measure the true COD if it is less than this value (Figure 2.5a). Second, the crack faces can only be identified to the nearest pixel. Often, the true position of the crack face will be somewhere in the interior of the pixel, rather than the edge. On average, it should be expected that the true position of the crack face is in the middle of the pixel which has been identified as a crack-face pixel. Therefore, the true COD should be expected to be 1 pixel larger (0.5 pixel at each crack face) than what can be experimentally measured (Figure 2.5b).

For both cases, resolution artifacts can be addressed by including a constant term in Eqns. 2.1 and 2.2 which is equal to the resolution limit ($0.3 - 0.4$ μm). At the low loads, this offset directly accounts for the first effect. At higher loads, the single pixel offset accounts for the average 1-pixel bias resulting from imprecise crack-face identifications. In both cases the single pixel correction yields excellent agreement between model and experiment (Figure 2.4). Thus, the observed discrepancies between the average behavior in the HFC system and the model-predicted behavior can be explained by the resolution limit. The resolution-corrected models will be used in the following discussion.

2.3.2 COD Measurements From In-SEM Observations

Crack opening displacements (CODs) in the low fiber content (LFC) specimen were larger than those in the high fiber content (HFC) specimen for the same stress state throughout the loading profile (Figure 2.3). This is a result of lower bridging tractions in the LFC system ($\Gamma_c = 1.2 \pm 0.5$ Jm^{-2} in LFC versus $\Gamma_c = 5.5 \pm 3.9$ Jm^{-2} in HFC) [172, 173].

In the LFC system, cracks with the largest openings, including the failure crack, occurred relatively far from their nearest neighboring cracks (>500–700 μm) compared to the average crack spacing at matrix crack saturation, $\approx 400\mu$m (Figure 2.6a). This observation is consistent with experiments reported by Chateau et al. for minicomposite specimens [20, 32], who found that cracks with large CODs were relatively far from other cracks. This can be understood by first observing that the debonded region of a crack that is spatially isolated from all other cracks is unlikely to overlap with another debonded region. This allows for the debonded region of that isolated crack to grow in a manner that is unconstrained by its proximity to the debonded region of an adjacent crack. As COD is a monotonic function of the debond length in the Hutchinson model, it follows that maximizing the size of the debond region maximizes the COD. In the HFC system the same correlation was not observed. This is hypothesized to be the result of the HFC system's highly discontinuous microstructure [33] and is the subject of ongoing investigation.

A significant number of CODs did not evolve according to the expected σ^2 dependence from fracture mechanics models. Their evolution clearly did not follow a σ^2 law, but instead appeared to more closely follow a linear or $\sigma^{1/2}$ law (Figure 2.3). Similar deviations from a σ^2 dependence have been previously observed by the authors, but the origin of these deviations was unclear [31]. In the LFC system, 75% (12/16) and 60% (6/10) of the measured cracks in sample 1 and 2 respectively did not follow a σ^2 dependence. In HFC system, 45% (10/22) and 57% (8/14) of the measured cracks in samples 1 and 2 respectively did not follow a σ^2 dependence. Cracks showing this anomalous behavior had either bifurcated or were in sufficiently close proximity to a neighboring crack, such that their debond lengths overlapped (as estimated from Eqn. 3). Therefore, we attribute the deviation from σ^2 to stem from the effects of overlapping debond lengths and crack bifurcations. It must be noted that it is usually assumed that cracks cannot form within the debond region of another crack, but this is only valid when the minicomposite is under a uniform tensile stress at all locations in the cross-section. In realistic systems, geometrical variations in the cross-section due to heterogeneous pore and fiber distributions could result in local stress states that would violate this assumption and permit crack formation within debonded regions.

In the four samples tested (two LFC and two HFC), the existence of nearby neighboring cracks/bifurcations (NNCB) was a necessary but not sufficient condition to cause the deviation from σ^2. All of the 62 (26 LFC and 36 HFC) cracks whose CODs deviated from σ^2 had NNCBs, but not all of cracks with NNCBs exhibited this deviation. There were case exceptions. Over all samples, 12 cracks with NNCBs (6 in LFC and 6 in HFC) still demonstrated the canonical σ^2 dependence. The mechanism that drove these case exceptions is currently unclear.

2.3.3 Comparison of Experimental Measurements to Model-Based CODs

In this section, we investigate how the microstructural assumptions of the Marshall and Hutchinson models influence their predictive capability. We demonstrate that matching model assumptions to specimen microstructure, specifically the state of pre-existing debonding, results in better prediction of the experimentally measured average COD behaviors. This is particularly important when interfacial parameters are higher, such as in the HFC system investigated herein (Figure 2.3). It is also demonstrated that simpler models incorporating less physics are likely to be more useful for lifetime prediction, as they provide more conservative estimates and are less likely to underpredict CODs.

In the LFC system, the Hutchinson model is more suited to predict *average* CODs as compared to the Marshall model (Figure 2.3). This is attributed to the initial state of fiber debonding in the microstructure. Fiber push-in testing revealed the majority (90%) of fibers tested were completely bonded to the matrix via the interphase. The assumptions of the Marshall model, specifically that fibers are already debonded, do not match the microstructure. As a result, the predictions of the Marshall model fall further above the statistical average of CODs in the LFC system (Figure 2.3). In contrast, the state of the microstructure more closely matches the assumptions of the Hutchinson model, where fibers are not assumed to be debonded prior to loading.

In the HFC system, the Marshall model is more appropriate for predicting *average* COD behavior (Figure 2.3). Part of the reason for this is due to the state of fiber bonding within the HFC system. During fiber push-in testing, it was discovered that only a small portion of fibers (43%) were bonded to the matrix via the interphase with the others debonded prior to loading. As a result, the Hutchinson model, which assumes bonded fibers, the crack closure effect due to energy absorbed during debonding predicted by the model is larger than what is expected in the minicomposite (Figure 2.3). Instead, since about half of the fibers are debonded prior to loading, the effects from debonding are reduced, and frictional coupling of the fibers to the matrix becomes more dominant for determining CODs, as is assumed in the Marshall model. Here it is noted that unbonded fibers were excluded from interfacial parameter calculations, meaning that the values of Γ_c input to each model are likely higher than the true global average. Reducing the value of Γ_C for the HFC system to 43% of its original value (i.e. the estimated ratio of fibers bonded) causes CODs to increase by no more than 0.1 μm. This indicates that the value of Γ_c is not the only factor which contributes to the underprediction by the Hutchinson model. Other factors which contribute are hypothesized to also be a result of the highly heterogeneous microstructure of the HFC system.

These findings highlight that canonical models can accurately predict COD trends when the microstructure matches model assumptions; this is true even when crack paths are non-linear [32], have bifurcated, or the crack has a nearby neighbor. Additionally, the assumption that all fibers bear load through the whole lifetime of the composite appears to be valid for COD prediction. If this assumption did not hold, then each model would tend to underestimate CODs. It should be noted that while such progressive fiber failure does not appear to have a significant impact on the evolution of CODs, it may play a role in the site of final failure [12].

2.3.4 Effects of Interfacial Parameters

While COD models accurately capture the average behavior of cracks, as described in Section 2.3.3, they fail to account for the range of CODs observed in experiment. To explain the observed ranges of CODs, it is helpful to consider the simplifications assumed by the models. In these models, and more broadly [174–176], architectural simplifications are used, such as: homogeneous interphase volumes, an unchanging microstructure, and uniformly distributed fibers. These simplifications do not hold for CMCs in practice. Processing considerations dictate that interphase volumes are non-negligible and often non-uniform, the microstructure evolves as damage accumulates, and fibers are rarely distributed uniformly.

Here, the effects of simplification of homogeneous interphase volumes are exoplored. It is shown that the largest CODs can be accounted for by permitting variable interfacial parameters, while the smallest CODs cannot. This suggests that some of the issues related to modeling CODs, such as neglecting progressive fiber failure, can be attributed to other factors such as constituent property variations. Similarly, the existence of CODs below the prediction envelope (Figure 2.7) suggests that the growth rate of some cracks is impeded by effects not related to constituent properties. With respect to the neglect of progressive fiber failure, it must be noted that it exists in real systems [12, 20], but cannot currently be deconvolved from local variations in interfacial parameters. Additionally, the effects of non-uniform fiber distributions cannot currently be captured with analytical models and are not considered here.

The assumption that composites can be described by a single valued interfacial property at every location in the minicomposite sample is only valid when interfacial properties come from a single statistical distribution. In the course of CMC processing, non-uniform interphase properties can arise at different locations in the minicomposite specimen due to heterogeneities in the processing environment. It follows that one cannot assume that interfacial properties will be drawn from the same distribution at each location in the minicomposite. For example, the LFC system has thin interphases in comparison to the HFC system, which has a significant interphase volume ($> 2\mu$m thickness or $\approx 25\%$ volume fraction). As a result, interfacial parameters are less likely to vary in the LFC system as there is less volume that can contain defects in the interphase and the coating thickness is relatively uniform. This is in contrast to the HFC interphase, which is large ($\approx 25\%$ by volume) and non-uniform (Figure 2.1). As a result, the averaged interfacial values used as inputs to the model may not represent the local values at each crack, particularly in the HFC system.

To explore the reality of local variations in both τ and Γ_c arising from non-uniform processing, the full range of interfacial values (described in section 2.2) was used as input into the Hutchinson model for the LFC system, and the Marshall model for the HFC system (Figure 2.7). This procedure provides bounds for the smallest and largest COD that is expected to be measured.

Sweeping the full range of interfacial parameters suggests that microstructural effects not related to constituent properties (e.g. crack geometry, heterogeneous fiber distribution) play a significant role in how cracks grow. If constituent properties were the main driver in COD growth rates, then it would be expected that the full range of measured CODs could be captured by using the full range of interfacial parameters as model inputs. In contrast to this, only the upper middle portion of measured CODs was captured with the full range of interfacial parameters. Therefore, a significant number of cracks had growth rates lower than those that could be predicted by variations in intrinsic constituent properties.

Impeded COD growth could not be solely attributed to the effects of constituent property variations and was instead attributed to restricted debond lengths and crack geometry. With respect to the limited debond length, many cracks that formed were not isolated. As previously discussed, this serves to limit the size of the debond region. Since CODs are dependent on the size of the debond region in the Hutchinson model, a limited debond length would limit CODs and cause them to grow less quickly than more isolated cracks. The second effect which is hypothesized to be the cause of restricted COD growth is a consequence of crack-growth geometry. In minicomposites, cracks initiate at the surface forming a circumferential ring which then grows inward towards the center [32]. Therefore, many cracks which present as through-thickness, when observed by surface-only SEM measurements, likely have significant matrix content which is uncracked in the bulk [32, 172]. This would restrict COD growth, due to the enhanced stiffness that the additional intact matrix provides past the proportional limit of the minicomposite [172]. Therefore, cracks with intact matrix below the surface would have CODs falling below model-predicted values.

While the method used to interrogate the effect of non-homogeneous interfacial properties provides bounds on what can be expected to be measured, the exact effects remain unknown. For example, a distribution of values exists within a single plane, but it is unclear how large deviations between planes should be. There are no known studies which investigate distributions of interfacial parameters across the whole gauge length, as well as in-plane surface of a minicomposite. Further work is needed to assess the degree to which non-uniform processing changes local interfacial parameters.

2.4 Conclusions

In this chapter, *in-situ* matrix crack opening displacements (CODs) were characterized for two systems of SiC/SiC minicomposites containing a low fiber volume fraction (LFC) and high fiber volume fraction (HFC). The measured CODs of both LFC and HFC were compared to two seminal models: (i) by Marshall [10] and (ii) Hutchinson and Jensen [11]. The CODs of a large proportions of cracks, 69% (18/26) and 50% (15/36) in the LFC and HFC system respectively, were observed to deviate from the canonical σ^2 dependence. Such cracks were in close proximity to a neighboring crack, or had bifurcated. Although this had been previously observed by the authors [31] for macrocomposite systems, this is the first known report in minicomposites. Further work is being conducted to understand this.

The predictive power of each COD model depended on how well the state of debonding in the real microstructure matched the assumption of the model, rather than the geometry of the crack plane or existence of NNCBs. When the state of debonding in a sample is unknown, or interfacial parameters are large, models that do not account for finite debond lengths (such as the Marshall model [10]) should be used, as they will provide conservative COD estimates. In the LFC system, the Hutchinson and Jensen model better reflected observed crack behavior because the majority of fiber/matrix interphases (90%) were intact. In contrast, 43% of fiber/matrix interphases in the HFC system were undamaged, which in turn led to led to the excellent agreement between experiment and the Marshall model where the effects of debonding were negligible compared to the effects of frictional shear resistance.

To account for spatial heterogeneities in the interphase resulting from realistic processing conditions, an envelope of CODs was created using the distribution of interfacial parameters which was experimentally measured. When this modification is made, the models adequately explained the upper/middle range of observed CODs but were unable to explain the lower range of CODs. These smaller-than-predicted CODs were attributed to the cracks being proximity to a neighboring crack, as well as undamaged matrix content below the surface which could not be characterized via SEM measurements.

Future exploration of these hypotheses will lead to better modeling choices and improved lifetime prediction of CMCs however, the ability to rapidly image subsurface damage represents a significant barrier to conducting such investigations. Studies which make use of laboratory scale x-ray computed tomography machines may take weeks to fully test a single sample. This renders their use infeasible, since many of these machines exist in user-facilities where acquiring such extended instrument time is not possible. While the capability for rapid imaging exists at synchrotron facilities, the accessibility of these tools is limited by long queues ($\approx$ 2 years) and small number of facilities (8 located in the USA according to the International Atomic Energy Agency). As a result, it is highly desirable to develop the capability to interrogate subsurface damage without the use of such tools. Consequently, the following chapters focus on the creation and use of trustworthy ML tools, which have the unique potential to enable large-scale subsurface damage mechanism identification in SiC/SiC CMCs.

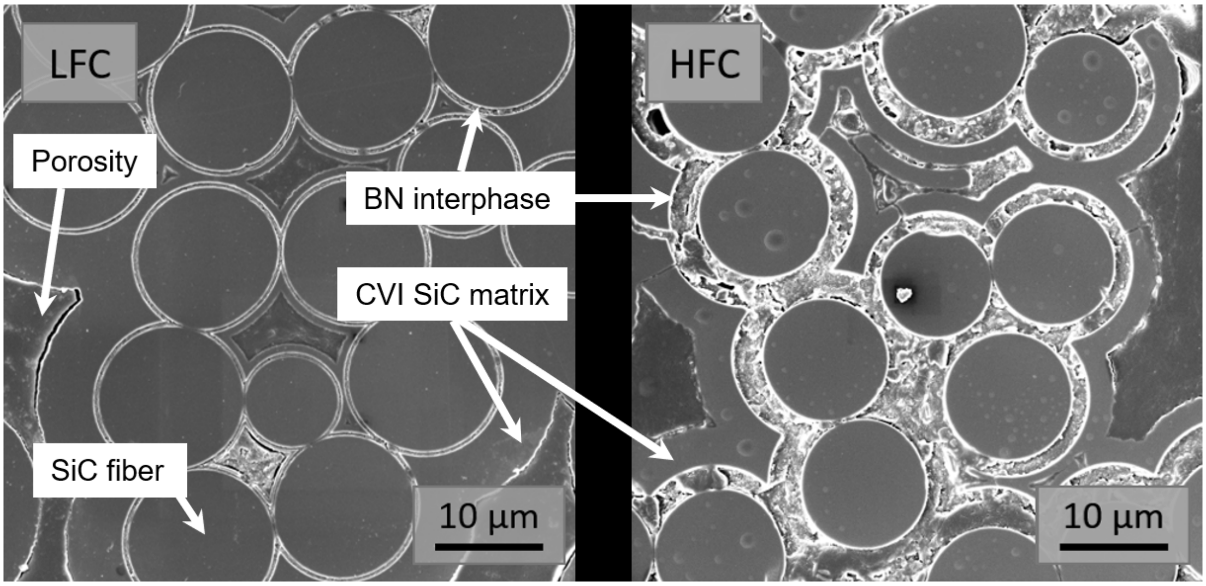

Figure 2.1: **Constituent distribution in LFC and HFC specimens.** In both systems, the fiber distribution was heterogeneous, in contrast to fundamental modeling assumptions. LFC and HFC specimens exhibited closely spaced regions of fibers, which resulted in uneven matrix and interphase deposition.

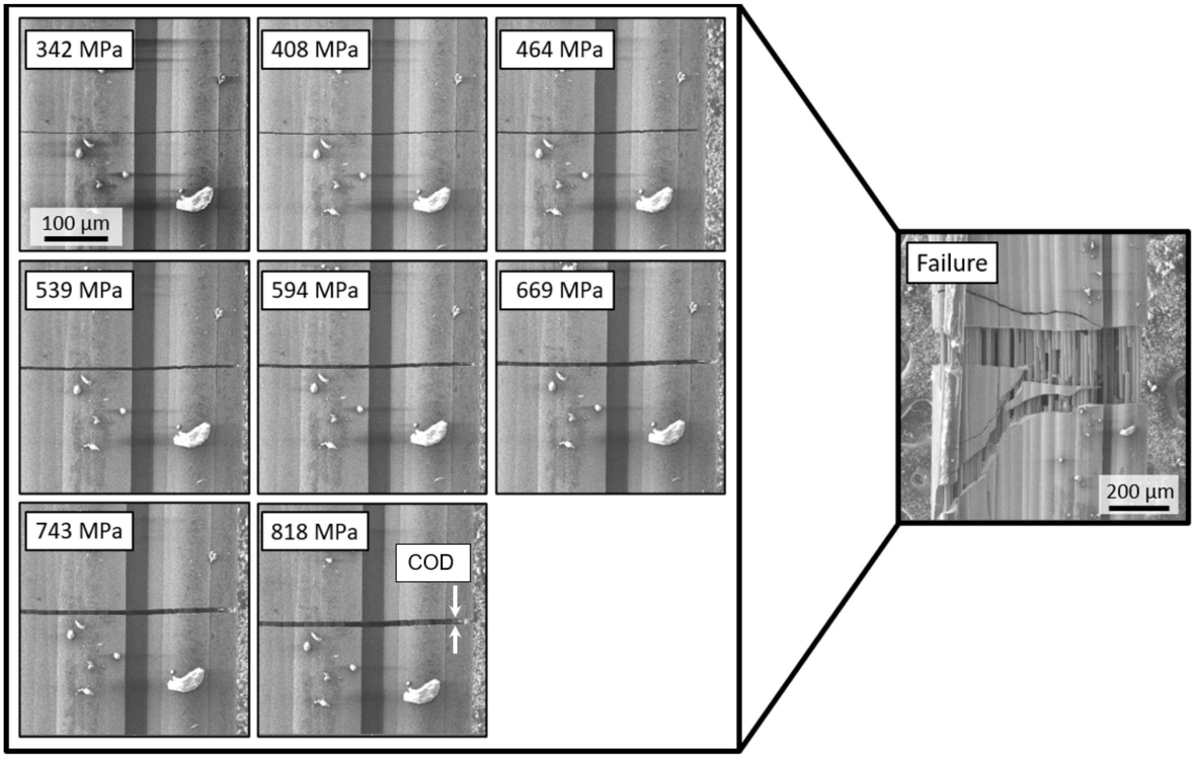

Figure 2.2: **The progressive crack opening behavior of the LFC failure crack.** By high-resolution SEM, *in situ* behavior of these CODs can be measured, which has the potential to advance understanding of why some cracks evolve to become more probabilistic locations for failure than other cracks.

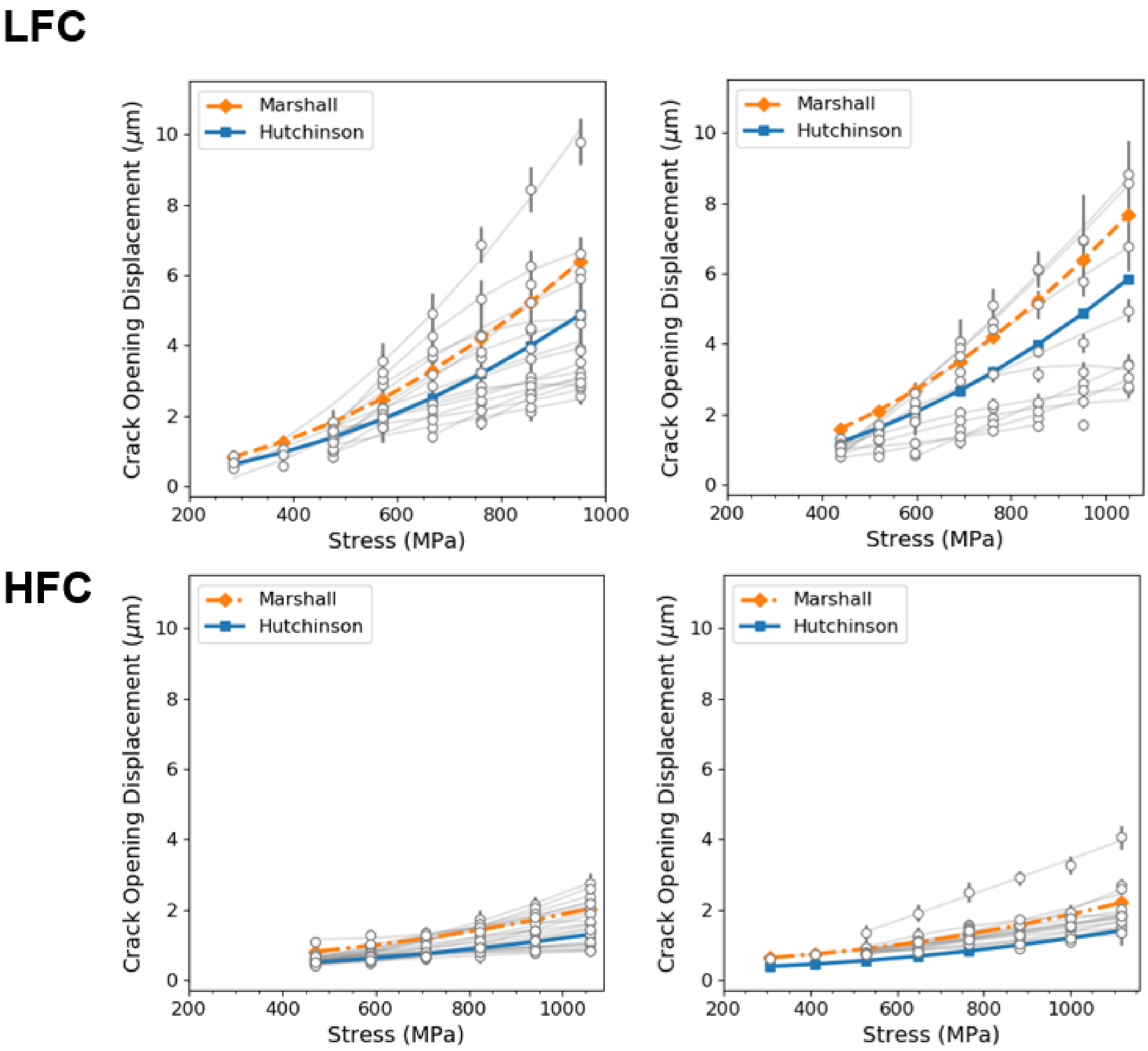

Figure 2.3: **Experimentally measured and resolution-corrected model-predicted CODs in two samples of the LFC and HFC systems.** Experimentally measured CODs for the LFC and HFC system are shown, with grey polynomial regression lines. CODs as predicted by [10] and [11] are overlaid. In the LFC system, the Hutchinson model accurately captures the average behavior of the cracks, while the Marshall model overpredicts CODs. In the HFC system, the Marshall model accurately predicts the average CODs while the Hutchinson model tends to underpredict average CODs. This is attributed to the different microstructures.

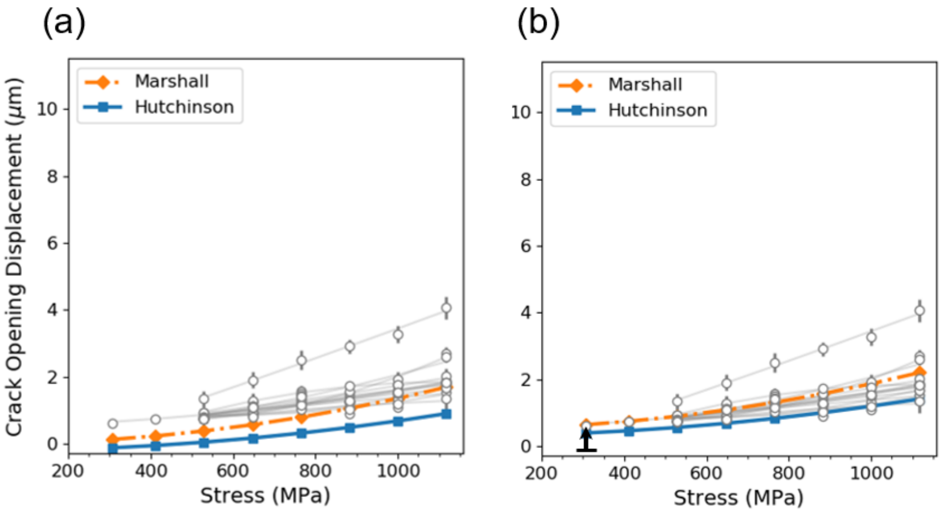

Figure 2.4: **Experimentally measured CODs for sample 2 of the HFC system with (a) original models and (b) resolution-corrected models overlaid.** The arrow denotes magnitude of the correction. Experimental CODs, particularly at the lowest stress states, measure 1-2 pixels in width. When this is the case, Measurements are, on average, a single pixel larger than what will be predicted by models (discussed in detail in Section 2.3.1).

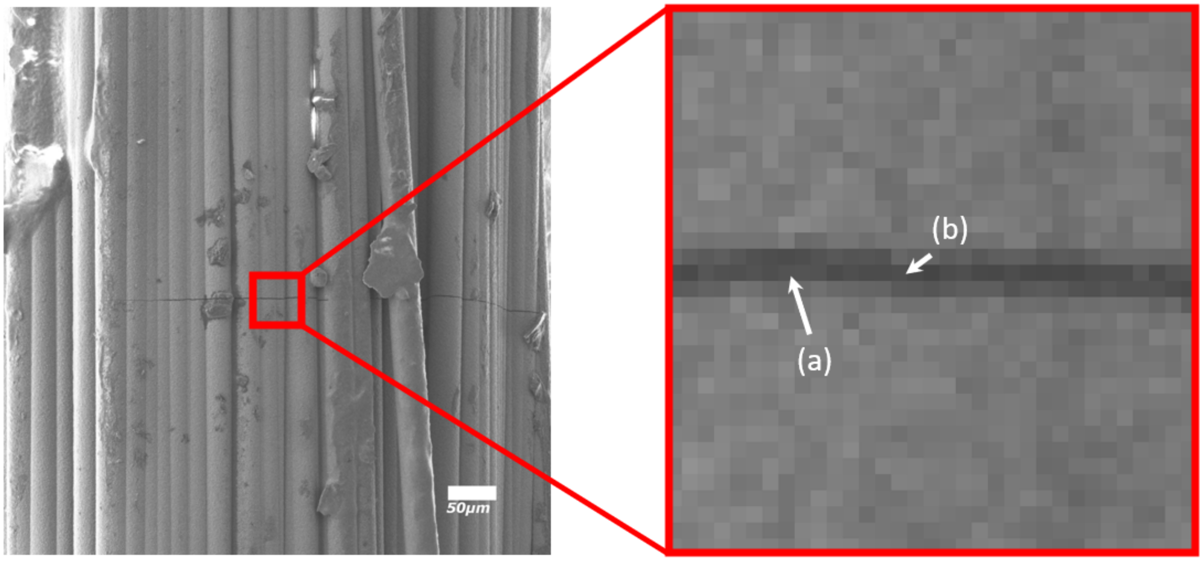

Figure 2.5: **Example of crack that can be identified, but cannot be accurately measured because the COD is below the resolution limit.** This presents issues related to identification of (a) the position of the crack face, whose true position resides somewhere in the interior of the pixel and (b) the COD when its value is less than the resolution limit (.449 μm).

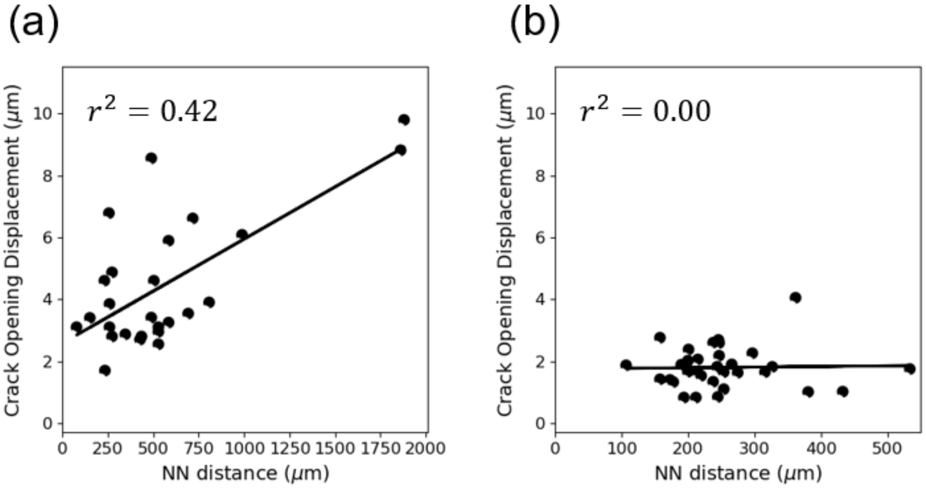

Figure 2.6: **Maximum crack opening displacement of all cracks in the (a) LFC and (b) HFC system as a function of distance to nearest neighboring crack.** (a) The CODs measured in the LFC system were found to be correlated to distance to neighboring cracks. (b) The same correlation was not observed in the HFC. This is hypothesized to be a result of the HFC's discontinuous microstructure, and is the subject of ongoing investigations.

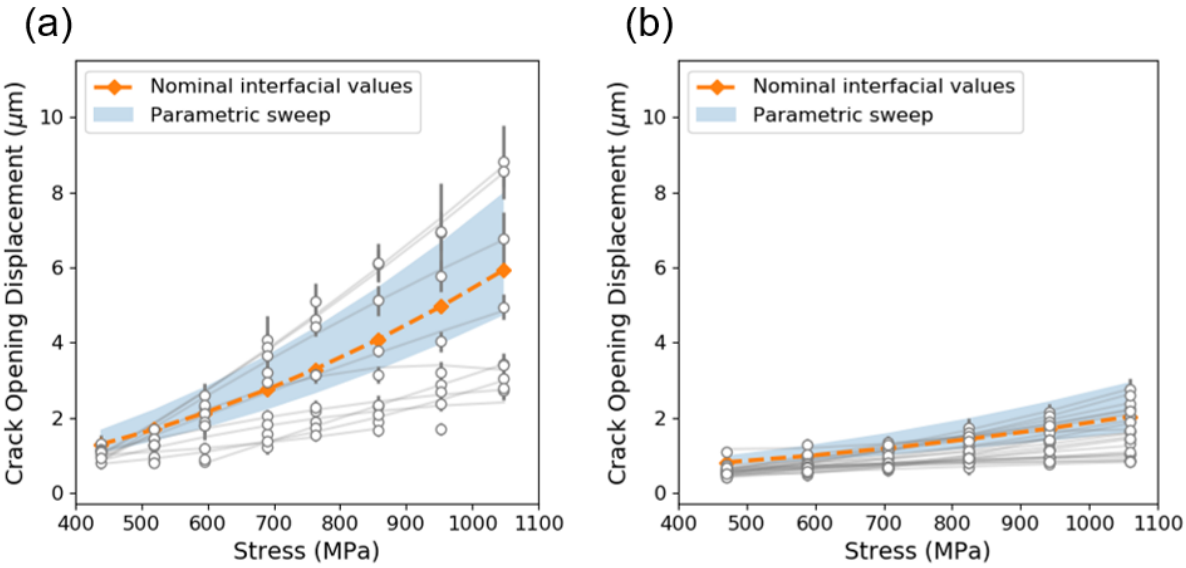

Figure 2.7: | Resolution-corrected model-predicted CODs using the measured range of interfacial values as inputs compared with experimental data. (a) The envelope Hutchinson-predicted CODs for the lower and upper interfacial values are shown alongside the nominally predicted CODs for the LFC system. (b) The envelope of Marshall-predicted CODs in the HFC system. While the range of interfacial parameters explains the upper middle range of measured CODs, a large number of cracks fall below model predictions. This is hypothesized to be a result of overlapping debond lengths and intact matrix below the surface of the minicomposite.

Chapter 3

ML-Enabled Damage Mechanism Identification From AEs in SiC/SiC Minicomposites

3.1 Scope

In the previous chapter, I have highlighted the need for novel tools to characterize subsurface damage in SiC/SiC composite structures. Among the most promising methods is acoustic emission (AE), however this technique is limited in that it cannot be used to identify damage mechanisms in elastically similar composites. Here, I will demonstrate that damage mechanism identification from acoustic emission (AE) signals generated in minicomposites with elastically similar constituents is possible. AE waveforms were generated by SiC/SiC ceramic matrix minicomposites (CMCs) loaded under uniaxial tension and recorded by four sensors (two models with each model placed at two ends). Signals were encoded with a modified partial power scheme and subsequently partitioned through spectral clustering. Matrix cracking and fiber failure were identified based on the frequency information contained in the AE event they produced, despite the similar constituent elastic properties of the matrix and fiber. Importantly, the resultant identification of AE events closely followed CMC damage chronology, wherein early matrix cracking is later followed by fiber breaks, even though the approach is fully domain-knowledge agnostic. Additionally, the partitions were highly precise across both the model and location of the sensors, and the partitioning was repeatable. The presented approach is promising for CMCs and other composite systems with elastically similar constituents.

3.2 Methods

3.2.1 Experimental Configuration

SiC/SiC minicomposites (Rolls Royce High Temperature Composites, Cyprus, CA) were loaded under uniaxial tension using a custom load frame built in-house with a 220 N load cell and a cross-head displacement rate of 0.120 mm/minute, corresponding to a nominal strain rate of 7.5×10^{-5} s^{-1}. Specimens consisted of 500 Hi-Nicalon Type S^{TM} SiC fibers (HNS fibers), a boron nitride (BN) interphase, and an overlayer of chemical vapor infiltrated (CVI) SiC matrix. The microstructural and damage characteristics of these minicomposites are described in Chapter 2.2 and [33], where they are referred to as low fiber content (LFC) minicomposites. Three tension tests on LFC minicomposites are presented here to demonstrate the ML framework.

AE activity was recorded using a four-channel fracture wave detector acquisition system (Digital Wave Corporation, Centennial, CO). The threshold voltage was set to 0.1 V, the number of pre-trigger points was set to 256, and the total length of signal captured was 1024 points at a rate of 10 MHz. Two sets of AE sensor models were mounted (Figure 3.1) on Duralco 132 epoxy tabs with vacuum grease (Figure 3.1b). The tab thickness was 0.35mm, as measured from the surface of the minicomposite to the sensor. Tabs were created using 3D-printed molds of 10mm in diameter.

The choice of coupling medium is significant for waveform transmission through an interface. The transmission coefficients are determined by how closely the acoustic impedance of the specimen and the couplant match. As the impedance mismatch between the two increases, the transmission coefficient decreases [154]. Of the available medium choices examined, Duralco 132 exhibited the lowest impedance mismatch to SiC/SiC.

Additionally, the thickness of the epoxy tabs controls the quality of data collected. Previous work has reported that transmission coefficients decrease as a function of coupling medium thickness[154]. Thinner epoxy tabs resulted in more consistent labeling (discussed in Chapter 3.3.3), which is hypothesized to be a result of the higher transmission coefficients at all frequencies as compared to thick (>1.5mm) epoxy tabs.

AE signals were recorded using two miniature S9225 piezoelectric AE transducers (Physical Acoustics, Princeton, NJ) with a broadband response of 300-1800 kHz and two B1025 transducers (Digital Wave Corporation, Centennial, CO) with a broadband response of 50-2000 kHz. The sensors were designated S9225-a, S9225-b, B1025-a, and B1025-b, where the I.D. letter corresponds to the position on the minicomposite (Figure 3.1). The acquisition system was synchronized; when one sensor was triggered, all sensors recorded waveform data simultaneously.

3.2.2 Data Processing

After acquisition, the raw AE data was cleaned to remove events not suitable as input to the ML framework. First, clearly identifiable noise events were removed (Figure 3.2a).

These events were either characterized by a single voltage spike in the time-domain of the waveform, or they presented with energies less than 0.001 V (low signal-to-noise ratio). Recorded waveforms that showed multiple damage events within the same time window (Figure 3.2b) were also not considered; less than 1% of all recorded waveforms were of type (b). If the signal saturated the sensor (Figures 3.2c and 3.2d), the event was not considered. The majority of waveforms removed were of types (a), (c), and (d). Then, a location analysis was performed to remove out-of-gauge events [145] using:

$$\text{location} = \frac{x}{2}\left(\frac{\Delta t}{\Delta t_x}\right) \tag{3.1}$$

where x is the sensor separation distance, Δt is the difference in time-of-arrival, and Δt_x is the difference in time-of-arrival for an out-of-gauge event as a function of the damage parameter. This analysis ensured that only signals arising from in-gauge damage events were analyzed. During this process it was found that the majority of type (a) waveforms ($> 90\%$) came from out-of-gauge events. This cleaning removed approximately 35% of all recorded AE waveforms, including out-of-gauge events.

3.3 Unsupervised Classification of Acoustic Spectra

AE data collection and preliminary filtering (Chapter 3.2.2) represent Step 1 of the general unsupervised AE-ML framework which proceeds as follows:

1. **Experimentation**: A number, n, waveforms are collected
2. **Feature Extraction**: Waveforms are represented in feature space by extracting d pertinent features (i.e. each waveform is represented by a d-dimensional vector)
3. **Machine Learning Algorithm**: An algorithm is selected to partition waveforms

into clusters that are representative of damage mechanisms

4. **Labeling and Error Analysis**: Post-clustering analysis is performed to assign damage mechanism labels to clusters and assess the validity of results

At each step, considerations must be made to ensure the framework functions properly. The following section describes Steps 2-4. The code created for this investigation utilizes the Scikit-learn toolbox [6] and is available without restriction [177].

3.3.1 Feature Extraction

After AE waveforms are collected, suitable representations to encode damage mechanism information must be determined. Appropriate waveform features are those that are more dependent on the generating damage mechanism than on extrinsic factors, such as propagation distance. Prior finite element analysis studies, supported by experimental evidence, indicate that partial power is one such feature, provided that signals are recorded in the near-field [67, 74].

Here, AE waveforms are encoded with a modified partial power scheme, where the i^{th} component of the feature vector is:

$$Feature_i = \frac{\int_{k_{i-1}}^{k_i} F[x(t)]dk}{\int_{k_0}^{k_d} F[x(t)]dk} \tag{3.2}$$

where $F[*]$ is the Fourier transform operator, $x(t)$ is the recorded signal, and d is the number of entries in the feature vector.

We set $k_0 = 200$ kHz and $k_d = 800$ kHz for all specimens. To determine the value of k_0 a parametric sweep from $k_0 = 50 - 350$ kHz in increments of 50 kHz was conducted. The value optimizing validity metrics (Chapter 3.3.3) was chosen. Including partial powers above 800 kHz did not improve clustering quality. This was attributed to the fact that the power of a frequency spectrum above 800 kHz approached zero and thus could not provide additional discriminating power.

The frequency bounds encompass the pre-amplifier bandpass on the Digital Wave system and the flat frequency response of the B1025 sensors. While the frequency range includes values outside the flat frequency response of the S9225 sensors, this does not impact discriminating power. Any partitions made by the ML algorithm result from differences between waveform characteristics. The only stipulation is that recorded waveforms should be clustered independently for each sensor to capture the shift in damage mechanism (i.e. the singular set of all waveforms are not clustered together).

Another parametric study was conducted to determine d, where d was swept from $d = 2$ to 45. When all other parameters are fixed, $d = 26$ ($\Delta k = 23$ kHz) optimized validity metrics (Chapter 3.3.3) for all specimens.

Though previous investigations have included the partial power approach as part of their representation schemes [49, 89, 115, 159, 169], the representation scheme described here is unique in that it uses a comparatively much finer resolution and *only* uses partial power. Typically the partial power bandwidth used is 200-600 kHz, whereas the approach herein uses a width of 23 kHz.

3.3.2 Spectral Clustering

Once AE data is properly represented, a suitable ML algorithm for clustering can be chosen. For this task, spectral clustering was used. This is an unsupervised learning technique that has been shown empirically to outperform k-means [76, 178] and is less restrictive with respect to assumptions about input data geometries.

Spectral clustering models the input dataset as a graph with nodes (data points) connected by edges whose weight is 1 if the nodes are nearest neighbors (NNs) and 0 if they are not. The algorithm finds the optimal place to remove edges and segment the original graph into a user-specified number of sub-graphs (i.e. clusters) [76].

Both the number of clusters and the number of NNs are considered hyperparameters (i.e. a set of user-selected parameters). One common method for estimating the number of clusters is through use of the eigengap heuristic, a measure of differences between successive eigenvalues of the graph Laplacian of the data [179]. However, noisy data can reduce differences between successive eigenvalues, which is often the case for AE data. In this case, the eigengap heuristic is not sufficient for determining the number of clusters. Instead, a parametric sweep from 2-5 clusters is performed and a drop in validity metrics is used to indicate the optimum number of clusters (Chapter 3.3.3). A steep drop is observed after 2 (Figure 3.3), corresponding with the hypothesis that matrix cracking and fiber failure events (the dominant damage mechanisms in SiC/SiC minicomposites) can be differentiated.

It is important to note there are other less dominant damage mechanisms active in minicomposites during loading (e.g., interfacial debonding and frictional sliding). It is well-established from a micromechanics frame that when matrix cracking occurs, there is simultaneous debonding and sliding in the crack wake; it is also understood, that when fibers fail, there is simultaneous fiber sliding and pullout [23, 180]

When dominant and non-dominant mechanisms occur simultaneously, their waveforms become superimposed [181]; when non-dominant mechanisms occur independently, they likely do so in quantities too small for recognition by the ML algorithm. As such, it is currently infeasible to isolate events resulting from damage to the BN interphase or determine which AE features are characteristic of such damage. This is expected to be a source of error. Moreover, the inability to discriminate interfacial damage from other types of damage is reflected by the steep drop in validity metrics seen after two clusters in Figure 3.3.

Similar to the determination of d and the number of clusters, a parametric study from NN$= 5 - 20$ was performed. The value of NN that optimized validity metrics slightly varied between studies (5, 5, 7 for three experiments respectively), however there was a range of values for NN that produced acceptable results. To demonstrate the effectiveness of our approach, the number of NN is fixed to be 5 for all specimens resulting in sub-optimal validity metrics (Section3.3.3) for Experiment 3, particularly between sensors B1025-a and S9225-a (Table 3).

3.3.3 Validity Metrics for Error Analysis

To evaluate the efficacy of our AE-ML framework, events at each sensor were clustered according to the steps described above, the results of which are called a *partition*. The desired outcome is that all partitions for AE data from a given specimen are the same and are independent of both the sensor model and sensor location.

To quantify partitioning success, the total matching rate is first considered, which is the percentage of events assigned the same label by the clustering routine. As clusters have unbalanced sizes (fewer fiber break events are expected than matrix crack events), it is possible to be unable to discriminate between damage mechanisms and retain high matching rates. For example, if 90% of AE events come from matrix cracks, then classifying *every* AE event as a matrix crack would yield a 90% matching rate, yet the ML framework would have no discriminating power.

Therefore, it is also useful to consider the permutation model of the adjusted Rand Index (ARI) which makes considerations for unbalanced cluster sizes[131, 134]. The ARI is an adjusted-for-chance version of the Rand Index (RI), a metric for comparing the similarity of a partition to the ground truth. First, the RI for two different partitions,

(A, B), is calculated as:

$$RI(A,B) = \frac{N_{11} + N_{00}}{\binom{N}{2}} \tag{3.3}$$

where N is the number of elements, N_{11} is the number of element pairs which are grouped into the same cluster in both partitions, and N_{00} is the number of element pairs that are grouped into different partitions in both A and B. The ARI is then calculated as [6]:

$$ARI(A,B) = \frac{RI(A,B) - \mathbb{E}[RI(A,B)]}{\max[RI(A,B)] - \mathbb{E}[RI(A,B)]} \tag{3.4}$$

where $\mathbb{E}[RI(A,B)]$ is the expected value of the RI under a random model. The ARI is bound between 0 and 1, with 0 corresponding to random label assignments, and 1 corresponding to perfectly matching labels.

This metric is useful for comparing a partition to the ground truth and is also useful for comparing two partitions that are assumed to be drawn from the same random model [134]. This makes it an effective tool to compare similarity between two partitions whose ground truth is not known *a priori*. ARI values exceeding 0.40 are correlated with high values of other classification metrics [133].

3.3.4 T-Distributed Stochastic Neighbor Embedding

A final, necessary step in this study was to confirm that the AE data forms identifiable clusters in the chosen feature space. To this end, t-distributed stochastic neighbor embedding (t-SNE) was employed. T-SNE is a manifold learning algorithm used to produce a low-dimensional visualization of high-dimensional data. Although t-SNE axes, inter-cluster separation, and cluster size have no intrinsic meaning, t-SNE has been empirically shown to be a powerful tool for the identification of natural cluster structures in high-dimensional data [9].

T-SNE maintains pairwise distances between the high-dimensional and low-dimensional representation of feature vectors, x_i and y_i, respectively [9]. For a given feature vector, x_i, t-SNE models pairwise distances in the high-dimensional representation according to a Gaussian probability distribution with standard deviation σ_i, centered at x_i. Under this model, the conditional probability of finding another feature vector x_j is then:

$$p_{j|i} = \frac{\exp(-|x_i - x_j|^2/2\sigma_i^2)}{\sum_{k\neq i}\exp(-|x_i - x_j|^2/2\sigma_i^2)} \tag{3.5}$$

and pairwise distances in the high-dimensional space are:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \tag{3.6}$$

The pairwise distances in the high dimension are then translated to similar pairwise distances in the low dimension, q_{ij}, which follow a student t-distribution with a single degree of freedom:

$$q_{ij} = \frac{(1 + |y_i - y_j|^2)^{-1}}{\sum_{k\neq l}(1 + |y_k - y_l|^2)^{-1}} \tag{3.7}$$

If the low-dimensional representation has correctly maintained the same high-dimensional pairwise distances, then $p_{ij} = q_{ij}$ for all pairs i, j.

An important consideration for t-SNE plots is the choice of "perplexity." This hyper-parameter estimates a global value of σ_i, as there is no single value of σ_i to describe all data points. Perplexity measures how much of the local structure is retained in the final low-dimensional map; as perplexity increases, local structure information is exchanged for global structure information [9]. Typical values of perplexity range from 5-50; a perplexity value of 15 is chosen as this best shows the cluster structure within our data.

3.4 Results and Discussion

Matching rates and ARIs for AE data from three SiC/SiC specimens are presented (Tables 1, 2, and 3) for partitions made from: (i) sensors of the same model fixed on opposite ends of the specimen gauge, and (ii) sensors of different models fixed at the same gauge location. Specimen 1 was found to have a misaligned sensor on its larger epoxy tab wherein it did not fully overlay the minicomposite and is included to show that the ARI metric can also be used to detect experimental issues such as these. The ARI of zero for sensor pairs that included sensor S9225-b are the result of this improper placement.

From the high values of matching rates and ARIs obtained across specimens, we conclude that labels are not assigned randomly between sensors based either on model or location. A corollary of this observation is that stochastic experimental effects that are known to influence frequency spectra, such as source-to-sensor distance [149] or proximity to a free surface [65], do not drive label assignment when using the approach presented herein.

The clustered events at each individual sensor exhibit distinct activities in the stress domain, which are strongly characteristic of how damage progresses in a CMC (shown for Specimen 3 in Figure 3.4). These clusters follow the established chronology of damage accumulation in CMCs, wherein cracks initiate in the SiC matrix early in the loading profile and evolve throughout the specimen lifetime. As such, the initially activated cluster is designated as matrix cracking. After the onset of major matrix cracking, fiber failure begins and fibers continue to fail until rupture [182]. The secondary cluster becomes significantly active at 70-85% of the ultimate tensile strength (UTS), which agrees with experimental observations of fiber failure in SiC/SiC [12].

To further explore the hypothesis that there are frequency trends which allow for discrimination between matrix cracking and fiber failure, it is useful to inspect the partial powers of AE signals (the input representation) as a function of stress. It was found that select frequency bands exhibited similar behaviors, as shown in Figure 3.5. Specifically, AE events occurring at stresses $> 70\%$ UTS exhibited tighter distributions in partial power. A two sample Kolmogorov-Smirnov test shows that the partial powers, for the selected frequency bands, sampled below 70% UTS come from a different distribution than the partial powers sampled above 70% UTS at a significance level of $\alpha = 0.01$ (Figure 3.5). This decrease in partial power scatter coincided with the activation of the fiber failure cluster, but it is currently unclear whether this correlation is (i) characteristic of the damage mechanism or (ii) characteristic of the differential strain-at-failure of the constituents [173, 183] (e.g. a weak fiber failing at low strain is indistinguishable from matrix cracking at the same strain). The second option is possible while still allowing for the stepped activity observed as the strain-to-failure of the fiber and matrix are disparate. The former possibility is more consequential, as it would indicate that the ML algorithm is *learning* the differences between dominant damage mechanisms. Given that the homogeneous orthotropic model of wave propagation does not predict meaningful frequency trends for elastically similar material systems, additional experimentation and explicit modeling of AE in SiC/SiC minicomposite geometries is needed to understand the physical origins of the observed trends.

The partitioning of each AE event into its associated cluster is the result of the intrinsic cluster-like structure and not an artifact of the chosen algorithm. This finding is evident in the visualization of the input feature vectors via t-SNE. In Figure 3.6, the left-hand column shows the raw feature vectors plotted via t-SNE, and the right-hand column shows the same feature vectors subsequently colored according to the labels assigned by our ML framework. In Figures 3.6a and 3.6b, the two clusters are sufficiently distinct to be visually identified in the unlabeled data. The unlabeled cluster structures in Figures 3.6c and 3.6d are less readily discernible, yet still evident. This behavior is likely a result of dimension reduction rather than an intrinsic lack of cluster structure, as validity metrics show the events are given the same labels.

3.5 Conclusions

Damage mechanism identification in SiC/SiC CMCs from acoustic emission data is of interest for both lifetime prediction and structural health monitoring. While computational frameworks exist for damage mechanism identification, these are predominantly successful only when the composite constituents are elastically dissimilar. As such, differentiating between the dominant damage mechanisms in SiC/SiC architectures has remained a long-standing challenge.

In this chapter, an AE-ML framework capable of overcoming these difficulties was developed and evaluated. A modified partial power representation scheme is adopted which allows inspection of local changes to frequencies. This representation scheme is combined with the spectral clustering algorithm, which is well-suited for partitioning AE data. This framework is then applied to waveforms collected by our unique 4-sensor configuration, which allows us to draw the following conclusions:

1. Damage mechanism identification in elastically similar composite systems is possible, when salient features are chosen to represent AE waveforms and a suitable ML approach is applied.

2. Partial power is a salient feature for damage mechanism discrimination. It is not significantly perturbed by stochastic experimental factors, as is evidenced by consistent labeling between sensors, independent of both location and model. Moreover, t-SNE plots show AE data intrinsically adopts cluster-like structures when repre-

sented by this scheme.

3. There are meaningful frequency trends, which are not predicted by the orthotropic model of wave propagation, which enable damage mechanism discrimination in SiC/SiC CMCs. Further investigation is needed to determine the physical basis of this.

This AE-ML framework is domain-knowledge agnostic, yet when contextualized with domain knowledge it is clear that cluster activities follow a domain-based assessment of accuracy, indicating promising robustness of this approach. Still, a full characterization of this framework's behavior is needed before it can be more broadly applied. While precision of clustering results is high in our model minicomposite system, it is unclear how precision will be affected when this framework is applied to more complex CMC architectures, where more damage mechanisms are active. Additionally, it is not known how specimen geometry, architecture, and scale influence AE features [151, 184]; this is an active area of research. Both of these questions will require testing large scale composites in conjunction with *in situ* microstructural observations for error quantification, and is the subject of future work. Finally, it is desirable to understand the drivers for observed trends in partial power to remove irrelevant frequency bands from the feature space and increase precision. Future investigations will aim to address this.

Table 1: **Validity metrics for Experiment 1**

Sensor pair	**Matching rate (%)**	**ARI**
B1025-a, B1025-b	96.4	0.78
B1025-a, S9225-a	92.1	0.62
B1025-b, S9225-b*	81.0	0.0
S9225-a, S9225-b*	73.0	0.0

* This test had a misaligned sensor on its larger epoxy tab wherein it did not fully overlay the minicomposite and is included to show that the ARI metric can be used to detect experimental issues such as these. The ARI of zero for sensor pairs that included sensor S9225-b are the result of this improper placement. (Chapter 3.2).

Table 2: **Validity metrics for Experiment 2**

Sensor pair	**Matching rate (%)**	**ARI**
B1025-a, B1025-b	87.1	0.42
B1025-a, S9225-a	82.9	0.31
B1025-b, S9225-b	86.1	0.45
S9225-a, S9225-b	84.0	0.40

Table 3: **Validity metrics for Experiment 3**

Sensor pair	**Matching rate (%)**	**ARI**
B1025-a, B1025-b	88.4	0.55
B1025-a, S9225-a	81.2	0.27
B1025-b, S9225-b	94.4	0.73
S9225-a, S9225-b	91.7	0.53

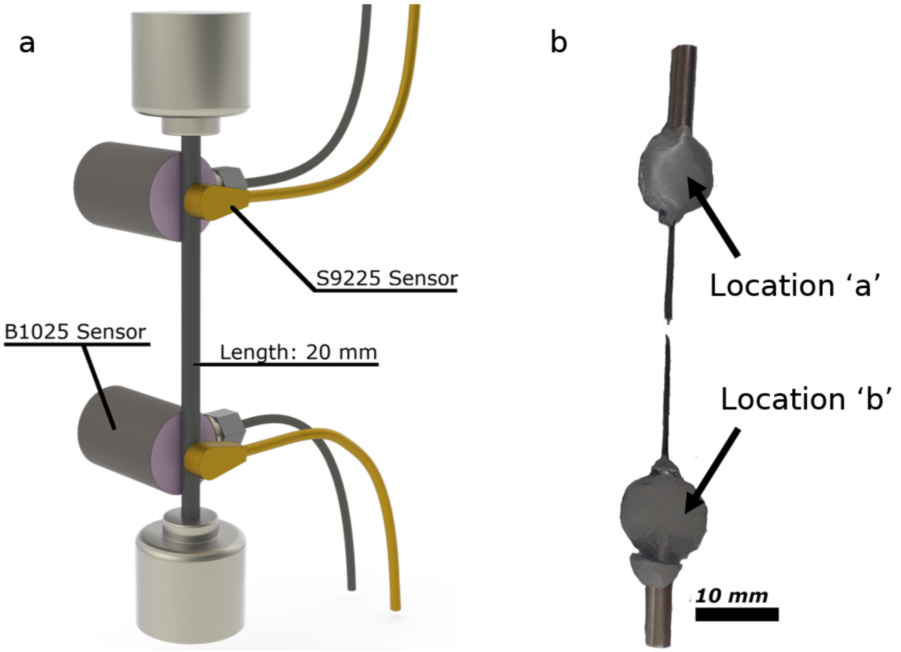

Figure 3.1: **Diagram of the experimental setup.** (a) Two sets of AE sensors were coupled to the minicomposite whose length is nominally 20mm. When any sensor was triggered, all sensors began recording, ensuring that each AE event could be correlated between all sensors. (b) A photograph of a sample. Epoxy tabs used to mount AE sensors are denoted with arrows.

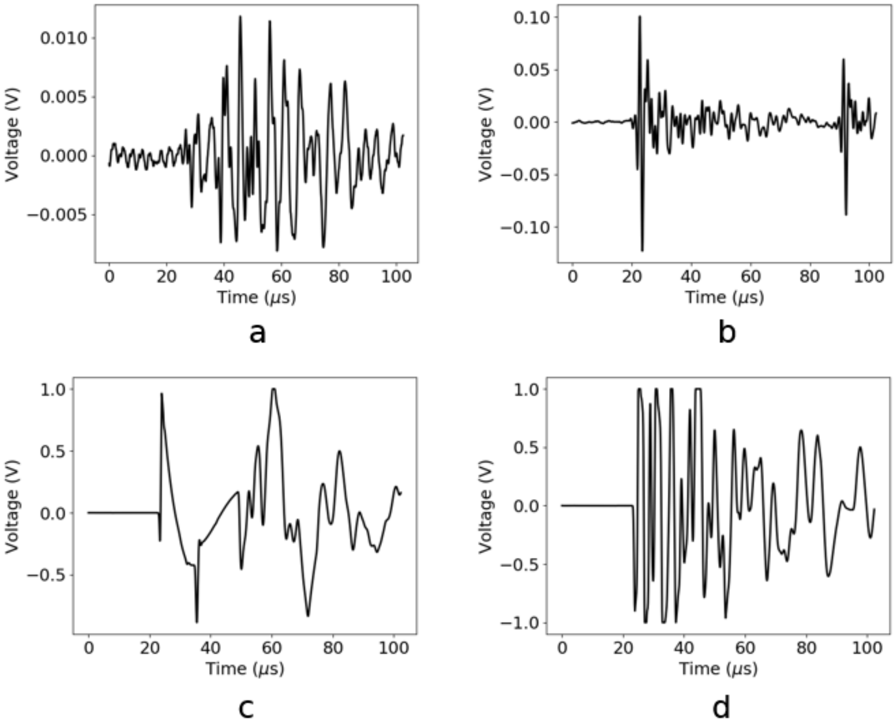

Figure 3.2: **Examples of removed waveforms.** (a) Low signal to noise ratio, (b) Two distinct damage events occurring in the same time window, and (c,d) events that saturated the AE sensor. The majority of removed waveforms were of type (a), (c), and (d). Less than 10 waveforms were of type (b) across the 3 specimens.

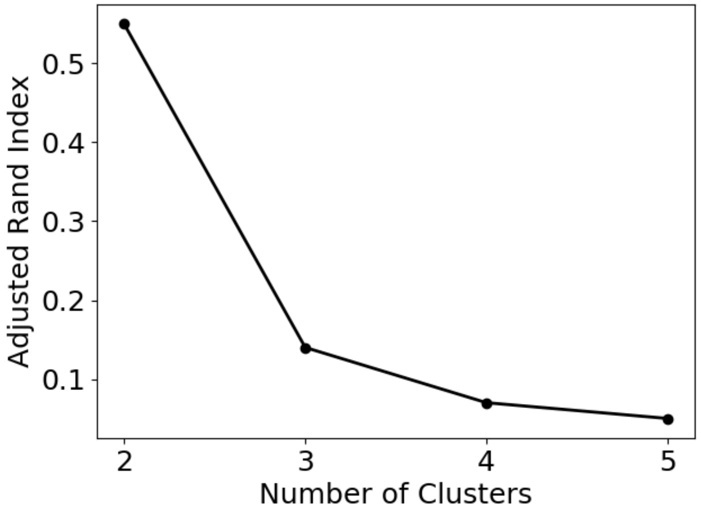

Figure 3.3: **Adjusted Rand Index as a function of the number of clusters.** When more than 2 clusters are used to initialize spectral clustering, the steep drop in ARI corresponds to a decrease in precision. As such, when more than 2 clusters are specified, the spectral clustering algorithm is forced to find clusters, which are not correlated with damage mechanisms. This drop occurs in all experiments and is corroborated by results shown in Figure 3.4.

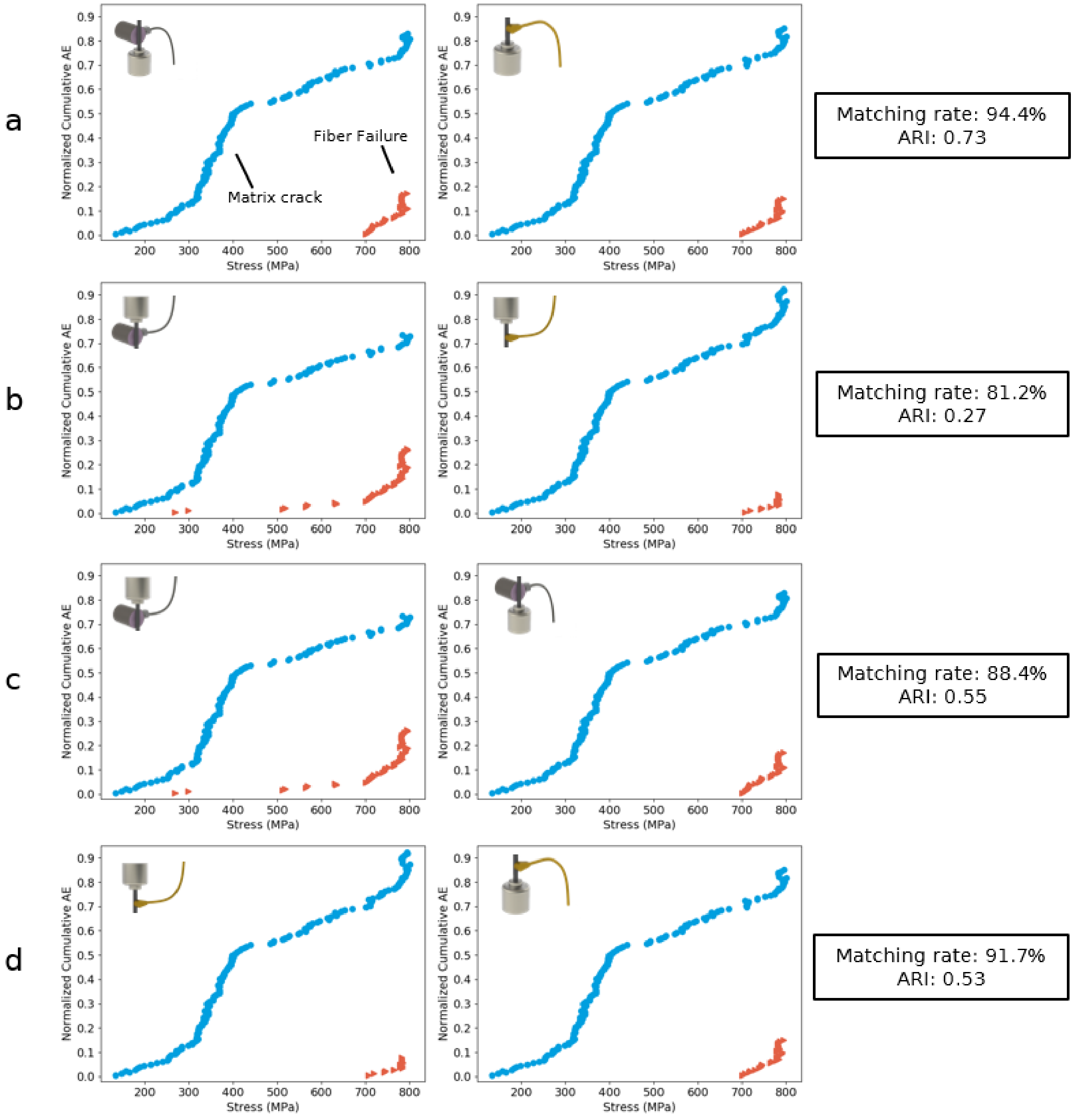

Figure 3.4: **Damage mechanism identification from AE signals.** AE waveforms were generated by SiC/SiC ceramic matrix composites (CMCs) loaded under uniaxial tension and recorded by four sensors; **(a)** B1025-b and S9225-b, **(b)** B1025-a and S9225-a, and at different locations **(c)** B1025-a and B1025-b and **(d)** S9225-a and S9225-b. The resultant identification of AE events closely follows CMC damage chronology, wherein early matrix cracking is later followed by fiber breaks. The cluster that becomes active at ≈ 85% of the UTS is labeled as fiber failure, consistent with experiment [12]. Cluster assignment of individual AE events was independent of both the sensor model and location of the sensors

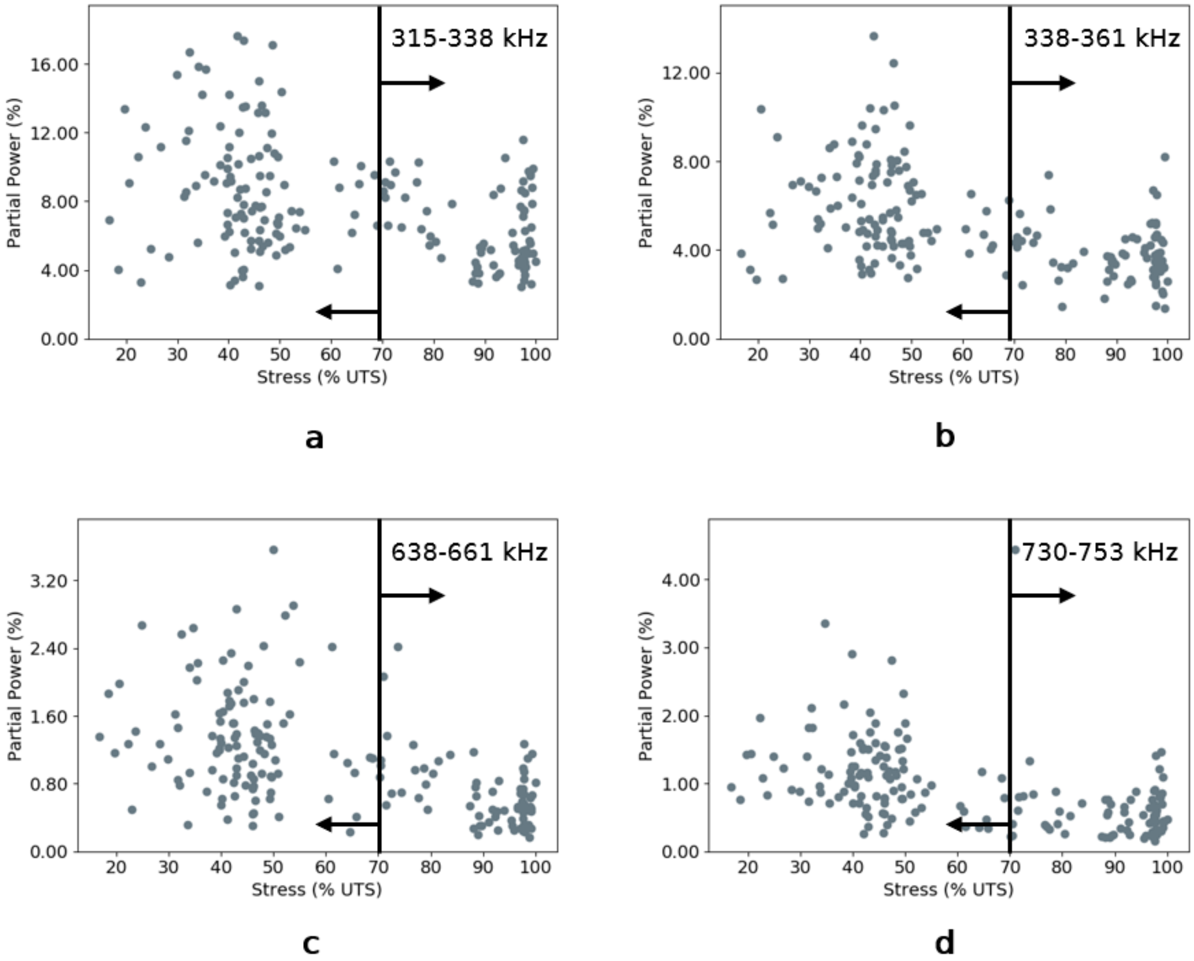

Figure 3.5: **The partial power for four selected frequency bands in Specimen 3.** Events occurring below 70% UTS are sampled from a different distribution than events occurring above 70% UTS at a significance level of $\alpha = 0.01$ in these bands. This is not predicted by the orthotropic model of wave propagation. We hypothesize that the shift in partial power is a result of a shift in active damage mechanism from matrix cracking to fiber failure; however, further experimentation and modeling is needed.

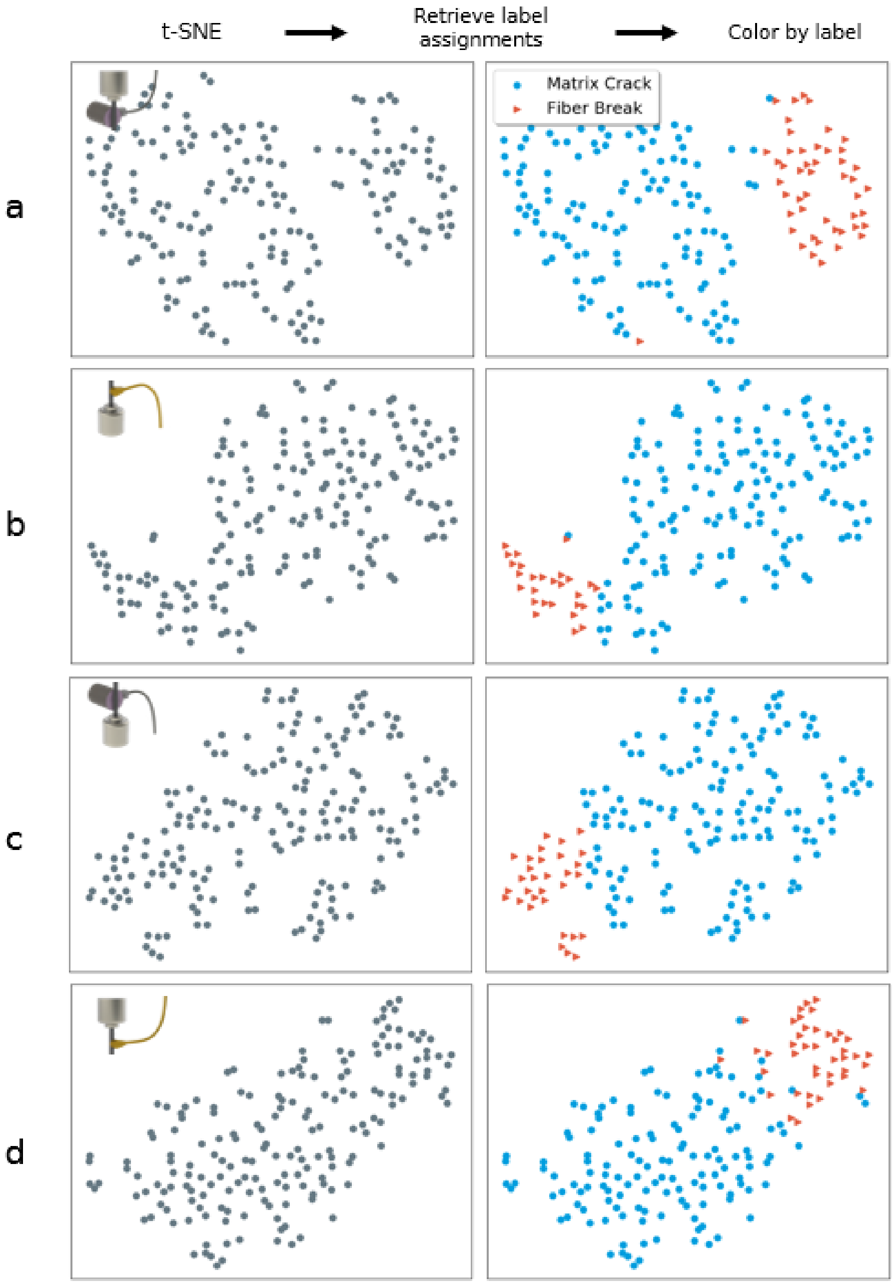

Figure 3.6: **Dimensionality reduction by t-SNE of Specimen 3 feature vectors.** Plots for the raw feature vectors are shown for each sensor on the left, and are subsequently colored on the right according to the labels given after spectral clustering as described in Figure 3.4. While the t-SNE axes have no intrinsic meaning, well–formed clusters indicate that partitions were made according to similarity between feature vectors and are not an artifact of the clustering routine.

Chapter 4

Quantitative benchmarking of acoustic emission machine learning frameworks for damage mechanism identification

4.1 Scope

In Chapter 1.4, a distinct lack of ground truth data sets was identified. This gap has precluded any quantitative benchmarking of framework accuracy and prevents optimizations for existing strategies AE-ML frameworks. In this chapter, I address this challenge by creating a set of experimentally generated signals which form a ground truth library. This ground truth data set is comprised of pencil lead breaks with known angles of incidence, θ, where each angle generates a unique frequency spectrum that changes continuously with θ. which could be analogous to attributes of acoustic emission signals generated from failure processes, such as those that occur in composites. Five frameworks are then applied to the ground truth dataset and benchmarked according to their ability to discriminate between two sets of signals with a fixed $\Delta\theta$. A discussion of their performance as related to choice of features is given, and a set of guidelines for best-practices for feature selection and standardized practices are proposed.

4.2 Methods

4.2.1 Data Collection

All pencil lead breaks (PLBs) were conducted with Pentel 0.5 mm HB leads and a nominal free lead length of 4 mm [14]. A Pentel GraphGear 500 mechanical drafting pencil was fixed to a custom-built, displacement controlled, load frame (Figure 4.1). The load frame was composed of a rotational stage, which allowed for angle adjustments in increments of 2° (corresponding manual angle measurement error is $\frac{1}{2}$ the unit of measurement, or $\pm 1°$), and two precision-adjust linear stages. The aluminum plate on which the PLBs were conducted had an unsupported span of 200.7 mm, width of 51.0 mm, and thickness of 1.2 mm.

PLBs were recorded at $20°, 22°, 26°, 30°$, $36°$, and $40°$. PLBs were generated by lowering the pencil via the linear Z stage until the lead fractured on the aluminum plate. For each angular condition, the rotational stage was fixed using a set screw and the set screw was loosened only to change angles. During an angle change, the linear X stage was adjusted to maintain a nominal distance of 25 mm from the PLB source to the sensor. Upon inspection of collected AE signals, some were found to be reflections. These reflections presented themselves as a second, low amplitude signal occurring immediately after the initial PLB, and were excluded from the data set. Due to the exclusion of these reflections, each angle has a differing number of signals, between 75-111. For the purposes of this analysis, only the first 75 signals of each angle were clustered by an AE-ML framework.

AE were recorded using a piezoelectric B1025 transducer (Digital Wave Corporation, Centennial, CO) with a broadband response of 50-2000 kHz (Figure 4.1). The threshold voltage was set to 0.1 V, the number of pre-trigger points was set to 256, and the total length of signal captured was 1024 points at a rate of 10 MHz. The sensor was fixed to the aluminum plate with an alligator clamp using vacuum grease as a coupling agent. The sensor was not remounted at any point during the experiment, meaning PLBs at all angles were conducted for a fixed sensor coupling. The authors note here that because the coupling is unchanging during all data collection, an absolute calibration of the sensor as described in [185] is not necessary. Additionally, all signals were collected within a single 3-hour span in a temperature controlled laboratory where environmental factors which might otherwise affect the absolute sensor calibration, such as temperature, were assumed to be unchanging. Unsupervised frameworks group signals according to differences in signal features, rather than the absolute values of those features. Since the absolute sensor calibration is unchanging, any differences in signal features can be attributed changes in the angle of incidence.

Signals collected at the reference angle $\theta_0 = 20^\circ$ and signals at a single benchmarking angle $\theta_b \in [22^\circ, 26^\circ, 30^\circ, 36^\circ, 40^\circ]$ were clustered using each of the frameworks described in Chapter 4.3, and relative discriminating power was assessed quantitatively using the procedure described below.

4.2.2 Quantitative Benchmarking

The permutation model of the adjusted Rand index (ARI) was used to benchmark frameworks. The ARI, which ranges from 0 to 1, measures accuracy of ML-calculated clusters as compared to the ground truth in a label-agnostic way. It compares the membership similarity of objects in the ML-calculated clustering, A, to the membership similarity of objects in the ground-truth clustering, B, and assigns a higher number if similarities are high [132]. In the context of this work, signals from θ_0 and θ_b are fed to an AE-ML framework. These signals are then assigned a label by the framework, either 0 or 1, depending on if the framework believes the signal should be grouped with θ_0 or θ_b. The ML-assigned label of each signal is then compared to the ground truth, the known angle at which the signal was collected. If the membership similarity of the ML-assigned labels and the true angles are similar, then the ARI will take on higher values.

The ARI is an adjusted-for-chance version of the Rand index (RI) and is calculated as [132]:

$$RI(A, B) = \frac{N_{11} + N_{00}}{\binom{N}{2}} \tag{4.1}$$

where N is the number of signals, N_{11} is the number of signal pairs which are grouped into the same cluster in A and B, and N_{00} is the number of signal pairs that are grouped into different clusters in both A and B. The ARI can then be calculated as [6, 134]:

$$ARI(A,B) = \frac{RI(A,B) - \mathbb{E}[RI(A,B)]}{\max[RI(A,B)] - \mathbb{E}[RI(A,B)]} \tag{4.2}$$

where $\mathbb{E}[RI(A,B)]$ is the expected value of the RI under a random model. The ARI is bound between 0 and 1, with 0 corresponding to random label assignments and 1 corresponding to perfectly matching labels. While many cluster similarity metrics exist, the ARI was chosen to compare partitions because it can be calculated for any number of clusters (provided the number of clusters in each partition is equal), and it accommodates unbalanced cluster sizes [131, 134]. Signals from θ_0 and θ_b were clustered by each framework. The value of $\Delta\theta = \theta_b - \theta_0$ at which the ARI vanished represents the point at which the framework has lost all discriminating power and is unable to identify differences between two groups of different signals.

4.3 Framework Descriptions and Accuracy Metrics

A general AE-ML framework pipeline has been described in Chapter 1.3, schematically shown in Figure 1.2. Following data collection (Figure 1.2a), the most important step in the framework is the selection of the feature set (Figure 1.2b). Waveforms can only be sorted according to their source mechanism *if* the feature set captures something fundamental about the waveform-mechanism relationship. Features can be classified as belonging to the time domain, frequency domain, or time-frequency domain. However, there is little consensus as to which category is best suited for damage mechanism identification. In fact, even when two frameworks leverage features within the same domain, their feature sets differ. Consequently, each framework uses a unique feature set, where d pertinent features are identified, extracted, and stored as a feature vector $\mathbf{v} \in \mathbb{R}^d$ (Figure 1.2b). The reader is referred to the original investigations for discussions on why particular feature sets were chosen [2, 49, 86, 186, 187].

Next, individual features of a feature vector may be re-scaled or re-mapped with a transformation (Figure 1.2c). Similar to the variations in feature sets, each framework utilizes a different set of pre-processing steps. Finally, the ML algorithm is applied to partition feature vectors by assigning them to clusters, where feature vectors in the same cluster are proximal under a chosen distance metric (Figure 1.2d).

The frameworks described in the following sub-sections follow this workflow and were adopted *directly* from literature. They were chosen to span the current space of diverse feature set types and ML algorithms [188]. The key differences between frameworks are the choice of feature sets, preprocessing steps, and ML algorithm. The specific choice of feature set, pre-processing steps, and ML algorithm are further summarized in Table 1. In Chapter 4.4, we provide key findings and discuss the impact of feature selection.

4.3.1 Base Framework

We define a Base Framework relative to subsequent frameworks, which are variations on this base (either by swapping out the feature set, ML algorithm or both). This framework employs a time-domain feature set as investigated in [86]:

1. average frequency (number of counts/signal length)
2. rise frequency (average frequency from signal start to maximum amplitude)
3. ln(energy)
4. ln(rise time/duration)
5. ln(amplitude/rise)

6. ln(amplitude/decay time)

7. ln(amplitude/average frequency)

The start and end time of an experimental signal was determined by the first and last crossing of a floating 10% voltage threshold.

Each feature was scaled by the maximum observed value of that feature, using the MaxAbsScaler method in [106], such that they fell in the range [-1,1]. A principal component analysis (PCA) transformation was performed, and principal components containing $\geq 95\%$ of the variance were retained. Distances, d, between any two feature vectors, $\mathbf{x}, \mathbf{y}$ were calculated using a modified Euclidean metric:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_i \lambda_i (x_i - y_i)^2} \tag{4.3}$$

where x_i and y_i are the ith vector components of the feature vectors in the PCA basis, and λ_i is the eigenvalue of the i^{th} PCA axis. As the Scikit-learn implementation of k-means enforced the standard Euclidean metric, a rescaling of feature vectors was required to accommodate the modified Euclidean metric:

$$x_i' = \sqrt{\lambda_i} x_i \tag{4.4}$$

It should be noted that the distance metric in Eqn. 2 differs from the standard PCA whitening approach, where distances along axes with large eigenvalues are contracted, rather than elongated [78].

K-means was then applied to the feature vectors. For a detailed description of the k-means algorithm, the reader is referred to [78]. Because k-means is not guaranteed to converge to an optimum solution, it is typically run many times and the initialization with the lowest value of loss function is taken [5]. To determine the number of re-initializations needed convergence checks were performed by increasing the number of re-initializations until the loss remained unchanged. The minimum objective function did not change after 2×10^3 re-initializations. To conservatively ensure a global optimum of the objective function had been reached, the number of re-initializations was set to 2×10^4. Similarly, an error tolerance of 0.0001 and 300 iterations were sufficient to ensure a local optimum was reached within a single initialization of k-means.

4.3.2 Agglomerative Framework

The agglomerative framework [186] used the feature set:

1. amplitude
2. peak frequency

Rather than partitioning feature vectors by k-means as in the Base framework, the Agglomerative framework uses a hierarchical agglomerative approach. In this approach, each data point is initially defined as a cluster. Clusters are then iteratively merged such that the chosen objective function (usually the sum of squared distances) is extremized. For a discussion of this algorithm, the reader is referred to [106, 186].

The linkage type, the parameter defining pairwise distances between points, was not reported in the original work. Here, each linkage type was tested and no linkage type outperformed another.

4.3.3 Spectral Framework

The Spectral framework [187] used the partial power feature set. The i^{th} component of the feature vector is:

$$x_i = \frac{\int_{k_{i-1}}^{k_i} F[s(t)]dk}{\int_{k_0}^{k_d} F[s(t)]dk} \tag{4.5}$$

where $F[*]$ is the Fourier transform operator, $s(t)$ is the recorded signal, k_i and k_{i-1} are the frequency bounds over which integration is performed, and d is the number of entries in the feature vector. We set $k_0 = 200$ kHz, $k_d = 800$ kHz, and $d = 23$. The width of integration bounds, $k_i - k_{i-1}$ was set to be equal for all i as in [187].

The sklearn implementation of spectral clustering was used to cluster the feature vectors [106]. A detailed explanation of the algorithm can be found in [76]. The ARPACK eigensolver was used and the number of nearest neighbors was set to $NN = 5$. To ensure cluster membership did not depend on initialization parameters, convergence checks were performed for error tolerance and maximum number of iterations. The cluster membership stabilized after 10 re-initializations. To conservatively ensure stability, the number of re-initializations was set to 100.

4.3.4 Frequency Framework

The frequency framework used a feature set in the *frequency* domain [49]:

1. average frequency
2. reverberation frequency
3. rise frequency
4. peak frequency

5. frequency centroid
6. weighted peak frequency
7. partial powers from 0-150 kHz, 150-300 kHz, 450-600 kHz, 600-900 kHz, and 900-1200 kHz

Features were independently normalized with the variance scaler, which centers features to have zero mean and scales them to unit variance. Feature vectors were then clustered with k-means. The same convergence checks as the Base framework were conducted, and identical parameters were sufficient for convergence.

4.3.5 WPT Framework

The wavelet packet transform (WPT) framework extracted features through application of a WPT [2]. Waveforms were subjected to a WPT on three levels using the Daubechies wavelet of order 2 as the mother wavelet. Fractional energies carried in each node were calculated, and the five least correlated values were retained. These, in addition to the waveform energies read by the AE software, were used as features. Feature vectors were then normalized with the maximum value scaler and subjected to PCA. Principal components containing $\geq 95\%$ of the variance were retained. The feature vectors were then partitioned via k-means, using the modified Euclidean metric (Eqn. 4.3).

Convergence checks were conducted and parameters identical to the Base framework were sufficient for convergence. It should be noted that the algorithm described by [120] and used by [2] is k-means *optimized* by a genetic algorithm. Thus, a fully converged k-means solution will not differ from a fully converged genetic solution.

4.4 Results and Discussion

The frequency content of PLB signals from our experimental configuration were able to be precisely controlled by varying the angle of incidence. Signals generated followed expectations from plate-wave theory [13, 14]; as the angle of incidence increased, the low-frequency components of the signal were observed to increase in power Figure 4.3. Moreover, little variation in PLB signals was observed within each angular condition. The mean signal and its standard deviation envelope are shown in Figure 4.4. As a result of the small variation between signals for a single angular condition, when signals are represented in feature space (Figure 4.2b), the standard deviation of those features is smaller than if signals had a large variability. If a benchmark set were constructed from more repeatable acoustic signals, it would be expected that signals at different angles would form tighter clusters in feature space, and subsequently the ARI of each framework at each value of $\Delta\theta$ would increase.

By comparing ARI values at a fixed value of $\Delta\theta$, it is possible to quantitatively evaluate the relative discriminating power of various frameworks. The frameworks listed in Table 1 were applied to group signals according to the procedure described in Chapter 4.2.2. The accuracy of each framework was plotted as a function of $\Delta\theta$ (Figure 4.5). The discriminating power of each framework increased with $\Delta\theta$ which can be attributed to increasing differences in the signal structure as a function of $\Delta\theta$. Frameworks exhibiting higher ARI values at lower values of $\Delta\theta$, such as the Spectral and Frequency frameworks, have higher overall discriminating power, and will likely be able to distinguish between damage mechanisms that emit similar signals.

4.4.1 Saliency of Features

Across the five frameworks no feature set was shared. While this is common in the broader context of AE-ML frameworks [188, 189], the lack of consensus raises an important question: ”What features *should* be used for the purposes of AE signal discrimination?”. Addressing this question is of utmost importance, since the discriminating power of a framework hinges on how signals are represented [81, 188].

For signal discrimination, both exclusion of noisy features and inclusion of useful features is necessary: a principle known as the Ugly Duckling theorem [81]. To highlight the degree to which this principle impacts discriminating power, features were parametrically excluded from the Frequency framework and Base framework. In the Frequency framework, ARI was maximized when clustering was performed using average frequency, rise frequency, and partial power from 150-300 kHz. When these three optimal features were used to encode signals, the ARI at $\Delta\theta = 2^\circ$ increased from 0.681 to 0.973, representing a change from modest to high discriminating power. A similar procedure was conducted for the Base framework, and when only the average frequency and log(amplitude/average frequency) were included, the ARI increased from 0.325 to 0.82.

While such parametric studies can yield insight into which features are useful for a specific dataset, they are less effective in identifying universally salient features. For this purpose, it is necessary to consider the physics of the emitting source on a case-by-case basis and when possible, exclude features that are dependent on external and uncontrollable factors unrelated to the source mechanism. For example, although amplitude is correlated to the angle in this dataset (Figure 4.6a), it should not be used for sorting signals from multi-phase materials because it is convolved with factors, such as crack area formed and the source to sensor distance, which are unrelated to individual mechanisms [190]. Similarly, even though rise time is commonly used in AE-ML frameworks, [87, 108, 110, 120, 139, 159], it is more strongly related to the source to sensor distance, due to the different velocities of the flexural and extensional wave components [34]. Consequently, two signals emitted from similar locations will have similar rise times, even if the emitting mechanisms are different (Figure 4.6b).

4.4.2 Limitations of the PLB Dataset

Intuitively, a framework with higher ARI values is a promising candidate for damage mechanism identification when signals from different mechanisms are expected to be similar. However, the degree to which performance on the PLB dataset translates to performance under more realistic conditions and material systems is unknown. Specifically, the PLB signals in this dataset are collected under the strictest possible conditions; signals are from a single source-to-sensor distance, sensor coupling, and source type (e.g. pencil lead), removing the effect of factors that influence a signal such as dispersion, attenuation, and absolute frequency response. ML approaches for mechanism identification must ultimately be robust to these effects. Although this dataset represents a first step towards quantitative benchmarking, a full characterization of framework performance under realistic conditions is still required.

Another limitation of the dataset we have collected is the angular resolution; the $\pm 1^\circ$ tolerance of the rotational stage has implications on the measured ARI. For example, if the true $\Delta\theta$ between two angular conditions was less than the reported value, due to the $\pm 1^\circ$ tolerance, signals generated at these angles would be more similar than expected. Consequently, the ARI measured would be lower than if the signals had been collected from a true angular condition with a larger value of $\Delta\theta$. The exact degree to which the ARI would change is highly dependent on how each feature varies with θ, and subject to any data dependent pre-processing, such as PCA, which would further impact framework performance.

Finally, also due to the angular resolution, the current experimental setup prevents collecting signals from values of $\Delta\theta < 2°$. In order to properly benchmark frameworks, there must be at least one value of $\Delta\theta$ where ARIs are not saturated at 1. For example, for $\Delta\theta = 20$, the Spectral, Frequency, and Base frameworks all perform equally well, but $\Delta\theta = 6$ allows for comparison of discriminating power (Figure 4.5). As the community continues to improve the discriminating power of frameworks, ARI values will increase. Consequently, it will become necessary to collect signals from values of $\Delta\theta < 2°$, below what we have allowed for in this study, to ensure frameworks' performances can be separated.

4.5 Conclusions

While many AE-ML frameworks have been developed and implemented, the lack of ground-truth datasets has restricted discussions of their strengths and limitations, particularly with respect to feature choice, and has prevented development of standardized quantitative benchmarking procedures. In this chapter, a method for quantitative benchmarking procedures has been proposed, and used to analyze the effectiveness of existing AE-ML frameworks. Below, considerations for the quantity of data in benchmarking sets, types of features that should be included in a framework, and transparent benchmarking practices are discussed.

The performance of an unsupervised framework is intrinsically tied to how well the sampled data represents its population distribution. In the context of AE-ML benchmarking datasets, it is critical to ensure enough signals have been collected to capture statistical variations. If too few waveforms are collected at any angle, it is unlikely that the sampling distribution will represent the population distribution of waveform features

(Figure 4.2b). Consequently, the addition on new waveforms will lead to spurious performance of an AE-ML framework. To ensure enough signals are in a benchmarking set, a framework's performance must be shown to be independent of the number of signals collected. For this benchmarking set, it is demonstrated that 75 signals per angle are sufficient to ensure the ARI values we calculate are independent of the amount of data (Figure 4.7).

Feature selection is of critical importance with respect to maximizing the discriminating power. As demonstrated in Chapter 4.4.1, the inclusion of non-salient features was directly correlated with poor framework performance. Despite the importance of feature selection, there is little discussion within the literature as to why certain features were chosen [188]. As a result, many modern frameworks continue to include non-salient features (e.g. rise time) which negatively impact framework performance.

Towards better feature selection, universal features should be prioritized, and when possible, choice of feature set should be explicitly motivated. If it is possible to construct cases where a given feature cannot reliably discriminate between signals emitted by two unique sources, then the feature is likely convolved with factors unrelated to the source mechanism and are therefore not universal. The use of such non-universal features must be treated with caution. For example, although small amplitudes and large rise times have been correlated with Mode II cracks, these features are not universal because they are also a strong function of the source-to-sensor distance [72]. This makes it possible to construct a dataset where unique signals can appear artificially similar resulting from little to no statistical difference between features (Figure 4.6).

Although universally salient features will not change between material systems or loading configurations, the values of the features might vary. For example, partial power appears to be a universally salient feature [49, 159, 187], but every frequency band does not provide equal discriminating power. As demonstrated by the parametric removal of

Frequency framework features, the frequency band from 150-300 kHz was the most useful for signal discrimination. In this case, 150-300 kHz was useful for discrimination between two PLB signals, however different frequency bands will be useful when the material system or loading conditions changes [67, 151].

Finally, publicly available standardized datasets should be used for quantitative benchmarking of frameworks. Although these types of benchmarking tools are common in other fields [191–194], they are absent in the AE community. Development and continued maintenance of benchmarking datasets will provide the tools necessary to assess the strengths and limitations of AE-ML frameworks and will allow for detailed discussions regarding the specific strengths and weaknesses. In turn, this will provide transparency and trust in the results obtained from such frameworks, promoting their broader use in both scientific and engineering applications.

Table 1: **Investigated framework summaries**

Framework	**Feature set**	**ML algorithm**
Base [86]	Time-domain	k-means
Agglomerative [186]	Mixed	Agglomerative
Spectral [187]	Frequency-domain	Spectral clustering
Frequency [49]	Frequency-domain	k-means
WPT [2]	WPT Energies	k-means

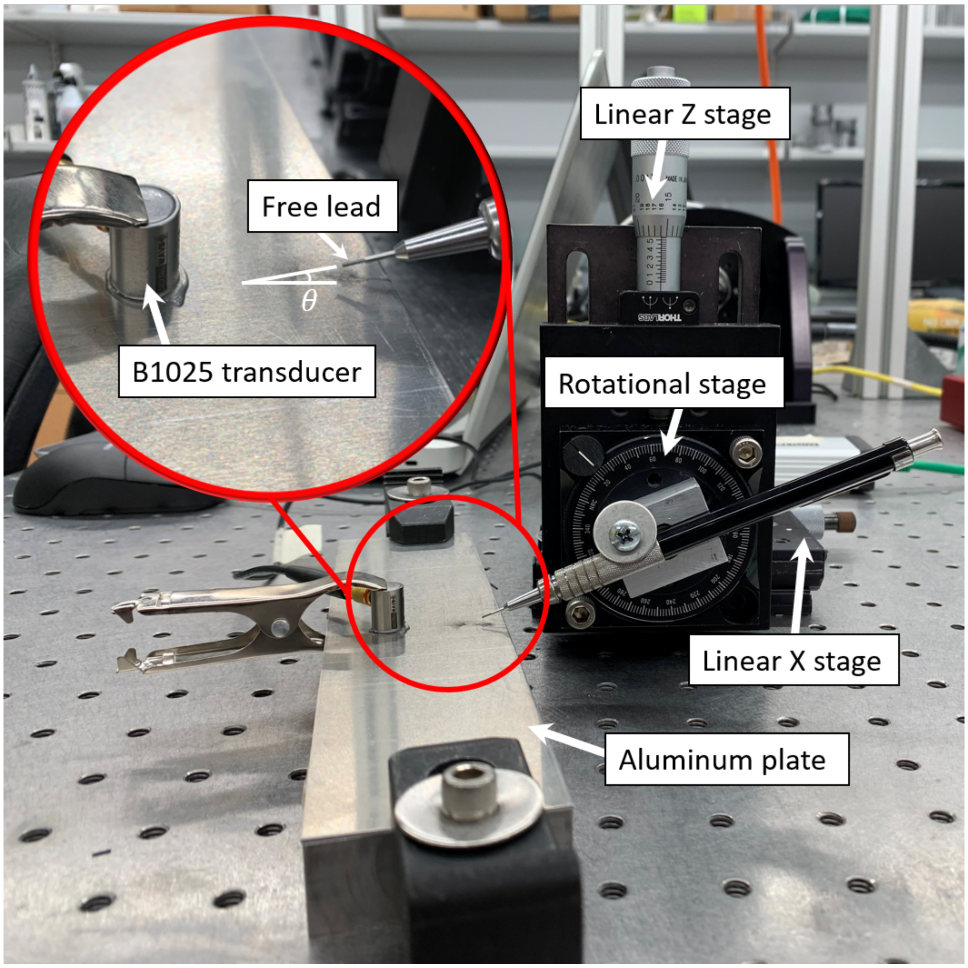

Figure 4.1: **Photograph of experimental setup.** A mechanical pencil is attached to a rotational stage which controls the angle of incidence θ. The linear X stage is used to position the tip of free lead at a consistent location on the aluminum plate. The linear Z stage is used to lower the pencil lead until fracture. The resultant waveform generated is recorded by the piezoelectric B1025 transducer, located approximately 25mm from the tip of free lead.

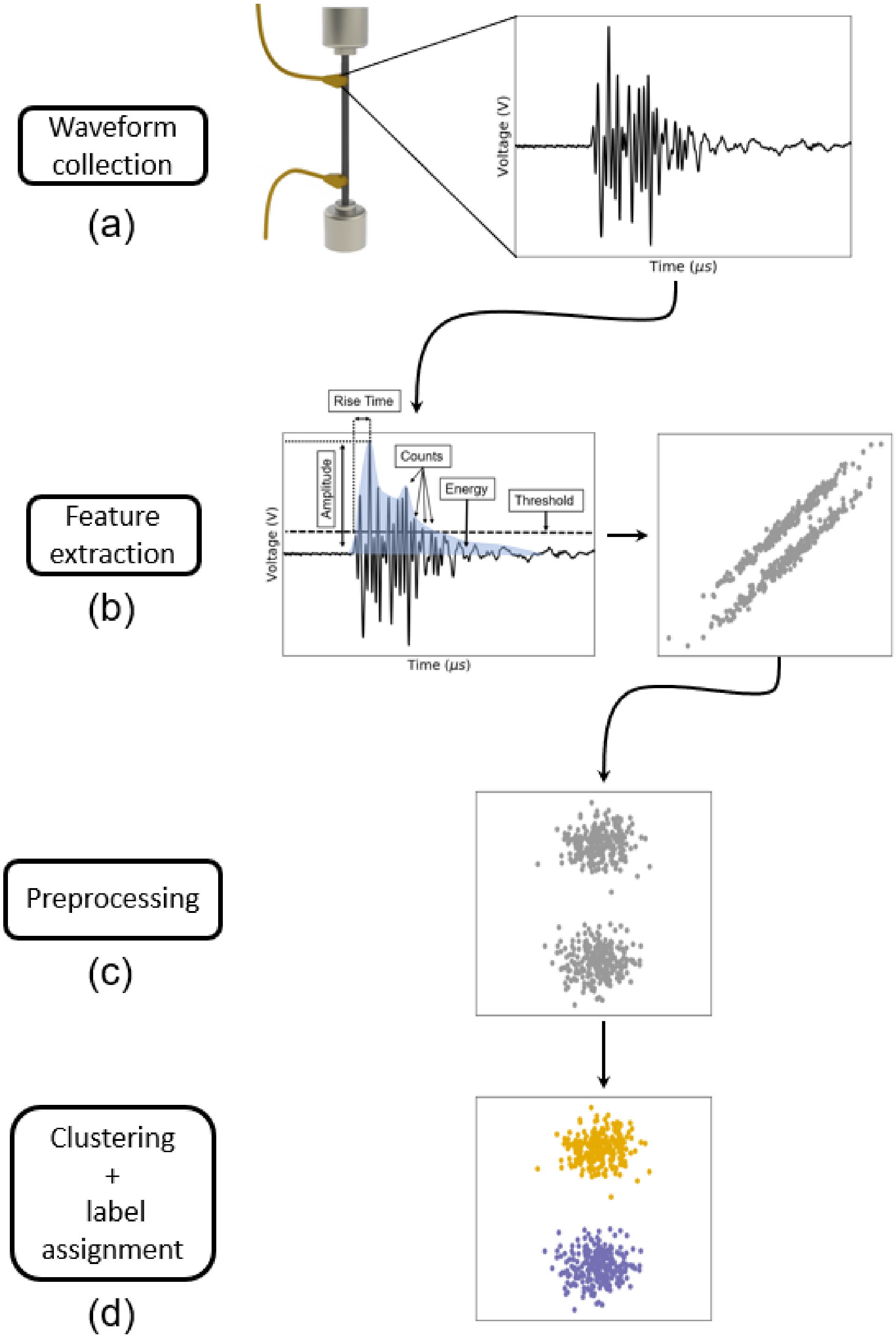

Figure 4.2: **Workflow diagram of an AE-ML framework.** (a) Waveforms are collected and (b) pertinent features are extracted from the waveforms, which are then represented as vectors in feature space. (c) Feature vectors can then be re-scaled and/or re-mapped before (d) the clustering algorithm is applied and feature vectors are labeled. Every AE-ML framework follows this procedure.

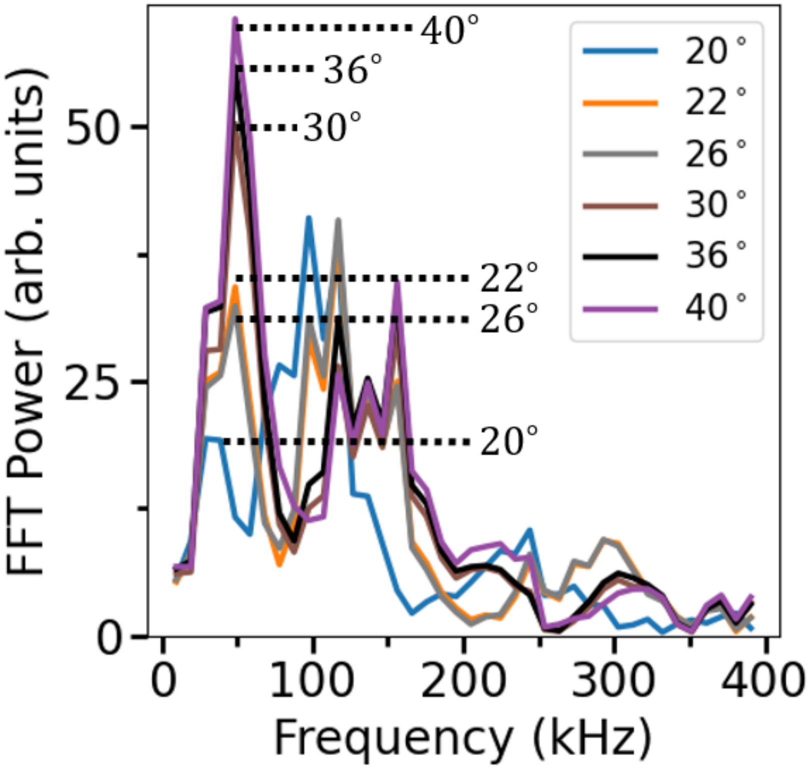

Figure 4.3: **Fourier transform (FFT) of average signals at each angular condition**. As the angle of incidence increases, the low frequency components increase in power following the findings of [13] and [14].

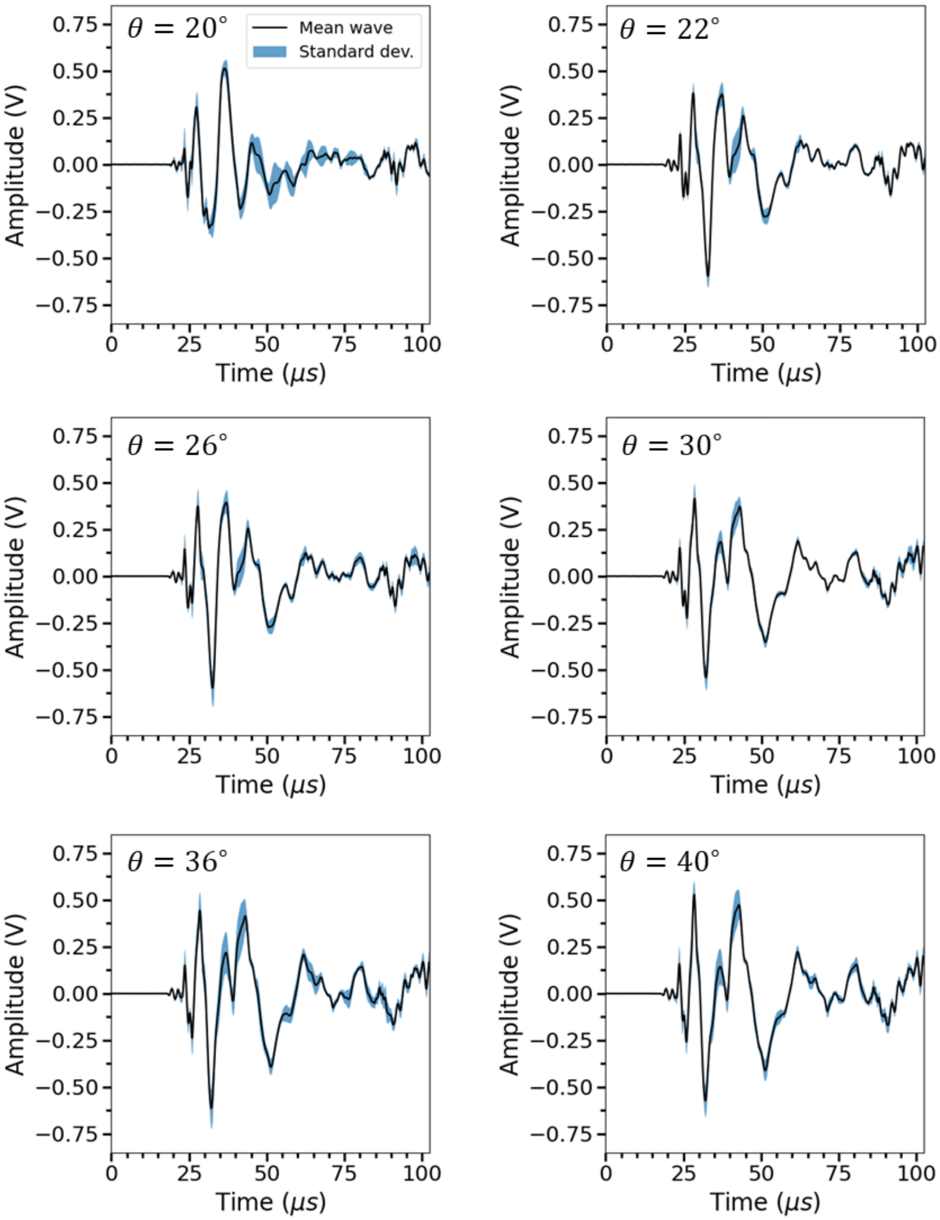

Figure 4.4: **The mean PLB signal and point-wise standard deviation at each angular condition.** Signals generated using the experimental fixture shown in Figure 4.1 were repeatable, while still containing variation that might be expected from signals collected during *in operando* health monitoring.

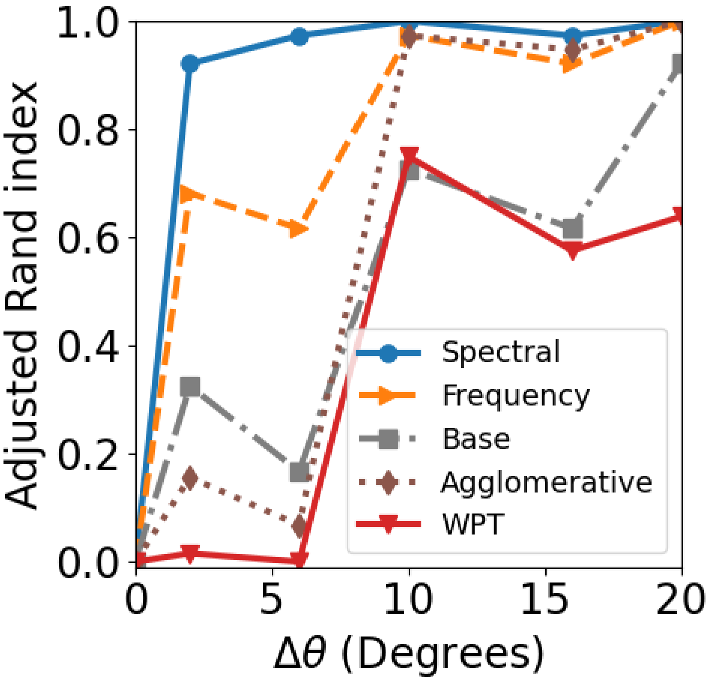

Figure 4.5: **The ARI of each framework as a function of $\Delta\theta$.** ARI values exceeding 0.4 are correlated with good discriminating power, whereas values near 0 correspond to no discriminating power. The discriminating power of each framework increases with $\Delta\theta$ and high ARIs at low values of $\Delta\theta$ are better suited for clustering signals whose differences are minor. The ability to directly compare accuracy between frameworks allows researchers to choose an appropriate framework for their specific needs.

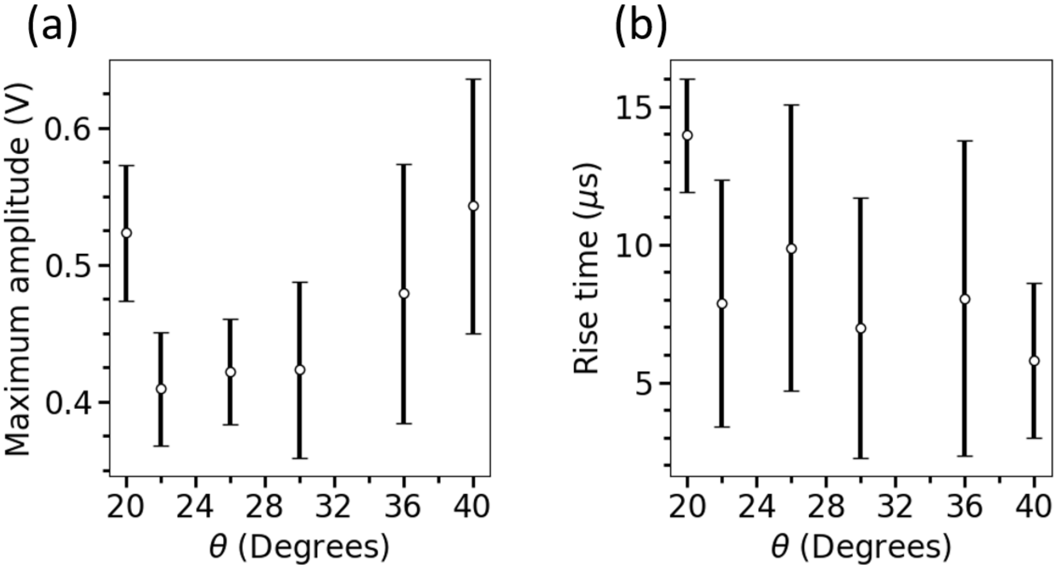

Figure 4.6: **The (a) average maximum amplitude and (b) average rise time of signals generated at each angle of incidence θ.** Error bars correspond to 1 standard deviation. There is no consistent difference between values in either feature. Because it is possible to construct many sets of unique signals with indistinguishable amplitudes and rise times, they should not be considered salient features and their use should be taken with caution.

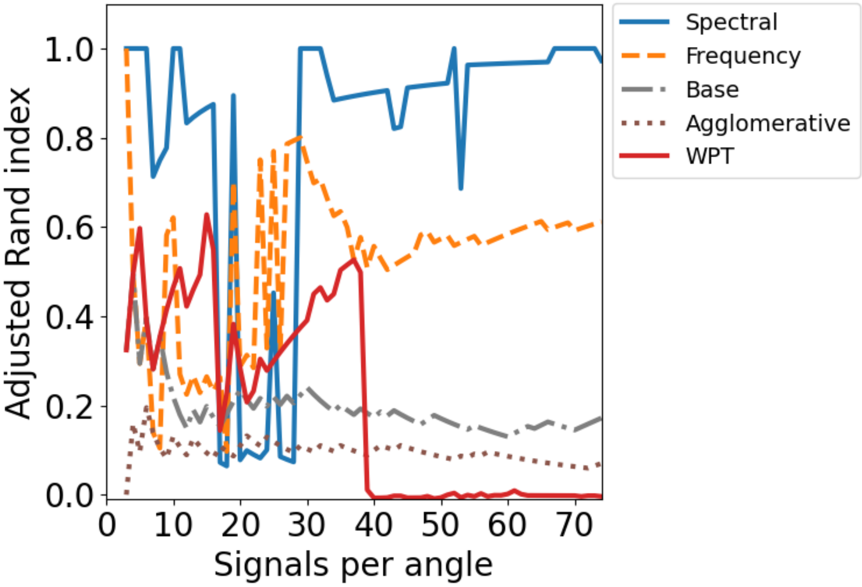

Figure 4.7: **Adjusted Rand index vs. number of signals per angle.** Signals from $\theta_0 = 20°$ and $\theta_b = 26°$ ($\Delta\theta_0 = 6°$) were clustered using an increasing number of signals per angle. As the number of signals increased, the performance of frameworks becomes independent to the addition of new signals, indicating enough data is present to capture stochastic waveform variations.

Chapter 5

Supervised Damage Mechanism Identification in SiC/SiC Composites

5.1 Scope

The work in Chapter 4 is a first step in developing the ability to benchmark AE-ML frameworks. However, the usefulness of this benchmarking set is limited because acoustic signals generated from PLBs are unlikely to share all characteristics with acoustic signals generated from damage in brittle composites like SiC/SiC. This presents two problems: the first is that while the library in Chapter 4 can be used to evaluate how well AE-ML frameworks perform relative to one another, it cannot be used to evaluate how well a framework will perform during real-time monitoring. This is because the benchmarking set was captured in an environment which controls for a large number of factors seen by in-service parts, such as external noise. The second problem is that this library cannot be used to train supervised ML damage identification models. Any model trained using the PLB library will learn characteristics of acoustic signals emitted by PLBs rather than damage mechanisms.

As a result of these shortcomings, ground truth libraries composed of signals from real damage are needed. However, gathering and curating these libraries is prohibitively expensive in time and money. Few optical methods can visualize subsurface damage, and none of them can do so with the necessary time resolution. For example, laboratory scale x-ray computed tomography studies may take as long as four days to acquire a single scan. When two or more scans are needed, acquiring the necessary instrument time quickly becomes intractable. The use of synchrotron light sources, which can scan composites at much higher rates, but are even more expensive and less accessible. As of the writing of this dissertation, there is only one facility equipped for *in-situ* AE mechanical testing [30, 195], and length of time needed to acquire the necessary beam time precludes its use.

This chapter presents and discusses initial work to gather labeled acoustic signals emitted during damage in CMCs to address these difficulties in library curation. SiC/SiC minicomposites are subjected to uniaxial tension and imaged via x-ray computed tomography. Signals are separated into two sets, based on general damage mechanism trends that are observed by XCT. An autoencoder architecture is then leveraged to perform anomaly detection to clean these two sets by identifying outliers in each group. Error quantification is performed via 100 repeats of 5-fold CV, and it is found that fiber break signals have unusually low reconstruction errors, indicating a high degree of homogeneity within these signals. This behavior is contextualized by the composite microstructure, lends support to hypotheses developed in Chapter 3, and motivates choices for future AE-ML frameworks.

5.2 Experimental Methodology

5.2.1 Materials

SiC/SiC minicomposite specimens were manufactured by Rolls Royce High Temperature Composites (Cyprus, CA) and consisted of a single-tow, 500 Hi-Nicalon Type S™(Nippon Carbon, Tokyo, Japan) SiC fibers, coated with a chemical vapor infiltrated (CVI) boron nitride (BN) interphase and an overlayer of CVI SiC matrix. The tested system was from the low fiber content (LFC) specimens described in Chapter 2. The nominal cross-sectional area was $A_{total} = 0.253$ mm^2 with a fiber volume fraction (V_f), matrix volume fraction (V_m), and interphase volume fraction (V_i) of 25.4%, 72.1%, and 2.4 % respectively calculated using the same method as [33, 173]. This simplified composite system was chosen because its damage progression characteristics (e.g. proportional limit, onset of matrix crack saturation, activation of fiber failures) are well characterized [33, 188].

5.2.2 Tabbing and Acoustic Emission

Prior to tensile testing, collection of AE, and imaging, the minicomposite was tabbed using a custom 3D printed structure consisting of a reservoir and sensor housing (Figure 5.1a). The minicomposite was inserted into the hollow tabs through the slot in the housing. The reservoir was subsequently filled with Duralco 132 epoxy until the epoxy was level with the reservoir walls and allowed to cure for a minimum of 24 hours. The tabs and sample were placed in a laser-cut acrylic aligner during the curing process to ensure a 10mm gauge length and minimal off-axis loading. Two B1025 transducers (Digital Wave Corporation, Centennial, CO) with a broadband response of 50-2000 kHz were then affixed to the sample in the housing area by a layer of Crystal Bond 509 coating the sensor housing (Figure 5.1b). This allowed for the removal of sensor wiring between load and image steps, without changing how the sensor was coupled to the sample. Thus, differences in waveform frequency content are a result of differences in damage mechanism, and not a result of differences in sensor coupling [80]. AE activity was recorded using a synchronized four-channel fracture wave detector acquisition system (Digital Wave Corporation, Centennial, CO); when one sensor was triggered, the other sensor recorded waveform data simultaneously.

5.2.3 Tensile Testing and X-ray Computed Tomography Imaging

To accommodate acoustic wiring in the XCT, the minicomposite had to be loaded and imaged in a two-step process. First, the minicomposite was loaded to 564 MPa, 66.9% ultimate tensile stress (UTS), under uniaxial tension using a Kammrath and Weiss microtensile stage equipped with a 500N load cell (Figure 5.1b). The initial maximum load stress was chosen because it had previously been identified as the matrix crack saturation load [36], and was confirmed to be the point of matrix crack saturation by the stagnation of incoming AE signals (Figure **(cm)** fig). The cross-head displacement rate was 0.5 μm/s, corresponding to a nominal strain rate of 5×10^{-5} s^{-1}. The sample was then unloaded to 0 MPa at the same rate. Acoustic sensor wiring was detached, leaving sensors in place, and the sample was transferred into a Deben CT5000 load frame (Figure 5.1c).

The sample was then loaded to 475 MPa and imaged in a Zeiss Xradia 520 Versa (Carl Zeiss AG, Oberkochen, Baden-Württemberg, Germany) at the Air Force Research Laboratory. Projections were taken with a 115s exposure time, and pixel size of 1.16 μm over a 1mm x 1mm field of view (FOV) (Figure 5.1d). Three FOVs were stitched together along the gauge length of the minicomposite, yielding an imaged section just exceeding 2.5 mm in length (due to overlap of FOVs). Collection took 3.5 days. Over the scan period, the sample relaxed from 475 MPa to 275 MPa. After imaging the minicomposite was transferred to the microtensile stage, acoustic wires were re-connected, and the sample was loaded to 731 MPa (86.8% UTS). No acoustic signals were recorded before 564 MPa, indicating that re-loading the specimen below its previously seen maximum stress does not induce more damage. This also serves as evidence that damage observed during imaging was not due to loading within the Deben CT5000. The stress-displacement curve for both load steps, and the final load-to-failure step is shown in Figure 5.3.

Scans were reconstructed using the Zeiss Scout-and-Scan software and segmented with Dragonfly ORS (Dragonfly 2022.2, Comet Technologies Canada Inc., Montreal, Canada). For each of the two scans, the deep learning toolkit in Dragonfly was leveraged to identify regions of matrix, fibers, matrix cracks, and fiber breaks (Figure 5.4). The segmented scans were then used to highlight individual fiber breaks, which were manually confirmed. Finally, a shared gauge section within both scans was identified and manually registered between the two load states. Fiber breaks in this shared section were counted, and are plotted as a function of position (Figure 5.5).

5.2.4 Acoustic Data Processing

After acquisition, the raw AE data was cleaned to remove waveforms with sufficiently low signal-to-noise ratios (energy < 0.001V), waveforms showing multiple damage events within the same time window, and signals which saturated the sensor according to the same procedure in Chapter 3.2.2. Then, a location analysis was performed to remove out-of-gauge events [145] using:

$$\text{location} = \frac{x}{2}\left(\frac{\Delta t}{\Delta t_x}\right) \tag{5.1}$$

where x is the sensor separation distance, Δt is the difference in time-of-arrival, and Δt_x is the difference in time-of-arrival for an out-of-gauge event as a function of the damage parameter. This analysis ensured that only signals arising from in-gauge damage events were analyzed. This cleaning removed approximately 20% of all recorded AE waveforms, including out-of-gauge events.

5.2.5 The Autoencoder

The autoencoder deep learning architecture is a neural network which acts as a surrogate for the identity function $I : \vec{x} \mapsto \vec{x}$ [75]. Typically, autoencoders consist of two elements: (1) The encoder function, $f : \vec{x} \mapsto f(\vec{x})$, and the decoder function, $g : f(\vec{x}) \mapsto \vec{x}$. During the training phase, the weights and biases of f and g are adjusted to minimize the reconstruction error, $L(\vec{x}, g(f(\vec{x})))$, where L is the chosen loss function. One of the more common autoencoder architectures is the undercomplete autoencoder, in which a fully connected feedforward network is broken into two parts: the encoder, f, which maps the input data to a latent space whose dimension is less than the input space, and the decoder g which maps the encoded representation back to the original input space (Fig. 5.6). By splitting the network this way, the encoder is able to represent the most salient features of the original high-dimensional input data in the compressed latent space, and the decoder is able to take the compressed representation back to its original form.

However, the ability to reliably compress and decompress is predicated on the assumption that the training data is representative of population distribution fed to the autoencoder during inference. If the autoencoder is shown test data which falls significantly outside the training distribution, it will fail to compress and reconstruct the test data with low error. This principle serves as the basis for their use as time-series anomaly detectors [196–200], and motivates their use for cleaning imperfect training libraries [201]. Here, we leverage this use of autoencoders to deconvolve matrix crack acoustic signals from fiber break acoustic signals. As discussed in Chapter 5.3, the acoustic signals obtained from load step 1 (564 MPa) are a mix of primarily matrix crack signals, with small quantities of fiber break signals. While using this relatively small and imperfect library to train a supervised classification model risks teaching the model incorrect attributes of matrix crack signals, it is possible to use the imperfect library to identify fiber break signals. Because fiber breaks are under-represented in acoustic signals collected from load step 1 (Chapter 5.3), they will have higher reconstruction errors compared to matrix cracks. The autoencoder anomaly detector architecture is schematically shown in Figure 5.6.

5.2.6 Training, Error Quantification, and Secondary Verification

Prior to training, each signal, $x_i(t)$ is normalized linearly such that the range of each signal falls between 0 and 1, this is equivalent to normalization by the MinMax scaler in Sklearn [6]. The signals are then split into two sets: the matrix set, comprised of signals gathered up to matrix crack saturation (load step 1, 0 to 564 MPa) and signals gathered after matrix crack saturation (load step 2, 564 to 731 MPa). These two sets will be referred to as the matrix set and the fiber set respectively, because the XCT observations indicate the acoustic signals in these sets are a result of the respective damage mechanisms. The autoencoder was trained on the matrix set for 500 epochs

(Figure 5.7a) using the Mean Absolute Error (MAE) loss function:

$$MAE = \frac{1}{n}\frac{1}{T}\sum_{i=1}^{n}\sum_{t=1}^{T}|x_i(t) - \hat{x_i}(t)| \tag{5.2}$$

where $x_i(t)$ is the value of signal i at time t, $\hat{x_i}(t)$ is the reconstruction of signal i, T is the total number of discrete time samples, and n is the number of signals. The formula in Eqn. 5.2 is the average point-wise absolute difference between the original signal and the reconstruction (Figure 5.7b). The Adam optimizer with default learning hyperparameters and a batch size of 32 was used for training. Training was ended at 500 epochs, a form of regularization known as early stopping; while the training loss continued to decrease past this point, the results were unchanged with longer training times (Figure 5.7a). To show the efficacy of signal reconstruction, an example waveform and reconstruction is shown in Figure 5.7b.

After training, the autoencoder can be used to compress and reconstruct signals in the fiber set. The MAE of a reconstruction is calculated, and the signal is labeled as an anomaly if the MAE exceeds a pre-set threshold. In order to choose this threshold appropriately, it is necessary to know the MAE distribution of signals seen during the training phase, i.e. matrix crack signals. Once this distribution is known, it can be reasonably assumed that values falling outside this distribution have not been seen during the training phase, and can therefore be considered anomalous, i.e. fiber break signals.

To estimate the MAE distribution of a matrix crack signal, 100 runs of 5-fold cross validation (5-fold CV) were performed on the matrix set. During a single run of 5-fold CV, the training dataset is shuffled, and split into 5 mutually exclusive folds. The autoencoder is then trained on four folds, with the fifth fold being left out to calculate the MAE. This process is then repeated, systematically holding out another fold each time [202]. Applying this procedure allows estimation of the reconstruction error distribution and helps inform the choice of threshold at which a signal is labeled an anomaly.

Once the error distribution has been estimated with 5-fold CV and a threshold is chosen, signals can be labeled. However, the fidelity of the labels must be checked. Although broad trends have been identified via XCT observations, the low temporal resolution of XCT prevents a more detailed analysis. Therefore, in order to assess the accuracy of the labels, each signal from the fiber set is also labeled by the spectral framework described in Chapter 3. The details of processing are identical to Chapter 3.2 with one exception, the frequency range used was lowered from 200-800 kHz to 50-550 kHz. This adjustment was necessary as the Crystal Bond layer (1-2mm thickness) between the sensor and sample attenuated the high frequency components of the signal. When this adjustment is made, the spectral-labeled signals followed the expected damage chronology (Figure 5.8), demonstrating that the correct hyperparameters were identified for this experimental configuration. The autoencoder labels were then compared to the spectral framework labels using the adjusted Rand index (described in Chapter 3.2).

5.3 Results and Discussion

5.3.1 X-ray Computed Tomography and Acoustic Emission

Acoustic signals during the first and second load step were correlated via XCT observations to matrix cracking and fiber break damage mechanisms, respectively. After initial loading to matrix crack saturation (564 MPa) the minicomposite was scanned over three 1×1mm FOVs. Due to a FOV stitching failure at this step by the automated program during the first scan, for this scan the FOVs had to be manually registered to the scan from the next load step. In order to minimize error due to manual registration, fiber breaks from 1 FOV were compared over two load steps (Figure 5.5). Up to matrix crack saturation, the majority of damage to the minicomposite was from matrix cracking, and the number of fiber breaks tended not to exceed 6-8 in a given region (Figure 5.5a). Additionally, fiber breaks clustered near the plane of a matrix crack, previously observed by [12, 30]. After matrix crack saturation, the quantity of both single and clustered fiber breaks increased (Figure 5.5b). Negligible damage to the matrix was observed during this load step, implying that acoustic signals are primarily the result of fiber failure.

When taken in context with results from the spectral framework, XCT observations provide evidence that it is not possible to identify fiber breaks when matrix cracking is a primary damage mechanism. To understand why this is the case, it is necessary to recognize that fiber breaks occurring at low loads are likely due to dynamic cracking effects where crack propagation through the matrix initiates simultaneous fiber breaks [20]. The acoustic signals from these simultaneous damage events are superimposed and when this is the case, the spectral framework is expected to assign labels according to the mechanism which contributes the most energy to the frequency spectrum [188].

For these superimposed signals it is likely that the energy of the matrix crack will always exceed the energy of the fiber break, contribute the majority of acoustic energy to the frequency spectrum, and drive label assignments. Considering a signal that is composed of a matrix crack and fiber break, whose energies are E_{matrix} and E_{fiber} respectively, a ratio of

$$\frac{E_{\text{matrix}}}{E_{\text{fiber}}} >> 1 \tag{5.3}$$

implies the signal will be labeled as a matrix crack. Similarly, if the ratio is less than unity,

$$\frac{E_{\text{matrix}}}{E_{\text{fiber}}} << 1 \tag{5.4}$$

the signal will be labeled as a fiber break. To estimate the range of values that this ratio can take, the acoustic energy of a signal is assumed to be proportional to the energy release rate of the constituent material, $\Gamma_{\text{matrix/fiber}}$, and the damaged area, $A^d_{\text{matrix/fiber}}$ [203].

$$E_{\text{matrix/fiber}} \propto \Gamma_{\text{matrix/fiber}} A^d_{\text{matrix/fiber}} \tag{5.5}$$

$$\Rightarrow \frac{E_{\text{matrix}}}{E_{\text{fiber}}} = \frac{\Gamma_{\text{matrix}} A^d_{\text{matrix}}}{\Gamma_{\text{fiber}} A^d_{\text{fiber}}} \tag{5.6}$$

If the energy release rates of the fibers and matrix are taken to be equivalent, then

$$\frac{E_{\text{matrix}}}{E_{\text{fiber}}} = \frac{A^d_{\text{matrix}}}{A^d_{\text{fiber}}} \tag{5.7}$$

A reasonable estimate for A^d_{matrix} is 5% of the in-plane matrix content (Figure 5.9), yielding a damaged area of $0.05 A_{total} V_m$. If a matrix crack of this size initiates a cluster of 25 fiber breaks, more than twice as large as the maximum observed clustered fiber breaks (Figure 5.5), the mechanism ratio becomes

$$\frac{E_{\text{matrix}}}{E_{\text{fiber}}} = \frac{0.05 A_{total} V_m}{25 \pi r_f^2} \tag{5.8}$$

Substituting the nominal fiber radius of $r_f = 5\mu\text{m}$:

$$\frac{E_{\text{matrix}}}{E_{\text{fiber}}} = 4.65 >> 1 \tag{5.9}$$

As such, even low-energy partial matrix cracking will dominate the energy landscape during early loads and control the labels of signals (Figure 5.9). Therefore it is only possible to identify fiber failures if the matrix crack mechanism is exhausted.

5.3.2 Autoencoder Signal Labeling

After acoustic signals had been correlated to load states and labeled with the spectral framework, the autoencoder was used to generate labels. To do this, the MAE distribution of unseen matrix crack signals was estimated by performing 100 runs of 5-fold CV on the matrix set (Figure 5.10a). The average 5-fold CV error over 100 runs was 0.0185, and the average of the standard deviation was 0.00256. With the empirical error distribution, an initial threshold was chosen to be 0.47, such that it marked the 95th percentile (Figure 5.10a). Choosing this threshold allows for the identification of fiber breaks with reasonable certainty, in that if a matrix crack was reconstructed with the autoencoder, there would only be a 5% chance its MAE would exceed this threshold [204]. Therefore, signals with MAEs exceeding this threshold are unlikely to be the result of a matrix crack and more likely to be from a fiber break. Once this threshold was chosen, signals from the fiber set were compressed and reconstructed by the autoencoder.

Surprisingly, every reconstructed signal fell below the 95% threshold, indicating that there were no fiber break signals contained within the fiber break set; a conclusion which contrasts with both optical XCT observations and spectral framework signal labels. These results prompted a secondary verification of the code base to ensure it was bug-free; no bugs were found. Following code verification, a threshold scan was conducted on the fiber set to determine if there was a threshold that existed which would yield the same labels as the spectral framework. A threshold which permitted for accurate signal labeling was MAE= 0.0075, just below the 5th percentile of MAEs in the 5-fold CV error distribution. More than this, signals with MAEs *below* this value, rather than above, were labeled as fiber breaks. This finding shows that the autoencoder was able to reconstruct fiber break signals to a high degree of fidelity.

To interpret this, it is important to remember that the fidelity of an autoencoder reconstruction depends on how well the training distribution represents the population distribution of training signals. With this in mind, there are two possible reasons for the unusually low reconstruction errors: 1) the matrix set was incorrectly assumed to be composed primarily of matrix cracks, or 2) the population distribution of fiber break signals is compact compared to matrix crack signals, and therefore few fiber break exemplars are needed to achieve low reconstruction errors. The first hypothesis can be dismissed. There are visual observations, along with the secondary verification from the spectral framework, which show limited fiber failures in the presence of extended matrix cracking (Figure 5.5a, Figure 5.8). Therefore, the more likely hypothesis is what is termed here as the compact distribution hypothesis.

Compact Distribution Hypothesis: The distribution in excited frequencies emitted by fiber break signals is compact relative to the distribution of matrix crack signals.

If this hypothesis is true, then a limited number of fiber break signals in the training set can be sufficiently large to represent the population distribution of fiber breaks. To test this hypothesis and confirm that the low MAEs are a result of compact fiber signal distribution, it is useful to treat the autoencoder labels as a function of the training set. When viewed this way, the the correspondence between the autoencoder and the spectral framework labels will increase as the number of fiber signals in the training set increases, the ARI (i.e. correspondence) between autoencoder labels and spectral framework labels also increases. This hypothesis was tested by:

1. Begin with an empty training set
2. Add the signal recorded at the next lowest stress to the training set
3. Shuffle the training set

4. Train a new, previously untrained autoencoder according to Chapter 5.2.6
5. Reconstruct signals from the fiber set and record their MAEs
6. Label signals with an MAE of < 0.0075 as fiber breaks
7. Measure correspondence with spectral framework labels using the ARI
8. Repeat 2-7 until all signals are contained in the training set

As signals were added to the training set at progressively higher stresses, and the likelihood of fiber break signals being included increased, the ability to reconstruct fiber signals increased as did the correspondence between autoencoder labels and spectral framework labels (Figure 5.11a). To further demonstrate that the distribution of fiber breaks is compact compared to matrix cracks, the above training protocol was modified to reverse the order in which signals were added to the training set. This is equivalent to steps 1-8 above where step 2 is replaced with:

2. Add the signal recorded at the **highest** stress to the training set

Following this procedure, the ARI between the autoencoder and the spectral framework labels immediately takes on a value exceeding 0.4, only dropping below this value due to the stochastic nature of training. This provides two key pieces of support for the compact distribution hypothesis. First, it shows faithful fiber break signal reconstructions can be achieved with few training exemplars, which subsequently implies a tight distribution of signals. Second, it shows that the low ARIs when fewer than 60 training signals are used in Figure 5.11a are not a result of the small training set size, and rather lack of fiber breaks contained within the training set.

5.4 Conclusions

While progress has been made towards AE-ML damage mechanism identification (DMI) in SiC/SiC composites, existing frameworks have substantial limitations which preclude their broader use in real-time structural health monitoring. For example, while the work in Chapter 3 presents the spectral framework, a method for DMI in SiC/SiC composites, the spectral framework cannot be used to identify damage mechanisms while a mechanical test is on-going. Since clustering methods will always find the number of clusters that is specified, the spectral framework can only accurately identify damage mechanisms if both matrix crack and fiber break signals are present which requires mechanical tests to be run to very high stresses $> 85\%$ UTS, or completion. Considering this, the spectral framework can only be used for DMI after the composite is close to failure or has already failed. Furthermore, due to the safety-critical nature of these technologies, the error rates of AE-ML frameworks must be robustly characterized in conditions that accurately represent operating environments. This requires libraries of ground-truth acoustic signals gathered under operating conditions, which are currently infeasible to acquire.

To address these challenges, SiC/SiC minicomposites were subjected to uniaxial tension and imaged via x-ray computed tomography. Signals were separated into two sets, based on general damage mechanism trends observed by XCT, and an autoencoder architecture was used to identify matrix cracks and fiber breaks. Error quantification, via 100 runs of 5-fold CV, was performed and when combined with XCT observations allowed us to conclude that:

1. Early fiber breaks are obscured by matrix crack signals, preventing identification from their acoustic signal
2. The distribution of fiber break acoustic signals is small compared to matrix crack

signals

3. Autoencoders can be leveraged to label signals under realistic operating conditions

These findings are used to support and explain previously developed hypotheses [20, 30, 187]. Moving forward, the obscuring of fiber break signals will be the largest challenge with implementation of any DMI framework. In full-scale composite layups, matrix cracking is a function of the layup geometry, and may never reach saturation [205]. In such cases, DMI is not possible without temporal deconvolution of signals. While not currently feasible, this capability will be needed to leverage AE-ML framework in full-scale composites and is an avenue for future work.

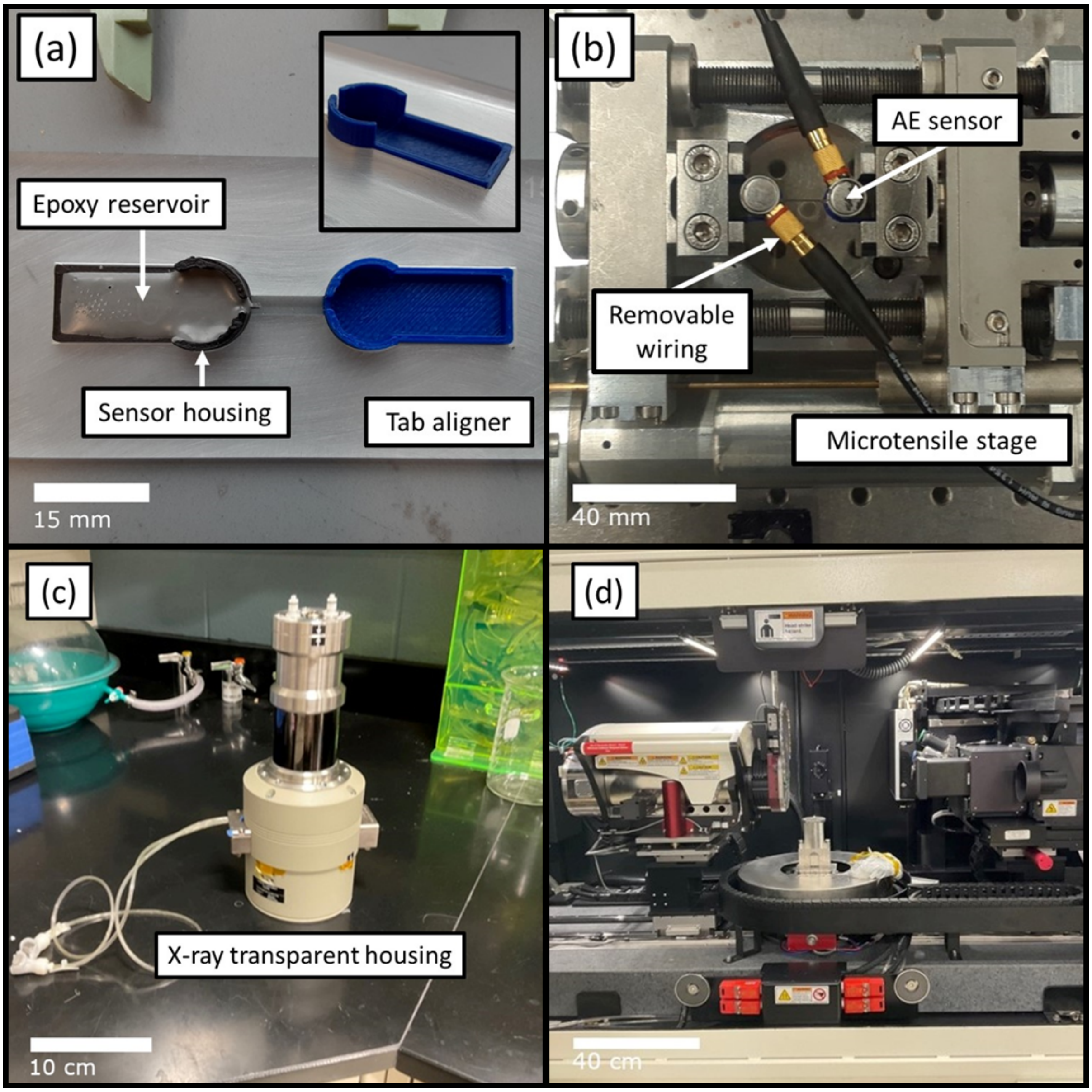

Figure 5.1: **Photograph of mechanical test.** (a) Top-down view of custom minicomposite grips, inset is an isometric view of the tab. Tabs featuring an epoxy reservoir and AE sensor housing were 3D printed. Tabs and sample were aligned in a laser-cut acrylic mold to mitigate off-axis loading. (b) AE sensors are attached to the minicomposite via Crystal Bond. Removable wiring allows transfer to XCT without sensor removal. (c) Side view of *in-situ* load frame with X-ray transparent housing. The lack of wiring ports necessitates disconnectable AE wiring. (d) Zeiss Xradia 520 CT scanner. The load frame sits in the center between the source and the detector.

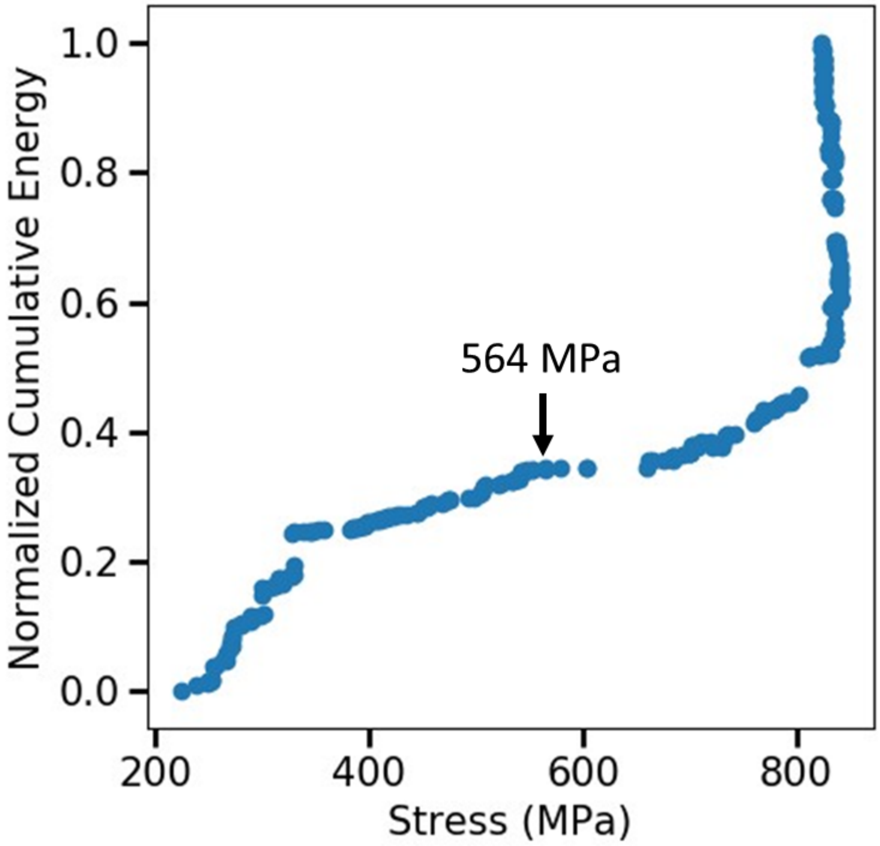

Figure 5.2: **Normalized Cumulative AE energy.** The minicomposite was loaded to 564 MPa during the first load step. Immediately after this load, there is a lack of AE activity, indicating that matrix crack saturation has been reached. Further loading activates fiber failures, and AE activity increases.

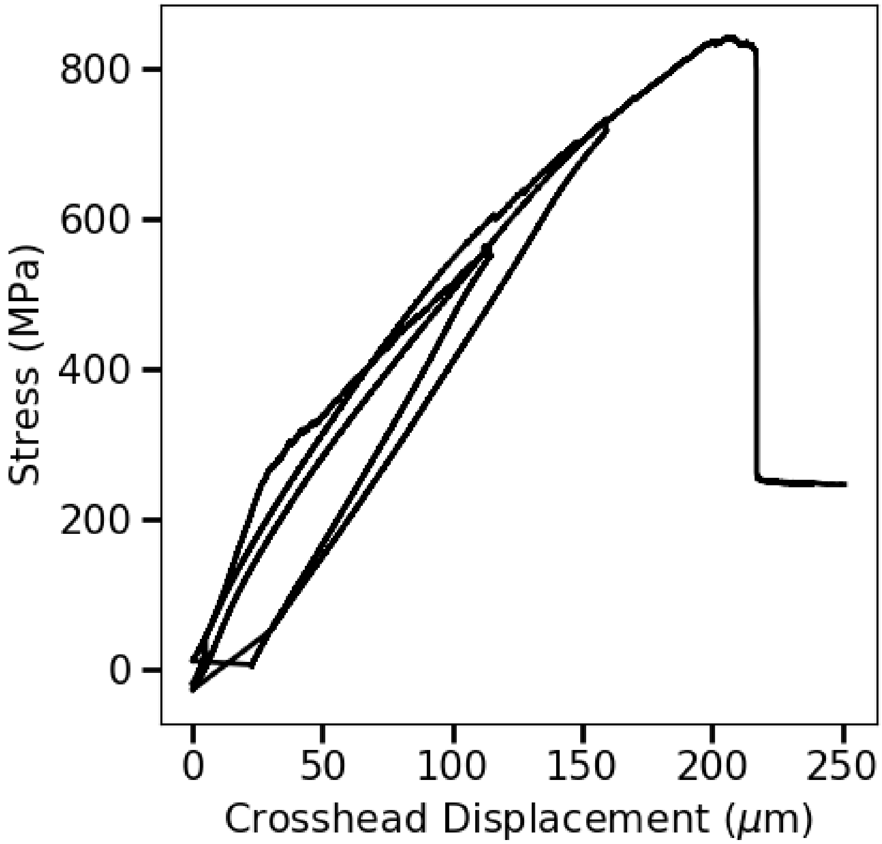

Figure 5.3: **Stress-Displacement of tested minicomposite.** During the first and second load step, the minicomposite was loaded to 564 MPa and 731 MPa respectively before being unloaded. XCT scans were taken after each load step. Following this, the sample was loaded to failure.

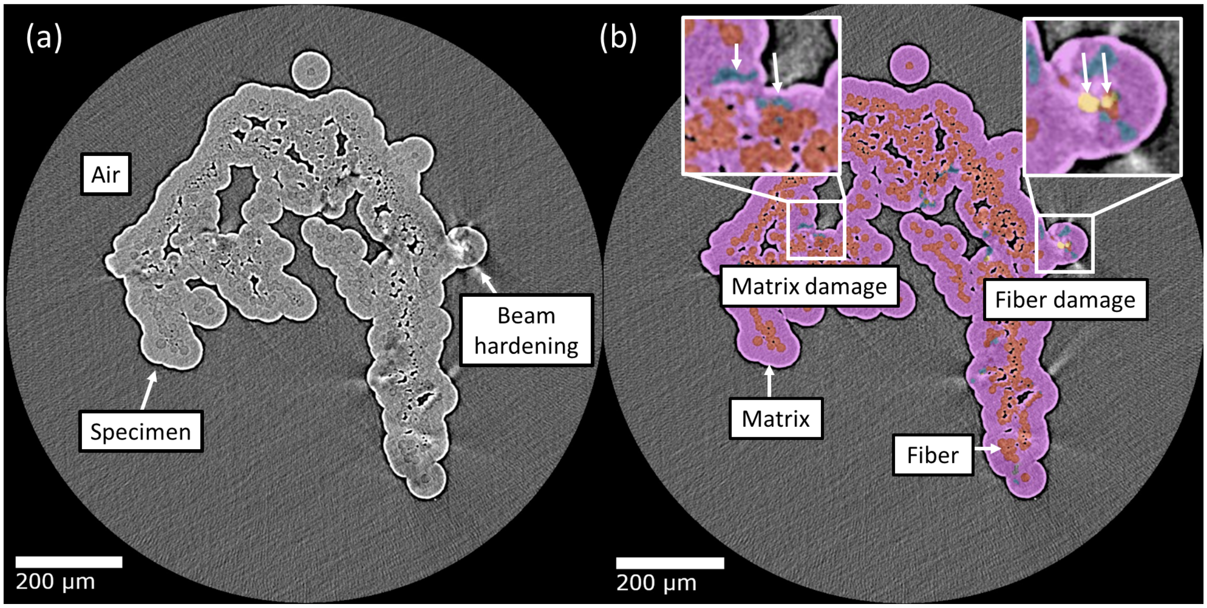

Figure 5.4: **Minicomposite XCT cross-section after loading to matrix crack saturation (564 MPa).** (a) After loading to 564 MPa, the specimen was transferred to the XCT, loaded to 475 MPa to re-open cracks, and scanned. (b) A U-net segmentation model was trained using the deep learning toolkit in Dragonfly. Matrix, fibers, matrix damage, and fiber damage are identified, allowing for rapid identification of

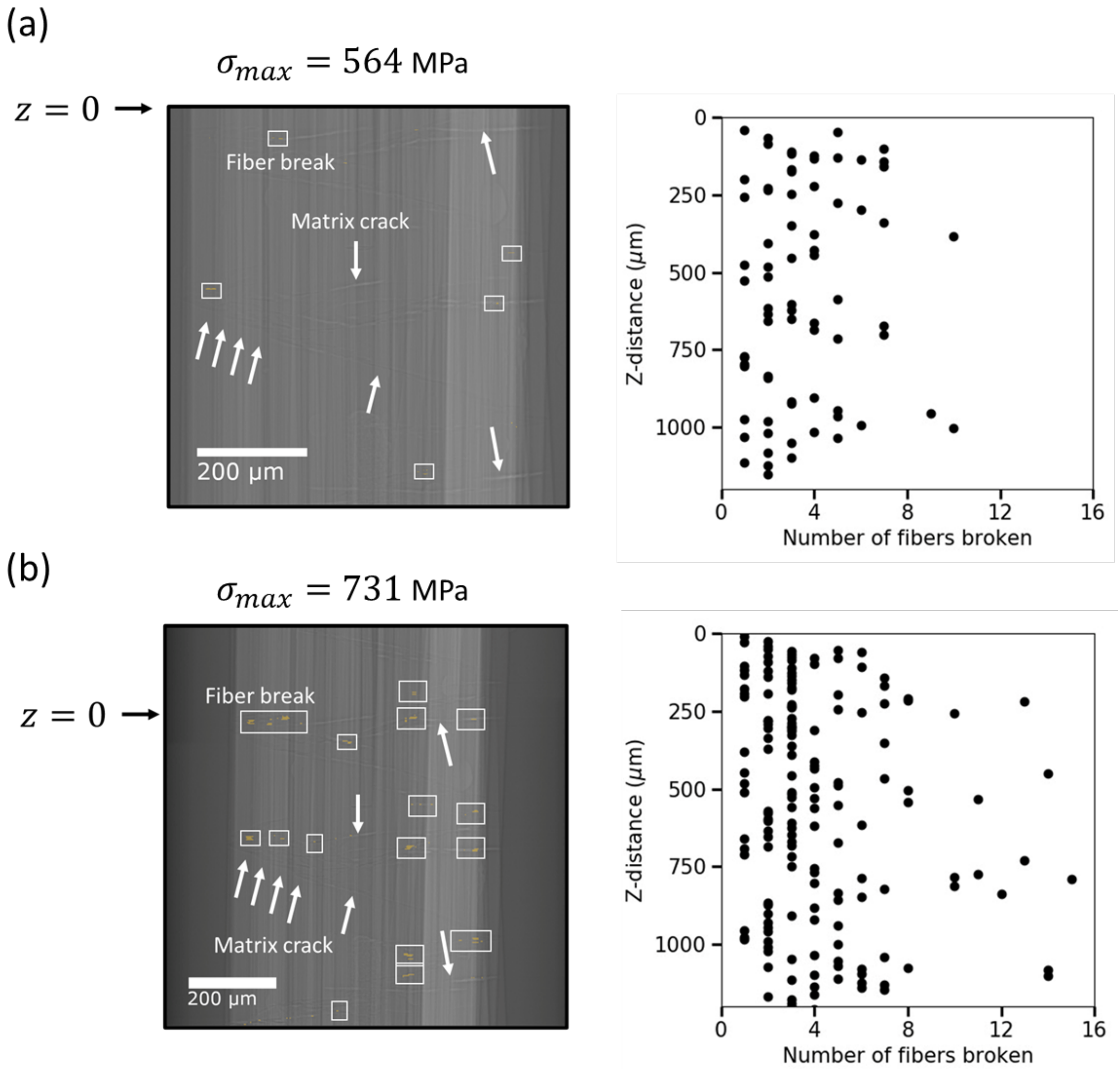

Figure 5.5: **Damaged minicomposite and its fiber break spatial distribution.** (a) Radiograph of minicomposite after loading to matrix crack saturation (564 MPa), with matrix cracks denoted by arrows and fiber breaks denoted by boxes. (b) Radiograph of the same minicomposite, noting the slight change in FOV, after loading to 731 MPa. During the second load step, fiber break activity increased while the extent of matrix cracking was relatively unchanged. Accordingly, the majority of acoustic signals gathered from the second load step result from fiber breaks.

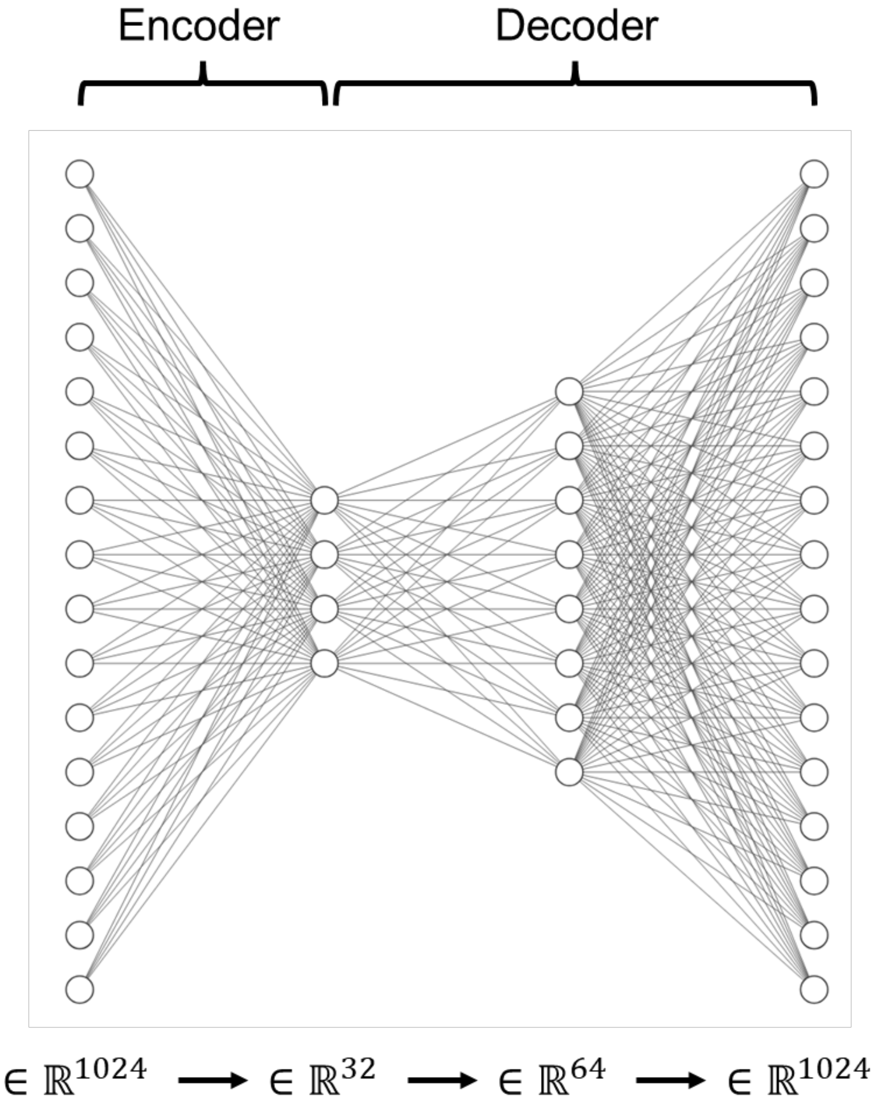

Figure 5.6: **Schematic diagram of the autoencoder architecture.** The encoder function maps the waveform to a 32 dimensional latent space in a single step before a 2-step decompression by the decoder function. Signals with large differences between the original waveform are considered anomalous.

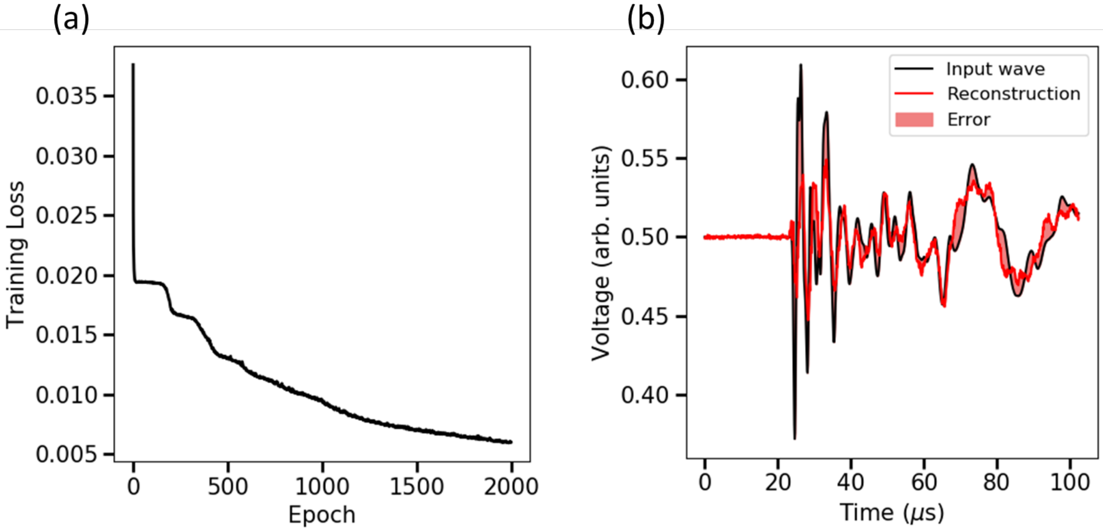

Figure 5.7: **Training loss history of the autoencoder and example of an autoencoder reconstruction.** (a) Training loss history of the autoencoder. Early stopping was implemented at 500 epochs as a form of regularization. Longer training times did not change autoencoder-generated signal labels; (b) Example of an input waveform, its autoencoder reconstruction, and error. Training protocol prevented overfitting while allowing for faithful waveform reconstruction

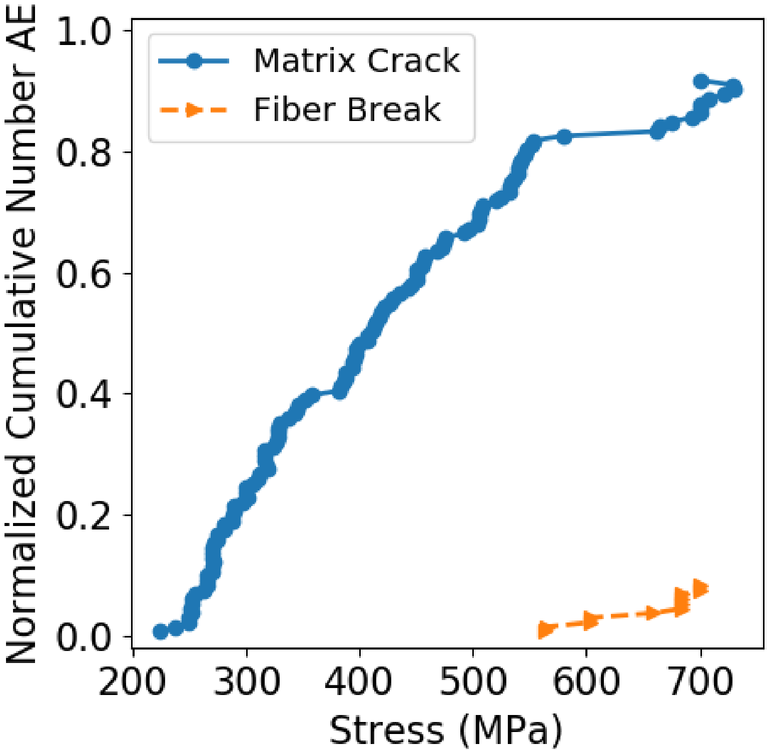

Figure 5.8: **Cumulative signals as clustered by the spectral framework.** Fiber break events become identifiable after matrix crack saturation. Due to the higher energy content of a matrix crack acoustic signal, fiber failures cannot be identified when matrix cracking is the dominant mechanism.

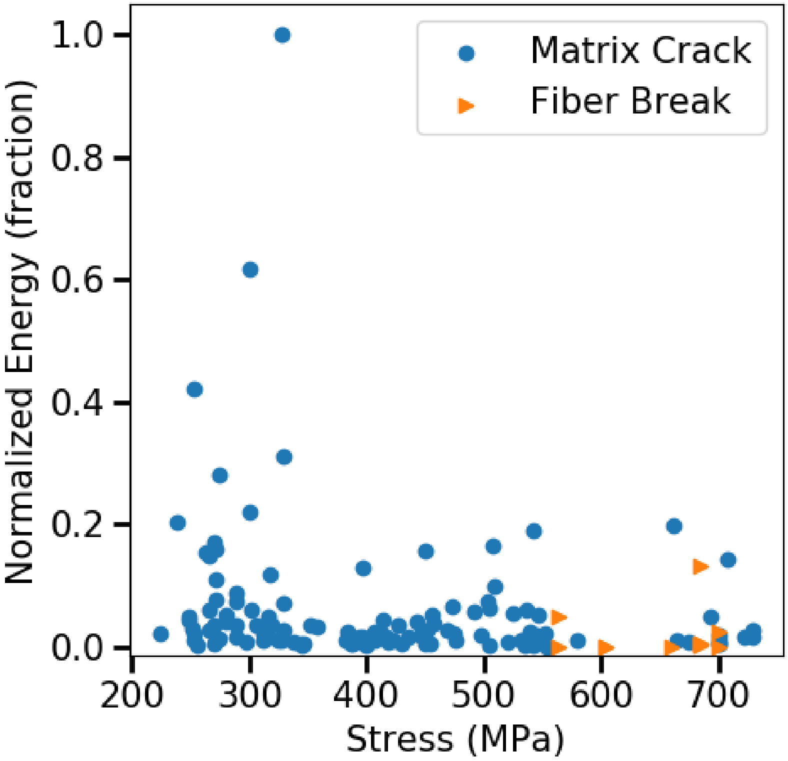

Figure 5.9: **Energy of acoustic signals normalized by the highest energy event as a function of stress.** Before matrix crack saturation, low energy microcracks obscure fiber break signals and drive label assignments by the spectral framework. After matrix crack saturation, fiber failures occur in the absence of microcracking, and can be identified by the spectral framework.

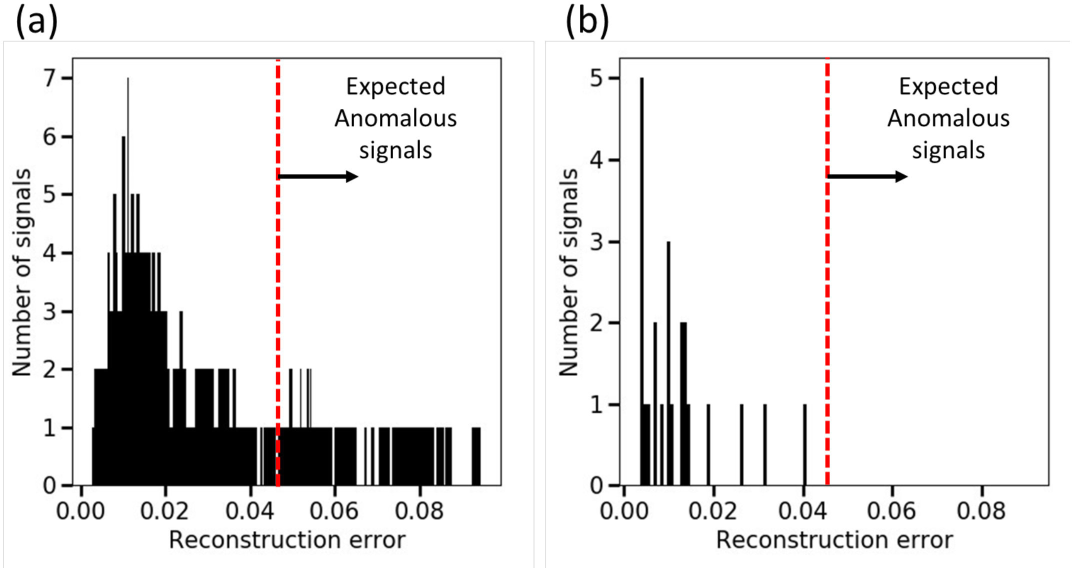

Figure 5.10: **The Mean Absolute Error (MAE) distribution of reconstructed signals.** (a) The empirical MAE distribution estimated by 100 runs of 5-fold CV. The dashed line demarks the 95% threshold. Signals exceeding this threshold are considered anomalous. (b) The MAE distribution of signals from the fiber set when reconstructed during inference. No signal exceeds the 95% threshold indicating no anomalous signals.

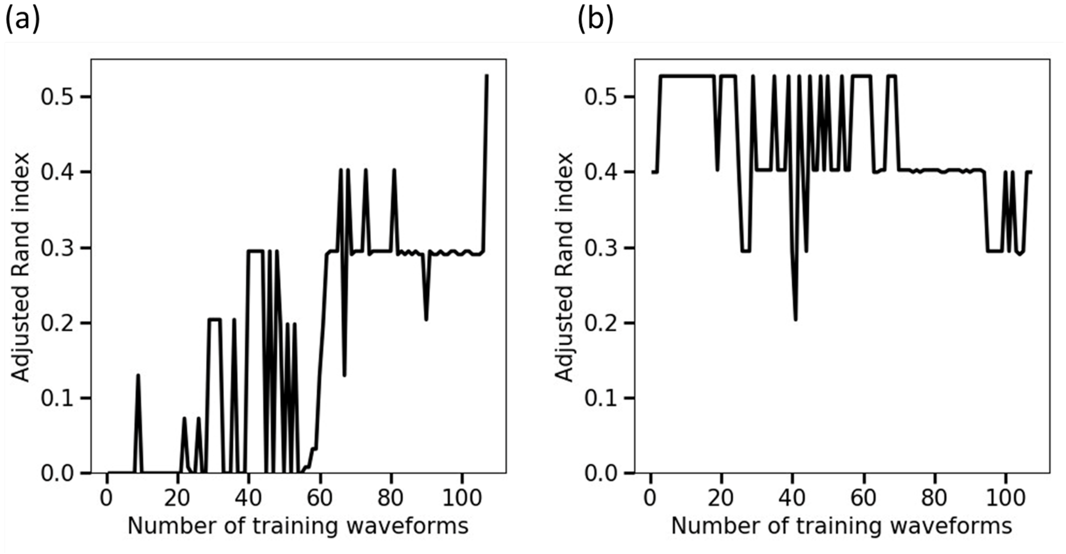

Figure 5.11: **The Adjusted Rand Index (ARI) between autoencoder labels and spectral labels as a function of the number of signals included in the training set.** Few fiber break exemplars are needed to enable low MAE-reconstructions. (a) Signals gathered at increasing stresses were progressively added to the training set. As signals at higher stresses were included, the likelihood of a fiber break signal being included increased, as did the ability to reconstruct fiber break signals. (b) Signals gathered at **decreasing** stresses were progressively added to the training set. This shows the ARI is dependent on the type of signal seen during training, not the number of signals contained in the training set.

Chapter 6

Conclusions and Recommendations

6.1 Conclusions and Impact

Fiber-reinforced ceramic matrix composites (CMCs) are industrially relevant and scientifically interesting structural materials for extreme environments. Their heterogeneous microstructure prevents unstable crack propagation and runaway failure of parts made from these materials. Yet, the same heterogeneous microstructure complicates lifetime prediction and limits the use of CMCs in safety-critical applications. The work in this dissertation: 1) explored the limitations of well-established damage models, 2) developed trustworthy machine learning (ML) tools for damage mechanism identification (DMI) 3) leveraged the newly developed ML tools together with x-ray computed tomography to precisely characterize damage progression in SiC/SiC minicomposites.

The experimental study in Chapter 2 demonstrated that common microstructural heterogeneities in SiC/SiC minicomposites led to substantial differences between observed crack opening displacements (CODs) and predictions from well-established monotonic tension models. Independent of microstructure, these heterogeneities resulted in incomplete and non-planar cracking, which in turn led to COD growth rates that deviated from the expected σ^2 dependence. Across four lengths of SiC/SiC minicomposites, with two different microstructures, the CODs of 62 cracks were measured as a function of stress. More than 50% of all observed cracks were found to deviate from this σ^2 dependence and fell below the range of CODs that could be explained by using any set of interfacial parameters. The results from this study substantiated hypotheses from previous literature that these underpredictions were the result of intact subsurface matrix content. However, this study was conducted via scanning electron microscope (SEM) observations which are unable to directly observe subsurface damage.

As a result, subsequent studies which aim to characterize subsurface damage require the use of high-fidelity subsurface interrogation techniques. However, there are few of these techniques that exist and even fewer that are easily accessible. Acoustic emission (AE) is one such method that satisfies both these requirements; it is capable of characterizing subsurface damage with high temporal resolution, experimental setups are cheap ($< \$10,000$), and the hardware is compatible with other high-resolution characterization methods such as SEM or tunneling electron microscopy. Yet despite the advantages of AE, the ability to identify damage mechanisms from their acoustic signal is underdeveloped, and developing this capability is critical if AE is to be fully leveraged for subsurface damage characterization. In Chapter 3, a novel machine learning framework for AE-based damage mechanism identification (DMI) was created. This framework, termed the spectral framework, is the first demonstration of AE-based DMI in elastically similar composites. It leverages a frequency-based representation scheme along with the spectral clustering algorithm to group waveforms based on their damage mechanism.

In Chapter 4 a standardized benchmarking data set was created and used to quantitatively rank the spectral framework against four state-of-the-art DMI frameworks. This data set was used to explore salient signal representation schemes, propose a set of guidelines for identification of salient features, and quantify the amount of data needed to ensure stable framework performance. Moreover, the benchmarking data set in this chapter offers the AE community a cheap and highly accessible method for creating and expanding a database which is suitable for benchmarking. The experimental setup used to gather the data set is replicable with common laboratory equipment and uses commercially available pencil leads to generate acoustic signals. Consequently, the number and type of acoustic signals within this data set can be readily increased, making this method a powerful tool for generating the large amounts of data needed for supervised ML approaches.

While the PLB data set is useful for assessing the relative discriminating power of DMI frameworks, it cannot be used to assess how a framework will perform when applied to full-scale composite structures. Moreover, the DMI frameworks in Chapter 4 are unable to be used for real-time DMI which limits their usefulness. Addressing these challenges in Chapter 5, an autoencoder-based anomaly detection approach was created and shown to enable real-time DMI. X-ray computed tomography observations were leveraged alongside AE to create a library of matrix crack and fiber break signals. When both the autoencoder approach and the spectral framework were used to label signals in this library, it was shown that the energy content of early matrix cracks signals prevented labeling of early fiber breaks. Additionally, uncertainty quantification was carried out which showed that the distribution of frequency content within fiber break signals is narrow. These findings indicate that without the ability to deconvolve overlapping fiber break and matrix crack signals, detection of fiber break onset is not possible in composite layups.

6.2 Future Outlooks

The work found throughout this dissertation raises several interesting questions and avenues for future work, related to both machine learning and high-fidelity mechanical characterization of composites. Based on the findings herein, paths forward which are expected to be of highest impact are:

1. Temporal deconvolution of acoustic signals
2. ML-enabled high throughput experimentation
3. ML for hypothesis development and testing

These points are discussed in the sections below.

6.2.1 Temporal Deconvolution of Acoustic Signals

In recent years, AE has been combined with optical characterization methods to enhance experiments by providing highly time-resolved data streams [12, 30, 33]. These investigations have shed light on the microstructural factors governing the initiation and accumulation of damage. However, many of these studies have been limited to unidirectional SiC/SiC structures. While findings in these studies can inform composite engineering, this dissertation demonstrated that the mechanics of unidirectional composites are not representative of what occurs in full-scale laminates. To achieve the goal of characterizing representative damage accumulation, AE needs to be integrated with optical studies conducted on full-scale composites. Full utilization of AE requires the ability to perform Damage Mechanism Identification (DMI). Yet, as demonstrated in Chapter 5, fiber breaks at low stresses cannot be identified due to the masking effect of matrix cracks. Therefore, to reliably conduct AE-based DMI, it is essential to develop the capability to deconvolute, or at the very least, identify overlapped matrix crack and fiber break signals. This is crucial for enabling high-fidelity multi-modal studies of representative composite structures.

6.2.2 ML-Enabled High Throughput Experimentation

In the past decade, increases to computing power and storage capacity have transformed the landscape of scientific research. Modern experiments can now explore macroscale areas and volumes of materials, while simultaneously gathering various types of data (e.g., mechanical, chemical, optical, etc.) [206]. While these advancements have resulted in a wealth of information-rich experimental data, it's important to note that each experiment is context-dependent, the underlying mechanics are often unique to specific systems, and obtaining a statistically significant number of observations can be challenging. Recent developments in large language models show promise in connecting seemingly unrelated experiments [207], and generative models have the potential to address issues related to low counting statistics [208, 209], however both of these approaches are in their infancy, and the experimental pipelines needed to complement and verify them are underdeveloped. As demonstrated in Chapter 5 and elsewhere [210, 211], ML has the ability to significantly reduce the time required for experimentation and analysis. This, in turn, permits for rapid collection and analysis of large numbers of observations. By thoughtfully integrating ML into the most time-consuming aspects of experimental pipelines, such as image segmentation, data analysis, or multi-modal data alignment, time costs can be substantially reduced enabling high-throughput experimentation.

6.2.3 ML For Hypothesis Development and Testing

Within experimental mechanics, ML is typically regarded as a tool for accelerating the experimental process, typically by automating time-consuming tasks as discussed earlier. However, ML is often overlooked as a tool which can be used to directly test scientific hypotheses; few studies exist in this application space [212–216]. For hypothesis testing, ML has proven valuable for finding hidden correlations within complex data sets that would otherwise be challenging to interpret. As an example, Gronbech et al. utilized a variational autoencoder (VAE) to show that variations in cancer cell gene expressions are only weakly linked to cell type [213]. In their study, they showed that classification accuracy was uncorrelated to how well the VAE had captured the multivariate transcription distribution. Similarly, Wang et al. used a deep neural network architecture to test the hypothesis that compound stability is more influenced by chemical composition than by crystal structure [216]. To support this, they demonstrated that their architecture could not predict stability when trained on crystal structure information, indicating that other factors, such as chemical composition, play a more significant role in determining compound stability. Finally, as demonstrated in Chapter 5, ML can be applied to test hypotheses about the acoustic behavior of fiber damage in SiC/SiC composites. In each of these cases, ML was used to enable testing of hypotheses that would otherwise be precluded, either due to the volume of data which needs to be parsed or due to the nature of the experiment itself.

Bibliography

[1] D. Baccar and D. Söffker, *Identification and classification of failure modes in laminated composites by using a multivariate statistical analysis of wavelet coefficients*, *Mechanical Systems and Signal Processing* **96** (2017) 77–87.

[2] E. Maillet, N. Godin, M. R'Mili, P. Reynaud, G. Fantozzi, and J. Lamon, *Damage monitoring and identification in SiC/SiC minicomposites using combined acousto-ultrasonics and acoustic emission*, *Composites Part A: Applied Science and Manufacturing* **57** (2014) 8–15.

[3] S. E. Hamdi, A. Le Duff, L. Simon, G. Plantier, A. Sourice, and M. Feuilloy, *Acoustic emission pattern recognition approach based on Hilbert-Huang transform for structural health monitoring in polymer-composite materials*, *Applied Acoustics* **74** (2013), no. 5 746–757.

[4] A. K. Jain, *Data clustering: 50 years beyond k-means*, *Pattern Recognition Letters* **31** (2010), no. 8 651–666.

[5] P. Fränti and S. Sieranoja, *How much can k-means be improved by using better initialization and repeats?*, *Pattern Recognition* **93** (2019) 95–112.

[6] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Édouard Duchesnay, *Scikit-learn: Machine learning in python*, *Journal of Machine Learning Research* **12** (2011), no. 85 2825–2830.

[7] N. Godin, S. Huguet, and R. Gaertner, *Integration of the Kohonen's self-organising map and k-means algorithm for the segmentation of the AE data collected during tensile tests on cross-ply composites*, *NDT and E International* **38** (2005), no. 4 299–309.

[8] L. Calabrese, G. Campanella, and E. Proverbio, *Use of Cluster Analysis of Acoustic Emission Signals in Evaluating Damage Severity in Concrete Structures*, *Journal of Acoustic Emission* **28** (2010) 129–141.

[9] L. van der Maaten and G. Hinton, *Visualizing data using t-SNE*, *Journal of Machine Learning Research* (2008), no. 9 2579–2605.

[10] D. B. Marshall, B. N. Cox, and A. G. Evans, *The Mechanics of Matrix Cracking in Brittle-Matrix Fiber Composites*, *Acta Materialia* **33** (1985), no. 11 2013–2021.

[11] J. W. Hutchinson and H. M. Jensen, *Models of fiber debonding and pullout in brittle composites with friction*, *Mechanics of materials* **9** (1990), no. 2 139–163.

[12] A. M. Hilmas, K. M. Sevener, and J. W. Halloran, *Damage evolution in SiC/SiC unidirectional composites by x-ray tomography*, *Journal of the American Ceramic Society* **103** (2020), no. 5 3436–3447.

[13] M. R. Gorman and W. H. Prosser, *AE source orientation by plate wave analysis*, *Journal of Acoustic Emission* **9** (1991), no. 4 283–288.

[14] M. G. R. Sause, *Investigation of Pencil-Lead Breaks as Acoustic Emission Sources*, *Journal of Acoustic Emission* **29** (2011) 184–196.

[15] A. EVANS and D. MARSHALL, *The Mechanical Behavior of Ceramic Matrix Composites*, in *Proceedings of The 7th International Conference On Fracture (ICF7)*, pp. 3593–3641, Pergamon, jan, 1989.

[16] W. E. Lee, M. Gilbert, S. T. Murphy, and R. W. Grimes, *Opportunities for advanced ceramics and composites in the nuclear sector*, *Journal of the American Ceramic Society* **96** (2013), no. 7 2005–2030.

[17] R. Naslain, O. Dugne, A. Guette, J. Sevely, C. R. Brosse, J.-P. F. Rocher, and J. Cotteret, *Boron nitride interphase in ceramic-matrix composites*, *Journal of the American Ceramic Society* **74** (1991) 2482–2488.

[18] F. W. Zok, *Ceramic-matrix composites enable revolutionary gains in turbine engine efficiency*, *Journal of the American Ceramic Society* **95** (2016), no. 5.

[19] J. Aveston, G. Cooper, and A. Kelly, *The Properties of Fiber Composites: Single and Multiple Fracture*, in *Proceedings of the The properties of the fibre composites*, pp. 15–26, 1971.

[20] C. Chateau, L. Gélébart, M. Bornert, J. Crépin, D. Caldemaison, and C. Sauder, *Modeling of Damage in Unidirectional Ceramic Matrix Composites and Multi-Scale Experimental Validation on Third Generation SiC/SiC Minicomposites*, *Journal of the Mechanics and Physics of Solids* **63** (2014), no. 1 298–319.

[21] B. F. Sorenson and R. Talreja, *Analysis of Damage in a Ceramic Matrix Composite*, *International Journal of Damage Mechanics* **2** (1993), no. 3 246–271.

[22] W. Curtin, *4.03 - stress strain behavior of brittle matrix composites*, in *Comprehensive Composite Materials* (A. Kelly and C. Zweben, eds.), pp. 47–76. Pergamon, Oxford, 2000.

[23] E. B. Callaway, P. G. Christodoulou, and F. W. Zok, *Deformation, rupture and sliding of fiber coatings in ceramic composites*, *Journal of the Mechanics and Physics of Solids* **132** (2019).

[24] L. Saucedo-Mora and T. J. Marrow, *Multi-Scale Damage Modelling in a Ceramic Matrix Composite using a Finite-Element Microstructure Meshfree Methodology*, *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **374** (jul, 2016) 20150276.

[25] A. Shojaei, G. Li, J. Fish, and P. Tan, *Multi-Scale Constitutive Modeling of Ceramic Matrix Composites by Continuum Damage Mechanics*, *International Journal of Solids and Structures* **51** (nov, 2014) 4068–4081.

[26] L. Borkowski and A. Chattopadhyay, *Multiscale Model of Woven Ceramic Matrix Composites Considering Manufacturing Induced Damage*, *Composite Structures* **126** (aug, 2015) 62–71.

[27] B. Budiansky, J. W. Hutchinson, and A. G. Evans, *Matrix fracture in fiber-reinforced ceramics*, *Journal of the Mechanics and Physics of Solids* **34** (1986), no. 2 167–189.

[28] D. S. Beyerle, S. M. Spearing, F. W. Zok, and A. G. Evans, *Damage and Failure in Unidirectional Ceramic-Matrix Composites*, *Journal of the American Ceramic Society* **75** (1992), no. 10 2719–2725.

[29] T. A. Parthasarathy, B. Cox, O. Sudre, C. Przybyla, and M. K. Cinibulk, *Modeling Environmentally Induced Property Degradation of SiC/BN/SiC Ceramic Matrix Composites*, *Journal of the American Ceramic Society* **101** (mar, 2018) 973–997.

[30] E. Maillet, A. Singhal, A. Hilmas, Y. Gao, Y. Zhou, G. Henson, and G. Wilson, *Combining in-situ synchrotron x-ray microtomography and acoustic emission to characterize damage evolution in ceramic matrix composites*, *Journal of the European Ceramic Society* **39** (2019), no. 13 3546–3556.

[31] K. M. Sevener, J. M. Tracy, Z. Chen, J. D. Kiser, and S. Daly, *Crack opening behavior in ceramic matrix composites*, *Journal of the American Ceramic Society* **100** (2017), no. 10 4734–4747, [0608246v3].

[32] C. Chateau, L. Gélébart, M. Bornert, J. Crépin, E. Boller, C. Sauder, and W. Ludwig, *In Situ X-Ray Microtomography Characterization of Damage in* SiC_f/SiC *Minicomposites*, *Composites Science and Technology* **71** (apr, 2011) 916–924.

[33] B. Swaminathan, N. McCarthy, A. Almansour, K. Sevener, T. Pollock, J. Kiser, and S. Daly, *Microscale characterization of damage accumulation in cmcs*, *Journal of the European Ceramic Society* **41** (2021), no. 5 3082–3093. Science of High-Temperature Ceramic-matrix Composites.

[34] G. N. Morscher and N. Godin, *Use of Acoustic Emission for Ceramic Matrix Composites*, in *Ceramic Matrix Composites: Materials, Modeling and Technology* (N. P. Bansal and J. Lamon, eds.), pp. 571–590. John Wiley & Sons, Inc., 1 ed., 2015.

[35] T. Whitlow, E. Jones, and C. Przybyla, *In-situ damage monitoring of a SiC/SiC ceramic matrix composite using acoustic emission and digital image correlation*, *Composite Structures* **158** (dec, 2016) 245–251.

[36] B. Swaminathan, N. McCarthy, A. Almansour, K. Sevener, T. Pollock, J. Kiser, and S. Daly, *Microscale characterization of damage accumulation in cmcs*, *Journal of the European Ceramic Society* **41** (2021), no. 5 3082–3093. Science of High-Temperature Ceramic-matrix Composites.

[37] M. Roth, E. Mojaev, E. Dul'kin, P. Gemeiner, and B. Dkhil, *Phase transition at a nanometer scale detected by acoustic emission within the cubic phase Pb(Zn1/3Nb2/3)O3-xPbTiO3 relaxor ferroelectrics*, *Physical Review Letters* **98** (2007), no. 26 1–4.

[38] J. M. Park and H. C. Kim, *The effects of attenuation and dispersion on the waveform analysis of acoustic emission*, *Journal of Physics D: Applied Physics* **22** (1989), no. 5 617–622.

[39] H. N. Wadley, C. B. Scruby, and J. H. Speake, *Acoustic emission for physical examination of metals*, *International Metals Reviews* **25** (jan, 1980) 41–62.

[40] J. H. Williams, H. Nayeb-Hashemi, and S. S. Lee, *Ultrasonic attenuation and velocity in AS/3501-6 graphite fiber composite*, *Journal of Nondestructive Evaluation* **1** (1980), no. 2 137–148.

[41] W. H. Prosser, J. Dorighi, and M. R. Gorman, *Extensional and Flexural Waves in a Thin-Walled Graphite/Epoxy Tube*, *Journal of Composite Materials* **26** (1992), no. 14 2016–2027.

[42] G. N. Morscher, J. Martinez-Fernandez, and M. J. Purdy, *Determination of interfacial properties using a single-fiber microcomposite test*, 1996.

[43] Q.-Q. Ni, K. Kurashiki, and M. Iwamoto, *AE Technique for Identification of Micro in CFRP Composites Failure Modes*, *Materials Science Research International* **7** (2001), no. 1 67–71.

[44] M. R. Gorman and S. M. Ziola, *Plate waves produced by transverse matrix cracking*, *Ultrasonics* **29** (1991), no. 3 245–251.

[45] M. Surgeon and M. Wevers, *Modal analysis of acoustic emission signals from CFRP laminates*, *NDT and E International* **32** (1999), no. 6 311–322.

[46] R. K. Elsley and L. J. Graham, *Pattern Recognition In Acoustic Emission Experiments*, in *Pattern Recognition and Acoustical Imaging*, vol. 0768, p. 285, 1987.

[47] P. J. de Groot, P. A. Wijnen, and R. B. Janssen, *Real-time frequency determination of acoustic emission for different fracture mechanisms in carbon/epoxy composites*, *Composites Science and Technology* **55** (1995), no. 4 405–412.

[48] M. Johnson and P. Gudmundson, *Broad-band transient recording and characterization of acoustic emission events in composite laminates*, *Composites Science and Technology* **60** (2000), no. 15 2803–2818.

[49] M. G. Sause, A. Gribov, A. R. Unwin, and S. Horn, *Pattern recognition approach to identify natural clusters of acoustic emission signals*, *Pattern Recognition Letters* **33** (2012), no. 1 17–23.

[50] D. Brewer, *HSR/EPM combustor materials development program*, *Materials Science and Engineering A* **261** (1999), no. 1-2 284–291.

[51] G. N. Morscher, J. Hurst, and D. Brewer, *Intermediate-Temperature Stress Rupture of a Woven Hi-Nicalon, BN-Interphase, SiC-Matrix Composite in Air*, *Ceramic Engineering and Science Proceedings* **23** (2000), no. 3 295–.

[52] D. E. Glass, *Ceramic matrix composite (CMC) thermal protection systems (TPS) and hot structures for hypersonic vehicles*, in *15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, no. March 2007, pp. 1–36, 2008.

[53] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015. Software available from tensorflow.org.

[54] N. Godin, S. Huguet, R. Gaertner, and L. Salmon, *Clustering of acoustic emission signals collected during tensile tests on unidirectional glass/polyester composite using supervised and unsupervised classifiers*, *NDT and E International* **37** (2004), no. 4 253–264.

[55] C. R. Ramirez-Jimenez, N. Papadakis, N. Reynolds, T. H. Gan, P. Purnell, and M. Pharaoh, *Identification of failure modes in glass/polypropylene composites by*

means of the primary frequency content of the acoustic emission event, *Composites Science and Technology* **64** (2004), no. 12 1819–1827.

[56] L. Li, S. V. Lomov, X. Yan, and V. Carvelli, *Cluster analysis of acoustic emission signals for 2D and 3D woven glass/epoxy composites*, *Composite Structures* **116** (2014), no. 1 286–299.

[57] D. Xu, P. F. Liu, J. G. Li, and Z. P. Chen, *Damage mode identification of adhesive composite joints under hygrothermal environment using acoustic emission and machine learning*, *Composite Structures* **211** (2019), no. December 2018 351–363.

[58] K. F. Graff, *Wave motion in elastic solids.* Ohio State University Press, 1975.

[59] M. C. Halbig, M. H. Jaskowiak, J. D. Kiser, and D. Zhu, *Evaluation of ceramic matrix composite technology for aircraft turbine engine applications*, *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 2013* (2013).

[60] W. Huang, S. I. Rokhlin, and Y. J. Wang, *Effect of fibre-matrix interphase on wave propagation along, and scattering from, multilayered fibres in composites. Transfer matrix approach*, *Ultrasonics* **33** (1995), no. 5 365–375.

[61] G. N. Morscher and A. L. Gyekenyesi, *The velocity and attenuation of acoustic emission waves in SiC/SiC composites loaded in tension*, *Composites Science and Technology* **62** (2002), no. 9 1171–1180.

[62] S. Biwa, Y. Watanabe, and N. Ohno, *Analysis of wave attenuation in unidirectional viscoelastic composites by a differential scheme*, *Composites Science and Technology* **63** (2003), no. 2 237–247.

[63] C. B. Scruby, H. N. Wadley, and J. J. Hill, *Dynamic elastic displacements at the surface of an elastic half-space due to defect sources*, *Journal of Physics D: Applied Physics* **16** (1983), no. 6 1069–1083.

[64] M. Hamstad, a. O'Gallagher, and J. Gary, *A wavelet transform applied to acoustic emission signals: part 2: source location*, *Journal of Acoustic Emission* **20** (2002), no. 853 62–82.

[65] M. A. Hamstad, *Acoustic emission signals generated by monopole (pencil-lead break) versus dipole sources: finite element modeling and experiments*, *Journal of Acoustic Emission* **25** (2007) 92–106.

[66] P. D. Wilcox, C. Lee, J. J. Scholey, M. I. Friswell, M. Wisnom, and B. Drinkwater, *Progress Towards a Forward Model of the Complete Acoustic Emission Process*, *Advanced Materials Research* **13-14** (2006) 69–76.

[67] M. G. Sause and S. Horn, *Simulation of acoustic emission in planar carbon fiber reinforced plastic specimens*, *Journal of Nondestructive Evaluation* **29** (2010), no. 2 123–142.

[68] M. G. R. Sause and S. R. Horn, *Influence of Specimen Geometry on Acoustic Emission Signals in Fiber*, *29th European Conference on Acoustic Emission Testing* (2010) 1–8.

[69] M. Y. Bhuiyan, J. Bao, B. Poddar, and V. Giurgiutiu, *Toward identifying crack-length-related resonances in acoustic emission waveforms for structural health monitoring applications*, *Structural Health Monitoring* **17** (2018), no. 3 577–585.

[70] A. M. Zelenyak, M. A. Hamstad, and M. G. Sause, *Modeling of acoustic emission signal propagation in waveguides*, *Sensors (Switzerland)* **15** (2015), no. 5 11805–11822.

[71] T. L. Gall, T. Monnier, C. Fusco, N. Godin, and S. E. Hebaz, *Towards quantitative acoustic emission by finite element modelling: Contribution of modal analysis and identification of pertinent descriptors*, *Applied Sciences (Switzerland)* **8** (2018), no. 12.

[72] D. G. Aggelis, T. Shiotani, A. Papacharalampopoulos, and D. Polyzos, *The influence of propagation path on elastic waves as measured by acoustic emission parameters*, *Structural Health Monitoring* **11** (2012), no. 3 359–366.

[73] I. E. Tabrizi, A. Kefal, J. S. M. Zanjani, C. Akalin, and M. Yildiz, *Experimental and numerical investigation on fracture behavior of glass/carbon fiber hybrid composites using acoustic emission method and refined zigzag theory*, *Composite Structures* **223** (2019), no. March 110971.

[74] W. Prosser, *Advanced AE techniques in composite materials Research*, *Journal of Acoustic Emission* **14** (1996), no. 3.

[75] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* MIT Press, 2016. `http://www.deeplearningbook.org`.

[76] A. Ng, M. Jordan, and Y. Weiss, *On spectral clustering: analysis and an algorithm*, *Advances in Neural Information Processing Systems,* **14** (2001).

[77] D. Reynolds, *Gaussian Mixture Models*, *Encyclopedia of biometrics* **741** (2009), no. 2 1–5.

[78] A. K. Jain, P. Murty, M, and J. Flynn, P, *Data clustering: a review*, *ACM Computing Surveys* **31** (1999), no. 3 264–323, [1101.1881].

[79] A. K. Jain, R. Duin, and J. Mao, *Statistical Pattern Recognition: A Review*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (2000), no. 1 4–37.

[80] C. Muir, N. Tulshibagwale, A. Furst, B. Swaminathan, A. Almansour, K. Sevener, M. Presby, J. Kiser, T. Pollock, S. Daly, *et. al.*, *Quantitative benchmarking of acoustic emission machine learning frameworks for damage mechanism identification*, *Integrating Materials and Manufacturing Innovation* **12** (2023), no. 1 70–81.

[81] S. Watanabe, *Pattern Recognition: Human and Mechanical.* John Wiley and Sons, 1985.

[82] W. Alchakra, K. Allaf, and J. M. Ville, *Acoustical emission technique applied to the characterisation of brittle materials*, *Applied Acoustics* **52** (1997), no. 1 53–69.

[83] M. Johnson, *Waveform based clustering and classification of AE transients in composite laminates using principal component analysis*, *NDT and E International* **35** (2002), no. 6 367–376.

[84] V. Kostopoulos, T. H. Loutas, A. Kontsos, G. Sotiriadis, and Y. Z. Pappas, *On the identification of the failure mechanisms in oxide/oxide composites using acoustic emission*, *NDT and E International* **36** (2003), no. 8 571–580.

[85] R. de Oliveira and A. T. Marques, *Health monitoring of FRP using acoustic emission and artificial neural networks*, *Computers and Structures* **86** (2008), no. 3-5 367–373.

[86] M. Moevus, N. Godin, M. R. Mili, D. Rouby, P. Reynaud, G. Fantozzi, and G. Farizy, *Analysis of damage mechanisms and associated acoustic emission in two SiC_f/[Si–B–C] composites exhibiting different tensile behaviours. Part II: unsupervised acoustic emission data clustering*, *Composites and Science Technology* **68** (2008), no. 6 1258–1265.

[87] R. Gutkin, C. J. Green, S. Vangrattanachai, S. T. Pinho, P. Robinson, and P. T. Curtis, *On acoustic emission for failure investigation in CFRP: Pattern recognition and peak frequency analyses*, *Mechanical Systems and Signal Processing* **25** (2011), no. 4 1393–1407.

[88] S. Momon, N. Godin, P. Reynaud, M. R'Mili, and G. Fantozzi, *Unsupervised and supervised classification of AE data collected during fatigue test on CMC at high temperature*, *Composites Part A: Applied Science and Manufacturing* **43** (2012), no. 2 254–260.

[89] M. Kempf, O. Skrabala, and V. Altstädt, *Reprint of: Acoustic emission analysis for characterisation of damage mechanisms in fibre reinforced thermosetting polyurethane and epoxy*, *Composites Part B: Engineering* **65** (2014) 117–123.

[90] E. Maillet and G. N. Morscher, *Waveform-based selection of acoustic emission events generated by damage in composite materials*, *Mechanical Systems and Signal Processing* **52** (2015), no. 53.

[91] L. Li, Y. Swolfs, I. Straumit, X. Yan, and S. V. Lomov, *Cluster analysis of acoustic emission signals for 2D and 3D woven carbon fiber/epoxy composites*, *Journal of Composite Materials* **50** (2016), no. 14 1921–1935.

[92] A. K. Das, D. Suthar, and C. K. Leung, *Machine learning based crack mode classification from unlabeled acoustic emission waveform features*, *Cement and Concrete Research* **121** (2019), no. January 42–57.

[93] N. E. Huang and Z. Wu, *A Review on Hilbert-Huang Transform : Method and Its Applications To Geophysical Studies*, *Reviews of Geophysics* **46** (2008), no. 2007 1–23.

[94] H. Suzuki, T. Kinjo, Y. Hayashi, M. Takemoto, K. Ono, and Y. Hayashi, *Wavelet transform of acoustic emission signals*, *Journal of Acoustic Emission* **14** (1996), no. 2 69–84.

[95] N. E. Huang, Z. Shen, and S. R. Long, *A New View of Nonlinear Water Waves: The Hilbert Spectrum*, *Annual Review of Fluid Mechanics* **31** (1999), no. 1 417–457.

[96] G. Qi, *Wavelet-based AE characterization of composite materials*, *NDT and E International* **33** (apr, 2000) 133–144.

[97] S. F. Wirtz, N. Beganovic, and D. Söffker, *Investigation of damage detectability in composites using frequency-based classification of acoustic emission measurements*, *Structural Health Monitoring* **18** (2019), no. 4 1207–1218.

[98] K. M. Bak, K. Kalaichelvan, G. K. Vijayaraghavan, and B. Sridhar, *Acoustic emission wavelet transform on adhesively bonded single-lap joints of composite laminate during tensile test*, *Journal of Reinforced Plastics and Composites* **32** (2013), no. 2 87–95.

[99] K. Arakawa and T. Matsuo, *Acoustic emission pattern recognition method utilizing elastic wave simulation*, *Materials Transactions* **58** (2017), no. 10 1411–1417.

[100] A. Satour, S. Montrésor, M. Bentahar, and F. Boubenider, *Wavelet Based Clustering of Acoustic Emission Hits to Characterize Damage Mechanisms in Composites*, *Journal of Nondestructive Evaluation* **39** (2020), no. 2 1–11.

[101] I. Daubechies, *The Wavelet Transform, Time-Frequency Localization and Signal Analysis*, *IEEE Transactions on Information Theory* **36** (1990), no. 5 961–1005.

[102] M. Fotouhi, M. Saeedifar, S. Sadeghi, M. Ahmadi Najafabadi, and G. Minak, *Investigation of the damage mechanisms for mode I delamination growth in foam core sandwich composites using acoustic emission*, *Structural Health Monitoring* **14** (2015), no. 3 265–280.

[103] N. Morizet, N. Godin, J. Tang, E. Maillet, M. Fregonese, and B. Normand, *Classification of acoustic emission signals using wavelets and Random Forests: Application to localized corrosion*, *Mechanical Systems and Signal Processing* **70-71** (2016) 1026–1037.

[104] The Mathworks, Inc., Natick, Massachusetts, *MATLAB version 9.9.0.1462360 (R2020b)*, 2020.

[105] D. Laszuk, "Python implementation of empirical mode decomposition algorithm." `https://github.com/laszukdawid/PyEMD`, 2017.

[106] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, *Nature Methods* **17** (2020) 261–272.

[107] Z. Yang, Z. Yu, C. Xie, and Y. Huang, *Application of Hilbert-Huang Transform to acoustic emission signal for burn feature extraction in surface grinding process*, *Measurement: Journal of the International Measurement Confederation* **47** (2014), no. 1 14–21.

[108] H. WenQin, L. Ying, G. AiJun, and F. G. Yuan, *Damage Modes Recognition and Hilbert-Huang Transform Analyses of CFRP Laminates Utilizing Acoustic Emission Technique*, *Applied Composite Materials* **23** (2016), no. 2 155–178.

[109] J. Xu, W. Wang, Q. Han, and X. Liu, *Damage pattern recognition and damage evolution analysis of unidirectional CFRP tendons under tensile loading using acoustic emission technology*, *Composite Structures* **238** (2020), no. December 2019.

[110] A. Marec, J. H. Thomas, and R. El Guerjouma, *Damage characterization of polymer-based composite materials: Multivariable analysis and wavelet transform for clustering acoustic emission data*, *Mechanical Systems and Signal Processing* **22** (2008), no. 6 1441–1464.

[111] K. Chidananda Gowda and G. Krishna, *Agglomerative clustering using the concept of mutual nearest neighbourhood*, *Pattern Recognition* **10** (1978), no. 2 105–112.

[112] S. Baraty, D. A. Simovici, and C. Zara, *The impact of triangular inequality violations on medoid-based clustering*, in *Foundations of Intelligent Systems* (M. Kryszkiewicz, H. Rybinski, A. Skowron, and Z. W. Raś, eds.), (Berlin, Heidelberg), pp. 280–289, Springer Berlin Heidelberg, 2011.

[113] Y. P. Reddy, P. Viswanath, and B. E. Reddy, *Semi-supervised single-link clustering method*, in *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–5, 2016.

[114] D. Rajendra, T. Knighton, A. Esterline, and M. J. Sundaresan, *Physics-based classification of acoustic emission waveforms*, in *Nondestructive Characterization for Composite Materials, Aerospace Engineering, Civil Infrastructure, and Homeland Security 2011*, vol. 7983, 2011.

[115] M. G. Sause and S. Horn, *Quantification of the uncertainty of pattern recognition approaches applied to acoustic emission signals*, *Journal of Nondestructive Evaluation* **32** (2013), no. 3 242–255.

[116] M. Moevus, D. Rouby, N. Godin, M. R'Mili, P. Reynaud, G. Fantozzi, and G. Farizy, *Analysis of damage mechanisms and associated acoustic emission in two SiC/[Si-B-C] composites exhibiting different tensile behaviours. Part I: Damage patterns and acoustic emission activity*, *Composites Science and Technology* **68** (2007), no. 6 1250–1257.

[117] R Core Team, *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2017.

[118] J. F. MacGregor and T. Kourti, *Statistical process control of multivariate processes*, *Control Engineering Practice* **3** (1995), no. 3 403–414.

[119] M. E. Celebi, H. A. Kingravi, and P. A. Vela, *A comparative study of efficient initialization methods for the k-means clustering algorithm*, *Expert Systems with Applications* **40** (2013), no. 1 200–210, [arXiv:1209.1960].

[120] A. Sibil, N. Godin, M. R'Mili, E. Maillet, and G. Fantozzi, *Optimization of acoustic emission data clustering by a genetic algorithm method*, *Journal of Nondestructive Evaluation* **31** (2012), no. 2 169–180.

[121] M. Fotouhi, H. Heidary, M. Ahmadi, and F. Pashmforoush, *Characterization of composite materials damage under quasi-static three-point bending test using wavelet and fuzzy C-means clustering*, *Journal of Composite Materials* **46** (2012), no. 15 1795–1808.

[122] Y. Mi, C. Zhu, X. Li, and D. Wu, *Acoustic emission study of effect of fiber weaving on properties of fiber-resin composite materials*, *Composite Structures* **237** (2020), no. January 111906.

[123] R. Mohammadi, M. A. Najafabadi, M. Saeedifar, J. Yousefi, and G. Minak, *Correlation of acoustic emission with finite element predicted damages in open-hole tensile laminated composites*, *Composites Part B: Engineering* **108** (2017) 427–435.

[124] M. Shateri, M. Ghaib, D. Svecova, and D. Thomson, *On acoustic emission for damage detection and failure prediction in fiber reinforced polymer rods using pattern recognition analysis*, *Smart Materials and Structures* **26** (2017), no. 6.

[125] A. Refahi Oskouei, H. Heidary, M. Ahmadi, and M. Farajpur, *Unsupervised acoustic emission data clustering for the analysis of damage mechanisms in glass/polyester composites*, *Materials and Design* **37** (2012) 416–422.

[126] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum Likelihood from Incomplete Data Via the EM Algorithm*, *Journal of the Royal Statistical Society: Series B (Methodological)* **39** (1977), no. 1 1–22.

[127] C. B. Do and S. Batzoglou, *What is the expectation maximization algorithm?*, *Nature Biotechnology* **26** (2008), no. 8 897–899.

[128] J. T. Tou, *Dynoc-A dynamic optimal cluster-seeking technique*, *International Journal of Computer & Information Sciences* **8** (1979), no. 6 541–547.

[129] D. L. Davies and D. W. Bouldin, *A Cluster Separation Measure*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1** (1979), no. 2 224–227.

[130] P. J. Rousseeuw, *Silhouettes: A graphical aid to the interpretation and validation of cluster analysis*, *Journal of Computational and Applied Mathematics* **20** (1987), no. C 53–65.

[131] L. Hubert and P. Arabie, *Comparing partitions*, *Journal of Classification* **2** (1985), no. 1 193–218.

[132] W. M. Rand, *Objective criteria for the evaluation of clustering methods*, *Journal of the American Statistical Association* **66** (1971), no. 336 846–850.

[133] J. M. Santos and M. Embrechts, *On the use of the adjusted rand index as a metric for evaluating supervised classification*, in *Artificial Neural Networks – ICANN 2009* (C. Alippi, M. Polycarpou, C. Panayiotou, and G. Ellinas, eds.), (Berlin, Heidelberg), pp. 175–184, Springer Berlin Heidelberg, 2009.

[134] A. J. Gates and Y. Y. Ahn, *The impact of random models on clustering similarity*, *Journal of Machine Learning Research* **18** (2017) 1–28.

[135] G. J. Mclachlan and S. Rathnayake, *On the number of components in a Gaussian mixture model*, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **4** (2014), no. 5 341–355.

[136] T. Kohonen, *The self-organizing map*, *Neurocomputing* **21** (1998), no. 1-3 1–6.

[137] A. K. Jain, *Artificial Neural Networks for Feature Extraction and Multivariate Data Projection*, *IEEE Transactions on Neural Networks* **6** (1995), no. 2 296–317.

[138] J. P. McCrory, S. K. Al-Jumaili, D. Crivelli, M. R. Pearson, M. J. Eaton, C. A. Featherston, M. Guagliano, K. M. Holford, and R. Pullin, *Damage classification in carbon fibre composites using acoustic emission: A comparison of three techniques*, *Composites Part B: Engineering* **68** (2015) 424–430.

[139] J. T. Kim, J. Sakong, S. C. Woo, J. Y. Kim, and T. W. Kim, *Determination of the damage mechanisms in armor structural materials via self-organizing map analysis*, *Journal of Mechanical Science and Technology* **32** (2018), no. 1 129–138.

[140] S. T. Roweis and L. K. Saul, *Nonlinear dimensionality reduction by locally linear embedding*, *science* **290** (2000), no. 5500 2323–2326.

[141] J. B. Tenenbaum, V. De Silva, and J. C. Langford, *A global geometric framework for nonlinear dimensionality reduction*, *science* **290** (2000), no. 5500 2319–2323.

[142] J. W. Sammon, *A nonlinear mapping for data structure analysis*, *IEEE Transactions on Computers* **C-18** (1969), no. 5 401–409.

[143] M. Wattenberg, F. Viégas, and I. Johnson, *How to use t-sne effectively*, *Distill* (2016).

[144] L. McInnes, J. Healy, and J. Melville, *Umap: Uniform manifold approximation and projection for dimension reduction*, 2020.

[145] G. N. Morscher, *Modal acoustic emission of damage accumulation in a woven SiC/SiC composite*, *Composites Science and Technology* **59** (1999), no. 5 687–697.

[146] Z. Wang, J. Ning, and H. Ren, *Frequency characteristics of the released stress wave by propagating cracks in brittle materials*, *Theoretical and Applied Fracture Mechanics* **96** (2018), no. April 72–82.

[147] J. Wei, H. Wang, B. Lin, T. Sui, F. Zhao, and S. Fang, *Acoustic emission signal of fiber-reinforced composite grinding: frequency components and damage pattern recognition*, *International Journal of Advanced Manufacturing Technology* **103** (2019), no. 1-4 1391–1401.

[148] J. J. Scholey, P. D. Wilcox, M. R. Wisnom, and M. I. Friswell, *Quantitative experimental measurements of matrix cracking and delamination using acoustic emission*, *Composites Part A: Applied Science and Manufacturing* **41** (may, 2010) 612–623.

[149] E. Maillet, C. Baker, G. N. Morscher, V. V. Pujar, and J. R. Lemanski, *Feasibility and limitations of damage identification in composite materials using acoustic emission*, *Composites Part A: Applied Science and Manufacturing* **75** (2015) 77–83.

[150] F. E. Oz, N. Ersoy, and S. V. Lomov, *Do high frequency acoustic emission events always represent fibre failure in CFRP laminates?*, *Composites Part A: Applied Science and Manufacturing* **103** (2017) 230–235.

[151] N. Ospitia, D. G. Aggelis, and E. Tsangouri, *Dimension effects on the acoustic behavior of TRC plates*, *Materials* **13** (2020), no. 4.

[152] E. Maillet, N. Godin, M. R'Mili, P. Reynaud, G. Fantozzi, J. Lamon, M. R'Mili, P. Reynaud, G. Fantozzi, and J. Lamon, *Real-time evaluation of energy attenuation: A novel approach to acoustic emission analysis for damage monitoring of ceramic matrix composites*, *Journal of the European Ceramic Society* **34** (jul, 2014) 1673–1679.

[153] P. Theobald, B. Zeqiri, and J. Avison, *Couplants and their influence on AE sensor sensitivity*, *Journal of Acoustic Emission* **26** (2008) 91–97.

[154] K. Ono, *Through-Transmission Characteristics of AE Sensor Couplants*, *J. Acoustic Emission* **34** (2017) 1.

[155] V. Kostopoulos, T. Loutas, and K. Dassios, *Fracture behavior and damage mechanisms identification of SiC/glass ceramic composites using AE monitoring*, *Composites Science and Technology* **67** (2007), no. 7-8 1740–1746.

[156] P. Lyu, L. Yao, X. Ma, G. An, G. Bai, and A. T. Augousti, *Correlation between failure mechanism and rupture lifetime of 2D-C / SiC under stress oxidation condition based on acoustic emission pattern recognition*, *Journal of the European Ceramic Society* (2020), no. June.

[157] S. Huguet, N. Godin, R. Gaertner, L. Salmon, and D. Villard, *Use of acoustic emission to identify damage modes in glass fibre reinforced polyester*, *Composites Science and Technology* **62** (2002), no. 10-11 1433–1444.

[158] W. Roundi, A. El Mahi, A. El Gharad, and J. L. Rebiere, *Acoustic emission monitoring of damage progression in Glass/Epoxy composites during static and fatigue tensile tests*, *Applied Acoustics* **132** (mar, 2018) 124–134.

[159] N. Guel, Z. Hamam, N. Godin, P. Reynaud, O. Caty, F. Bouillon, and A. Paillassa, *Data merging of AE sensors with different frequency resolution for the detection and identification of damage in oxide-based ceramic matrix composites*, *Materials* **13** (2020), no. 20 1–22.

[160] A. Monti, A. El Mahi, Z. Jendli, and L. Guillaumat, *Mechanical behaviour and damage mechanisms analysis of a flax-fibre reinforced composite by acoustic emission*, *Composites Part A: Applied Science and Manufacturing* **90** (nov, 2016) 100–110.

[161] A. Anastassopoulos and T. Philippidis, *Clustering Methodology for the Evaluation of Acoustic Emission from Composites*, *Journal of Acoustic Emission* **13** (1995), no. 1 11–22.

[162] A. Bussiba, M. Kupiec, S. Ifergane, R. Piat, and T. Böhlke, *Damage evolution and fracture events sequence in various composites by acoustic emission technique*, *Composites Science and Technology* **68** (2008), no. 5 1144–1155.

[163] P. fei Zhang, W. Zhou, H. fei Yin, and Y. jing Shang, *Progressive damage analysis of three-dimensional braided composites under flexural load by micro-CT and acoustic emission*, *Composite Structures* **226** (2019), no. April 111196.

[164] W. Zhou, R. Qin, K. ning Han, Z. yuan Wei, and L. H. Ma, *Progressive damage visualization and tensile failure analysis of three-dimensional braided composites by acoustic emission and micro-CT*, *Polymer Testing* **93** (jan, 2021) 106881.

[165] A. Farhidzadeh, A. C. Mpalaskas, T. E. Matikas, H. Farhidzadeh, and D. G. Aggelis, *Fracture mode identification in cementitious materials using supervised pattern recognition of acoustic emission features*, *Construction and Building Materials* **67** (2014), no. PART B 129–138.

[166] H. Tat, J. Wu, M. Pike, J. Schaefer, V. P. Pauca, and R. Li, *Machine learning for acoustic emission signatures in composite laminates*, *32nd Technical Conference of the American Society for Composites 2017* **2** (2017) 1235–1251.

[167] Y. Z. Pappas, Y. P. Markopoulos, and V. Kostopoulos, *Failure mechanisms analysis of 2D carbon/carbon using acoustic emission monitoring*, *NDT and E International* **31** (1998), no. 3 157–163.

[168] S. K. Chelliah, P. Parameswaran, S. Ramasamy, A. Vellayaraj, and S. Subramanian, *Optimization of acoustic emission parameters to discriminate failure modes in glass–epoxy composite laminates using pattern recognition*, *Structural Health Monitoring* **18** (2019), no. 4 1253–1267.

[169] B. Wisner, K. Mazur, V. Perumal, K. P. Baxevanakis, L. An, G. Feng, and A. Kontsos, *Acoustic emission signal processing framework to identify fracture in aluminum alloys*, *Engineering Fracture Mechanics* **210** (2019), no. April 2018 367–380.

[170] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, *Fiji: an open-source platform for biological-image analysis*, *Nature Methods* **9** (Jul, 2012) 676–682.

[171] B. Budiansky, A. G. Evans, and J. W. Hutchinson, *BUDIANSKY, 1995.pdf*, *International Journal of Solids and Structures* **32** (1994) 315–328.

[172] R. K. Goldberg, A. S. Almansour, and R. M. Sullivan, *Analytical Simulation of Effects of Local Mechanisms on Tensile Response of Ceramic Matrix Minicomposites*, *NASA/TM-20210012652* (2021).

[173] A. Almansour, E. Maillet, S. Ramasamy, and G. N. Morscher, *Effect of fiber content on single tow SiC minicomposite mechanical and damage properties using acoustic emission*, *Journal of the European Ceramic Society* **35** (2015), no. 13 3389–3399.

[174] C. P. Yang and F. Jia, *Crack opening model for unidirectional ceramic matrix composites at elevated temperature*, *Ceramics International* **44** (oct, 2018) 17167–17173.

[175] J.-M. Domergue, E. Vagaggini, and A. G. Evans, *Relationships between Hysteresis Measurements and the Constituent Properties of Ceramic Matrix Composites: II, Experimental Studies on Unidirectional Materials*, *Journal of the American Ceramic Society* **78** (oct, 1995) 2721–2731.

[176] N. Lissart and J. Lamon, *Damage and Failure in Ceramic Matrix Minicomposites: Experimental Study and Model*, *Acta Materialia* **45** (mar, 1997) 1025–1044.

[177] C. Muir, "Spectral clustering for ae." `https://github.com/Muir-UCSB/AE-ML_Framework`, 2021.

[178] M. Z. Rodriguez, C. H. Comin, D. Casanova, O. M. Bruno, D. R. Amancio, L. d. F. Costa, and F. A. Rodrigues, *Clustering algorithms: a comparative approach*, *PLOS ONE* **14** (01, 2019) 1–34.

[179] M. Afzalan and F. Jazizadeh, *An automated spectral clustering for multi-scale data*, *Neurocomputing* **347** (2019) 94–108, [arXiv:1902.0199].

[180] D. B. Marshall and W. C. Oliver, *Measurement of interfacial mechanical properties in fiber-reinforced ceramic composites*, *Journal of the American Ceramic Society* **70** (1987), no. 8 542–548.

[181] B. Swaminathan, N. McCarthy, A. Almansour, K. Sevener, A. Musaffar, T. Pollock, J. Kiser, and S. Daly, *Interpreting acoustic energy emission in sic/sic minicomposites through modeling of fracture surface areas*, *Journal of the European Ceramic Society* (2021).

[182] W. A. Curtin, B. K. Ahn, and N. Takeda, *Modeling brittle and tough stress-strain behavior in unidirectional ceramic matrix composites*, *Acta Materialia* **46** (1998), no. 10 3409–3420.

[183] E. B. Callaway and F. W. Zok, *Strengths of ceramic fiber bundles: Theory and practice*, *Journal of the American Ceramic Society* **100** (2017), no. 11 5306–5317, [https://ceramics.onlinelibrary.wiley.com/doi/pdf/10.1111/jace.15062].

[184] M. G. Sause, *Acoustic Emission Source Identification in Large Scale Fibre Reinforced Composites*, in *Czech Society for Nondestructive Testing 32nd European Conference on Acoustic Emission Testing*, (Prague , Czech Republic), pp. 125–136, 2016.

[185] ASTM, *Standard Practice for Secondary Calibration of Acoustic Emission Sensors*, *ASTM International , Designation: E1781/E1781M - 13, West Conshohocken, PA 19428-2959. United States,* **03** (2020), no. September 1–7.

[186] M. Saeedifar, M. A. Najafabadi, D. Zarouchas, H. H. Toudeshky, and M. Jalalvand, *Clustering of interlaminar and intralaminar damages in laminated composites under indentation loading using Acoustic Emission*, *Composites Part B: Engineering* **144** (2018), no. December 2017 206–219.

[187] C. Muir, B. Swaminathan, K. Fields, A. S. Almansour, K. Sevener, C. Smith, M. Presby, J. D. Kiser, T. M. Pollock, and S. Daly, *A machine learning framework for damage mechanism identification from acoustic emissions in unidirectional SiC/SiC composites*, *npj Computational Materials* **7** (2021), no. 1 1–10.

[188] C. Muir, B. Swaminathan, A. Almansour, K. Sevener, C. Smith, M. Presby, J. Kiser, T. Pollock, and S. Daly, *Damage mechanism identification in composites via machine learning and acoustic emission*, *npj Computational Materials* **7** (2021), no. 1 1–15.

[189] M. Saeedifar and D. Zarouchas, *Damage characterization of laminated composites using acoustic emission: A review*, *Composites Part B: Engineering* **195** (2020), no. December 2019 108039.

[190] B. Swaminathan, N. R. McCarthy, A. S. Almansour, K. Sevener, A. K. Musaffar, T. M. Pollock, J. D. Kiser, and S. Daly, *Interpreting acoustic energy emission in SiC/SiC minicomposites through modeling of fracture surface areas*, *Journal of the European Ceramic Society* **41** (2021), no. 14 6883–6893.

[191] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, *ImageNet: A large-scale hierarchical image database*, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, June, 2009. ISSN: 1063-6919.

[192] H. Xiao, K. Rasul, and R. Vollgraf, *Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms*, *arXiv:1708.07747* (Sept., 2017).

[193] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, 2015.

[194] A. Radford, L. Metz, and S. Chintala, *Unsupervised representation learning with deep convolutional generative adversarial networks*, 2015.

[195] H. A. Bale, A. Haboub, A. A. Macdowell, J. R. Nasiatka, D. Y. Parkinson, B. N. Cox, D. B. Marshall, and R. O. Ritchie, *Real-time quantitative imaging of failure events in materials under load at temperatures above 1,600°C*, *Nature Materials* **12** (jan, 2013) 40–46.

[196] P. Liu, X. Sun, Y. Han, Z. He, W. Zhang, and C. Wu, *Arrhythmia classification of lstm autoencoder based on time series anomaly detection*, *Biomedical Signal Processing and Control* **71** (1, 2022).

[197] M. Thill, W. Konen, H. Wang, and T. Bäck, *Temporal convolutional autoencoder for unsupervised anomaly detection in time series*, *Applied Soft Computing* **112** (11, 2021).

[198] H. Homayouni, I. Ray, S. Ghosh, S. Gondalia, and M. G. Kahn, *Anomaly detection in covid-19 time-series data*, *SN Computer Science* **2** (7, 2021).

[199] C. Zhou and R. C. Paffenroth, *Anomaly detection with robust deep autoencoders*, vol. Part F129685, pp. 665–674, Association for Computing Machinery, 8, 2017.

[200] Z. Ghrib, R. Jaziri, and R. Romdhane, *Hybrid approach for anomaly detection in time series data*, 2020.

[201] R. R. Mauritz, F. P. J. Nijweide, J. Goseling, and M. van Keulen, *A probabilistic database approach to autoencoder-based data cleaning*, *arXiv* (6, 2021).

[202] P. Refaeilzadeh, L. Tang, and H. Liu, *Cross-Validation*, *Encyclopedia of database systems* **5** (2009) 532–538.

[203] B. Swaminathan, N. R. McCarthy, A. S. Almansour, K. M. Sevener, A. K. Musaffar, J. D. Kiser, T. M. Pollock, and S. Daly, *Interpreting acoustic energy emission in SiC/SiC minicomp osites through modeling of fracture surface areas*, *Journal of the European Ceramic Society* (2021).

[204] A. Borghesi, A. Bartolini, M. Lombardi, M. Milano, and L. Benini, *Anomaly detection using autoencoders in high performance computing systems*, *Proceedings of the AAAI Conference on Artificial Intelligence* **33** (jul, 2019) 9428–9433.

[205] J. A. Nairn, *Matrix microcracking in composites*, *Polymer matrix composites* **2** (2000) 403–432.

[206] N. Brodnik, C. Muir, N. Tulshibagwale, J. Rossin, M. Echlin, C. Hamel, S. Kramer, T. Pollock, J. Kiser, C. Smith, and S. Daly, *Perspective: Machine learning in experimental solid mechanics*, *Journal of the Mechanics and Physics of Solids* **173** (2023) 105231.

[207] N. R. Brodnik, S. Carton, C. Muir, S. Ghosh, D. Downey, M. P. Echlin, T. M. Pollock, and S. Daly, *Perspective: Large language models in applied mechanics*, *Journal of Applied Mechanics* **90** (2023), no. 10.

[208] D. K. Jangid, N. R. Brodnik, M. P. Echlin, T. M. Pollock, S. H. Daly, and B. Manjunath, *Q-rbsa: High-resolution 3d ebsd map generation using an efficient quaternion transformer network*, *arXiv preprint arXiv:2303.10722* (2023).

[209] D. K. Jangid, N. R. Brodnik, M. P. Echlin, S. Daly, T. Pollock, and B. Manjunath, *Titanium 3d microstructure for physics-based generative models: A dataset and primer*, in *1st Workshop on the Synergy of Scientific and Machine Learning Modeling @ ICML2023*, 2023.

[210] A. Badran, D. Marshall, Z. Legault, R. Makovetsky, B. Provencher, N. Piché, and M. Marsh, *Automated segmentation of computed tomography images of fiber-reinforced composites by deep learning*, *Journal of Materials Science* **55** (2020) 16273–16289.

[211] A. E. Gongora, B. Xu, W. Perry, C. Okoye, P. Riley, K. G. Reyes, E. F. Morgan, and K. A. Brown, *A Bayesian experimental autonomous researcher for mechanical design*, *Science Advances* **6** (2020), no. 15.

[212] Z. Chen and S. Daly, *Deformation twin identification in magnesium through clustering and computer vision*, *Materials Science and Engineering A* **736** (2018), no. August 61–75.

[213] C. H. Grønbech, M. F. Vording, P. N. Timshel, C. K. Sønderby, T. H. Pers, and O. Winther, *Scvae: Variational auto-encoders for single-cell gene expression data*, *Bioinformatics* **36** (8, 2020) 4415–4422.

[214] L. M. Weber, M. Nowicka, C. Soneson, and M. D. Robinson, *diffcyt: Differential discovery in high-dimensional cytometry via high-resolution clustering*, *Communications Biology* **2** (12, 2019).

[215] N. Samusik, Z. Good, M. H. Spitzer, K. L. Davis, and G. P. Nolan, *Automated mapping of phenotype space with single-cell data*, *Nature Methods* **13** (6, 2016) 493–496.

[216] Y. Wang, X.-J. Zhang, F. Xia, E. A. Olivetti, S. D. Wilson, R. Seshadri, and J. M. Rondinelli, *Learning the crystal structure genome for property classification*, *Phys. Rev. Res.* **4** (Apr, 2022) 023029.