Nonlinear System Identification Methods for Characterizing Amplitude Dependent Modal Properties

by

Benjamin John Moldenhauer

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

DOCTOR OF PHILOSOPHY

(Engineering Mechanics)

at the

UNIVERSITY OF WISCONSIN - MADISON

2022

Abstract

The ability to leverage nonlinearity is becoming ever more necessary as modern structures face increasingly demanding design requirements and extreme environments. To incorporate nonlinear behavior into the design process, it must first be characterized experimentally and represented in terms of an identified model form. This is accomplished with nonlinear system identification methods in which signal processing techniques are used to analyze nonlinear system responses and determine corresponding model parameters. Jointed structures exhibit nonlinear stiffness and energy dissipation that is typically represented in terms of quasi-linear modal properties, which are amplitude dependent extensions of the underlying linear natural frequencies and damping ratios. Many current methods for determining quasi-linear parameters operate on the derivatives of the amplitude and phase of an oscillating response signal. Approaches for estimating these quantities are very sensitive to noise and typically require additional steps to achieve reasonable results. This dissertation addresses these issues by detailing three new signal processing techniques for use in determining quasi-linear modal properties.

The first contribution is a method for estimating amplitude and phase called the Short-time Hilbert Transform (STHT). This process is a generalization of the existing Hilbert Transform that combines it with the Short-time Fourier Transform to extract individual oscillations from the signal with time-frequency masking and to suppress end effect errors that arise in the results. While the STHT does still exhibit end effects, they only locally impact the edge and are removed from the rest of the signal. The included

case studies show that the amplitude and phase from the STHT are more accurate than those from the Hilbert Transform.

The second contribution is an additional technique for characterizing oscillations that utilizes nonlinear optimization to fit piecewise polynomial representations of the amplitude and phase to local sections of the signal. Individual components of the signal can be reconstructed by performing the optimization in the frequency domain where specific frequency content can be minimized. While this process is more computationally expensive than the STHT, it completely avoids the end effect issues that are inevitable in any implementation of the Hilbert Transform. The case studies demonstrate how the optimization can be used to extract and integrate individual responses from measured acceleration signals.

The third contribution is a new approach to determining quasi-linear modal parameters called QL-LSQ that operates directly on the measured signals instead of estimated derivatives of the amplitude and phase. This is accomplished by representing the stiffness and damping as B-spline curves that are fit to the response and force with linear least squares regression. The amplitude dependent representation of the quasi-linear modal properties can be directly computed by defining the B-splines with respect to the response amplitude. In the cases studies, the STHT and optimization process are first utilized to extract the necessary response and force signals and form the associated amplitude. QL-LSQ is then used to determine the spline curve representations of the quasi-linear parameters which are shown to be smoother than those from existing methods. The last application demonstrates the effectiveness of QL-LSQ for forced response by using it to characterize the effects of modal coupling in experimental data.

Acknowledgements

I would like to thank my advisor, Dr. Matthew Allen, for being a constant source of guidance and support, and allowing me the opportunity to perform this research and explore the possibilities of nonlinear structural dynamics. His patience and generosity are inspiring, and I look forward to working with him in the future.

Thank you to the Ph. D. committee for taking the time to review this dissertation, attend the defense, and provide valuable feedback. I would like to extend special thanks to Dr. Jacob Notbohm for taking on the role of advisor after Dr. Allen left UW-Madison and allowing me to complete my research and degree.

Thank you to my friends and family for the encouragement and support over the course of the last nine years while I've been at UW-Madison as an undergrad and graduate student. I had the pleasure of meeting many wonderful colleagues, and the comradery and support they provided through the trials and tribulations of higher education is much appreciated.

I am grateful for receiving financial support as a 3M Fellow and from the J. Gordon Baker Fellowship. Special thanks are extended to Dr. Dan Roettgen and Sandia National Labs for supporting the research throughout the doctoral process.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. This work describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Contents

1. Introduction

1.1. Motivation

Nonlinear analysis techniques have become a critical component of the design process as a means of producing efficient structures that are able to survive in challenging environments. While linear techniques are well understood and have been applied to great effect [1], their use inherently limits the scope of the design to operation within the linear regime. When faced with demanding requirements, this approach tends to result in over-built structures suffering from the inelegant solution of simply adding more material to increase strength and resilience. This tactic is effective up until the design objectives are simply impossible with linear techniques, in which case the only feasible path forward is to incorporate nonlinear behavior into the design process.

A modern instance of this is hypersonic aircraft, a prominent example of which is the NASA X-43A concept shown in Figure 1. These aircraft must operate under extreme conditions brought on by traveling at speeds in excess of Mach 5. Of note are the skin panels, which must withstand significant pressure and temperature gradients across the exterior of the aircraft. A linear design is impractical as it would require panels that are very thick and stiff, drastically increasing the weight of the aircraft. Nonlinear analysis techniques are required to accurately characterize and model the thin and light skin panel that must be used, as they exhibit significant nonlinear dynamic behavior [2], including large deformations [3], [4] and buckling [5].

Figure 1: Conceptual rendering of the NASA X-43A hypersonic aircraft.

An area of nonlinear structural dynamics that has recently garnered significant attention is modeling the dynamic behavior of bolted joints [6], [7]. While linear analysis techniques are suitable for low-level responses, these can yield predictions that are off by several orders of magnitude if employed at high-level [8]. This type of behavior is shown in Figure 2, which depicts frequency response functions from a structure containing a jointed interface [9]. High- and low-level impact tests are given by the dots and the solid line, respectively. Compared to the more linear low-level impact, the high-level impact exhibits noticeable nonlinear characteristics in the form of a decrease in frequency and a somewhat counterintuitive decrease in amplitude by nearly a factor of two, indicative of increased damping in the system.

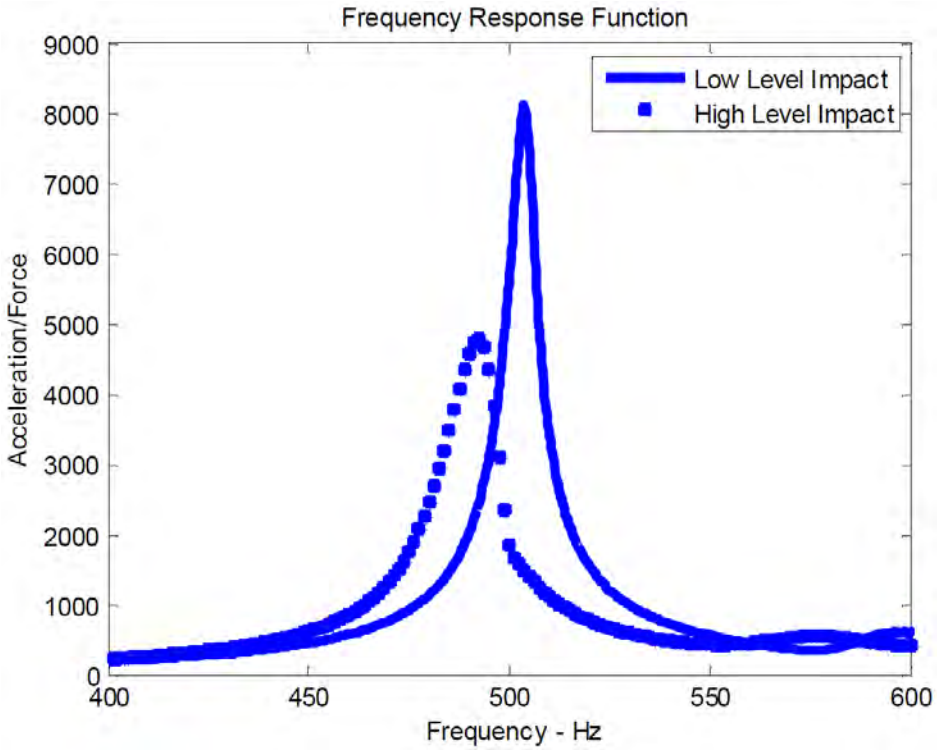

Figure 2: Typical behavior of structures with bolted joints, where high-level excitation can significantly activate the nonlinearity of a jointed interface [9].

An explanation for this behavior requires a close examination of the physics occurring within the jointed interface. As depicted in Figure 3, assume two plates of material are joined by a tightened bolt. In the region marked in red near the through-hole, the normal stresses due to the compressive load between the bolt head and the nut cause the contacting surfaces to statically stick together. If a tangential load is applied, marked in the figure by the horizontal arrows, regions further from the bolt have insufficient normal tractions to prevent minute displacements from occurring between the surfaces. This behavior, called micro-slip, results in small variations in joint stiffness and significant amounts of energy dissipation due to sliding friction [10], [11]. The curves in Figure 2 are indicative of these nonlinear response characteristics.

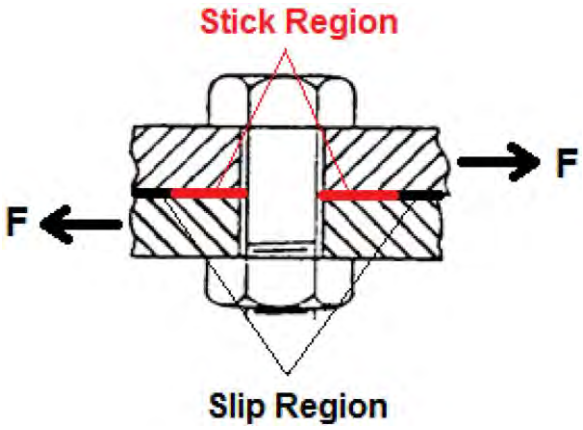

Figure 3: Diagram of a typical bolted joint. The stuck region is highlighted in red, while the microslip area is beyond that in black [10].

During the dynamic response of a structure, motion of material near a bolted joint will induce the tangential loads that cause micro-slip [12]. Assuming the response is harmonic for simplicity, the extent of the micro-slip region oscillates as the material periodically displaces and slips, causing the stiffness and energy dissipation of the joint to vary over a cycle of response. While it can be argued that the physical interactions governing the exact material behavior could theoretically be described by a system of nonlinear differential equations, the complexities in forming such a mathematical model are currently beyond our understanding [6]. A viable alternative technique that has been used to approximately model joint nonlinearity is the method of averaging [13]. In this process, nonlinear dynamic systems are characterized in terms of slowly varying parameters that capture the general trends of the nonlinearity while neglecting rapid, small scale variances. An example of this is shown in Figure 4 which depicts a nonlinear system and its averaged representation, given by the blue and orange lines, respectively. The averaged solution is a slowly decaying curve, neglecting the superimposed small oscillation present in the original system.

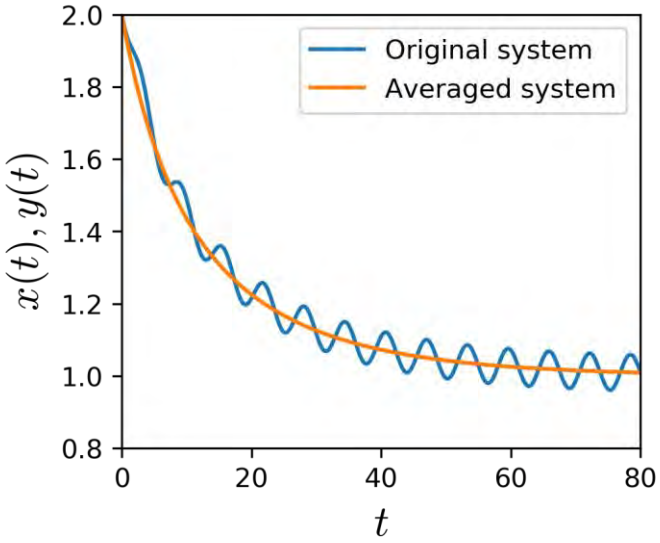

Figure 4: Method of Averaging example. The averaged system, in orange, captures the underlying slowly varying drift in the original system, given in blue, while neglecting the small oscillations [14].

To characterize joint nonlinearity in terms of the method of averaging, slowly varying counterparts to the oscillating material motion and the micro-slip region size must be determined. Suitable options have been found to be the oscillation amplitude of the response and the average of the stiffness and energy dissipation computed over a cycle of response [15]. From these values a quasi-linear modal model for the joint can be constructed in terms of amplitude dependent effective natural frequency and damping ratio [16], [17]. This can be represented as in Eq. (1), in which a single degree-of-freedom (SDOF) modal equation of motion (EOM) has been augmented by an amplitude dependence, A, on the natural frequency and damping ratio.

$$\ddot{x}(t) + 2\zeta(A)\omega_{\mathrm{n}}(A)\dot{x}(t) + \omega_{\mathrm{n}}^2(A)x(t) = f(t) \tag{1}$$

At adequately low response amplitude the joint nonlinearity is weakly excited and the amplitude dependent parameters converge with the underlying linear natural frequency and damping ratio of the structure, reducing Eq. (1) to a standard linear

EOM. As response amplitude increases, the nonlinear effects of the dynamic micro-slip region become more significant, driving the effective natural frequency and damping ratio away from the linear modal values.

In practice, these amplitude dependent nonlinear modal parameters can be represented as in Figure 5, which shows several experimental results from a jointed benchmark structure [18], [19]. On the left side, the natural frequency consistently decreases at greater amplitudes, signifying a stiffness softening nonlinearity. At lower amplitudes, the frequency levels off and approaches a constant linear value. Similar behavior is exhibited by the damping, in that it deviates upward from a linear value as amplitude increases, reaching a maximum value nearly three times greater. To generate these amplitude dependent curves from measured response data, nonlinear system identification techniques are utilized. The following section presents a primer on various topics in this field.

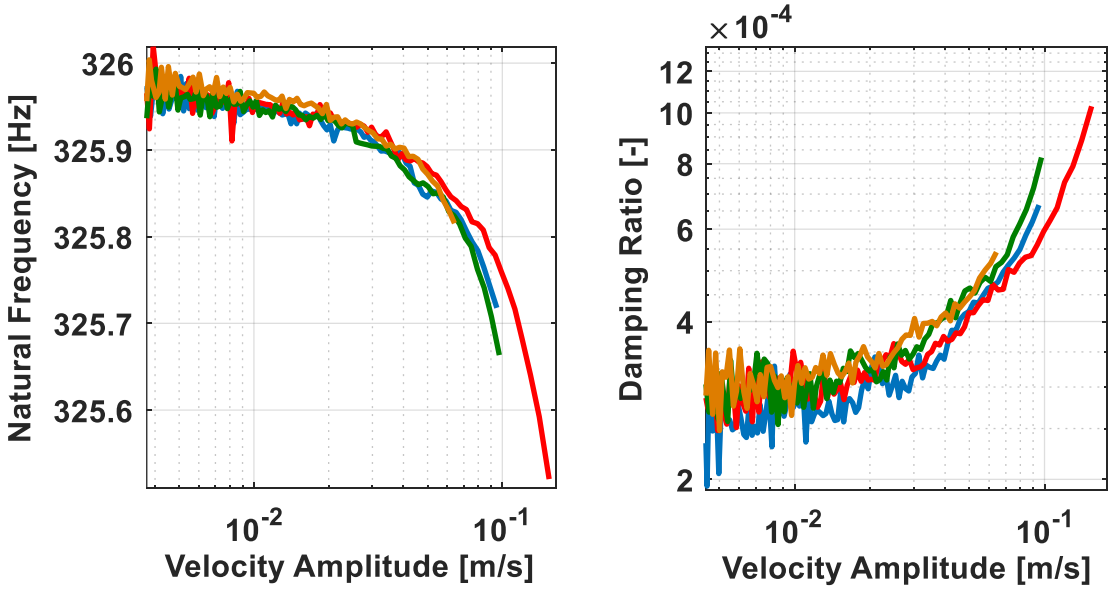

Figure 5: Typical amplitude dependent natural frequency and damping ratio curves. The various lines display results from impacts at a range of force levels [19].

1.2. Nonlinear System Identification Techniques

In structural dynamics, nonlinear system identification (NLSID) techniques encompass a vast array of methods developed over the last 50 years for the purpose of characterizing some aspect of complicated structural response. A thorough review tracking the evolution of NLSID approaches is presented in [20], [21], and guides to selecting and implementing various procedures are given in [22], [23]. In general, the goal of a NLSID method is to form a mathematical model that reproduces the behavior of some nonlinear dynamical system. As no single modelling technique can conceivably capture every possible nonlinear behavior, prior knowledge must be used to assume a specific model form that will best represent the physical characteristics of the system. This is as opposed to black-box modelling techniques which require no assumptions, but in return offer no insight into the nonlinear physics occurring.

In assuming a model form, the mathematical basis can be either parametric or nonparametric. Parametric models are characterized by a set number of variables that can be tuned to alter the behavior exhibited by the model to best match the observed nonlinear system. For bolted joint nonlinearity, a review of the most commonly used parametric models are given in [24], with notable examples being the 4-parameter Iwan model [25], the Bouc-Wen model [26], and the Valanis model [7]. Conversely, nonparametric models are typically defined in terms of an arbitrary number of polynomial expressions. These are fit to the observed behavior by assigning coefficients that minimize the residual between the model and data, usually through some form of least squares regression calculation.

A majority of nonparametric NLSID approaches assume that the mathematical model can be constructed by introducing nonlinear polynomial response terms into an initially linear EOM. Likely the most well-known example of this type of nonlinear EOM is the Duffing Oscillator [27], which includes a cubic displacement term to augment the apparent stiffness of the system. Established methods of this sort that process time domain data include the Restoring Force Surface (RFS) method from Masri and Caughey [28] and Nonlinear Auto-Regressive Moving Average with eXogeneous input (NARMAX) from Leontaritis and Billings [29], [30]. Formulations based on frequency domain data are Nonlinear Identification through Feedback of the Output (NIFO) from Adams and Allemang [31], the Conditioned Reverse Path (CRP) method from Richards and Singh [32], and a recent development that combines those two, Nonlinear Identification through eXtended Outputs (NIXO) from Kwarta [33].

The common thread running through each of these techniques is the assumption that the observed nonlinear behavior can be accurately represented by a set of nonlinear polynomials, which for convenience are usually taken to be quadratic and cubic terms of velocity and displacement. While straightforward, this approach is rather arbitrary and is only physically justifiable in certain experimental systems, such as with the geometric nonlinearity present in the ECL benchmark structure [34]. For general nonlinear systems, the inherent issue with these methods is that a model can be formed from any combination of polynomial terms, and there is little indication as to under what conditions, if any, the model will accurately reproduce the nonlinear behavior of the system. This difficultly is explored in detail with RFS in the next section.

An alternative class of nonparametric NLSID techniques instead assume a model form akin to the quasi-linear EOM derived in the previous section and shown in Eq. (1). These methods are essentially signal processing techniques that extract amplitude dependent modal parameters from measurements of a nonlinear system responding to some excitation. The overarching approach with these is to assume that the response can be represented by a nonstationary oscillation in terms of its amplitude and phase. By computing the rate of change of both, the instantaneous natural frequency and damping ratio can be determined at each point in time and amplitude.

The signal processing techniques for performing this analysis are explored in detail in Section 1.2.2; a brief summary is given here. Typical approaches generally fall into two main classifications. The first, time-frequency methods, represent the measured response in terms of the frequency content present at each time sample, with the desired instantaneous values determined by tracking how that distribution evolves through time. The second set of methods are those that directly process recorded time domain data. In the most basic sense, this entails inferring frequency and damping from zero crossings and peak values in the signal. A more advanced approach is to use the Hilbert Transform to estimate the amplitude and phase of the signal. The numerical derivatives of each are then used to compute the instantaneous frequency and damping. While relatively straightforward and computationally efficient, these techniques tend to have difficulties with signals that contain significant noise and/or multiple oscillations, requiring intensive pre- and post-processing to overcome.

In the body of this work, a set of new NLSID techniques are proposed for the propose of identifying amplitude dependent modal parameters from measured nonlinear system response. The following sections detail existing techniques, RFS and FREEVIB/ FORCEVIB, and explore difficulties in each that warrant the creation of the proposed methods.

1.2.1. The Restoring Force Surface Method

RFS is a very attractive technique due to its intuitive derivation and ease of implementation. The basis for the method is to assume that the nonlinear system to be modeled can be represented by the single degree-of-freedom equation of motion given in Eq. (2), where $\ddot{x}(t)$ is acceleration, $f(t)$ is external forcing, and $F_{\mathrm{NL}}\big(\dot{x}(t), x(t)\big)$, deemed the restoring force, characterizes all linear and nonlinear stiffness and damping properties of the system.

$$\ddot{x}(t) + F_{\mathrm{NL}}\big(\dot{x}(t), x(t)\big) = f(t) \tag{2}$$

Typically, the form of the restoring force is assumed to be a combination of response terms, including the standard linear stiffness and viscous damping as well as nonlinear quadratic and cubic polynomials of displacement and velocity. This series of terms representing F_{NL} is then shown as:

$$F_{\mathrm{NL}}\big(\dot{x}(t), x(t)\big) = c_0\dot{x}(t) + c_1|\dot{x}(t)|\dot{x}(t) + c_2\dot{x}^3(t) + k_0x(t) + k_1|x(t)|x(t) + k_2x^3(t) \tag{3}$$

If the external force, $f(t)$, acceleration, $\ddot{x}(t)$, velocity, $\dot{x}(t)$, displacement, $x(t)$, and linear coefficients, c_0 and k_0, are known, these terms may be collected on the right

side of the equation, leaving the nonlinear polynomials on the left. Putting this expression into a matrix form yields:

$$[|\dot{x}(t)|\dot{x}(t) \quad \dot{x}^3(t) \quad |x(t)|x(t) \quad x^3(t)] \begin{bmatrix} c_1 \\ c_2 \\ k_1 \\ k_2 \end{bmatrix} = f(t) - \ddot{x}(t) - c_0\dot{x}(t) - k_0 x(t) \tag{4}$$

In practice, each time-dependent variable is a tall column vector of values, sampled at discrete points in time. Collapsing the tall matrix on the left into $\mathbf{P}$ and the vector on the right into $\mathbf{U}$ simplifies the expression to be displayed as:

$$\mathbf{P} \begin{bmatrix} c_1 \\ c_2 \\ k_1 \\ k_2 \end{bmatrix} = \mathbf{U} \tag{5}$$

The polynomial coefficients may now be determined via linear least squares regression by left-multiplying both sides by the pseudo-inverse of $\mathbf{P}$, represented below with the dagger operator, †, and defined as $\mathbf{P}^\dagger = (\mathbf{P^T P})^{-1}\mathbf{P^T}$.

$$\begin{bmatrix} c_1 \\ c_2 \\ k_1 \\ k_2 \end{bmatrix} = \mathbf{P}^\dagger \mathbf{U} \tag{6}$$

With the coefficient values, the nonlinear EOM may be utilized in a numerical integration scheme to simulate the response of the modeled system to an arbitrary external forcing. Several examples of applying this technique to model real structures are described in [9], [35]–[37].

While RFS generates a nonlinear EOM that can be readily integrated, the inherent difficultly is that the form of that equation must be assumed a priori. In the derivation above, a series of nonlinear polynomial terms were used, and a model was

formed by determining the coefficients with a least squares solution, but there is no guarantee that the model accurately emulates the nonlinear dynamics of the system.

This issue was explored by the author in [38], in which RFS was implemented in an attempt to model the nonlinear response of a structure with a jointed interface. The structure is described in detail in [9]. Since the polynomial model form for general experimental data cannot be definitively determined, a Monte Carlo simulation was devised in which, in each iteration, a model form would be randomly generated and fit to the experimental restoring force. The models could contain anywhere from one to forty nonlinear terms and, to investigate the use of non-integer polynomial orders, exponents between 1.1 and 5 at 0.1 increments.

In the most basic sense, it should be expected that the generated models would be capable of reproducing the nonlinear behavior that was used to form them, when subjected to the same excitation. Thus, to evaluate the accuracy of the RFS models, their dynamic response to the measured forcing was numerically simulated. As the nonlinear properties of interest are the amplitude dependent natural frequency and damping ratio, these quantities were computed from both the experimental and simulated responses. Error metrics were calculated for the least squares fit of the model to the restoring force, and between the instantaneous frequency and damping values from the simulation and experiment.

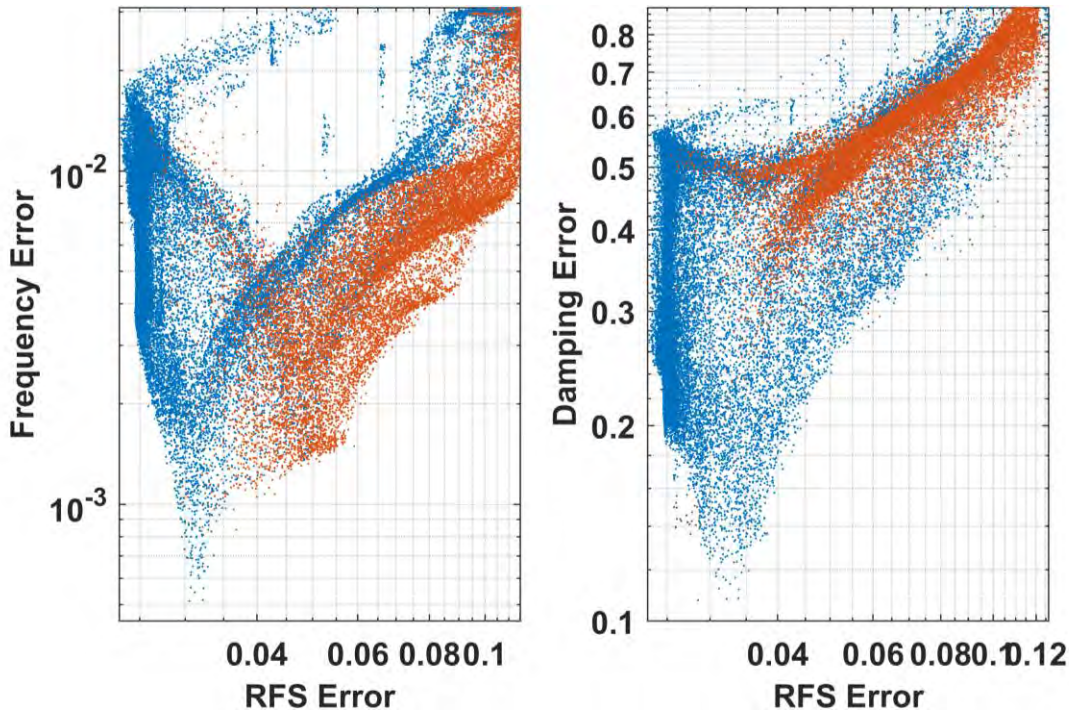

Figure 6: Frequency and Damping error vs restoring force error from 100,000 Monte Carlo simulations. The orange dots show the distribution of results from models containing only integer exponents, while the blue dots are from models that allow fractional exponents.

The scatter plots in Figure 6 show the results of 100,000 different model forms, each from a separate Monte Carlo iteration. The left plot shows the distribution of frequency error verses restoring force error, while the right plot displays the associated damping error. In each, the blue dots are results from models that allowed fractional polynomial exponents, while the orange dots are for models with exclusively integer exponents. It is noteworthy that, in general, the fractional exponents produced more models that accurately fit the restoring force, as shown by a denser cluster of blue points on the left side of each plot compared to the orange points. However, this distribution also indicates a discrepancy between the ability of the models to fit the restoring force, and their accuracy in simulating the correct nonlinear characteristics of the experimental system.

If a direct relationship between the error metrics existed, a clear trend would be visible from the lower left to the upper right in the point distributions. Instead, no points landed in the lower left corner and the majority of the blue dots are clustered on the far left, displaying great fits to the nonlinear restoring force but middling frequency and damping accuracy. While some models were able to present the correct frequency behavior, given by the slightly prominent minimum in frequency error, these show a greater restoring force error than most of the models tested. In the damping plot, there is no minimum as none of the models closely followed the measured nonlinear curve. Thus, even after generating and trialing tens of thousands of random model forms with fractional exponents, none were found to closely reproduce both the amplitude dependent frequency and damping behavior observed in the measured data. This seems to indicate that a model's ability to fit the restoring force does little to ensure that it properly exhibits the nonlinear characteristics of the observed system response.

While NLSID techniques based on response polynomials are simple and intuitive, the requirement that the nonlinear model form must be assumed a priori introduces a significant level of uncertainty into the generated nonlinear model. The RFS case study above highlights one such method and demonstrates how challenging it can be to produce an arbitrary polynomial model that accurately reproduces nonlinear characteristic of an experimental structure. It was, in part, these difficulties that compelled the author to develop the proposed NLSID methods that are the topic of this dissertation.

1.2.2. Instantaneous Frequency and Damping

Oscillating signals are encountered in a many different fields, notably AM/ FM radio communications, signal processing, biology [39], [40], particle physics [41], and of primary relevance here, structural dynamics. The simplest representation of an oscillation is in terms of its amplitude and phase. Assuming the form shown in Eq. (7), $A(t)$ dictates the amplitude envelope and $\psi(t)$ defines the phase angle at time t.

$$y(t) = A(t)\cos\big(\psi(t)\big) \tag{7}$$

Typically, when analyzing some measured oscillation, the current value of the phase is unimportant. A much more useful metric for characterizing the signal is the instantaneous frequency, ω, which can be defined as the derivative of the phase with respect to time.

$$\omega(t) = \frac{d}{dt}\psi(t) \tag{8}$$

Many techniques for estimating instantaneous frequency have been developed. A thorough review of these methods compiled by Boashash is presented in [42], [43]. For analyzing nonlinear structural dynamics measurements, where the goal is to determine the quasi-linear amplitude dependent natural frequency and damping ratio, the most widely used approaches are time and time-frequency methods.

In general, time-frequency methods produce an array of values characterizing the frequency content in the signal of interest at each instant in time. Common numerical processes for computing these values are the Short-Time Fourier Transform [44] and the Wavelet Transform [45]. Significant oscillations in the signal

produce prominent ridges that progress through the time-frequency content. To extract precise estimates of the instantaneous frequency and amplitude, a technique known as synchrosqueezing can be employed [46]–[48], as well as a variety of other methods described in [49], [50]. Figure 7 depicts a typical result from time-frequency analysis [51]. The plot on the left presents a time signal that is sweeping in frequency while also changing in amplitude. The plot on the right, referred to as a spectrogram, is a visual representation of the frequency content verses time, as computed with a Short-Time Fourier Transform. The extracted instantaneous frequency is highlighted by the red line following the ridge in the spectrogram. Several examples that implement these processes for structural dynamics NLSID are in [52]–[54].

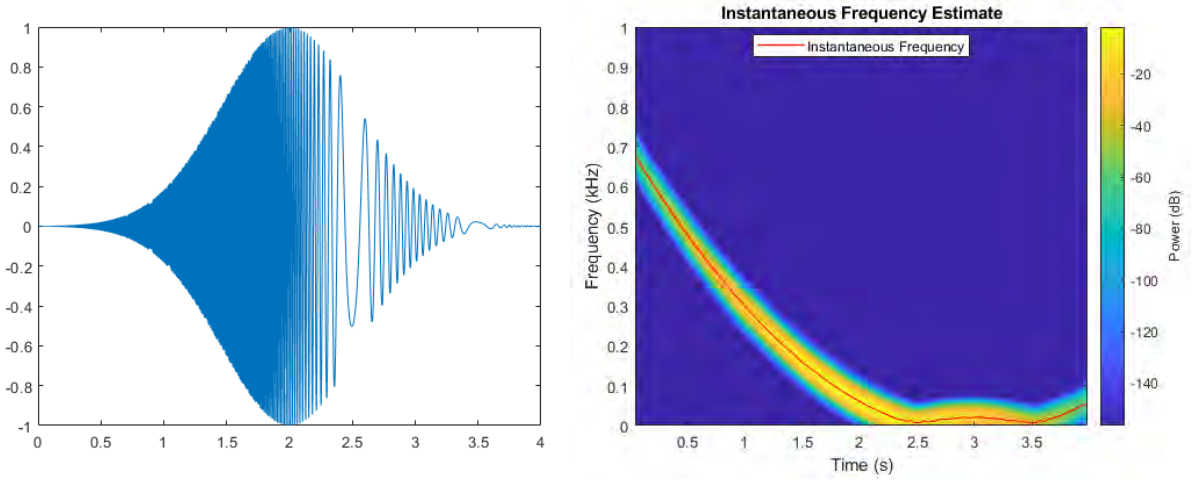

Figure 7: Left - A time domain signal with shifting frequency and amplitude. Right - Spectrogram of the signal showing how the amplitude and frequency evolve through time. [51]

The other main class of methods used in structural dynamics are those that produce instantaneous values directly from the supplied time signals. Most of these methods are formulated to extract natural frequencies and damping ratios from free decay response. The simplest approach is to infer frequency from approximate zero

crossings and damping from log decrement, as done in [55]. While incredibly straightforward, this approach requires a high sample rate and low noise to ensure the selected points are near the true zeros and peaks. A slightly more sophisticated methodology that bypasses some of those difficulties is to employ a smoothing and interpolation scheme to estimate precise peaks and zero crossings, a process utilized by Jin in [56] and Londoño in [57]. In [58], Goyder proposed an algorithm that instead considers every point in the oscillation and curve fits natural frequency and damping terms to the signal with nonlinear optimization. A review of several other methods for estimating nonlinear modal parameters from free decay signals is presented in [59].

A prevalent time domain method for determining instantaneous frequency and damping is through utilization of the Hilbert Transform. The Hilbert Transform is essentially a filter that shifts the phasing of a signal by one quarter of a period, putting it 90° out of phase with the original data. This process is defined mathematically in terms of a convolution as in Eq. (9), where $\boldsymbol{\mathcal{H}}$ is the Hilbert Transform, $y(t)$ is an arbitrary signal, and $\hat{y}(t)$ is its phase shifted counterpart.

$$\boldsymbol{\mathcal{H}}\big(y(t)\big) = \hat{y}(t) = \frac{1}{\pi t} * y(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y(\tau)}{t-\tau}\, d\tau \tag{9}$$

The result, $\hat{y}(t)$, can be interpreted as approximating an imaginary complement to $y(t)$. By combining both into a complex quantity, the original real valued signal can be represented as a complex exponential as in Eq. (10). The amplitude, $A(t)$, and phase, $\psi(t)$, are the magnitude and angle of $y(t) + i\hat{y}(t)$, as shown in Eq. (11) and Eq. (12), respectively [60], [61].

$$y(t) + i\hat{y}(t) = A(t)e^{i\psi(t)} \tag{10}$$

$$A(t) = \sqrt{y^2(t) + \hat{y}^2(t)} \tag{11}$$

$$\psi(t) = \tan^{-1}\left(\frac{\hat{y}(t)}{y(t)}\right) \tag{12}$$

From this result, and assuming the quasi-linear model form in Eq. (1), Feldman derived the FREEVIB [62] method, which generates instantaneous representations of natural frequency and damping ratio as in Eq. (13) and Eq. (14), respectively, where the time dependence in each term as been omitted for clarity. These expressions are simple functions of the amplitude and instantaneous frequency, and the time derivatives of both. For forced response, forcing terms can be added to these expressions as derived by Feldman in the FORCEVIB [63] method.

$$\omega_n = \sqrt{\omega^2 - \frac{\ddot{A}}{A} + \frac{2\dot{A}^2}{A^2} + \frac{\dot{A}\dot{\omega}}{A\omega}} \tag{13}$$

$$\zeta = \frac{1}{\omega_n}\left(-\frac{\dot{A}}{A} - \frac{\dot{\omega}}{2\omega}\right) \tag{14}$$

The most significant obstacle when implementing this process is noise in the Hilbert Transform result, $\hat{y}(t)$, as it will propagate into the estimates for amplitude and phase and be consequently amplified in their respective derivatives. Even if the original signal, $y(t)$, is noise free, the Hilbert Transform itself can introduce noise into $\hat{y}(t)$.

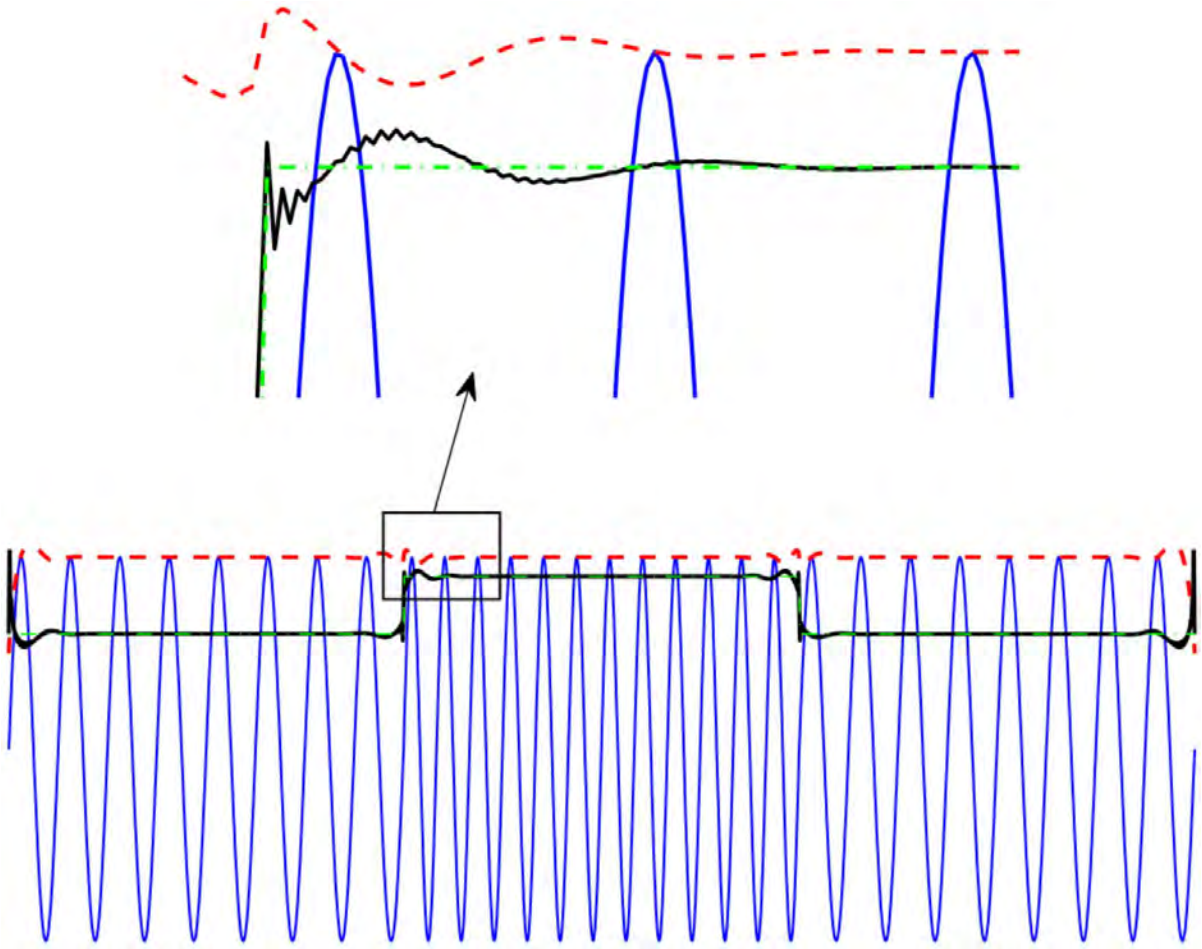

Figure 8: Hilbert Transform error around an abrupt step in frequency [64]. The blue line is the signal, the green line is the true frequency, the dashed line is the estimated amplitude envelope, and the solid black line is the estimated instantaneous frequency.

This behavior is shown in Figure 8, in which the results from the Hilbert Transform of a signal with a step change in frequency are displayed. In the plot, the solid blue line is the signal, the dash-dot green line is the true frequency, the dashed red line is the estimated amplitude envelope, and the solid black line is the estimated instantaneous frequency. On either side of the step change, undulations are present in both the estimated amplitude and frequency, as highlighted in the zoomed portion of the figure. Such errors occur about any abrupt changes in the signal, akin to the

Gibbs Phenomenon in the Fourier Transform. This also leads to the significant end-effects present at the beginning and end of the signal. To produce usable results when this behavior is encountered in practice, the noise must be smoothed from the data. This is typically a user intensive process of tuning bandpass filters or curve fitting algorithms [60], [65].

An additional limitation of the Hilbert Transform and FREEVIB/ FORCEVIB is that the processed signal must be monoharmonic for the results to represent the physical properties of a structural resonance. If the signal contains multiple simultaneous oscillations, the Hilbert Transform will return a superposition of each amplitude and phase. To correct this, Feldman proposed the Hilbert Vibration Decomposition (HVD) [66] which uses filters to iteratively separate the most dominant component from the superimposed frequency and envelope. An example using HVD to identify nonlinear aspects of a beam with a moving mass is given in [67], [68]. An additional method for separating the mono-components of a combined signal is the Hilbert-Huang Transform [69], in which a technique known as empirical mode decomposition (EMD) is used to separate the signal into a set of distinct oscillations, called intrinsic mode functions (IMF), that may then be independently analyzed with the Hilbert Transform. Detailed descriptions of EMD can be found in [70], [71], and a comparison of EMD and HVD is given in [72]. This procedure has been used extensively for analyzing structural response, such as in [73]–[76]. However, a common difficulty with EMD is that extracting the IMFs is not entirely trivial or automatic. A significant amount of user interaction may be required to extract well

behaved mono-harmonics that contain frequency content of a single oscillation. A potential solution to this is given in [77], in which the wavelet transform is used in an optimization scheme to ensure that each IMF is monoharmonic. Additional applications that combine EMD with time-frequency techniques are found in [78]–[80].

The Hilbert Transform is a powerful tool that is widely used in many fields due to its conceptual simplicity and computational efficiency. For structural measurements, free or forced, it can be used to calculate instantaneous natural frequency and damping ratio. However, this process is functionally limited to smooth signals that are monoharmonic, as significant noise is introduced if these conditions are not met. While many methods to overcome these limitations have been proposed, such as HVD, EMD and bandpass filtering, these may entail a great deal of manual pre- and post-processing to produce satisfactory results. These difficulties compelled the author to explore the development of alternative methods, as detailed in the body of this work.

1.3. Scope of the Dissertation

This dissertation presents three new NLSID techniques for use in determining amplitude dependent modal parameters. The first is a generalization of the Hilbert Transform that uses the Short-time Fourier Transform to produce more accurate amplitude and phase results. The second technique is also for characterizing nonstationary oscillations but is a more sophisticated process that employs nonlinear optimization to curve fit polynomial representations of amplitude and phase to a signal. The output from these methods can then be used in the third contribution, which is a process for determining the quasi-linear natural frequency and damping ratio in terms of B-spline curves directly from the measured response and force.

The first section of Chapter 2 presents a detailed description of the Hilbert Transform process in order to provide a robust theoretical basis for why errors are so often encountered in its results. A typical method for performing the Hilbert Transform is by applying a transfer function to the frequency spectrum of the signal directly computed via the Fourier Transform. If the Short-time Fourier Transform is used instead, where the signal is first divided into overlapping windowed segments before determining the spectrum of each, the Hilbert Transform can be applied to isolated sections of the response that can then be reassembled into a unified result for the entire signal. This modified Hilbert Transform process proposed in this work is the so-called Short-time Hilbert Transform (STHT). Determining the amplitude and phase of the signal with this method suppresses the ringing artifacts to only effect the signal in

the local area about any sudden changes. Additionally, mode decomposition can also be performed through time-frequency masking that is applied to the segment spectrums simultaneously with the Hilbert Transform transfer function.

This approach is first demonstrated on a simple linear free decay signal, which shows that the STHT is more accurate than the standard Hilbert Transform and produces negligible error in the central portion of the response. The second numerical example is a complicated multicomponent signal with three harmonics sweeping in frequency. While existing decomposition techniques EMD and VMD fail to correctly characterize this signal, the STHT is shown to accurately extract each component and be resilient to noise. The method is then applied to experimental data from a nonlinear benchmark structure, where quasi-linear natural frequency and damping curves for individual modal responses are extracted from a complicated free decay response.

Chapter 3 then details an alternative approach for estimating the amplitude and phase of a nonstationary oscillation that avoids the Hilbert Transform completely. This is accomplished by assuming the amplitude and phase are represented as polynomials in time such that nonlinear optimization can be used to curve fit these with local regression to segments of the measured data. To ensure a computationally efficient solution, the analytical form of the objective function gradient is defined for use in a nonlinear least squares algorithm. An interesting feature of this process is that the error minimization in the optimization can be computed in terms of either time or frequency domain data. This property can be leveraged in scenarios where the signal contains multiple superimposed oscillations. By restricting the optimization to

only consider a certain subset of the frequency content, individual oscillations can be fit and extracted from the signal in a process similar to time-frequency masking. From the resultant polynomial expressions for amplitude and phase, several characteristics of the nonlinear system that produced the signal can be determined. First, whether the data is acceleration, velocity, or displacement, the fit oscillation can be analytically integrated and/or differentiated to yield exact representations of the other two response quantities. And second, an implementation of FREEVIB/ FORCEVIB that utilizes the amplitude and phase polynomials can be formed.

This process is demonstrated on a series of case studies, starting with the simulated response of a 4-parameter Iwan Model. The main benefit of this model is that there are closed form expressions for its amplitude dependent nonlinear behavior, such that exact quasi-linear parameters can be compared to. The optimization is shown to accurately fit the response and force and reasonably emulate the nonlinear behavior. It is then used on an experimental free decay measurement from impact testing of the S4 Beam, in which individual modal responses are extracted from the response and characterized. The final case is experimental forced response of the Cylinder-Plate-Beam (CPB) structure in which the optimization process accurately filters out higher harmonics and forms representations of the quasi-linear parameters.

In Chapter 4, a new approach is presented for determining amplitude dependent natural frequencies and damping ratios for structural resonances undergoing both free and forced response. The method, denoted as QL-LSQ, is derived from the quasi-linear equation of motion in which the stiffness and damping

coefficients on the displacement and velocity are assumed to be variable instead of constants. By representing these coefficients as B-spline curves, they can be determined via linear least squares regression between the response and force terms in the equation of motion, similar to the model fitting procedure used in RFS. This process forms amplitude dependent modal parameters comparable to FREEVIB/ FORCEVIB, but directly computes them from the measured response and force values, avoiding the need to estimate the amplitude and phase derivatives, and does not require that the force be harmonic, or the response be steady state.

Section 4.2 first presents a review of relevant background information on utilizing penalized B-splines as that formulation is then used to derive the QL-LSQ process. B-splines offer an incredible amount of flexibility and adaptability that is demonstrated through revisiting the case studies from the previous chapters. The first is the simulated response of the Iwan Model to a sine beat excitation, such that the response exhibits periods of both forced and free response. QL-LSQ very accurately produces the amplitude dependent behavior of the Iwan Model throughout the response. Then, the extracted resonant free decay signals from the S4 beam response as determined by the optimization process are used by QL-LSQ to generate the associated quasi-linear parameters. Compared to a collection of existing methods that utilize FREEVIB, QL-LSQ generates a curve that is much smoother while still exhibiting the same nonlinear behavior. The experimental forced response of the CPB is then reexamined, where the STHT is used to filter out higher harmonics and generate and estimate of the response amplitude. From those results, QL-LSQ

produces quasi-linear parameters exhibit less noise than those from FORCEVIB. As a final application of QL-LSQ, data from the CPB which exhibits nonlinear modal coupling effects is characterized and viewed as a function of two modal amplitudes. Finally, Chapter 5 provides an overview of the conclusions of the dissertation and avenues of future work.

2. The Short-Time Hilbert Transform

2.1. Introduction

Creating efficient structures that can survive in demanding environments is possible through leveraging nonlinearity. To incorporate nonlinear behavior into the design process, it must first be characterized experimentally and represented in terms of an identified model form. This is accomplished with nonlinear system identification methods, which are signal processing techniques that analyze measured nonlinear response signals and determine corresponding model parameters. An area of nonlinear structural dynamics that has been a topic of great interest and development is the modelling of structures that contain bolted joints [6]. This type of nonlinear behavior is typically characterized in an averaged sense, where it is assumed that variations in the stiffness and energy dissipation of the structure can be represented by a quasi-linear natural frequency and damping ratio that vary with respect to the amplitude of the response. A common approach for determining these parameters is the FREEVIB method [62], which operates on the derivatives of the instantaneous amplitude and phase of the structural response. As amplitude and phase cannot be directly measured, these quantities must be estimated from the observed response. The most widely used process for accomplishing this is with the Hilbert Transform [60].

The Hilbert Transform is a linear transformation that generates a phase-quadrature complement to an input function by applying a 90° phase delay [81]. These

signals are then taken as the real and imaginary parts of a complex-valued function known as an analytic signal. While this process was originally developed for examining the properties of complex-valued functions, it was later adapted to the analysis of harmonic signals by Gabor in [82], in which stationary oscillations represented as analytic signals were used to quantify the inherent uncertainty in determining the frequency of a signal relative to the length of time it is observed. This was then expanded upon to incorporate nonstationary oscillations for use in the telecommunications field, in which information is transmitted in the form of a modulation signal that is combined with a carrier wave that operates at a much higher frequency [83]. The encoded information can then be reconstructed in a receiver by extracting the applied modulation from the carrier wave. Using the Hilbert Transform as a theoretical basis, Bedrosian and Nuttall defined limitations on the modulation to ensure that it can be correctly recovered. These conditions, conventionally known as the Bedrosian Theorem [84] and the Nuttall Theorem [85], dictate that the modulation is separable from the carrier wave in the Hilbert Transform if and only if the frequency spectrums of the modulation and carrier are bandlimited and do not overlap. If these conditions are violated, the output from the Hilbert Transform does not match the expected phase-quadrature signal that was used to define the modulation. Detailed discussions on the consequences of violating the Bedrosian and Nuttall Theorems and investigations into the subsequent errors are given in [86], [87].

The Hilbert Transform has more recently found broad use as a general signal processing tool [88]. After computing the Hilbert Transform of an oscillating signal, an

estimate of its amplitude and phase are simply the magnitude and angle of the associated analytic signal. For many applications, this amplitude and phase is only physically meaningful if the processed signal contains a single harmonic function. Thus, the Hilbert Transform is often used in conjunction with sophisticated mode decomposition techniques which are able to extract individual harmonics from a multicomponent signal. A vast number of decomposition methods have been proposed, with a few of the most well-known being Empirical Mode Decomposition (EMD) [69], Analytical Mode Decomposition (AMD) [89], Variational Mode Decomposition (VMD) [90], and Hilbert Vibration Decomposition (HVD) [66]. While each of these produce isolated monoharmonic signals, the amplitude and phase results from the Hilbert Transform typically exhibit significant errors. These typically come in the form of end effects that primarily distort the beginning and end of the signal, where the usual solution is to simply truncate the signal and remove the corrupted sections [60].

The cause of these errors can be understood in terms of the Bedrosian and Nuttall Theorems by taking the amplitude to be the modulation signal and the phase to represent the carrier wave. For an arbitrary signal, such as a measured structural response, the start and end states of the signal are likely not exactly periodic, such that its frequency spectrum contains significant leakage. Consequently, the spectrums of the amplitude and phase are not bandlimited and separate, and the Hilbert Transform generates an analytic signal in which the magnitude and angle do not cleanly produce the desired amplitude and phase [91]. The Hilbert Transform is

typically computed through the Fourier Transform, where a phase shifting transfer function is applied to the frequency spectrum of the complete input signal. If there are any sudden changes in the signal, such as nonperiodic edges or an impact, its spectrum contains significant broadband leakage. Portions of this leakage are incorrectly modified by the Hilbert Transform transfer function, such that the resultant analytic signal in the time domain contains errors known as ringing artifacts or Gibbs Phenomenon [92]. In general, these ringing artifacts propagate with a decaying amplitude throughout the entire signal, where the maximum value is proportional to and is at the location of the sudden change. Therefore, the errors in the estimated amplitude and phase can be reduced by manipulating the signal to suppress the spectral leakage.

To that end, this work proposes a modified Hilbert Transform process called the Short-time Hilbert Transform (STHT) which utilizes the Short-time Fourier Transform (STFT) instead of the conventional Fourier Transform. The framework of the STFT allows for a signal to be divided into windowed segments, the spectral content of each to be individually manipulated, and then the result to be transferred back to the time domain and assembled into a unified signal [44]. Applying a windowing function to each segment forces them into a pseudo-periodic state which reduces the amount of broadband leakage in the frequency spectrum that is due to nonperiodic edge conditions. This then minimizes time domain end effects due to any modifications that were made to the spectrum, such that they can be removed from the assembled signal by defining a degree of overlap between adjacent segments.

The Hilbert Transform is incorporated into this process by simply applying its transfer function to the spectrum of each segment. These individual analytic signals can then be assembled to form a phase-quadrature complement to the total input signal. The benefit of using the STFT to form this result is that errors are localized. If the original input signal contains any sudden changes, the resultant ringing artifacts are confined to only those segments that also contain that sudden change instead of propagating throughout the entire signal.

An additional aspect of the STFT that can be leveraged in the STHT is the ability to apply other modifications to the spectral content beyond just the Hilbert Transform. This allows for mode decomposition to be performed through time-frequency masking, which is a filtering technique commonly used in signal processing for audio applications [93]. With this method, a specific harmonic in a multicomponent signal can be extracted by removing the frequency content of other components from the spectrum of each segment. Used in conjunction with the Hilbert Transform, the resultant assembled signal contains the desired harmonic component and its phase-quadrature compliment.

Several other alternatives to the conventional Hilbert Transform approach have also been suggested. Dishan formulated an algorithm for utilizing the wavelet transform to compute the Hilbert Transform of a signal in [94]. This approach was utilized in [95] and found to yield good results, provided the wavelet parameters were properly selected. Work by H. Olkkonen, Pesola, and J. T. Olkkonen in [96] demonstrated how the Discrete Cosine Transform could be implemented to compute

the Hilbert Transform of a signal. While this method is more efficient than the standard Fourier Transform approach and allowed for noise removal via truncation of the transform coefficients, significant ringing artifacts are still present. There is also historical precedence for directly applying a windowing function to a finite impulse response digital filter Hilbert Transformer to reduce errors in the results [88]. Several recent approaches from the field of geophysics have incorporated time-frequency analysis into the Hilbert Transform, such as the Generalized Hilbert Transform from Luo et al. [97] in which the STFT coefficients of a signal are first weighted by their norm, such that low amplitude content is suppressed in an attempt to remove noise, where the Hilbert Transform result is then computed by directly summing the weighted STFT coefficients. This process was expanded upon by Zhang et al. in [98] with the inclusion of additional weighting terms. In the work of Lu and Zhang in [99], they formulate a so-called Windowed Hilbert Transform that initially resembles the STHT proposed here, but their process then employs a magnitude-based filter that scales the amplitude of the STFT coefficients similar to that used in the Generalized Hilbert Transform. Consequently, their approach can only recover the phase in each segment as the magnitude information has been distorted beyond use. An extension of the Windowed Hilbert Transform by Sattari in [100] incorporates an adaptive fast sparse S-transform to produce higher resolution results. This was then extended again by Kakhki, Mansur, and Aghazade in [101] to incorporate the magnitude weighting for increased resilience to noise. An issue present in these alternatives is that each is a variation of the Hilbert Transform and are thus subject to the Bedrosian and Nuttall

Theorems. To avoid this completely, Huang et al. [102] propose that the amplitude envelope can be directly determined by fitting a spline curve to the oscillation peaks in a monoharmonic signal. After normalizing the signal by this curve, the phase can be recovered via a direct quadrature computation. While this does circumvent certain limitations of the Hilbert Transform, the process lacks solid theoretical foundation and is prone to significant errors if the signal contains noise.

The remainder of the chapter is structured as follows. Contained in Section 2.2 is a detailed review of information pertinent to understanding the Hilbert Transform in the context of structural dynamics, as this seems to be lacking in the currently available literature. This begins with a discussion of the relation between the phase-quadrature representation of an oscillating signal and its amplitude and phase, and then how these are used in characterizing nonlinear structural behavior. After deriving the theoretical background for the Hilbert Transform and analytic signals in the time and frequency domains, it is shown that placing the analytic signal into polar form produces an estimate of the amplitude and phase of a signal. To determine why these estimations often contain significant end effects, the error between the Hilbert Transform output and the desired phase-quadrature signal is examined. This reveals that spectral leakage causes the signals to violate the Bedrosian Identity and errors to form. It is then shown that applying a windowing function suppressing the leakage and reduces the subsequent errors. From this basis, the STHT is then detailed in Section 2.3, beginning with a derivation of the STFT. After describing how to apply the Hilbert Transform transfer function and the time-frequency masking, the phase-

quadrature signal is assembled in the time domain with the weighted overlap-add synthesis method. As a numerical demonstration of the STHT, a linear free decay response is analyzed such that there are known truth values for the amplitude, natural frequency, and damping ratio. While both the Hilbert Transform and the STHT exhibit similar end effects, the STHT results have negligible error throughout the central portion of the response where the Hilbert Transform shows significant ringing artifacts. A second numerical case study is then given in the form of a complicated multicomponent signal that contains three harmonics that sweep through coincident frequency ranges while also decaying in amplitude. The STHT is able to accurately extract each component from the signal, while both the EMD and VMD methods fail. Noise is then added to the signal to demonstrate that legible results can still be produced by the STHT under such circumstances. For an experimental test case, a nonlinear free decay from a benchmark structure is presented. With the STHT, quasi-linear natural frequency and damping ratio curves for individual modes can be extract from the complicated, multicomponent response. A set of overall conclusions and findings are then given to close the chapter.

2.2. Review of The Hilbert transform

2.2.1. Background – Amplitude, Phase, and Nonlinear System ID

The simplest form of a harmonic signal is an oscillation with time invariant amplitude and frequency. This type of signal, commonly known as a stationary oscillation, is typically represented as in Eq. (15), in which t is time and $A_0, \psi_0 \,\&\, \omega_0$

are real-valued constants representing respectively the amplitude, phase offset, and frequency.

$$s(t) = A_0 \cos(\omega_0 t + \psi_0) \tag{15}$$

In many applications, such as electronic circuit analysis and communications theory [83], it is advantageous to instead represent stationary oscillations in the complex-valued form shown in Eq. (16) where i is the unit imaginary number, $\mathrm{i} = \sqrt{-1}$. In this equation, the real part is equivalent to Eq. (15) while the imaginary component is its phase-quadrature complement, exhibiting the same amplitude but shifted in phase by 90°. The benefit of this formulation is that the real and imaginary components are orthogonal, such that the oscillation can be represented in a vector space defined by the complex plane.

$$p(t) = A_0 \cos(\omega_0 t + \psi_0) + \mathrm{i} A_0 \sin(\omega_0 t + \psi_0) \tag{16}$$

Utilizing Euler's formula, Eq. (16) can be equivalently represented in polar form by a complex exponential as in Eq. (17), where e is the natural exponential function. In this format, known as the phasor representation, oscillating functions can be interpreted as a constant magnitude vector rotating in the complex plane at a fixed rate with initial position defined by the amplitude and phase offset.

$$p(t) = A_0 \mathrm{e}^{\mathrm{i}(\omega_0 t + \psi_0)} \tag{17}$$

To characterize nonstationary oscillations, the phasor concept can be generalized to incorporate arbitrary time varying representations of amplitude and phase. This more general formulation is shown in Eq. (7), where $A(t)$ and $\psi(t)$ define a time-dependent amplitude envelope and phase angle. These are combined into a

complex quantity $q(t)$, which can be represented in rectangular or polar form through Euler's formula, such that the orthogonal real and imaginary components are a phase-quadrature pair that exhibit the same amplitude envelope.

$$q(t) = A(t)\cos\big(\psi(t)\big) + \mathrm{i}A(t)\sin\big(\psi(t)\big) = A(t)\mathrm{e}^{\mathrm{i}\psi(t)} \tag{18}$$

Within this time variant framework, the concept of frequency is typically adapted to signify the instantaneous rate of change of the phase angle. As depicted in Eq. (19), this instantaneous frequency, $\omega(t)$, is defined as the first derivative of the phase, $\psi(t)$, with respect to time [92].

$$\omega(t) = \frac{d}{dt}\psi(t) \tag{19}$$

In dynamic systems analysis, the response to an input can typically be regarded as being a nonstationary oscillation. The simplest type of response signal is the free decay of a linear single-degree-of-freedom (SDOF) system, in which the closed-form solution of the displacement is shown in Eq. (20) [103]. In this equation, the initial displacement and velocity conditions are x_0 & $\dot{x}_0$, and the natural frequency, damping ratio, and damped natural frequency are ω_n, ζ, and $\omega_\mathrm{d} = \omega_\mathrm{n}\sqrt{1-\zeta^2}$.

$$x(t) = \mathrm{e}^{-\zeta\omega_\mathrm{n}t}\left(x_0\cos(\omega_\mathrm{d}t) + \frac{\dot{x}_0 + \zeta\omega_\mathrm{n}x_0}{\omega_\mathrm{d}}\sin(\omega_\mathrm{d}t)\right) \tag{20}$$

The expression above can be equivalently represented in terms of a general nonstationary oscillation with an amplitude envelope and phase angle as shown in Eq. (21), where $A(t)$ and $\psi(t)$ are then as given in Eq. (22). In this form the initial displacement and velocity conditions translate to an initial amplitude and phase angle,

and the evolution of the response through time is dictated by the decaying exponential term in the amplitude and the damped natural frequency in the phase.

$$x(t) = A(t)\cos\big(\psi(t)\big) \tag{21}$$

$$\begin{aligned} A(t) &= \sqrt{x_0^2 + \left(\frac{\dot{x}_0 + \zeta\omega_\mathrm{n} x_0}{\omega_\mathrm{d}}\right)^2}\, \mathrm{e}^{-\zeta\omega_\mathrm{n} t} \\ \psi(t) &= \omega_\mathrm{d} t - \arctan\left(\frac{\dot{x}_0 + \zeta\omega_\mathrm{n} x_0}{x_0\omega_\mathrm{d}}\right) \end{aligned} \tag{22}$$

While the response in Eq. (21) is a real-valued, oscillating signal, it can be put into the generalized phasor form by combining it with an imaginary, phase-quadrature complement as demonstrated in Eq. (7). This is illustrated in Figure 9, in which a decaying oscillation, with $\omega_\mathrm{n} = 10\ \mathrm{Hz}$ and $\zeta = 5\%$, is represented as a complex-valued signal. In the left-side plot, the signal is given in polar form in the complex plane. As time progresses, the amplitude decays from an initial value and converges toward the origin, forming the spiral shown as the solid, black line. The complex value at any point in time can be represented in terms of a vector with magnitude equal to $A(t)$ and angle relative to the positive real axis given by $\psi(t)$. This is visualized in the figure by the red arrow, which would decrease in length while rotating counterclockwise about the origin at angular rate $\omega(t)$. As it evolves through time, the components of this vector projected onto the real and imaginary axes give the signal in rectangular form. This is shown in the right-side plot, in which the real part is the solid blue line, and the phase-quadrature imaginary part is the dotted orange line. The oscillatory behavior in both,

defined by $\cos(\psi(t))$ and $\sin(\psi(t))$ respectively, is bounded by the amplitude envelope, $A(t)$, given as the black dash-dot line.

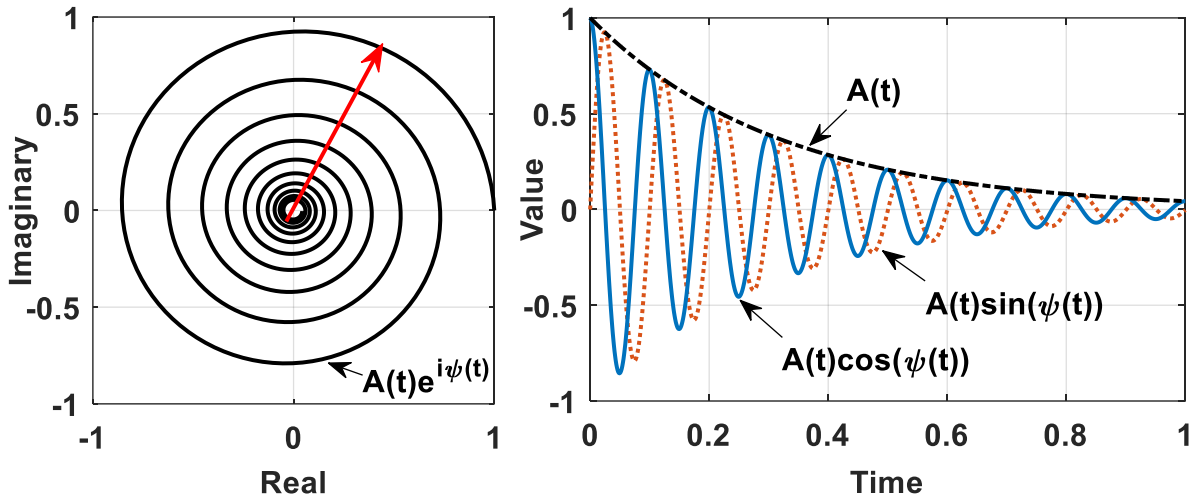

Figure 9: Interpretations of a linear free decay as a complex-valued nonstationary oscillation. (Left) The decaying oscillation is shown in the complex plane as a spiral that converges toward the origin from an initial value. The red arrow depicts the generalized phasor representation in vector form, with decaying magnitude $A(t)$ and counterclockwise rotation about the origin at angular rate $\omega(t)$. (Right) The components of the decaying oscillation as it evolves through time. The oscillatory behavior of the orthogonal real and imaginary parts, defined by $cos(\psi(t))$ and $sin(\psi(t))$ respectively, is bounded by the amplitude envelope $A(t)$.

Leveraging this generalized phasor representation of nonstationary oscillations for a nonlinear structural dynamics application, Feldman derived the FREEVIB [62] approach to nonlinear system identification of structures undergoing free response, and FORCEVIB [63] for forced response. With these methods, nonlinear representations of stiffness and viscous dissipation can be determined as a function of the response in terms of the instantaneous amplitude, phase, and the derivatives of each with respect to time as shown in Eq. (23) and Eq. (24). In these expressions, $\omega_{\mathrm{n}}(A)$ and $\zeta(A)$ are amplitude dependent, quasi-linear versions of modal natural

frequency and damping ratio, respectively, and $A(t), \dot{A}(t)$ & $\ddot{A}(t)$ are the instantaneous amplitude and its first two derivatives with respect to time. Similarly, $\omega(t)$ & $\dot{\omega}(t)$ are the instantaneous frequency and its derivative, which are each the first and second derivatives of the instantaneous phase. In the $f(t)/x(t)$ term, $x(t)$ is the displacement response of the system and $f(t)$ is an externally applied force, where both are in a complex-valued form such that the real and imaginary components are phase-quadrature complements.

$$\omega_{\mathrm{n}}(A) = \sqrt{\omega^2(t) - \frac{\ddot{A}(t)}{A(t)} + \frac{2\dot{A}^2(t)}{A^2(t)} + \frac{\dot{A}(t)\dot{\omega}(t)}{A(t)\omega(t)} + \mathrm{Re}\left(\frac{f(t)}{x(t)}\right) - \frac{\dot{A}(t)}{A(t)\omega(t)}\mathrm{Im}\left(\frac{f(t)}{x(t)}\right)} \quad (23)$$

$$\zeta(A) = \frac{1}{\omega_{\mathrm{n}}(A)}\left(-\frac{\dot{A}(t)}{A(t)} - \frac{\dot{\omega}(t)}{2\omega(t)} + \frac{1}{2\omega(t)}\mathrm{Im}\left(\frac{f(t)}{x(t)}\right)\right) \quad (24)$$

In this complete form, the FREEVIB equations include up to second order derivatives of amplitude and phase which are susceptible to significant noise errors in practice. To avoid this, these terms are typically neglected under the assumption that their effect on the desired stiffness and dissipation is negligible [104]. In this reduced form, the FREEVIB method is described by Eq. (25) and Eq. (26).

$$\omega_{\mathrm{n}}(A) \cong \sqrt{\omega^2(t) + 2\frac{\dot{A}^2(t)}{A^2(t)}} \quad (25)$$

$$\zeta(A) \cong \frac{-\dot{A}(t)}{\omega_{\mathrm{n}}(A)\,A(t)} \quad (26)$$

Utilizing FREEVIB to characterize a realistic system is complicated by the fact that instantaneous amplitude and phase are not directly observable quantities and

must be inferred from the measurable structural response. In the theoretical SDOF linear free response in Eq. (20) the initial conditions, system parameters, and model form are all known quantities, where $A(t)$ and $\psi(t)$ can be explicitly defined from the analytical solution for the response. Conversely, when attempting to characterize observed nonlinear behavior of some physical structure via an appropriate nonlinear model form, the only quantity that can be measured is typically the real-valued system response. Assuming it to be a general nonstationary oscillation of the form $A(t)\cos\big(\psi(t)\big)$, a typical approach for determining the amplitude and phase is to form the complex quantity in Eq. (7), in which the measured signal is the observable real-valued component, $\Re\big(q(t)\big) = A(t)\cos\big(\psi(t)\big)$, and an approximation of its phase-quadrature complement can be taken as the missing imaginary part, $\Im\big(q(t)\big) = A(t)\sin\big(\psi(t)\big)$. From this estimate of the real and imaginary components of $q(t)$, the desired amplitude and phase can be determined by casting it into polar form, where $A(t)$ and $\psi(t)$ are the magnitude and angle of $q(t)$ as shown in Eq. (27) and Eq. (28).

$$A(t) = |q(t)| = \sqrt{\Re\big(q(t)\big)^2 + \Im\big(q(t)\big)^2} \tag{27}$$

$$\psi(t) = \angle q(t) = \arctan\left(\frac{\Im\big(q(t)\big)}{\Re\big(q(t)\big)}\right) \tag{28}$$

Consequently, using FREEVIB is contingent upon applying some signal processing technique to obtain an estimated phase-quadrature complement to the measured structural response. In Feldman's derivations, this is accomplished through use of the Hilbert Transform [60].

2.2.2. Theory of the Hilbert Transform and Analytic Signals

The Hilbert Transform is a linear transformation that applies a 90° phase delay to the input function. Initially proposed by David Hilbert in his work regarding properties of complex-valued functions [81], the Hilbert Transform is defined mathematically in terms of the convolution in Eq. (9), where $\boldsymbol{\mathcal{H}}$ denotes the Hilbert Transform operator and $y(t)$ is an arbitrary, real-valued input. The convolution is defined as the Cauchy principal value of the integral, denoted as $\mathrm{p.v.}\int$, to avoid the singularity at $\tau = t$ [105].

$$\boldsymbol{\mathcal{H}}\big(y(t)\big) = \frac{1}{\pi}\,\mathrm{p.v.}\int_{-\infty}^{+\infty}\frac{y(\tau)}{t-\tau}\,d\tau = \frac{1}{\pi t} * y(t) \tag{29}$$

The purpose of the Hilbert Transform lies in the construction of the so-called analytic signal, which is a complex-valued function with real and imaginary components related by the Hilbert Transform. Since the Hilbert Transform imparts a 90° phase shift, the output theoretically represents a phase-quadrature complement to the input that can be taken as the imaginary component of a combined complex-valued signal. This is shown in Eq. (30), where $z(t)$ is the complex-valued analytic signal comprised of the input signal, $y(t)$, and its Hilbert Transform, $\boldsymbol{\mathcal{H}}\big(y(t)\big)$ [106].

$$z(t) = y(t) + \mathrm{i}\,\boldsymbol{\mathcal{H}}\big(y(t)\big) \tag{30}$$

To glean additional insight into analytic signals and form an alternative approach to applying the Hilbert Transform, the analytic signal can be reformulated in the frequency domain. First, from the definition of an analytic signal in Eq. (30), substitute in the convolution form of the Hilbert Transform for $\boldsymbol{\mathcal{H}}\big(y(t)\big)$.

$$z(t) = y(t) + \mathrm{i}\left(\frac{1}{\pi t} * y(t)\right) \tag{31}$$

Apply the continuous-time Fourier Transform to all terms as shown in Eq. (32), where the Fourier Transform is represented by $\boldsymbol{\mathcal{F}}$. In this form, the Hilbert Transform convolution has been converted into a product of the Fourier Transforms of the function $1/\pi t$ and the input signal.

$$\boldsymbol{\mathcal{F}}\big(z(t)\big) = \boldsymbol{\mathcal{F}}\big(y(t)\big) + \mathrm{i}\left(\boldsymbol{\mathcal{F}}\left(\frac{1}{\pi t}\right) \cdot \boldsymbol{\mathcal{F}}\big(y(t)\big)\right) \tag{32}$$

The Fourier Transform of $1/\pi t$ is given below in Eq. (33), in which sgn is the sign, or signum, function [106]. This shows that the Hilbert Transform is realized in the frequency domain as a simple transfer function that applies a phase shift to the Fourier coefficients of the input signal, where the positive frequency coefficients are phase shifted by -90° and the negative frequency coefficients are shifted by +90°. The magnitude of each coefficient is unchanged except for the DC term which is reduced to zero.

$$\boldsymbol{\mathcal{F}}\left(\frac{1}{\pi t}\right) = -\mathrm{i}\,\mathrm{sgn}(\omega) = \begin{cases} \mathrm{e}^{-\mathrm{i}\frac{\pi}{2}} & \omega > 0 \\ 0 & \omega = 0 \\ \mathrm{e}^{\mathrm{i}\frac{\pi}{2}} & \omega < 0 \end{cases} \tag{33}$$

Substituting this result into Eq. (32) produces Eq. (34), where the capitalized variables, $Y(\omega)\ \&\ Z(\omega)$, represent the Fourier Transform of the input and the analytic signal, respectively.

$$Z(\omega) = Y(\omega) + \mathrm{i}\Big(\big(-\mathrm{i}\,\mathrm{sgn}(\omega)\big) \cdot Y(\omega)\Big) \tag{34}$$

Since it is the imaginary component of the analytic signal that incorporates the Hilbert Transform, the frequency domain Hilbert Transformer in Eq. (34) is multiplied by i. This i functionally applies an additional +90° phase shift that returns the positive frequencies back to +0° while further shifting the negative frequencies to +180°. Using the identity $\mathrm{i}^2 = -1$, the transfer function then becomes $\mathrm{sgn}(\omega)$, as given in Eq. (35).

$$Z(\omega) = Y(\omega) + \big(\mathrm{sgn}(\omega) \cdot Y(\omega)\big) \tag{35}$$

Consolidating terms and expanding the sign function produces Eq. (36), revealing that an additional property of analytic signals is that their Fourier Transform must vanish in the negative frequencies. This is inherently tied to the definition of the analytic signal, in which the input and its Hilbert Transform are the real and imaginary components. As shown in Eq. (35), in the frequency domain this relation leads to the imaginary component of the analytic signal containing the same spectral content as the input but with no DC component and the negative frequency coefficients multiplied by -1. When these are combined to form the analytic signal from the input, the positive frequency coefficients are doubled, the DC value is retained, and the negative frequency coefficients inevitably reduce to zero. The time domain representation of the analytic signal defined in Eq. (30) can then be obtained via the inverse Fourier Transform, $y(t) + \mathrm{i}\,\boldsymbol{\mathcal{H}}\big(y(t)\big) = \boldsymbol{\mathcal{F}}^{-1}\big(Z(\omega)\big)$.

$$Z(\omega) = \begin{cases} 2Y(\omega) & \omega > 0 \\ Y(\omega) & \omega = 0 \\ 0 & \omega < 0 \end{cases} \tag{36}$$

For use in practical signal processing applications, the above derivations formulated in continuous, infinite time can be translated into a comparable discrete

and finite time domain. An analytic-like complex-valued discrete and finite signal can be constructed that exhibits the same properties as the analytic signal, where the real part is equivalent to the input signal, the real and imaginary components are orthogonal and related by the Hilbert Transform, and the Fourier Transform vanishes for negative frequencies [107]. Likewise, a finite duration Hilbert Transform emulates the ideal phase shifting behavior defined as a convolution over an infinite domain in Eq. (9). In the time domain this can be accomplished using finite impulse response (FIR) filters that directly approximate the convolution of the input signal with $1/\pi t$. Alternatively, by transferring the input signal to the frequency domain with the discrete Fourier Transform (DFT), the Hilbert Transform phase shift can be applied as a simple transfer function akin to Eq. (33) [88]. In this work the focus is on the frequency domain approach, such that the analytic-like signal can be computed by directly leveraging Eq. (36) where the spectrum of the input signal is doubled for positive frequencies and reduced to zero for negative frequencies.

In this scenario, the real-valued input to the discrete-time Hilbert Transform is a signal $y[n]$ that has been sampled at constant time intervals. The time sample indices are $n = 0,1,\cdots,N-1$, where the total number of samples is N and the overall duration of the measurement is T. The DFT of $y[n]$ is represented as $Y[k]$ with frequency sample indices $k = 0,1,\cdots,N-1$, and can be computed numerically via the Fast Fourier Transform (FFT) algorithm. If N is an even integer, the discrete frequency bins are given by $\omega[k] = \frac{1}{T}\left[0,1,\cdots,\frac{N}{2}-1,-\frac{N}{2},\cdots,-1\right]$, and if N is odd, $\omega[k] =$

$\frac{1}{T}\left[0,1,\cdots,\frac{N}{2}-\frac{1}{2},-\frac{N}{2}+\frac{1}{2},\cdots,-1\right]$. To construct the spectrum of the analytic-like signal, $Z[k]$, an analogues DFT version of Eq. (36) can be formed, as shown in Eq. (37) for N being an even integer. Note that in this case, the Nyquist frequency bin, $k = N/2$, is retained and unaltered akin to the DC term even though it is generally considered to be a negative frequency component. If N is an odd integer, there is no Nyquist frequency coefficient and the positive frequency coefficients to double are from $k = 1,\cdots,(N-1)/2$, and the negative frequency bins to set to zero are $k = (N-1)/2 + 1,\cdots,N-1$ [107].

$$Z[k] = \begin{cases} Y[0] & k = 0 \\ 2Y[k] & k = 1,\cdots,N/2-1 \\ Y[N/2] & k = N/2 \\ 0 & k = N/2+1,\cdots,N-1 \end{cases} \tag{37}$$

The inverse DFT of $Z[k]$, which can be computed numerically with the inverse Fast Fourier Transform (IFFT) algorithm, produces the analytic-like signal, $z[n]$, which contains the original input $y[n]$ as its real part, and the discrete Hilbert Transform of this input, $\boldsymbol{\mathcal{H}}_{\mathbf{d}}(y[n])$, as its imaginary part [107].

$$z[n] = y[n] + \mathrm{i}\boldsymbol{\mathcal{H}}_{\mathbf{d}}(y[n]) \tag{38}$$

Assuming the necessary computational resources for determining the FFT of $y[n]$ are accessible, this approach offers a simple method for simultaneously computing the discrete Hilbert Transform of the input signal and forming the associated analytic-like signal.

2.2.3. Estimating the Amplitude and Phase of Harmonic Signals

The Hilbert Transform was introduced in the preceding section with the input, $y(t)$, being some real-valued function and the output, $\boldsymbol{\mathcal{H}}\big(y(t)\big)$, being its phase-quadrature complement. While this general formulation proved useful for presenting the fundamental properties of analytic signals, the focus of this work is analyzing structural dynamics responses in which the expected input is likely to be a nonstationary oscillation. As previously described in Eq. (7), this type of harmonic signal can be characterized by two arbitrary, time-varying, real-valued functions in the form of an amplitude envelope, $A(t)$, and phase angle, $\psi(t)$. These were combined to produce a complex-valued quantity, $q(t)$, which can be represented in polar form as a complex exponential, or in rectangular form through Euler's formula with real and imaginary components that are both subject to the same amplitude envelope and are a phase-quadrature pair.

The underlying issue in the FREEVIB and FORCEVIB methods detailed above was that the necessary amplitude, $A(t)$, and phase, $\psi(t)$, are not directly observable and must be estimated from the response of the dynamic system. This can be accomplished with the Hilbert Transform by taking the measured dynamic response to be the input signal, and assuming that these are representative of the real part of $q(t)$ as shown in Eq. (39).

$$y(t) = A(t)\cos\big(\psi(t)\big) \tag{39}$$

If the Hilbert Transform of this input replicates the imaginary component of $q(t)$, the associated analytic signal, $z(t)$, is equivalent to $q(t)$ and can be cast into polar form to generate the desired amplitude and phase. The primary complication in this process

is that this input signal is the product of two functions, $A(t)$ and $\cos\big(\psi(t)\big)$, such that the output in general is a phase-quadrature complement to the combined term. To reproduce the imaginary part of $q(t)$, the phase shifting effect of the Hilbert Transform must exclusively transform the cosine term to sine while passing the amplitude through unchanged. This can be accomplished through the Bedrosian Identity [84], which simplifies the Hilbert Transform of the combined signal into the product of the amplitude envelope with the Hilbert Transform of the cosine of the phase. While this is a convenient result, the Bedrosian Identity is only valid if certain conditions on the input signal are satisfied. These conditions are explored in detail in the next section. For now, operating under the assumption that the Bedrosian Identity is valid, the Hilbert Transform of the input signal defined in Eq. (39) produces the correct imaginary part of $q(t)$ as shown in Eq. (40), and the subsequent analytic signal is equivalent to the true quadrature signal, $z(t) = q(t)$.

$$\boldsymbol{\mathcal{H}}\big(y(t)\big) \cong A(t)\sin\big(\psi(t)\big) \tag{40}$$

Thus, the complex quantity formed from the input and its Hilbert Transform can be cast from rectangular to polar form as demonstrated in Eq. (41), where the amplitude, $A(t)$, and phase, $\psi(t)$, are the magnitude and angle of the analytic signal, $z(t)$, as shown in Eq. (42) and Eq. (43), respectively [88].

$$z(t) = y(t) + \mathrm{i}\,\boldsymbol{\mathcal{H}}\big(y(t)\big) \cong A(t)\mathrm{e}^{\mathrm{i}\psi(t)} \tag{41}$$

$$A(t) \cong |z(t)| = \sqrt{y(t)^2 + \boldsymbol{\mathcal{H}}\big(y(t)\big)^2} \tag{42}$$

$$\psi(t) \cong \angle z(t) = \arctan\left(\frac{\boldsymbol{\mathcal{H}}\big(y(t)\big)}{y(t)}\right) \tag{43}$$

These derivations can also be adapted to the discrete and finite time domain detailed in the previous section, where the analytic-like signal, $z[n]$, is formed from a sampled input signal, $y[n]$, and its discrete Hilbert Transform, $\boldsymbol{\mathcal{H}}_{\mathbf{d}}(y[n])$. Assuming the Bedrosian Identity is still valid in this scenario, the magnitude and angle of the complex values in $z[n]$ produce estimates of the amplitude, $A[n]$, and phase, $\psi[n]$, at each sample point [108].

As a demonstration of how this procedure is typically applied to a measured structural response, the discrete Hilbert Transform can be used to estimate the phase-quadrature complement to the linear SDOF free response previously described in Eq. (20) and shown above in Figure 9. After creating a discretized version of the response values defined by Eq. (20), its associated analytic-like signal can be formed in the frequency domain using Eq. (37). Inverting the spectrum to the time domain produces the complex-valued signal $z[n]$, where the real part is the observable free response, the imaginary part is its discrete Hilbert Transform, and the estimated amplitude and phase are the magnitude and angle of these complex values.

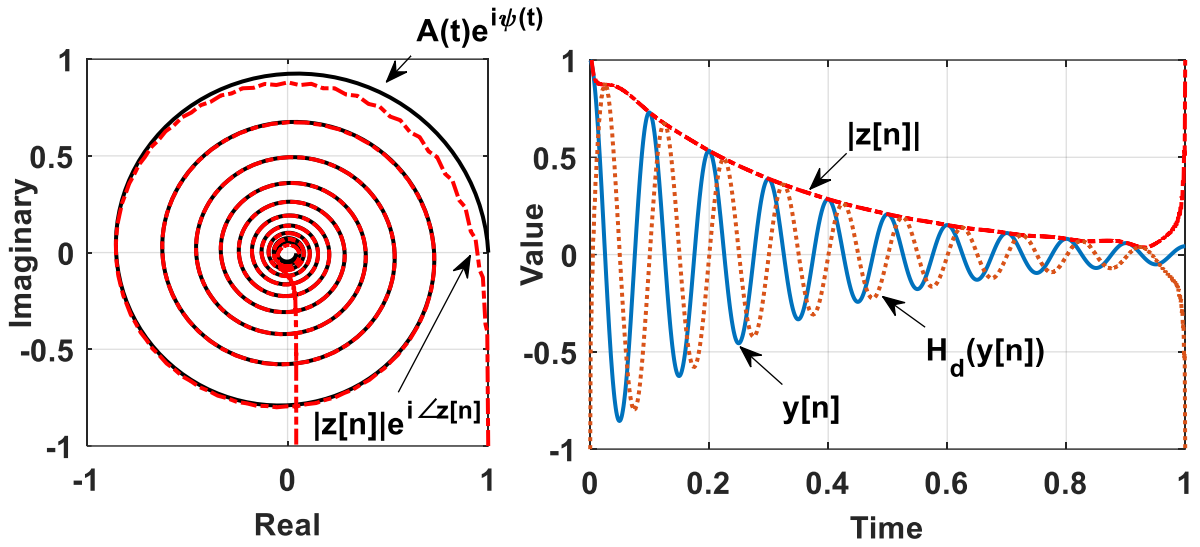

Figure 10: Estimating the phase-quadrature complement to a linear free decay with the discrete Hilbert Transform. In both the polar and rectangular forms, there are prominent end effects where the results deviate from the expected true phase-quadrature signal.

The results of this computation are shown in Figure 10, where the right plot shows the components of the analytic-like signal in rectangular form with respect to time, with the real part as the solid blue line, the imaginary part as the orange dotted line, and their magnitude as the red dash-dot line. While the central portion of the analytic-like signal appears relatively correct when compared to the expected exact quadrature curves seen on the right side of Figure 9, the beginning and end contain large, sudden spikes contaminating both the imaginary and magnitude curves. These end effects are also visible in the left plot, which shows the computed analytic-like signal in polar form in the complex plane as the dash-dot red line and the exact complex-valued quadrature signal from the left side of Figure 9 as the solid black line. While the spiral behavior of each complex signal overlays through the middle of the response, the end effects are visible here as the ends of the analytic-like signal deviating toward large negative imaginary values. These signals are also shown in the

frequency domain in Figure 11. Both curves have the same peak frequency but exhibit drastically different leakage profiles. While the true quadrature signal contains significant spectral content in the negative frequency band, the negative frequency components of the analytic signal are zero.

To understand this discrepancy, recall that, in general, the Hilbert Transform produces a phase-quadrature complement to the complete input signal, $\mathcal{H}\big(y(t)\big) = \mathcal{H}\big(A(t)\cos\big(\psi(t)\big)\big)$, such that the negative frequency components of the subsequent analytic signal are zero. Conversely, the complex signal shown in Eq. (7) is defined in terms of Euler's formula, where the real and imaginary components are phase-quadrature complements that properly portray the expected amplitude and phase. The Bedrosian Identity essentially describes the specific situation in which these two definitions produce the same result. For the free decay response, the true quadrature signal includes significant negative frequency components that are required to correctly represent the amplitude and phase. As the analytic-like signal cannot reproduce these values, the Bedrosian Identity is not a valid assumption in this scenario. Instead, the discrete Hilbert Transform result contains end effects which are a time domain artifact of applying the Hilbert Transform frequency domain transfer function that causes the negative frequency components of the associated analytic-like signal to reduce to zero.

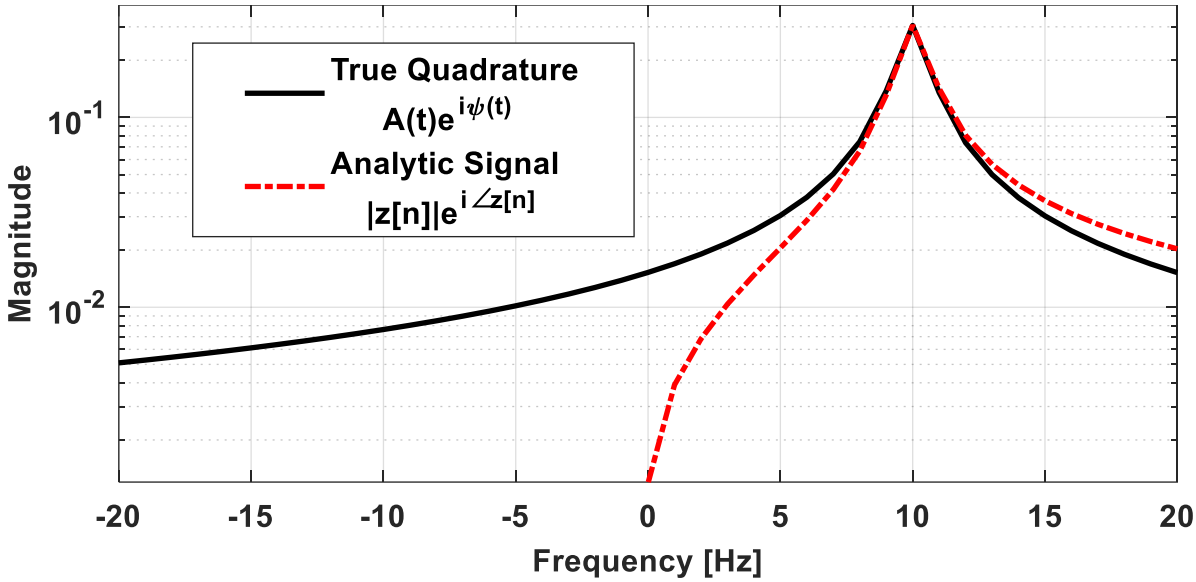

Figure 11: Comparing the linear free decay true quadrature and analytic signals in the frequency domain. The true quadrature signal contains significant spectral content in the negative frequency band. It is impossible for the analytic signal to reproduce this, as the negative frequency components are zero by the definition of the Hilbert Transform.

2.2.4. The Accuracy of the Bedrosian Identity

While the process for estimating the amplitude and phase of some sampled, finite signal with the discrete Hilbert Transform is both conceptually simple and numerically efficient, it is also subject to significant errors if the Bedrosian Identity is invalid. The simple example above shows that this identity is not applicable to even the most basic structural dynamic response and the discrete Hilbert Transform result is contaminated by significant end effects. As errors such as these are very commonly observed when utilizing the Hilbert Transform in general signal processing tasks, this implies that most practical applications of the Hilbert Transform are on signals where the Bedrosian Identity does not hold. To determine the conditions under which use of the Bedrosian Identity is a valid assumption, this section investigates the error between the general

output from the Hilbert Transform and the phase-quadrature complement required to correctly produce the amplitude and phase.

In terms of the free decay detailed in the previous section, the difference between the desired quadrature signal and the discrete Hilbert Transform is shown in Figure 12, where the true phase-quadrature signal is the solid blue line, the discrete Hilbert Transform approximation is the dash-dot orange line, and the error between them is the dotted yellow line.

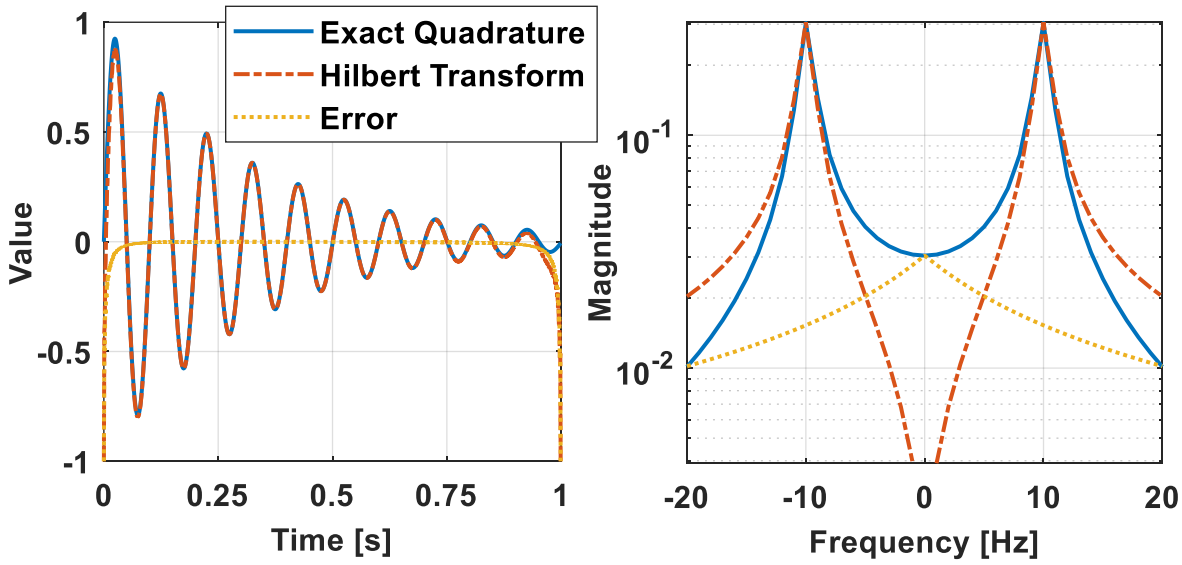

Figure 12: The exact phase-quadrature signal compared to the discrete Hilbert Transform approximation. (Left) The imaginary component of the true quadrature signal and the Hilbert Transform result in the time domain. (Right) The same curves, but in the frequency domain. The error is predominately a low frequency signal that is symmetric in the positive and negative frequency bands.

The end effects are visible in the time domain in the left plot, where the error is near zero throughout the central portion of the oscillations while rapidly increasing in magnitude at the start and end. The right plot depicts the same signals in the frequency domain, where the error is found to have a peak at zero Hertz and

symmetrically propagates into the positive and negative frequency bands, showing that the error is equally distributed in both the positive and negative frequency ranges.

An expression that quantifies this error can be constructed based on the difference between the general result of the Hilbert Transform and the exact quadrature signal. By identifying when the error approaches zero, the conditions under which the Bedrosian Identity is valid can be established. An error metric such as this was first developed by Nuttall in [85] in the context of communications theory to expand upon the original definition of the Bedrosian Identity [84]. The following is an adaptation of Nuttall's derivation, where it has been generalized to better function in a structural dynamics framework.

For a general nonstationary oscillation, the Hilbert Transform input signal in Eq. (39) took the form $y(t) = A(t)\cos\big(\psi(t)\big)$. To accurately recover the amplitude and phase from the magnitude and angle of the subsequent analytic signal, the output from the Hilbert Transform must be $\boldsymbol{\mathcal{H}}\big(y(t)\big) = A(t)\sin\big(\psi(t)\big)$. The error between this desired output and the Hilbert Transform of the input can be quantified in terms of their total squared difference.

$$E = \int_{-\infty}^{+\infty} \left|\boldsymbol{\mathcal{H}}\big(A(t)\cos\big(\psi(t)\big)\big) - A(t)\sin\big(\psi(t)\big)\right|^2 dt \tag{44}$$

Through Parseval's Theorem, the argument of E can be equivalently represented in the frequency domain. The Hilbert Transform term then becomes a product of the Hilbert Transform transfer function with the Fourier Transform of the input.

$$E = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|\left(-\mathrm{i}\,\mathrm{sgn}(\omega)\right)\cdot \boldsymbol{\mathcal{F}}\left(A(t)\cos\left(\psi(t)\right)\right) - \boldsymbol{\mathcal{F}}\left(A(t)\sin\left(\psi(t)\right)\right)\right|^2 d\omega \qquad (45)$$

As the input and quadrature time domain signals are strictly real-valued, their Fourier Transforms are Hermitian, and the positive and negative frequency components of each spectrum are complex conjugates. This can be more explicitly represented in the expression by expanding the cosine and sine terms into complex exponentials with Euler's Formulas.

$$E = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|\left(-\mathrm{i}\,\mathrm{sgn}(\omega)\right)\cdot \boldsymbol{\mathcal{F}}\left(A(t)\cdot\tfrac{1}{2}\left(\mathrm{e}^{\mathrm{i}\psi(t)} + \mathrm{e}^{-\mathrm{i}\psi(t)}\right)\right) - \boldsymbol{\mathcal{F}}\left(A(t)\cdot\tfrac{1}{2\mathrm{i}}\left(\mathrm{e}^{\mathrm{i}\psi(t)} - \mathrm{e}^{-\mathrm{i}\psi(t)}\right)\right)\right|^2 d\omega \qquad (46)$$

This format shows that the Fourier Transforms of the input signal and the desired quadrature result can both be represented in terms of the spectrums of the complete quadrature signal, $\boldsymbol{\mathcal{F}}\left(A(t)\mathrm{e}^{\mathrm{i}\psi(t)}\right)$, and its complex conjugate, $\boldsymbol{\mathcal{F}}\left(A(t)\mathrm{e}^{-\mathrm{i}\psi(t)}\right)$. For convenience, these terms are respectively represented as the Hermitian functions $f(\omega)$ and $f^*(\omega)$ as shown in Eq. (47). As $A(t)\mathrm{e}^{\mathrm{i}\psi(t)}$ and $A(t)\mathrm{e}^{-\mathrm{i}\psi(t)}$ are complex valued in the time domain their frequency spectrums are single-sided, where $f(\omega)$ is primarily contained within the positive frequency components, and its complex conjugate, $f^*(\omega)$, is primarily in the negative frequency components.

$$f(\omega) = \boldsymbol{\mathcal{F}}\left(A(t)\mathrm{e}^{\mathrm{i}\psi(t)}\right) \quad \& \quad f^*(\omega) = \boldsymbol{\mathcal{F}}\left(A(t)\mathrm{e}^{-\mathrm{i}\psi(t)}\right) \qquad (47)$$

Substituting $f(\omega)$ and $f^*(\omega)$ into Eq. (46) yields Eq. (48).

$$E = \frac{1}{8\pi}\int_{-\infty}^{+\infty} \left|\left(-\mathrm{i}\,\mathrm{sgn}(\omega)\right)\cdot\left(f(\omega) + f^*(\omega)\right) + \mathrm{i}\left(f(\omega) - f^*(\omega)\right)\right|^2 d\omega \qquad (48)$$

This shows that the Hermitian spectrum of the input signal is composed of the single-sided spectrums of the desired quadrature signal and its complex conjugate, such that $\boldsymbol{\mathcal{F}}\big(A(t)\cos\big(\psi(t)\big)\big) = \frac{1}{2}f(\omega) + \frac{1}{2}f^*(\omega)$. To exactly transform this input into the desired phase-quadrature output, the $f(\omega)$ term must be multiplied by $-\mathrm{i}$, and the $f^*(\omega)$ term must be multiplied by i, as in $\boldsymbol{\mathcal{F}}\big(A(t)\sin\big(\psi(t)\big)\big) = -\mathrm{i}\frac{1}{2}f(\omega) + \mathrm{i}\frac{1}{2}f^*(\omega)$. Meanwhile, the Hilbert Transform applies the phase shift in terms of $-\mathrm{i}\,\mathrm{sgn}(\omega)$, where the positive frequency components are multiplied by $-\mathrm{i}$, and the negative components are multiplied by i. To resolve the $\mathrm{sgn}(\omega)$ term, the integral can be split into separate expressions for the negative and positive sides of the frequency spectrum.

$$E = \frac{1}{8\pi}\int_{-\infty}^{0}\left|\mathrm{i}\big(f(\omega) + f^*(\omega)\big) + \mathrm{i}\big(f(\omega) - f^*(\omega)\big)\right|^2 d\omega + \frac{1}{8\pi}\int_{0}^{+\infty}\left|-\mathrm{i}\big(f(\omega) + f^*(\omega)\big) + \mathrm{i}\big(f(\omega) - f^*(\omega)\big)\right|^2 d\omega \tag{49}$$

The error term can then be simplified as shown in Eq. (50).

$$E = \frac{1}{2\pi}\int_{-\infty}^{0}|\mathrm{i}f(\omega)|^2 d\omega + \frac{1}{2\pi}\int_{0}^{+\infty}|-\mathrm{i}f^*(\omega)|^2 d\omega \tag{50}$$

This expression reveals that, if $f(\omega)$ and $f^*(\omega)$ extend into the negative and positive frequency bands, respectively, there are residual errors due to the Hilbert Transform. Any portion of $f(\omega)$ that falls into the negative frequency range is incorrectly phase shifted by i, and similarly any portion of $f^*(\omega)$ on the positive side is improperly shifted by $-\mathrm{i}$. For the linear free decay, these components of the error term are shown explicitly in Figure 13. In this plot, the spectrum of the real-valued free decay signal is the black dash-dot line, and the spectrums of the true quadrature signal and its

complex conjugate, $f(\omega)$ and $f^*(\omega)$, are given by the solid blue and the dashed orange curves, respectively. The portion of each that crosses zero Hertz and receives the wrong phase shift from the Hilbert Transform is marked by the respective arrow and label. These rogue segments, which are at a maximum at 0 [Hz] and symmetric in magnitude in the positive and negative frequency bands, are exactly the error visible in the right side of Figure 12.

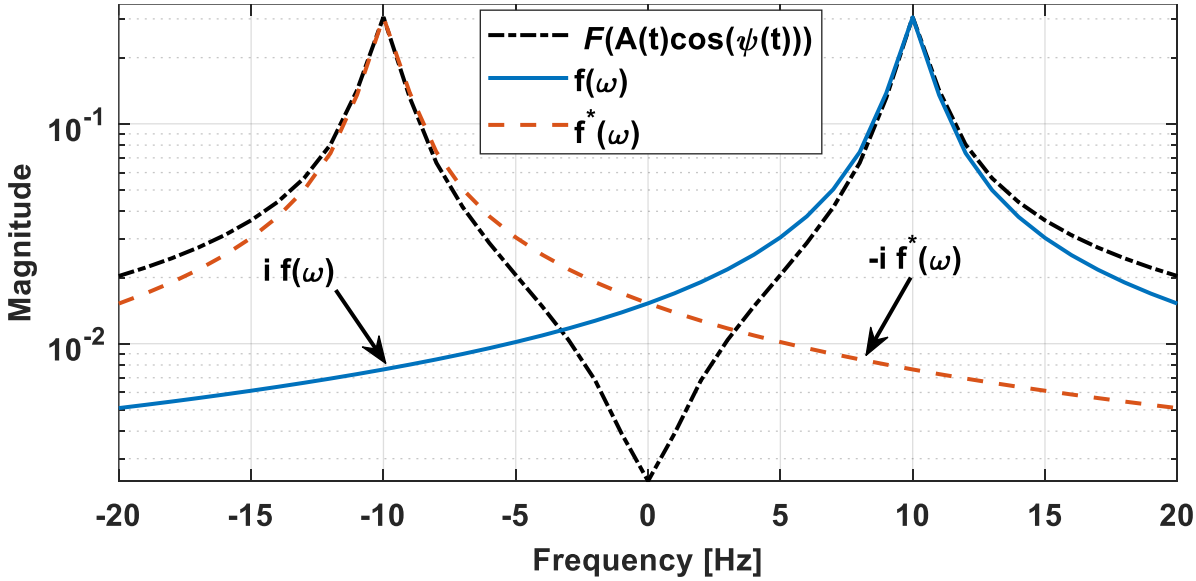

Figure 13: The Hermitian frequency spectrum of the real-valued free decay input signal, and the single-sided spectrums of the exact quadrature signal and its complex conjugate that can be combined to form it. When applying the Hilbert Transform to the free decay, the segments of each quadrature spectrum that cross over 0 [Hz] receive the marked incorrect phase shifts.

The primary issue with the error metric given in Eq. (50) is that it cannot be used to give quantitative results in practical applications, since it is dependent on the unknown spectrum of the desired complete quadrature signal. However, the derivation above can still be used to determine conditions for when the Bedrosian Identity is valid in general, which is when the error vanishes. For this to occur, $f(\omega)$ and $f^*(\omega)$ must be completely confined to the positive and negative frequency bands, respectively,

such that the Hilbert Transform transfer function, $-\mathrm{i}\,\mathrm{sgn}(\omega)$, exactly replicates the spectral phasing of the desired quadrature signal. Since $f(\omega)$ and $f^*(\omega)$ collectively represent the Hermitian spectrum of the real-valued input signal, if the spectrum on the positive and negative sides vanishes as it approaches 0 Hertz, then the Fourier Transform functionally separates the input signal into the phase-quadrature compliment and its complex conjugate, where each is fully contained within their respective frequency bands. Thus, the Bedrosian Identity is valid for an arbitrary input if its frequency content vanishes at frequencies near 0 Hertz. However, in realistic signals leakage and noise will generally produce significant frequency components in this range, leading to end effect errors in the Hilbert Transform output.

The accuracy of the phase-quadrature output from the Hilbert Transform can be increased by mitigating the magnitude of frequency content near 0 Hertz in the input signal. In many cases, this content is comprised of spectral leakage due to bringing a nonperiodic time signal to the frequency domain. Suppressing this leakage can be achieved by simply applying a windowing function to the signal. A window function intended for spectral analysis is generally smooth and symmetric about a central maximum while tapering to zero at the edges. By multiplying the nonperiodic signal by the window, it is forced into a pseudo-periodic form which exhibits reduced leakage. This is demonstrated in Figure 14, which shows the linear free decay from the previous section as the solid blue line. In the time domain on the left, the free decay is visibly nonperiodic, as it starts at an initial value of 1 and converges toward zero as time progresses. Applying a windowing function to this signal produces the

orange dash-dot line, which begins and ends near zero. The windowing function utilized in this scenario is a 4-term Blackman-Harris window [109], which is shown in the left plot as the black dotted line. The spectrums for the standard free decay and its windowed counterpart are shown in the right plot. The frequency content near 0 Hertz is approximately three orders of magnitude lower in the windowed signal compared to the original free decay.

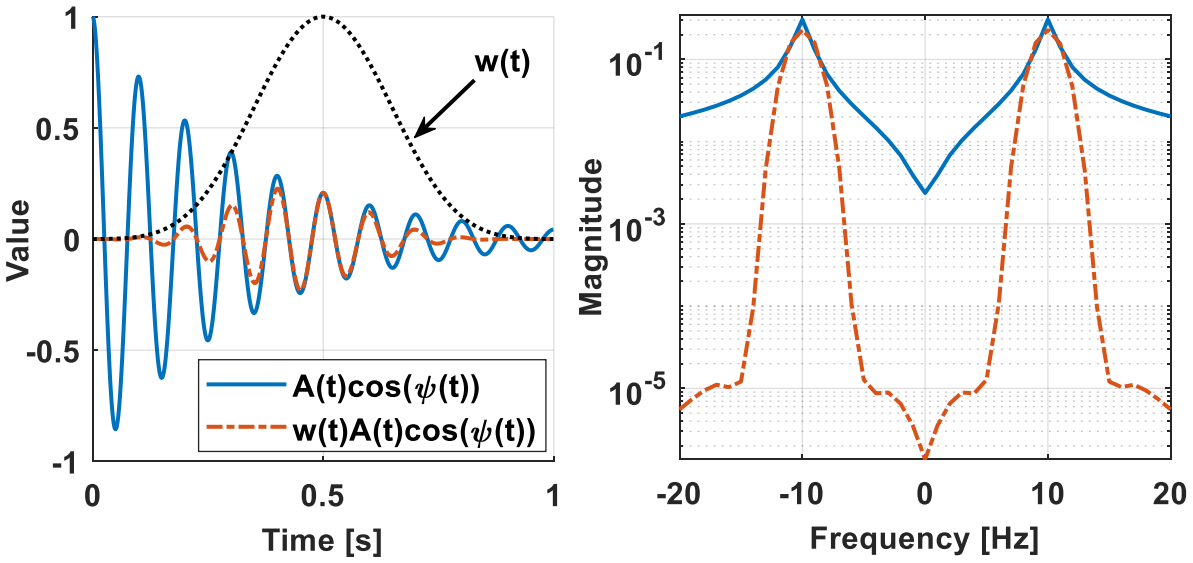

Figure 14: Spectral leakage can be suppressed by applying a windowing function to a nonperiodic signal. For the free decay signal, a 4-term Blackman-Harris window reduces the frequency content near 0 Hertz by approximately three orders of magnitude.

The discrete Hilbert Transform of the windowed free decay signal can be computed in the frequency domain as done for the unalter free decay signal in the previous section. The results of this procedure are shown in Figure 15, which depicts the Hilbert Transform output along with a windowed version of the exact quadrature signal, and the error between the two curves. These signals are shown in the time domain in the top left plot, where the oscillations seemingly overlay with no

discernable error. To properly view the error curve, it is also shown separately in the bottom left plot. While the shape of the error curve is similar to that of the regular free decay result in Figure 12, in that it is near zero through the central portion of the signal and peaks at the ends, it is about four orders of magnitude smaller. This is also observed in the frequency domain plot on the right, in which the spectrum of the error is on the order of 10^{-6}, while the error spectrum for the free decay in Figure 12 is on the order of 10^{-2}. Since the exact quadrature result is known for the free decay signal, the expression for the error E can be utilized to quantify the accuracy of these Hilbert Transform results. The error for the original free decay signal shown in Figure 12 is $E = 4.7$, while the error after applying the windowing function is $E = 2.1 \cdot 10^{-8}$.

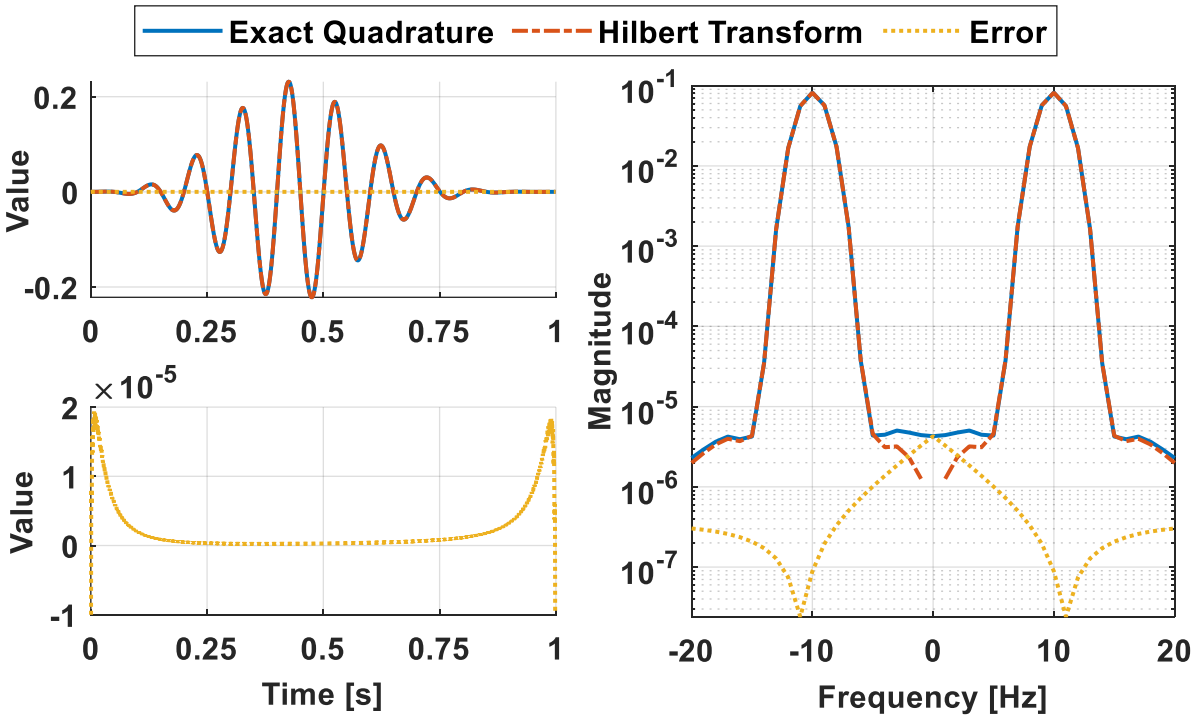

Figure 15: The exact phase-quadrature signal compared to the discrete Hilbert Transform approximation for the windowed free decay signal. While end effects are still present in the error curve, they are approximately four orders of magnitude smaller than those observed in the Hilbert Transform of the standard free decay signal.

This section detailed the derivation of an expression for the error between the general output from the Hilbert Transform and the expected phase-quadrature complement defined by the Bedrosian Identity. By defining and quantifying this error, the condition for the Bedrosian Identity to be valid and accurate could be identified. It was found that, if the frequency content of the input signal vanishes at frequencies near 0 Hertz, then the correct phase shift is applied to the spectrum by the Hilbert Transform and the result is equivalent to the true quadrature signal. However, in practical applications the input signal is likely to not be precisely periodic, such that significant spectral leakage will cause the Bedrosian Identity to fail and end effects to contaminate the Hilbert Transform output. To suppress the leakage and dramatically reduce the presence of end effects, a windowing function can be applied to the input signal. However, this alters the output signal, such that the subsequently estimated amplitude is predominately in the shape of the applied windowing function. The next section presents a means of accounting for the window and reconstructing the phase-quadrature signal to accurately produce the amplitude and phase with minimal errors.

2.3. Theory of The Short-time Hilbert Transform

The Hilbert Transform is frequently used as a signal processing tool to estimate the amplitude and phase of an arbitrary nonstationary oscillation. However, the results are often corrupted by end effects, where the most common solution is to simply truncate the result and discard that portion of the signal. The previous section presented a detailed review of pertinent background information on the Hilbert

Transform. It was shown that the accuracy of the estimated amplitude and phase are entirely dependent on the validity of the Bedrosian Identity, and that the Bedrosian Identity is not accurate when the signal contains significant frequency content near 0 Hertz. This typically occurs with measured signals due to spectral leakage, caused by the signal being a finite duration that cannot be rendered into an infinitely smoothly periodic form. Consequently, the beginning and end of the signal likely have discontinuities which manifest as leakage in the frequency domain, causing the Bedrosian Identity to fail, and end effects to form in the Hilbert Transform result. It was then demonstrated that applying a windowing function to the signal suppresses the leakage and mitigates the subsequent end effects at the cost of altering the shape of the signal. To generate a correctly shaped result while also taking advantage of the benefits of windowing, this section introduces the Short-time Hilbert Transform (STHT), in which the Hilbert Transform is integrated into the framework of the Short-time Fourier Transform.

2.3.1. The Short-time Fourier Transform

The Short-time Fourier Transform (STFT) is a time-frequency analysis tool commonly used to determine how the frequency content in a signal evolves through time. This is done by segmenting the signal into overlapping sections, applying a windowing function to each, and then utilizing the Fourier Transform to transfer each to the frequency domain. For a comprehensive derivation and analysis of the STFT, see [44]. For the purposes here, the signal is assumed to be in the finite and discretized form of a measured structural response, previously defined as $y[n]$. To

compute the STFT representation of this signal, it must first be segmented and windowed as shown in Eq. (51).

$$\tilde{\mathbf{y}}[m,n] = y[n]w_{\mathrm{a}}[n - mR] \tag{51}$$

In this scenario, the signal is divided into M segments, where $m = 0,1,\ldots,M-1$ are the segment indices. Each segment is of length L, and the amount of overlap with adjacent segments is defined by the hop size R, where in general, $R < L < N$. The windowing function, also known as the analysis window in this context, is contained within $w_{\mathrm{a}}[n]$ and is represented by L nonzero samples, the position of which are shifted by R samples for each segment m. These values produce the matrix $\tilde{\mathbf{y}}[m,n]$ containing m columns and n rows, where each column is a separate windowed segment of the signal $y[n]$. This matrix can then be transferred to the frequency domain by computing the DFT of each column via the FFT algorithm, producing the STFT of the signal, $\tilde{\mathbf{Y}}[m,k]$ as in Eq. (52). In this process, it is typically advantages to pad each column with zeros such that the length, N, is increased to be a power of 2. Doing so improves the computational efficiency of the FFT algorithm while also increasing the frequency resolution of the result.

$$\tilde{\mathbf{Y}}[m,k] = \mathrm{DFT}(\tilde{\mathbf{y}}[m,n]) \tag{52}$$

The segment length and analysis window type should be chosen based on the content in $y[n]$. Generally, shorter segments give finer time resolution while longer segments give finer frequency resolution. For an arbitrary signal, the segment length should be a balance of those, where it is short enough to adequately capture any nonstationary behavior but long enough to clearly convey what that behavior is. For

the STHT, the minimum segment length is set by what type of windowing function is selected. The primary difference between analysis windows are the frequency domain characteristics, which are the magnitude of the side lobes and the width of the main lobe relative to the number of oscillation cycles within the window. For the Hilbert Transform to produce minimal errors, the main lobe must not fall across 0 Hertz and the side lobes should be as small as possible. The 4-term Blackman-Harris window is utilized in this work as it exhibits negligible side lobes across a wide frequency range but requires at least four oscillation cycles to produce an acceptable main lobe width. Other windows have narrower main lobes at the cost of larger side lobes, such as the Hann window which only needs two oscillation cycles.

A last point of note is regarding options in defining the segments at the beginning and end of the signal. While the segments can be constructed to simply start and end with the first and last samples, this could lead to issues later in the synthesis step. A common strategy is to artificially extend the signal, such that segments can be defined that similarly extend beyond the signal bounds. This can be done by simply repeating copies of the signal in the periodic sense, placing mirrored copies that potentially better match the edge conditions, or by just padding the signal with zeros.

2.3.2. Integrating the Hilbert Transform into the STFT

The Hilbert Transform can be incorporated into the STFT by transforming the spectrum of each input signal segment into the spectrum of the associated analytic signal. This is done by adapting the frequency domain transformation defined in Eq. (37) to operate on each column of $\widetilde{\mathbf{Y}}[m,k]$ and produce the matrix $\tilde{\mathbf{Z}}[m,k]$ as shown in Eq. (53). In this expression, every DC and Nyquist coefficient is retained, the positive frequency components are doubled, and the negative frequency components are reduced to zero.

$$\tilde{\mathbf{Z}}[m,k] = \begin{cases} \widetilde{\mathbf{Y}}[m,0] & k = 0 \\ 2\widetilde{\mathbf{Y}}[m,k] & k = 1,\cdots,N/2-1 \\ \widetilde{\mathbf{Y}}[m,N/2] & k = N/2 \\ 0 & k = N/2+1,\cdots,N-1 \end{cases} \tag{53}$$

2.3.3. Mode Decomposition via Time-Frequency Masking

In addition to forming the analytic signal, the spectra of the input signal in $\widetilde{\mathbf{Y}}[m,k]$ can be further modified through time-frequency masking [93]. In this process, weights are applied to portions of the STFT matrix to amplify or diminish the magnitude of specific frequency coefficients. These weights can be defined in terms of a binary matrix, $\mathbf{M}[m,k]$, as shown in Eq. (54), which contains all zeros except for indices which correspond to spectral content that is to be retained.

$$\mathbf{M}[m,k] = \begin{cases} 1 & [m,k] \text{ to retain} \\ 0 & \text{Otherwise} \end{cases} \tag{54}$$

As the analytic signal transformation in Eq. (53) already sets the negative frequency components to zero, the functionality of the mask matrix $\mathbf{M}[m,k]$ is to remove portions of the remaining frequency information. This can be used as a form of mode

decomposition for signals which contain multiple significant oscillatory components. Assuming the oscillations are separable in the STFT representation of the signal, the masking matrix $\mathbf{M}[m,k]$ can be strategically constructed such that only frequency information related to a specific oscillation is kept and all else due to unwanted oscillations and/or noise is eliminated. The mask can then function as a time-dependent ideal bandpass filter in which the bounds of the pass band are separately defined for each segment. If $\mathbf{M}[m,k]$ is defined as a matrix of ones, then $\tilde{\mathbf{Z}}[m,k]$ is unaltered and each segment is transformed into the standard analytic-like signal form.

2.3.4. Assembling the Phase-Quadrature Signal

After modifying $\widetilde{\mathbf{Y}}[m,k]$ as detailed above, the resulting phase-quadrature signal can be rendered in the time domain by the inverting the STFT through a process called synthesis. Of the various forms of synthesis that have been proposed, the approach utilized here is the Weighted Overlap-Add (WOLA) method [110]. The first step is to apply the inverse DFT to the masked analytic signal spectrums as shown in Eq. (55), where the matrices $\tilde{\mathbf{Z}}[m,k]$ and $\mathbf{M}[m,k]$ are multiplied elementwise.

$$\tilde{\mathbf{z}}[m,n] = \mathrm{iDFT}\big(\tilde{\mathbf{Z}}[m,k] \cdot \mathbf{M}[m,k]\big) \tag{55}$$

This produces the matrix $\tilde{\mathbf{z}}[m,n]$, where each column is a complex-valued signal with the real and imaginary parts related by the discrete Hilbert Transform. Assuming $\mathbf{M}[m,k]$ has been defined such that all the significant frequency content of at least one oscillation has been inverted, each column of $\tilde{\mathbf{z}}[m,n]$ contains a signal that is centered at the shifted location of the applied window and is in the shape of the analysis window.

Each signal also includes small end effects at the edges of the segment, as previously seen in Figure 15.

In the context of the STFT and synthesis, these errors are known as artifacts and are understood to be a consequence of manipulating the frequency spectrum, which is done here by the Hilbert Transform and masking. In the WOLA method a second windowing function, known as the synthesis window, is applied as a weighting to each signal segment after computing the inverse DFT. If the synthesis window is the same as or similar to the analysis window, then the artifacts are suppressed. The artifacts can also be completely removed from each segment by defining a synthesis window that is shorter than the analysis window. To prevent gaps in the synthesized signal, the difference between the window lengths must be less than the amount of overlap provided by the hop size.

The overlap-add method is then used to assemble the weighted signal segments into a unified phase-quadrature result, $\hat{z}[n]$. This is shown in Eq. (56), where $w_s[n]$ is the synthesis window and is constructed similar to the analysis window. In this expression, $\hat{z}[n]$ is given as the ratio of two separately summed quantities. In the numerator, each column of $\tilde{\mathbf{z}}[m,n]$ is weighted by the synthesis window and then combined. Depending on the hop size and the exact form of the analysis and synthesis windows, this produces a signal that is likely distorted by the applied windows. This is corrected by the normalization term in the denominator, which accounts for the combined influence of both windows in each segment.

$$\hat{z}[n] = \frac{\sum_{m=0}^{M} \tilde{\mathbf{z}}[m,n]\, w_{\mathrm{s}}[n - mR]}{\sum_{m=0}^{M} w_{\mathrm{a}}[n - mR]\, w_{\mathrm{s}}[n - mR]} \tag{56}$$

When a desired quadrature signal contains nonzero negative frequency components, it is impossible for the Hilbert Transform to produce an analytic signal that emulates this behavior due to the requirement that the negative frequency components must be zero. While each isolated segment in $\tilde{\mathbf{z}}[m,n]$ is an analytic-like signal, the synthesis process in Eq. (56) does not preserve this condition and the assembled phase-quadrature signal $\hat{z}[n]$ is not analytic. This allows $\hat{z}[n]$ to produce a more accurate approximation of the desired quadrature signal as it is no longer constrained to the limitations of analytic signals that lead to ringing artifacts.

2.3.5. Summary

The STHT is an adaptation of the conventional frequency domain-based Hilbert Transform that replaces the Fourier Transform with the Short-time Fourier Transform. With the Fourier Transform, the spectrum of the total signal is used to form a phase-quadrature compliment to the input. If the signal is not periodic, leakage in the frequency domain leads to prominent end effects in the time domain that corrupt the edges of the result and permeate through the entire signal. In the Short-time Fourier Transform, the input signal is divided into a set of overlapping segments which are then windowed to suppress leakage in the frequency domain. After applying the Hilbert Transform transfer function to each segment and defining an appropriate masking matrix, they can be transformed back to the time domain and assembled into a phase-quadrature complement to the input signal. This process is summarized in

Figure 16. While the end effects are mitigated, they are not removed entirely. Segments that contain discontinuities that generate leakage will still exhibit errors. However, these errors are localized and confined to only those segments and do not propagate throughout the signal as a whole. The primary improvements offered by the STHT over the conventional Hilbert Transform are localized errors instead of global, and the ability to perform additional operations on the STFT spectrum such as mode decomposition through time-frequency masking.

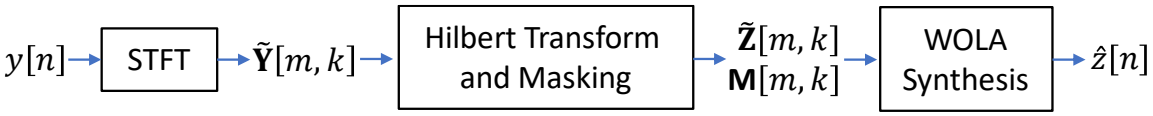

Figure 16: The components of the Short-time Hilbert Transform process

2.4. Numerical Demonstrations

This section presents two numerical case studies to demonstrate the STHT process. The first is a continuation of the linear free decay example utilized in the earlier section detailing various aspects of the Hilbert Transform. Estimates of the amplitude, natural frequency, and damping ratio are formed from the STHT as well as the typical DFT-based Hilbert Transform and a FIR filter Hilbert Transformer. The second case is a complicated multicomponent signal which contains three oscillations with decaying amplitudes that are also sweeping through overlapping frequency bands. This offers a difficult mode decomposition scenario which cannot be accomplished with typical bandpass filters and is problematic even for more sophisticated methods such as EMD and VMD. With the STHT, the individual

harmonic components can be individually extracted from the signal through time-frequency masking while simultaneously applying the Hilbert Transform to form estimates of the amplitude and phase. Noise is then added to the signal to demonstrate that the STHT is able to produce reasonable results under such conditions.

2.4.1. Linear Free Decay

As a demonstration of the STHT process and to provide a direct comparison to the standard Hilbert Transform, this section utilizes both to estimate the amplitude, natural frequency, and damping ratio of a simulated response. The signal is the linear free decay used throughout the previous sections of this work and initially shown in Figure 9. This displacement response defined in Eq. (20) was evaluated over a time span of 1 [s] at a sample rate of 1000 [Hz], with initial conditions $A_0 = 1\ [\mathrm{m}]$ and $\psi_0 = 0\ [\mathrm{Rad}]$, and system parameters $\omega_\mathrm{n} = 10\ [\mathrm{Hz}]$ and $\zeta = 5\ [\%]$. To employ the STHT on this signal, the segment size for the STFT was set to 0.4 [s], the hop size was one sample, and the 4-term Blackman-Harris window was used. Copies of the response were appended to the beginning and end such that the periodic edge conditions observed in the previous sections are maintained. The synthesis window was formed from a C^∞-smooth trigonometric blending function [111] that was 10% of the segment length. As an additional point of comparison, an FIR filter Hilbert Transformer was also applied to the response to estimate the phase-quadrature signal [88]. The filter length was the same as the STFT segment size, and a Kaiser window was applied to the filter to taper the edges.

The estimated amplitude from each approach is shown in Figure 17, in which the truth value is the solid blue line, the conventional result where the DFT-based Hilbert Transform (HT-DFT) is applied to the whole signal is shown with the yellow dashed line, the FIR filter Hilbert Transformer result (HT-FIR) is the black dotted line, and the STHT estimate is given as the orange dash-dot curve. In the full view, all estimates exhibit similar end effects and seem to follow the truth value through the central portion of the response. However, the zoomed in plots show the discrepancy between each result in detail. In the right-side zoom plot, the HT-DFT result shows the most significant ringing artifacts that propagate into the response. While the HT-FIR result has less significant errors, the STHT shows the least and converges to the truth value in the region marked by the arrow. The left zoom plot shows that the HT-DFT curve exhibits some degree of ringing artifacts throughout the response, the HT-FIR result is slightly offset from the correct amplitude, and that the STHT curve accurately follows the true amplitude with no visible error. The STHT result converges to the exact truth curve from 0.2 [s] to 0.8 [s] because segments that are defined between those points do not contain the signal edges. These are simply a windowed version of the smoothly varying response where any edge artifacts from the Hilbert Transform are removed in the WOLA assembly process, yielding a practically error-free phase-quadrature result.

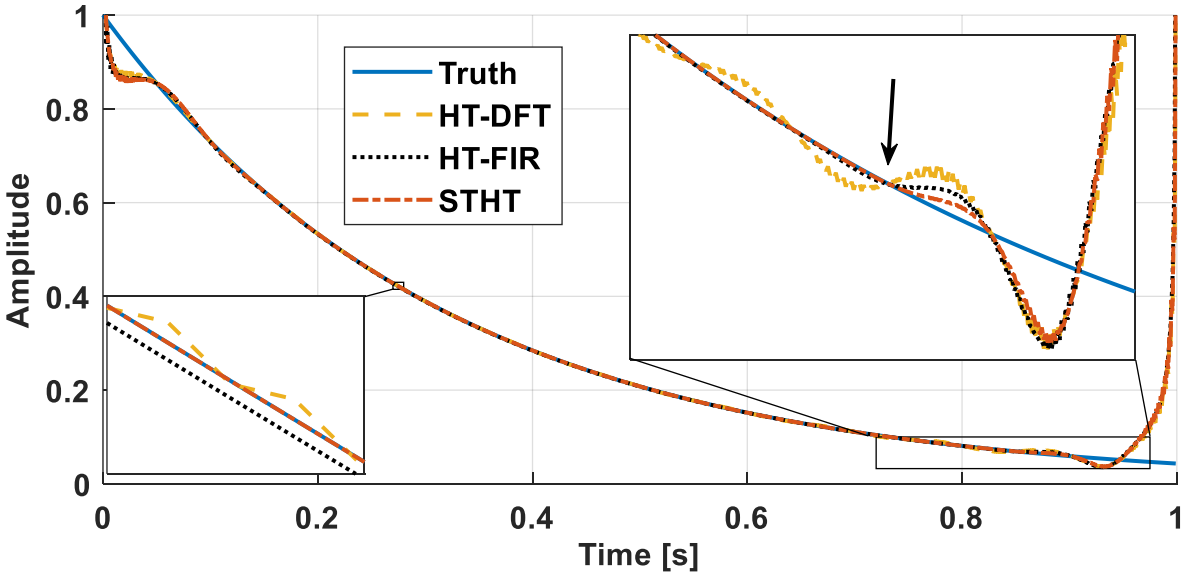

Figure 17: The amplitude of the linear free decay and the estimations from the Short-time Hilbert Transform (STHT) and the Hilbert Transform as computed via the DFT and an FIR filter. The ringing artifacts that ripple through the HT-DFT curve are suppressed in the STHT result, as noted by the arrow in the zoomed in section. While the HT-FIR curve is more accurate than the HT-DFT result, it is slightly offset from the true value and shows more ringing than the STHT.

The FREEVIB method detailed in Eq. (25) and Eq. (26) can be used to form estimates of the natural frequency and damping ratio of the system. To generate the required derivatives, a central difference approach was applied to the amplitude and phase results from each of the methods described above. The computed natural frequency and damping ratio curves are shown in the left and right sides of Figure 18, respectively. As these results are derived from derivative values, the errors are more pronounced. The HT-DFT curve shows significant ringing artifacts throughout the response, especially in the latter half. While the HT-FIR result is more accurate, it exhibits larger errors than the STHT curve which again overlays with the truth values between 0.2 [s] and 0.8 [s].

This example demonstrated that the STHT process is able to eliminate ringing artifacts in segments away from the signal edges. The more conventional DFT-based Hilbert Transform displayed the most significant error, with ringing that propagated throughout the response. While the FIR filter approach showed better results, it was less accurate than the STHT.

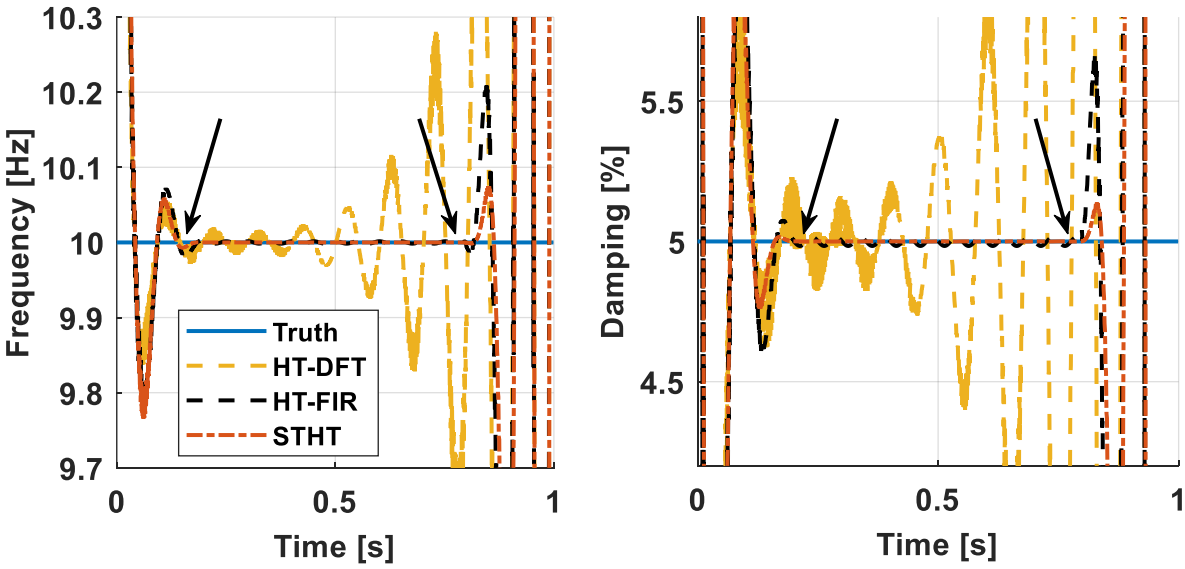

Figure 18: Natural Frequency (left) and Damping Ratio (right) of the linear SDOF system with estimates derived from the derivatives of the amplitude and phase results from the STHT and the Hilbert Transform computed via the DFT and an FIR filter. The STHT result closely overlays with the true values in the central region of the response between the arrows. The HT-FIR curve shows slight ringing in this region, while the HT-DFT contains significantly more error throughout the response.

2.4.2. Multicomponent Signal

In this section, a multicomponent signal is presented to demonstrate the adaptability of the STHT process to more complicated scenarios. The signal is composed of a frequency sweep and its second and third harmonics, each decaying from different initial amplitudes. An expression for this signal is given in Eq. (57), where $x_1(t), x_2(t)$ & $x_3(t)$ are the individual oscillatory components, $\sigma(t)$ is additive noise,

and $x(t)$ is the complete signal formed by summing these terms. The frequency of the oscillations is then $50 + t + 0.9t^2$ and its second and third harmonics.

$$\begin{aligned} x(t) &= x_1(t) + x_2(t) + x_3(t) + \sigma(t) \\ x_1(t) &= 1.00e^{-0.1t}\cos\left(2\pi(50t + 0.5t^2 + 0.3t^3)\right) \\ x_2(t) &= 0.75e^{-0.1t}\cos\left(4\pi(50t + 0.5t^2 + 0.3t^3)\right) \\ x_3(t) &= 0.50e^{-0.1t}\cos\left(6\pi(50t + 0.5t^2 + 0.3t^3)\right) \end{aligned} \tag{57}$$

For this test case, $x(t)$ was rendered over a time span of 10 [s] at a sample rate of 5000 [Hz]. In the initial analysis of this signal the noise is set to zero, but the impact of the noise will be investigated at the end of the section. The resultant time series is shown on the left side of Figure 19. This signal displays highly asymmetric behavior about zero due to the interference between the harmonic components as shown in detail in the zoomed plot window. To extract and characterize each component with the STHT, the STFT must first be computed. This was done with a segment size of 0.5 [s], an overlap of 75%, and the 4-term Blackman-Harris analysis window. To mitigate end effects, mirrored copies of the signal were appended to the start and end. A plot of the magnitude of the resulting frequency coefficients, called a spectrogram, is shown on the right side of Figure 19, where darker color denotes low amplitude and brighter is higher. The components of the signal are visible as the three high amplitude ridges that sweep to higher frequencies as time progresses.

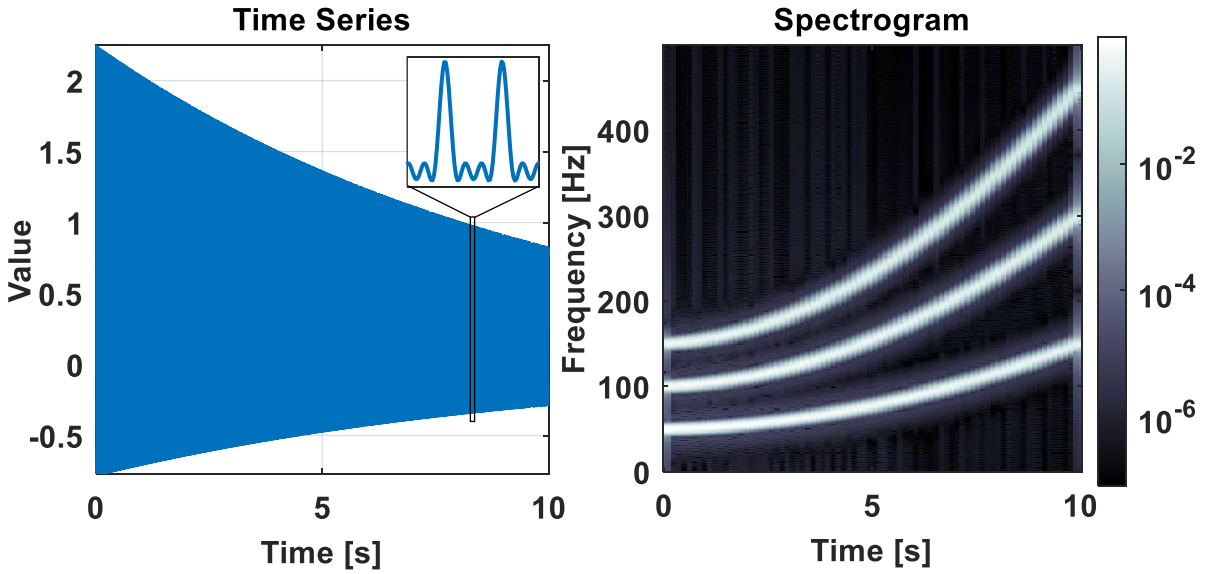

Figure 19: The multicomponent signal in the time domain (left) and the time-frequency domain (right). As the signal is composed of harmonics, interference between each result in the highly asymmetric profile in the time series. The STFT spectrogram representation of the signal shows the frequency sweep behavior of each harmonic as a high amplitude ridge that progresses through time.

To separate the individual components of the signal and form estimates of the frequency and amplitude of each, time-frequency masking is used to selectively remove portions of the STFT result. Each harmonic can be individually assembled in the time domain by constructing mask matrices that only retain frequency coefficients that are part of the associated ridge. In this case, the masks are defined to retain coefficients ± 20 [Hz] from the frequency of that component. This is shown on the left side of Figure 20, in which the colored lines enclose the frequency content that is retained in each mask matrix. These regions functionally act as an ideal bandpass filter with a pass band of 40 [Hz] that shifts its center frequency in each segment to properly track the ridge. The time series for each masked region is then assembled with the WOLA method and are shown in the plot on the right side of Figure 20. The color of each signal corresponds to the color of the masked region in the spectrogram

view. In the zoomed in plot, each assembled harmonic component is shown to be cleanly separated.

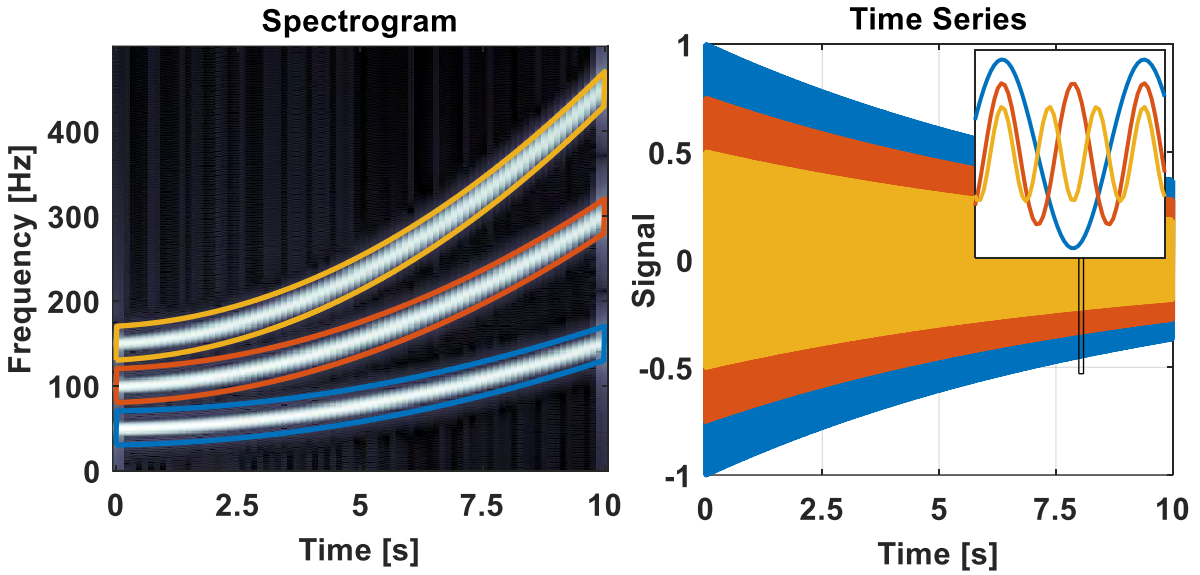

Figure 20: (Left) The portions of the STFT retained by each mask are denoted by the colored bounding boxes. (Right) The associated assembled time domain signals. Each harmonic component is visible in the zoomed plot.

As each of the assembled time series are in a complex-valued form, the magnitude can be taken as the amplitude and the frequency is the central difference derivative of the unwrapped phase angle. The results of this process for each extracted component are shown in Figure 21, with the amplitude on the left and the frequency on the right. The truth values are given as solid black lines, while the estimated curves are given as dashed lines in the associated color of the assembled time series shown in the right side of Figure 20. There is no discernable error between the estimated and the true values in the plot, as the amplitude and frequency errors are both on the order of 10^{-4} [%].

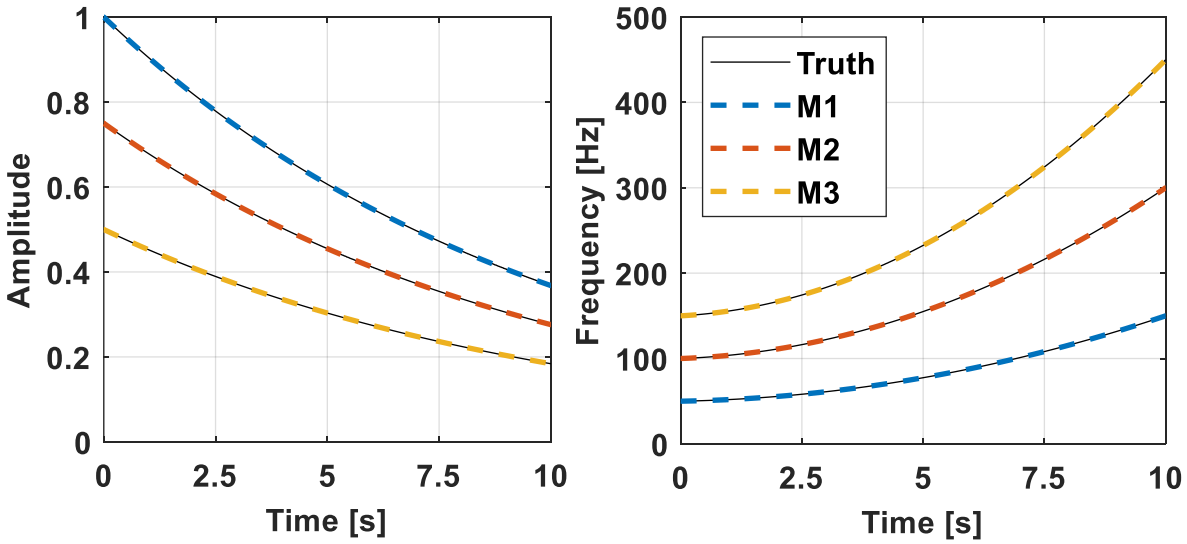

Figure 21: The amplitude (left) and frequency (right) of the components extracted via the STHT. The truth values are shown as black solid lines, and the estimated values are dashed line with colors respective of the mask regions.

Analyzing this signal with existing decomposition methods produces much more varied results. The approaches employed here are EMD [69] and VMD [90], where the specific coded implementations utilized are those supplied within MATLAB 2021a. Both algorithms yield a set of oscillatory components called intrinsic mode functions (IMF) which are extracted from the signal. The DFT-based Hilbert Transform is then applied to each IMF to form estimates of the amplitude and phase. Applying EMD to the multicomponent signal in this test case generates two significant IMFs with the resultant amplitude and frequency of each shown in Figure 22. EMD is notoriously susceptible to mode mixing, which is when an IMF contains multiple oscillations at unique frequencies [112]. The first IMF is heavily affected by this phenomenon, where it is primarily composed of the second and third harmonics leading to the large variations in amplitude and frequency seen in Figure 22 and in

detail in the zoomed plots. While the second IMF is largely the first hormonic it also contains frequency content from the second and third, such that the amplitude is offset from any true value and the frequency is oscillating about the first harmonic. In applying VMD, three IMFs were formed that produced the amplitude and frequency results shown in Figure 23. These also exhibit mode mixing as evidenced by the high variability in each curve. Though the IMFs are somewhat accurate near the midpoint of the signal, each switch to a different harmonic in the latter half and show significant ringing artifacts at the start and end. While EMD and VMD are unable to correctly decompose the harmonics, the STHT accurately extracted and characterized all components of the signal.

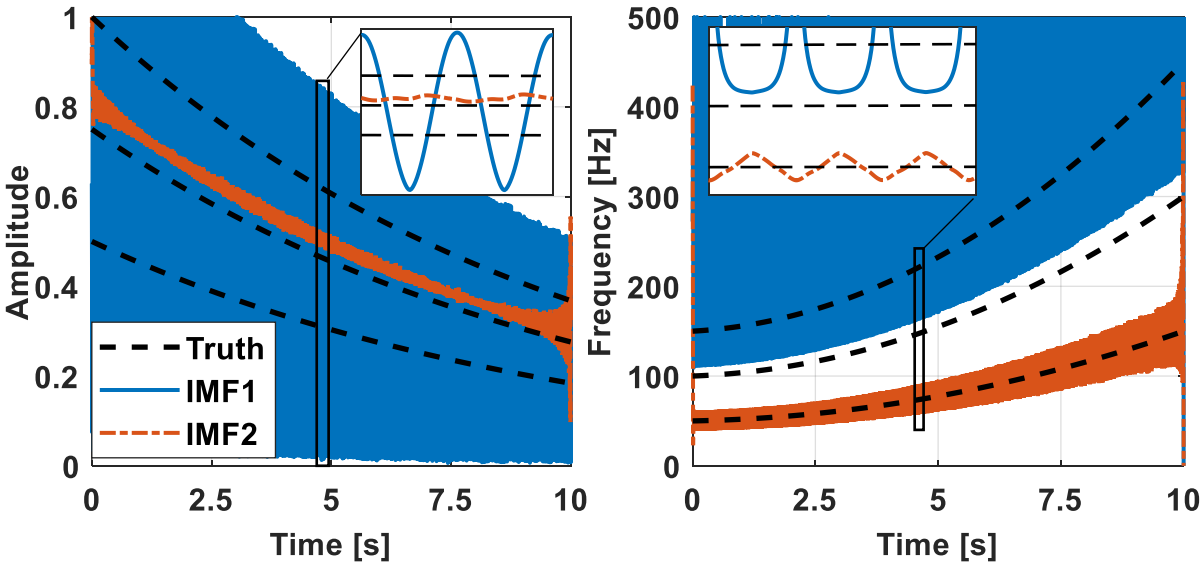

Figure 22: Amplitude (left) and Frequency (right) estimated from the IMFs generated by EMD. Both IMFs contain a mix of the three harmonics in the signal, leading to the inaccuracies in the results.

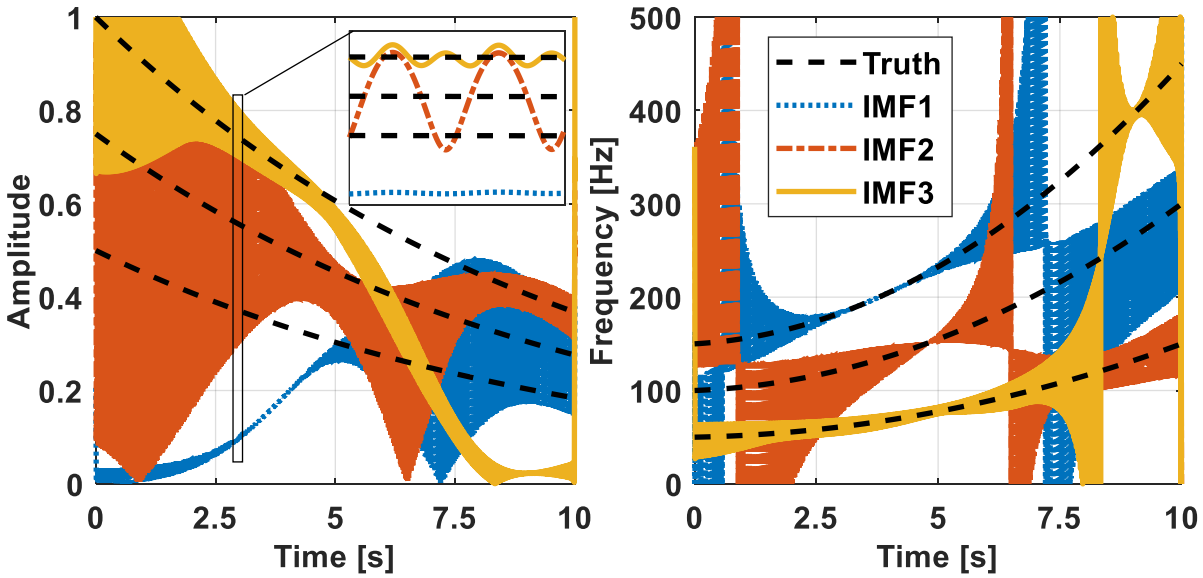

Figure 23: Amplitude (left) and Frequency (right) estimated from the IMFs generated by VMD. While somewhat accurate in the middle of the signal, each IMF has large errors at the start and also switch to different harmonics in the latter half.

To demonstrate the effect of noise on the STHT results, the $\sigma(t)$ term in Eq. (57) is now defined to apply Gaussian white noise to the multicomponent signal. The strength of the noise is set such that the signal-to-noise ratio (SNR) of $x(t)$ to $\sigma(t)$ is 6 [dB]. For each component $x_1(t), x_2(t)$ & $x_3(t)$ relative to $\sigma(t)$, the SNRs are 4.22 [dB], 2.84 [dB], and 1.51 [dB], respectively. The STFT of the signal is computed with the same parameters as used previously for the clean case. With the noise floor of the signal raised, the low amplitude sides of the time-frequency ridges are corrupted. To reduce the amount of noise that is included in the signal assembly process, the width of the masking regions is reduced to ± 6 [Hz] on either side of the ridge peak.

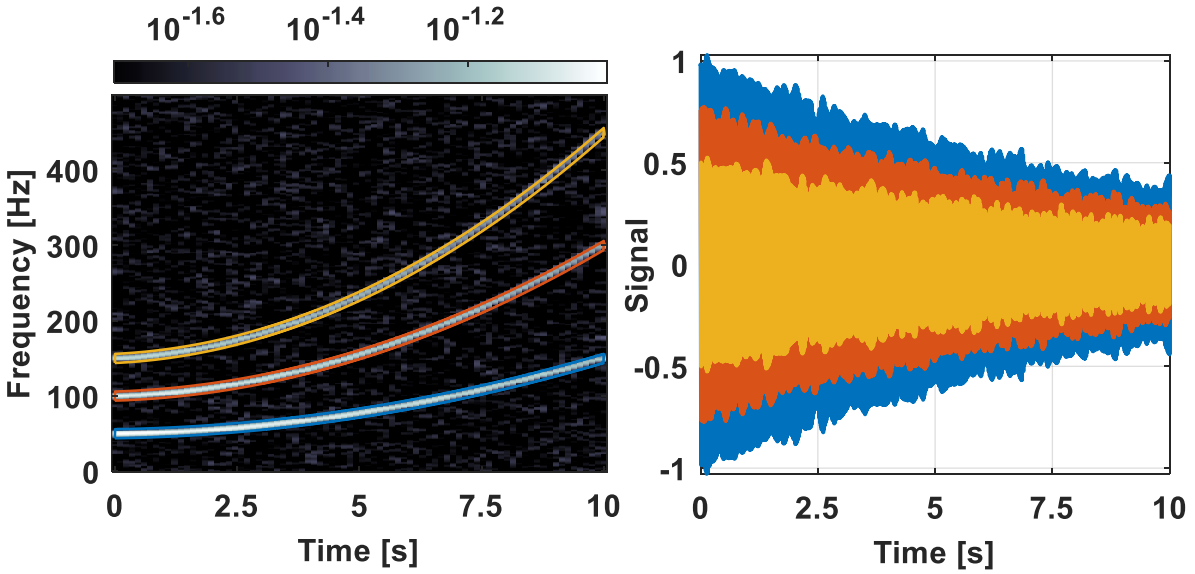

Figure 24: (Left) The spectrogram of the noisy multicomponent signal and the bounds of each masking region for the STHT. (Right) The assembled time domain signal from each masked region of the spectrogram.

The spectrogram of the noisy signal along with the narrower masking regions are shown on the left side of Figure 24, as well as the associated assembled time domain signal from each mask in the right-side plot. Compared to their noise-free counterparts, $x(t), x_1(t), x_2(t)$ & $x_3(t)$, the extracted signals have SNRs of 23.22 [dB], 26.14 [dB], 23.17 [dB], and 18.84 [dB], respectively. The estimated amplitude and frequency of each component from the STHT is given in Figure 25. In order from the lowest to highest frequency component, the mean errors for amplitude are 3.1%, 4.2%, and 6.9%, and for frequency are 0.25%, 0.29%, and 0.41%.

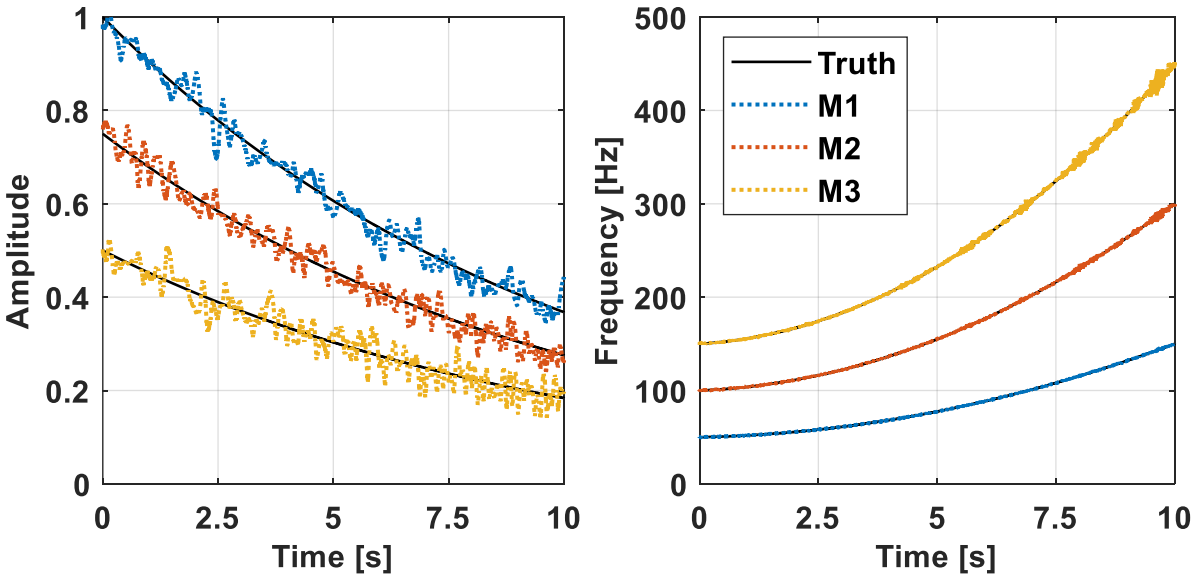

Figure 25: The amplitude (left) and frequency (right) of the noisy signal components as estimated by the STHT.

In this numerical example, a multicomponent signal was used to demonstrate the adaptability of the STHT to complicated signals that are difficult to process with existing methods. Based on the spectrogram representation of the signal, bounding regions for time-frequency masking can be defined about each prominent ridge. By individually inverting the frequency content in each masked region back to the time domain to be assembled, the harmonic components of the signal can be extracted and characterized in terms of their amplitude and frequency. Meanwhile, for this signal the EMD and VMD methods of mode decomposition produced IMFs that exhibit significant mode mixing such that the amplitude and frequency from the Hilbert Transform vary wildly. And lastly, it was shown that the STHT can generate reasonable amplitude and frequency estimates even in the presence of strong noise.

2.5. Experimental Demonstration

This section details an experimental case study in which the STHT is utilized to form amplitude dependent natural frequency and damping ratio curves from a measured nonlinear free decay signal. The response was generated by the nonlinear benchmark structure known as the S4 Beam, which consists of two nominally identical stainless-steel beams that are bolted together at raised surfaces on each end. While initial investigations into this design were done by Singh et. al in [18], the specific experimental setup presented here is from more recent work done by Wall in [19]. Images of the structure and instrumentation are shown in Figure 26. The beams are suspended from bungees to approximate free boundary conditions and the response is recorded by an array of accelerometers positioned at the ends and middle of the beams. The data analyzed in this section is from a modal hammer impact performed in the center of the beam, marked by the red dot and arrow in Figure 26, with the drive point response measured by an accelerometer positioned on the direct opposite side of the beam.

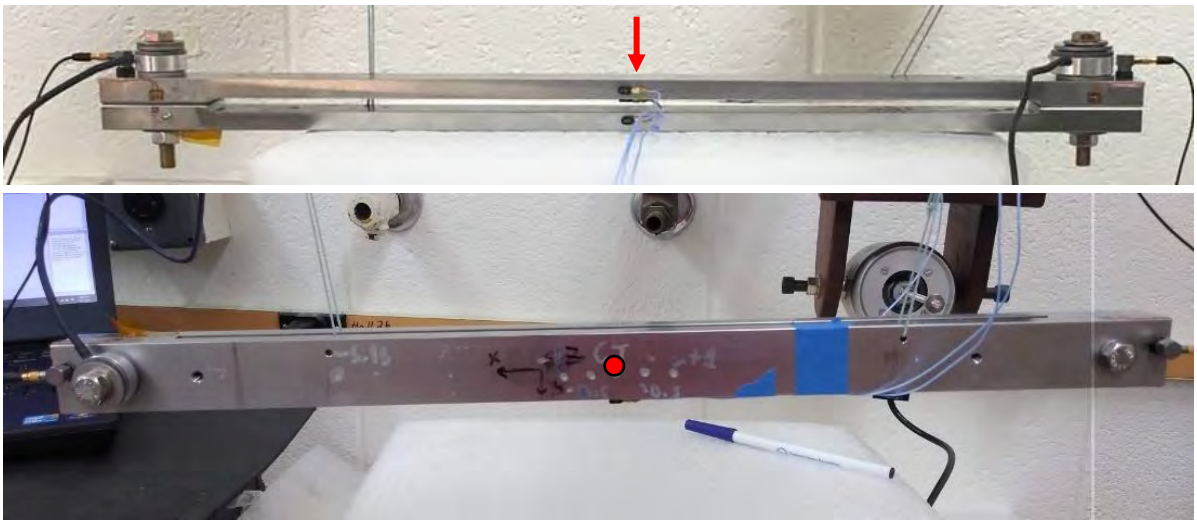

Figure 26: (Top) The S4 Beam structure, composed of two nominally identical steel beams joined by bolts at the raised ends. (Bottom) The experimental setup to conduct the hammer tests in which the beam is instrumented with accelerometers and suspended from bungees. For the data utilized here, the impact was applied at the location marked in red by the dot and arrow, with the response recorded by an accelerometer on the opposite side of the beam. [19]

The measured drive point acceleration is shown in the top left of Figure 27 and displays typical impulse excitation behavior, going from a rest state to a sudden high-amplitude response that exponentially decays as the energy from the impact dissipates from the system. A DFT of this signal is plotted in the bottom left, in which the prominent peaks show that there are six modes that were significantly excited by the impact. At the start of the transient, the most dominant components of the response are the pair of modes near 1300 [Hz] and 1500 [Hz]. This is detailed in the upper zoom window of the time series, where the beating pattern due to interference between the close modes is visible. As the ringdown progresses the higher order resonances quickly decay, and the response appears to be primarily composed of the first elastic mode near 280 [Hz] as shown in the lower zoom window.

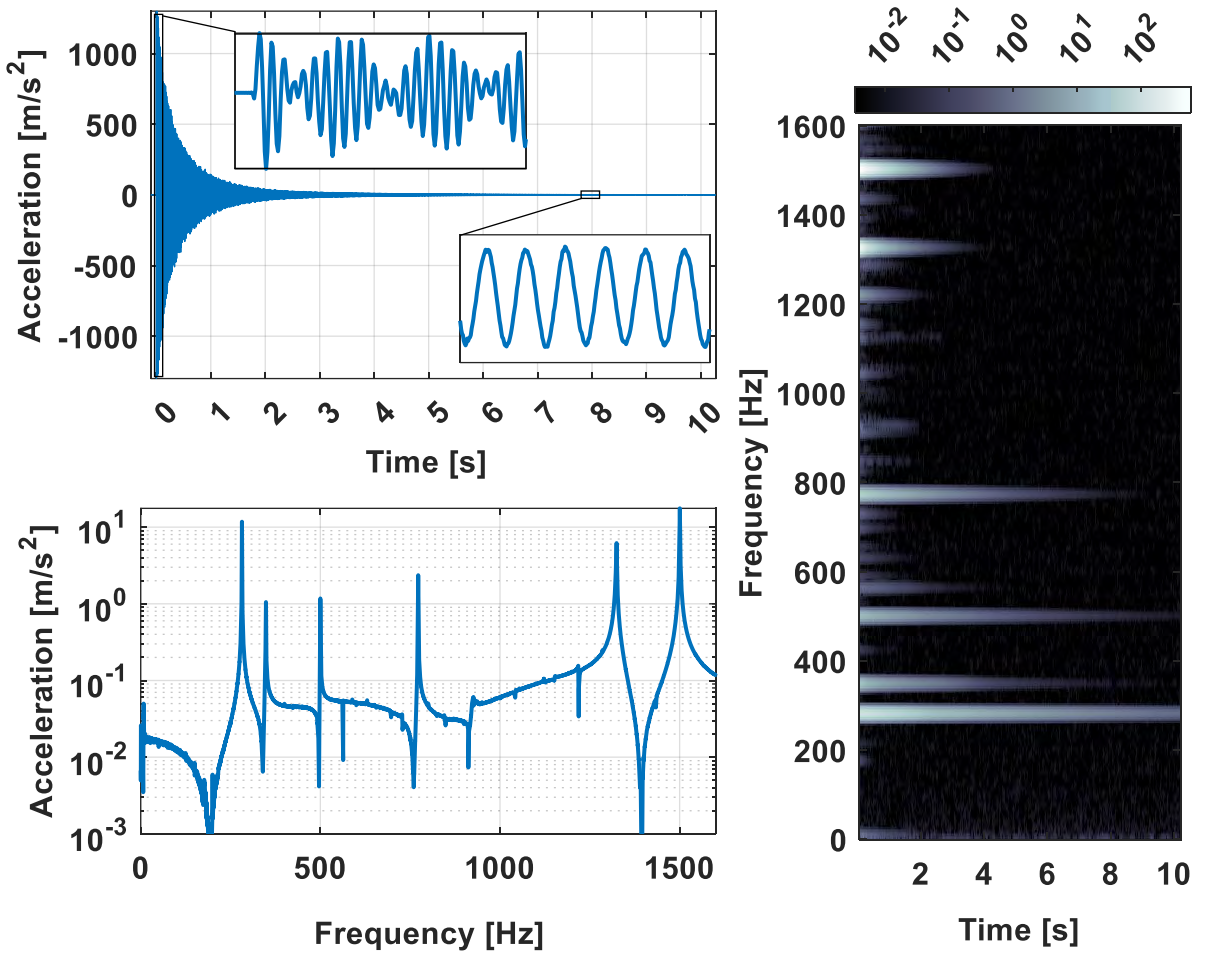

Figure 27: The recorded acceleration of the S4 beam near the drive point in the time domain (Left Top) and frequency domain (Left Bottom). (Right) Spectrogram showing the STFT of the measured S4 Beam response. Each modal response is denoted by a ridge that slowly fades as time progresses.

To utilize the STHT to characterize the modal responses in this signal, the STFT must first be computed. This was done with a segment size of 0.15 [s], an overlap of 75%, and the 4-term Blackman-Harris analysis window. The first segment was set to begin at the first peak of the response to avoid including the impulse and the ringing artifacts it would generate. These parameters produced the STFT shown as the spectrogram on the right side of Figure 27. Each modal response seen as a

frequency peak in the DFT is now represented by a time-frequency ridge that decays in amplitude as time progresses.

Extracting individual modal responses from the free decay is accomplished by constructing a time-frequency mask that retains a single ridge while removing all other frequency content. For this case study the masks were defined by manually selecting points to set the upper and lower extents of the passband region. The masking region selected for the first mode is shown on the left side of Figure 28, where the red dots mark the manually selected points and the blue lines are linearly interpolated between each to enclose the region. With this time-frequency mask, the time series of the first mode is then assembled with the WOLA synthesis method. The synthesis window is a trigonometric blending function that is 50% of the segment length to remove any residual artifacts at the edges. While this discards the outer 25% of each segment, the 75% overlap leads to a contiguous result that is smoothly blended and complete. The real part of the assembled signal along with the response and the residual between the two are shown in the time and frequency domains on the right side of Figure 28. The zoom plots of the time series show that while the first mode is subdominant early in the response, its lower frequency leads to a slower decay rate such that it is the only active mode late in the response. In the frequency domain, the STHT is shown to have very accurately extracted that first mode as the result exhibits no other peaks and the residual contains no trace of the first peak.

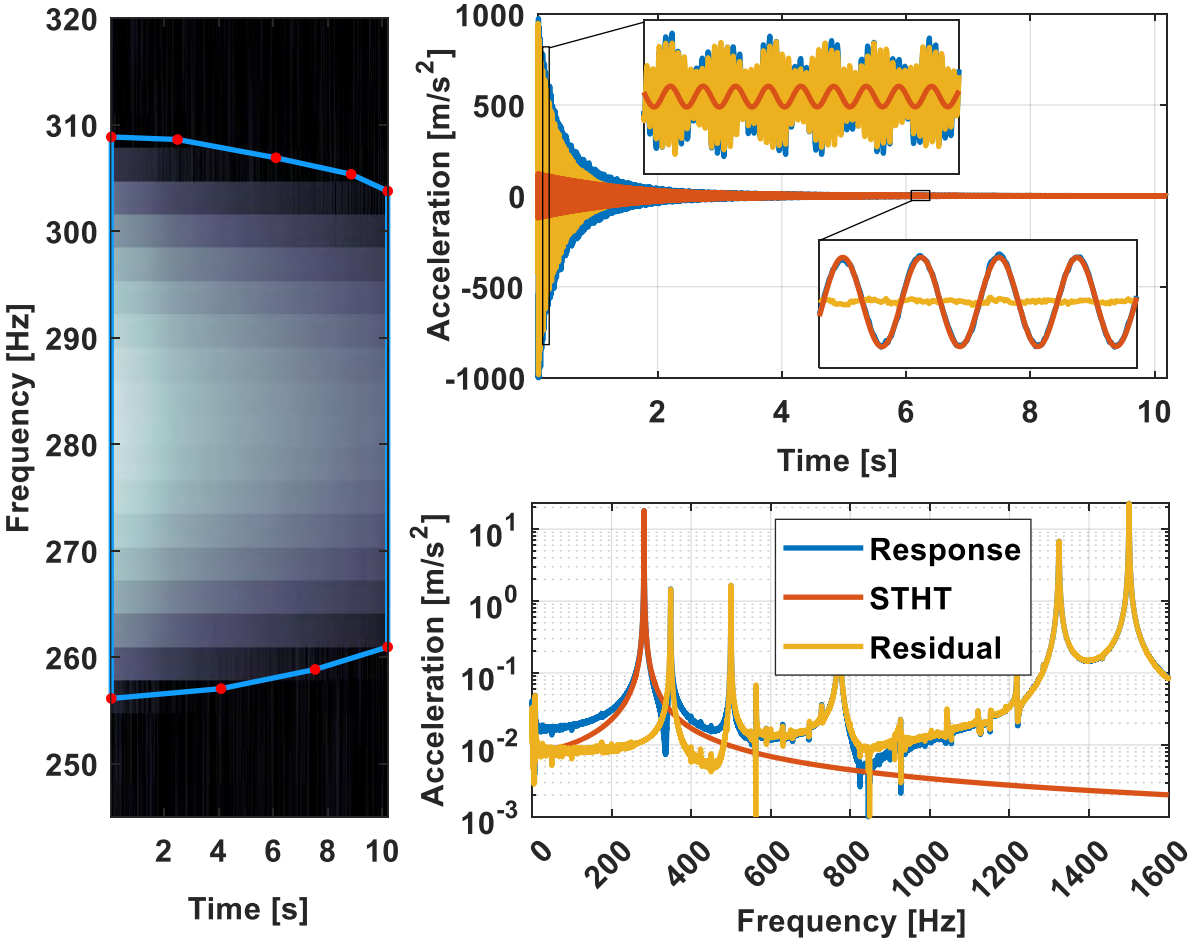

Figure 28: (Left) The section of the spectrogram about the first mode and the selected lines defining the masking region about the ridge. (Right Top) The oscillation assembled from the masked region along with the original signal and the residual difference between them. (Right Bottom) The DFT of the signals shows that the first mode is accurately extracted.

Extracting the other five modal responses is accomplished by following the same procedure of defining an appropriate masking region about each prominent ridge in the spectrogram and synthesizing the associated oscillation in the time domain. The result of this process is displayed in Figure 29, in which the DFT of the real part of each assembled modal response is plotted as a solid line. These perfectly overlay with the peaks present in the measured response signal, which is the dash-dotted line. The black dotted line represents what remains after removing the six

modal components from the response signal. This shows that each mode is completely removed from the response, as the residual is primarily composed of low amplitude peaks from weakly excited modes that were active for a short duration just after the impact.

The VMD and EMD methods can also be applied to the measured response to compare the mode decomposition performance of each to the result produced by the time-frequency masking in the STHT. For both methods, the implementations of each algorithm that is supplied with MATLAB 2021a were again utilized. The results of these mode decompositions are given in Figure 30, in which the DFT of the IMFs from VMD are in the left plot and those from EMD are on the right. While VMD cleanly separates most of the modal peaks into isolated IMFs, it is unable to do so for the second mode. This peak is primarily included with the IMF that also contains the first mode but is present in all IMFs to some extent as shown in the zoomed window. Applying EMD to the response signal produces two IMFs that exhibit significant mode mixing. While the first IMF primarily captures all frequency content above 1000 [Hz], it contains portions of the lower order modes as well. It also generates significant noise below 1000 [Hz] that then corrupts the modes in the second IMF.

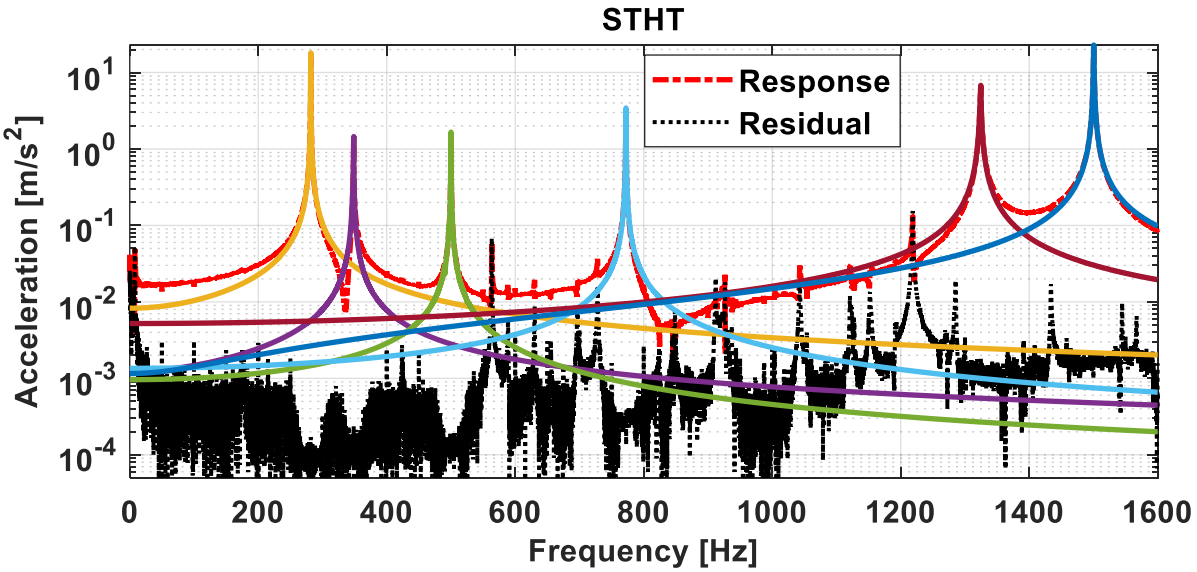

Figure 29: Result of extracting the six significant modal responses. The residual is primarily composed of weakly excited modes that quickly decay after the impact.

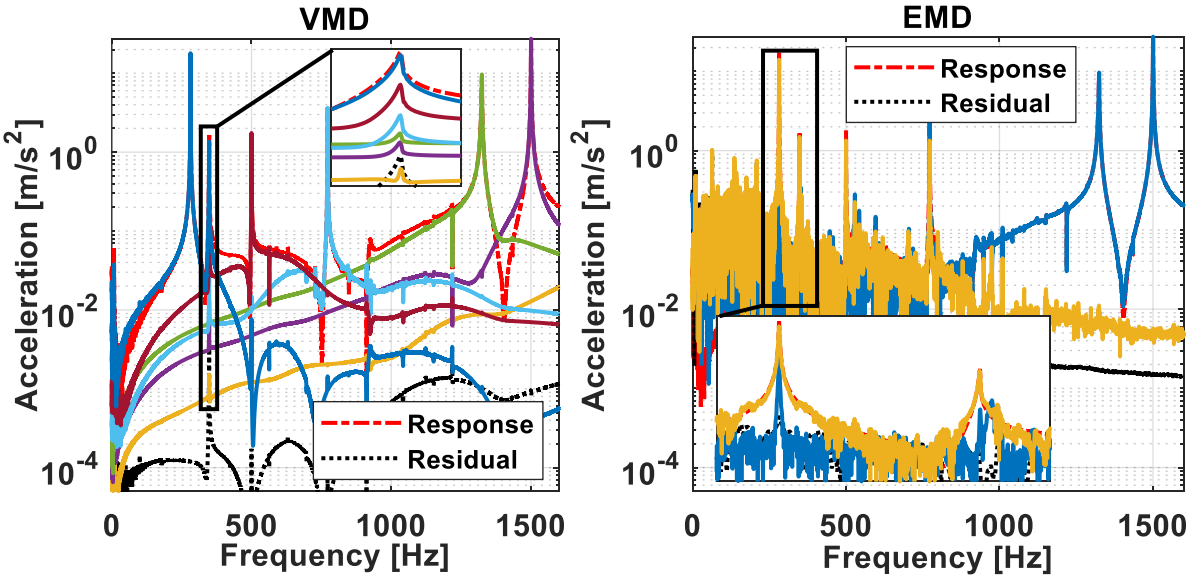

Figure 30: The result of applying VMD (Left) and EMD (Right) to the response. VMD generates IMFs that capture the modal responses fairly accurately, except for the second mode which is primarily mixed with the first mode. EMD produces two IMFs that do not clearly contain any individual mode while also corrupting the signal with additional noise.

The FREEVIB equations detailed in Eq. (25) and Eq. (26) are now used to determine amplitude dependent natural frequency and damping ratios from the identified modal responses. For the STHT results, each modal response is already in

a complex-valued form such that the amplitude and phase are simply their magnitude and angle. The DFT-based Hilbert Transform is applied to the VMD and EMD results to determine amplitude and phase results for each IMF. To attain a cleaner FREEVIB output, measurement noise is smoothed from the amplitude and phase curves prior to computing their derivatives via central differences. The smoothing is applied identically to all curves and is performed by a third degree Savitzky-Golay filter [113] with a span of 0.5 [s]. This process produced the curves shown in Figure 31, in which the left plots are natural frequency, the right are damping ratio, the top is for the first mode, and the bottom are for the second mode. The STHT cleanly separated each mode such that distinct results are generated for each plot. These curves display clear nonlinearity in the form of increasing and decreasing natural frequency and damping ratio behavior with respect to the amplitude of the modal response. For VMD and EMD, the results from the IMF that most closely resembled the first mode are shown in those plots. As that IMF also contains the second mode due to mode mixing, neither approach produces a curve that can be placed into the bottom plots. Comparing the curves for the first mode, the VMD results appear to follow the same underlying trend as the STHT curve but exhibit significant noise due to the mode mixing with the second mode. Even with smoothing the EMD result is incredibly noisy and completely incomprehensible.

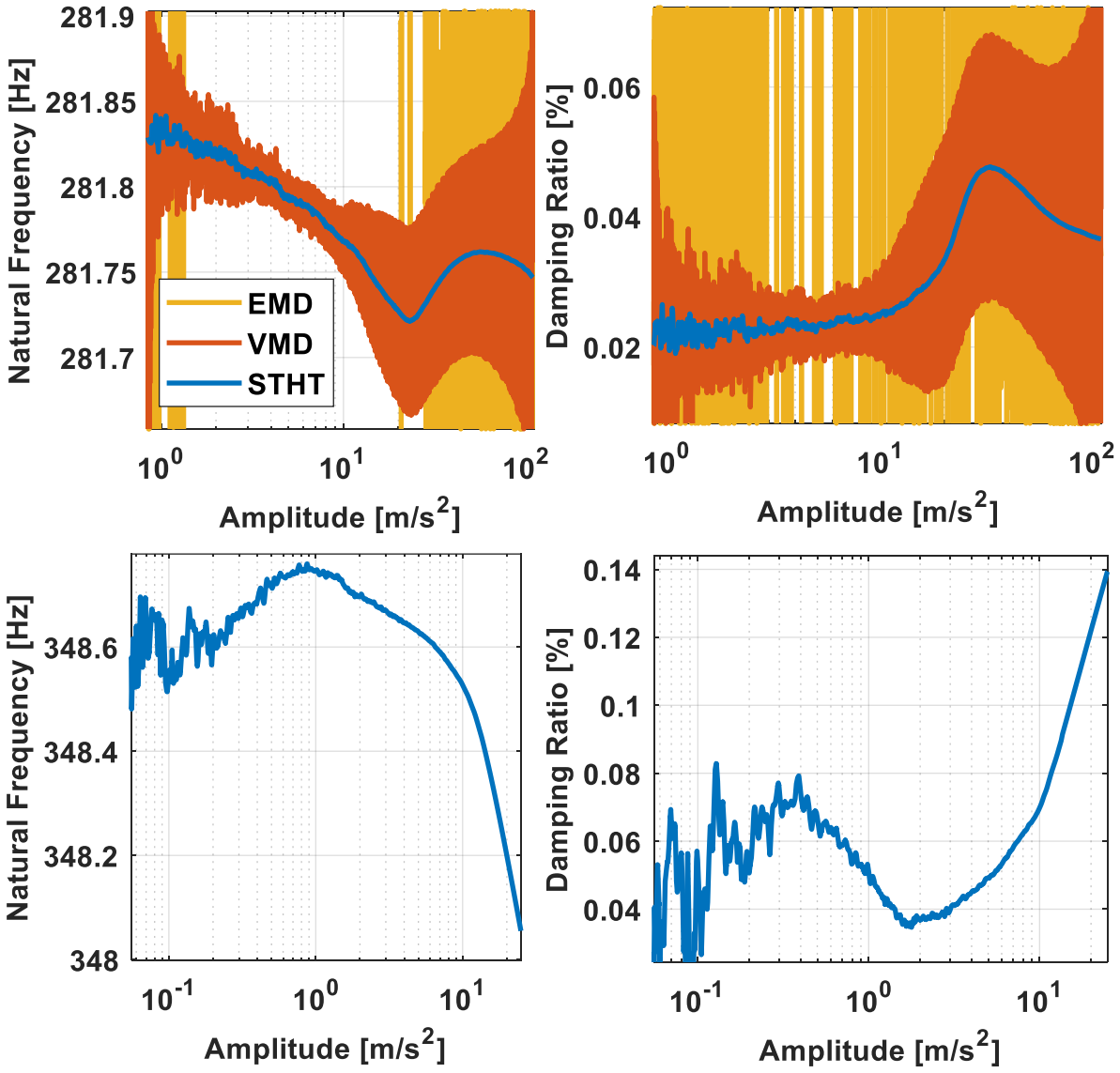

Figure 31: Amplitude dependent natural frequency (Left) and damping ratio (Right) for the first mode (Top) and the second mode (Bottom). The STHT is able to cleanly separate and individually extract both modes and produce the curves characterizing the nonlinearity in each. Both VMD and EMD produce IMFs with mode mixing that either approximately capture the nonlinear behavior with noise, or not at all.

2.6. Conclusion

This paper introduced the Short-time Hilbert Transform as an improved approach to estimating the amplitude and phase of a nonstationary oscillation. The conventional Hilbert Transform is very popular due to its conceptual simplicity and computational efficiency. However, the results it produces very commonly include ringing artifacts in the form of end effect errors. The thorough review presented here of the theoretical background for the Hilbert Transform demonstrated that the ringing artifacts are a consequence of frequency domain leakage from the Fourier Transform being incorrectly phase shifted by the Hilbert Transform transfer function. Therefore, the ringing artifacts can be suppressed by reducing the amount of leakage exhibited by the signal. That was accomplished with the STHT in this work by instead utilizing the Short-time Fourier Transform for computing the Hilbert Transform. The STFT offers an existing framework for dividing a signal into smaller windowed segments, performing modifications to the spectral content, and then reassembling the altered time signal. In this process, the Hilbert Transform is applied in the modify-spectral-content step as a simple phase shifting transfer function as normal. Segmenting the signal localizes ringing artifacts to only impact the segments that contain the cause of the leakage. Any segment that only contains smoothly varying oscillations exhibits negligible end effects as the applied window minimizes leakage caused by nonperiodic edge conditions. Furthermore, this process allows for additional modifications to be made to the spectral content in each segment, such as performing mode decomposition via time-frequency masking. The mask can be strategically

constructed to essentially function as a time-dependent ideal bandpass filter with a passband that is separately defined for each segment. By setting the mask to only retain frequency information related to a specific component, the associated oscillation can be extracted from the signal and synthesized in the time domain.

This process was first demonstrated on a linear free decay to show the improved accuracy of the results compared to the standard Hilbert Transform. The second numerical case study was a complicated multicomponent signal that contained three harmonics that swept through a wide frequency range while also exhibiting decaying amplitudes. While the EMD and VMD methods failed to separate the components, the STHT accurately extracted each harmonic from the signal. A strong Gaussian white noise was then added to the signal. By narrowing the passband of the time-frequency masks, reasonable results for each harmonic were able to be recovered. The last case study presented experimental data in the form of a measured nonlinear free decay from a benchmark structure. By computing a STFT of the response signal, it could be visualized as a spectrogram. This allows for the time-frequency mask regions to be easily defined by selecting points on the plot and manually constructing the bounds of the mask matrix. With this process, the amplitude and phase of the individual modal responses could be synthesized from the complete measured response. These were then used to generate quasi-linear natural frequency and damping ratio curves that characterize the nonlinear behavior exhibited by the structure.

The STHT is a simple generalization of the standard Hilbert Transform process that exchanges use of the Fourier Transform with the Short-time Fourier Transform. Doing so provides a flexible framework that can be easily adapted to accurately characterize highly nonlinear signals that could not be processed with more conventional approaches such as bandpass filtering and the Hilbert Transform. In its simplest form, the STHT is a variation of the usual inverse STFT process, where only the positive frequency content is reconstructed in the time domain instead of the complete spectrum. Estimates of amplitude and phase generated by the STHT are only locally impacted by ringing artifacts instead of globally. While this does not entirely remove ringing artifacts, their effect on the signal is suppressed and results are improved. An avenue for future work would be to investigate sophisticated signal extension processes. These are methods that analyze the content at the signal edge and generate an artificial continuation that smoothly extend the signal such that the discontinuity and the associated end effects are eliminated. Alternatively, the optimization process explored in the next chapter could also be used to overcome the lingering issues with ringing artifacts present in the STHT.

3. Characterizing Oscillations with Nonlinear Optimization

3.1. Introduction

In this chapter, a new NLSID method is proposed that circumvents issues present in current methods for determining instantaneous system parameters. Instead of estimating the complex form of the signal with the Hilbert Transform, nonlinear optimization is used to directly curve fit representations of amplitude and phase to sampled oscillations. This is inspired by work done by Goyder and Lancereau [58], in which they directly curve fit a linear free response model form to segments of a nonlinear signal. By incrementing this window down the time series, a piece-wise linear representation of the nonlinear behavior of the natural frequency and damping can be constructed. Additionally, several other works describe a similar procedure. Early work by Smith in [114] was developed during the Space Shuttle dynamic testing program for identifying linear modal parameters from curve fitting measured transients. Also, Hallauer in [115] uses a MATLAB nonlinear least squares algorithm to fit measured response data with viscous damping and Coulomb friction parameters.

The method proposed here extends these processes in several ways. The representations for amplitude and phase are generalized to polynomial expressions of arbitrary order, allowing for nonstationary oscillations to be directly represented. To maximize the numerical efficiency of this approach, the exact form of the gradient to be used in the nonlinear optimization is derived. In performing the nonlinear least squares solution, the previous works minimized the residual between time domain

vectors. While first considering that approach, this work demonstrates that the data can be transformed into the frequency domain, and the error minimization can be computed on information within a specific frequency band. This can be leveraged to curve fit individual signals from data containing multiple prominent oscillations, without the need for filtering or other pre-processing techniques. After completing the optimization, the resultant fit can be used to represent the integrals and/or derivatives of the oscillation. And finally, an implementation of FREEVIB/ FORCEVIB is given in terms of the curve fitting results, producing natural frequency and damping ratio during free and forced response, respectively. Following these derivations, several case studies are presented to demonstrate the various aspects of this NLSID method on numerical and experimental data.

3.2. Theoretical Derivations

3.2.1. Amplitude and Phase Curve Fitting

In the most general sense, the proposed curve fitting method seeks to form a best-fit approximation, $\tilde{y}(t)$, to some existing oscillating signal, $y(t)$.

$$\tilde{y}(t) \cong y(t) \tag{58}$$

An appropriate expression for $\tilde{y}(t)$ is a general nonstationary oscillation with time dependent terms for amplitude, $\beta(t)$, and phase, $\alpha(t)$.

$$\tilde{y}(t) = \mathrm{e}^{\beta(t)} \cos\big(\alpha(t)\big) \tag{59}$$

Convenient forms for $\alpha(t)$ and $\beta(t)$ are polynomials of time with arbitrary order.

$$\alpha(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \cdots \tag{60}$$

$$\beta(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \cdots \tag{61}$$

The function to be used in the curve fitting process is then given below in Eq. (62), where the optimization variables are the $\alpha(t)$ and $\beta(t)$ polynomial coefficients.

$$\tilde{y}(t) = \mathrm{e}^{\beta_0+\beta_1 t+\beta_2 t^2+\cdots} \cos(\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots) \tag{62}$$

While this renders a model form that could be directly implemented in a nonlinear curve fitting algorithm, the problem can be simplified by restructuring Eq. (62) such that the lowest order polynomial terms, α_0 and β_0, can be determined outside of the optimization. This is done by first separating the exponential into the product of two terms, $\mathrm{e}^{\beta_0}\mathrm{e}^{\beta_1 t+\beta_2 t^2+\cdots}$, where e^{β_0} is the amplitude at $t = 0$, and $\mathrm{e}^{\beta_1 t+\beta_2 t^2+\cdots}$ describes the shape of the amplitude envelope. Additionally, α_0, which is the phase offset of the oscillation, is removed from the cosine with the Angle-Sum Identity:

$$\cos(\theta + \phi) = \cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi) \tag{63}$$

This transforms the single cosine in Eq. (62) into an equivalent form that is the superposition of a sine and cosine with scaling coefficients based on α_0.

$$\begin{aligned}\tilde{y}(t) = \mathrm{e}^{\beta_1 t+\beta_2 t^2+\cdots}\big(&\mathrm{e}^{\beta_0}\cos(\alpha_0)\cos(\alpha_1 t + \alpha_2 t^2 + \cdots) \\ &- \mathrm{e}^{\beta_0}\sin(\alpha_0)\sin(\alpha_1 t + \alpha_2 t^2 + \cdots)\big)\end{aligned} \tag{64}$$

Without any loss of generality, the sine and cosine coefficients may be condensed into the arbitrary constants A and B:

$$A = \mathrm{e}^{\beta_0}\cos(\alpha_0) \tag{65}$$

$$B = -\mathrm{e}^{\beta_0}\sin(\alpha_0) \tag{66}$$

The subsequent expression can be placed into a matrix form as in Eq. (67). While Eq. (62) is a nonlinear function of α_0 and β_0, these terms are equivalently represented linearly in Eq. (67) as A and B, reducing the number of optimization variables by two.

$$\tilde{y}(t) = \mathrm{e}^{\beta_1 t+\beta_2 t^2+\cdots}[\cos(\alpha_1 t + \alpha_2 t^2 + \cdots) \quad \sin(\alpha_1 t + \alpha_2 t^2 + \cdots)] \begin{bmatrix} A \\ B \end{bmatrix} \tag{67}$$

For ease of notation, gather the remaining polynomial coefficients into vector $\boldsymbol{\kappa}$:

$$\boldsymbol{\kappa} = [\beta_1 \quad \beta_2 \quad \cdots \quad \alpha_1 \quad \alpha_2 \quad \cdots] \tag{68}$$

Additionally, it is beneficial to represent the time-dependent terms in Eq. (67) as:

$$\mathbf{h}(\boldsymbol{\kappa}, t) = e^{\beta_1 t+\beta_2 t^2+\cdots}[\cos(\alpha_1 t + \alpha_2 t^2 + \cdots) \quad \sin(\alpha_1 t + \alpha_2 t^2 + \cdots)] \tag{69}$$

This results in the equation for the synthesized oscillation to be represented as:

$$\tilde{y}(t) = \mathbf{h}(\boldsymbol{\kappa}, t) \begin{bmatrix} A \\ B \end{bmatrix} \tag{70}$$

In practice, this expression for $\tilde{y}(t)$ will be used to fit data that has been sampled at N_t discrete points in time. Therefore, the curve fit to the current set of data points is given by Eq. (71), where $\mathbf{y}$ is a vector of recorded values with length N_t, and $\tilde{\mathbf{y}}$ is the best-fit approximation to those values.

$$\tilde{\mathbf{y}} \cong \mathbf{y} \tag{71}$$

Subsequently, by discretizing Eq. (69) at the points in time associated with the values in $\mathbf{y}$, $\mathbf{h}(\boldsymbol{\kappa}, t)$ becomes a matrix, $\mathbf{H}(\boldsymbol{\kappa})$, with dimensions $[N_\mathrm{t} \times 2]$.

$$\tilde{\mathbf{y}} = \mathbf{H}(\boldsymbol{\kappa}) \begin{bmatrix} A \\ B \end{bmatrix} \tag{72}$$

A and B are now determined through linear least squares between $\mathbf{H}(\boldsymbol{\kappa})$ and $\mathbf{y}$. This is shown mathematically in Eq. (73) in terms of a matrix pseudo-inverse, but it should

be emphasized that the linear least squares solution should be computed with whatever numerical method is most convenient and accurate.

$$\begin{bmatrix} A \\ B \end{bmatrix} = \mathbf{H}(\boldsymbol{\kappa})^{\dagger}\mathbf{y} \tag{73}$$

The expression for the curve fit to $\mathbf{y}$ is now given as:

$$\tilde{\mathbf{y}} = \mathbf{H}(\boldsymbol{\kappa})(\mathbf{H}(\boldsymbol{\kappa})^{\dagger}\mathbf{y}) \tag{74}$$

The optimal values of the polynomial coefficients in $\boldsymbol{\kappa}$, $\alpha_1, \alpha_2, \ldots$ and $\beta_1, \beta_2, \ldots$, are found by minimizing the residual between the original data, $\mathbf{y}$, and the expression for $\tilde{\mathbf{y}}$ given in Eq. (74). This can be performed by a nonlinear least squares algorithm, such as Trust-Region-Reflective [116], [117] or Levenberg-Marquardt [118], [119]. These algorithms are modified implementations of Newton's method, specifically structured to numerically solve nonlinear least squares problems. As such, the gradient of the objective function is required at each iteration. While this can be approximated with central difference, doing so decreases the algorithm efficiency as many more function evaluations must be performed. To avoid this unnecessary computational burden, the form of the gradient can be analytically derived by calculating the partial derivative of $\tilde{\mathbf{y}}$ with respect to each optimization variable. Equation (75) depicts the general form of the partial derivative of Eq. (72) with respect to the terms in $\boldsymbol{\kappa}$. Note that, while A and B are constants, their value is dependent on the current iteration of the optimization variables in $\boldsymbol{\kappa}$, as given by Eq. (73).

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \boldsymbol{\kappa}} = \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}}\begin{bmatrix} A \\ B \end{bmatrix} + \mathbf{H}(\boldsymbol{\kappa})\frac{\partial \begin{bmatrix} A \\ B \end{bmatrix}}{\partial \boldsymbol{\kappa}} \tag{75}$$

The first term in Eq. (75) is the partial derivative of Eq. (69), evaluated at the current time points. This is given below in Eq. (76) with respect to the α_n coefficients, where n is the index of the coefficient, in which the sine and cosine terms are now their respective derivatives and a t^n term has come out of each. Similarly, the partial derivative with respect to the β_n terms is given in Eq. (77), where the t^n term has simply come down from the amplitude exponential.

$$\frac{\partial \mathbf{h}(\boldsymbol{\kappa}, t)}{\partial \alpha_n} = t^n e^{\beta_1 t + \beta_2 t^2 + \cdots} [-\sin(\alpha_1 t + \alpha_2 t^2 + \cdots) \quad \cos(\alpha_1 t + \alpha_2 t^2 + \cdots)] \tag{76}$$

$$\frac{\partial \mathbf{h}(\boldsymbol{\kappa}, t)}{\partial \beta_n} = t^n e^{\beta_1 t + \beta_2 t^2 + \cdots} [\cos(\alpha_1 t + \alpha_2 t^2 + \cdots) \quad \sin(\alpha_1 t + \alpha_2 t^2 + \cdots)] \tag{77}$$

The second partial term in Eq. (75), the partial derivative of A and B, is less straightforward as it requires the derivative of a linear least squares solution. This can be defined in terms of the pseudo-inverse notation used in Eq. (73):

$$\frac{\partial \begin{bmatrix} A \\ B \end{bmatrix}}{\partial \boldsymbol{\kappa}} = \frac{\partial \mathbf{H}(\boldsymbol{\kappa})^\dagger}{\partial \boldsymbol{\kappa}} \mathbf{y} \tag{78}$$

To derive the derivative of the matrix pseudo-inverse [120], begin with a property of the pseudo-inverse given in Eq. (79), where the product of the matrix and its pseudo-inverse is an identity. The partial derivative of this expression with respect to $\boldsymbol{\kappa}$ is determined in Eq. (80) by applying the chain rule. This can then be rearranged as in Eq. (81), providing an expression for the partial derivative of $\mathbf{H}(\boldsymbol{\kappa})^\dagger$.

$$\mathbf{H}(\boldsymbol{\kappa})^\dagger \mathbf{H}(\boldsymbol{\kappa}) = \mathbf{I} \tag{79}$$

$$\mathbf{H}(\boldsymbol{\kappa})^\dagger \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{H}(\boldsymbol{\kappa})^\dagger}{\partial \boldsymbol{\kappa}} \mathbf{H}(\boldsymbol{\kappa}) = 0 \tag{80}$$

$$\frac{\partial \mathbf{H}(\boldsymbol{\kappa})^{\dagger}}{\partial \boldsymbol{\kappa}} = -\mathbf{H}(\boldsymbol{\kappa})^{\dagger} \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \mathbf{H}(\boldsymbol{\kappa})^{\dagger} \tag{81}$$

With this result, Eq. (78) can be reformulated as:

$$\frac{\partial \begin{bmatrix} A \\ B \end{bmatrix}}{\partial \boldsymbol{\kappa}} = -\mathbf{H}(\boldsymbol{\kappa})^{\dagger} \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \mathbf{H}(\boldsymbol{\kappa})^{\dagger} \mathbf{y} = -\mathbf{H}(\boldsymbol{\kappa})^{\dagger} \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \begin{bmatrix} A \\ B \end{bmatrix} \tag{82}$$

The partial derivatives in Eq. (75) may then be represented concisely as in Eq. (83), where the individual partials with respect to each term in $\boldsymbol{\kappa}$ are evaluated in Eq. (76) for $\alpha_1, \alpha_2, \ldots$ and Eq. (77) for $\beta_1, \beta_2, \ldots$.

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \boldsymbol{\kappa}} = \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \begin{bmatrix} A \\ B \end{bmatrix} - \mathbf{H}(\boldsymbol{\kappa}) \left(\mathbf{H}(\boldsymbol{\kappa})^{\dagger} \frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \begin{bmatrix} A \\ B \end{bmatrix} \right) \tag{83}$$

The nonlinear least squares curve fit posed in Eq. (74), along with this expression for the gradient, can be used to directly synthesize a best-fit approximation of the amplitude and phase of a sampled oscillation. However, it is typically impractical to accurately fit an entire measurement with a single polynomial. Thus, it may be necessary to use local regression [121], in which the signal is partitioned into smaller sections to be fit separately by low order polynomials. A smooth result can be constructed by defining a degree of overlap between adjacent sections, and then averaging the results.

3.2.2. Frequency Domain Formulation

The process for curve fitting the amplitude and phase of an oscillation described in the previous section is posed in the time domain, where the difference between $\mathbf{y}$ and $\tilde{\mathbf{y}}$ is minimized directly in terms of the sampled values at each discrete time point. Alternatively, the Fourier Transform can be utilized to transfer this

information to the frequency domain where equivalent operations can be performed between Fourier coefficients. This allows for the optimization to be confined to operate within a specific frequency range, restricting the fit to consider only information that is relevant to the oscillation of interest.

To perform the optimization in the frequency domain, the Fourier Transform, denoted by $\mathcal{F}$, is applied to the sampled oscillation $\mathbf{y}$ and the matrix $\mathbf{H}(\boldsymbol{\kappa})$ as in Eq. (84) and Eq. (85), respectively.

$$\mathbf{Y}(\boldsymbol{\omega}) = \mathcal{F}(\mathbf{y}) \tag{84}$$

$$\boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}) = \mathcal{F}\big(\mathbf{H}(\boldsymbol{\kappa})\big) \tag{85}$$

The expression for the frequency domain curve fit is then formed as:

$$\widetilde{\mathbf{Y}}(\boldsymbol{\omega}) = \boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}) \begin{bmatrix} A \\ B \end{bmatrix} \tag{86}$$

While the total frequency range is denoted by $\boldsymbol{\omega}$, in the following expressions let $\boldsymbol{\omega}_{\mathrm{t}}$ represent a truncated region of those frequencies that contains information pertinent to the oscillation of interest. The linear least squares fit for A and B is then:

$$\begin{bmatrix} A \\ B \end{bmatrix} = \boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_{\mathrm{t}})^{\dagger} \mathbf{Y}(\boldsymbol{\omega}_{\mathrm{t}}) \tag{87}$$

The mathematical form of the curve fit is then given by Eq. (88), where the difference between $\mathbf{Y}(\boldsymbol{\omega}_{\mathrm{t}})$ and $\widetilde{\mathbf{Y}}(\boldsymbol{\omega}_{\mathrm{t}})$ is minimized.

$$\widetilde{\mathbf{Y}}(\boldsymbol{\omega}_{\mathrm{t}}) = \boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_{\mathrm{t}}) \left(\boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_{\mathrm{t}})^{\dagger} \, \mathbf{Y}(\boldsymbol{\omega}_{\mathrm{t}}) \right) \tag{88}$$

The gradient in the frequency domain is calculated similarly to the time domain derivation above, where the partial derivatives with respect to the optimization

variables are first computed with Eq. (76) for $\alpha_1, \alpha_2, \ldots$ and Eq. (77) for $\beta_1, \beta_2, \ldots$, and then simply converted to the frequency domain with the Fourier Transform.

$$\boldsymbol{\mathcal{M}}_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}, \boldsymbol{\omega}) = \mathcal{F}\left(\frac{\partial \mathbf{H}(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}}\right) \tag{89}$$

This renders the necessary frequency domain gradients as:

$$\frac{\partial \widetilde{\mathbf{Y}}(\boldsymbol{\omega}_t)}{\partial \boldsymbol{\kappa}} = \boldsymbol{\mathcal{M}}_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_t)\begin{bmatrix} A \\ B \end{bmatrix} - \boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_t)\left(\boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_t)^{\dagger}\boldsymbol{\mathcal{M}}_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_t)\begin{bmatrix} A \\ B \end{bmatrix}\right) \tag{90}$$

The procedure described above functionally minimizes the difference between a set of Fourier coefficients, which are complex numbers defining the amplitude and phase of the associated frequency bin. However, linear and nonlinear least squares algorithms may be unable to correctly handle complex values. A practical workaround that preserves the information in the original complex vectors is to simply stack their real and imaginary components:

$$\mathcal{F}(\omega) \rightarrow \begin{bmatrix} \mathrm{Re}\big(\mathcal{F}(\omega)\big) \\ \mathrm{Im}\big(\mathcal{F}(\omega)\big) \end{bmatrix} \tag{91}$$

This process is performed on each of the Fourier Transform results, $\mathbf{Y}(\boldsymbol{\omega}_t)$, $\boldsymbol{\mathcal{M}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_t)$, and $\boldsymbol{\mathcal{M}}_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}, \boldsymbol{\omega}_t)$, prior to computing $\widetilde{\mathbf{Y}}(\boldsymbol{\omega}_t)$ and the gradient.

It should be noted that the Fourier Transform above can be interpreted as a weighting on the time domain curve fit that simply redistributes the time data to be in terms of its frequency content, and hence the accuracy of the algorithm is not affected by leakage or other factors arising from the FFT. Specifically, the Fourier Transform could be represented as a matrix multiplication by the FFT matrix, $\mathbf{F}$.

$$\mathbf{Y}(\boldsymbol{\omega}) = \mathcal{F}(\mathbf{y}) = \mathbf{F}\mathbf{y} \tag{92}$$

To select the frequencies of interest, one would pre-multiply by a matrix $\mathbf{E}$ that is populated with zeros except at the frequencies to be retained, so $\mathbf{Y}(\boldsymbol{\omega}_t) = \mathbf{E}\mathbf{F}\mathbf{y}$ and the least squares problem simply becomes:

$$\begin{bmatrix} A \\ B \end{bmatrix} = \left(\mathbf{E}\mathbf{F}\mathbf{H}(\boldsymbol{\kappa})\right)^{\dagger} \mathbf{E}\mathbf{F}\mathbf{y} \tag{93}$$

where the pseudo-inverse in Eq. (87) now involves a weighting matrix $\mathbf{W} = \mathbf{F}^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\,\mathbf{E}\mathbf{F}$.

3.2.3. Multi-Harmonic Signals

In practice, it is unlikely that a sampled signal will contain a single pure oscillation with negligible noise. Typical physical measurements can contain multiple superimposed modal responses, each exhibiting distinct physical characteristics that are all simultaneously changing. Other nonlinear effects may also be present, such as higher order harmonics of the oscillations and coupling between modes. Extracting properties of a single oscillation from this type of complicated response usually entails a great deal of pre- and post-processing, like bandpass filtering and data smoothing. A benefit of the curve fitting process detailed in the previous sections is that it can directly fit the amplitude and phase of each oscillation without the need for filtering or other techniques that may negatively impact and corrupt the desired data.

The most straightforward means of fitting a single oscillation among many is to perform the optimization in the frequency domain. Attempting the fit in the time domain is only viable if the desired oscillation is significantly larger in amplitude than any other. This is due to the objective of the curve fit being to produce the smallest residual,

which is achieved by fitting and removing the largest component of the signal. Thus, by transferring to the frequency domain and restricting the fit to the Fourier coefficients in the vicinity of the oscillation frequency, other oscillations may be avoided. However, a potential complication with performing the optimization in the frequency domain is leakage in the Fourier coefficients.

If the measured signal, $\mathbf{y}$, contains multiple significant oscillations, the frequency band of interest may contain enough overlapping leakage to obscure the desired oscillation, making it impossible to accurately fit. To suppress the leakage in this scenario, a windowing function can be applied to each time domain vector prior to computing its Fourier Transform. Such windows artificially force the time signal to be periodic by smoothly decreasing its amplitude to zero at the extremities. A side effect of this behavior is that it effectively applies a weighting in the shape of the window to the least squares fit. As the data near the edges of the time segment is now close to zero, the greatest effect in minimizing the residual is accomplished by accurately fitting the content in the center of the time segment that was relatively unaffected by the window.

After completing the curve fit to the current oscillation, additional oscillations may be extracted from the resultant residual by employing the same procedure of windowing and transferring to the frequency domain. This can be performed successively until all desired oscillations have been characterized. However, if a scenario arises in which two or more oscillations cannot be independently fit, an alternative method is to expand the optimization to fit parameters to multiple

oscillations simultaneously. Doing so significantly increases the size of the optimization problem but allows for the combined effect of the oscillations to be considered in minimizing the residual.

3.2.4. Integration & Differentiation

From the polynomial representations of amplitude and phase generated in the previous sections, the integrals and/or derivatives of the fit oscillation can be analytically defined. For example, if the measured oscillation is of acceleration, representations of the associated velocity and displacement can be calculated. This is derived by first assuming that some function of time, $q(t)$, can be represented as a complex exponential, $\mathrm{e}^{\theta(t)}$, where $\theta(t)$ is an arbitrary time dependent complex number with finite real and imaginary components.

$$q(t) = \mathrm{e}^{\theta(t)} \tag{94}$$

The first derivative of this term with respect to time is simply the original complex exponential with a complex coefficient, $\dot{\theta}(t)$, that alters its amplitude and phase.

$$\dot{q}(t) = \dot{\theta}(t)\mathrm{e}^{\theta(t)} = \dot{\theta}(t)q(t) \tag{95}$$

This means that, if $\dot{\theta}(t)$ and $\dot{q}(t)$ are known quantities, calculating the integral of $\dot{q}(t)$ is equivalent to dividing it by $\dot{\theta}(t)$, reversing the amplitude scale and phase shift applied during the derivative.

$$q(t) = \frac{\dot{q}(t)}{\dot{\theta}(t)} \tag{96}$$

Similarly, the second derivative of $q(t)$ is the original complex exponential, but now with a complex coefficient based on $\dot{\theta}(t)$ and $\ddot{\theta}(t)$.

$$\ddot{q}(t) = \left(\ddot{\theta}(t) + \dot{\theta}^2(t)\right) \mathrm{e}^{\theta(t)} = \left(\ddot{\theta}(t) + \dot{\theta}^2(t)\right) q(t) \tag{97}$$

Through the same process as above, $\ddot{q}(t)$ can be transformed into $q(t)$ by dividing by the $\dot{\theta}(t)$ and $\ddot{\theta}(t)$ terms accrued through the two derivatives.

$$q(t) = \frac{\ddot{q}(t)}{\ddot{\theta}(t) + \dot{\theta}^2(t)} \tag{98}$$

Thus, through a combination of the steps above, if both $\dot{\theta}(t)$ and $\ddot{\theta}(t)$ are known, one of $q(t)$, $\dot{q}(t)$ or $\ddot{q}(t)$ can be used to reconstruct the other two.

In terms of applying this to the curve fit result, the oscillation is first represented as a complex exponential, $z(\boldsymbol{\kappa}, t)$:

$$z(\boldsymbol{\kappa}, t) = (A - iB)\mathrm{e}^{\beta_1 t + \beta_2 t^2 + \cdots i(\alpha_1 t + \alpha_2 t^2 + \cdots)} \tag{99}$$

Note that the real part of $z(\boldsymbol{\kappa}, t)$ is equivalent to $\tilde{y}(t)$:

$$\mathrm{Re}\big(z(\boldsymbol{\kappa}, t)\big) = \tilde{y}(t) = \mathbf{h}(\boldsymbol{\kappa}, t) \begin{bmatrix} A \\ B \end{bmatrix} \tag{100}$$

Since $\alpha(t)$ and $\beta(t)$ were taken to be simple polynomials, their derivatives with respect to time are:

$$\begin{aligned} \dot{\alpha}(t) &= \alpha_1 + 2\alpha_2 t + \cdots & \dot{\beta}(t) &= \beta_1 + 2\beta_2 t + \cdots \\ \ddot{\alpha}(t) &= 2\alpha_2 + 6\alpha_3 t + \cdots & \ddot{\beta}(t) &= 2\beta_2 + 6\beta_3 t + \cdots \end{aligned} \tag{101}$$

These terms can then be consolidated into $\dot{\theta}(t)$ and $\ddot{\theta}(t)$:

$$\begin{aligned} \dot{\theta}(t) &= \dot{\beta}(t) + \mathrm{i}\dot{\alpha}(t) \\ \ddot{\theta}(t) &= \ddot{\beta}(t) + \mathrm{i}\ddot{\alpha}(t) \end{aligned} \tag{102}$$

Thus, assuming $q(t)$ is displacement, $\dot{q}(t)$ is velocity, and $\ddot{q}(t)$ is acceleration, if the oscillation fit by $z(\boldsymbol{\kappa}, t)$ was displacement, then:

$$q(t) = z(\boldsymbol{\kappa}, t)$$
$$\dot{q}(t) = \dot{\theta}(t) z(\boldsymbol{\kappa}, t) \tag{103}$$
$$\ddot{q}(t) = \left(\ddot{\theta}(t) + \dot{\theta}^2(t)\right) z(\boldsymbol{\kappa}, t)$$

Similarly, if $z(\boldsymbol{\kappa}, t)$ is velocity:

$$q(t) = \frac{z(\boldsymbol{\kappa}, t)}{\dot{\theta}(t)}$$
$$\dot{q}(t) = z(\boldsymbol{\kappa}, t) \tag{104}$$
$$\ddot{q}(t) = \frac{\ddot{\theta}(t) + \dot{\theta}^2(t)}{\dot{\theta}(t)} z(\boldsymbol{\kappa}, t)$$

And finally, if $z(\boldsymbol{\kappa}, t)$ is acceleration:

$$q(t) = \frac{z(\boldsymbol{\kappa}, t)}{\ddot{\theta}(t) + \dot{\theta}^2(t)}$$
$$\dot{q}(t) = \frac{\dot{\theta}(t)}{\ddot{\theta}(t) + \dot{\theta}^2(t)} z(\boldsymbol{\kappa}, t) \tag{105}$$
$$\ddot{q}(t) = z(\boldsymbol{\kappa}, t)$$

3.2.5. Natural Frequency & Damping Ratio

While the curve fit provides the amplitude and phase of the oscillation, these do not directly characterize any useful aspects of the physical system that generated the signal. In structural dynamics, nonlinear systems can typically be identified in terms of how the modal natural frequencies and damping ratios shift as a function of response amplitude. To calculate these quantities, an implementation of Feldman's FREEVIB [62] and FORCEVIB [63] methods will be utilized. These algorithms are for

free and forced response respectively and primarily operate on the derivatives of the amplitude and phase of the oscillation. Feldman derived these from the perspective that the Hilbert Transform and narrow bandpass filters would be used to estimate the amplitude and phase and their derivatives. Since the proposed curve fitting method directly produces polynomial representations of the amplitude and phase, their respective derivatives are simply calculated as in Eq. (101). This bypasses any additional data processing typically required when using the Hilbert Transform, such as filtering the signal to be monoharmonic and smoothing noise from the derivatives.

To derive FREEVIB and FORCEVIB in terms of the optimization results, first assume that the nonlinear system can be characterized by the quasi-linear EOM model form defined in Eq. (1). Instead of functions of amplitude, for now assume that the natural frequency and damping ratio terms are functions of time.

$$\ddot{x}(t) + 2\zeta(t)\omega_n(t)\dot{x}(t) + \omega_n^2(t)x(t) = f(t) \tag{106}$$

Assuming the response and forcing terms are known, there is one equation containing two unknowns. To formulate expressions for $\omega_n(t)$ and $\zeta(t)$, Feldman proposed that the EOM be expanded from real space into the complex plane such that two equations can be formed from the real and imaginary components. For the acceleration, velocity, displacement, and external forcing assume that $\ddot{q}(t)$, $\dot{q}(t)$, $q(t)$, and $p(t)$ represent complex quantities where the real part of each is equal to $\ddot{x}(t)$, $\dot{x}(t)$, $x(t)$, and $f(t)$, respectively. Substituting these into Eq. (106) results in:

$$\ddot{q}(t) + 2\zeta(t)\omega_n(t)\dot{q}(t) + \omega_n^2(t)q(t) = p(t) \tag{107}$$

While Feldman assumed that the complex representations of the responses and force could be estimated with the Hilbert Transform, here instead we will adopt a formulation based on the polynomial expressions used in the nonlinear optimization process detailed in the previous sections. As $p(t)$ is also required in a complex form, note that an additional optimized curve fit will need to be performed on the sampled external forcing when applicable. Using the expressions for $\dot{q}(t)$ in Eq. (95) and $\ddot{q}(t)$ in Eq. (97), Eq. (107) becomes:

$$\left(\ddot{\theta}(t) + \dot{\theta}^2(t)\right) q(t) + 2\zeta(t)\omega_n(t)\dot{\theta}(t)q(t) + \omega_n^2(t)q(t) = p(t) \tag{108}$$

After dividing through by $q(t)$, this is further expanded with the expressions for $\dot{\theta}(t)$ and $\ddot{\theta}(t)$ in Eq. (102).

$$\begin{gathered}\left(\ddot{\beta}(t) + i\ddot{\alpha}(t)\right) + \left(\dot{\beta}^2(t) + 2i\dot{\alpha}(t)\dot{\beta}(t) - \dot{\alpha}^2(t)\right) + \\ 2\zeta(t)\omega_n(t)\left(\dot{\beta}(t) + i\dot{\alpha}(t)\right) + \omega_n^2(t) = \frac{p(t)}{q(t)}\end{gathered} \tag{109}$$

Since $p(t)$ and $q(t)$ are in a complex form, it is interesting to note that their ratio, $p(t)/q(t)$, defines the relative magnitude and phasing between the force and displacement, effectively quantifying the energy transfer into the system due to the external forcing. Separating the real and imaginary parts produces Eq. (110) and Eq. (111), respectively.

$$\ddot{\beta}(t) + \dot{\beta}^2(t) - \dot{\alpha}^2(t) + 2\zeta(t)\omega_n(t)\dot{\beta}(t) + \omega_n^2(t) = \mathrm{Re}\left(\frac{p(t)}{q(t)}\right) \tag{110}$$

$$\ddot{\alpha}(t) + 2\dot{\alpha}(t)\dot{\beta}(t) + 2\zeta(t)\omega_n(t)\dot{\alpha}(t) = \mathrm{Im}\left(\frac{p(t)}{q(t)}\right) \tag{111}$$

Assuming that $\dot{\alpha}(t)$, $\dot{\beta}(t)$, $\ddot{\alpha}(t)$, $\ddot{\beta}(t)$, and $q(t)$ are known quantities from the oscillation curve fit, and the force $p(t)$ is also known in a complex form, these two equations can be solved for $\zeta(t)$ and $\omega_n(t)$:

$$\omega_n(t) = \sqrt{\dot{\alpha}^2(t) + \dot{\beta}^2(t) - \ddot{\beta}(t) + \ddot{\alpha}(t)\frac{\dot{\beta}(t)}{\dot{\alpha}(t)} - \frac{\dot{\beta}(t)}{\dot{\alpha}(t)}\mathrm{Im}\left(\frac{p(t)}{q(t)}\right) + \mathrm{Re}\left(\frac{p(t)}{q(t)}\right)} \quad (112)$$

$$\zeta(t) = \frac{1}{2\omega_n(t)}\left(-2\dot{\beta}(t) - \frac{\ddot{\alpha}(t)}{\dot{\alpha}(t)} + \frac{1}{\dot{\alpha}(t)}\mathrm{Im}\left(\frac{p(t)}{q(t)}\right)\right) \quad (113)$$

While these are calculated directly with respect to time, this relation can be exchanged for the associated response amplitude returned by the curve fit at each time sample to produce the amplitude dependent natural frequency and damping ratio.

As a final note, it should be acknowledged that these expressions are incredibly general, in that, as long as all the required complex response and force information can be accurately determined, any free or forced SDOF response can be processed. However, most use cases of this type of analysis have been for free decay where it is assumed that the nonlinear behavior can be fit in piecewise linear fashion with local regression; notable examples being [58], [65]. To apply the proposed curve fitting method in these scenarios, the amplitude and phase are assumed to be linear first order polynomials, $\alpha(t) = \alpha_0 + \alpha_1 t$ and $\beta(t) = \beta_0 + \beta_1 t$, where the first derivatives are $\dot{\alpha}(t) = \alpha_1$ and $\dot{\beta}(t) = \beta_1$, and the second derivatives are consequently zero. After disregarding the forcing terms, as this is modeling free decay, the above expressions for instantaneous natural frequency and damping ratio reduce to the following, where

the time dependence has been dropped since the results are constant over the current interval.

$$\omega_n = \sqrt{\alpha_1^2 + \beta_1^2} \tag{114}$$

$$\zeta = -\frac{\beta_1}{\omega_n} \tag{115}$$

Rearranging Eq. (114) and Eq. (115) to be in terms of α_1 and β_1 yields:

$$\alpha_1 = \omega_n\sqrt{1-\zeta^2} = \omega_d \tag{116}$$

$$\beta_1 = -\zeta\omega_n \tag{117}$$

Now in this case, recall that the equation for the optimized curve fit is as given in Eq. (118), where A and B are defined by the initial conditions of the response.

$$\tilde{y}(t) = e^{\beta_1 t}(A\cos(\alpha_1 t) - B\sin(\alpha_1 t)) \tag{118}$$

Along with Eq. (116) and Eq. (117), this is then equivalent to the analytical solution for the free response of a linear SDOF EOM [103], shown in Eq. (119).

$$x(t) = e^{-\zeta\omega_n t}\left(x_0\cos(\omega_d t) + \frac{\dot{x}_0 + \zeta\omega_n x_0}{\omega_d}\sin(\omega_d t)\right) \tag{119}$$

This demonstrates that the general expressions in Eq. (112) and Eq. (113) reduce to the conventional linear modal parameters after making the requisite assumptions.

3.3. Case Studies

3.3.1. Application to Forced Modal Iwan Model

To demonstrate various aspects of the NLSID method described above, this section details a numerical case study investigating simulated response from an SDOF Modal Iwan Model [122]. The Iwan Model is typically used to model the effects of microslip in bolted joints [65] as both exhibit a strong damping nonlinearity and moderate shift in natural frequency, making it an ideal benchmark for evaluating the proposed NLSID method. It is also noteworthy that the underlying EOM for the Iwan Model is hysteretic and not easily modelled by polynomials of displacement and velocity, making most other NLSID techniques unsuitable in this scenario. While several parametrizations for the Iwan Model have been proposed, implemented here is the 4-parameter Iwan Model proposed by Segalman in [25]. A benefit of this formulation of the Iwan Model is that, in the microslip regime, Segalman derived closed form expressions for the amplitude dependent natural frequency and damping ratio which provide truth data to compare the NLSID results to. For this case study, the model parameters given in Table 1 were utilized. These values result in a linear natural frequency of 1 [Hz] and damping ratio of 1%.

Table 1: Parameters used in the SDOF Modal Iwan Model Numerical Case Study

Parameter	F_S	K_T	χ	β	K_∞
Value	400 [N]	25.27 [N/m]	-0.35 [-]	10 [-]	14.21 [N/m]

The Iwan Model was excited by a sine beat [123], [124], which is formed by applying a windowing function to a stationary sinusoid. By selecting a window that suppresses leakage, such as a 4-term Blackman-Harris window [125], this type of forcing corresponds to a very concentrated pulse of energy applied to a narrow frequency band. The excitation can then be focused onto specific resonances, driving them to high response amplitudes and causing a subsequent free decay after the force is removed [126].

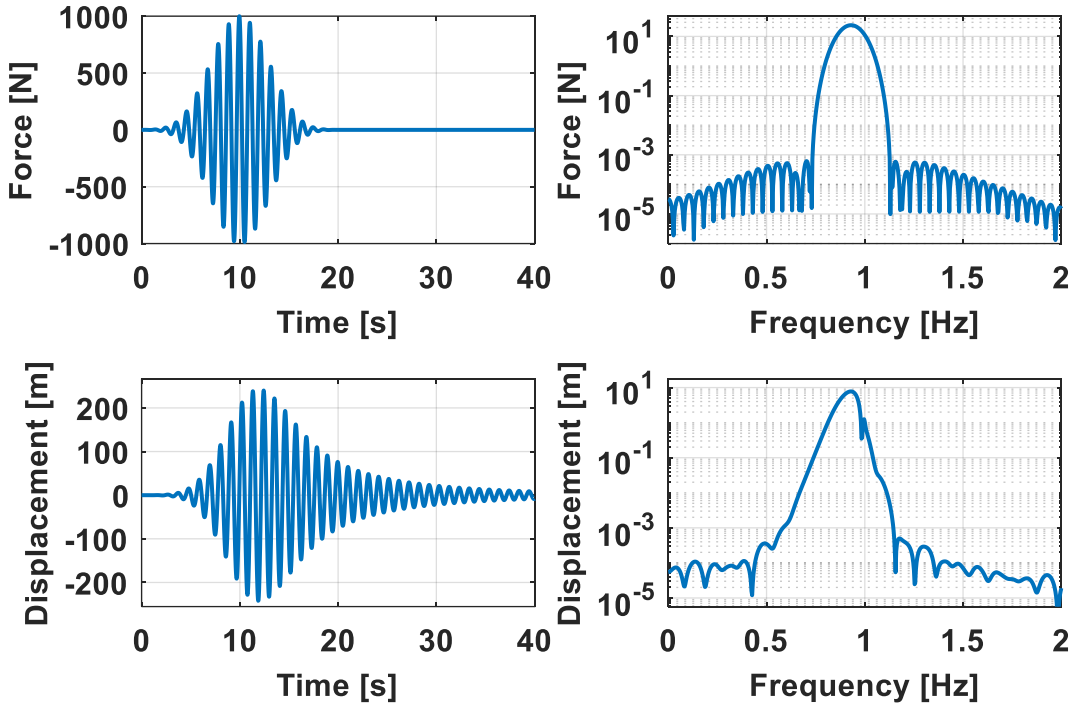

Figure 32: Top - The applied force, in the form of a sine beat, in the time and frequency domain. Bottom - The resultant response of the Iwan model in time and frequency domains.

To simulate the response of the Iwan Model to the sine beat excitation, the integration process developed by Shetty in [127] was utilized. The simulation results are shown in Figure 32, in which the time series of the applied forcing is depicted in the top left plot while its corresponding frequency content is in the top right.

This sine beat is centered at 0.93 [Hz], has a frequency band width of 0.2 [Hz], and ends at 20 [s]. The resultant displacement is given in the bottom plots, with the time series exhibiting similar behavior as the force, but with a clear ring down after the excitation concludes. However, when comparing the frequency domain representations, the response is quite distorted, indicative of the nonlinear aspects of the simulated system.

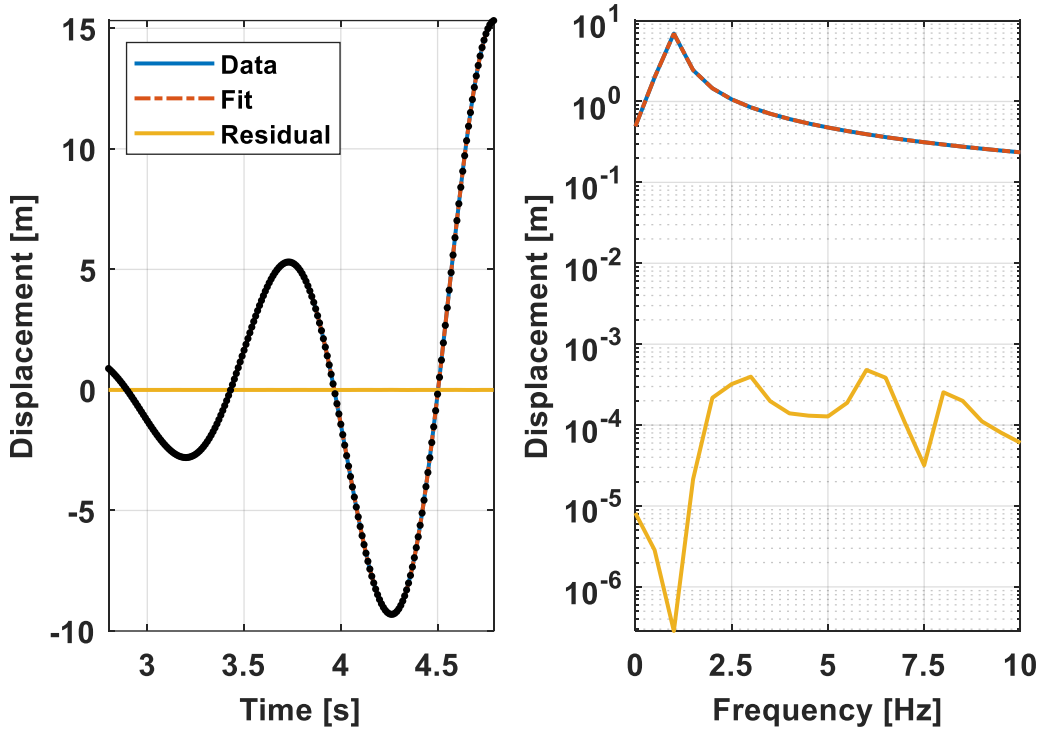

Figure 33: An excerpt from the optimization showing the resultant curve fit to a segment of the displacement. The data points utilized in computing the residual are highlighted as black dots.

To curve fit the simulated displacement with the optimization based NLSID method, the amplitude and phase were represented as fifth order polynomials and time segments 2 [s] long with 90% overlap were chosen. As the response is primarily monoharmonic, the least squares solution was performed directly on the time series. The optimized fit to one of the segments of response is displayed in Figure 33. On the

left side, the time vectors are shown, with the response as a solid blue line, the fit as a dashed orange line, the residual in yellow, and the points used in the optimization highlighted as black dots. The accuracy of the curve fit is exemplified prominently in the frequency domain plot on the right, where the residual is between four and six orders of magnitude less than the response and fit.

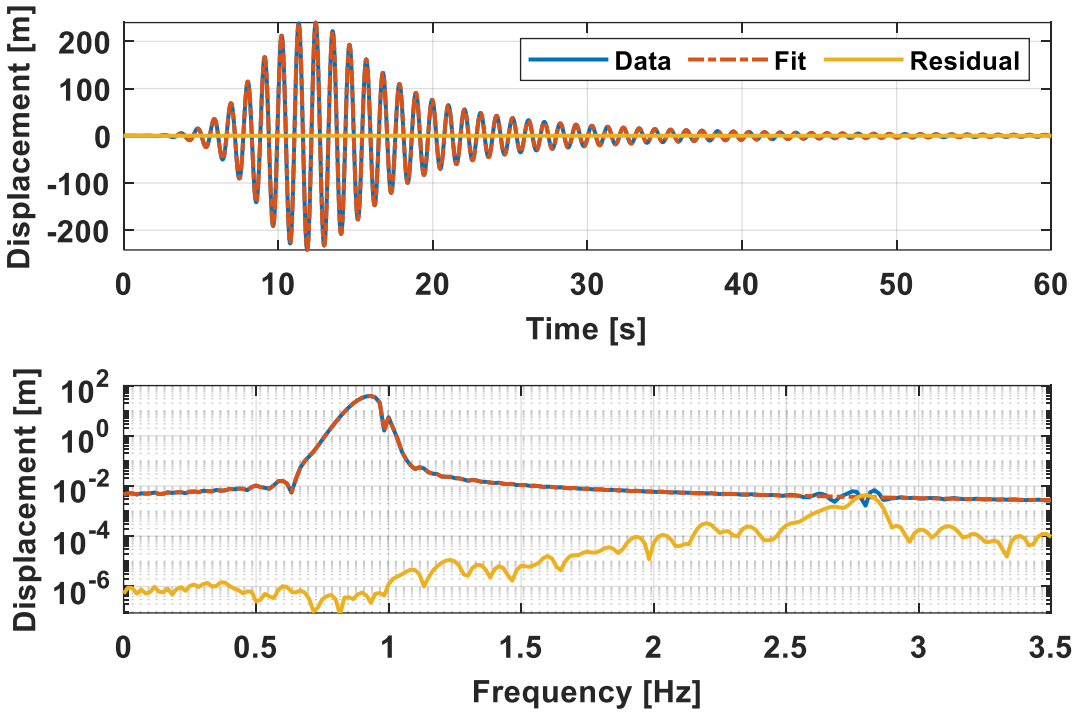

Figure 34: Overall fit to the simulated displacement of the Iwan model. The residual under the main peak is approximately eight orders of magnitude less than the response.

After computing the curve fit in each time segment, the overlapping results were averaged to produce an overall fit to the entire displacement time series. A comparison between the response and the fit is given in Figure 34, where again the response is in blue, the fit is orange, and the residual is yellow. The frequency domain residual under the main peak of the displacement is approximately six orders of magnitude less than the response. The most prominent feature in the residual is a third harmonic of the

main response, visible between 2.5 [Hz] and 3 [Hz]. While harmonics such as this are common in nonlinear responses, they cannot be modeled with the quasi-linear model form and are essentially noise in this case. It is then a beneficial outcome that the curve fit has effectively filtered it from the data, characterizing only the nonstationary oscillation of the fundamental harmonic.

As this is a numerical case study with simulated responses, the acceleration, velocity, and displacement are all known quantities. To demonstrate that the optimization can approximate the derivatives of the displacement, the resultant curve fit polynomials were used to form estimates of the velocity and acceleration. These are compared to the truth data from the simulation in Figure 35, in which the velocity is shown on the left and the acceleration is on the right. While neither are as accurate as the fit to the displacement, the residuals near the main peaks are at least three orders of magnitude less than the truth data. The error in these estimates is likely caused by small discrepancies in the higher order polynomial terms, as these comprise the derivatives of the amplitude and phase that are significant in forming the representations of velocity and acceleration.

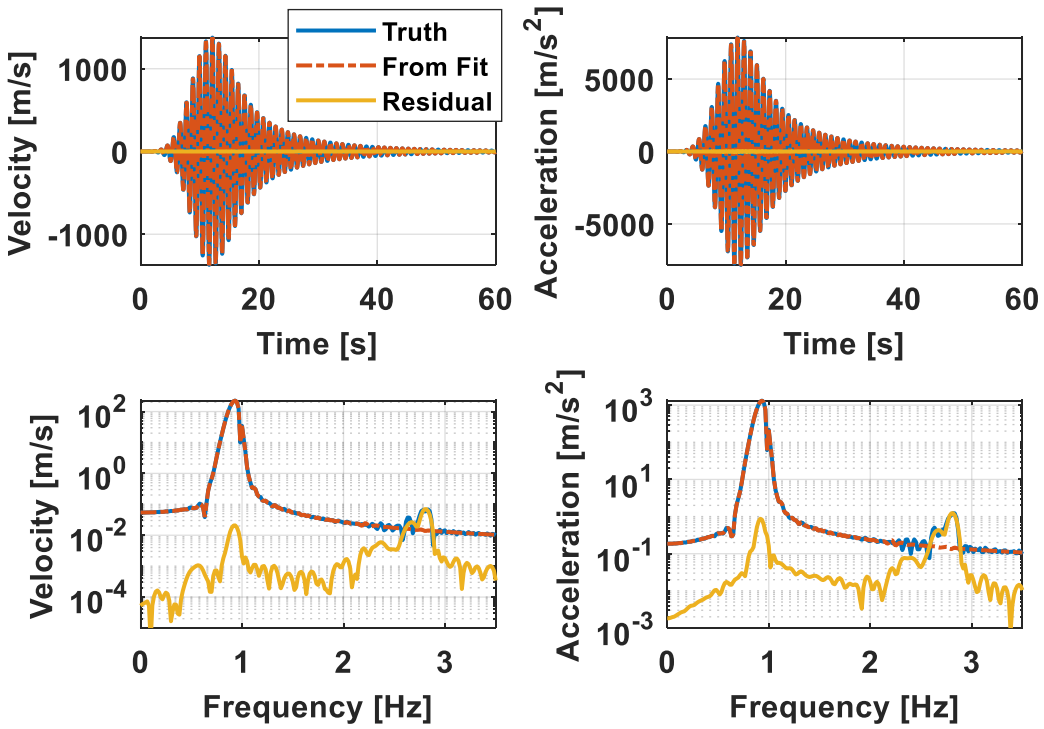

Figure 35: Comparison of the velocity and acceleration directly from the integration verses estimated from the optimized curve fit to the displacement.

Most of the response is in free decay, meaning that the instantaneous natural frequency and damping ratio can be determined directly from the curve fit to the displacement with FREEVIB. However, during the initial 20 seconds the system is actively excited by the applied sine beat. To calculate the instantaneous values in this region, the general expression for FORCEVIB is required in which the force must be represented as a complex number. As the applied force is defined by a vector of real values, the complex form was determined with the optimization curve fitting procedure. The same time segment size and polynomial order was used to fit the force and displacement. The resultant fit is shown in Figure 36, which displays the same level of accuracy as had when fitting the displacement.

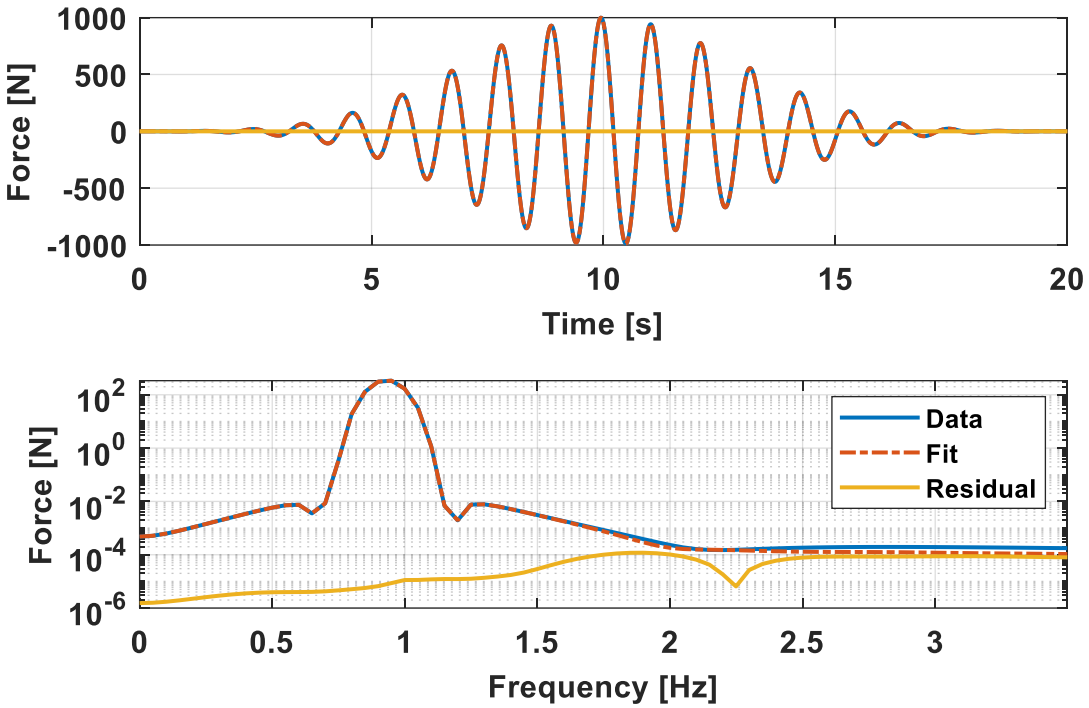

Figure 36: The resultant fit to the force applied to the Iwan model. This result is required in the FORCEVIB method, where the force must be known in a complex form.

From the curve fits to the displacement and force, all terms in Eq. (112) and Eq. (113) are known such that the natural frequency and damping ratio of the Iwan Model can be determined with respect to time and response amplitude. The results are shown in Figure 37, in which the left plots are frequency, the right plots are damping, the top is verses time, and the bottom is verses velocity amplitude. In each, the truth solution from the closed form equations is represented by the blue line, while the FREEVIB/ FORCEVIB result from the curve fit is shown as a dashed orange line. In the time dependent curves, the natural frequency lowers by approximately 10% as the response reaches the peak amplitude during the sine beat. In the subsequent ring down, the frequency curve converges back to the low amplitude linear value. The damping ratio displays a similar behavior, except increases with amplitude to a value

nearly five times larger than the linear values. The amplitude dependence in each is more apparent in the lower plots, where both curves trace out consistent backbones during the initial ramp up and ensuing ringdown.

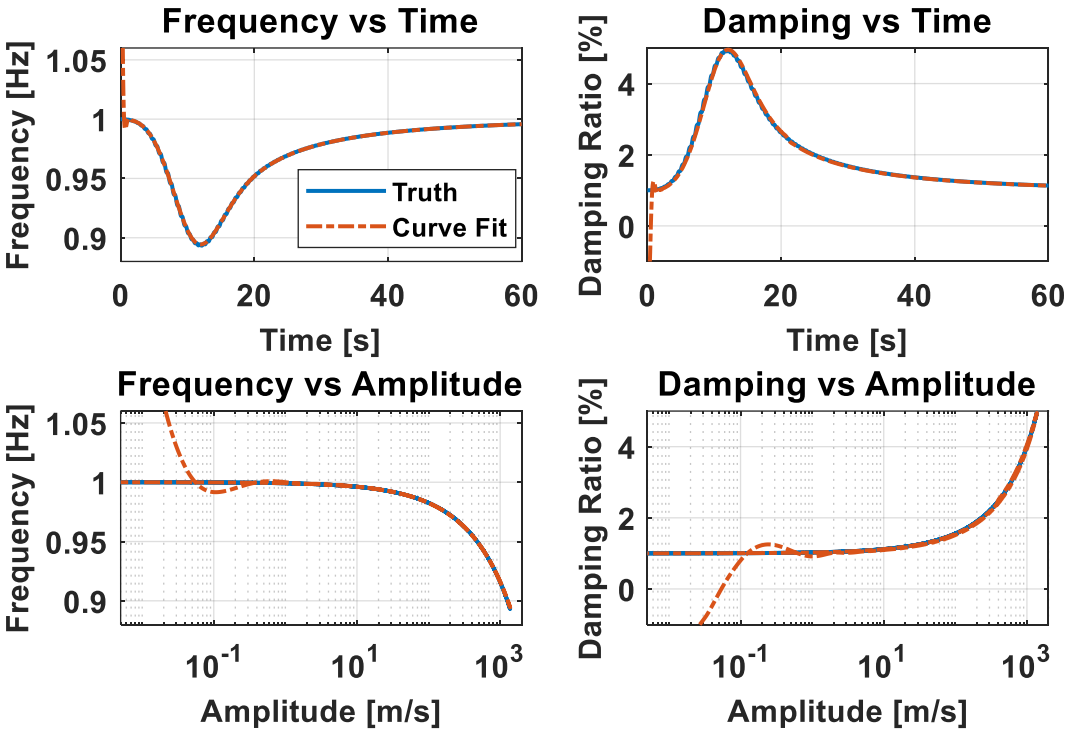

Figure 37: Results from optimization based FORCEVIB, compared to the truth solution for the simulated Iwan model. The only significant error is an initial end effect.

The only significant error is an initial end-effect at the very start of the simulation. This is due to a combination of representing the amplitude and phase as polynomials and the curve fit being computed in terms of a least squares solution. To produce the smallest residual, the optimization is naturally biased toward accurately fitting regions of the data with the largest amplitude as these have the most significant impact on the resultant error. Consequently, if the oscillation in the current time segment exhibits a rapid change in amplitude, the best-fit polynomials will be biased toward the data with the larger values. At the very start of the simulation, the ramping

force causes the response to quickly increase from zero, leading to the polynomial fits being more accurate toward the data later in the segment. In the context of the curve fit, this is the correct result as the residual is then equally small throughout the segment. However, since the natural frequency and damping ratio are based mainly on the derivatives of the polynomials, their error is prominent and amplified, leading to the significant end-effects visible in Figure 37.

This numerical case study demonstrated many key aspects of the proposed NLSID method. As the response was from a simulated 4-parameter Iwan Model, the nonlinear behavior has known truth data to compare the identified model to. First, the curve fit to the amplitude and phase of the displacement yielded errors six orders of magnitude less than the response. From that result, representations of the velocity and acceleration were shown to be reasonably accurate. Then, from curve fits to the displacement and force, the instantaneous natural frequency and damping ratio were calculated with FORCEVIB while the force was active and FREEVIB during the subsequent free decay. When compared to the known truth data, the estimated curves are very accurate apart from an initial end-effect. While this is an idealized case with perfectly clean data, the favorable comparisons to the known nonlinear behavior of the Iwan Model point to this NLSID approach potentially being very effective for identifying a quasi-linear representation of a nonlinear system with bolted joints. This is evaluated in the next section, where the NLSID method is used to characterize the nonlinear response of a jointed experimental benchmark structure.

3.3.2. Application to Experimental Impact Response

In structural dynamics, a common testing technique is to excite a structure with a modal hammer and analyze the resulting transient free decay. Impact testing is incredibly simple to set up and perform and the subsequent data analysis can be completed with many straightforward and well-established procedures. A typical method for determining amplitude dependent natural frequency and damping ratio from an impulse response is with the Hilbert Transform and FREEVIB [60], [62]. However, in implementing this approach, the user must grapple with the limitations of the Hilbert Transform in that the signal must be monoharmonic and errors will occur near sudden changes in the oscillation. These complications are significant here because the hammer impact will likely excite multiple modal responses that will unavoidably begin abruptly. A standard means of mitigating these difficulties is though modal and/or bandpass filtering the signal down to individual oscillations and discarding regions of the data corrupted by end-effects. While these are functional solutions, they reduce the amount of usable information that can be gleaned from the measured responses. Alternatively, the optimization based NLSID technique detailed in the previous section could be used to directly extract and characterize the various modal responses from the free decay, avoiding the detrimental effects of filters and the Hilbert Transform. In this section, measured nonlinear response from an experimental system is analyzed, demonstrating how the proposed NLSID technique can be utilized to produce improved results for a real jointed structure over a more conventional approach involving the Hilbert Transform.

This case study considers the nonlinear benchmark structure known as the S4 Beam that was explored previously in section 2.5. The structure and test setup as portrayed in [19] was shown in Figure 26. The measured acceleration response due to the modal hammer impact is shown in Figure 38, in which six significant modal response peaks are visible in the frequency domain. The zoomed in plots of the time series show that the early transient is dominated by the pair of modes near 1300 [Hz] and 1500 [Hz], while the later portion of the ringdown is primarily the first elastic mode near 280 [Hz].

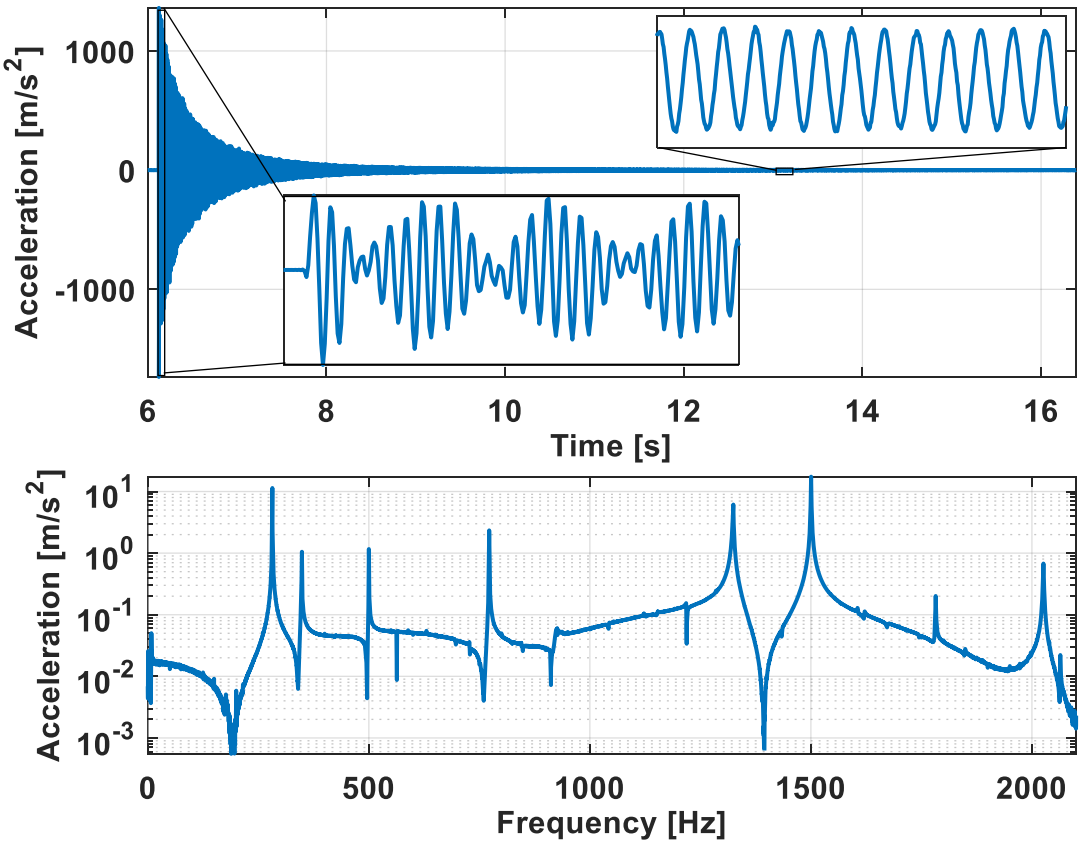

Figure 38: The recorded acceleration of the S4 beam near the drive point. The accelerometer is located on the opposite side of where the hit was applied to the beam.

To characterize the response of the first mode during the ringdown, the measured signal was divided into segments 0.25 [s] long with 95% overlap. In each, third order polynomials for the amplitude and phase were curve fit to the data in the frequency domain. As there are several prominent oscillations, a 4-term Blackman-Harris window [125] was applied to the segments before computing the Fourier Transform. This limits the amount of leakage from other oscillations in the frequency bins about the signal of interest.

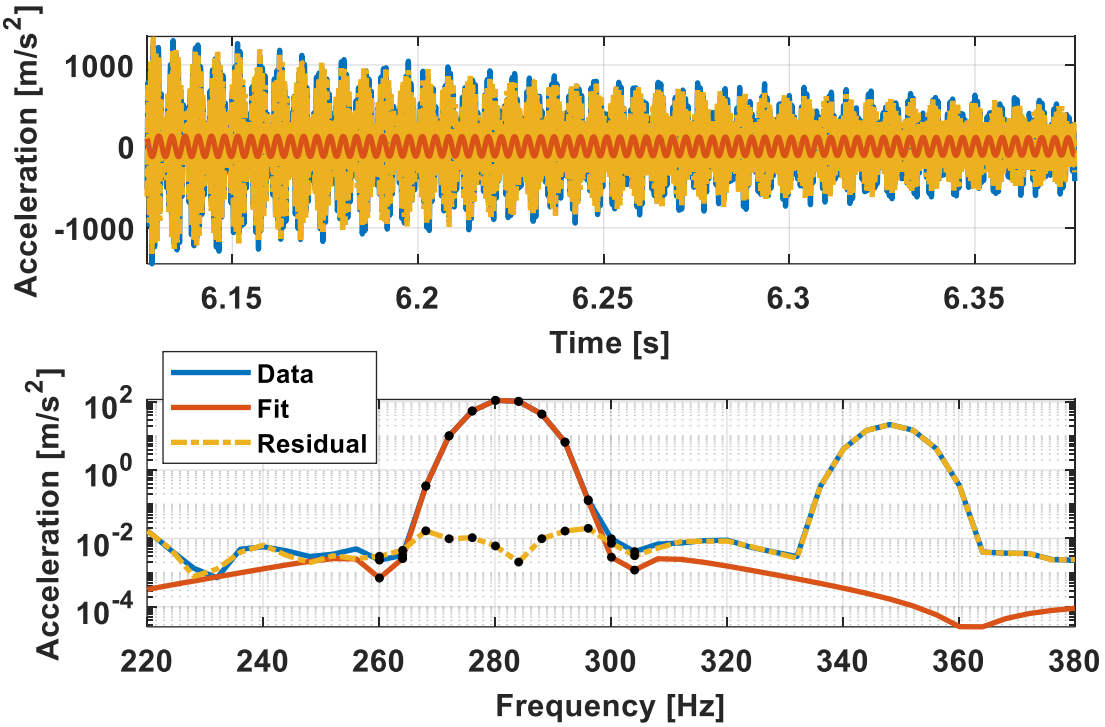

Figure 39: The optimized curve fit to the oscillation of mode 1 in the first time segment. The top shows the time vectors while the bottom shows the frequency domain fit. The black dots mark which Fourier coefficients were used to perform the curve fit. In this instance, a Blackman-Harris Window was applied to the time vector before computing the Fourier coefficients.

The result of the optimization in the first time segment is shown above in Figure 39. The lower plot displays the windowed frequency domain information, where the data points used in minimizing the residual are marked by black dots that encompass the

first resonance. In the upper plot, the resulting oscillation from the best-fit amplitude and phase polynomials is overlaid on the original signal, along with the residual between them. It is clear from this representation that the response of the first mode is much lower amplitude than the rest of the data in this segment, which is primarily dominated by the modes near 1300 [Hz] and 1500 [Hz] noted earlier. While the curve fit was performed easily in the frequency domain, achieving the same fit in the time domain would be quite difficult as there the minimum residual would be found by fitting those higher amplitude responses. The only caveat to using the windowed Fourier coefficients in the optimization is that it is then effectively a weighted least squares fit. Since the window artificially suppressed the signal near the segment edges, the fit is less accurate there and more accurate in the center.

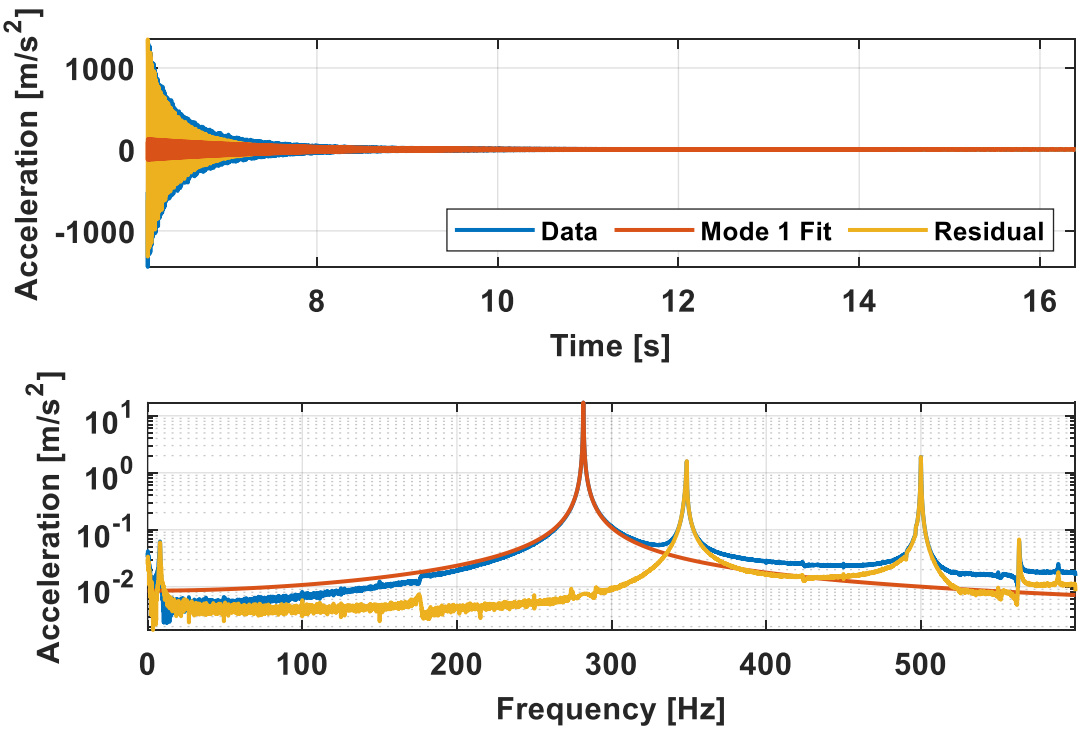

Figure 40: Resultant fit to mode 1 after performing the individual fits in each segment and averaging the portions that overlap.

After performing the optimization in each time segment and averaging results in overlapping sections, the overall fit to the free decay of the first mode is found as shown in Figure 40. It can be seen in the frequency domain representation in the lower plot that the first mode has been cleanly extracted from the signal, as there is essentially no indication of it left in the residual.

A similar process was implemented to characterize the remaining modal responses. After completing an optimized fit to the current oscillation throughout the time series, the residual was taken to be the data from which the next curve fit would be performed. Thus, best-fit estimates of each oscillation were consecutively computed and subtracted from the original measured response until all significant frequency peaks were removed, leaving only noise and resonances that were too weakly excited to merit characterization.

The final result of this procedure is shown below in Figure 41. In the top plot, the separate curve fits to each resonance peak are overlaid on the measured response, showing the variety of signals present in the data that can be individually extracted. The bottom plot compares the original recorded signal shown in Figure 38 to the summation of all curve fits and gives the subsequent residual. This demonstrates the accuracy achieved by the optimized curve fits, as the residual is simply the noise floor of the measurement intermixed with a multitude of very low amplitude response peaks.

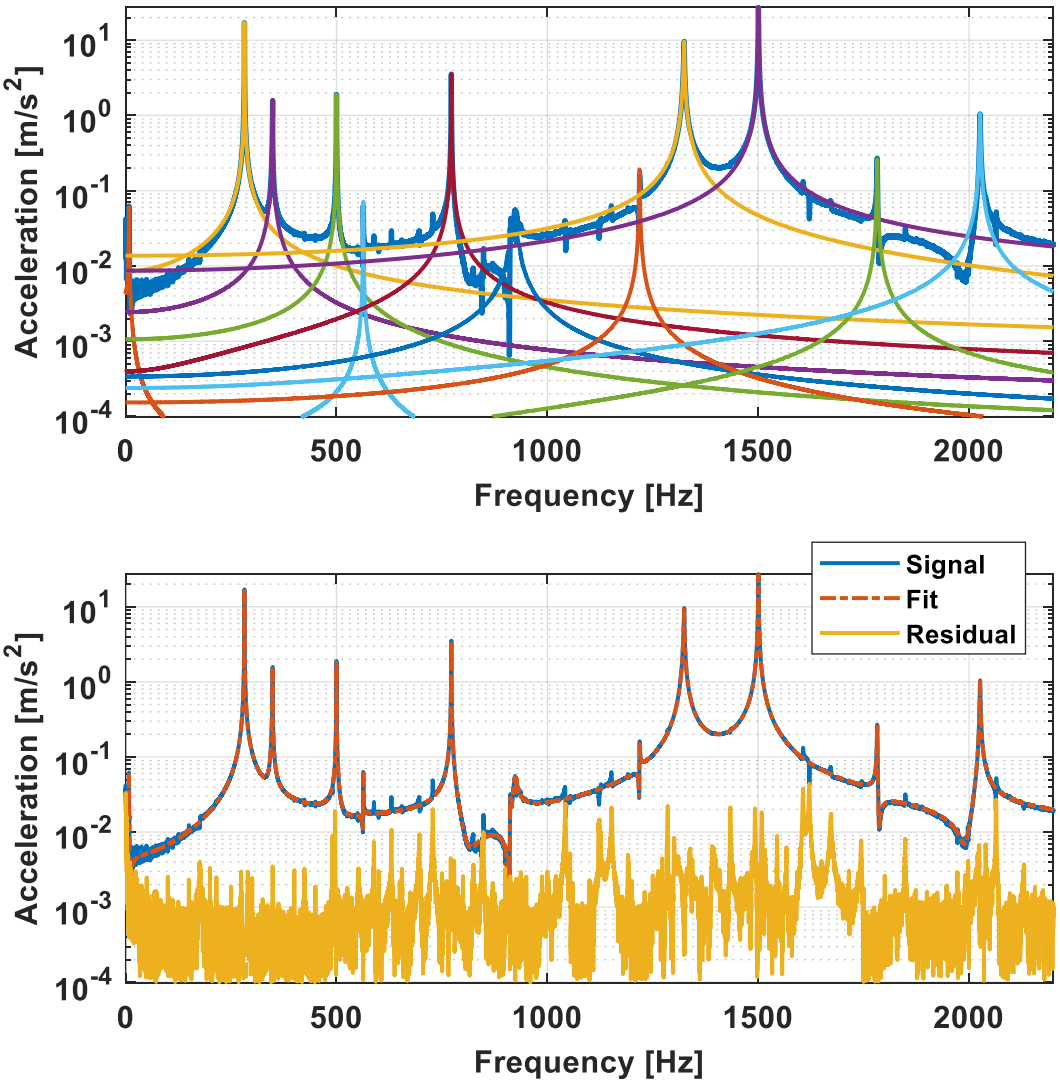

Figure 41: Top - Overlay of all curve fit oscillations. Bottom - Composite of all fits vs the recorded data shown previously in Figure 38.

The apparent accuracy of the curve fits to the measured response lends confidence in the associated quality of the identified quasi-linear modal parameters for each resonance. The amplitude dependent natural frequency and damping ratio for the first and second elastic modes of the S4 Beam are shown in Figure 42 with the frequency on the left and the damping on the right. Additionally, an image of the deflection shape of each mode from a finite element model of the structure is shown

over the left side. In each plot, the results from the optimized curve fits are given by an orange line. These are overlaid on a set of blue points that represent values determined by Wall in [19], using an algorithm akin to Feldman's FREEVIB detailed in [128]. In that process, the response measurements are modal and bandpass filtered down to a single nonstationary oscillation, with the Hilbert Transform then used to estimate the amplitude and phase. To avoid the accompanying end effects, the beginning and end of the signal are simply discarded. As the natural frequency and damping ratio are dependent on the derivatives of the amplitude and phase, any noise present in the signal can significantly corrupt the results. To prevent this, a technique proposed in [104] is employed in which smoothed derivatives are estimated from the slope of piecewise linear curves fit to the amplitude and phase with local regression. This approach is somewhat similar to the optimization method proposed here but is limited to forming piecewise linear models of the nonlinear signal, producing a single data point in each segment as opposed to the solution at every time sample provided by the optimization. Nonetheless, Hilbert Transform methods such as what was used by Wall in [19] are popular due to the simplicity and computational efficiency of the underlying algorithm. This typically comes at the cost of a great deal of user interaction to tune the aspects of the solution process to ensure the results are acceptably smooth and accurate. In its current form, the optimization process suffers from similar drawbacks, in that trial and error is used to select an optimal constant size for the time segments. However, the benefit here is improved accuracy in the fits, avoidance of Hilbert Transform errors, and solutions at every time sample.

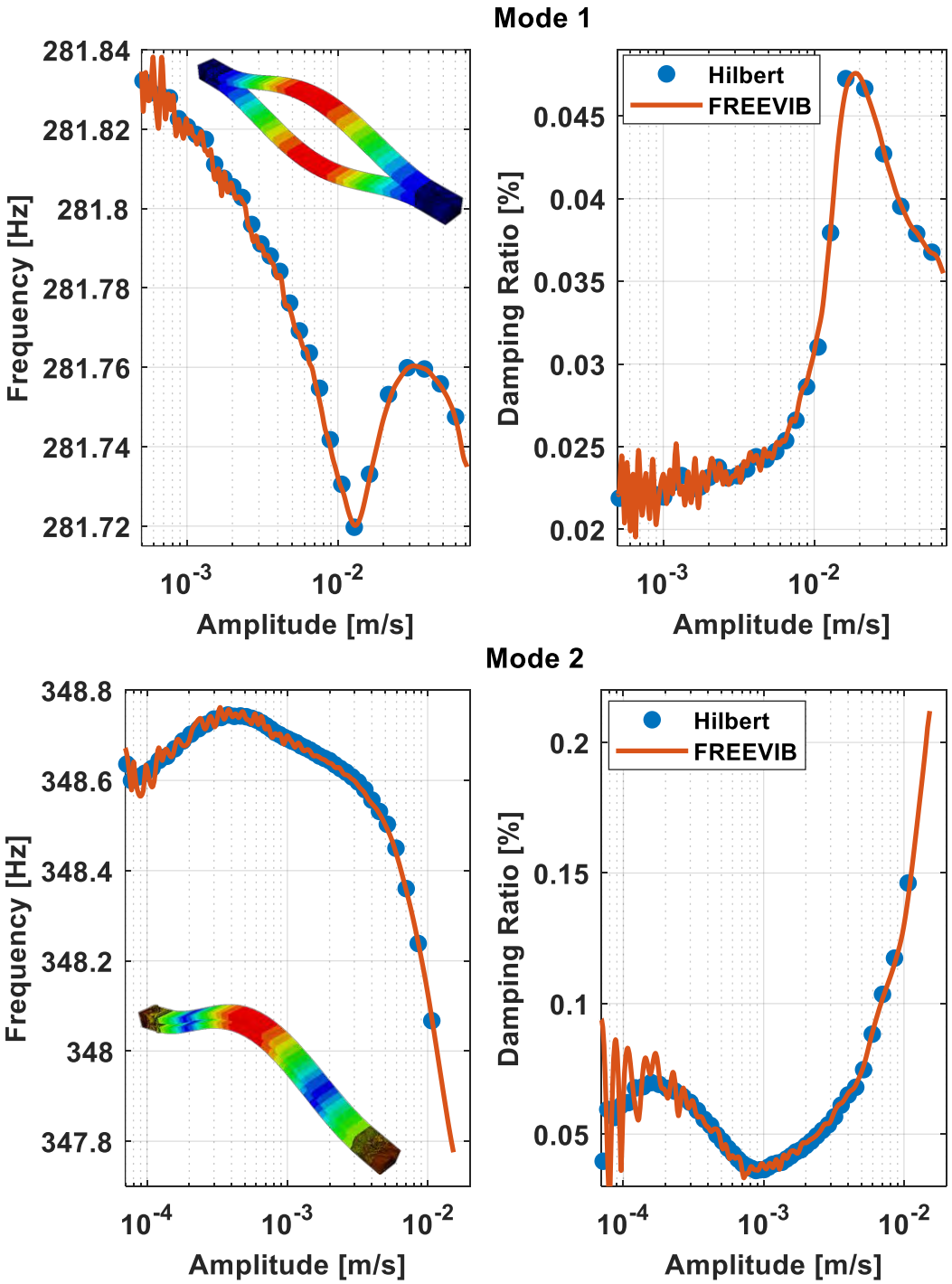

Figure 42: Amplitude dependent natural frequency (Left) and damping ratio (Right) for elastic mode 1 (Top) & mode 2 (Bottom) of the S4 Beam. In each, the orange curve represents the optimized curve fit results, while the blue dots are from a Hilbert Transform algorithm used by Wall in [19]. To demonstrate the deflection shape of the resonances, an image of each mode shape from a finite element model is shown superimposed over the left plots.

The data plotted in Figure 42 exemplify the benefits of the optimization over the Hilbert Transform. While both solution methods are in good agreement as to the natural frequency and damping ratio at specific amplitude levels, the optimization provides a dense set of data points with much finer amplitude resolution. An interpolation scheme could be used to fill in the gaps between the Hilbert results, but this would generate points that are not directly based on the response data like the optimization and could produce inaccurate and spurious results. Additionally, more of the initial high amplitude information is captured by the nonlinear curve fit, given by the orange lines extending further to the right than the blue dots in each plot, especially those for the second mode in Figure 42. It is also noteworthy that at lower amplitudes the optimization results appear quite noisy, with significant variance about the Hilbert points. This is due to the nonlinear curve fit operating on time segments with a constant size that was chosen to produce the smallest overall residual. Thus, the segment size is biased toward accurately capturing the high amplitude, initial portion of the response that is rapidly varying and necessitates relatively short segments. As the signal decays to lower amplitudes with characteristics that vary more slowly, larger time segments would be well suited to producing smooth results while the response approaches the noise floor of the measurement. Addressing this issue of constant segment size is a topic for future work.

This case study explored characterizing the nonlinear behavior of a bolted-joint benchmark structure in terms of quasi-linear natural frequencies and damping ratios from free response measured after an impact excitation. The FREEVIB method

provides a well-established process for extracting these quasi-linear modal properties based on the derivatives of the amplitude and phase of the nonstationary oscillation representing the structural response. A typical approach for determining the required derivatives utilizes the Hilbert Transform to first estimate the amplitude and phase. While simple and efficient, this technique usually requires significant pre- and post-processing and user interaction in the form of filtering and smoothing operations to ensure the derivatives are not corrupted by noise and will produce accurate natural frequency and damping ratio estimates. However, these difficulties can be avoided by utilizing the optimization-based NLSID technique proposed in this work to form best-fit polynomial representations of the amplitude and phase of each modal response. This approach directly produces smoothed results at every time point in the fit and can properly handle sections of the response that would typically be corrupted and discarded in Hilbert Transform methods. The results presented by the proposed NLSID technique represent a notable improvement over those from a comparable Hilbert Transform based approach.

3.3.3. Application to Experimental Forced Response

Generating high amplitude structural response that significantly activates nonlinearity in bolted joints typically requires the use of forced excitation from an electrodynamic shaker. These shakers are constructed similarly to a loudspeaker, in which a coil of wire is positioned around a central permanent magnet. Passing an amplified voltage signal through the wire generates an electromagnetic field that

interacts with the magnet to produce axial motion. In a speaker, this is used to create soundwaves by driving a lightweight cone that is attached to the magnet. Shakers instead feature a stiff armature with a connection point through which axial force from the electromagnetic interactions can be transmitted to an external structure.

By altering the supplied voltage signal, the shaker output can be tailored to many different experimental scenarios such as random noise excitation for linear tests and harmonic forcing for nonlinear data. Impact hammer testing does not offer the same level of adaptability, as the main variable there is the bandwidth, or how high in frequency the impulse excites. This is dictated by the hammer tip, with stiffer materials reaching higher frequencies. Consequently, hammer testing will always be functionally limited to providing broadband excitation. While this has the potential to excite many resonances and provide free decay measurements that are relatively simple to analyze, the energy from the impulse is distributed over a wide frequency band, such that each modal response is only excited to a relatively low amplitude. This can be problematic for characterizing nonlinear structural behavior as it will be weakly excited. Conversely, in shaker testing a harmonic voltage signal can be used to focus the excitation into a narrow frequency band. By appropriately defining the frequency of the forcing, the energy supplied by the shaker can be concentrated into individual resonances, driving them to the high response amplitudes required to significantly activate the nonlinearity in the structure. However, a common difficulty with this form of nonlinear testing is that the shaker connection introduces a potentially significant boundary condition into the dynamic response of the structure. Properly accounting

for this and correcting the effect it has on the identified nonlinear parameters is not always simple or straightforward.

This section presents a case study demonstrating how the proposed curve fitting procedure can be used to identify nonlinear characteristics of an experimental system subjected to forced excitation. The specific structure explored here, known as the Cylinder-Plate-Beam (CPB), is a bolted joint testbed that has been used extensively in nonlinear dynamics research at Sandia National Laboratories. Several studies involving the CPB are detailed in [9], [37], [38], [126]. The CPB, shown in Figure 43, is an assembly of three aluminum components. As per the name, these include a hollow cylinder, a circular plate, and a rectangular beam. The plate is attached to an open end of the cylinder via eight steel bolts, forming a large continuous jointed interface. Extending from the center of the plate, the beam is connected via two bolts and an epoxy resin. The predominant source of nonlinearity in the dynamic response of the structure is assumed to arise from the distributed interface between the cylinder and plate. The primary means of exciting this area is through the cantilever modes of the beam, which cause the plate to flex and slip against the cylinder face.

The data analyzed in this section was collected from the experimental setup depicted in Figure 43. The CPB is suspended from bungees and paracord to approximate free boundary conditions. An array of 26 triaxial accelerometers are positioned across the surface of the structure to measure the dynamic response. The precise location of each was selected using finite element model mode shapes to ensure each resonance under 1600 [Hz] can be independently observed. Excitation

is provided by an electrodynamic shaker connected to the side of the cylinder opposite the plate and the beam, as highlighted by the red circle in Figure 43. Physically linking the shaker armature to the cylinder is a stinger, which is a length of 1/8" diameter cylindrical steel wire. The axial force transmitted between the shaker and CPB through the stinger is measured by a force transducer mounted to the side of the cylinder.

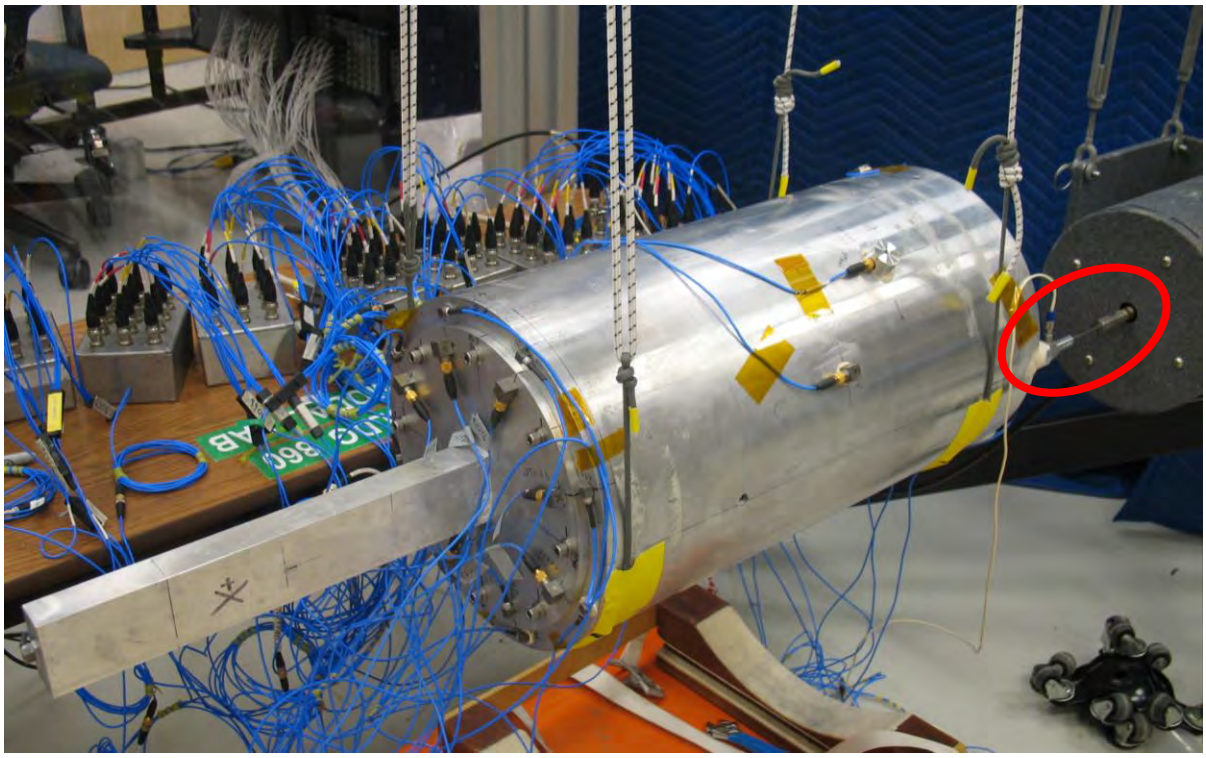

Figure 43: The Cylinder-Plate-Beam structure and the experimental setup for applying excitation and recording the response. The connection between the shaker and CPB is circled in red.

As a precursor to the nonlinear experiments, low level burst random tests were done to determine the linear modal parameters of the CPB. These use a voltage signal composed of short pulses of very low amplitude broadband noise. Averaging many instances of the ensuing response produces clean, linear data that can be processed with linear modal analysis techniques [1]. The specific algorithm used here was SMAC [129]–[131], developed by Mayes at Sandia National Labs. The primary purpose of

these tests was to determine the linear mode shapes of the structure for use in modal filtering the nonlinear test data. Assuming the mode shapes are constant, the process of modal filtering condenses the measured physical signals into a set of approximately uncoupled SDOF modal responses and forces [132], [133]. Additionally, the linear natural frequency and damping ratio identified from these low-level tests establish a baseline point of comparison for the quasi-linear modal parameters determined from the high-level nonlinear response data. This case study investigates the first elastic mode of the CPB, which is essentially a first order cantilever mode of the beam. From the linear test data, this mode was found to have a natural frequency of 128.4 [Hz] and damping ratio of 0.43 [%].

Nonlinear experimental measurements were collected by exciting the CPB with harmonic forcing from the shaker in the form of sine beats. As detailed previously in the numerical case study, sine beats excite a narrow frequency band, allowing for the energy supplied by the shaker to be focused into individual resonances, driving them to high response amplitudes. In that idealized numerical scenario, the nonlinear system exhibits pure free decay once the sine beat force goes to zero. This can only be approximately emulated in the experimental setup used here as the stinger connects the CPB and shaker throughout the excitation and subsequent response. To account for the effects of this boundary condition on the free response of the CPB, the force transmitted through the stinger during the ringdown must be recorded such that FORCEVIB can be used for the entire duration of the measurement.

Through trial and error, it was determined that the largest modal response was generated by defining a sine beat centered at 118 [Hz] with a bandwidth of 8 [Hz]. After modal filtering the measured response and force, the amplitude and phase of both were determined via the nonlinear optimization process detailed in Section 3.2. For the force, cubic polynomials were fit to the amplitude and phase in time segments 0.05 [s] long with 90% overlap. The same parameters were used to fit the response, except for slightly larger segments of 0.075 [s]. As the response and force are primarily composed of a single dominant oscillation, the nonlinear least squares operations could be performed in the time domain, minimizing the residual directly between the time vectors in each segment. The modal force and the resultant curve fit are shown below in Figure 44, while the modal acceleration and curve fit result are in Figure 45. For this measurement, the shaker sine beat w.as actively exciting the CPB structure from 0.1 to 0.5 seconds. After 0.5 [s], the response continues to smoothly decay while the force exhibits a low amplitude oscillation due to the CPB motion transferring energy through the stinger back to the shaker. The accuracy of the curve fits to the force and response are most evident in the frequency domain plots. In each, the fit has cleanly extracted the fundamental harmonic of the signal, effectively filtering out the higher order components. While the higher harmonics are part of the nonlinear dynamics of the excited system, they cannot be modeled by the averaged quasi-linear modal parameters used here. Consequently, in terms of the FORCEVIB identification method, the harmonics are essentially noise that will only corrupt the results and should be removed from the signal.

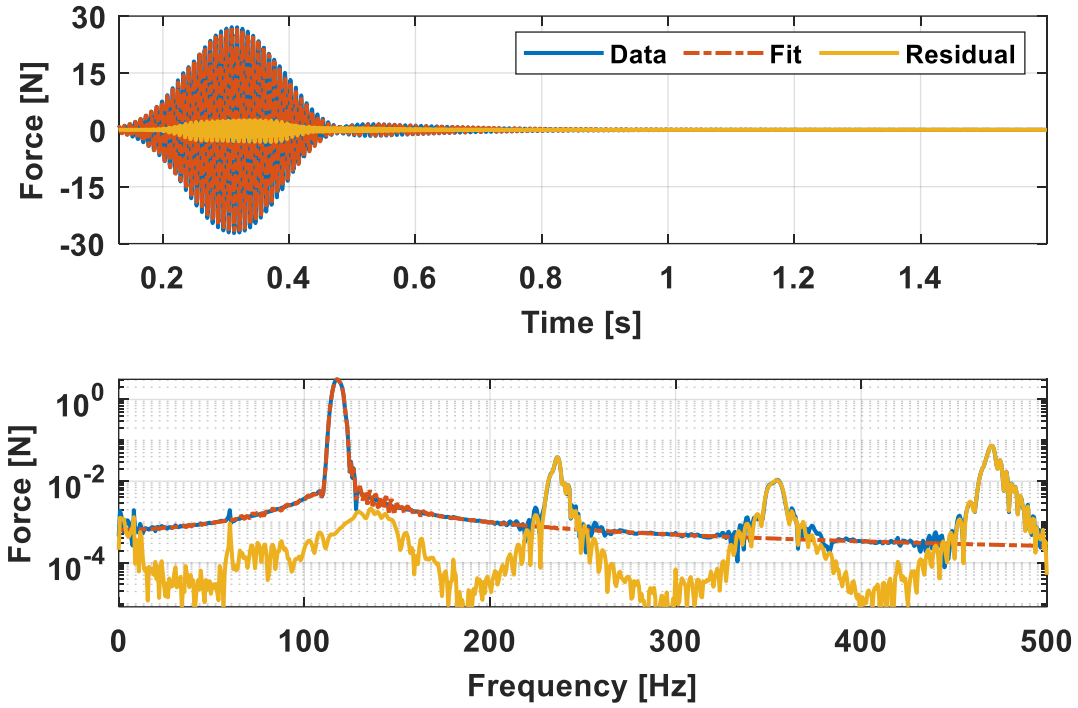

Figure 44: The recorded modal force applied to the CPB and the result of curve fitting the amplitude and phase of the fundamental harmonic.

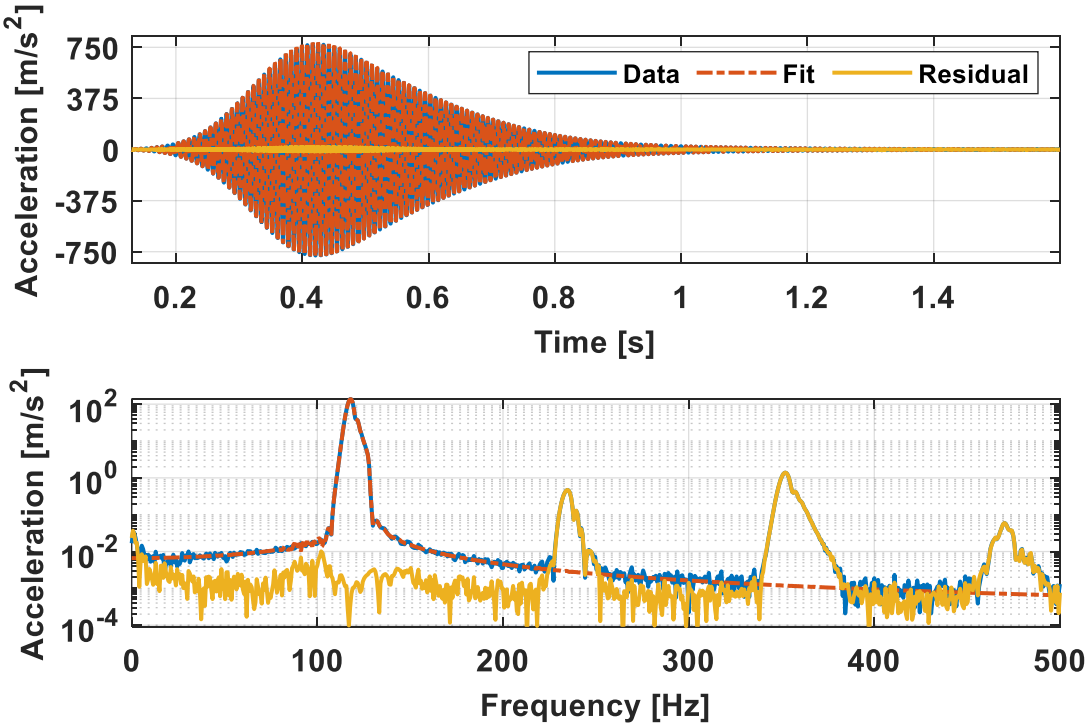

Figure 45: The measured modal acceleration of the first CPB mode and the result of curve fitting the amplitude and phase of the fundamental harmonic.

From the optimized curve fits to the modal force and response, the quasi-linear natural frequency and damping ratio can be determined with Eq. (112) and Eq. (113), respectively. These results are given in Figure 46, in which the natural frequency is on the left, the damping ratio on the right, and verses time on the top, and verses modal velocity amplitude on the bottom. In each plot, the linear value from the low level burst random test is given by a horizontal black line, the result from FREEVIB is in blue, and the curve from FORCEVIB is in orange. The result from FREEVIB has been included for comparison because the original intent with this experimental data was to simply assume that the response could be approximated as free decay once the sine beat ends and the shaker no longer actively excites the CPB. This approach involves simpler data analysis since only the response is considered, but that is at the cost of essentially ignoring the boundary condition imposed by the stinger and shaker. As FORCEVIB incorporates the measured force transference into the quasi-linear model, the boundary condition can be appropriately accounted for.

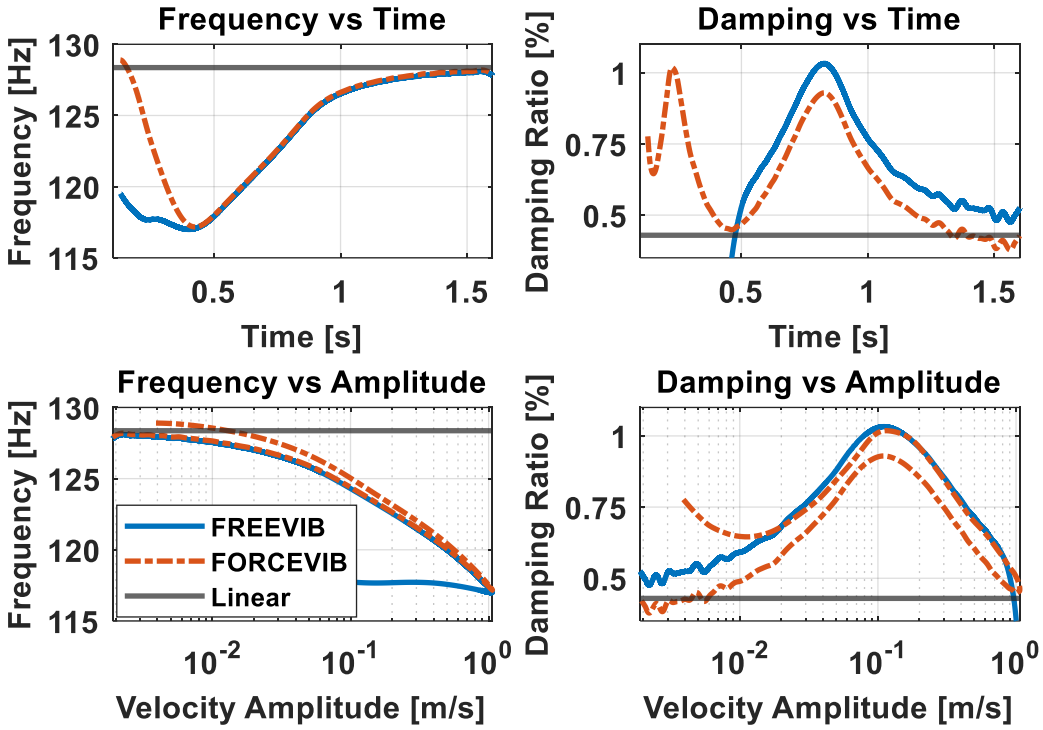

Figure 46: Quasi-linear Natural Frequency and Damping Ratio of mode 1 of the CPB as functions of time and modal velocity amplitude.

For the natural frequency in the left plots, the result from FORCEVIB converges toward the linear value during both the initial ramp up and subsequent ringdown and traces out a consistent amplitude dependent backbone curve. While the result from FREEVIB closely agrees with FORCEVIB after the sine beat ends at 0.5 [s], it significantly diverges from it prior to that point highlighting the effect the active forcing has on the identified parameter. In the damping verses time plot, before 0.5 [s] FREEVIB gives a nonphysical negative damping ratio due to the increasing response amplitude. After 0.5 [s] FORCEVIB and FREEVIB exhibit similar curves, with FORCEVIB giving a lower estimate that seems to converge closer to the linear value. This is due to FORCEVIB accounting for the energy transferred back into the shaker during the ringdown, whereas FREEVIB attributes that energy loss to additional

dissipation in the structure, and consequently overestimates the damping ratio. The damping is more affected by the boundary condition than the natural frequency because the stinger negligibly effects the stiffness of the CPB, while the shaker acts as a significant source of energy dissipation compared to the viscous and friction loses in the CPB structure. Lastly, in the lower right plot, the FORCEVIB damping ratio curve exhibits a similar trend during both the excitation and decay, except for an initial end effect. As observed previously in the numerical case study, this error is due to the polynomial fit performing poorly on the low amplitude data at the start of the measurement, where both the force and response are rapidly increasing, and the least squares fit is naturally biased toward accurately fitting the higher amplitude data.

In this case study, quasi-linear modal parameters were determined by curve fitting the amplitude and phase of forced response from an experimental structure subjected to sine beat excitation. It was shown that if the boundary conditions due to the shaker connection are neglected during the ringdown, FREEVIB yields acceptable natural frequency results but overestimates the damping ratio. Correctly accounting for the effects of the shaker requires the use of FORCEVIB. While this process is slightly more complicated, as the force must be accurately measured and curve fit, it can provide good results throughout the experiment, during both the active forcing and the subsequent ringdown.

3.4. Conclusion

This chapter detailed a new NLSID technique that utilizes nonlinear optimization to determine the amplitude and phase of a nonstationary oscillation. Using local regression, low order polynomial representations of the amplitude and phase are formed in short segments of the overall signal. While minimizing the residual between the optimized oscillation and the original data can be completed directly in the time domain, equivalent calculations may be performed in the frequency domain with Fourier coefficients. This approach can be leveraged to limit the curve fit to only considering information that is relevant to the current oscillation of interest. Additionally, by windowing the time data to suppress leakage in the Fourier coefficients, individual signals can be extracted from data containing many superimposed oscillations, regardless of their respective amplitudes and without the use of filters or other techniques for decomposing the data. From the optimized polynomials, analytical expressions can be used to form the integrals and derivatives of the fit oscillation and produce the quasi-linear natural frequency and damping ratio via a formulation of FREEVIB/ FORCEVIB. Three case studies were then presented, demonstrating how this NLSID method can be applied to various scenarios and the accuracy of the results in each.

The first explored a numerical study of a modal Iwan model excited by a sine beat, such that it exhibits sections of both forced and free response. After curve fitting the applied force and the ensuing response, the resultant quasi-linear natural frequency and damping ratio curves agreed extremely well with the truth data for the

Iwan model, except for an initial end effect. This error is due to the least squares process more heavily weighting data that is higher in amplitude. Since the force and response start from zero and rapidly increase, the earliest section of each will be the least accurately represented by the best-fit polynomials.

In the second case study, measured experimental response from an impact test of the S4 Beam structure was analyzed. While this data contains many superimposed oscillations, it was demonstrated that individual modal responses can be accurately extracted and characterized with the curve fitting procedure. The resultant frequency and damping curves were compared to results from an algorithm based on the Hilbert Transform, showing how the optimized values provide more information and avoid the problematic end effects that arise during the computation of the Hilbert Transform.

The final case examined experimental forced response in the form of exciting the CPB structure via sine beats from an electrodynamic shaker. It was demonstrated that, by curve fitting both the measured modal force and response, the FORCEVIB method can be used to correctly account for the boundary condition imposed by the connection with the shaker. The optimization results traced out consistent frequency and damping backbone curves during both the ramp up and ringdown, except for an initial end effect, similar to what was observed in the numerical Iwan case.

The proposed NLSID method exhibits several notable improvements over comparable approaches. The Hilbert Transform is limited to accurately handling only monoharmonic signals that are periodic and smooth, such that significant noise and

end effects arise if these conditions are not met. Meanwhile, by determining the amplitude and phase via optimized polynomial expressions, the results are guaranteed to be smooth. This also functionally limits the curve fit results to what can be modeled by relatively low order polynomials, which can be problematic when the actual properties of the oscillation of interest vary more rapidly than can be emulated by the defined polynomials. However, this is likely a fringe case in structural dynamics NLSID since most testing methods that produce harmonic response, either free from impacts or forced from sine beats, sweeps, or dwells, do not also cause abrupt or rapid changes during the response. Thus, this proposed NLSID technique should be a useful addition to the toolbox of existing methods, presenting an accurate alternative for use in a wide variety of nonlinear response scenarios. An avenue for future work is to adapt the solution process to directly optimizing the amplitude and phase in terms of spline curves. This framework would automatically generate a cohesive, global result from the local regression fits, instead of the approach detailed here of individually computing a polynomial for each segment and averaging the overlapping results.

4. Quasi-Linear Parameters from Least Squares Regression

4.1. Introduction

The ability to leverage nonlinear behavior is becoming ever more necessary as modern structures face increasingly demanding design requirements and challenging environments. However, to take advantage of nonlinearity, it must first be experimentally characterized and modelled. This is accomplished through nonlinear system identification (NLSID) techniques, which process measured data and fit it to some prescribed model form. In structural dynamics, a common approach is to represent observed nonlinear behavior in terms of quasi-linear modal parameters. These quantities characterize variations in the stiffness and energy dissipation of the structural resonances as amplitude dependent extensions of the standard linear natural frequency and damping ratio.

Many NLSID techniques for determining these quasi-linear parameters are based on Feldman's FREEVIB [62] and FORCEVIB [63] methods. These approaches are derived from the single-degree-of-freedom (SDOF) quasi-linear modal equation of motion (EOM) shown in Eq. (1), where $\ddot{x}(t), \dot{x}(t), x(t)$ & $f(t)$ are the acceleration, velocity, displacement, and external force, respectively. The quasi-linear natural frequency, $\omega_{\mathrm{n}}(A)$, and damping ratio, $\zeta(A)$, are functions of the amplitude of the response, $A(t)$.

$$\ddot{x}(t) + 2\zeta(A)\omega_{\mathrm{n}}(A)\dot{x}(t) + \omega_{\mathrm{n}}^2(A)x(t) = f(t) \tag{120}$$

In Feldman's derivations, the quasi-linear natural frequency and damping ratio are determined by assuming that the response and force are represented in a complex exponential form in terms of a time varying amplitude and phase. This means that the EOM in Eq. (1) can be split into real and imaginary components that can be rearranged to yield expressions for $\zeta(A)$ and $\omega_{\mathrm{n}}(A)$, where FREEVIB and FORCEVIB are versions of the equations for free and forced response, respectively.

The primary difficulty in applying these methods is that the equations for $\zeta(A)$ and $\omega_{\mathrm{n}}(A)$ depend on the derivatives of the amplitude and phase of the response. In addition to complicating the theoretical presentation of the method, an inherent issue with this process is that amplitude and phase are not intrinsic, unique properties of a real signal. As there are an infinite number of amplitudes and phases that can be combined to yield a specific real component, estimating them from a real signal will always carry some degree of uncertainty and difficulty. This is highlighted with regard to the Hilbert Transform, which is very commonly used to estimate amplitude and phase, in that the results can be significantly corrupted by noise if the signal is not monoharmonic, smoothly varying, and periodic [60]. Thus, this process typically also requires the use of filtering and smoothing methods to produce comprehensible results that are physically meaningful [104], [128]. Some approaches attempt to avoid these issues by foregoing the Hilbert Transform, such as the Peak Finding and Fitting (PFF) algorithm [56]. In this method, the change in the location and height of successive oscillation peaks in a free decay signal are used as discrete, local

estimates of the amplitude and phase derivatives. A process shown in [57] is similar, but estimates the phase derivative based on time between zero-crossings.

This work details a new NLSID method that functionally fills the same role as FREEVIB and FORCEVIB but operates directly on the measured response and force signals instead of estimated amplitude and phase values. When utilizing the quasi-linear EOM as the model form in a NLSID method, the identification process is a matter of determining two unknowns, the natural frequency and damping ratio, from the single equation, the EOM. While Feldman did this by assuming a complex form to get a second equation, the alternative explored here directly determines the two unknowns from the single equation via linear least squares regression, entirely bypassing the need to estimate the amplitude and phase. In this new approach, denoted as Quasi-Linear - Least SQuares (QL-LSQ), the quasi-linear stiffness and damping coefficients on the velocity and displacement are modelled as B-spline curves that are simultaneously fit with least squares to the nonlinear restoring force, which is the difference between the force and the acceleration. The B-spline framework is incredibly adaptable. An extension highlighted here is penalized B-splines, which improve the noise resilience of the curve by adding a roughness penalty to the spline. Additionally, the B-spline basis can be defined in terms of the amplitude of the response, such that the amplitude dependence of the quasi-linear parameters can be directly determined from the data.

This NLSID process bears a strong resemblance to that of the Restoring Force Surface (RFS) method [28], in that linear least squares regression is used to fit a

model to the nonlinear restoring force of the system. However, the model form assumed in each method is drastically different. In RFS, a model is typically composed of nonlinear polynomials of the displacement and velocity which are assumed to be representative of the nonlinearity present in the system. Meanwhile, QL-LSQ expands the coefficients of the displacement and velocity into spline curves that characterize the nonlinear behavior exhibited by the restoring force in terms of effective quasi-linear stiffness and damping values. There are other methods that have been proposed for determining quasi-linear modal parameters that are also alternatives to FREEVIB and FORCEVIB. A related approach from Sapsis et al. [134] determines effective stiffness and damping values based on an estimate of the energy in a system during unforced response. However, their approach is dependent of forming amplitude envelopes and their derivatives, making it susceptible to some of the same difficulties as FREEVIB. Forced response is a very active area of experimental nonlinear structural dynamics. A prominent method is phase resonance testing, in which control modules are used to tune an applied harmonic excitation such that it tracks the nonlinear behavior of individual resonances and drives them to high response levels. While the quasi-linear natural frequency is directly the excitation frequency, several different methods for determining the quasi-linear damping ratio have been proposed. One is the Extended Periodic Motion Concept from Krack [135] in which the damping is inferred from a ratio of the applied force to the resultant velocity in a steady state response. This method was investigated experimentally in [136], demonstrated on system undergoing base excitation in [137], and explored with regard to different control methods in [138]. A

second method is an extension of the Single Nonlinear Resonant Mode formula by Kwarta in [139] in which the quasi-linear natural frequency and damping ratio are determined based on an iterative relation between the amplitude of the response, the forcing frequency, and the relative phase difference of the force and response at steady state.

The remainder of this chapter is structured as follows. The following section, 4.2, establishes the theoretical basis for the QL-LSQ method. As splines are a key component of the process, the first portion provides a background on the formulation of penalized B-splines and details how these are used for curve fitting. This is followed by the main development of this chapter, in which the B-spline framework is utilized for determining quasi-linear modal parameters directly from the applied force and measured response of a nonlinear system. After defining the components of QL-LSQ, it is demonstrated in 4.3 by revisiting the test cases that were explored in the previous chapters. This includes the numerical case study of the Iwan Model excited by a sine beat, the impact response of the S4 beam, and the forced response of the CPB. In each of these cases, the quasi-linear parameters produced by QL-LSQ are compared to those computed by existing methods that utilize implementations of FREEVIB and FORCEVIB. Section 4.4 details an extension of the CPB test case in which QL-LSQ is used to illustrate the effects of modal coupling between the excited resonances. Finally, some conclusions as to the efficacy of using QL-LSQ for characterizing nonlinear structural behavior are given.

4.2. Theoretical Derivation of QL-LSQ

4.2.1. Penalized B-splines

A spline is a piecewise polynomial representation of a function. Splines are particularly useful for constructing curves that are complicated on a global scale but can be accurately represented by relatively low order polynomials at a local level. An efficient framework for constructing splines is through the use of B-splines, or basis splines, in which the complete curve is formed from a set of basis functions. A general process for generating the basis functions is known as De Boor's algorithm [140]. The intricacies of the algorithm are beyond the scope of this work as it is already widely implemented in many programming packages, such as the *spcol* function in the MATLAB Curve Fitting Toolbox. To briefly touch of the parameters of the algorithm, the position and shape of the basis functions is dictated by a sequence of locations called knots and the desired polynomial order of the spline. Each basis function is itself a simple spline curve that spans as many knots as the order of the spline with piecewise polynomials of degree one less than the spline order. To form the spline curve, the individual basis functions are scaled by as associated coefficient called a control point and then summed. In this process, the basis functions are blending functions that determine how the spline curve transitions between the control point values. This is shown in Eq. (121), where $\boldsymbol{\mu}[m]$ is the assembled spline curve, α_n are the control points, and $\mathbf{b}_n[m]$ are the accompanying basis functions. The number of sample points in the generated spline curve is M with indices $m = 1{:}\,M$, and in the summation, there are N B-spline bases with indices $n = 1{:}\,N$.

$$\boldsymbol{\mu}[m] = \sum_{\boldsymbol{n=1}}^{\boldsymbol{N}} \alpha_n \mathbf{b}_n[m] \tag{121}$$

A convenient alternative representation of the assembled spline is in the matrix formulation given in Eq. (122), in which the basis functions are the columns of the matrix $\mathbf{B}$ and the control points are collected into the vector $\boldsymbol{\alpha}$. The dimensions of $\mathbf{B}$ are $[M \times N]$, and $\boldsymbol{\alpha}$ are $[N \times 1]$.

$$\boldsymbol{\mu}[m] = \mathbf{B}\boldsymbol{\alpha} = \begin{bmatrix} b_1[1] & b_2[1] & \cdots & b_N[1] \\ b_1[2] & b_2[2] & \cdots & b_N[2] \\ \vdots & \vdots & \ddots & \vdots \\ b_1[M] & b_2[M] & \cdots & b_N[M] \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \tag{122}$$

The components of the B-spline formulation are demonstrated in Figure 47, which shows a 4th order spline on a uniform distribution of 16 knots. These parameters produce a matrix $\mathbf{B}$ containing 12 basis functions which are given as the multicolored, bell-shaped curves that each span four knots and are composed of third-degree polynomial pieces. The control points in $\boldsymbol{\alpha}$ are shown as the like-colored dots directly above the peak of the associated basis function. In this example all of the control points have the default value of 1, such that the summation of the basis functions produces the trivial spline curve represented by the black dash-dot line. Constructing more complex curves is done by simply varying the control point values. As the basis functions are defined over a limited range of knot locations and smoothly blend together in the summation, the spline can be controlled at a local level by each control point while also exhibiting a continuous second order derivative. However, this sort of precise control over the curve is only possible between knots 4 and 13, marked by the vertical black lines. In this central region of the spline, it is sufficiently supported by the

overlapping basis functions to produce any piecewise cubic polynomial segment between the knots. The local control of the spline deteriorates beyond this range as the density of superimposed basis functions reduces.

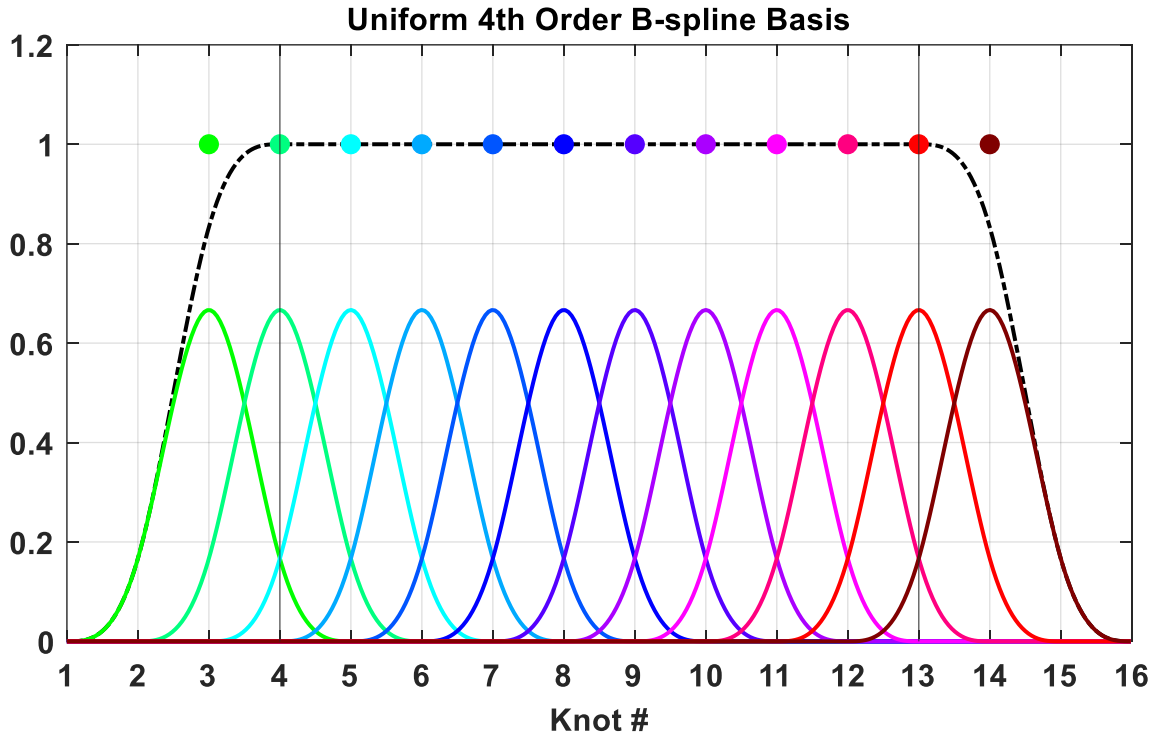

Figure 47: The components of a uniform 4th order B-spline. Each basis function is given as a bell-shaped curve and its associated control point is a same-colored dot positioned directly above the peak. These form the spline curve shown as the black dash-dot line. The usable range of the spline is between the vertical black lines where it is properly supported by the basis functions.

A common application of B-splines is curve fitting, in which a spline is constructed to approximate a set of M sampled data points, $\mathbf{y}$.

$$\mathbf{B\alpha} \cong \mathbf{y} \tag{123}$$

This is accomplished by first choosing a set of knot locations and the desired spline order to form the associated basis functions in $\mathbf{B}$. These can be fit to $\mathbf{y}$ by determining control point values, $\boldsymbol{\alpha}$, with respect to the objective function shown in Eq. (124). The

first term in this expression quantifies the residual between the data and the spline curve, where $\mathbf{W}$ is a diagonal matrix of positive scalar weights. This can be used to influence the fit by amplifying or suppressing residual values at certain samples, such as increasing the importance of key data points or removing outliners. By default, the weights are all 1 and $\mathbf{W}$ is an identity matrix that has no effect of the residuals.

$$Q = \|\mathbf{W}(\mathbf{y} - \mathbf{B}\boldsymbol{\alpha})\|^2 + \boldsymbol{\lambda}\|\mathbf{D}\boldsymbol{\alpha}\|^2 \tag{124}$$

The inclusion of second term, $\boldsymbol{\lambda}\|\mathbf{D}\boldsymbol{\alpha}\|^2$, was proposed by Eilers and Marx [141] as a means of efficiently enforcing a smoothing onto the spline curve. They refer to this formulation as penalized B-splines, or P-splines, as it penalizes the roughness of the control point values in $\boldsymbol{\alpha}$. The form of the smoothness is dictated by the penalty matrix $\mathbf{D}$ and the strength of the penalty is scaled by the parameter λ. While $\mathbf{D}$ can be customized to enforce a wide variety of penalty formulations, the most straightforward approach is a difference relation between adjacent control points. Penalty matrices that apply a first and second order difference are shown in Eq. (125). A simple method for generating these matrices is creating an identity matrix of size N and recursively applying a first order difference computation to each column. In MATLAB notation this is represented as $\mathbf{D}_1 = diff\big(eye(N)\big)$, $\mathbf{D}_2 = diff(\mathbf{D}_1)$, and so on for higher order difference matrices.

$$\mathbf{D_1} = \begin{bmatrix} \mathbf{-1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{-1} & \mathbf{1} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \qquad \mathbf{D_2} = \begin{bmatrix} \mathbf{1} & \mathbf{-2} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{1} & \mathbf{-2} & \mathbf{1} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-2} & \mathbf{1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{125}$$

These forms of the difference matrix are used in the definition of P-splines by Eilers and Marx. However, for this formulation of the penalty to be valid, the spline must be a uniform B-spline where the basis functions are constructed from an equally spaced sequence of knot locations, as demonstrated in Figure 47. To account for an arbitrary distribution of knots, work by Li and Cao [142] generalizes this process by showing that a weighting based on the relative position of the knots can be included in the computation of difference matrix so that it is valid for non-uniform B-splines. This method of applying discrete penalties to the control point values is computationally simpler than the more conventional approach to smoothing a spline curve by evaluating and minimizing its second derivative. Additionally, the functionality of the penalty is dependent on the number of basis functions and not on the order of the spline, allowing for the difference order of the penalty matrix to be higher than the spline order. For example, while first and second order splines do not have a second derivative that can be minimized to smooth the curve, a second order difference matrix can be easily defined that enforces the associated smoothing.

With each term in Eq. (124) defined, the optimal control point values that minimize the objective function can be determined via the linear least squares regression solution shown in Eq. (126).

$$\boldsymbol{\alpha} = (\mathbf{B}'\mathbf{W}\mathbf{B} + \lambda\mathbf{D}'\mathbf{D})^{-1}\mathbf{B}'\mathbf{W}\mathbf{y} \tag{126}$$

The quality of the spline curve formed from these values in $\boldsymbol{\alpha}$ is dictated by the user-defined parameters of the P-spline formulation, which are the weights, the spline order and knot sequence, and the penalty matrix and scalar. Eilers and Marx offer guidance

on selecting these parameters in [141]. They suggest that a 4^{th} order spline with a second order difference penalty matrix will be generally applicable to a wide range of applications as it exhibits a continuous second derivative that is smoothed by the discrete penalty on the control points. The most straightforward means of influencing the resultant curve is by varying the significance of the penalty with λ. If the value of λ approaches zero, the penalty is removed from the curve, and it becomes a standard B-spline fit. At the other extreme, if λ is set to be arbitrarily large then the P-spline converges to whatever type of curve is modelled by the penalty matrix. The first and second order difference matrices in Eq. (125) for instance lead to a constant-valued horizontal line, and a line with constant slope, respectively. Several methods for automatically determining a reasonable value for λ are described in [141].

Eilers and Marx stress that the most significant advantage of P-splines over B-splines is the relative importance of the knot sequence. When using standard B-splines, the resultant curve is heavily dependent on setting appropriate knot locations that adequately capture the behavior of the data. Significant effort is typically required when attempting to optimize the number of knots and their location. However, Eilers and Marx argue that this can be avoided with P-splines by simply utilizing a dense distribution of knots that form a rich basis that can be controlled via the discrete penalty on the control points. If this type of dense knot sequence is used for standard B-splines, basis functions can be generated in a range that does not contain any data points, leading to a column of zeros in $\mathbf{B}$ and $\mathbf{B}'\mathbf{W}\mathbf{B}$ in Eq. (126) to be singular. P-splines do not suffer from this numerical instability since the penalty ensures that every

basis function is at least supported by the adjacent control points and $\mathbf{B}'\mathbf{W}\mathbf{B} + \lambda\mathbf{D}'\mathbf{D}$ is always full rank and invertible. Therefore, while any large number of knots could be used with P-splines, Eilers and Marx suggest that starting with 50 basis functions should provide a decently flexible spline curve while still being computationally efficient.

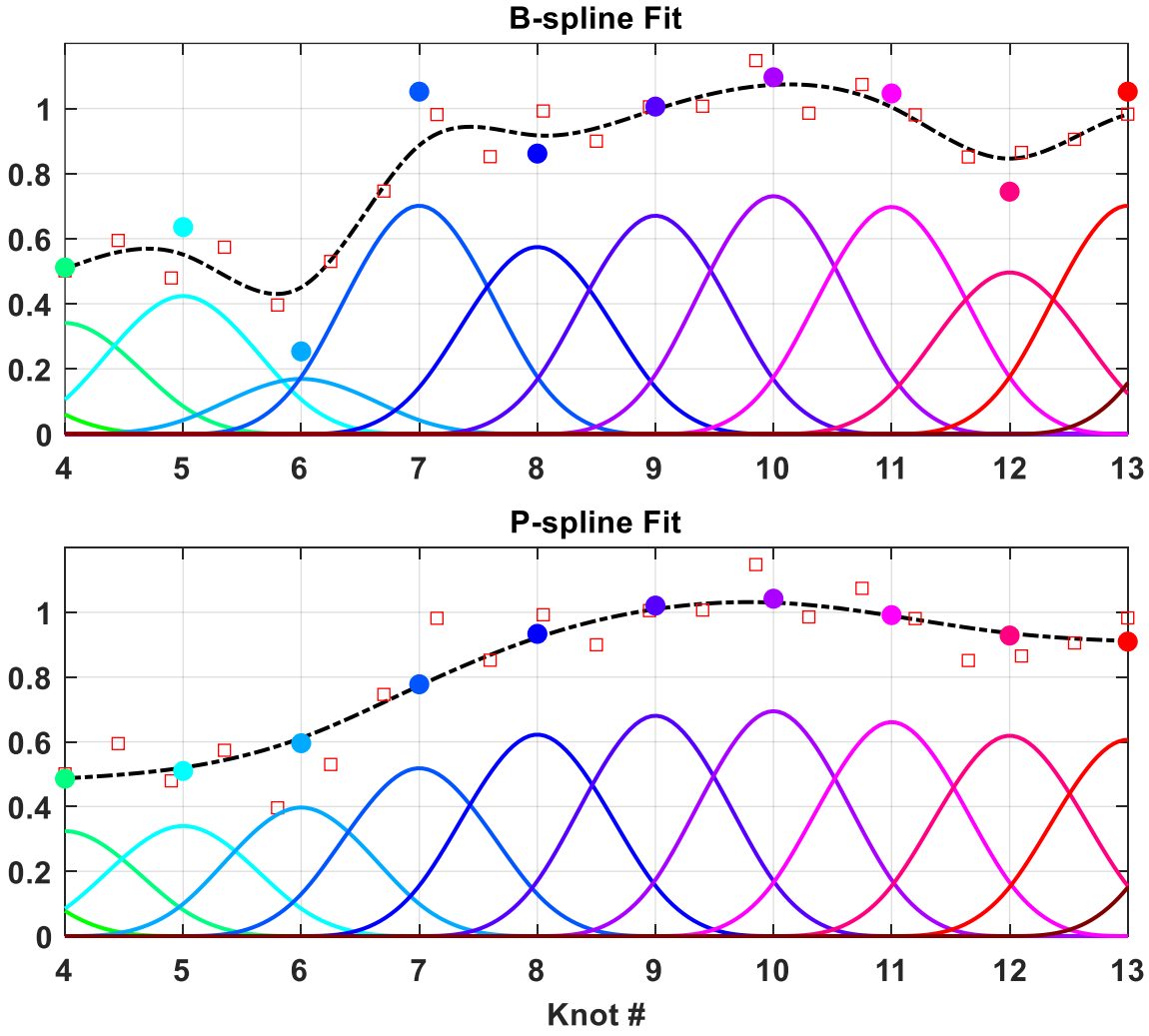

Figure 48: B-spline (top) and P-spline (bottom) curve fits to a distribution of data points. A second order difference penalty in the P-spline eliminates the large variations observed in the standard B-spline.

To demonstrate the differences of B-splines and P-splines, both are used to curve fit a set of data points as shown in the top and bottom plots in Figure 48. In this example, the 4th order uniform B-spline basis that was previously shown in Figure 47 is utilized to form the spline curves. The data points being curve fit are the same in both plots and are represented by the red squares. Note that the data is situated in the central region of the spline between knots 4 and 13, since beyond there the spline is poorly supported as explained above. While the standard B-spline fit in the top plot closely follows the data points, it exhibits large variations in the value of the control points and the resultant curve, especially near the sixth knot location. To determine the control points for the P-spline curve in the bottom plot, a second order difference penalty was used with $\lambda = 1$. This produces a curve that smoothly tracks through the data points, where the significant variations observed in the B-spline fit are eliminated by the penalty reducing the difference between adjacent control point values.

4.2.2. Determining Quasi-linear Modal Parameters with P-splines

The QL-LSQ method is based on representing quasi-linear modal parameters as splines. This NLSID technique is derived from the SDOF quasi-linear modal EOM in Eq. (1), in which the stiffness and viscous damping terms are assumed to be variables instead of constants. In order to determine values for these parameters, the other quantities in the expression, the acceleration, velocity, displacement, and external force, must be known. With experimental data, it is likely that only one of the response quantities is measured, such that numerical integration and/or differentiation

schemes must be used to compute the other two. Additionally, if the response and force are multimodal or multiharmonic, modal filtering and harmonic decomposition methods must be used to separate the data into distinct SDOF systems that can be individually characterized.

The model form that is fit to each SDOF response is shown in discrete form in Eq. (127), in which $m = 1{:}\,M$ are the sample indices and $\ddot{x}[m]$, $\dot{x}[m]$, $x[m]$, and $f[m]$ are vectors containing the measured system response and the applied force. In this expression, the viscous damping, $c[m]$, and stiffness, $k[m]$, are also in a discrete form that can vary from sample to sample. The circled dot operator, $\odot$, denotes element-wise matrix multiplication, also known as the Hadamard product.

$$\ddot{x}[m] + c[m] \odot \dot{x}[m] + k[m] \odot x[m] = f[m] \tag{127}$$

To determine the stiffness and damping, both are first represented in terms of B-spline basis functions and control points as shown in Eq. (128). Each B-spline formulation can be defined completely independently with distinct knot locations and spline order. In the expression, the damping bases is given as $\mathbf{B}_{\mathrm{c}}[m, n_{\mathrm{c}}]$ with control points $\alpha_{\mathrm{c}}[n_{\mathrm{c}}]$, where $n_{\mathrm{c}} = 1{:}\,N_{\mathrm{c}}$ are the indices of each basis. The stiffness is similarly defined, with $\mathbf{B}_{\mathrm{k}}[m, n_{\mathrm{k}}]$, $\alpha_{\mathrm{k}}[n_{\mathrm{k}}]$, and $n_{\mathrm{k}} = 1{:}\,N_{\mathrm{k}}$.

$$\begin{aligned} c[m] &= \mathbf{B}_{\mathrm{c}}[m, n_{\mathrm{c}}] * \alpha_{\mathrm{c}}[n_{\mathrm{c}}] \\ k[m] &= \mathbf{B}_{\mathrm{k}}[m, n_{\mathrm{k}}] * \alpha_{\mathrm{k}}[n_{\mathrm{k}}] \end{aligned} \tag{128}$$

Substituting this form of the stiffness and damping into Eq. (127) and moving the acceleration term to the right side produces:

$$(\mathbf{B}_{\mathrm{c}}[m, n_{\mathrm{c}}] * \alpha_{\mathrm{c}}[n_{\mathrm{c}}]) \odot \dot{x}[m] + (\mathbf{B}_{\mathrm{k}}[m, n_{\mathrm{k}}] * \alpha_{\mathrm{k}}[n_{\mathrm{k}}]) \odot x[m] \cong f[m] - \ddot{x}[m] \quad (129)$$

The left side of this expression can be consolidated into a matrix representation as shown in Eq. (130), where the basis matrices are concatenated into a single matrix that has dimensions $[M \times (N_{\mathrm{c}} + N_{\mathrm{k}})]$, and the associated control points are stacked into a vector with dimensions $[(N_{\mathrm{c}} + N_{\mathrm{k}}) \times 1]$. On the right side is the difference of the force and acceleration vectors, which is a quantity commonly known as the restoring force.

$$[\mathbf{B}_{\mathrm{c}}[m, n_{\mathrm{c}}] \odot \dot{x}[m] \quad \mathbf{B}_{\mathrm{k}}[m, n_{\mathrm{k}}] \odot x[m]] \begin{bmatrix} \alpha_{\mathrm{c}}[n_{\mathrm{c}}] \\ \alpha_{\mathrm{k}}[n_{\mathrm{k}}] \end{bmatrix} \cong f[m] - \ddot{x}[m] \quad (130)$$

The control point values for both the damping and stiffness are now simultaneously determined via linear least squares regression as shown in Eq. (131), which is an adaptation of the general solution for P-spline control points given in Eq. (126).

$$\begin{bmatrix} \alpha_{\mathrm{c}}[n_{\mathrm{c}}] \\ \alpha_{\mathrm{k}}[n_{\mathrm{k}}] \end{bmatrix} = (\mathbf{B}'\mathbf{W}\mathbf{B} + \mathbf{D}'\mathbf{D})^{-1}\mathbf{B}'\mathbf{W}(f[m] - \ddot{x}[m]) \quad (131)$$

In this expression, the restoring force is the data points that are being curve fit. The weighting matrix, $\mathbf{W}$, is an identity by default but can be set by the user to influence the fit to the restoring force. The B-spline basis matrix, $\mathbf{B}$, is as shown in Eq. (132) and is composed of the concatenated B-spline basis matrices for the damping and stiffness that are multiplied element-wise by the velocity and displacement vectors, respectively. In this scenario, the splines essentially form a best-fit local representation of the nonlinearity in the response in terms of an effective stiffness and damping.

These combine with the displacement and velocity to model whatever nonlinear behavior is exhibited by the restoring force.

$$\mathbf{B} = [\mathbf{B}_{\mathrm{c}}[m, n_{\mathrm{c}}] \odot \dot{x}[m] \quad \mathbf{B}_{\mathrm{k}}[m, n_{\mathrm{k}}] \odot x[m]] \tag{132}$$

To enforce a smoothing onto the damping and stiffness control points, the P-spline discrete difference penalty can be applied to each. As these are modelled by separate B-splines, independent difference matrices $\mathbf{D_c}$ and $\mathbf{D_k}$ can be formed for $\alpha_{\mathrm{c}}[n_{\mathrm{c}}]$ and $\alpha_{\mathrm{k}}[n_{\mathrm{k}}]$, respectively. The complete penalty matrix, $\mathbf{D}$, is constructed as shown in Eq. (133), in which the individual difference matrices are combined in a block diagonal form where $\mathbf{0}_{\boldsymbol{N_{\mathrm{c}}}}$ and $\mathbf{0}_{\boldsymbol{N_{\mathrm{k}}}}$ are appropriately sized zero matrices. The penalty scalars, $\sqrt{\lambda_{\mathrm{c}}}$ and $\sqrt{\lambda_{\mathrm{k}}}$, are inserted as square roots since the $\mathbf{D}'\mathbf{D}$ term in Eq. (131) will square the values.

$$\mathbf{D} = \begin{bmatrix} \sqrt{\lambda_{\mathrm{c}}}\mathbf{D_c} & \mathbf{0}_{\boldsymbol{N_{\mathrm{k}}}} \\ \mathbf{0}_{\boldsymbol{N_{\mathrm{c}}}} & \sqrt{\lambda_{\mathrm{k}}}\mathbf{D_k} \end{bmatrix} \tag{133}$$

With the control point values produced by the least squares fit in Eq. (131), the spline representations of the effective viscous damping, $c[m]$, and stiffness, $k[m]$, are formed as defined in Eq. (128). In standard linear structural dynamics [103], the viscous damping and stiffness are related to the natural frequency and damping ratio through $c = 2\zeta\omega_{\mathrm{n}}$ and $k = \omega_{\mathrm{n}}^2$. Assuming this relation can be extended to also describe quasi-linear modal parameters, Eq. (134) shows expressions for the quasi-linear natural frequency and damping ratio based on the computed spline forms of stiffness and viscous damping.

$$\omega_{\mathrm{n}}[m] = \sqrt{k[m]}$$
$$\zeta[m] = \frac{c[m]}{2\omega_{\mathrm{n}}[m]} \tag{134}$$

By default, the dependent variable for the sampled response and force is likely time, where successive indices in m are in chronological order and progress forward in time. If the B-spline knot sequences are defined relative to these points in time, then the resultant $c[m]$ and $k[m]$ spline curves are also formed with respect to time. While this representation of the nonlinearity as a function of time can be informative, it is often more useful to model the nonlinear behavior as quasi-linear parameters that are dependent on the response amplitude. In QL-LSQ, the amplitude dependent curves can be directly determined by defining the splines with respect to the amplitude of the response instead of time. A simple approach for this is to sort the response and force vectors relative to an identified vector of response amplitude values at each sample, $A[m]$, such that $m = 1$ is the sample with the lowest amplitude and $m = M$ has the highest amplitude. Setting the knot sequences relative to these reordered vectors then generates B-spline basis functions which span the amplitude range of the response, instead of the amount of time it was recorded over. This process could be utilized to generate splines that are dependent on any quantity that the response and force can be sorted with respect to, such as the instantaneous value of the response itself or the amplitude of other resonances. This aspect of the QL-LSQ method is highlighted in the next section, in which it is used to characterize a simulated nonlinear system.

4.3. Case Studies Revisited

4.3.1. Application to Forced Modal Iwan Model

To demonstrate the use of QL-LSQ on a nonlinear system with known truth data, this section revisits the Modal Iwan Model previously explored in Section 3.3.1. The same Iwan parameters listed in Table 1 and sine beat forcing and simulated response shown in Figure 32 are used here. By utilizing the same data set, the natural frequency and damping ratio results from QL-LSQ can be directly compared to the truth values and those from the implementation of FREEVIB/ FORCEVIB used in the optimization process detailed in Chapter 3. While that approach required both the response and force to be curve fit to estimate the amplitude and phase of each, the solution process for QL-LSQ is much more straightforward as it operates directly on the real valued response and force vectors. In this numerical case study, the acceleration, velocity, and displacement are all taken to be known from the simulation results. While this is unrealistic as precisely measuring all three in an experimental setting is incredibly uncommon, this is done to avoid ambiguity in the results due to errors introduced through numerical integration and/or differentiation.

To set up the integrated response and the force such that the quasi-linear parameters can be determined directly as B-splines with respect to amplitude, the vectors are sorted with respect to the velocity amplitude returned by the integration scheme [127]. This is shown in Figure 49, in which the displacement is shown with respect to time on the left, and with respect to the velocity amplitude on the right. When represented with respect to time, the response initially starts as linear and

rapidly becomes more nonlinear as the sine beat causes the response amplitude to increase. The system then slowly rings down and converges back toward linear behavior. By viewing the response as a function of the amplitude, the sorting process has effectively folded the initial ramp up and subsequent ring down portions of the sine beat response on top of each other. This allows for the ramp up and ring down portions of the response to be characterized simultaneously. Since the nonlinearity is amplitude dependent, it does not matter when the response occurred, only what the amplitude was at that time.

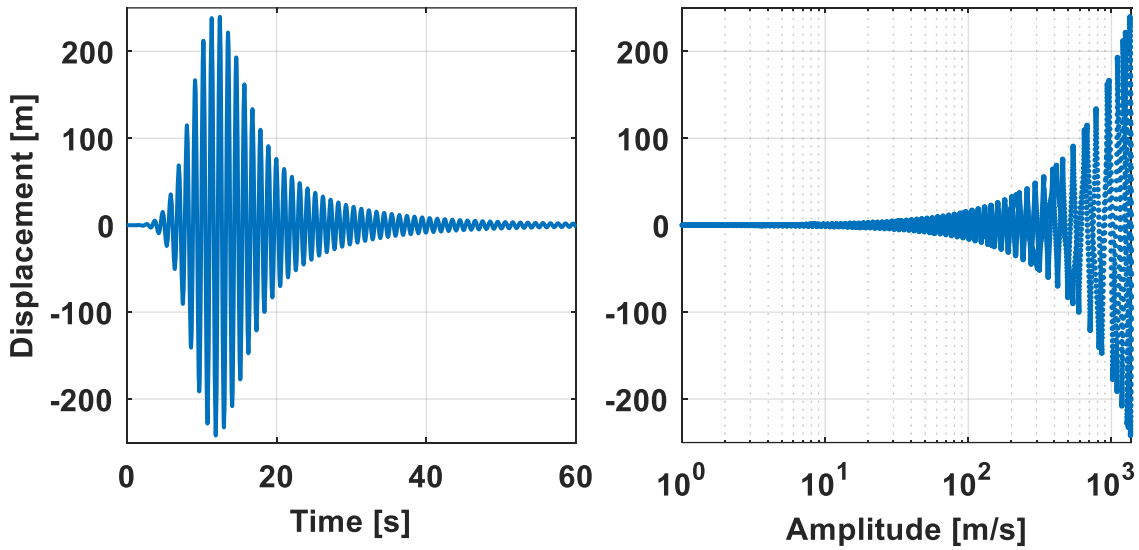

Figure 49: (Left) The displacement response of the Iwan Model vs time. (Right) The same response but sorted with respect to the velocity amplitude.

To determine the quasi-linear stiffness and damping of the simulated Iwan Model response with QL-LSQ, a set of B-spline basis functions were defined with respect to the velocity amplitude. In this case, the same basis was used for both the stiffness and damping curves and consisted of 4^{th} order basis functions constructed from a uniform sequence of 40 knots. The same penalty matrix was also used for both,

which enforced a second order difference penalty onto the control points with a scaler value of $\lambda = 0.001$. The control points defining the stiffness and viscous damping splines are then computed with least squares regression using Eq. (131). This produces the curves shown in Figure 50. Each B-spline basis and its associated control point are shown as the multicolored curves and dots, and the resultant spline curve is shown as a black dash-dot line.

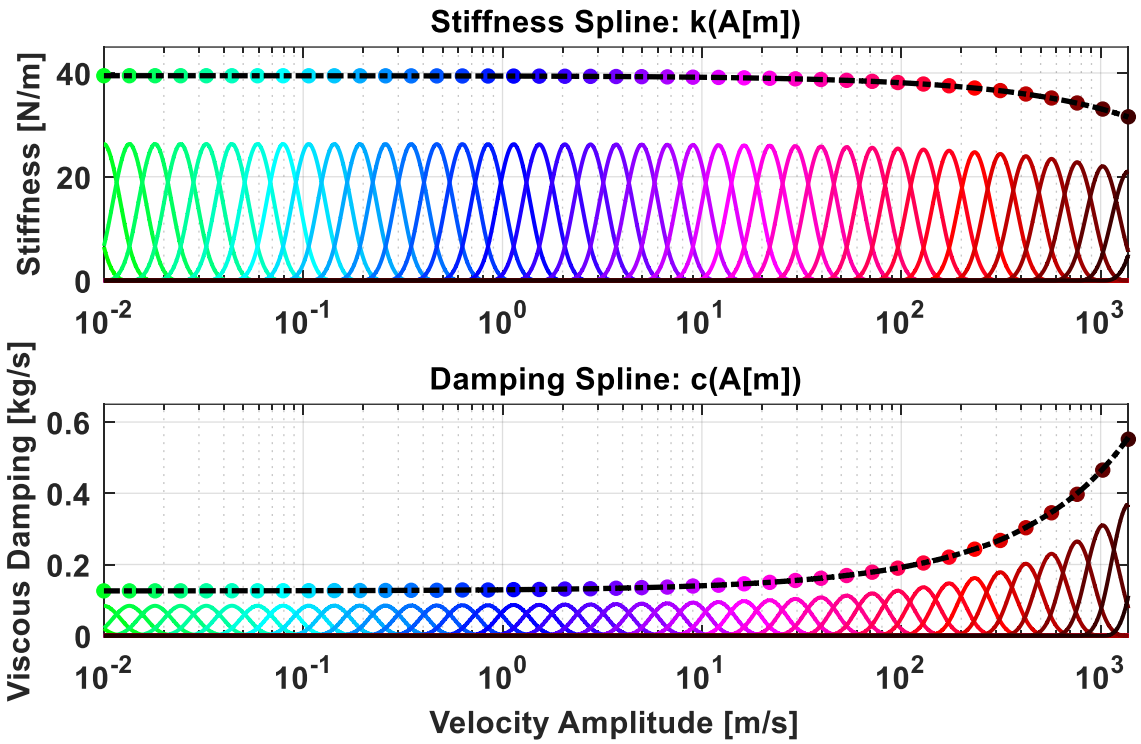

Figure 50: The resultant B-spline curves for the quasi-linear stiffness (top) and viscous damping (bottom) of the Iwan Model.

Using Eq. (134), the quasi-linear natural frequency and damping ratio can be recovered from these spline curves. The result is displayed in Figure 51, which repeats the data in Figure 37 but with additional dashed purple lines representing the QL-LSQ result. To represent the QL-LSQ results with respect to time, they are simply reordered

by reversing the sorting that was applied to the response and force vectors to order them as functions of amplitude. In each plot, the QL-LSQ curves accurately overlay with the true values for the nonlinear behavior of the simulated Iwan Model.

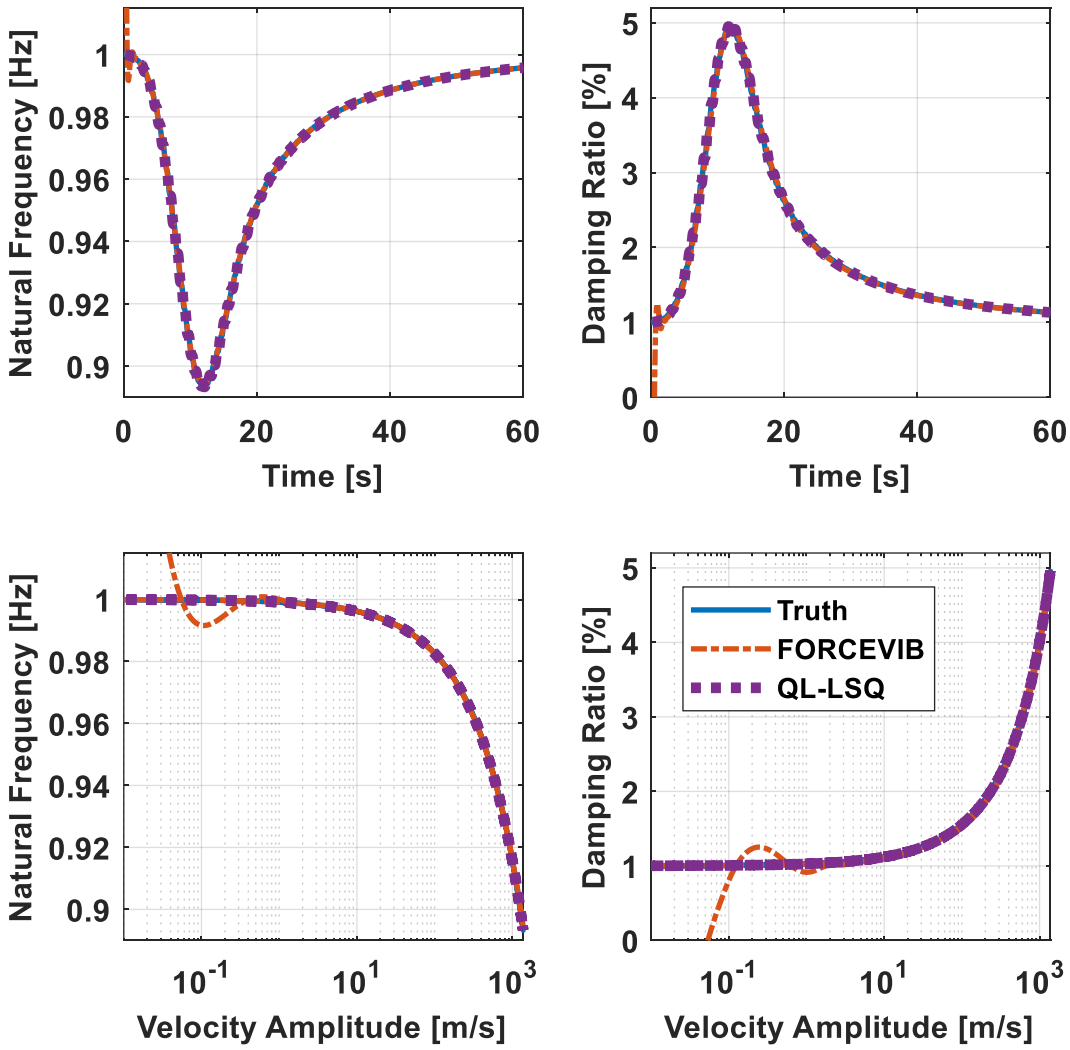

Figure 51: Comparison of Frequency and Damping vs Time and Velocity Amplitude from the exact truth values for this Modal Iwan Model, the Curve Fit, and QL-LSQ.

Furthermore, the QL-LSQ model does not exhibit the significant initial end effects present in the FREEVIB/ FORCEVIB result from the previous chapter. This is due to the fact that the FREEVIB/ FORCEVIB methods are heavily dependent on accurately estimating derivatives of the amplitude and phase, where any error is amplified and becomes more pronounced. Conversely, QL-LSQ operates directly on the response and force with respect to the amplitude, such that the initial ramp up of the response is characterized by the same portion of the spline that applies to the later ringdown. This separation of the temporal and physical behavior of the system allows QL-LSQ to produce a more accurate result.

In this numerical case study, QL-LSQ was shown to accurately model the simulated response of a Modal Iwan Model. The generated quasi-linear modal parameters very closely match the truth data for the nonlinear system while avoiding the initial end effect that pollutes the FORCEVIB curve due to its reliance on derivatives of the amplitude and phase.

4.3.2. Application to S4 Beam Impact Response

In this case study, the nonlinear experimental free decay from impact testing of the S4 Beam is characterized with the QL-LSQ method and compared to the results returned by other NLSID methods that specialize in free response. This data was previously explored in Section 2.5 with the STHT and Section 3.3.2 with the polynomial optimization process. The test structure and experimental setup are depicted in Figure 26 and the recorded drive point acceleration response from a modal hammer impact is shown in Figure 27.

To utilize QL-LSQ, the modal responses must be individually extracted from the physical response and integrated to generate the necessary SDOF acceleration, velocity, and displacement. These response quantities that were formed with the optimization process in the previous chapter are used here. The accelerations from the optimized polynomial fits to the first two elastic modes are shown in Figure 52.

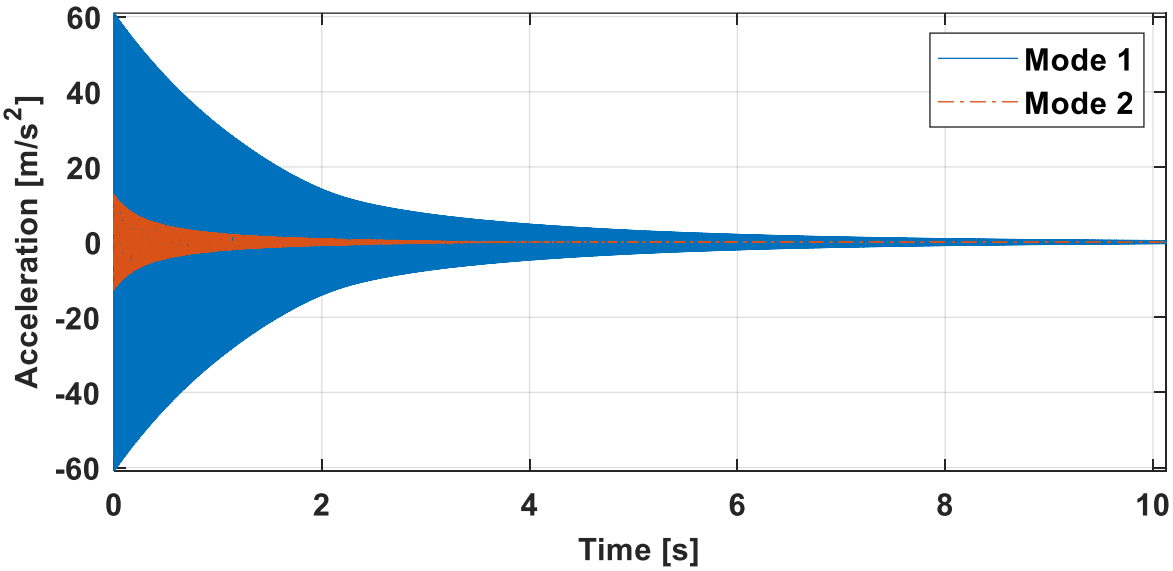

Figure 52: The acceleration response of the first two modes of the S4 beam that were extracted from the measured signal with the polynomial optimization process.

To determine the quasi-linear modal parameters of each resonance with QL-LSQ, the parameters for the B-spline curves are first set. Since the response is a free decay, the amplitude is only decreasing as time progresses, such that reordering the responses to be with respect to increasing amplitude is accomplished by simply flipping each vector. In this case, the difference between forming the splines relative to time verses amplitude is due to the nonlinearity in the system causing the amplitude to decay at different rates as time progresses. Constructing the knot sequence with respect to the amplitude characterizes the system response purely in terms of its

dependence on the amplitude and compensates for any temporal effect that could be captured by the spline curves. To that end, 4th order B-spline basis functions were constructed for each parameter from 40 equally spaced knots that span the amplitude range of the modal responses. Penalty smoothing was implemented in the form of standard second order difference matrices and λ values for the stiffness and damping splines of 10^{-9} & 10^{-2} for the first mode and 10^{-11} & $5 \cdot 10^{-5}$ for the second mode.

The resultant quasi-linear natural frequencies and damping ratios for both modes are shown in Figure 53 as solid black lines. Each plot also contains estimates of the amplitude dependent parameters for other NLSID methods. The blue solid line is from the PFF algorithm [56] which operates on estimates of location and height of oscillation peaks. The green dots are results published by M. Wall [19] using an implementation of the Hilbert Transform and FREEVIB in conjunction with a smoothing spline to estimate the amplitude and phase derivatives. The orange dotted line show the results from the STHT previously shown in Figure 31, and the yellow dashed line is from the optimization results shown in Figure 42. Each of these other approaches are different implementations of FREEVIB and are thus sensitive to noise in the signal during the numerical differentiation process to determine the frequency and damping. This is demonstrated in each curve from those methods, where the quasi-linear parameters get increasingly noisy as the amplitude decreases. While all curves are relatively smooth and agree at higher amplitudes, the QL-LSQ curves remain smooth throughout the amplitude range due to the flexibility and resilience to noise of the penalized B-spline framework.

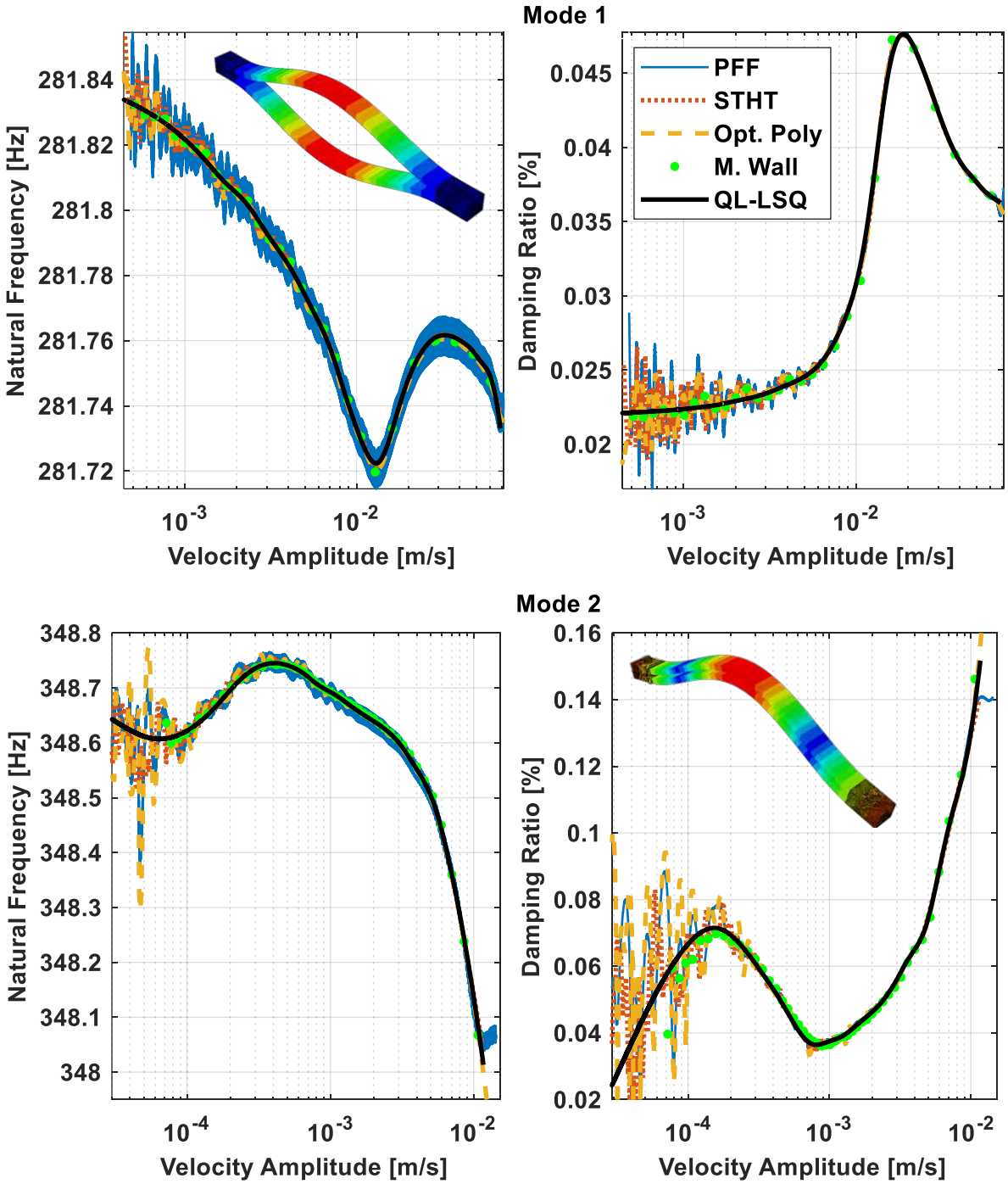

Figure 53: Amplitude dependent natural frequency (Left) and damping ratio (Right) for elastic mode 1 (Top) & mode 2 (Bottom) of the S4 Beam. In each, the green dots are from a Hilbert Transform based algorithm used by M. Wall in [19], the blue line is from the PFF algorithm, the orange dotted curves represent the STHT result, the yellow dashed line is from the optimized polynomial curve fit, and the solid black lines are from QL-LSQ. To demonstrate the deflection shape of the resonances, an image of each mode shape from a finite element model is shown superimposed over the plots.

4.3.3. Application to CPB Forced Response

The final case study presented here is an application of QL-LSQ to experimentally measured forced response of a benchmark structure. To compare against the FORCEVIB results from the optimization process explored in the previous chapter, this section revisits the Cylinder-Plate-Beam testbed and associated data detailed in Section 3.3.3. The experimental setup and CPB structure are shown in Figure 43, while the measured sine beat force and modal acceleration are plotted in Figure 44 and Figure 45. As QL-LSQ also requires the velocity and displacement, these were computed by numerically integrating the measured acceleration in the frequency domain. This is done by first computing the discrete Fourier coefficients of the acceleration with the FFT algorithm. The resulting vector of complex coefficients represent the amplitude and phase of a series of complex exponentials, $e^{i\omega t}$, oscillating at frequencies ω. Integrating the acceleration then amounts to dividing each coefficient by $i\omega$ to produce the velocity, and $-\omega^2$ for displacement. The time vector of each is found by computing the inverse FFT of the integrated acceleration coefficients.

The STHT from Chapter 2 is then used to extract the fundamental harmonic component of the modal responses and force and also estimate the amplitude of the integrated velocity signal. To compute the STFT of the signals for use in the STHT, a segment size of $0.1\,[\mathrm{s}]$ was used with an overlap of 75%. The corresponding spectrogram of the acceleration is shown in Figure 54. To decompose the fundamental harmonic, a time-frequency mask was constructed to filter out the higher

harmonics. The masking region is overlaid on the spectrogram where the selected points shown in red form the region enclosed by the blue line. Synthesizing the frequency content contained in the mask generates the time signal below the spectrogram. This process was performed on the responses and force to reduce each to monoharmonic signals and to form an accurate estimate of the velocity amplitude.

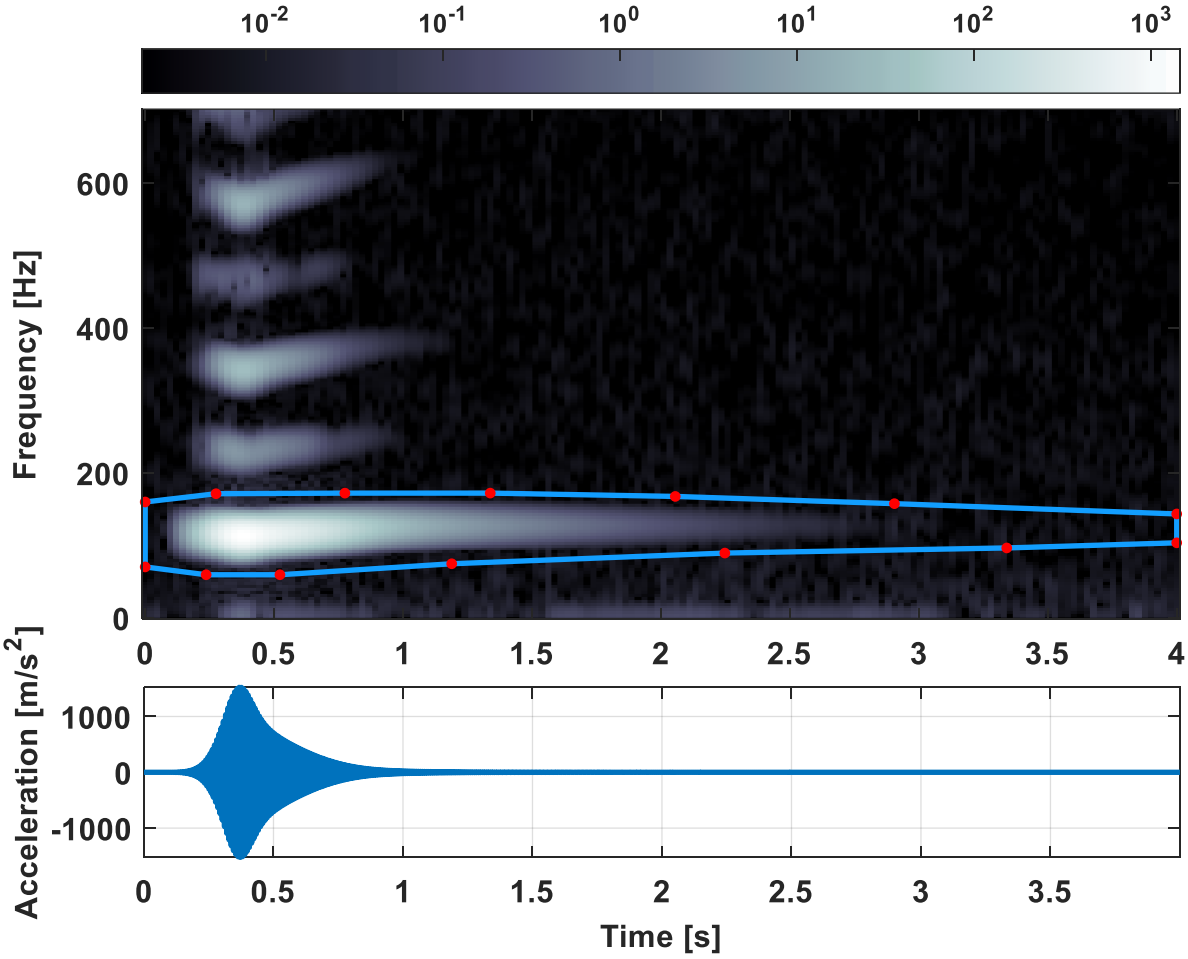

Figure 54: Extracting the fundamental harmonic of the CPB response with the STHT. Reassembling the enclosed section of the spectrogram in the time domain produces the shown time signal.

To characterize the nonlinear response of the first elastic CPB mode with QL-LSQ, the responses and force were reordered with respect to the velocity amplitude and a 4th order B-spline basis was formed from a uniform distribution of 40 knots that

span the amplitude. A second order difference matrix was used for the roughness penalty on the control point values, and λ values for the stiffness and damping splines were set to 10^{-8} and 0.1, respectively.

The resultant natural frequency and damping ratio curves are shown in Figure 55. In each plot, the QL-LSQ result is given by the orange curve and the blue dashed line is the FORCEVIB result from Figure 46 as computed by the optimized curve fitting process. The natural frequency estimates from both NLSID methods are in good agreement as the curves closely overlay in the time and amplitude plots. In comparing the damping ratio results, QL-LSQ does not exhibit the significant initial end effect present in the optimized curve fit and instead correctly converges toward the linear value at low amplitudes. Additionally, since FORCEVIB inherently operates on a sample-to-sample basis, it produces slightly different results at the same amplitude values during the initial ramp up and the subsequent ring down of the response and forms a loop in the amplitude dependent plots. Meanwhile, by ordering the data and defining the B-spline bases with respect to the amplitude, QL-LSQ generates a single-valued curve that simultaneously fits the portions of the response that have the same amplitude. The curve is essentially a weighted average of the data points near that amplitude value, where the least squares result is biased toward characterizing the nonlinear behavior in the form exhibited by the greatest number of data points in the response. Since the ringdown takes more time than the ramp up, the FORCEVIB result is closer to that portion of the QL-LSQ curve.

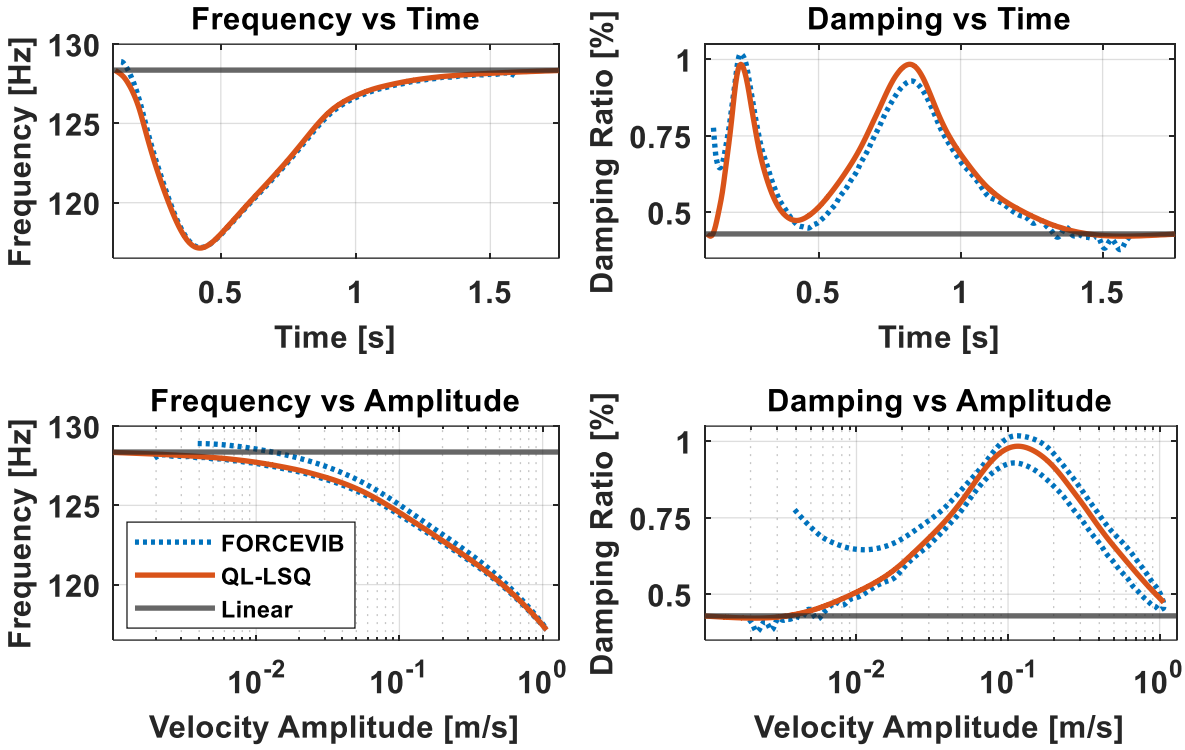

Figure 55: Comparison of Frequency and Damping vs Time and Velocity Amplitude from FORCEVIB, QL-LSQ, and the linear value from a separate, low amplitude test.

In addition to the single sine beat test analyzed above, several more at various force amplitudes were also performed. The quasi-linear natural frequency and damping ratio curves produced by QL-LSQ for each measurement are overlaid in Figure 56. In the bottom left plot, the natural frequency results trace out a consistent amplitude dependent backbone curve. Likewise, the damping ratio curves all display the same trend when plotted verses amplitude, demonstrating the repeatability of the measurements and the consistency of the QL-LSQ results. It is also noteworthy that, even though the lowest level test, given by the green line, is exhibiting forced response at 118 [Hz] throughout the measurement, QL-LSQ calculates the natural frequency and damping ratio to be quite close to the linear values.

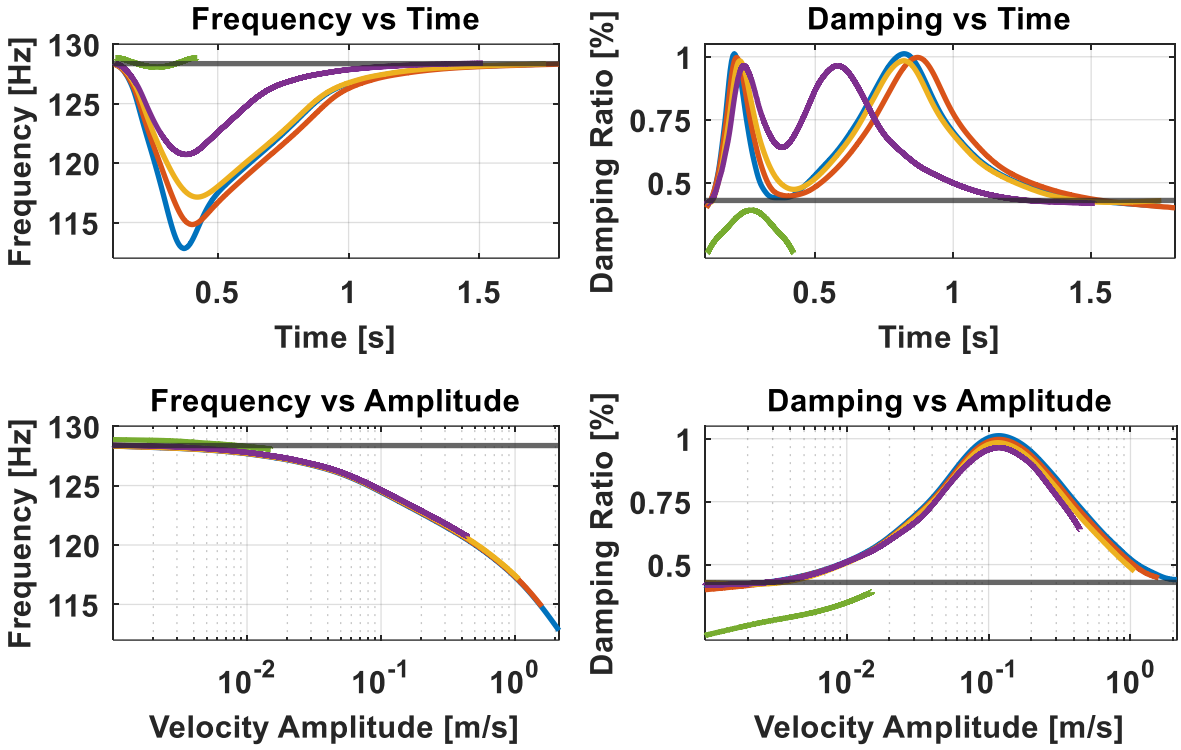

Figure 56: Frequency and Damping vs Time and Velocity Amplitude from QL-LSQ applied to several experimental measurements at various forcing amplitudes.

In this section, the QL-LSQ method was used to characterize the nonlinear behavior of the CPB structure undergoing forced response. After numerically integrating the measured acceleration to produce the associated velocity and displacement, the STHT was used to filter out the higher harmonics and estimate the amplitude. Linear least squares regression could be used to form B-spline representations of the quasi-linear natural frequency and damping ratio. Compared to the FORCEVIB results from the optimization process, QL-LSQ yields similar curves but does not produce the significant initial end effect present in the optimization curve and is also able to generate a single-valued curve with respect to amplitude instead of a loop.

4.4. Characterizing Modal Coupling with QL-LSQ

4.4.1. Introduction

Nonlinearity is often encountered in structural dynamics, and while many system identification techniques have been explored, these existing approaches are ill equipped to properly address many nonlinear behaviors that are observed in practice. Although great progress has been made in developing methods for analyzing isolated nonlinear resonances, these methods typically assume that the modal properties are uncoupled, in that the response of one mode does not have any effect on the behavior of another. This is a suitable assumption when analyzing weakly nonlinear systems in which variations in the stiffness and damping of each mode can be accurately characterized as a univariate function of its response amplitude. However, this approach fails when the effects of modal coupling are no longer negligible, and individual modal properties vary with respect to multiple modal responses. To that end, this section explores a method for efficiently characterizing the coupling effects between two nonlinear resonances in terms of the relative modal response amplitude of each. This is done by simultaneously exciting the relevant structural modes at a variety of amplitude combinations and extracting amplitude dependent natural frequency and damping ratio curves for each mode from the response measurements. The effect of the coupling between the resonances can be visualized and described in a three-dimensional space by representing each quasi-linear modal parameter as a function of both modal amplitudes. In this section, this approach is demonstrated utilizing experimentally measured nonlinear response data

from the CPB. Coupling behavior is observed by concurrently exciting pairs of modes of the structure that each activate the joint nonlinearity. Representing the identified quasi-linear parameters from the measured response data in the three-dimensional space mapped out by the associated modal response amplitudes offers an intuitive and concise means of characterizing and visualizing this complex nonlinear structural behavior.

4.4.2. Background on Modal Coupling

Within structural dynamics, linear modal analysis is the most well understood and utilized approach for characterizing the dynamic behavior of a system. Linear methods typically amount to determining constant valued natural frequencies, damping ratios, and mode shapes, $[\omega_{\mathrm{n}}, \zeta, \phi]$, that collectively describe the dynamic motion of the structure as a superposition of uncoupled and orthogonal modal responses. While linear techniques can be extremely useful, all real systems are nonlinear in some regard and the highly idealized representation produced by linear parameters is likely only accurate at low excitation levels. As a means of extending the concept of linear modal parameters beyond the linear regime, the nonlinear behavior exhibited by the stiffness and energy dissipation of the structure can be modelled in terms of equivalent natural frequencies and damping ratios that are a function of the amplitude of the structural response, $[\omega_{\mathrm{n}}(A), \zeta(A), \phi]$. These quasi-linear parameters present an intuitive extension to their linear counterparts, converging to the underlying constant, linear value at low amplitudes, while diverging at higher response amplitudes to emulate the nonlinear behavior of the structure in an

averaged sense. Typically, quasi-linear parameters are assumed to be a function purely of their respective modal response amplitude. In certain structures, such as those with bolted joints, some identified quasi-linear parameters display inconsistent trends with respect to their modal amplitude [143]. Usually, this anomalous behavior is attributed to being measurement error or noise, or some ambiguous modal coupling effect that cannot be reliably modelled or accounted for as it does not fit the assumption that the modes are uncoupled.

In an effort to understand and model observed modal coupling, the quasi-linear parameters may be considered to be a function of two modal amplitudes, $[\omega_{\mathrm{n}}(A_1, A_2), \zeta(A_1, A_2), \phi]$, such that, instead of tracing out a univariate backbone curve in 2D space, a representative surface is formed in 3D space, or if more than two modal amplitudes are considered, a higher dimensional plane. This approach was investigated by Haslam et al. in [144], in which numerical models with known nonlinearity were examined. Two methods for forming a 3D modal coupling surface were explored, the first was to utilize the Restoring Force Surface method to form coupled nonlinear equations of motion that attempt to emulate the nonlinear behavior illustrated by the modal coupling surface. While this approach showed some promise in those numerical cases, it has subsequently been shown to be very ineffective in modeling experimentally measured modal coupling data [38]. The second method was to directly fit an interpolating plane to a set of identified quasi-linear curves that are represented as a function of two modal amplitudes. This is the approach that is

investigated in the following section on a set of experimentally measured data that displays behavior suspected to be a result of modal coupling.

4.4.3. Modal Coupling in the CPB

The data presented here was collected from the aforementioned CPB structure and experimental test setup that was shown in Figure 43. This data was originally measured by the author in [38] in an effort to use the RFS method to model the observed modal coupling behavior. As detailed in Section 1.2.1, that approach proved less than ideal and prompted the development of the present work. In the response explored in Section 3.3.3 and Section 4.3.3, a single sine beat excitation was applied to significantly excite only the first mode of the CPB. Additional test cases included sine beats at various magnitudes applied to the first three elastic modes of the CPB individually and simultaneously in pairs. These modes are essentially the beam moving in three perpendicular directions: first order cantilever modes in the soft and stiff directions and translating axially upon the first order drum mode of the plate.

The measured CPB responses were reduced to single-degree-of-freedom modal responses via modal filtering with the linear mode shapes. While the quasi-linear parameters are assumed to be coupled, the modal filter operates upon the mode shapes, which should remain roughly constant and independent such that the modally filtered responses are approximately SDOF. As in Section 4.3.3, the modal acceleration was integrated in the frequency domain to produce the associated modal velocity and displacement. To reduce the responses and force to monoharmonic signals and to estimate the amplitude, the STHT was used as previously shown in

Figure 54, in which a time-frequency masking region was defined about the fundamental harmonic of the signal to filter out noise and higher harmonic components. The QL-LSQ process was then utilized to calculate the quasi-linear modal parameters from these filtered responses. For this experimental data, the response and force vectors are kept in order with respect to time, such that the knot sequences for the B-spline bases are also formed relative to the time span of the measurement. The process used in the case studies in Section 4.3, where the vectors are reordered with respect to the response amplitude, cannot be used here since these responses are assumed to be a function of two different amplitudes. If the quasi-linear parameters are determined with respect to one of the amplitudes, the resulting spline curves from the least squares fit will be an average of the nonlinear modal coupling behavior that is the focus of this section.

Shown below in Figure 57 are identified quasi-linear parameters of the third elastic mode as a function of the amplitude of the third mode. The underlying linear values for each, as determined from a low-level burst random noise test, were $\omega_{\mathrm{n}} = 544.8\ [\mathrm{Hz}]$ and $\zeta = 0.275\ [\%]$. In the plots, results from five experiments are shown. The two yellow curves are from sine beats that were applied to only provide excitation in the frequency range about the third mode such that its isolated nonlinear behavior could be characterized. These results trace out fairly consistent backbone curves in which the natural frequency displays a softening, and the damping ratio has an increasing then decreasing trend. The three blue curves are from a set of experiments in which the excitation consisted of a superposition of two sine beats that were defined

to excite both the first and third modes at magnitude ratios of 1:1, 2:1, and 1:2 to produce responses that exhibit various levels of modal coupling. The portion of the curves from the initial ramp up segment of the response, marked by the green arrows, agrees with the results from isolated Mode 3 response. However, the later portion of the coupled curves, marked by the red arrows, drastically deviates from the isolated nonlinear behavior exhibited by the third mode.

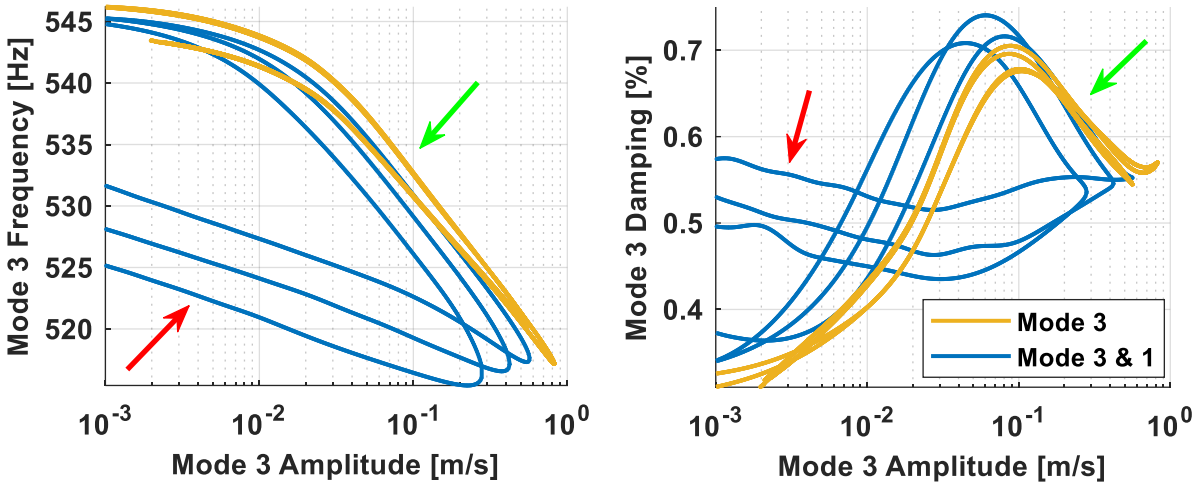

Figure 57: Quasi-linear behavior of the third CPB mode as a function of the mode 3 amplitude. The curves from when Modes 3 and 1 are simultaneously excited exhibit significantly different nonlinear behavior compared to when Mode 3 is excited in isolation.

These same curves are shown with respect to the Mode 1 amplitude in Figure 58. In this representation the behavior in the curves from the combined first and third mode excitation, marked again by the red arrow, now closely overlay, signifying a dependence on the amplitude of the first mode instead of the third. This is as opposed to the isolated Mode 3 curves, again marked by the green arrow, which show no consistent trend relative to the first mode amplitude, signifying that these are indeed dependent only on the amplitude of the third mode.

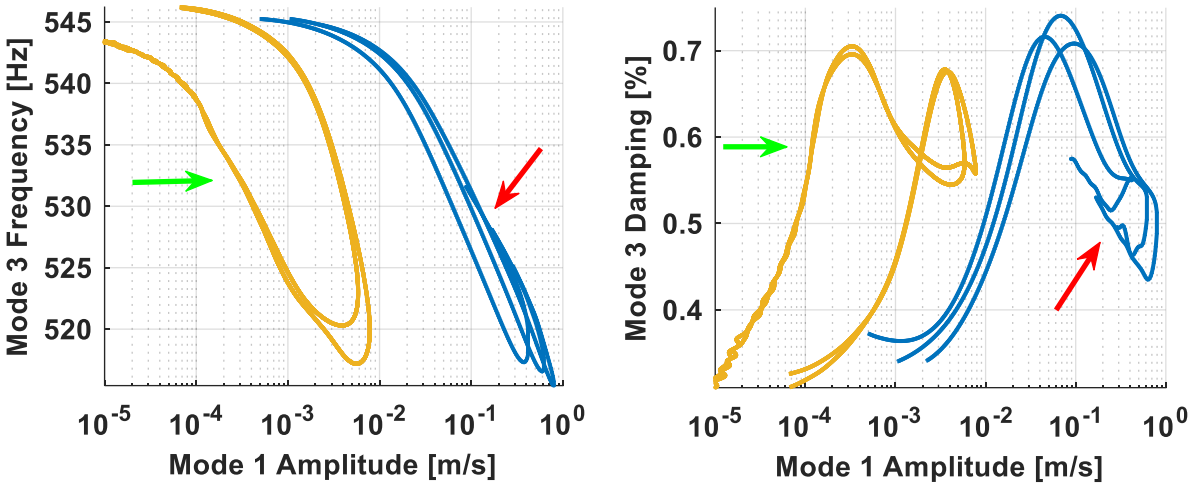

Figure 58: The quasi-linear modal parameters of the third CPB mode as a function of the amplitude of the first mode. The nonlinear behavior in the combined excitation cases exhibit modal coupling in the form of a dependence on the Mode 1 amplitude.

These quasi-linear parameter curves are shown in a combined 3D space in Figure 59. The trajectory of the curves through the space defined by the amplitudes is shown beneath each in a darker color. In the natural frequency results in the top plot, the curves begin near the linear value and soften to lower frequencies at higher amplitudes with respect to both the Mode 1 and Mode 3 amplitudes. In the damping, the curves begin near the linear value before displaying a sort of bell shape with the increasing amplitudes. While the isolated Mode 3 result follows the same trend with increasing and decreasing amplitude, the cases with the simultaneous Mode 1 response ring down with higher damping ratios. These results show that the quasi-linear behavior of the third mode is very sensitive to the excitation of the first mode. This is likely a result of both modes heavily involving motion of the beam, which warps the plate and distorts the contact between the plate and cylinder, causing variations in the frictional effects and perceived stiffness of the interface.

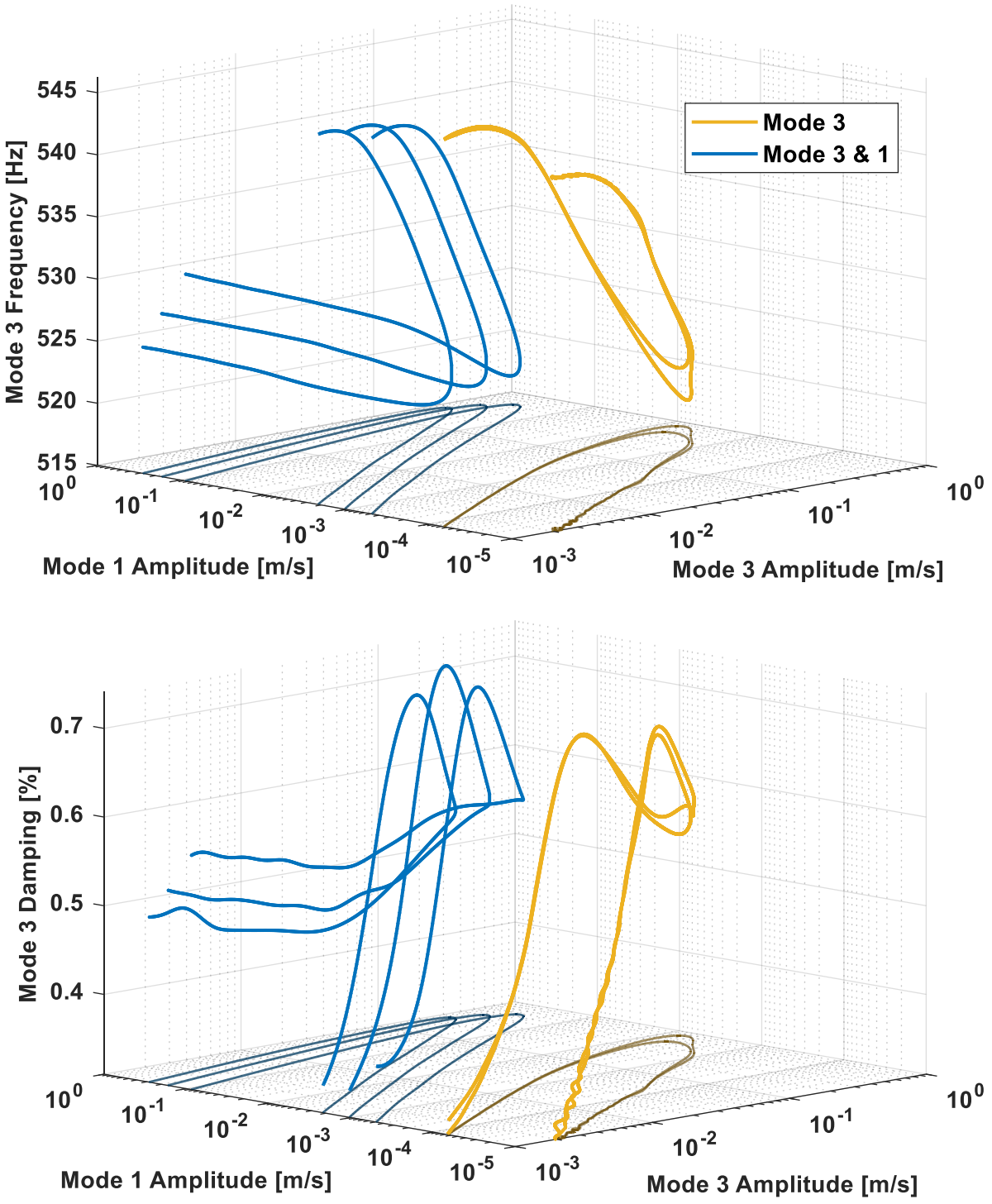

Figure 59:The quasi-linear parameters of CPB Mode 3 in a 3D space defined by the response amplitudes of mode 1 and mode 3, with Natural Frequency on top and Damping Ratio on the bottom.

Modal coupling is a complicated dynamic phenomenon that has proved challenging to identify and model, to the extent that coupling is often avoided by simply assuming that it is negligible. In systems where that is not feasible, techniques are needed that present the nonlinear behavior of the modal coupling in an intuitive format. This work demonstrated the use of one such approach on experimentally identified quasi-linear parameters that exhibit the effects of modal coupling. By plotting the varying natural frequency and damping ratio in a 3D space with respect to multiple modal responses, the combined amplitude dependence of the parameters can be visualized and more easily interpreted.

4.5. Conclusion

This chapter detailed a new nonparametric NLSID technique which utilizes the quasi-linear model form. The primary objective of this method, here called QL-LSQ, is to determine equivalent modal natural frequencies and damping ratios that represent amplitude dependent averaged nonlinear behavior. In QL-LSQ, this behavior is modelled in terms of variations in stiffness and viscous damping, represented by penalized B-spline curves. No assumptions have been made as to the form or physical processes in the nonlinearity, only that the resultant averaged effects are reasonably smooth and vary slowly enough to be accurately captured by relatively low-order piecewise polynomials. Fitting the model to measured nonlinear response data is done by determining the control point values through a linear least squares solution between the velocity and displacement and the difference between the acceleration

and force. By reordering the vectors and forming the B-spline bases with respect to the amplitude of the response, the amplitude dependent form of the resultant spline curves for the stiffness and damping can be determined directly from the data.

The results produced by QL-LSQ are comparable to those from Feldman's FREEVIB/ FORCEVIB methods as they are also derived from the quasi-linear model form. While QL-LSQ determines stiffness and damping in terms of B-splines fit with a least squares solution, Feldman's approach was to assume the quasi-linear EOM is in a complex form such that the real and imaginary components can be separated to form two equations to solve for the two unknowns. The inherent difficulty in this process is that it requires the amplitude, phase, and derivatives of each to be estimated from the measured data, which can be a complicated and troublesome procedure. QL-LSQ then presents a notable simplification over FREEVIB/ FORCEVIB as it operates directly on the real-valued response and force signals. Whereas Feldman's approach requires the use of temperamental processes like the Hilbert Transform to estimate amplitude and phase, QL-LSQ exchanges this for the simpler process of numerical integration/ differentiation.

The case studies explored in the previous chapters were revisited here to demonstrate how the QL-LSQ process may be applied in various scenarios and to compare the results to those from FREEVIB/ FORCEVIB. First the numerical Iwan Model simulation data from Section 3.3.1 was analyzed, showing how QL-LSQ produces quasi-linear parameters that are in very good agreement to the known truth values. Additionally, QL-LSQ does not exhibit the significant initial end effect present

in the FORCEVIB result. The next case examined experimental free response of the S4 Beam first shown in Section 2.5. As this data set contains multiple simultaneously decaying oscillations, the extracted and integrated signals from the optimization process in Section 3.3.2 were used with QL-LSQ to characterize the nonlinearity in the first two modes. The resultant quasi-linear natural frequency and damping ratio curves from QL-LSQ are in good agreement with those from FREEVIB-based methods at high amplitude, while also proving to be smoother and more resilient to noise at low amplitude. Lastly, the experimental forced response of the CPB structure presents an instance where the measured force signal must be accurately accounted for to produce physically meaningful quasi-linear parameters. The QL-LSQ results are again in good agreement with the comparable FORCEVIB result, demonstrating the accuracy achieved by the QL-LSQ method while avoiding the amplitude and phase derivatives required in FORCEVIB. Lastly, QL-LSQ was used to characterize additional measurements from the CPB that exhibit modal coupling. By representing the quasi-linear modal parameters as a function of two modal amplitudes, the effect that each has on the nonlinear behavior of a single mode can be visualized and more easily interpreted.

The QL-LSQ method presents an attractive alternative to Feldman's FREEVIB/ FORCEVIB approaches, as it provides quasi-linear natural frequency and damping ratio results directly from the measured response and force signals. This simplifies the data analysis while increasing the resilience to noise, providing a notable improvement over existing algorithms that utilize FREEVIB/ FORCEVIB.

5. **Conclusion & Future Work**

This dissertation presented a set of new NLSID techniques that offer notable improvements over existing approaches used in structural dynamics for determining amplitude dependent modal properties. In Chapter 2, the STHT was introduced as a generalization of the process for numerically computing the Hilbert Transform. By utilizing the STFT instead of the Fourier Transform, the ringing artifacts caused by leakage in the signal are suppressed, and time-frequency masking can be used for simple but effective mode decomposition if the signal components are separable. While the case studies showed that this approach produces more accurate amplitude and phase results compared to the Hilbert Transform, the STHT will still exhibit ringing artifacts in the local vicinity of any sharp changes in the signal due to inherent limitations of the Hilbert Transform.

An alternative was detailed in Chapter 3, in which nonlinear optimization was used to curve fit polynomial representations of amplitude and phase to a sampled oscillation. This proposed process avoids many of the issues encountered in implementations of the Hilbert Transform. The preprocessing steps that are typically used to ensure that the signal is monoharmonic, like bandpass filtering, EMD, or VMD, are not needed for the optimization as each oscillation can be individually fit directly from the original multi-harmonic signal. Additionally, the optimized polynomials can be used to represent the integrals and derivatives of the oscillation, and the result is much less sensitive to noise as the amplitude and phase are modeled by smooth polynomials that are curve fit to the data in a least squares sense. However, compared

to the STHT, utilizing nonlinear optimization is much more computationally expensive. Therefore, which approach is used would be decided on a case-by-case basis. The STHT can be utilized if the residual end effects are negligible, and the oscillation components are separable via time-frequency masking. Otherwise, the optimization process can be implemented in cases where there is significant noise that will pollute the STFT segments with leakage or to extract oscillations that are close in frequency. As a future extension to these processes, a combined approach could be taken where the STHT is utilized to form an estimate of the amplitude and phase that are then used as an initial guess for each in the optimization, leading to a more stable and faster solution. This could also be implemented in a spline framework, in which spline curves fit to the STHT results are then refined by the optimization process.

Chapter 4 presented a new method for determining amplitude dependent modal parameters called QL-LSQ. In this process, the quasi-linear stiffness and viscous damping are represented as B-spline curves that are fit to the nonlinear restoring force with linear least squares regression. The main advantage of QL-LSQ over existing methods is the adaptability to a wide range of response conditions and the fact that it does not require amplitude and phase derivatives as part of the solution process. Computing those quantities is typically very sensitive to noise, requiring additional smoothing and filtering operations to mitigate. In QL-LSQ, the penalized B-splines automatically average through any noise and are highly customizable. This provides a notable improvement over existing algorithms as it simplifies the data analysis while also increasing the resilience to noise.

The overall contribution of this dissertation is a comprehensive solution for characterizing quasi-linear modal parameters in which the STHT and optimization process are used as preprocessing steps for QL-LSQ. The proposed methods individually have certain weaknesses that can be accounted for with the others. QL-LSQ produces smooth amplitude dependent natural frequency and damping ratio curves, but requires an estimate of the response amplitude, and the acceleration, velocity, displacement, and force signals must all be known in a SDOF and monoharmonic form. While these quantities can be determined with the STHT and/or optimization process, the case studies for those individual methods demonstrated the difficulties in utilizing FREEVIB/ FORCEVIB. Employing a combination of the three proposed methods yields superior results that avoid pitfalls in each, where the measured signals characterized by the STHT and optimization are then passed to QL-LSQ to determine the quasi-linear modal properties.

The experimental case studies in Section 4.3 exemplify this framework. For the S4 Beam data, the optimization process was used to extract and integrate the modal responses and estimate the amplitude. The amplitude dependent natural frequencies and damping ratios computed with QL-LSQ exhibited significantly less noise than those from the FREEVIB-based methods while also tracking the same underlying trend in the nonlinear behavior. To analyze the CPB forced response, the STHT was used to filter out the higher harmonic components of the responses and force and estimate the velocity amplitude. The subsequent quasi-linear parameter curves follow the same trend as the those from FORCEVIB but do so again with less noise and in

terms of a univariate curve that simultaneously accounts for the nonlinearity during the ramp up and ringdown of the response. Each of these case studies focused on modelling nonlinear behavior that is observed in structures containing bolted joints. An avenue for future applications of QL-LSQ is to investigate if other types of nonlinearities, such as viscoelastic, material, and geometric, could also be represented by a model form consisting of B-splines fit using least squares regression.

In the final case study, the STHT and QL-LSQ were used to generate quasi-linear modal parameter curves from CPB response data that exhibited effects of modal coupling. Currently, modal coupling is very poorly understood and difficult to model. This case demonstrated that the identified parameter curves can be represented as a function of two modal amplitudes in a 3D space as a means of visualizing and interpreting the effects of modal coupling in the CPB. This is a promising approach that can be explored further in the future. To quantitatively characterize the coupling between resonances, a surface could be formed in the 3D space that maps how different modal amplitudes effect the properties of one resonance. This surface could be formed by interpolating between the individual 3D curves, or by directly forming it in terms of multivariate B-splines that span both amplitude ranges. The most significant issue with either approach is that data is needed throughout the amplitude space to properly define the surface. This would likely require devising a specialized testing methodology that combines sine beats and dwells to efficiently collect response data that occurs at a wide range of relative amplitudes to fill in more areas of the 3D surface plot.

Publications of Doctoral Work

Conference Papers

- B. Moldenhauer *et al.*, “Influences of Modal Coupling on Experimentally Extracted Nonlinear Modal Models,” 2020, pp. 189–204. doi: 10.1007/978-3-030-12391-8_25.
- B. Moldenhauer, D. R. Roettgen, and B. Pacini, “Implementing the Restoring Force Surface Method to Fit Experimentally Measured Modal Coupling Effects,” in *Nonlinear Structures & Systems, Volume 1*, 2021. doi: 10.1007/978-3-030-47626-7_12.
- B. Moldenhauer, M. Allen, and D. Roettgen, “Variation of the Restoring Force Surface Method to Estimate Nonlinear Stiffness and Damping Parameters.,” Sandia National Laboratories, SAND2020-7255C, Jul. 2020. [Online]. Available: https://www.osti.gov/biblio/1809200

Conference Presentations

- B. Moldenhauer, A. Singh, M. Allen, and D. Roettgen, “Extensions to a Method for Characterizing Instantaneous Frequency and Damping of Nonlinear Systems,” 39th International Modal Analysis Conference, 2021
- B. Moldenhauer, M. Allen, and D. Roettgen, “A Method for Estimating the Amplitude and Phase of Oscillating Signals,” 40th International Modal Analysis Conference, 2022
- B. Moldenhauer, M. Allen, and D. Roettgen, “Characterizing Experimental Nonlinear Modal Coupling with 3D Surfaces,” 40th International Modal Analysis Conference, 2022

Journal Papers to be Submitted

Committee Note: These two works have drafts that are in a complete state and will be submitted before the final dissertation is submitted to UW-Madison.

- B. J. Moldenhauer, M. S. Allen, and D. R. Roettgen, “The Short-time Hilbert Transform,” *Mechanical Systems and Signal Processing* (To be Submitted)
- B. J. Moldenhauer, M. S. Allen, and D. R. Roettgen, “Quasi-Linear Stiffness and Damping Parameters from Arbitrary Forcing and Response,” *Mechanical Systems and Signal Processing* (To be Submitted)

Journal Papers in Preparation

Committee Note: These two are still in active development and will be explored further during my post-doc and by Dr. Allen.

- B. J. Moldenhauer, M. S. Allen, and D. R. Roettgen, “Algorithm for Determining Optimized Spline Representations of Amplitude and Phase for Nonstationary Oscillations,” *Mechanical Systems and Signal Processing* (In preparation)
- B. J. Moldenhauer, M. S. Allen, and D. R. Roettgen, “Characterizing Experimental Nonlinear Modal Coupling as Surface Maps,” *Mechanical Systems and Signal Processing* (In preparation)

References

[1] D. J. Ewins, *Modal Testing: Theory, Practice and Application*. Baldock, England: Research Studies Press, 2000.

[2] R. W. Gordon and J. J. Hollkamp, “Reduced-order Models for Acoustic Response Prediction,” Air Force Research Laboratory, AFRL-RB-WP-TR-2011-3040, Dayton, OH, AFRL-RB-WP-TR-2011-3040, Jul. 2011.

[3] M. A. Crisfield, *Nonlinear finite element analysis of solids and structures. Volume 1: Essentials*. New York: John Wiley and Sons, 1991.

[4] M. A. Crisfield, *Nonlinear finite element analysis of solids and structures. Volume 2: Advanced Topics*. New York: John Wiley and Sons, 1991.

[5] W.-Y. Tseng and J. Dugundji, “Nonlinear Vibrations of a Buckled Beam Under Harmonic Excitation,” *J. Appl. Mech.*, vol. 38, no. 2, pp. 467–476, Jun. 1971, doi: 10.1115/1.3408799.

[6] M. Brake, *The Mechanics of Jointed Structures: Recent Research and Open Challenges for Developing Predictive Models for Structural Dynamics*. 2018. doi: 10.1007/978-3-319-56818-8.

[7] L. Gaul and J. Lenz, “Nonlinear dynamics of structures assembled by bolted joints,” *Acta Mech.*, vol. 125, no. 1–4, Art. no. 1–4, 1997.

[8] D. R. Roettgen and M. S. Allen, “Nonlinear characterization of a bolted, industrial structure using a modal framework,” *Mech. Syst. Signal Process.*, vol. 84, pp. 152–170, 2017, doi: 10.1016/j.ymssp.2015.11.010.

[9] R. L. Mayes, B. R. Pacini, and D. R. Roettgen, “A Modal Model to Simulate Typical Structural Dynamic Nonlinearity,” Jan. 2016.

[10] D. Segalman, “An Initial Overview of Iwan Modeling for Mecanical Joints,” Sandia SAND Report SAND2001-0811, Mar. 2001.

[11] E. E. Ungar, “The status of engineering knowledge concerning the damping of built-up structures,” *J. Sound Vib.*, vol. 26, no. 1, Art. no. 1, Jan. 1973, doi: 10.1016/S0022-460X(73)80210-X.

[12] M. Brøns, T. A. Kasper, G. Chauda, S. W. B. Klaassen, C. W. Schwingshackl, and M. R. W. Brake, “Experimental Investigation of Local Dynamics in a Bolted Lap Joint Using Digital Image Correlation,” *J. Vib. Acoust.*, vol. 142, no. 051114, Jul. 2020, doi: 10.1115/1.4047699.

[13] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*. Springer Science & Business Media, 2007.

[14] EdRSantos, *Comparison between a perturbed logistic growth equation and the averaged over an amount of time.* 2018. Accessed: Feb. 01, 2021. [Online]. Available: https://commons.wikimedia.org/wiki/File:Logistic_growth_equation.png#filelinks

[15] H. Wentzel, “Modelling of frictional joints in dynamically loaded structures: a review,” 2006.

[16] S. R. Ibrahim, “Time-domain quasilinear identification of nonlinear dynamic systems,” *AIAA J.*, vol. 22, no. 6, pp. 817–823, Jun. 1984, doi: 10.2514/3.8686.

[17] S. J. Elliott, M. G. Tehrani, and R. S. Langley, “Nonlinear damping and quasi-linear modelling,” *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, vol. 373, no. 2051, p. 20140402, Sep. 2015, doi: 10.1098/rsta.2014.0402.

[18] A. Singh *et al.*, “Experimental Characterization of a new Benchmark Structure for Prediction of Damping Nonlinearity,” Orlando, Florida, 2018. [Online]. Available: http://sd.engr.wisc.edu/wp-uploads/2018/04/31_sin.pdf

[19] M. Wall, M. S. Allen, and R. J. Kuether, “Observations of Modal Coupling due to Bolted Joints in an Experimental Benchmark Structure,” *Mech. Syst. Signal Process.*, p. 23, 2020.

[20] G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval, “Past, present and future of nonlinear system identification in structural dynamics,” *Mech. Syst. Signal Process.*, vol. 20, no. 3, pp. 505–592, Apr. 2006, doi: 10.1016/j.ymssp.2005.04.008.

[21] J. P. Noël and G. Kerschen, “Nonlinear system identification in structural dynamics: 10 more years of progress,” *Mech. Syst. Signal Process.*, vol. 83, pp. 2–35, Jan. 2017, doi: 10.1016/j.ymssp.2016.07.020.

[22] J. Schoukens and L. Ljung, *Nonlinear System Identification: A User-Oriented Roadmap*. 2019.

[23] L. Ljung, “Perspectives on system identification,” *Annu. Rev. Control*, vol. 34, no. 1, pp. 1–12, Apr. 2010, doi: 10.1016/j.arcontrol.2009.12.001.

[24] A. T. Mathis, N. N. Balaji, R. J. Kuether, A. R. Brink, M. R. W. Brake, and D. D. Quinn, “A Review of Damping Models for Structures With Mechanical Joints,” *Appl. Mech. Rev.*, vol. 72, no. 040802, Jul. 2020, doi: 10.1115/1.4047707.

[25] D. J. Segalman, “A Four-Parameter Iwan Model for Lap-Type Joints,” *J. Appl. Mech.*, vol. 72, no. 5, Art. no. 5, Sep. 2005.

[26] Y. Wen, “Method for Random Vibration of Hysteretic Systems,” *J. Eng. Mech.-Asce*, vol. 102, pp. 249–263, 1976.

[27] Hamel, “Georg Duffing, Ingenieur: Erzwungene Schwingungen bei veränderlicher Eigenfrequenz und ihre technische Bedeutung. Sammlung Vieweg. Heft 41/42, Braunschweig 1918. VI+134 S,” *ZAMM - J. Appl. Math. Mech. Z. Für Angew. Math. Mech.*, vol. 1, no. 1, pp. 72–73, 1921, doi: https://doi.org/10.1002/zamm.19210010109.

[28] S. F. Masri and T. K. Caughey, “A Nonparametric Identification Technique for Nonlinear Dynamic Problems,” *J. Appl. Mech.*, vol. 46, no. 2, pp. 433–447, Jun. 1979, doi: 10.1115/1.3424568.

[29] I. J. Leontaritis and S. A. Billings, “Input-Output Parametric Models for Non-Linear Systems Part I: Deterministic Non-Linear Systems,” *Int. J. Control*, vol. 41, no. 2, pp. 303–328, Feb. 1985, doi: 10.1080/0020718508961129.

[30] I. J. Leontaritis and S. A. Billings, “Input-Output Parametric Models for Non-Linear Systems Part Ii: Stochastic Non-Linear Systems,” *Int. J. Control*, vol. 41, no. 2, pp. 329–344, Feb. 1985, doi: 10.1080/0020718508961130.

[31] R. J. Allemang and D. L. Brown, “A Unified Matrix Polynomial Approach to Modal Identification,” *J. Sound Vib.*, vol. 211, no. 3, Art. no. 3, 1998.

[32] C. M. Richards and R. Singh, “Identification of Multi-Degree Non-Linear Systems Under Random Excitations by the ‘REVERSE PATH’ Spectral Method,” *J. Sound Vib.*, vol. 213, pp. 673–708, Jun. 1998, doi: 10.1006/jsvi.1998.1522.

[33] M. Kwarta and M. Allen, “Extensions to NIFO and CRP to Estimate Frequency-Independent Nonlinear Parameters,” 2019. doi: 10.1007/978-3-030-47709-7_9.

[34] F. Thouverez, “PRESENTATION OF THE ECL BENCHMARK,” *Mech. Syst. Signal Process.*, vol. 17, no. 1, pp. 195–202, Jan. 2003, doi: 10.1006/mssp.2002.1560.

[35] M. Allen, H. Sumali, and D. Epp, “Piecewise-linear restoring force surfaces for semi-nonparametric identification of nonlinear systems,” *Nonlinear Dyn.*, vol. 54, pp. 123–135, 2008.

[36] M. Allen, H. Sumali, and D. S. Epp, “Restoring Force Surface Analysis of Nonlinear Vibration Data from Micro-Cantilever Beams,” Nov. 2006.

[37] B. R. Pacini, R. L. Mayes, B. C. Owens, and R. Schultz, “Nonlinear Finite Element Model Updating, Part I: Experimental Techniques and Nonlinear Modal Model Parameter Extraction,” Garden Grove, CA, 2017.

[38] B. Moldenhauer, D. R. Roettgen, and B. Pacini, “Implementing the Restoring Force Surface Method to Fit Experimentally Measured Modal Coupling Effects,” in *Nonlinear Structures & Systems, Volume 1*, 2021. doi: 10.1007/978-3-030-47626-7_12.

[39] Y. Kopsinis, E. Aboutanios, D. A. Waters, and S. McLaughlin, “Investigation of bat echolocation calls using high resolution spectrogram and instantaneous frequency based analysis,” in *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, Aug. 2009, pp. 557–560. doi: 10.1109/SSP.2009.5278516.

[40] Y. Fu and L. N. Kloepper, “A systematic method for isolating, tracking and discriminating time-frequency components of bat echolocation calls,” *J. Acoust. Soc. Am.*, vol. 143, no. 2, pp. 716–726, Feb. 2018, doi: 10.1121/1.5023205.

[41] J. Gagnon, E. Goulielmakis, and V. S. Yakovlev, “The accurate FROG characterization of attosecond pulses from streaking measurements,” *Appl. Phys. B*, vol. 92, no. 1, pp. 25–32, Jul. 2008, doi: 10.1007/s00340-008-3063-x.

[42] B. Boashash, Ed., “Chapter 4 - Advanced Time-Frequency Signal and System Analysis,” in *Time-Frequency Signal Analysis and Processing (Second Edition)*, Oxford: Academic Press, 2016, pp. 141–236. doi: 10.1016/B978-0-12-398499-9.00004-2.

[43] B. Boashash, Ed., “Chapter 10 - Instantaneous Frequency Estimation and Localization,” in *Time-Frequency Signal Analysis and Processing (Second Edition)*, Oxford: Academic Press, 2016, pp. 575–635. doi: 10.1016/B978-0-12-398499-9.00010-8.

[44] J. Allen, “Short term spectral analysis, synthesis, and modification by discrete Fourier transform,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 25, no. 3, pp. 235–238, Jun. 1977, doi: 10.1109/TASSP.1977.1162950.

[45] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, Second Edition*. CRC Press, 2017.

[46] F. Auger *et al.*, “Time-Frequency Reassignment and Synchrosqueezing: An Overview,” *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013, doi: 10.1109/MSP.2013.2265316.

[47] D. Iatsenko, P. V. E. McClintock, and A. Stefanovska, “Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms,” *Digit. Signal Process.*, vol. 42, pp. 1–26, Jul. 2015, doi: 10.1016/j.dsp.2015.03.004.

[48] S. Meignen, T. Oberlin, and D.-H. Pham, “Synchrosqueezing transforms: From low- to high-frequency modulations and perspectives,” *Comptes Rendus Phys.*, vol. 20, no. 5, pp. 449–460, Jul. 2019, doi: 10.1016/j.crhy.2019.07.001.

[49] D. Iatsenko, P. V. E. McClintock, and A. Stefanovska, “Extraction of instantaneous frequencies from ridges in time–frequency representations of signals,” *Signal Process.*, vol. 125, pp. 290–303, Aug. 2016, doi: 10.1016/j.sigpro.2016.01.024.

[50] Y. Abdoush, J. A. Garcia-Molina, and G. E. Corazza, “Adaptive instantaneous frequency estimation based on time-frequency distributions with derivative approximation,” *Signal Process.*, vol. 160, pp. 99–105, Jul. 2019, doi: 10.1016/j.sigpro.2019.01.027.

[51] MathWorks, “Estimate instantaneous frequency - MATLAB instfreq,” 2018. https://www.mathworks.com/help/signal/ref/instfreq.html (accessed Jan. 24, 2021).

[52] Robert. J. Kuether and M. R. W. Brake, “Instantaneous Frequency and Damping from Transient Ring-Down Data,” Sandia National Laboratories, Sandia SAND Report SAND2015-9173C, 2015. [Online]. Available: https://doi.org/10.1007/978-3-319-56818-8_21

[53] K. Wijesundara, C. Negulescu, E. Foerster, and D. Monfort, “Identification of modal parameters of ambient excitation Structures using continuous wavelet transform,” Sep. 2012.

[54] K. Czarnecki, “The instantaneous frequency rate spectrogram,” *Mech. Syst. Signal Process.*, vol. 66–67, pp. 361–373, Jan. 2016, doi: 10.1016/j.ymssp.2015.06.029.

[55] J. L. Zapico-Valle, M. García-Diéguez, and R. Alonso-Camblor, “Nonlinear modal identification of a steel frame,” *Eng. Struct.*, vol. 56, pp. 246–259, Nov. 2013, doi: 10.1016/j.engstruct.2013.04.026.

[56] M. Jin, W. Chen, M. R. W. Brake, and H. Song, “Identification of Instantaneous Frequency and Damping From Transient Decay Data,” *J. Vib. Acoust.*, vol. 142, no. 051111, Jun. 2020, doi: 10.1115/1.4047416.

[57] J. M. Londoño, S. A. Neild, and J. E. Cooper, “Identification of backbone curves of nonlinear systems from resonance decay responses,” *J. Sound Vib.*, vol. 348, pp. 224–238, Jul. 2015, doi: 10.1016/j.jsv.2015.03.015.

[58] H. Goyder and D. Lancereau, “Methods for the Measurement of Non-Linear Damping and Frequency in Built-Up Structures,” Aug. 2017, p. V008T12A013. doi: 10.1115/DETC2017-67007.

[59] M. Jin, M. R. W. Brake, and H. Song, “Comparison of nonlinear system identification methods for free decay measurements with application to jointed structures,” *J. Sound Vib.*, vol. 453, pp. 268–293, Aug. 2019, doi: 10.1016/j.jsv.2019.04.021.
[60] M. Feldman, *Hilbert Transform Applications in Mechanical Vibration*. Chichester: Wiley, 2011.
[61] M. Feldman, “Hilbert transform methods for nonparametric identification of nonlinear time varying vibration systems,” *Mech. Syst. Signal Process.*, vol. 47, no. 1, pp. 66–77, Aug. 2014, doi: 10.1016/j.ymssp.2012.09.003.
[62] M. Feldman, “Non-linear system Vibration Analysis using Hilbert Transform - I. Free Vibration Analysis Method ‘FREEVIB,’” *Mech. Syst. Signal Process.*, vol. 8, no. 2, pp. 119–127, 1994.
[63] M. Feldman, “Non-linear system Vibration Analysis using Hilbert Transform - II. Forced Vibration Analysis Method ‘FORCEVIB,’” *Mech. Syst. Signal Process.*, vol. 8, no. 3, pp. 309–318, 1994.
[64] M. Feldman, “Hilbert transform in vibration analysis,” *Mech. Syst. Signal Process.*, vol. 25, no. 3, pp. 735–802, Apr. 2011, doi: 10.1016/j.ymssp.2010.07.018.
[65] B. Deaner, M. S. Allen, M. J. Starr, D. J. Segalman, and H. Sumali, “Application of Viscous and Iwan Modal Damping Models to Experimental Measurements From Bolted Structures,” *ASME J. Vib. Acoust.*, vol. 137, p. 12, 2015, doi: 10.1115/1.4029074.
[66] M. Feldman, “Time-varying vibration decomposition and analysis based on the Hilbert transform,” *J. Sound Vib.*, vol. 295, no. 3, pp. 518–530, Aug. 2006, doi: 10.1016/j.jsv.2005.12.058.
[67] M. Bertha and J.-C. Golinval, “Modal identification of time-varying systems using Hilbert transform and signal decomposition,” 2014, Accessed: Jan. 10, 2021. [Online]. Available: https://orbi.uliege.be/handle/2268/172196
[68] M. Bertha and J. C. Golinval, “Identification of a Time-Varying Beam Using Hilbert Vibration Decomposition,” in *Topics in Modal Analysis II, Volume 8*, Cham, 2014, pp. 71–81. doi: 10.1007/978-3-319-04774-4_7.
[69] N. E. Huang *et al.*, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998, doi: 10.1098/rspa.1998.0193.
[70] S. Braun and M. Feldman, “Decomposition of non-stationary signals into varying time scales: Some aspects of the EMD and HVD methods,” *Mech. Syst. Signal Process.*, vol. 25, no. 7, pp. 2608–2630, Oct. 2011, doi: 10.1016/j.ymssp.2011.04.005.
[71] Y. Lei, J. Lin, Z. He, and M. J. Zuo, “A review on empirical mode decomposition in fault diagnosis of rotating machinery,” *Mech. Syst. Signal Process.*, vol. 35, no. 1, pp. 108–126, Feb. 2013, doi: 10.1016/j.ymssp.2012.09.015.
[72] Y. Huang, “On the difference between empirical mode decomposition and Hilbert vibration decomposition for earthquake motion records,” 2012.

[73] U. Boz and M. Eriten, “Nonlinear system identification of soft materials based on Hilbert transform,” *J. Sound Vib.*, vol. 447, pp. 205–220, May 2019, doi: 10.1016/j.jsv.2019.01.025.

[74] P. F. Pai, “Time–frequency characterization of nonlinear normal modes and challenges in nonlinearity identification of dynamical systems,” *Mech. Syst. Signal Process.*, vol. 25, no. 7, pp. 2358–2374, Oct. 2011, doi: 10.1016/j.ymssp.2011.02.013.

[75] M. Kurt, M. Eriten, D. M. McFarland, L. A. Bergman, and A. F. Vakakis, “Strongly nonlinear beats in the dynamics of an elastic system with a strong local stiffness nonlinearity: Analysis and identification,” *J. Sound Vib.*, vol. 333, no. 7, pp. 2054–2072, Mar. 2014, doi: 10.1016/j.jsv.2013.11.021.

[76] P. Frank Pai, B.-A. Nguyen, and M. J. Sundaresan, “Nonlinearity identification by time-domain-only signal processing,” *Int. J. Non-Linear Mech.*, vol. 54, pp. 85–98, Sep. 2013, doi: 10.1016/j.ijnonlinmec.2013.04.002.

[77] K. J. Moore, M. Kurt, M. Eriten, D. M. McFarland, L. A. Bergman, and A. F. Vakakis, “Wavelet-bounded empirical mode decomposition for measured time series analysis,” *Mech. Syst. Signal Process.*, vol. 99, pp. 14–29, Jan. 2018, doi: 10.1016/j.ymssp.2017.06.005.

[78] M. Eriten, M. Kurt, G. Luo, D. Michael McFarland, L. A. Bergman, and A. F. Vakakis, “Nonlinear system identification of frictional effects in a beam with a bolted joint connection,” *Mech. Syst. Signal Process.*, vol. 39, no. 1, pp. 245–264, Aug. 2013, doi: 10.1016/j.ymssp.2013.03.003.

[79] P. F. Pai and A. N. Palazotto, “Detection and identification of nonlinearities by amplitude and frequency modulation analysis,” *Mech. Syst. Signal Process.*, vol. 22, no. 5, pp. 1107–1132, Jul. 2008, doi: 10.1016/j.ymssp.2007.11.006.

[80] A. F. Vakakis, L. A. Bergman, D. M. McFarland, Y. S. Lee, and M. Kurt, “Current efforts towards a non-linear system identification methodology of broad applicability,” *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 225, no. 11, pp. 2497–2515, Nov. 2011, doi: 10.1177/0954406211417217.

[81] D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen.* Leipzig: B.G. Teubner, 1912.

[82] D. Gabor, “Theory of communication. Part 1: The analysis of information,” *J. Inst. Electr. Eng. - Part III Radio Commun. Eng.*, vol. 93, no. 26, pp. 429–441, Nov. 1946, doi: 10.1049/ji-3-2.1946.0074.

[83] E. Bedrosian, “The Analytic Signal Representation of Modulated Waveforms,” *Proc. IRE*, vol. 50, no. 10, pp. 2071–2076, Oct. 1962, doi: 10.1109/JRPROC.1962.288236.

[84] E. Bedrosian, “A product theorem for Hilbert transforms,” *Proc. IEEE*, vol. 51, no. 5, pp. 868–869, May 1963, doi: 10.1109/PROC.1963.2308.

[85] A. H. Nuttall and E. Bedrosian, “On the quadrature approximation to the Hilbert transform of modulated signals,” *Proc. IEEE*, vol. 54, no. 10, pp. 1458–1459, Oct. 1966, doi: 10.1109/PROC.1966.5138.

[86] C. Pachaud, T. Gerber, N. Martin, and C. Mailhes, “Consequences of non-respect of the Bedrosian theorem when demodulating,” in *CM & MFPT 2013*,

Krakow, Poland, Jun. 2013, p. 13. [Online]. Available: https://oatao.univ-toulouse.fr/15163/1/pachaud_15163.pdf

[87] M. R. Shahriar, P. Borghesani, R. B. Randall, and A. C. C. Tan, “An assessment of envelope-based demodulation in case of proximity of carrier and modulation frequencies,” *Mech. Syst. Signal Process.*, vol. 96, pp. 176–200, Nov. 2017, doi: 10.1016/j.ymssp.2017.04.020.

[88] S. L. Hahn, *Hilbert Transforms in Signal Processing*. Boston: Artech House, 1996.

[89] Z. Wang and G. Chen, “Analytical mode decomposition with Hilbert transform for modal parameter identification of buildings under ambient vibration,” *Eng. Struct.*, vol. 59, pp. 173–184, Feb. 2014, doi: 10.1016/j.engstruct.2013.10.020.

[90] K. Dragomiretskiy and D. Zosso, “Variational Mode Decomposition,” *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014, doi: 10.1109/TSP.2013.2288675.

[91] N. E. Huang and S. S. Shen, Eds., *Hilbert-Huang transform and its applications*, 2nd edition. [Hackensack] New Jersey: World Scientific, 2014.

[92] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. Amsterdam Heidelberg: Academic Press, 2016.

[93] M. Müller, *Fundamentals of Music Processing: Using Python and Jupyter Notebooks*. Cham: Springer International Publishing, 2021. doi: 10.1007/978-3-030-69808-9.

[94] H. Dishan, “A Wavelet-Based Algorithm for the Hilbert Transform,” *Mech. Syst. Signal Process.*, vol. 10, no. 2, pp. 125–134, Mar. 1996, doi: 10.1006/mssp.1996.0009.

[95] Y.-T. Sheen, “3D spectral analysis for vibration signals by wavelet-based demodulation,” *Mech. Syst. Signal Process.*, vol. 20, no. 4, pp. 843–853, May 2006, doi: 10.1016/j.ymssp.2005.08.031.

[96] H. Olkkonen, P. Pesola, and J. T. Olkkonen, “Computation of Hilbert Transform via Discrete Cosine Transform,” *J. Signal Inf. Process.*, vol. 01, no. 01, pp. 18–23, 2010, doi: 10.4236/jsip.2010.11002.

[97] Y. Luo, S. Al-Dossary, M. Marhoon, and M. Alfaraj, “Generalized Hilbert transform and its applications in geophysics,” *Lead. Edge*, vol. 22, no. 3, pp. 198–202, Mar. 2003, doi: 10.1190/1.1564522.

[98] G. Zhang, Y. Li, T. Wang, H. Du, F. Luo, and Y. Zhan, “Extended Hilbert Transform and Application for Seismic Attributes,” *Earth Space Sci.*, vol. 6, no. 5, pp. 873–886, May 2019, doi: 10.1029/2019EA000551.

[99] W. Lu and C.-K. Zhang, “Robust estimation of instantaneous phase using a time-frequency adaptive filter,” *GEOPHYSICS*, vol. 78, no. 1, pp. O1–O7, Jan. 2013, doi: 10.1190/geo2011-0435.1.

[100] H. Sattari, “High-resolution seismic complex trace analysis by adaptive fast sparse S-transform,” *GEOPHYSICS*, vol. 82, pp. V51–V67, Jan. 2017, doi: 10.1190/GEO2015-0425.1.

[101] M. K. Kakhki, W. J. Mansur, and K. Aghazadeh, “Robust and high-resolution seismic complex trace analysis,” *Acta Geophys.*, vol. 68, no. 6, pp. 1689–1701, Dec. 2020, doi: 10.1007/s11600-020-00499-w.

[102] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank, “On Instantaneous Frequency,” *Adv. Adapt. Data Anal.*, vol. 01, no. 02, pp. 177–229, Apr. 2009, doi: 10.1142/S1793536909000096.

[103] J. H. Ginsberg, *Mechanical and Structural Vibrations: Theory and Applications*, 1st ed. New York: John Wiley and Sons, 2001.

[104] H. Sumali and R. A. Kellogg, “Calculating Damping from Ring-Down Using Hilbert Transform and Curve Fitting.,” Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), SAND2011-1960C, Mar. 2011. Accessed: Feb. 23, 2021. [Online]. Available: https://www.osti.gov/biblio/1108644-calculating-damping-from-ring-down-using-hilbert-transform-curve-fitting

[105] G. H. Hardy, “On Hilbert transforms,” *Q. J. Math.*, vol. os-3, no. 1, pp. 102–112, Jan. 1932, doi: 10.1093/qmath/os-3.1.102.

[106] F. R. Kschischang, “The Hilbert Transform.” Mar. 10, 2015. Accessed: May 12, 2022. [Online]. Available: https://www.comm.utoronto.ca/frank/notes/hilbert.pdf

[107] L. Marple, “Computing the discrete-time ‘analytic’ signal via FFT,” *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2600–2603, Sep. 1999, doi: 10.1109/78.782222.

[108] H. Li, L. Li, and T. Qian, “Discrete-time analytic signals and Bedrosian product theorems,” *Digit. Signal Process.*, vol. 20, no. 4, pp. 982–990, Jul. 2010, doi: 10.1016/j.dsp.2009.11.002.

[109] F. J. Harris, “On the Use of Windows for Harmonic Analysis With the Discrete Fourier Transform,” *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978, doi: 10.1109/PROC.1978.10837.

[110] J. O. Smith, *Spectral Audio Signal Processing*. Stanford, Calif: Stanford University, CCRMA, 2011.

[111] H. Olofsen, “Blending functions based on trigonometric and polynomial approximations of the Fabius function,” *Nor. IKT-Konf. Forsk. Og Utdanning*, Nov. 2019, Accessed: Feb. 15, 2022. [Online]. Available: https://ojs.bibsys.no/index.php/NIK/article/view/644

[112] B. Xu, Y. Sheng, P. Li, Q. Cheng, and J. Wu, “Causes and Classification of EMD Mode Mixing,” *Vibroengineering PROCEDIA*, vol. 22, pp. 158–164, Mar. 2019, doi: 10.21595/vp.2018.20250.

[113] Abraham. Savitzky and M. J. E. Golay, “Smoothing and Differentiation of Data by Simplified Least Squares Procedures.,” *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964, doi: 10.1021/ac60214a047.

[114] W. Smith, “Least-squares time-domain method for simultaneous identification of vibration parameters from multiple free-response records,” in *22nd Structures, Structural Dynamics and Materials Conference*, 0 vols., American Institute of Aeronautics and Astronautics, 1981. doi: 10.2514/6.1981-530.

[115] W. Hallauer, W. Slemp, and R. Kapania, “MATLAB® Curve-fitting for Estimation of Structural Dynamic Parameters,” Oct. 05, 2007.

[116] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust Region Methods*. Society for Industrial and Applied Mathematics, 2000.
[117] J. J. Moré and D. C. Sorensen, “Computing a Trust Region Step,” *SIAM J. Sci. Stat. Comput.*, vol. 4, no. 3, pp. 553–572, Sep. 1983, doi: 10.1137/0904038.
[118] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” *Q. Appl. Math.*, vol. 2, no. 2, pp. 164–168, 1944, doi: 10.1090/qam/10666.
[119] D. W. Marquardt, “An Algorithm for Least-Squares Estimation of Nonlinear Parameters,” *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, Jun. 1963, doi: 10.1137/0111030.
[120] R. J. Barnes, “Matrix Differentiation,” *Springs J.*, 2006, [Online]. Available: https://atmos.washington.edu/~dennis/MatrixCalculus.pdf
[121] J. Fox and S. Weisberg, *An R Companion to Applied Regression*. SAGE Publications, 2018. [Online]. Available: https://books.google.com/books?id=k_NrDwAAQBAJ
[122] W. D. Iwan, “A Distributed-Element Model for Hysteresis and Its Steady-State Dynamic Response,” *J. Appl. Mech.*, vol. 33, no. 4, pp. 893–900, Dec. 1966, doi: 10.1115/1.3625199.
[123] E. Fischer, “Sine beat vibration testing related to earthquake response spectra,” 1971, no. 2.
[124] A. Morrone, “Seismic vibration testing with sine beats,” *Nucl. Eng. Des.*, vol. 24, no. 3, pp. 344–356, Mar. 1973, doi: 10.1016/0029-5493(73)90005-8.
[125] F. Harris, “On the Use of Windows for Harmonic Analysis With the Discrete Fourier Transform,” *Proc. IEEE*, vol. 66, pp. 51–83, Feb. 1978, doi: 10.1109/PROC.1978.10837.
[126] D. Roettgen, B. R. Pacini, and R. Mayes, “Techniques for Nonlinear Identification and Maximizing Modal Response,” in *Nonlinear Structures and Systems, Volume 1*, Cham, 2020, pp. 173–188. doi: 10.1007/978-3-030-12391-8_24.
[127] D. Shetty and M. S. Allen, “Fast Simulation of a Single Degree-of-Freedom system consisting of an Iwan element using the Method of Averaging,” *ASME J. Vib. Acoust.*, vol. 142, no. 5, Oct. 2020, doi: https://doi.org/10.1115/1.4047070.
[128] B. J. Deaner, M. S. Allen, M. J. Starr, and D. J. Segalman, “Investigation of Modal Iwan Models for Structures with Bolted Joints,” in *Topics in Experimental Dynamic Substructuring, Volume 2*, New York, NY, 2014, pp. 9–25. doi: 10.1007/978-1-4614-6540-9_2.
[129] R. L. Mayes and S. E. Klenke, “The SMAC Modal Parameter Extraction Package,” 1999, pp. 812–818.
[130] R. L. Mayes and S. E. Klenke, “Automation and Other Extensions of the SMAC Modal Parameter Extraction Package,” 2000.
[131] R. L. Mayes, L. R. Dorrell, and S. E. Klenke, “Applications of the Automated SMAC Modal Parameter Extraction Package,” 2000.
[132] Q. Zhang, R. J. Allemang, and D. L. Brown, “Modal Filter: Concept and Applications,” 1990, pp. 487–496.

[133] S. M. Spottswood and R. J. Allemang, “On the investigation of some parameter identification and experimental modal filtering issues for nonlinear reduced order models,” *Exp. Mech.*, vol. 47, no. 4, Art. no. 4, 2007.
[134] T. P. Sapsis, D. Dane Quinn, A. F. Vakakis, and L. A. Bergman, “Effective Stiffening and Damping Enhancement of Structures With Strongly Nonlinear Local Attachments,” *J. Vib. Acoust.*, vol. 134, no. 011016, Jan. 2012, doi: 10.1115/1.4005005.
[135] M. Krack, “Nonlinear modal analysis of nonconservative systems: Extension of the periodic motion concept,” *Comput. Struct.*, vol. 154, pp. 59–71, Jul. 2015, doi: 10.1016/j.compstruc.2015.03.008.
[136] M. Scheel, S. Peter, R. I. Leine, and M. Krack, “A phase resonance approach for modal testing of structures with nonlinear dissipation,” *J. Sound Vib.*, vol. 435, pp. 56–73, Nov. 2018, doi: 10.1016/j.jsv.2018.07.010.
[137] F. Müller, L. Woiwode, J. Gross, M. Scheel, and M. Krack, “Nonlinear damping quantification from phase-resonant tests under base excitation,” *Mech. Syst. Signal Process.*, vol. 177, p. 109170, Sep. 2022, doi: 10.1016/j.ymssp.2022.109170.
[138] G. Abeloos *et al.*, “A consistency analysis of phase-locked-loop testing and control-based continuation for a geometrically nonlinear frictional system,” *Mech. Syst. Signal Process.*, vol. 170, p. 108820, May 2022, doi: 10.1016/j.ymssp.2022.108820.
[139] M. Kwarta and M. S. Allen, “Nonlinear Normal Mode backbone estimation with near-resonant steady state inputs,” *Mech. Syst. Signal Process.*, vol. 162, p. 108046, Jan. 2022, doi: 10.1016/j.ymssp.2021.108046.
[140] C. De Boor, *A Practical Guide to Splines*, Rev. ed. New York: Springer, 2001.
[141] P. H. C. Eilers and B. D. Marx, *Practical Smoothing: The Joys of P-Splines*. Cambridge, UK ; New York, NY: Cambridge University Press, 2021.
[142] Z. Li and J. Cao, “General P-Splines for Non-Uniform B-Splines.” arXiv, Apr. 07, 2022. doi: 10.48550/arXiv.2201.06808.
[143] B. Moldenhauer *et al.*, “Influences of Modal Coupling on Experimentally Extracted Nonlinear Modal Models,” 2020, pp. 189–204. doi: 10.1007/978-3-030-12391-8_25.
[144] A. H. Haslam *et al.*, “Non-linear system identification in the presence of modal coupling,” in *Nonlinear Dynamics, Conference Proceedings of the 36th IMAC*, Orlando, FL, 2017, p. 20.
[145] B. Moldenhauer, M. Allen, and D. Roettgen, “Variation of the Restoring Force Surface Method to Estimate Nonlinear Stiffness and Damping Parameters.,” Sandia National Laboratories, SAND2020-7255C, Jul. 2020. [Online]. Available: https://www.osti.gov/biblio/1809200
[146] D. Shetty and M. S. Allen, “A General Iwan Element Derived from Quasi-Static Force-Displacement Data,” Feb. 2021, p. 9.