


Statistical Partial Wavefield Imaging Using Lamb Wave Signals

Journal Title
XX(X):1–15
©The Author(s) 2017
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/


Joel B. Harley^{1,2} and Chen Ciang Chia³

¹Department of Electrical and Computer Engineering, University of Utah, USA

²Department of Mechanical Engineering, University of Utah, USA

³Department of Aerospace Engineering, Universiti Putra Malaysia, MY

Abstract

This paper presents a baseline-free, model-driven, statistical damage detection and imaging framework for guided waves measured from partial (i.e., non-dense) wavefield scans. Wavefield analysis is an effective non-contact technique for nondestructive evaluation. Yet, there are several limitations to practically implementing wavefield methods. These limitations include slow data acquisition and a lack of statistical reliability. Our approach addresses both of these challenges. We use sparse wavenumber analysis, sparse wavenumber synthesis, and data-fitting optimization to accurately model damage-free wavefield data. We then combine this model with matched field processing to image damage from a small number of partial wavefield measurements. We further derive a hypothesis test based on extreme value theory to statistically detect damage. We test our framework with Lamb wave measurements from a steel plate. With 70 experimental wavefield measurements, we achieve an empirical probability of damage detection of more than 98%, an empirical probability of false alarm of less than 0.17%, and an accurate image of the damage.

Keywords

guided waves, wavefield, statistics, modeling, imaging, matched field processing

Introduction

In the last decade, the acquisition and analysis of the full, or dense, wavefields (generated using bulk or guided waves) has garnered significant interest for non-contact nondestructive evaluation. Researchers use full wavefield analysis to study wave propagation in structures and detect damage. Full wavefield analysis has been successfully applied to both isotropic metals^{1–4} and anisotropic composite materials^{5–7}. These studies have presented several methods for detecting and locating damage, such as through local energy⁵ or local wavenumber⁶ computation. A significant advantage of these full wavefield analysis methods is that they do not require prior baseline (damage-free) data and can resolve small features of damage with a high spatial resolution wavefield. Yet while full wavefield analysis provides significant benefits, it has two significant challenges: slow data acquisition as well as a need for statistical reliability and better data analysis⁸.

Wavefields are often generated with a piezoelectric transducer and then measured at scanning locations with a laser doppler vibrometer (LDV)¹. Alternatively, wavefields can be generated at scanning locations with a pulse laser and then measured from a piezoelectric transducer^{9,10}. For LDV scans, many measurements are often acquired at each location and averaged to overcome poor signal-to-noise ratios. As a result, the scanning process can take significant time. For example, densely scanning a 10 cm by 10 cm region can take more than 12 hours. To test larger structures, much faster processes are necessary.

Researchers have proposed several methods to improve the speed of wavefield acquisition. These include methods based on remote beamforming¹¹, which initially determine local regions that contain damage before performing a full wavefield acquisition. This approach significantly reduces the scan area but has limited resolution¹² and no statistical reliability. Other methods use time and spatial information to predict, or interpolate, the wavefield at the locations not sampled. These algorithms use *partial* wavefields that can be acquired at a rapid rate. Many of these methods utilize algorithms from compressive sensing^{7,13} to perform the interpolation. While they can accurately interpolate wavefields, they are currently limited in their ability to reliably detect, isolate, and locate weak damage wavefields.

In addition, statistically reliable approaches for wavefield analysis and guided wave analysis are immature. While many guided wave localization techniques are based on statistical principles and methods, these statistical principles are often not rigorously applied. For example, delay-and-sum¹⁴ and matched field processing¹⁵ are both fundamentally based on the statistical matched filter^{16,17}. Similarly, guided wave methods for maximum likelihood estimation¹⁸ and Bayesian sensor optimization¹⁹ are directly derived from statistical

Corresponding author:

Department of Electrical and Computer Engineering University of Utah
50 S. Central Campus Dr. Salt Lake City, UT 84112, USA

Email: joel.harley@utah.edu

assumptions. While based on statistical approaches, these methods currently do not provide thresholds or hypothesis tests that ensure statistical significance for their applications.

To improve acquisition time and provide statistical reliability, this paper introduces a statistical detection and imaging framework that extends and improves partial wavefield analysis methods. Our framework uses sparse wavenumber analysis²⁰ to create a damage-free model of the wavefield. We optimally fit this model to measured data to remove errors and uncertainties. We then subtract the fitted, damage-free model from our data to isolate the damage signatures. An image of the damage is then generated with data-driven matched field processing¹⁵. We also introduce a statistical test based on extreme value theory²¹ to ensure statistically significant damage detection and imaging.

Compared with previous work, our approach provides several contributions. Specifically, sparse wavenumber analysis and sparse wavenumber synthesis, which are typically used for localization¹⁵, are now applied to create a damage-free baseline from data containing damage. The scale transform, which is typically used for temperature compensation²², is now applied with a shift estimator and amplitude estimator to correct for wavefield variations between our data and the synthesized baseline. In addition, we present a new constant rate of false alarm matched field processor as well as a statistical analysis framework for defining statistically significant values in a localization image. These methods improve the scan time of current wavefield imaging, allow us to account for effects of local damage on global wavefields, and provide a statistical method for distinguishing damage from non-damage.

Similar to previous work with phased arrays²³, our framework can be used in conjunction with full wavefield imaging. That is, a partial wavefield may be used to rapidly detect damage as well as determine statistically significant regions of concern across the structure. These regions may then be further assessed using a dense, full wavefield scan. The dense scan can resolve small features of the damage.

We test our method with a 10 cm by 10 cm dense wavefield of Lamb waves from a steel plate with a circular 2 mm diameter half-thickness notch at its center. From the dense wavefield, we perform a Monte Carlo analysis with randomly chosen locations that form a set of 600 partial wavefields. The Monte Carlo analysis is repeated with partial wavefields containing 10 measurements to 170 measurements.

With 40 wavefield measurements (1000 times fewer measurements than in the original dense wavefield), we achieve an average localization error below 2 mm (the diameter of the notch). With 70 wavefield measurements (more than 500 times fewer measurements than in the original dense wavefield), we achieve an empirical probability of detection of greater than 98%, an empirical probability of false alarm of less than 0.17%, and identify a 5 mm by 2.5 mm statistically significant region of the plate that contains the 2 mm diameter notch. This small region can then be rapidly evaluated with a dense, full wavefield scan.

Outline

This paper is organized as follows. First, we discuss the challenges with modeling wavefields due to velocity uncertainty, weak anisotropy, and speckle. We then discuss

how we achieve accurate wavefield models through sparse wavenumber analysis and several optimization strategies to reduce the previously discussed errors. These methods allow us to analyze wavefields without prior baseline data. We then describe our statistical framework that integrates matched field processing and extreme value theory for the statistical detection and imaging of damage. We then describe our experimental setup with which we test our framework, and we present the results from these experiments.

Wavefield Modeling Challenges

In this paper, we focus on analyzing partial wavefield scans. We define the density of wavefields according to pixel spacing Δs and the shortest wavelength in the data λ_{\min} . We define a partial wavefield to satisfy

$$\Delta s > \lambda_N = \lambda_{\min}/2. \quad (1)$$

That is, a partial wavefield does not satisfy the spatial Nyquist criteria. We refer to λ_N as the Nyquist wavelength.

Figure 1 shows the local wavefield energy (where $\lambda_{\min} \approx 5\text{mm}$ and the wavelength with the largest amplitude is $\approx 10\text{mm}$) from a full/dense scan ($\Delta s < \lambda_N$), a critically sampled scan ($\Delta s \approx \lambda_N$), and a partial scan ($\Delta s > \lambda_N$). The local energy displays the squared sum of each measurement over time. This is a common approach for locating damage from full wavefields⁵.

In the 40,000 measurement densely sampled scan ($\Delta s = 0.2\lambda_N$), the damage (centered in the frame) is clearly visible. In the 1600 measurement critically sampled scan ($\Delta s = \lambda_N$), the damage is visible but the resolution is poor. In the 100 measurement partially sampled scan ($\Delta s = 4\lambda_N$), the damage is not visible. The partial scan requires 400 times fewer measurements than the dense scan and 16 times fewer measurements than the critically sampled scan. This translates to large data acquisition speed improvements. In this paper, we focus on processing similar partial wavefields.

We distinguish partial wavefield methods from sparse array methods based on two characteristics. First, sparse arrays typically imply an *in situ* system with multiple sensors that use baseline signals to locate damage. In contrast, partial wavefields do not use *in situ* systems and do not require baseline measurements. Second, sparse arrays are often defined by the number of sensors on the structure. In contrast, a partial wavefield is better defined by the average density or average spacing of the scan points rather than the total number. In general, the benefits and limitations for partial wavefields have not been thoroughly explored in the literature.

To isolate damage in a partial wavefield without a prior damage-free baseline, we create a damage-free sparse wavenumber model^{20,24} of Lamb waves. Note that, as is common with modeling, our initial model does not perfectly match the experimental measurements. This is due to many small uncertainties regarding our knowledge of the wavefield. While these uncertainties may not be visibly significant, small differences between our model wavefield and the true wavefield can hide damage signatures that are often 100 times weaker than the direct wavefield. For this reason, we incorporate several optimization techniques to better fit our model with the data.

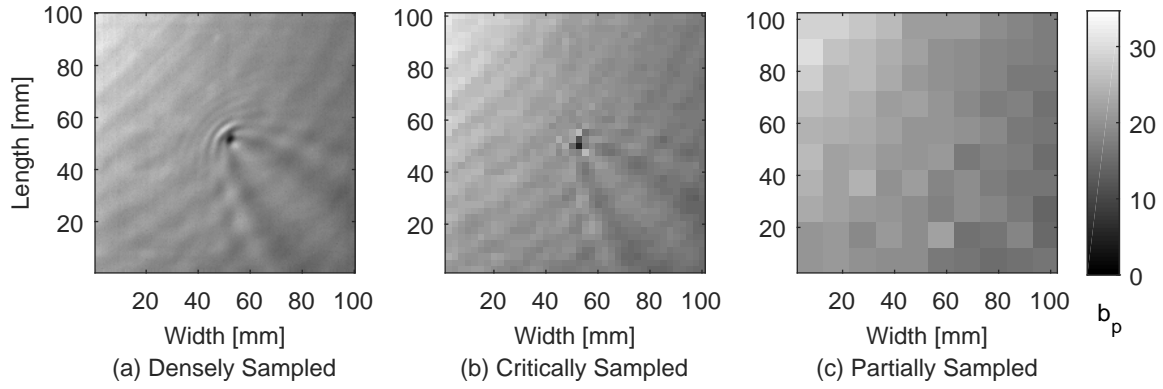


Figure 1. Illustration of the spatial energy of one wavefield with three levels of spatial sampling according to the pixel spacing Δs and minimum wavelength λ_{\min} of the waves. For this wavefield, $\lambda_N \approx 2.5$ mm. (a) Densely sampled wavefield, where $\Delta s < \lambda_N$. (b) Critically sampled wavefield, where $\Delta s \approx \lambda_N$. (c) Partially sampled wavefield, where $\Delta s > \lambda_N$. Our work focuses on partially sampled wavefields, when the spatial energy wavefield is too undersampled to reliably detect damage.

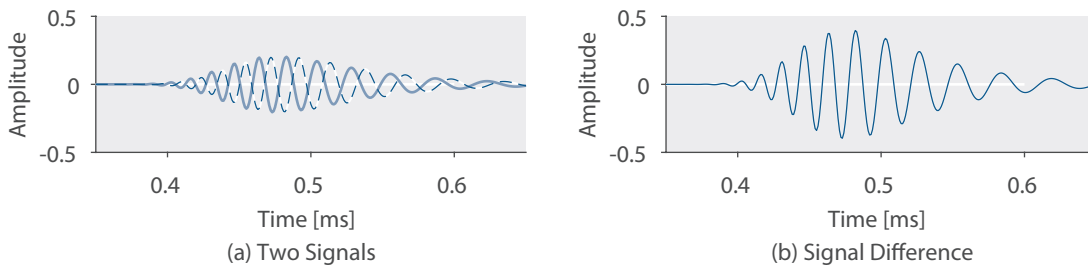


Figure 2. Illustration of the effects of velocity error on guided wave signals. (a) Two simulated guided waves signal where the light-colored signal has a 1% slower velocity than the dark-colored signal. (b) The difference of the two signals, which has a maximum amplitude similar to the original data.

In this section, we discuss sources of uncertainty and modeling error that we observe in the experimental data. We discuss modeling errors in wave velocity, amplitude, anisotropy, and random speckle. In general, these differences are caused by natural characteristics of the material that we cannot reliably account for in modeling.

Velocity and Amplitude Uncertainty

Velocity and amplitude uncertainty is common in ultrasonic guided wave data analysis^{20,25}. Velocity and amplitude both vary with material properties²⁶, temperature²⁷, plate thickness²⁸, and other factors. Since guided wave velocity and amplitude is frequency-dependent (i.e., dispersive), the errors will also vary with frequency. Velocity errors are known to create significant differences in wavefields. Figure 2(a) illustrates portions of two anti-symmetric Lamb wave modes and their subtraction. Both measurements travel a distance of 1.00 m but have a 1% difference in velocity. As shown in Figure 2(b), the difference has almost the same maximum amplitude as the measurements.

In the next section, we learn the frequency-dependent phase velocity characteristics of our experimental Lamb wave data through sparse wavenumber analysis²⁰. This information is incorporated into our wavefield model.

Anisotropy

When modeling guided waves in metal, we often assume the wavefield to be isotropic. However, weak anisotropy

can be created by the natural grain structure of the metal²⁹ and by the direction of the rolling process on the metal sheets³⁰. As a result, certain wave directions will travel at slightly different speeds. In addition, when using lasers to excite or measure wavefields, anisotropy can be artificially created by the measurement system. If there is a small angle deviation away from normal between the plate and laser, the measurements will be taken at slightly different locations than intended. When processing the data, if the measurement grid is not adjusted for this error, the waves will appear to have weak anisotropic behavior.

Weak anisotropy is difficult to visually identify but can create significant differences in data. Figure 3 illustrates weak anisotropy in a measured wavefield from a steel plate. The thin circular line represents a perfect circle with the center at a transducer. When overlaid on the measured wavefield, we see a close, but imperfect resemblance. The perfect circle is closer to the wavefield zero crossing at a 45 degree angle from the transducer. Yet, near the edges of scanning area, the perfect circle is closer to the crest of the wavefield. Much like small errors in velocity, this can create large errors in subtraction.

In the next section, we account for anisotropy error by fitting our model with each measurement. Specifically, we find the optimal shifts and the optimal time-stretches that best match the model data to our measurements. We then adjust our model accordingly.

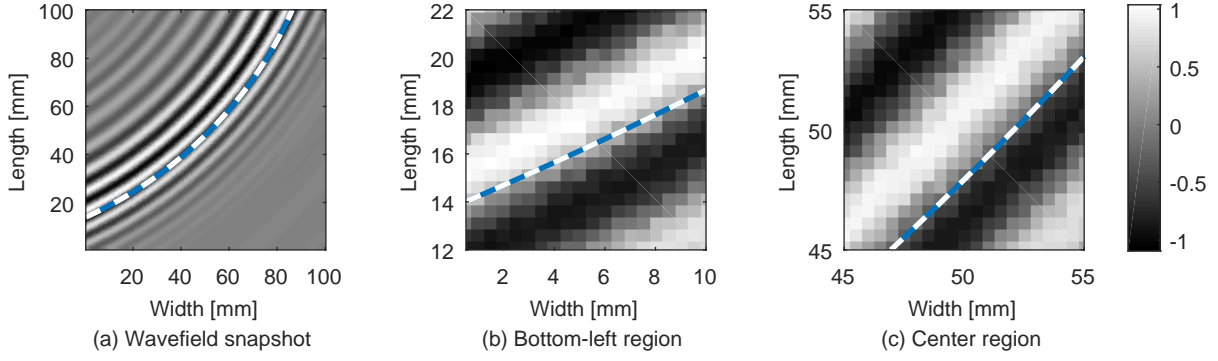


Figure 3. Illustration of weak anisotropy in experimental data. (a) Dense wavefield superimposed with a circle centered around the transmitter location at coordinates $(-70 \text{ mm}, 170 \text{ mm})$. (b) Bottom-left region of the circle, where the circle is near the crest of the wave. (c) Center region of the circle, where the circle is at the zero-crossing of the wave.

Speckle

Speckle noise refers to effects of a random rough or granular structure that diffuses waves and creates interference in data³¹. At the image or pixel level, this causes a random lightening and darkening of the pixel values across the wavefield. For a single time measurement, this can cause small amplitude variations or small shifts in the time-domain. Figure 4 illustrates speckle noise in a wavefield from a steel plate. When we magnify the experimental wavefield, the waves show a weak granular, speckle noise pattern rather than a smooth transition between the crests and troughs of the wave. Like the other sources of error, this effect is weak but can cause significant errors in model subtraction.

In the next section, we account for speckle error by fitting our model to each measurement, similar to our approach for anisotropy. In addition to determining the optimal shift and time-stretch behavior, we optimally re-scale local amplitudes of each measurement to match with our model.

Other Uncertainties

While not prevalent in the data presented in this paper, there are many other possible sources of uncertainty in wavefield data. These uncertainties include source directionality³², multiple-scattering³³, mechanical stresses³⁴, viscoelastic attenuation³⁵, and others. These uncertainties create direction-dependent, inhomogeneous amplitude and velocity variations and can potentially create additional wavefronts in the data. The methodology we describe in the following section can partially address these uncertainties, but additional future research is necessary to optimally align data with these and other complex effects.

Methodology: Isolating Damage Signatures

To isolate damage signatures with minimal error between our model and the experimental data, we apply several optimization techniques. Our methodology has six steps:

1. Extract global dispersion curves from the data
2. Create a model from these dispersion curves
3. Optimally shift-align the model and data
4. Optimally stretch-align the model and data
5. Optimally amplitude-align the model and data

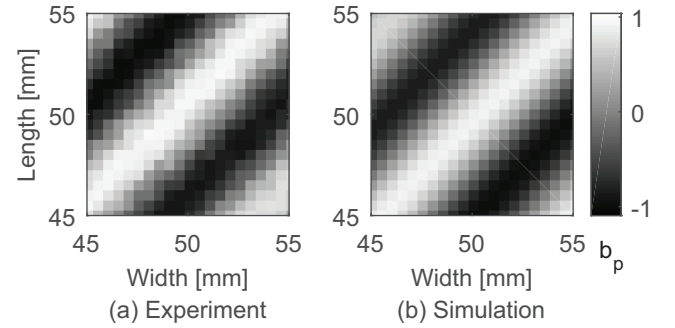


Figure 4. Illustration of speckle in a measured wavefield. (a) Experimental wavefield with speckle (i.e., lightening and darkening of the pixels). (b) Simulated wavefield without speckle.

6. Subtract the model from the data

We describe each step in the following subsections.

Figure 5 summarizes the sparse wavenumber analysis and optimization strategies described in this section. We assume our data consists of M partial wavefield measurements and Q frequencies. In this block diagram, \mathbf{X} is an $M \times Q$ matrix of frequency-domain measurements \mathbf{x}_m (where $1 \leq m \leq M$). Each measurement represents one pixel in the wavefield, although the measurements (or pixels) do not have to be spatially arranged in a uniform grid.

The matrix $\hat{\mathbf{V}}$ is a sparse $N \times Q$ dispersion curve matrix with N chosen wavenumbers. The vector $\hat{\mathbf{x}}_m$ is our unaligned, undamaged model for measurement m . The vector $\hat{\mathbf{x}}_m^{(a)}$ is the aligned, undamaged model for measurement m . The vector \mathbf{z}_m is the model-subtracted data that we use to detect and locate damage.

Step 1: Compute Global Dispersion Curves

Sparse wavenumber analysis assumes our Lamb wave measurements fit a general, far-field model defined by²⁰

$$X(\omega) = \sum_n \sqrt{\frac{1}{k_n(\omega)r}} G_n(\omega) e^{-jk_n(\omega)r}. \quad (2)$$

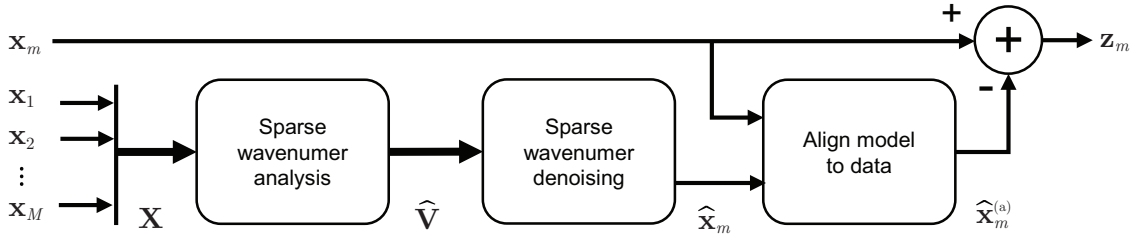


Figure 5. Illustration of our data alignment process for partial wavefield measurement \mathbf{x}_m .

In (2), ω represents angular frequency, r represents the distance traveled by the wave from its source, and $k_n(\omega)$ represents the propagating wavenumber as a function of frequency ω and mode n . Each mode exhibits geometric spreading (i.e., $1/\sqrt{k_n(\omega)r}$) and a frequency response (i.e., $G_n(\omega)$). The wavenumber function $k_n(\omega)$ is often referred to as the dispersion curves of the medium. Overall, this model assumes Lamb waves travel as a sum of multiple isotropic wave modes and each mode is modeled as a plane wave with a frequency-dependent and mode-dependent wavenumber $k_n(\omega)$. Note that the wavenumber is inversely proportional to the wave's phase velocity.

Sparse wavenumber analysis utilizes (2) expressed in a linear algebraic representation

$$\mathbf{X} = \Phi \mathbf{V}. \quad (3)$$

In this expression, \mathbf{X} is an $M \times Q$ matrix of frequency-domain measurements, \mathbf{V} is a sparse $N \times Q$ dispersion curve matrix, and Φ is an $M \times N$ propagation matrix such that

$$\mathbf{X} = [X_m(\omega_q)]_{mq} \quad (4)$$

$$\mathbf{V} = [G_n(\omega_q)]_{nq} \quad (5)$$

$$\Phi = \left[\sqrt{\frac{1}{\kappa_n r_m}} e^{-j\kappa_n r_m} \right]_{mn}. \quad (6)$$

The function $X_m(\omega_q)$ represents our frequency-domain Lamb wave data across $1 \leq m \leq M$ measurements and $1 \leq q \leq Q$ frequencies. The variable κ_n is a discretization of the wavenumber domain across $1 \leq n \leq N$ user-chosen wavenumbers. The function $G_n(\omega_q)$ is the dispersion curve in frequency-wavenumber space for $1 \leq n \leq N$ wavenumbers and $1 \leq q \leq Q$ frequencies. Figure 6 illustrates values of the \mathbf{V} matrix that correspond to the dispersion curves of Lamb waves.

Since \mathbf{V} has few non-zero values, we use algorithms from compressive sensing³⁶, such as orthogonal matching pursuit³⁷, to recover \mathbf{V} from knowledge of \mathbf{X} and Φ . This process is known as sparse wavenumber analysis. The detailed process of implementing sparse wavenumber analysis has been discussed extensively in prior work^{20,38}. We refer to our recovered dispersion curves as $\hat{\mathbf{V}}$.

Step 2: Construct Model

Once the sparse wavenumber analysis dispersion curve estimate $\hat{\mathbf{V}}$ is known, we use it to re-solve

$$\hat{\mathbf{X}} = \Phi \hat{\mathbf{V}}. \quad (7)$$

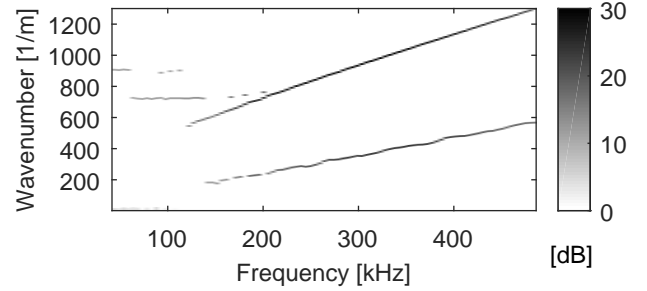


Figure 6. Frequency-wavenumber (i.e., dispersion curves) representation of our experimental wavefield. Frequencies below 100 kHz do not exhibit waves with significant strength.

This process is known as sparse wavenumber denoising (or, more generally, sparse wavenumber synthesis)²⁰. We refer to it as denoising because the process cannot reconstruct information (e.g., noise, damage signatures, anisotropy, etc.) that are not represented in the model in (2). As a result, any information created by damage is not reconstructed in $\hat{\mathbf{X}}$, and $\hat{\mathbf{X}}$ represents a damage-free model of \mathbf{X} . Note that extensive damage and noise can degrade the performance of this process, but this degradation can be reduced by increasing the number of measurements³⁸.

Our model $\hat{\mathbf{X}}$ possesses the velocities and amplitudes defined by the dispersion curves in $\hat{\mathbf{V}}$. Note that to synthesize a baseline signal that can successfully isolate damage signatures requires extremely accurate dispersion curves. In general, theoretical dispersion curves from the Rayleigh-Lamb equation²⁸ are not sufficient for this purpose due to significant uncertainties. In particular, the frequency-dependent complex amplitudes for each wave mode are heavily influenced by the sensors and their effects are generally unknown without measuring them.

As a result, experimentally obtained dispersion curves are essential to our method. Even so, additional errors and uncertainties due to anisotropy and speckle, which dispersion curves do not characterize, may cause our model $\hat{\mathbf{X}}$ to not perfectly match our data \mathbf{X} . Therefore, we apply additional optimizations to further align our model and data.

Step 3: Realign Data (Shift)

Let $\hat{x}_m(t)$ and $\hat{X}_m(\omega)$ be the time-domain and frequency-domain representations of our model signal m (i.e., $\hat{X}_m(\omega)$ represents the m -th column $\hat{\mathbf{X}}$). Let $x_m(t)$ and $X_m(\omega)$ be the corresponding signals from our measured data set. In this step, we shift-align these two signals.

Specifically, we shift our model $\hat{x}_m(t - \tau)$ by some time delay τ such that

$$\tau = \arg \min_{\tau} \int_{-\infty}^{\infty} |x_m(t) - \hat{x}_m(t - \tau)|^2 dt, \quad (8)$$

or equivalently,

$$\tau = \arg \min_{\tau} \int_{-\infty}^{\infty} |X_m(\omega) - \hat{X}_m(\omega)e^{-j\tau\omega}|^2 d\omega. \quad (9)$$

The optimal shift τ is determined by solving this least squares optimization in the frequency domain. We solve the problem in the frequency domain since τ in (9) can be any real number and is not restricted to an integer sample as in (8). This optimization problem is not generally convex (i.e., it may have local minima). However, since the model and data are very similar, the global minimum solution is likely to be locally convex around $\tau = 0$. After τ is determined, the model signal is shifted accordingly. This is repeated for each model signal and data signal pair m .

Note that the shift alignment is intended to adjust for small shifts between the data and the model. Therefore, to remove the possibility of large, erroneous shifts, we recommend the optimization be constrained to only consider shifts smaller than a quarter of a period. In our setup with a 250 kHz center frequency signal and a 5 MHz sampling rate, this is equivalent to approximately constraining τ to be between -0.001 ms to 0.001 ms, or -5 to 5 samples. In our results, all of the optimal shift estimates are less than 0.0005 ms, or 2.5 samples.

Step 4: Realign Data (Stretch)

Optimal time shifting compensates for errors due to anisotropy and speckle. Yet, an error in velocity or distance for guided wave signals is often also represented by a time-stretch^{22,39}. Therefore, to further align our model and data, we optimally time-stretch our model $\hat{x}_m(\alpha t - \tau)$ by some stretching factor α such that

$$\alpha = \arg \min_{\alpha} \int_{-\infty}^{\infty} |x_m(t) - \hat{x}_m(\alpha t - \tau)|^2 dt. \quad (10)$$

There are a few common approaches for performing this optimization. Baseline signal stretch⁴⁰ aligns these signals by stretching $\hat{x}_m(t - \tau)$ by variety of α factors and then finding the α that minimizes the error. The scale transform approach²² performs a similar optimization in the scale domain, which allows the algorithm to be much more computationally efficient.

The scale transform solves a slightly different optimization than in (10) such that $x_m(t)$ and $\hat{x}_m(\alpha t - \tau)$ are energy normalized²². After α is determined, the model signal is stretched accordingly. This is repeated for each model signal and data signal pair m . We optimize for shifting and stretching individually for computational efficiency. The shift estimation is accomplished with the fast Fourier transform. The stretch estimate is accomplished with the fast Mellin transform²². Other optimizations can be applied to estimate both factors simultaneously, but for our experimental data, does not provide significant improvement.

Note that the stretch alignment is intended to account for small stretches between the data and the model. Therefore,

to remove the possibility of erroneous stretching, we recommend the optimization be constrained to not consider stretch factors $1 - (T_0/4)/L < \alpha < 1 + (T_0/4)/L$, where T_0 is a period of the signal and L is the total length of the signal. This assures that no part of the data is shifted by more than a quarter of a period. In our setup with a 250 kHz center frequency signal and approximately 20 ms length signal, this is equivalent to approximately $0.995 < \alpha < 1.005$. In our results, all of the optimal stretch factor estimates satisfy $0.9999 < \alpha < 1.0001$.

Step 5: Amplitude Alignment

Uncertainties, such as speckle, cause local time variations in signal amplitude that differ from our model. To remove these differences, we determine a function $\beta(t)$ such that when multiplied by $\hat{x}_m(\alpha t - \tau)$ results in an optimal match with our data $x_m(t)$. Note that this amplitude correction must be accomplished with caution to avoid over-fitting to the damage signature. Therefore, we choose $\beta(t)$ such that

$$\beta(T) = \frac{\int_{-\infty}^{\infty} w(t, T) x(t) \hat{x}(\alpha t - \tau) dt}{\int_{-\infty}^{\infty} |w(t, T) \hat{x}(\alpha t - \tau)|^2 dt}, \quad (11)$$

where $w(t, T)$ is a rectangular window of width W

$$w(t, T) = \begin{cases} 1 & , \quad T - W/2 \leq t \leq T + W/2 \\ 0 & , \quad \text{otherwise} \end{cases} \quad (12)$$

In (11), the T value represents the center of the data window. Therefore, $\beta(T)$ represents an amplitude adjustment estimate for a range of data centered around time T and with a time width of W . Mathematically, $\beta(T)$ represents the least-squares estimate corresponding to the solution of

$$\beta(T) = \arg \min_{\beta(T)} \int_{T-W/2}^{T+W/2} |x(t) - \beta(T) \hat{x}(\alpha t - \tau)|^2 dt \quad (13)$$

for a moving window of width W that is centered around time T . Once $\beta(T)$ is determined, we multiply it with our model to create an updated model. This is repeated for each model signal and data signal pair m .

Amplitude alignment minimizes the least-squares error between windowed sections of $x(t)$ and $\hat{x}(\alpha t - \tau)$ for each sample. Therefore, we can inadvertently remove damage signatures, particularly for a relatively small window size (smaller than a period). For relatively large windows (larger than a period), the optimization can increase the error between $x(t)$ and $\hat{x}(\alpha t - \tau)$ since it minimizes the error across the window rather than at the sample. This increase in error is usually small since the least-squares amplitude estimate needs to be consistent across the entire window.

Step 6: Model Subtraction

The previous steps find the global dispersion curves and optimal parameters τ , α , and $\beta(t)$ such that each measurement of our data $x_m(t)$ best fits our modified model $\beta(t) \hat{x}_m(\alpha t - \tau)$. To then isolate the damage signature, we perform model subtraction

$$z_m(t) = x_m(t) - \beta(t) \hat{x}_m(\alpha t - \tau). \quad (14)$$

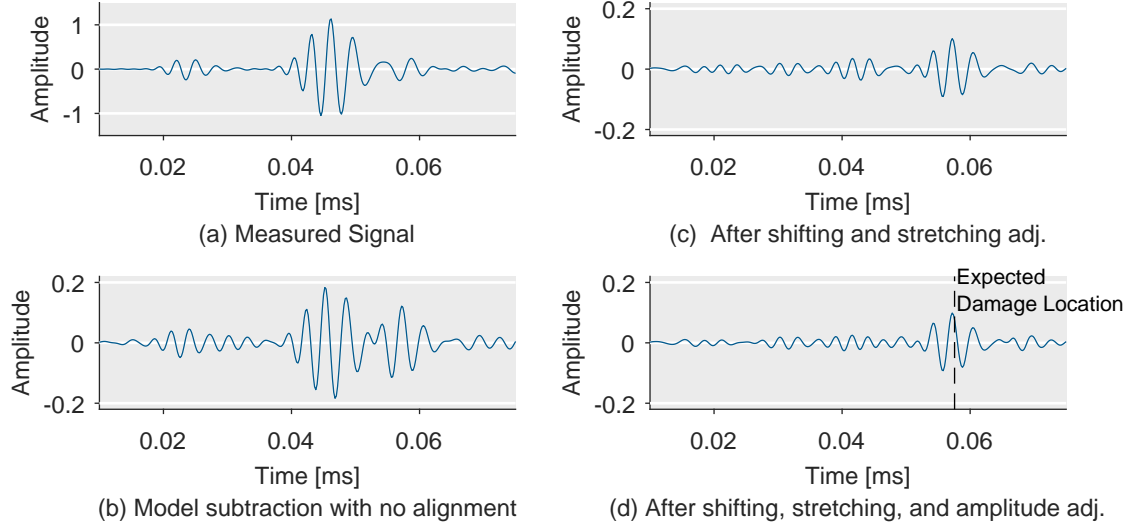


Figure 7. An illustration of the effects of our alignment procedure. (a) Raw measurement with no alignment. (b) Measurement after subtracting our sparse wavenumber model. (c) Subtraction after shift and stretch alignment. (d) Subtraction after shift, stretch, and amplitude alignment. This plot clearly shows the damage reflection at 0.55 ms that is less prominent in (b).

The model-subtracted signals $z_m(t)$ will therefore contain information not fitted to our model. In ideal, noise-free conditions, $z_m(t)$ would contain only the damage signature.

Figure 7 illustrates the effects of our optimal alignment techniques. Figure 7(a) shows one Lamb wave measurement before model subtraction and Figure 7(b) illustrates the signal after model subtraction (with no alignment). In both plots, we observe a large, similar pulse at around 0.045 ms, indicating possible poor model subtraction. Figure 7(c) illustrates the subtraction after shift-aligning and stretch-aligning the model. In this plot, the large pulse is removed and a smaller pulse around 0.055 ms, which corresponds to damage, is now more prominent. Figure 7(d) illustrates the subtraction after amplitude alignment (with a $W = 0.02$ ms window) as well. This further reduces some undesirable data before the 0.055 ms pulse.

Statistical Detection and Localization

We detect and locate damage in the structure through matched field processing^{17,38}. Matched field processing is a model-driven localization framework that compares the signal we expect from damage at pixel p (i.e., a matched field model) with the actual data received after baseline or model subtraction. This comparison is accomplished through a matched field processor, an extension of the matched filter⁴¹.

Matched field processing has been used extensively in underwater acoustics⁴², radar⁴³, and recently introduced to structural health monitoring³⁸. An advantage of matched field processing is its flexibility and extensibility since it is based on the matched filter^{16,41}. In this paper, we use a data-driven matched field processor³⁸ that obtains its model from sparse wavenumber analysis.

Figure 8 summarizes the matched field processing methodology that we describe in this section. In this block diagram, \mathbf{X} is an $M \times Q$ matrix of frequency-domain measurements, \mathbf{Z} is an $M \times Q$ matrix of frequency-domain

model-subtracted measurements, $\hat{\mathbf{V}}$ is a sparse $N \times Q$ dispersion curve matrix, and \mathbf{Y}_p is an $M \times Q$ matrix of our matched field model for pixel p . The b_p value represents our statistical measure of damage at pixel p while $b_p^{(b)}$ represents an estimate of statistical bias at pixel p .

For our statistical analysis, we assume the data matrix, the model-subtracted data matrix, and the matched field model matrix are represented by collections of vectors such that

$$\mathbf{X} = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \mathbf{x}_3^T \ \dots \ \mathbf{x}_M^T]^T \quad (15)$$

$$\mathbf{Z} = [\mathbf{z}_1^T \ \mathbf{z}_2^T \ \mathbf{z}_3^T \ \dots \ \mathbf{z}_M^T]^T \quad (16)$$

$$\mathbf{Y}_p = [\mathbf{y}_{p,1}^T \ \mathbf{y}_{p,2}^T \ \mathbf{y}_{p,3}^T \ \dots \ \mathbf{y}_{p,M}^T]^T. \quad (17)$$

In these expressions, \mathbf{x}_m , \mathbf{z}_m , and $\mathbf{y}_{p,m}$ represent the m -th frequency-domain measurement in the data matrix, the model-subtracted data matrix, and the matched field model matrix, respectively.

The Matched Field Model

Our matched field model \mathbf{Y}_p for pixel p is created from our sparse wavenumber analysis dispersion curves $\hat{\mathbf{V}}$. We assume damage acts as a collection of point scatterers. Therefore, we model signals that travel distances from each transmit location on the plate to each pixel p in our image combined with the distances from each pixel p in our image to each receiver or measurement location. In this paper, we consider M transmit locations across the plate and a single receiver. This is theoretically equivalent to considering one transmitter and M receivers. This damage model, which is also the basis for many sparse array and phased array methods, is ideal for locating damage that reflects waves in all directions, such as cracks⁴⁴ and holes^{15,45}. Similar models have also been applied to locate weaker reflectors, such as composite delaminations^{23,46}.

The sum of these distances from the transmitters to the receiver for each pixel p and each measurement m

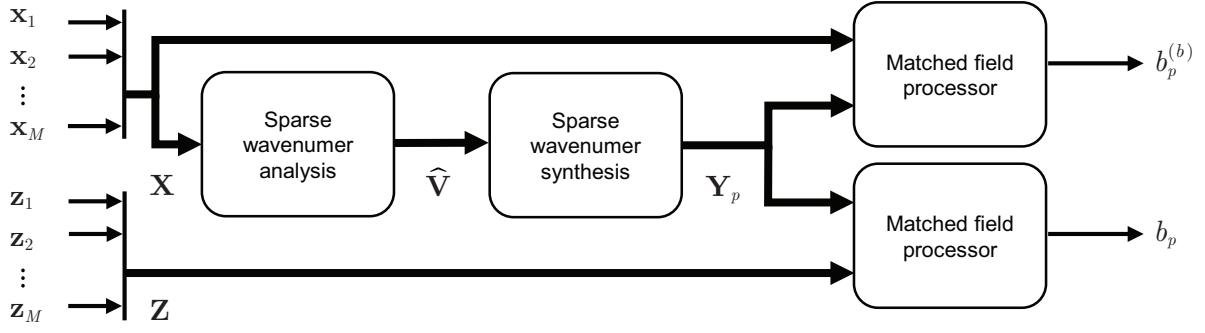


Figure 8. Illustration of our statistical matched field processor from our measured partial wavefield \mathbf{X} and our model-subtracted data \mathbf{Z} for pixel p . The output of this process is the image value, or ambiguity surface, b_p at pixel p and the potential bias value $b_p^{(b)}$ at pixel p .

are defined as $\rho_{p,m}$. These distances are used by sparse wavenumber synthesis to create the matched field model. Specifically, we solve

$$\tilde{\Phi}_p = \left[\sqrt{\frac{1}{\kappa_n \rho_{p,m}}} e^{-j\kappa_n \rho_{p,m}} \right]_{mn} \quad (18)$$

$$\mathbf{Y}_p = \tilde{\Phi}_p \hat{\mathbf{V}}. \quad (19)$$

In this scenario, $\tilde{\Phi}_p$ is a matrix with the same design as Φ in (4), but the distances now correspond to the sum of the distances from transmitter m , to location p , and then to our receiver. The process for building this model is discussed in more depth in ³⁸.

The Matched Field Processor

We define b_p as the intensity of our image (also known as the ambiguity surface) at pixel p . We choose a matched field processor to compute b_p based on statistical principles. Specifically, we compute each pixel intensity b_p as

$$b_p = M \frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2} \quad (20)$$

$$\hat{\mu}_p^2 = \left| \frac{1}{M} \sum_{m=1}^M \mathbf{y}_{p,m}^H \mathbf{z}_m \right|^2 \quad (21)$$

$$\hat{\sigma}_p^2 = \frac{1}{M-1} \sum_{m=1}^M \left| \mathbf{y}_{p,m}^H \mathbf{z}_m - \frac{1}{M} \sum_{m=1}^M \mathbf{y}_{p,m}^H \mathbf{z}_m \right|^2 \quad (22)$$

Conceptually, b_p is a measure of similarity between the model-subtracted data \mathbf{z}_m and our model $\mathbf{y}_{p,m}$ for damage at a given pixel p . When b_p is large, our model closely matches the data and there is a high likelihood that damage is present at that pixel.

Statistically, b_p is the ratio of the squared sample mean to sample variance of $\mathbf{y}_{p,m}^H \mathbf{z}_m$. The value $\mathbf{y}_{p,m}^H \mathbf{z}_m$ is the inner product between our model-subtracted data and our matched field model. The inner product is large when the data \mathbf{z}_m closely matches the model $\mathbf{y}_{p,m}$. Therefore, the pixel intensity b_p is high and denotes damage when the inner product mean is high and the variance is low. If the variance across measurements is high, then the result is considered uncertain and b_p is small.

Note that matched field processing analyzes the raw, high frequency information. This is in contrast with other

guided wave localization methods, such as delay-and-sum^{45,47}, that process a strictly positive envelope of the data. Compared with envelope-based approaches, we can achieve significantly higher resolutions¹⁵. In addition, errors in baseline subtraction cause fewer false alarms since the errors have a high probability to sum incoherently (i.e., summing negative and positive values that reduce the magnitude) at a given pixel while values at the damage location have a high probability to sum coherently (i.e., summing values that constructively increase the magnitude). Finally, since the errors sum incoherently, we can assume pixels outside of the damage region can be represented by a zero-mean random variable. This is an important assumption for our statistical analysis.

We define the estimated location of the damage as the pixel p that maximizes b_p ,

$$p^* = \arg \max_p b_p. \quad (23)$$

Note that while p^* always exists, it may not be a statistically significant value. Therefore, we derive a hypothesis test to ensure statistical significance of the maximum value of b_p .

Hypothesis Test

In (20), the expression for b_p is a type of CFAR (constant rate of false alarm) matched filter⁴⁸. The CFAR matched filter is an optimal statistic for testing the hypothesis

$$\mathcal{H}_0 : \mu_p = 0 \quad (24)$$

$$\mathcal{H}_1 : \mu_p \neq 0, \quad (25)$$

where $\mathbf{y}_{p,m}^H \mathbf{z}_m \sim N(\mu_p, \sigma_p^2)$ and σ_p^2 is an unknown quantity. Using this statistic assumes that when damage is absent (i.e., $\mu_p = 0$), our model subtraction \mathbf{z}_m is characterized as white Gaussian noise with equal unknown variances across each measurement m . Note that this assumption is relaxed in the next subsection.

In our scenario, the null hypothesis \mathcal{H}_0 represents the damage absent scenario and the alternative hypothesis \mathcal{H}_1 represents the damage present scenario. Statistically, we could choose the alternative hypothesis when

$$b_p > \gamma_0, \quad (26)$$

where γ_0 is a statistical threshold based on a desired false alarm rate and the distribution of b_p .

While we could use this statistical test to detect damage at each pixel, it would be ineffective. For example, if we had 10,000 pixels and a desired false alarm rate of 1%, we would expect 100 pixels to surpass the threshold γ . Furthermore, the pixel intensities are not independent of each other. Consequentially, the true number of pixels above the threshold could be much larger.

Therefore, for a collection of pixels, we instead choose the alternative hypothesis when

$$\max_p b_p > \gamma, \quad (27)$$

where γ is statistical threshold based on a desired false alarm rate and the distribution of $\max_p b_p$. When damage is absent, each b_p is distributed according to a Fisher distribution (or F-distribution) with degrees of freedom of 2 and $2(M-1)$, i.e., $F(2, 2(M-1))$. The 2's are introduced because the frequency domain has complex values. Then according to extreme value theory, the maximum of an F-distributed random variable follows an extreme value distribution (type 1), also known as a Gumbel distribution²¹. Therefore, the statistic $\max_p b_p$ is distributed according to a Gumbel distribution.

The relationship between the F-distribution parameters and the corresponding Gumbel distribution parameters does not have known closed form. Therefore, we determine the parameters and the threshold γ for the Gumbel distribution through a Monte Carlo simulation. Across 1000 Monte Carlo trials, we compute the maximum value from 50,000 random variables distributed according to an $F(2, 2(M-1))$ distribution. We then use these 1000 maximum values to estimate the Gumbel parameters through maximum likelihood estimation⁴⁹. Finally, we invert the cumulative Gumbel distribution function to find a threshold γ for a given probability of false alarm.

Figure 9 illustrates our estimated thresholds as a function of the number of measurements M and our desired false alarm rate. In this paper, we use a false alarm rate of 0.05. For this false alarm rate, we can fit the curve in Figure 9 to a power function (with $R^2 = 0.98$) such that the threshold can be computed as a function of M where

$$\gamma(M) = 271.5M^{-1.28} + 13.66. \quad (28)$$

Therefore, for $M = 100$ measurements, the threshold is approximately $\gamma = 14.40$.

Estimating Worst-case Bias

In our hypothesis test, we assume $\mathbf{y}_{p,m}^H \mathbf{z}_m$ has a zero mean when damage is absent. Yet, this is generally false due to imperfect model subtraction. When there is no damage, the model-subtracted data \mathbf{z}_m does not show a pure noise signal. Instead, part of the original data remains and that residual information will correlate with our matched field model $\mathbf{y}_{p,m}$ to produce a biased result in b_p .

Therefore, we estimate the worst-case bias for b_p and offset our threshold γ with this estimate. If this bias is due to imperfect subtraction, we assume the worst-case bias occurs when there is no subtraction (i.e., $\mathbf{z}_m = \mathbf{x}_m$ for each m). Under this assumption, the matched field processing result

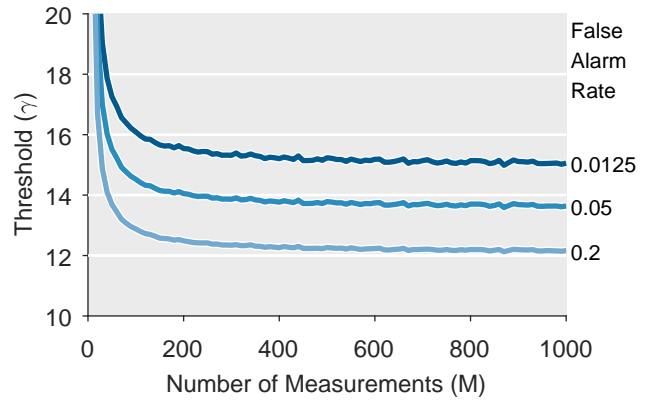


Figure 9. The approximate statistical threshold (computed using Monte Carlo estimation) for a Gumbel distribution as a function of the number of measurements M and false alarm rate.

with the worst-case bias can be computed as

$$b_p^{(b)} = M \frac{\hat{\mu}_p^2}{\hat{\sigma}_p^2} \quad (29)$$

$$\hat{\mu}_p^2 = \left| \frac{1}{M} \sum_{m=1}^M \mathbf{y}_{p,m}^H \mathbf{x}_m \right|^2 \quad (30)$$

$$\hat{\sigma}_p^2 = \frac{1}{M-1} \sum_{m=1}^M \left| \mathbf{y}_{p,m}^H \mathbf{x}_m - \frac{1}{M} \sum_{m=1}^M \mathbf{y}_{p,m}^H \mathbf{x}_m \right|^2 \quad (31)$$

Note that the image $b_p^{(b)}$ represents bias, but the true bias image is likely different. As a result, we consider only the worst-case value of $b_p^{(b)}$, i.e., $\max_p b_p^{(b)}$, as our measure of bias.

We use the worst-case bias estimate to adjust our detection threshold. However, this estimate is a very conservative (i.e., high) estimate of the bias. The maximum value is also not a robust estimate of the bias since a single new value can significantly change it. Therefore, we define a looser estimate that provides a higher, and more robust, probability of detection at the expense of a higher probability of error. Specifically, we consider the α -percentile value of $b_p^{(b)}$, i.e., the value of $b_p^{(b)}$ such that α percent of the $b_p^{(b)}$ values are smaller. We refer to this statistic as $b_{p(\alpha)}^{(b)}$. When α is 100%, $b_{p(\alpha)}^{(b)} = \max_p b_p^{(b)}$.

With an estimate of our bias, we now choose the alternative hypothesis when

$$\max_p b_p > \gamma + b_{p(\alpha)}^{(b)}, \quad (32)$$

where the bias correction $b_{p(\alpha)}^{(b)}$ offsets the original statistical threshold and makes the test more conservative. As a result, the desired false alarm rate is no longer equal to the empirical probability of false alarm. Instead, when α is 100%, we label the desired false alarm rate as the worst-case false alarm rate. That is, we achieve the desired false alarm rate when the true bias is equal to our worst-case estimate.

Experimental Setup

We test our methodology on dense wavefield data collected from a 100 mm by 100 mm region at the center of a 2 mm thick SAE 304 stainless steel plate. The plate is sufficiently large (1000 mm by 1000 mm) such that no reflections are observed in the wavefield. We perform our partial wavefield analysis with subsets of the dense wavefield. The full scan is measured across a uniform grid with a 0.5 mm interval (or pitch) in both the horizontal and vertical directions. This results in a 200 by 200 pixel grid.

A scanning Q-switched Nd:YAG diode-pumped solid state pulsed laser (Advanced Optowave, custom made) is used to non-destructively generate ultrasonic waves in the specimen through local thermoelastic expansion-contraction. The laser has wavelength, beam diameter at the laser exit port, divergence, and pulse duration of 1064 nm, 0.7 mm, 1.6 mrad, and 30 ns, respectively. Scanning is performed with a galvanometric scanner (SCANLAB, SCANcube 7) at a pulse repetition frequency of 100 Hz and approximately 13 ns of duration per pulse. Wave displacements are kept within the linear-elastic regime of the specimen material by using relatively low laser pulse energy (1 mJ) so that elastodynamic reciprocity holds. Due to acoustic reciprocity, waves traveling from the laser excitation point and sensed by a permanently bonded sensor can be interpreted as if generated at the sensor position and traveled to the laser excitation point. Reconstructing the full wavefield from signals generated at each spatial location produces much cleaner, high signal-to-noise ratio⁵⁰ measurements than taken in the reciprocal manner⁹.

The wavefield is then measured using one temporally bonded piezoelectric sensor (Fuji Ceramics, M304A). The sensor is located at $(x = -70 \text{ mm}, y = 170 \text{ mm})$ in our coordinate axis. Note that this is outside of our scanning region. The transmitted signals are narrowly bandpass filtered between approximately 200 kHz to 300 kHz so that they contain only S0 and A0 modes. Each spatial point records 500 samples with a sampling rate of 5 MHz.

To test and evaluate our method, we perform electrical discharge machining to introduce a half-thickness circular hole into the plate with a 2 mm diameter. The center of the hole is at approximately $(x = 52.7 \text{ mm}, y = 50.6 \text{ mm})$ on our grid. The hole is positioned on the side of the plate opposite of our laser. We performed a dense scan of the plate with and without this damage.

Analysis Setup

We assess our methods and data through a Monte Carlo process. We randomly choose M measurements from the dense wavefield, creating a partial wavefield. With these M measurements, we create and align our model with the data, perform matched field processing, and test our statistical hypothesis to detect damage. We repeat process for different numbers of measurements M , using datasets with and without damage. For each M , this process is repeated 600 times so that we can study the statistical reliability of our framework.

Before analyzing the data, each signal is zero-padded to 1024 samples to increase the resolution in the frequency domain. When working in the frequency domain, only

frequencies between approximately 49 kHz and 490 kHz are analyzed, resulting in $Q = 90$ frequency samples. We limit our frequencies to improve the computational speed of our analysis. Our results are not significantly affected by this choice since the measured data is weak outside the chosen frequency range. Note that this frequency range is larger than the signals 3dB bandwidth so that we use all of the available signal frequencies, both inside and outside of the 3dB band. When we perform sparse wavenumber analysis, we use $N = 2164$ uniformly spaced wavenumbers between 2 m^{-1} and 1300 m^{-1} . For our amplitude adjustment algorithm, we use a window with a width of $W = 100 \text{ ms}$.

Results

In this section, we discuss results from applying our method to the experimental data. We start by illustrating our localization results from matched field processing. We then demonstrate our detection results from our hypothesis test.

Localization

Matched field processing. Figures 10(a)-(c) illustrate our matched field processing images (or ambiguity surfaces) b_p with $M = 100$ measurements. The white squares represent the measurement locations. In the image, the intensities shown are capped at our statistical threshold with the bias correction of $\alpha = 95\%$.

Figure 10(a) shows the results when there is no damage. In the image, there is no significantly large values, particularly when compared to our threshold. Note though that some regions could potentially be mistaken for damage if no threshold was defined and the intensity scale was adjusted.

Figure 10(b) shows the result when damage is present. In the image, the damage is clearly visible in the center of our region. Figure 10(c) illustrates a magnified replica of the image. We observe that the largest values are contained within a single region and have well-defined boundaries. Compared with phased array methods, which are also used to quickly scan regions of a structure, our results achieve finer resolutions²³. This result is with consistent with theory since sparsely separated localization methods are not restricted by Rayleigh criteria¹².

Matched field processing with statistical threshold. Figures 10(d)-(f) illustrate our matched field processing images (or ambiguity surfaces) b_p after thresholding the results to remove all values below our statistical threshold with a bias correction of $\alpha = 95\%$. Therefore, these images illustrate all of the statistically significant values generated by matched field processing.

Figure 10(d) shows the results when there is no damage. In this scenario, the image is entirely zero values. That is, no damage is detected within the region of interest.

Figure 10(e) shows the result when damage is present. The statistically significant damage is clearly visible in the center region of the plate. Figure 10(f) illustrates a magnified replica of the image. The statistically significant values are contained within a single region and have well-defined boundaries. This result shows a statistically significant region that can be densely scanned to resolve small features in the damage region. Current imaging techniques with phased

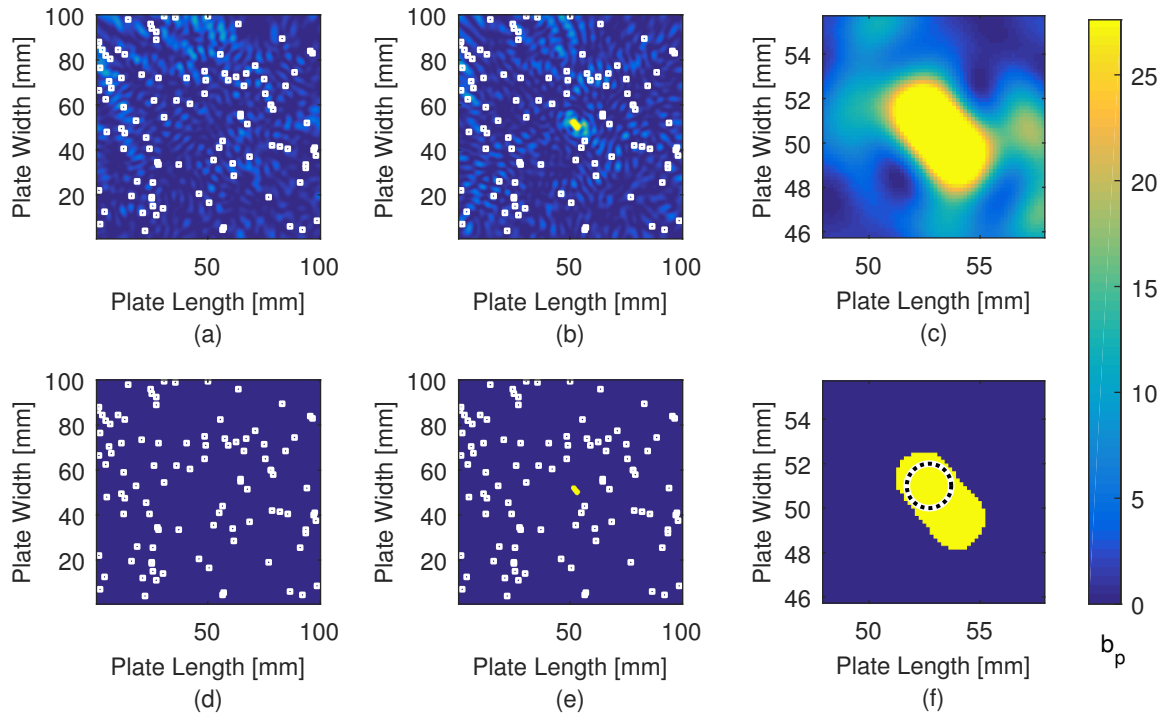


Figure 10. Illustrations of images produced by matched field processing when (a) damage is absent and (b) when damage is present. Subfigure (c) shows the same plot as (b), but magnified around the damage location. Subfigures (d)-(f) illustrate the same plots as (a)-(c), respectively, but values below our statistical threshold with 95% bias correction are removed. White squares represent $M = 100$ randomly chosen measurement locations.

array and sparse arrays do not provide the same statistical guarantees.

In Figure 10(f), a dotted line is shown to represent the approximate true location of the hole. The statistically significant values overlap with the hole and could be considered a rough image of the damage. We produce additional statistically significant values outside the true area of the damage due to acoustic shadowing and diffraction^{51–53}. These effects could potentially be reduced or removed by using multiple receivers to gain multiple perspectives on the damage or improving the damage model. Prior work has shown the ability to improve localization performance by incorporating diffusion⁵⁴ and scattering⁵⁵ properties in matched field processing and similar localization methods.

Matched field processing without alignment. Figure 11 illustrates the effects of not applying our alignment process before matched field processing. We use $M = 100$ measurements in this example. Figure 11(a) shows the localization of damage when no alignment is applied. In this circumstance, the damage is barely visible and no statistically significant values are shown because the damage signatures are not significantly prominent in the model-subtracted data. In Figure 11(b), shift and stretch alignment has been applied and statistically significant damage is visible. In Figure 11(c), amplitude alignment has also been applied. The amplitude alignment makes the intensity values more uniform across the damage region.

Localization performance. Figure 12 illustrates the localization distance error of matched field processing as a function of the number of measurements M . Specifically, the error is the distance between our estimated damage location

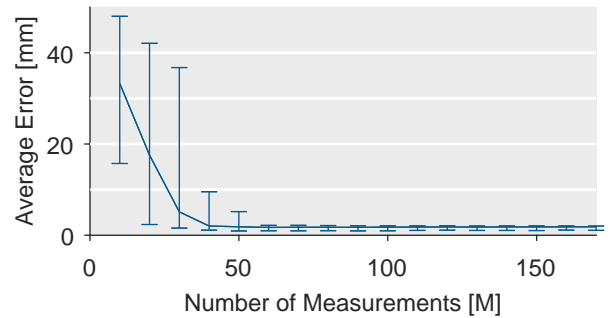


Figure 12. The mean localization error, distance from estimated damage location and the true damage location, as a function of number of measurements M . The error bars represent plus and minus one standard deviation away from the mean.

p^* and the center of the hole. The result shows that our mean localization error is below 2 mm (the diameter of the hole) with 40 measurements ($\Delta s \approx 6.3\lambda_N$) or more. For 30 measurements ($\Delta s \approx 7.3\lambda_N$), our mean localization error is below 5 mm (2.5 times the diameter of the hole). This performance is comparable to 8 sensor sparse array guided wave methods that can collect and process 28 unique measurements. Note that unlike typical sparse array methods, our approach does not use prior baseline data or an *in situ* sensor setup. For large M , the error remains relatively steady between 1.7 mm and 1.9 mm. Due to shadowing and diffusion, the maximum value appears slightly behind the true damage location.

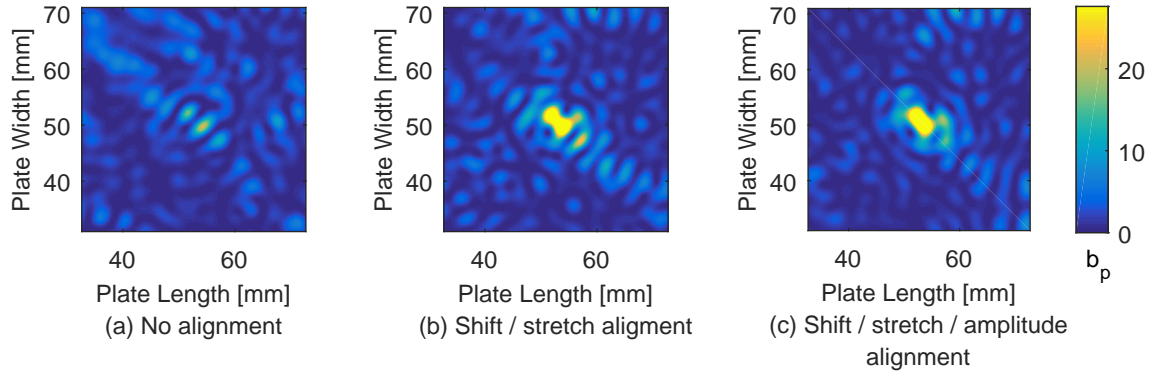


Figure 11. Illustrations of the effect of signal alignment on matched field processing on our plate with damage. (a) Shows the result with no alignment. (b) Shows the result with shift and stretching alignment. (c) Shows the result with shift, stretch, and amplitude alignment.

Statistical Detection

Bias Computation. Figure 13 illustrates example bias images from applying matched field processing to the original data. Figure 13(a) shows the result when damage is absent. Figure 13(b) shows the result when damage is present.

Based on these plots, the presence of damage, which creates a much weaker signal than the direct wavefield, does not significantly affect the bias image. The plots also show that bias increases as we get closer to receiver. This is reasonable since pixels closer to the receiver are associated with shorter travel times and higher amplitudes when compared with pixels far from the receiver. The α -percentile highest values from these images are used as the value of $b_{p(\alpha)}^{(b)}$.

Statistical Threshold Computation. Figure 14 illustrates the mean maximum image intensity value $\max b_p$ when damage is present and when damage is absent. In the figure, we compare these values with three thresholds as a function of M .

Figure 14(a) compares our maximum image intensities to γ with no bias correction (α of 0%). The maximum intensity from the plate with damage increases rapidly and the threshold value converges as M increases. The maximum intensity is, on average, greater than our threshold for $M \geq 30$. This corresponds well with our localization results in Figure 12. However, the maximum intensity with damage absent also steadily increases due to bias. On average, this damage-absent intensity value is greater than our threshold for $M \geq 140$. As a result, bias correction is necessary.

Figure 14(b) compares our maximum image intensities to γ with a bias correction of $\alpha = 100\%$. For small M , the threshold is dominated by γ . For large M , the threshold is dominated by the bias correction and increases with M . Since our bias correction is very conservative, we achieve no false alarms in this scenario. The maximum intensity with damage present, on average, is greater than the average threshold when $M \geq 60$.

Figure 14(c) compares our maximum image intensities to γ with a bias correction of $\alpha = 95\%$. For large M , the threshold is still dominated by the bias correction and

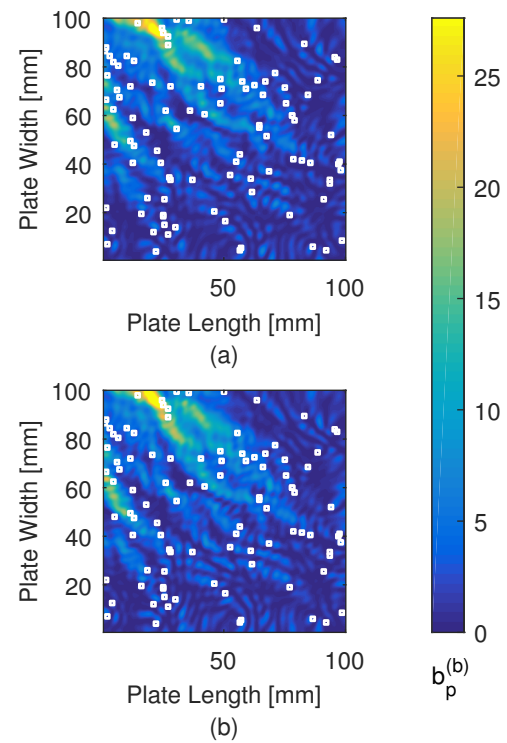


Figure 13. Two example bias images generated from plate data. (a) Shows the result using data from the damage-absent plate. (b) Shows the result using data from damage-present plate. White squares represent $M = 100$ randomly chosen measurement locations.

increases with M . We observe no more than one false alarm across 600 trials (an empirical false alarm rate of ≤ 0.0017 , or 0.17%). Furthermore, the maximum intensity with damage present, on average, is now greater than the average threshold when $M \geq 40$. As a result, we can achieve significantly better detection accuracy with a relatively insignificant increase in false alarm rate by decreasing α from 100% to 95%.

Detection Performance. Figure 15(a) illustrates our hypothesis test's empirical probability of detection and Figure 15(b) shows our empirical probability of false alarm as functions of the number of measurements M .

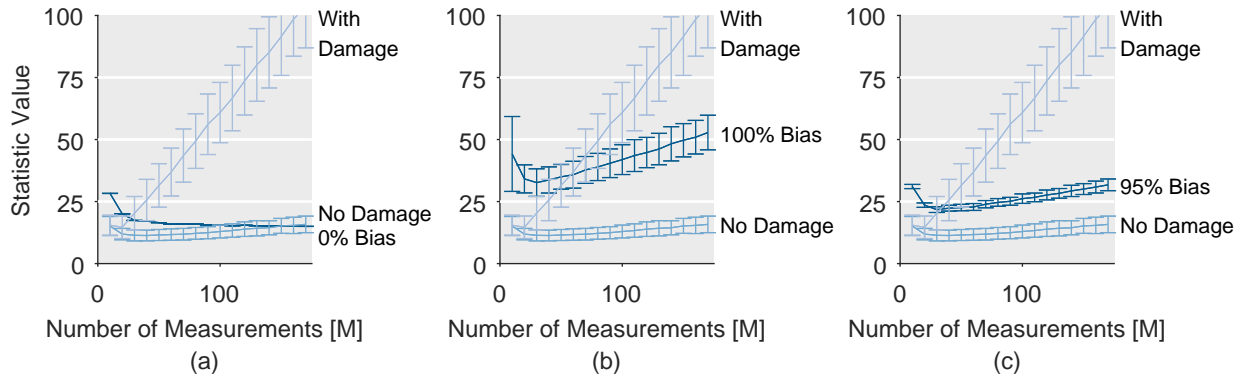


Figure 14. Three plots comparing the thresholds of data with the measured maximum intensities as a function of the number of measurements M . The “with damage” refers to the maximum intensities from the damage-present plate. The “no damage” label refers to the maximum intensities from the damage-absent plate. The $\alpha\%$ threshold refers to the threshold with $\alpha\%$ bias correction. (a) Compares the values with the 0% bias correction. (b) Compares the values with the 100% bias correction. (c) Compares the values with the 95% bias correction.

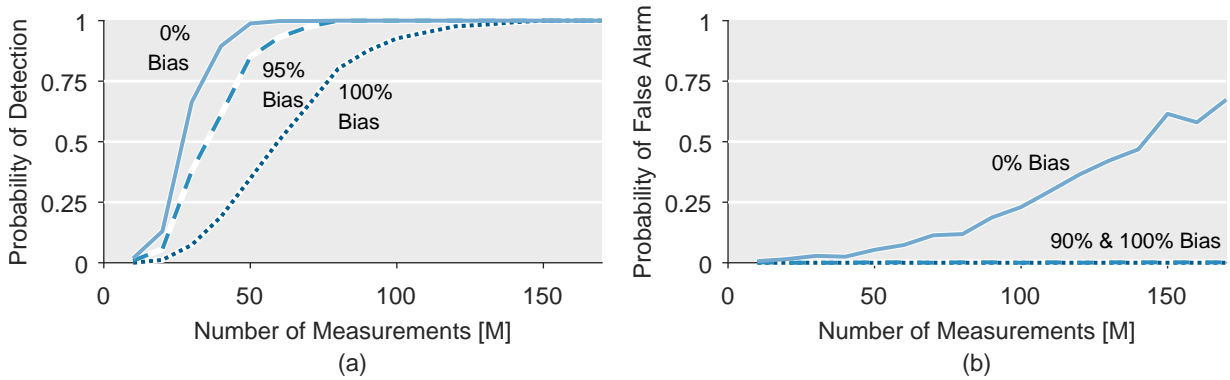


Figure 15. Empirical probability of detection in (a) and empirical probability of false alarm in (b) as a function of number of measurements M for three thresholds using a 100% bias correction, 95% bias correction, and 0% bias correction.

The empirical probability of detection is computed from 600 Monte Carlo trials for each M . We illustrate the empirical probabilities for the threshold of γ with no bias correction, a bias correction of $\alpha = 95\%$, and a bias correction of $\alpha = 100\%$.

The results in Figure 15(a) show that the empirical probability of detection monotonically increases with M for all three thresholds. For no bias correction, the probability of detection is greater than 0.98 for $M \geq 50$ ($\Delta s \gtrsim 5.7\lambda_N$). For the most conservative threshold (α of 100%), the curve shifts significantly to the right and achieves $\geq 98\%$ detection accuracy with $M \geq 140$ ($\Delta s \gtrsim 3.4\lambda_N$). Choosing an α of 95% significantly improves the result, achieving a $\geq 98\%$ detection accuracy using $M \geq 70$ ($\Delta s \gtrsim 4.8\lambda_N$).

The results in Figure 15(b) show that the empirical probability of false alarm monotonically increases with M for only the 0% bias. The threshold with 100% bias correction observes no false alarms. The threshold with 95% bias correction intermittently observes empirical false alarm rates of 0.0017. Note that we have tested the 100% and 95% thresholds for values of M as large as 1000 and do not observe larger false alarm rates.

Conclusions

This paper presents a framework for statistically detecting and imaging damage with partial wavefields. The framework

requires no prior baseline measurements and instead uses a sparse wavenumber model of the wavefield to perform model subtraction. We presented several algorithms to optimally fit the sparse wavenumber model with our data before the subtraction. We then presented a statistical detection and imaging strategy based on integrating matched field processing with extreme value theory.

The results demonstrate that we can achieve a nearly perfect statistical damage detection ($\geq 98\%$ detection rate with $\leq 0.17\%$ false alarm rate) with as few as 70 partial wavefield measurements. This number is significantly smaller (approximately 570 times smaller) than the approximately 40,000 that would be measured in a dense wavefield. Furthermore, the average spatial sampling interval is approximately 4.8 times larger than the Nyquist wavelength. As a result, we can detect damage in a fraction of the time needed by dense wavefield methods.

Even with so few measurements, we also still achieve a high resolution image of the damage. Neglecting the effects of shadowing and diffusion, our matched field processing results achieve accurate images of the size and shape of the damage. Furthermore, we achieve accurate localization of the damage with as few as 40 measurements with an average spatial sampling interval as large as 6.3 times larger than the Nyquist wavelength.

There are many opportunities to expand this work. Specifically, we can improve the imaging of damage by incorporating diffusion and scattering models into matched field processing. Alternative computational imaging methods may also improve the results. In addition, we can extend this framework to function with other wavefield models^{24,56} since our sparse wavenumber model is restricted to mostly-isotropic wavefields with no reflections.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grant no. CMMI-1562838.

References

1. Staszewski WJ, Lee BC, Mallet L et al. Structural health monitoring using scanning laser vibrometry: I. Lamb wave sensing. *Smart Mater Struct* 2004; 13(2): 251–260.
2. Ruzzene M. Frequency–wavenumber domain filtering for improved damage visualization. *Smart Mater Struct* 2007; 16(6): 2116.
3. Yu L, Leckey CAC and Tian Z. Study on crack scattering in aluminum plates with lamb wave frequency–wavenumber analysis. *Smart Mater Struct* 2013; 22(6): 065019.
4. Dawson AJ, Michaels JE and Michaels TE. Isolation of ultrasonic scattering by wavefield baseline subtraction. *Mech Syst Signal Process* 2016; 70–71: 891–903.
5. Sohn H, Dutta D, Yang JY et al. Automated detection of delamination and disbond from wavefield images obtained using a scanning laser vibrometer. *Smart Mater Struct* 2011; 20(4): 045017.
6. Mesnil O, Leckey CAC and Ruzzene M. Instantaneous and local wavenumber estimations for damage quantification in composites. *Structural Health Monitoring* 2015; 14(3): 193–204.
7. Ianni TD, Marchi LD, Perelli A et al. Compressive sensing of full wave field data for structural health monitoring applications. *IEEE Trans Ultrason Ferroelectr Freq Control* 2015; 62(7): 1373–1383.
8. Michaels JE, Chimenti DE and Bond LJ. Ultrasonic wavefield imaging: Research tool or emerging NDE method? *AIP Conf Proc* 2017; 1806(1): 020001.
9. Lee JR, Shin HJ, Chia CC et al. Long distance laser ultrasonic propagation imaging system for damage visualization. *Opt Lasers Eng* 2011; 49(12): 1361–1371.
10. Lee JR, Chia CC, Park CY et al. Laser ultrasonic anomalous wave propagation imaging method with adjacent wave subtraction: Algorithm. *Opt Laser Technol* 2012; 44(5): 1507–1515.
11. Yu L and Tian Z. Guided wave phased array beamforming and imaging in composite plates. *Ultrasonics* 2016; 68: 43–53.
12. Haimovich AM, Blum RS and Cimini LJ. MIMO radar with widely separated antennas. *IEEE Signal Process Mag* 2008; 25(1): 116–129.
13. Mesnil O and Ruzzene M. Sparse wavefield reconstruction and source detection using compressed sensing. *Ultrasonics* 2016; 67: 94–104.
14. Michaels JE. Detection, localization and characterization of damage in plates with an in situ array of spatially distributed ultrasonic sensors. *Smart Mater Struct* 2008; 17(3): 035035.
15. Harley JB and Moura JMF. Data-driven matched field processing for lamb wave structural health monitoring. *J Acoust Soc Am* 2014; 135(3): 1231–1244.
16. Giannakis GB and Tsatsanis MK. Signal detection and classification using matched filtering and higher order statistics. *IEEE Trans Acoust* 1990; 38(7): 1284–1296.
17. Baggeroer AB, Kuperman WA and Schmidt H. Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem. *J Acoust Soc Am* 1988; 83(2): 571.
18. Flynn EB, Todd MD, Wilcox PD et al. Maximum-likelihood estimation of damage location in guided-wave structural health monitoring. *P Roy Soc A Mat* 2011; .
19. Flynn EB and Todd MD. A bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. *Mech Syst Signal Process* 2010; 24(4): 891–903.
20. Harley JB and Moura JMF. Sparse recovery of the multimodal and dispersive characteristics of lamb waves. *J Acoust Soc Am* 2013; 133(5): 2732–2745.
21. Castillo E. *Extreme Value Theory in Engineering*. Elsevier, 2012.
22. Harley JB and Moura JMF. Scale transform signal processing for optimal ultrasonic temperature compensation. *IEEE Trans Ultrason Ferroelectr Freq Control* 2012; 59(10): 2226–2236.
23. Tian Z, Yu L and Leckey C. Rapid guided wave delamination detection and quantification in composites using global-local sensing. *Smart Mater Struct* 2016; 25(8): 085042.
24. Harley JB. Predictive guided wave models through sparse modal representations. *Proc IEEE* 2016; 104(8): 1604–1619.
25. Hall JS and Michaels JE. Model-based parameter estimation for characterizing wave propagation in a homogeneous medium. *Inverse Probl* 2011; 27(3): 035002.
26. Wu TT and Liu YH. Inverse determinations of thickness and elastic properties of a bonding layer using laser-generated surface waves. *Ultrasonics* 1999; 37(1): 23–30.
27. Raghavan A and Cesnik CES. Effects of elevated temperature on guided-wave structural health monitoring. *J Intell Mater Syst Struct* 2008; 19(12): 1383–1398.
28. Graff K. *Wave Motion in Elastic Solids*. 1st ed. ed. New York: Dover Publications, 1975.
29. Thomsen L. Weak elastic anisotropy. *Geophysics* 1986; .
30. Wang T, Zhu T, Sun J et al. Influence of rolling directions on microstructure, mechanical properties and anisotropy of Mg-5Li-1Al-0.5Y alloy. *J Magnesium Alloys* 2015; 3(4): 345–351.
31. Andria G, Attivissimo F, Lanzolla AML et al. A suitable threshold for speckle reduction in ultrasound images. *IEEE Trans Instrum Meas* 2013; 62(8): 2270–2279.
32. Senesi M and Ruzzene M. A frequency selective acoustic transducer for directional lamb wave sensing. *J Acoust Soc Am* 2011; 130(4): 1899–1907.
33. Stähler SC, Sens-Schönfelder C and Niederleithinger E. Monitoring stress changes in a concrete bridge with coda wave interferometry. *J Acoust Soc Am* 2011; 129(4): 1945–1952.
34. Gandhi N, Michaels JE and Lee SJ. Acoustoelastic lamb wave propagation in biaxially stressed plates. *J Acoust Soc Am* 2012; 132(3): 1284–1293.
35. Schubert KJ and Herrmann AS. On attenuation and measurement of lamb waves in viscoelastic composites. *Compos Struct* 2011; 94(1): 177–185.

36. Donoho DL. Compressed sensing. *IEEE Trans Inf Theory* 2006; 52(4): 1289–1306.
37. Tropp JA. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans Inf Theory* 2004; 50(10): 2231–2242.
38. Harley JB and Moura JMF. Dispersion curve recovery with orthogonal matching pursuit. *J Acoust Soc Am* 2015; 137(1): EL1–EL7.
39. Lu Y and Michaels JE. A methodology for structural health monitoring with diffuse ultrasonic waves in the presence of temperature variations. *Ultrasonics* 2005; 43(9): 717–731.
40. Croxford AJ, Moll J, Wilcox PD et al. Efficient temperature compensation strategies for guided wave structural health monitoring. *Ultrasonics* 2010; 50(4-5): 517–528.
41. Dosso SE and Wilmut MJ. Maximum-likelihood and other processors for incoherent and coherent matched-field localization. *J Acoust Soc Am* 2012; 132(4): 2273–2285.
42. Baggeroer AB, Kuperman WA and Mikhalevsky PN. An overview of matched field methods in ocean acoustics. *IEEE J Oceanic Eng* 1993; 18(4): 401–424.
43. Papazoglou M and Krolik JL. Matched-field estimation of aircraft altitude from multiple over-the-horizon radar revisits. *IEEE Trans Signal Process* 1999; 47(4): 966–976.
44. Chen X, Michaels JE, Lee SJ et al. Load-differential imaging for detection and localization of fatigue cracks using lamb waves. *NDT E Int* 2012; 51: 142–149.
45. Hall J and Michaels JE. Minimum variance ultrasonic imaging applied to an in situ sparse guided wave array. *IEEE Trans Ultrason Ferroelectr Freq Control* 2010; 57(10): 2311–2323.
46. Sharif-Khodaei Z and Aliabadi MH. Assessment of delay-and-sum algorithms for damage detection in aluminium and composite plates. *Smart Mater Struct* 2014; 23(7): 075007.
47. Clarke T and Cawley P. Enhancing the defect localization capability of a guided wave SHM system applied to a complex structure. *Struct Health Monit* 2010; 10(3): 247–259.
48. Scharf LL and Kraut S. Geometries, invariances, and SNR interpretations of matched and adaptive subspace detectors. *TS Traitement du signal* 1998; 15(6): 527–534.
49. Mahdi S and Cenac M. Estimating parameters of gumbel distribution using the methods of moments, probability weighted moments and maximum likelihood. *Revista de Matemática: Teoría y Aplicaciones* 2012; 12(1).
50. Chia CC, Lee JR and Park CY. Radome health management based on synthesized impact detection, laser ultrasonic spectral imaging, and wavelet-transformed ultrasonic propagation imaging methods. *Composites Part B* 2012; 43(8): 2898–2906.
51. Mckeen JCP and Hinder MK. Lamb wave scattering from a through hole. *J Sound Vib* 1999; 224(5): 843–862.
52. Grahm T. Lamb wave scattering from a circular partly through-thickness hole in a plate. *Wave Motion* 2003; 37(1): 63–80.
53. Quaegebeur N, Bouslama N, Bilodeau M et al. Guided wave scattering by geometrical change or damage: Application to characterization of fatigue crack and machined notch. *Ultrasonics* 2017; 73: 187–195.
54. Singh V, Knisely KE, Yönlak SH et al. Non-line-of-sight sound source localization using matched-field processing. *J Acoust Soc Am* 2012; 131(1): 292–302.
55. Williams WB, Michaels TE, Michaels JE et al. Estimation and application of 2-D scattering matrices for sparse array imaging of simulated damage in composite panels. *AIP Conf Proc* 2017; 1806(1): 020014.
56. Supreet Alguri K and Harley JB. Consolidating guided wave simulations and experimental data: a dictionary learning approach. In *SPIE Smart Structures and Materials + Non-destructive Evaluation and Health Monitoring*. International Society for Optics and Photonics, pp. 98050Y–98050Y–10.