Modal Analysis with Compressive Measurements

Jae Young Park, Michael B. Wakin, and Anna C. Gilbert

Abstract—Structural Health Monitoring (SHM) systems are critical for monitoring aging infrastructure (such as buildings or bridges) in a cost-effective manner. Such systems typically involve collections of battery-operated wireless sensors that sample vibration data over time. After the data is transmitted to a central node, modal analysis can be used to detect damage in the structure. In this paper, we propose and study three frameworks for Compressive Sensing (CS) in SHM systems; these methods are intended to minimize power consumption by allowing the data to be sampled and/or transmitted more efficiently. At the central node, all of these frameworks involve a very simple technique for estimating the structure’s mode shapes without requiring a traditional CS reconstruction of the vibration signals; all that is needed is to compute a simple Singular Value Decomposition. We provide theoretical justification (including measurement bounds) for each of these techniques based on the equations of motion describing a simplified Multiple-Degree-Of-Freedom (MDOF) system, and we support our proposed techniques using simulations based on synthetic and real data.

Index Terms—Compressive Sensing, modal analysis, Structural Health Monitoring, singular value decomposition

I. INTRODUCTION

A. Structural Health Monitoring systems

Over the past decade, more than 5 million commercial buildings [1], 130 million housing units [2], and 0.6 million bridges [3] have been built in the United States. In any structure, damage caused over time by continuous use is inevitable. In order to maintain safely operable structures for as long as possible, periodic inspections are a must. When damage is detected, some structures can be repaired, while others must be taken out of service immediately.

Due to the quantity, size, and complexity of structures, the task of inspection is labor intensive, costly, and time consuming. Consequently, there have been significant efforts in the structural engineering community to automate this process. Structural Health Monitoring (SHM) systems are precisely designed to address this issue.

Although the details of SHM systems vary, some features are shared among many methods [4], [5]. A typical SHM system monitors an in-service structure in real-time. To do so, it makes use of a network of sensors installed on the structure to collect vibration data for damage detection. This may include strain data, acceleration data, velocity data, or displacement data. The acquired data from each sensor is transmitted over a network to the central data repository where damage detection algorithms are run to detect, localize, or classify possible damage in the structure.

An important part of damage detection is a process called modal analysis. This process is used to infer properties such as the modal frequencies, mode shapes, and modal damping ratios of the structure. Such modal parameters describe the vibrational characteristics when external forces such as wind, earthquakes, or vehicle loadings are applied to the structure. For example, if a structure is forced to vibrate close to a modal frequency, the shape of the structure’s vibration will be dominated by the corresponding mode shape. This vibration will eventually die out in the absence of external force, and the modal damping ratio will determine rate of decay.

Many damage detection algorithms make use of modal parameters to detect, localize, and assess the severity of damage. Briefly speaking, these methods rely on the notion that when a structure is damaged, its modal parameters will change. Assuming that one has the modal parameters from the time when the structure was healthy, these can be compared to the current estimates of modal parameters to judge whether or not damage has occurred. A comprehensive survey of damage detection methods is presented in [4], [5].

B. Wireless SHM systems

In the early designs of SHM systems, sensors were linked via coaxial cables that provided reliable communication to the central data repository. Power cables were coupled with the coaxial cables to provide the power to run the sensors. Unfortunately, it was often impractical to install a dense network of sensors along with all of the requisite cables. When only a few sensors could be installed, the accuracy of damage detection and analysis would be sacrificed.

As a way to overcome this issue, researchers have proposed to deploy wireless sensors on structures to acquire and transmit data to the central repository. With the rapid advancement of wireless technology and the ability to build sensors at low cost, it has become possible to quickly deploy a much denser network of sensors for a given budget.

The challenges of wireless sensor networks in SHM are quite different from those of wired sensor networks. In particular, a primary challenge in a wireless network is to maximize the life expectancy of the batteries that power the sensors. From this perspective, there are important questions that should be considered when designing a wireless sensor network for SHM: Should we compress the time data before we transmit it to the central data repository? This would allow us to transmit less information but would require additional power for processing. Or, would we save more battery life by sending the raw data itself without processing it at all?

In [6], the authors presented thorough answers to these questions in real world applications involving a certain wire-
less sensor. The authors found that on average, significant savings in battery can be expected by locally compressing and processing the data first compared to sending the raw time data. The main reason for this was that the radio drains much more power (almost 3× more) than the on-board processor. As long as the execution time of the local algorithm is reasonably short, it would be more beneficial to first process and compress, then transmit less information.

Another factor that contributes to the draining of the battery is the sampling process. The Nyquist-Shannon sampling theorem states that when sampling a signal, the sampling rate should be proportional to the maximum frequency content in the signal. For the purpose of modal analysis the highest frequency content is dictated by the highest modal frequency of the underlying structure. Intuitively, structures that are stiff and light will generally possess high modal frequencies, and for such structures we may have to sample at relatively high rates. Unfortunately, faster sampling requires more power. The total length of the sampling time span also plays an important role in optimizing the battery life. Obviously, it would be best to sample only for a certain amount of time and turn the sensor off once a sufficient amount of data has been collected.

### C. Compressive Sensing in wireless SHM systems

In light of these observations, we believe that wireless sensors equipped with Compressive Sensing (CS) architectures will be a perfect fit to improve the efficiency and accuracy of wireless SHM systems. The theory of CS has in recent years offered the great promise of efficiently capturing essential signal information with low-rate sampling protocols, often below the minimum rate required by the Nyquist sampling theorem [7], [8], [9]. The use of CS can dramatically reduce both the complexity of a sensor and the amount of data that must be stored and/or transmitted downstream.

Employing CS in wireless SHM systems would allow compressed data to be acquired directly without the need for local processing. This enables power savings in several ways: First, because the data can be collected at a rate lower than the Nyquist rate the physical demands on the sensing hardware will be lower. Second, because there is no need for local processing, there is no computational burden at the sensor. Third, because the data is acquired in compressed form, the amount of information that must be wirelessly transmitted to the central data repository is minimized. Finally, all of the sensors can acquire the time data in a completely disjoint fashion,[1] eliminating the need for sensors to communicate while compressing their data.

Mathematically, denoting the continuous-time displacement signal at each sensor as \( u_i(t) \), where \( i = 1, \ldots, N \), and \( N \) denotes the number of nodes, a CS architecture employing non-uniform sampling (which is just one possible CS protocol [9]) would simply sample at time points \( t_1, \ldots, t_M \) and transmit the resulting samples \( v_i(t_1), \ldots, v_i(t_M) \). The compression would come from the fact that \( M \) is smaller than the number of Nyquist samples obtained over an appropriately chosen total sampling time span. Typically, one would choose \( t_1, \ldots, t_M \) randomly.

Typically, when we undersample a signal compared to its Nyquist rate, we must solve an underdetermined system of equations in order to reconstruct the original signal (this requires a sparse model for the signal in some basis). There have been several papers involving the application of CS to SHM systems. In [10], the authors implement a prototype wireless sensor that computes the compressed measurements locally after the wireless sensor has acquired the time data. By sending both compressed measurements and the raw time measurements to the central node, the authors evaluate the performance of CS reconstruction of the raw data. It is shown that a number of measurements \( M \approx 0.8L \) is needed for an accurate reconstruction of the raw data, where \( L \) represents the length of the original Nyquist-rate sample vector. Once the time data is reconstructed, conventional modal analysis techniques are carried out that aid in damage detection. In [11], the authors also reconstruct the original time data from CS measurements obtained at each sensor. Similar to [10], the authors claim that a relatively large number of measurements are required for accurate reconstruction.

The main reason that the above methods require so many measurements is because the signals are simply not sparse enough in the Discrete Wavelet Transform (DWT) or Discrete Fourier Transform (DFT). This suggests that the model of sparsity may not be sufficiently strong for the individual signals that arise in modal analysis. A joint signal model for the entire signal ensemble could serve as a better model to exploit the correlations that potentially exist across the signals.

At this point it is worth asking whether signal reconstruction is necessary at all when employing CS in wireless SHM systems. The only reason that the above methods attempt to reconstruct the original signals is that conventional modal analysis tools are designed to operate on signal samples captured at or above the Nyquist rate. There are two reasons to question whether this approach is optimal. The first reason is the potential loss of performance. The frequency content within the signals plays an important role in the subsequent modal analysis. For example, a popular modal analysis algorithm proposed in [12] is very sensitive to the accuracy of each frequency component of the signal. As we anticipate noise in our acquired signals, the reconstruction of each signal will also be noisy. These errors will propagate to the modal analysis step, which could potentially lead to misidentification and errors in the estimated modal parameters.

The second reason is the added computational complexity. Taking the above method for example, using an off-the-self CS reconstruction algorithm presented in [13], the total computational complexity for \( N \) signals will scale as \( N L \log^2(L) \). Added to this will be the computation required for the actual modal analysis. One can view the reconstruction step as being somewhat wasteful as it is carried out only to enable the use of conventional modal analysis algorithms.

We use the word “compression” to reflect the fact that fewer samples are transmitted. A deeper analysis—beyond the scope of this paper—would consider specific binary encodings of the raw and CS data and quantify the actual compression savings in terms of bits.

---

[1] In this paper, we assume that the time samples obtained at different sensors are all synchronized in time.
D. Contributions

The main objective of this paper is to propose a novel method for directly extracting the mode shapes from CS measurements without the need for reconstructing the individual time signals. Our proposed method differentiates itself from the previously proposed CS-based methods in that it exploits the joint signal structure that arises from the equations describing a simplified (no damping and free-decay) Multiple-Degree-Of-Freedom (MDOF) system.

Our method is as simple as computing the Singular Value Decomposition (SVD) of the signal matrix obtained by stacking each sample vector \( \{ v_1 \} = \{ v_1(t_1), \ldots , v_1(t_M) \}^* \) into a matrix denoted as \( [V] = \{ \{ v_1 \}, \ldots , \{ v_n \} \}^* \). Here, the superscript * denotes the conjugate transpose operator (we explain the use of complex-valued data in Section III-A). We evaluate the performance of this method both when \( t_1, \ldots , t_M \) are chosen deterministically with uniform spacing and when \( t_1, \ldots , t_M \) are chosen uniformly at random. We also evaluate the performance when we compute the SVD of \([Y]\) which is formed by choosing \( t_1, \ldots , t_M \) as uniform deterministic time points to form \([V]\) and subsequently multiplying \([V]\) by an \( M \times M' (M' < M) \) random matrix \([\Phi]\) such that \( [Y] = [V][\Phi] \).

For each sampling method, we give sufficient conditions on the required sampling rate, the total sampling time span, and the total number of measurements for accurate recovery of mode shape vectors. Our analysis reveals that the requisite sampling rate for uniform sampling can be lower than the Nyquist rate, but that the required number of samples is structure dependent. For random sampling, our analysis reveals that the required number of samples is structure independent and that we can achieve the same recovery guarantee as for uniform sampling, but the number of samples has a slightly increased dependence on the number of sensor nodes. Finally, our analysis for the scheme involving uniform sampling followed by random matrix multiplication shows that the requisite number of measurements (the number of columns of \( \Phi \)) is dependent on the rank of \([V]\). At the end of this paper, we present promising simulation results showing that our methods can accurately estimate the mode shapes using a number of samples or measurements that is only a small fraction of the original signal length.

II. BACKGROUND

In this section, we give an introduction to the frequently used mathematical model that governs the motion of structures. These equations form the basis of our proposed method. We begin with the simple Single-Degree-Of-Freedom (SDOF) system and then move on to the MDOF system. Following what is standard in the structural dynamics community, we use \( \{ x \} \) to denote a vector \( x \) and \( [A] \) to denote a matrix \( A \). We denote the \( l \)th entry of \( \{ x \} \) as \( \{ x \}(l) \), and the entry of \( [A] \) in the \( l \)th row and \( m \)th column as \( [A]_{l,m} \). Furthermore, we reserve \( i = \sqrt{-1} \) to denote the imaginary unit.

A. Single-degree-of-freedom system

An SDOF system under no external force can be described by the following differential equation:

\[
m \ddot{x}(t) + C \dot{x}(t) + kx(t) = 0,
\]

where \( m, C, \) and \( k \) denote the mass, damping, and stiffness parameters of the underlying system. To solve for the displacement signal \( x(t) \) that satisfies the above equation, let us assume a solution of the form \( x(t) = Ae^{\imath \omega t} \), where \( A, s \in \mathbb{C} \). Then, \( \ddot{x}(t) = As^2e^{\imath \omega t}, \dot{x}(t) = Ase^{\imath \omega t}, \) and \( x(t) = Ae^{\imath \omega t} \), and by plugging in these expressions into equation (1) we get \( (ms^2 + Cs + k) x(t) = 0 \). Since this needs to be satisfied for all \( t \), it must be that \( ms^2 + Cs + k = 0 \), and it is easy to see that \( s = \frac{-C+\sqrt{C^2-4mk}}{2m} \). In the structural dynamics community it is customary to rewrite this as

\[
s = -\xi \omega_0 \pm \omega_0 \sqrt{\xi^2 - 1},
\]

where \( \omega_0 = \sqrt{\frac{k}{m}} \) and \( \xi = \frac{C}{2m \omega_0} \) represent the natural frequency and damping ratio, respectively. As we can see, the natural frequencies and damping ratios will always be positive, and depending on the value of \( \xi, s \) may be real or complex and there may be one or two possible solutions.

In this paper, we restrict ourselves to the case when there is no damping, i.e., \( C = 0 \), and thus \( \xi = 0 \). From equation (2) we can see that when \( \xi = 0 \), we have two purely imaginary solutions \( s_1 = i\omega_0 \), and \( s_2 = -i\omega_0 \). Thus, \( x_1(t) = Ae^{i\omega_0 t} \) and \( x_2(t) = Be^{-i\omega_0 t} \) are both eligible solutions to equation (1).

In fact, any solution to the above equation can be expressed as a linear combination of \( x_1(t) \) and \( x_2(t) \), such that the general solution can be written as \( x(t) = Ae^{i\omega_0 t} + Be^{-i\omega_0 t} = (A + B) \cos(\omega_0 t) + i(A - B) \sin(\omega_0 t) \).

Furthermore, noting that \( x(0) = A + B \) and \( \dot{x}(0) = i(A - B)\omega_0 \), it follows that \( x(t) = x(0) \cos(\omega_0 t) + \frac{x(0)}{\omega_0} \sin(\omega_0 t) \).

Since we want to deal with real valued signals \( x(t) \), this demands that \( B = A' \). Denoting \( A = a + ib \), we can once again rewrite the solution as \( x(t) = 2a \cos(\omega_0 t) - 2b \sin(\omega_0 t) \).

Finally, because any linear combination of sines and cosines with the same frequency is also a sine wave with the same frequency, we can rewrite this as \( x(t) = \rho \sin(\omega_0 t + \theta) \), where

\[
\rho = 2\sqrt{a^2 + b^2} \quad \text{and} \quad \theta = \begin{cases} \arcsin \left( \frac{a}{\sqrt{a^2 + b^2}} \right), & b \leq 0, \\ \pi - \arcsin \left( \frac{a}{\sqrt{a^2 + b^2}} \right), & b > 0. \end{cases}
\]

B. Multiple-degree-of-freedom system

Similar to the SDOF system, an \( N \)-degree MDOF system\(^3\) can be formulated as

\[
[M]\{\ddot{u}(t)\} + [C]\{\dot{u}(t)\} + [K]\{u(t)\} = \{0(t)\},
\]

where \([M]\) is an \( N \times N \) diagonal mass matrix, \([C]\) is a symmetric \( N \times N \) damping matrix, \([K]\) is an \( N \times N \) symmetric matrix.

\(^3\)Theoretically, a structure will have infinitely many degrees of freedom. However, the number of mode shapes that one can detect is equal to the number of sensors placed on the structure. In the following, whenever we deal with an \( N \)-degree MDOF system, we are implicitly assuming that we have \( N \) sensor nodes deployed on the structure.
stiffness matrix, and \( \{u(t)\} \) is an \( N \times 1 \) vector of displacement signals. Note that \( \{u(t)\} = \{u_1(t), \ldots, u_N(t)\} \), and each \( \{u(t)\}(l) = u_l(t), l = 1, \ldots, N, \) is a displacement signal. One can view \( u_l(t) \) as the signal being observed at the \( l \)th sensor node.

Again we consider an undamped system and set \( [C] = [0] \). This simplifies the above equation to \( [M]\{\ddot{u}(t)\} + [K]\{u(t)\} = \{0(t)\} \). Let us assume \( \{u(t)\} = \{\psi\}Ae^{i\omega t} \) to be a solution to this equation, where \( \omega, t \in \mathbb{R} \). Here, \( \{\psi\} \) is an \( N \times 1 \) spatial vector that is independent of time; we define it to have unit energy, i.e., \( \|\{\psi\}\|^2 = 1 \) (we can assume this as the normalization can be absorbed into the scalar variable \( A \)). Plugging in the appropriate derivative to the above expression we get \( (-[M]\omega^2 + [K])\{\psi\}Ae^{i\omega t} = \{0(t)\} \). Since this must hold for all values of \( t \) it must be that

\[
\{[K] - \omega^2[M]\}\{\psi\} = \{0\}.
\]

The above represents a generalized eigenvalue problem and our objective is to find pairs of \( \omega^2 \) and \( \{\psi\} \) that satisfy this equation. Notice the similarity to the conventional eigenvalue problem, which corresponds to the case when \( [M] = I \). To solve the above problem, as in the conventional eigenvalue problem, one starts off by computing the generalized eigenvalue \( \omega^2 \) that satisfies

\[
\det\{[K] - \omega^2[M]\} = 0.
\]

Assuming that this does not vanish as a function of \( \omega^2 \) and that \( M \) is full rank, the left hand side of (5) will represent an \( N \)th order polynomial and there will be \( N \) generalized eigenvalues \( \omega_1, \ldots, \omega_N \), which are known as the modal frequencies, and \( N \) corresponding generalized eigenvectors \( \{\psi_1\}, \ldots, \{\psi_N\} \), which are known as the mode shape vectors. Each modal frequency will be real and positive, and the mode shape vectors will be orthonormal to one another. Without loss of generality, we assume the frequencies are sorted such that \( \omega_1 \geq \omega_2 \geq \cdots \geq \omega_N > 0 \).

It is clear that each \( \{\psi_n\}A_ne^{i\omega_nt}, n = 1, \ldots, N, \) will be a valid solution to the MDOF system equation. Furthermore, as in the SDOF system \( \{\psi_n\}B_ne^{-i\omega_nt} \) will also be a valid solution and thus for each \( n \) a complete solution will be of the form \( \{\psi_n\}(A_ne^{i\omega_nt} + B_ne^{-i\omega_nt}) \). We can guarantee this solution is real by ensuring that \( \{\psi_n\} \) is real (there exists a real eigenvector that satisfies the above equation given that the mass and stiffness matrices are real and symmetric) and \( B_n = A_n^* \). Thus, as with the SDOF case we can rewrite this solution as \( \{\psi_n\}\rho_n\sin(\omega_nt + \theta_n) \), where \( \rho_n \) and \( \theta_n \) are as defined in (3). Finally, it is easy to see that all linear combinations of this solution are valid solutions to the MDOF system equation and thus the general solution is of the form

\[
\{u(t)\} = \sum_{n=1}^{N} \{\psi_n\}\rho_n\sin(\omega_nt + \theta_n).
\]

In the structural dynamics community this is known as the modal superposition equation.

### III. Problem Formulation

#### A. The analytic signal of \( \{u(t)\} \)

Our framework and analysis will involve sampling what is known as the analytic signal of \( \{u(t)\} \) [13].

**Definition 1** (Definition 1.2.1, [14]). A signal \( v(t) \) is said to be analytic iff \( V(f) = 0 \) for \( f < 0 \), where \( V(f) \) is the Fourier transform of \( v(t) \).

An analytic signal can be obtained by removing all negative frequencies in a given signal. The analytic signal is a frequently used representation in mathematics, signal processing, and communications; in some problems (such as ours) it can simplify the mathematical manipulations.

To discuss the analytic signal of \( \{u(t)\} \), let us examine each entry in \( \{u(t)\} \), i.e., \( u_l(t) \). Based on the derivation in Section II-B each \( u_l(t) \) can be written as

\[
u_l(t) = \sum_{n=1}^{N}\{\psi_n\}(l)(A_ne^{i\omega_n t} + A_n^*e^{-i\omega_n t}).
\]

Thus, the analytic signal of \( u_l(t) \), represented as \( v_l(t) \), is simply

\[
v_l(t) = \sum_{n=1}^{N}\{\psi_n\}(l)A_ne^{i\omega_n t},
\]

and the analytic signal of the entire vector \( \{u(t)\} \), denoted as \( \{v(t)\} \), can be written as

\[
\{v(t)\} = \sum_{n=1}^{N}\{\psi_n\}A_ne^{i\omega_n t}.
\]

Note that \( \{v(t)\} \) is no longer real but complex.

Obtaining an analytic signal in practice involves the application of a Hilbert transform. However, detailed discussion of this matter is out of scope of this paper and we will refer interested readers to [15] for more information. For the remainder of this paper, we will assume that we have successfully extracted the analytic signal from each \( u_l(t) \). Thus, all derivations from here onwards will be in terms of \( \{v(t)\} \).

#### B. The relationship to the SVD

We can write the modal superposition equation (7) in matrix-vector multiplication format as

\[
\{v(t)\} = [\Psi][\Gamma]\{s(t)\},
\]

where \( [\Psi] = \{\{\psi_1\}, \{\psi_2\}, \ldots, \{\psi_N\}\} \) denotes the \( N \times N \) mode shape vector matrix, which as mentioned before, has orthonormal columns,

\[
[\Gamma] = \sqrt{M} \cdot \begin{bmatrix}
A_1 & 0 & \cdots & 0 \\
0 & A_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_N
\end{bmatrix}
\]
denotes an $N \times N$ diagonal matrix, and

$$\{s(t)\} = \frac{1}{\sqrt{M}} \cdot \begin{bmatrix} e^{i\omega_1 t} \\ e^{i\omega_2 t} \\ \vdots \\ e^{i\omega_N t} \end{bmatrix}$$

denotes an $N \times 1$ modal coordinate vector.

In order to see how the SVD could be useful in extracting the modal parameters, let us suppose that we sample each row of $\{v(t)\}$ at $M$ distinct points in time $t_1, \ldots, t_M$. We assume $M \geq N$ and denote the resulting $N \times M$ data matrix as

$$[V] = \begin{bmatrix} v_1(t_1) & v_1(t_2) & \cdots & v_1(t_M) \\ v_2(t_1) & v_2(t_2) & \cdots & v_2(t_M) \\ \vdots & \vdots & \ddots & \vdots \\ v_N(t_1) & v_N(t_2) & \cdots & v_N(t_M) \end{bmatrix} \in \mathbb{C}^{N \times M}. \quad (8)$$

The sampling of $\{v(t)\}$ at $t_1, \ldots, t_M$ implies the sampling of $\{s(t)\}$ at the exact same time points which leads us to define

$$[S] = \frac{1}{\sqrt{M}} \cdot \begin{bmatrix} e^{i\omega_1 t_1} & e^{i\omega_1 t_2} & \cdots & e^{i\omega_1 t_M} \\ e^{i\omega_2 t_1} & e^{i\omega_2 t_2} & \cdots & e^{i\omega_2 t_M} \\ \vdots & \vdots & \ddots & \vdots \\ e^{i\omega_N t_1} & e^{i\omega_N t_2} & \cdots & e^{i\omega_N t_M} \end{bmatrix} \in \mathbb{C}^{N \times M}$$

and allows us to write the matrix of samples as

$$[V] = [\Psi][\Gamma][S]. \quad (9)$$

Equation (2) makes explicit the relationship between the SVD and the modal parameters. We know that $[\Psi]$ is a square matrix with orthonormal columns, and $[\Gamma]$ is a diagonal matrix. Hypothetically, if $[S]$ happened to be a matrix with orthogonal (or orthonormal) rows, then equation (9) would precisely describe the SVD of $[V]$. In that case, one could obtain the modal parameters by simply computing the SVD of $[V]$! As an example, the rows of $[S]$ would be perfectly orthogonal if they happened to equal $N$ distinct length-$M$ DFT vectors. This could be ensured if we sampled at uniform times $t_m = T_s(m - 1)$, where $T_s$ is a sampling interval, $m \in \{1, \ldots, M\}$, and $M \geq N$, but it would require the modal frequencies to lie on a grid such that $\omega_n = \frac{2\pi k_n}{MT_s}$, where $k_n \in \{1, \ldots, M\}$. If these conditions were satisfied, the SVD of $[V]$ would exactly recover the modal parameters. Unfortunately, this is an unrealistic model for the purpose of modal analysis because the modal frequencies will typically not lie on a grid.

If we drop the assumption that the modal frequencies lie on a grid, the problem becomes much more complicated, and in general the rows of $[S]$ will not be orthogonal. However, our main results, presented in Section IV, rely on characterizing sampling strategies that ensure the rows of $[S]$ will still be nearly orthogonal and showing in these situations that the mode shape vectors can be accurately estimated by computing the SVD of $[V]$.

C. SVD in modal analysis

Among the many techniques that have been proposed for modal analysis, we briefly mention a few that also make use of the SVD. The Ibrahim Time Domain (ITD) [10] method shares a number of similarities with our proposed method. This method also begins with the MDOF model and sets up an equation that relates the modal parameters to the observations akin to (6). After further algebraic manipulations, a set of equations reveals that one can extract the modal parameters via an eigendecomposition. Based on this observation, the ITD method obtains estimates of the modal parameters by computing the eigendecomposition of a matrix that is a function of the observed data matrix. In the Frequency Domain Decomposition (FDD) [12] method, the observations are used to compute cross power spectral density estimates. These cross power spectral density estimates are collectively a 3D data cube that consists of a 2D cross spectral matrix at each frequency. Given these estimates, the SVD is used to extract the singular vectors of the 2D cross spectral matrices. These singular vectors provide estimates of the mode shape vectors. The Eigensystem Realization Algorithm (ERA) [17] method makes use of ideas in control theory and sets up the problem with a state-space equation of an MDOF system. The main use of the SVD in this method is to decompose the Hankel matrix that can be constructed from measured impulse response data. The singular vectors and singular values of the Hankel matrix are then manipulated in order to form a system matrix describing the underlying system. The estimates of the modal parameters are then computed via an eigendecomposition on this system matrix.

All of these methods implicitly assume that the observations are sampled at uniform intervals at or above the Nyquist rate. Thus, these methods may not be directly applicable when the observations are sampled in a random fashion. Furthermore, to the best of our knowledge none of these methods are accompanied by error analysis or instructions on how long to sample the vibration signal. In the next section, we present our proposed method along with detailed analysis providing error bounds and sufficient conditions on how to sample in order to guarantee faithful recovery of the mode shape vectors.

IV. MAIN RESULTS

In this section, we present our main results. We propose three measurement schemes—uniform time sampling, random time sampling, and uniform time sampling followed by a random matrix multiplication—and for each measurement scheme we provide a sufficient condition for the accurate recovery of mode shapes via the SVD. Proofs of all of our results appear in the Appendix.

The main focus of this paper is the recovery of the mode shape vectors. However, in Section V we do provide a short discussion along with promising simulation results concerning the recovery of the modal frequencies.

In the results that follow, we use $\delta_{\text{min}}$ and $\delta_{\text{max}}$ to denote lower and upper bounds on the minimum and maximum separation of the modal frequencies. In other words, we assume that $\delta_{\text{min}} \leq |\omega_i - \omega_n|$ and $\delta_{\text{max}} \geq |\omega_i - \omega_n|$. Furthermore, we use $t_{\text{max}}$ to denote the total sampling time span. Finally for $a, b \in [0, 1]$, the quantity $D(a||b) := a(\log(a) - \log(b)) + (1-a)(\log(1-a) - \log(1-b))$ is known
and leads to a better preservation of the mode shapes but orthogonal; a small choice of variable \( \delta \) implies more orthogonal rows and thus are out of our control. In order to guarantee a small error in the \( n \)th mode shape when \( |A_n| \) is close to some other \( |A_l| \), one would need to make \( \epsilon \) smaller.

Turning our attention to the sampling parameters, the above theorem essentially tells us that we need to sample for a time span that is inversely proportional to the minimum spacing between the modal frequencies. Thus, the smaller the minimum spacing between the modal frequencies, the longer we must sample to get an accurate estimate. Also, since \( T_s = \frac{\pi}{\delta_{\text{max}} \omega_{\text{max}}} \), the maximum spacing between modal frequencies determines how fast we need to sample. Comparing this sampling interval to the Nyquist sampling interval which would be \( T_0 = \frac{\pi}{\omega_{\text{max}}} \), it is interesting to note that \( T_s > T_0 \). This suggests that for the purpose of mode shape extraction, we can potentially sample at a rate lower than the Nyquist rate and still accurately recover the mode shapes. However, it is important to note that in order to sample with \( T_s \) we must know in advance the maximum separation between modal frequencies. In scenarios where \( \delta_{\text{max}} \) is unknown it would be more reasonable to sample at a sufficiently small interval (\( T_s \approx T_0 \)) to ensure the sampling conditions are satisfied. Finally, note that the condition on \( M \) is fairly satisfactory in its logarithmic dependence on \( N \) and its linear dependence on \( \frac{1}{\epsilon} \) (this assumes the left hand term dominates in (11)). However, it also scales with the ratio \( \delta_{\text{max}} \), which depends on the structure. For some structures this ratio could in fact be large, and in the absence of additional information about the structure, one may need to assume this ratio is large. This motivates our second sampling strategy, which appears below.

2) Random time sampling: Let us now suppose that we sample \( \{v(t)\} \) uniformly at \( M \) time points \( t_1, \ldots, t_M \) chosen uniformly at random in the time interval \([0, t_{\text{max}}]\). We can establish the following theorem.

**Theorem 3.** Let \([V] = [\Psi][\Gamma][S]\) be as given in (9) describing an \( N \)-degree-of-freedom system sampled according to the random sampling scheme described above. Suppose we sample for a total time span of at least

\[
t_{\text{max}} \geq \frac{40(\log(N/2) + 1.01)}{\epsilon \omega_{\text{min}}} \tag{13}
\]

and within this time span suppose we take a number of measurements satisfying

\[
M > \max \left( \frac{\log(N) + \log(2/\tau)}{\min(D_1, D_2)}, N \right) \tag{14}
\]

where

\[
D_1 = D((1 + \epsilon)/N)((1 + \epsilon)/10)/N), \quad D_2 = D((1 - \epsilon)/N)((1 - \epsilon)/10)/N). \tag{15}
\]

Then with probability at least \( 1 - \tau \) all of the mode shape estimates \( \{\hat{\Psi}\} \) obtained via Algorithm 1 will satisfy (12).

This result for random time sampling looks somewhat similar to Theorem 2 for uniform time sampling. The recovery guarantee is the same and the required time span differs

---

**Algorithm 1: Pseudo-code for mode shape estimation**

**Input:** Data matrix \([V]\) as defined in (8)

**Output:** \([\hat{\Psi}]\) (left singular vectors of \([V]\))

\[\text{SVD}(V) = [\Psi][\Gamma][S]\]

as the binary information divergence, or the Kullback-Leibler divergence [18].

**A. Uniform time sampling and random time sampling**

Our proposed method for recovering the mode shape vectors from uniform time samples or from random time samples is very simple and is described in Algorithm 1. In particular, our method simply computes the SVD of \([V]\) and returns the matrix of left singular vectors \([\hat{\Psi}] = \{\hat{\psi}_1, \ldots, \hat{\psi}_N\}\) as estimates of the true mode shape matrix \([\Psi]\). One point to note about this algorithm is that because \([V]\) is \( N \times M \), where we assume \( M \geq N \), the dimensions of \([\Psi]\), \([\Gamma]\), and \([S]\) will be \( N \times N \), \( N \times M \), and \( M \times M \), respectively. These differ from the dimensions of their respective counterparts in (9). Taking a closer look and noting that only \( N \) diagonal entries in \([\Gamma]\) are non-zero, we can compute the truncated SVD to obtain the desired dimensions.

1) Uniform time sampling: Let us now suppose that we sample \( \{v(t)\} \) uniformly in time with a uniform sampling interval denoted by \( T_s \). The sampling times are given by \( t_m = (m-1)T_s \), \( m = 1, \ldots, M \). We are therefore sampling within the time span \([0, t_{\text{max}}]\), where \( t_{\text{max}} := (M-1)T_s \). We can establish the following theorem.

**Theorem 2.** Let \([V] = [\Psi][\Gamma][S]\) be as given in (9) describing an \( N \)-degree-of-freedom system sampled according to the uniform sampling scheme described above. For \( 0 < \epsilon < 1 \), suppose we sample for a total time span of at least

\[
t_{\text{max}} \geq \frac{2\pi(\log(N/2) + 1.01)}{\epsilon \delta_{\text{min}}} \tag{10}
\]

with sampling interval \( T_s = \frac{\pi}{\delta_{\text{max}}} \) and ensure that \( M \geq N \). Or, equivalently, suppose we take

\[
M \geq \max \left( \frac{2\log(N/2) + 1.01}{\epsilon} \delta_{\text{max}}, 1, N \right) \tag{11}
\]

total samples with the sampling interval \( T_s = \frac{\pi}{\delta_{\text{max}}} \). Then, the mode shape estimates \( \{\hat{\psi}\} \) obtained via Algorithm 1 satisfy

\[
\|\{\hat{\psi}_n\} - \{\hat{\psi}_n\}\| \leq \min \left\{ \sqrt{2}, \frac{\epsilon \sqrt{1 + \epsilon}}{\sqrt{1 - \epsilon}} \text{sepn}(\epsilon) \right\}, \tag{12}
\]

where

\[
\text{sepn}(\epsilon) = \max_{l \neq n} \min_{c \in [-1,1]} \left\{ ||A_l||A_n||, \sqrt{2} \| A_l \| A_n \| 2 + 1 + \epsilon c \| A_n \| \right\}.
\]

In (12) we see that the error in the \( n \)th estimated mode shape vector mainly depends on \( \epsilon \) and what is essentially the minimum separation between \( |A_n| \) and all other \( |A_l| \). The variable \( \epsilon \) controls how close the rows of \([S]\) are to being orthogonal; a small choice of \( \epsilon \) implies more orthogonal rows and leads to a better preservation of the mode shapes but requires more samples. Furthermore, the bigger the separation between \( |A_n| \) and all other \( |A_l| \), the better our estimate. Note that the parameters \( |A_n| \) are dependent on the underlying structure and thus are out of our control.
only by a constant. A critical difference, however, is that the requisite number of samples \( M \) no longer depends on the ratio \( \frac{\delta_{\text{max}}}{\delta_{\text{min}}} \). However, when \( N \) is large and \( \epsilon \) is small, the denominator in (14) will scale like \( \frac{1}{\sqrt{N}} \), and so the requisite number of samples will scale like \( \frac{N \log(N)}{\delta^2} \). This represents a stronger dependence on \( N \) compared to what appears in (11), but only by a logarithmic factor (because the right hand term in (11) scales like \( N \)). It also represents a stronger dependence on \( \epsilon \) compared to what appears in (11).

Ultimately, we see that in some cases random time sampling could provide a significant reduction in the number of samples for systems where \( \frac{\delta_{\text{max}}}{\delta_{\text{min}}} \) is large or unknown. For a given problem, the better choice between uniform and random time sampling may depend on the particular circumstances and the parameters of the system under study.

B. Uniform sampling followed by random matrix multiplication

The last measurement scheme that we consider involves taking uniform time samples and compressing these via multiplication by a random matrix. More specifically, let us form \([V]\) following the uniform sampling scheme as discussed in Section V-A with \( T_{\text{max}} \) and \( T_s \) as given in Theorem 2. Subsequently, we construct a random \( M \times M' \) matrix \([\Phi]\) and compute the \( N \times M' \) matrix \([Y] = [V]\[\Phi]\) of compressed measurements. We are specifically interested in cases where \( M' < M \), i.e., when \([Y]\) has fewer columns than \([V]\). We also note that \([\Phi]\) can be applied individually to each row of \([V]\) and the resulting measurements can be concatenated to form \([Y]\). This means that this CS measurement scheme can be performed sensor-by-sensor in a SHM system.

To state our results, we write the truncated SVD of \([Y]\) analogously to that of \([V]\) as \([Y] = [\Psi][\Gamma][\Sigma]\). The matrices \([\Psi] = \{\psi_1, \ldots, \psi_N\}\), \([\Gamma]\) and \([\Sigma]\) will be \( N \times N \), \( N \times M' \), and \( M' \times M' \), respectively. We also require the following definition.

**Definition 4.** A \( Q \times m \) random matrix \([\Phi]\) is said to satisfy the distributional JL property if for any fixed \( \{x\} \in \mathbb{C}^Q \) and any \( 0 < \epsilon < 1 \),

\[
\Pr \left[ \| (\Phi^*) \{x\} \|_2^2 - \| \{x\} \|_2^2 \right] > \epsilon \| \{x\} \|_2^2 \leq 4 e^{-m f(\epsilon)},
\]

where \( f(\epsilon) > 0 \) is a constant depending only on \( \epsilon \).

For most random matrices satisfying the distributional JL property, the functional dependence on \( \epsilon \), \( f(\epsilon) \), is quadratic in \( \epsilon \) as \( \epsilon \rightarrow 0 \). There are a variety of random matrix constructions known to possess the distributional JL property. Notably, random matrices populated with independent and identically distributed (i.i.d.) subgaussian entries will possess this property [19]. Subgaussian distributions include suitably scaled Gaussian and \( \pm 1 \) Bernoulli random variables.

We are now ready to state our next theorem.

**Theorem 5.** Let \([V] = [\Psi][\Gamma][\Sigma]\) be as given in (2) describing an \( N \)-degree-of-freedom system sampled according to the uniform sampling scheme with \( T_s \) and \( T_{\text{max}} \) as required by

\[
M' \geq \frac{2k \log(42/\epsilon) + \log(4/\delta)}{f(\epsilon/\sqrt{2})},
\]

where \( k \leq N \) represents the rank of \([V]\) and \( \epsilon \) represents the distortion factor of \([\Phi]\). Let \([Y] = [V][\Phi]\) and let \([\Gamma]\) and \([\Phi]\) denote the estimated singular values and left singular vectors of \([Y]\) returned by Algorithm 2 when we provide \([Y]\) in place of \([V]\) as the input matrix. Then, with probability exceeding \( 1 - \delta \) the mode shape estimates \([\hat{\Psi}]\) satisfy the following bound:

\[
\|\{\phi_n\} - \{\hat{\phi}_n\}\|_2 \leq \min \left\{ \sqrt{2}, \frac{\sqrt{1 + \epsilon}}{\sqrt{1 - \epsilon}} \cdot \text{sep}_n(\epsilon) + \frac{\epsilon' \sqrt{1 + \epsilon'}}{\sqrt{1 - \epsilon'}} \cdot \text{sep}'_n(\epsilon') \right\},
\]

where

\[
\text{sep}_n(\epsilon) = \max_{l \neq n} \min_{c \in [-1,1]} \{ |\sigma_n^2 - \sigma_l^2(1 + c \epsilon')| \}
\]

and \( \sigma_n \) are the singular values of \([V]\).

The error bound (15) looks similar to those appearing in Theorems 2 and 3 except that instead of having one term we now have two terms. The first term is essentially the error due to the uniform time sampling matrix \([V]\) and the second term is the error due to the multiplication by a random matrix \([\Phi]\). The required number of columns in the random matrix \([\Phi]\) is dependent on the rank \( k \) of \([V]\). The higher the rank of \([V]\), the more measurements we need; in the worse case one could assume \( k = N \). One could easily envision a scenario where this measurement scheme could be useful. For example, suppose we are dealing with a structure that has high \( \delta_{\text{max}} \) and small \( \delta_{\text{min}} \) (or suppose that we do not know these quantities and so we conservatively suppose they are large and small, respectively). This means that we need to take a large number \( M \) of uniform samples using a small sampling interval \( T_s \) over a long time duration \( t_{\text{max}} \). In such scenarios, one could choose to post-process the signals to reduce the number of measurements by multiplying each sample vector with \([\Phi]\). Another similar scenario is when we have a conventional uniform sensor over which we do not have control over the sampling interval. Again, we may use \([\Phi]\) as a way to reduce the number of measurements. In Section V-A we demonstrate the usefulness of this sampling scheme with a synthetic example.

V. EXPERIMENTAL RESULTS

A. Experiments with synthetic data

We begin by demonstrating the effectiveness of the various methods using an idealized synthetic dataset. The system that we consider is a 4-degree-of-freedom structure with no damping and under free vibration. For reproducibility the modal vectors are the eigenvectors of the following symmetric matrix:

\[
[Z] = \begin{bmatrix}
2 & -1 & 0 & 0 \\
0 & 2 & -1 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1
\end{bmatrix}.
\]
We use these mode shapes throughout this subsection. We also use the following $[\Gamma]$ matrix throughout this subsection:

$$[\Gamma] = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0.45 & 0 & 0 \\
0 & 0 & 0.15 & 0 \\
0 & 0 & 0 & 0.01 \\
\end{bmatrix}.$$ 

The modal superposition equation can be written as

$$\{v(t)\} = [\Psi] \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0.45 & 0 & 0 \\
0 & 0 & 0.15 & 0 \\
0 & 0 & 0 & 0.01 \\
\end{bmatrix} \begin{bmatrix}
e^{i\omega_1t} \\
e^{i\omega_2t} \\
e^{i\omega_3t} \\
e^{i\omega_4t} \\
\end{bmatrix} \{\omega(t)\} \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0.45 & 0 & 0 \\
0 & 0 & 0.15 & 0 \\
0 & 0 & 0 & 0.01 \\
\end{bmatrix}$$

where the modal frequencies $\omega_1, \ldots, \omega_4$ remain to be chosen.

In our first experiment, we demonstrate the uniform and random time sampling methods by plotting the errors of each of the four estimated mode shapes obtained from the SVD of the sampled matrix $[V]$. For both methods, we set $\omega_1 = 2.1\pi$, $\omega_2 = 4.28\pi$, $\omega_3 = 6.02\pi$, and $\omega_4 = 8.24\pi$ rad/s. For the uniform time sampling scheme, we use a sampling interval of $T_s = 0.1$s, which is just slightly faster than what our theorem prescribes. Using this fixed rate, we collect samples over a total time span of duration $t_{\text{max}}$, and we repeat the experiment for $t_{\text{max}} = [0 : T_s : 2]s$ (the value of $M$ therefore increases with $t_{\text{max}}$). For each value of $t_{\text{max}}$ and for $n = 1, \ldots, 4$, we plot the $\ell_2$ error $\|\{\psi_n\} - \{\hat{\psi}_n\}\|_2$ between the ground truth mode shape vector $\{\psi_n\}$ and the corresponding estimated vector $\{\hat{\psi}_n\}$ produced using the SVD on the data matrix $[V]$. The results are shown in the black curves in Figure 1(a). For the random time sampling scheme, we use the same values of $t_{\text{max}}$ and sample $\{v(t)\}$ uniformly at random within the interval $[0, t_{\text{max}}]$. For each value of $t_{\text{max}}$, the total number of samples $M$ we collect is chosen to equal the corresponding value of $M$ used for uniform time sampling above. The errors of the mode shape estimates are shown in the blue curves in Figure 1(a). We see that overall, the performance of the two sampling schemes is comparable. This is in agreement with Theorems 2 and 3 as they both suggest the same reconstruction guarantees given that we satisfy the sampling conditions.

Our second experiment highlights the role played by $\delta_{\text{min}}$ (the minimum separation of the modal frequencies). We consider two sets of modal frequencies. The first set is the same one used in our first experiment; for this set $\delta_{\text{min}} = 1.74\pi$, and recall the results plotted in Figure 1(a). For our second set we use $\omega_1 = 2.1\pi$, $\omega_2 = 4.28\pi$, $\omega_3 = 4.6\pi$, and $\omega_4 = 8.24\pi$ rad/s. This set has a smaller minimum separation between the modal frequencies; in particular, $\delta_{\text{min}} = 0.32\pi$. Based on Theorems 2 and 3 we anticipate the need to sample for a longer time span (larger $t_{\text{max}}$) when $\delta_{\text{min}}$ is smaller. The mode shape errors using the second set of modal frequencies are plotted in Figure 1(b). Comparing to the results from the first set, we see that a longer sampling duration $t_{\text{max}}$ is indeed needed to achieve comparable accuracy in estimating the mode shapes.

For uniform time sampling, when $T_s$ is fixed, then increasing $t_{\text{max}}$ will automatically require more samples $M$ to be acquired. For random time sampling, however, $t_{\text{max}}$ and $M$ can be chosen independently of one another. While in Figure 1 for each $t_{\text{max}}$ we have always used the same value of $M$ for random sampling as we used for uniform sampling, this is not actually necessary. In fact, our Theorem 3 suggests that, for random sampling when $\delta_{\text{min}}$ is small, we can increase $t_{\text{max}}$ without increasing the number of samples $M$. To demonstrate this, we conduct a third experiment, and for this we use the second (more closely spaced) set of modal frequencies above. For several values of $M$, we collect $M$ samples using both uniform time sampling (for which $t_{\text{max}}$ will be determined by $M$) and random time sampling. For each value of $M$ with random time sampling, however, we choose $t_{\text{max}}$ to be 2 seconds longer than the value of $t_{\text{max}}$ used for uniform sampling with the same value of $M$. The results are shown in Figure 2. We see that simply by increasing $t_{\text{max}}$ without affecting $M$, the random time sampling scheme can accommodate the decreased value of $\delta_{\text{min}}$.

To motivate our fourth experiment, consider a scenario...
where we are limited in the number of samples we can transmit to the central data repository. If our sensor is limited to collecting uniform time samples, then we may not be able to sample and transmit at a fast enough rate to avoid aliasing and accurately recover the mode shape vectors. In a scenario such as this, one way to improve the performance would be to first sample uniformly at a high rate but then to multiply the high rate sample vector by a random compressive matrix so that the amount of transmitted data is reduced. To illustrate this, we set the modal frequencies to be $\omega_1 = 10.6\pi$, $\omega_2 = 106.2\pi$, $\omega_3 = 200.8\pi$, and $\omega_4 = 360\pi$ rad/s, and we sample for a total time span of $t_{\text{max}} = 2s$. For these modal frequencies, in order to avoid aliasing, the uniform sampling interval would need to satisfy $T_s \leq \frac{\omega_{\text{max}}}{2\omega_{\text{max}}} = 0.0028s$. We first obtain a data matrix by sampling with $T_s = 0.0629s$ (a sub-Nyquist rate) over the total sampling time span $t_{\text{max}} = 2s$. This gives us in total 32 samples. We then obtain a second data matrix by sampling with $T_s = 0.002s$ (a super-Nyquist rate) and then multiplying the sample vector by a random compressive matrix $\Phi$ to produce $M' = 32$ measurements. We compute the left singular vectors for each of these data matrices to estimate the mode shapes. For the first matrix, we see errors of $\|\{\hat{\psi}_1\} - \{\psi_1\}\|_2 = 0.39832$, $\|\{\hat{\psi}_2\} - \{\psi_2\}\|_2 = 0.71303$, $\|\{\hat{\psi}_3\} - \{\psi_3\}\|_2 = 0.58411$, and $\|\{\hat{\psi}_4\} - \{\psi_4\}\|_2 = 0.03080$. For the second matrix, we see errors of $\|\{\hat{\psi}_1\} - \{\psi_1\}\|_2 = 0.05722$, $\|\{\hat{\psi}_2\} - \{\psi_2\}\|_2 = 0.08615$, $\|\{\hat{\psi}_3\} - \{\psi_3\}\|_2 = 0.03706$, and $\|\{\hat{\psi}_4\} - \{\psi_4\}\|_2 = 0.00351$. These results illustrate the utility of random matrix multiplication for reducing the dimensionality of a uniform sample vector.

As a final experiment on the synthetic data, we demonstrate a simple way to estimate the modal frequencies from the data matrix. We consider a uniform sampling scenario and set the modal frequencies as $\omega_1 = 6.24\pi$, $\omega_2 = 20.50\pi$, $\omega_3 = 30.06\pi$, and $\omega_4 = 40.22\pi$ rad/s. We sample with $T_s = 0.03s$, which is slightly faster than what Theorem 2 prescribes, and we set $t_{\text{max}} = 6.03s$. After constructing the sampled data matrix $[V]$, we compute the SVD. In order to estimate the modal frequencies we focus on the output matrix $[\hat{S}]$. Referring to \[\Phi\], we see that each row of $[\hat{S}]$ should approximately contain a complex exponential corresponding to one of the four modal frequencies. One simple way to extract each frequency is to compute the Fourier transform for each row of $[\hat{S}]$ and to identify the frequency with maximum magnitude. Figure 3 shows the magnitude plot of the FFT for each row of $[\hat{S}]$. Extracting the peak from each row, we obtain frequency estimates of $\hat{\omega}_1 = 6.3018\pi$, $\hat{\omega}_2 = 20.5638\pi$, $\hat{\omega}_3 = 30.1824\pi$, $\hat{\omega}_4 = 40.4643\pi$ rad/s. The accuracy of these estimates is naturally limited in that the frequency resolution will be inversely proportional to $t_{\text{max}}$. Finally, we note that when estimating modal frequencies, it is important that we avoid aliasing in $[S]$ in order to identify the correct modal frequencies. This experiment represents a promising proof of concept that modal frequencies can be estimated from compressive measurements; we reserve more detailed analysis of this problem for a future paper.

B. Experiments with real data

We conclude by presenting simulation results using vibration data collected from a bridge in Ypsilanti, MI. On this bridge, there are $N = 18$ wireless nodes, each of which is equipped with an accelerometer. The relative layout of the sensors is shown in Figure 4. Each sensor measures acceleration data and transmits it to the central node for analysis. Because this data contains noise and is collected from a real bridge, which does have damping and is not necessarily in free vibration mode, we feel this represents an interesting test for our methods when the assumptions of our current theoretical results are violated.

The data that is available to us from each sensor is a set of real-valued $M = 3000$ samples collected uniformly in time at a rate faster than the Nyquist rate. We stack this data into a real-valued data matrix we call $[U]$ (note that we do not assume samples of the analytic signals are available), and we test the effectiveness of multiplying $[U]$ by a random matrix $[\Phi]$ and then computing the SVD of the compressed matrix $[Y] = [U][\Phi]$. For a point of comparison, we test a method that
is similar to one presented in [10], [11]. That method, which we refer to as “CS+FDD,” involves reconstructing each length-\(M\) signal \(\{u_l\}\) from the random Gaussian measurements \(\{y_l\} = \{\Phi\}^*\{u_l\}\). These signals are reconstructed one-by-one and then fed into the Frequency Domain Decomposition (FDD) method [12] for modal analysis. The reconstruction of each \(\{u_l\}\) is accomplished by solving
\[
\min \{\alpha_l\} \quad \text{s.t.} \quad \{y_l\} = [\Phi]^*[W]\{\alpha_l\},
\]
where \([W]\) represents a DWT matrix (we saw similar results with the DFT), and we let the reconstructed \(\{u_l\} = [W]\{\alpha_l\}\).

For both the proposed SVD method and the CS+FDD method, we take \(M = 50\) measurements of each \(\{u_l\}\) using the same measurement matrix \([\Phi]\) for all \(l\). Since we do not know the true mode shapes of the structure, we use as a benchmark the three dominant mode shapes returned when FDD is applied to the original uncompressed data matrix \([U]\). The results are presented in Figure 3. As we can see, the mode shapes estimated using CS+FDD (plotted in black) are not particularly close to the mode shapes returned when FDD is applied on the original data (plotted in blue). This is apparently because it is difficult to accurately reconstruct any individual signal \(\{u_l\} \in \mathbb{R}^{3000}\) from just 50 random measurements. In contrast, however, when we apply the SVD to the compressed data matrix \([Y]\), the estimated mode shapes (plotted in red) provide much better approximations to the true FDD mode shapes.

These simulation results indeed support our theoretical results that the SVD of the data matrix \([Y]\) can return accurate estimates of the true mode shape vectors. We emphasize again that the dataset in this simulation is real-valued, contains noise, and is collected from a real structure with damping; technically, none of this is covered by the assumptions of our current theoretical results. The fact that our method was nevertheless able to successfully estimate the mode shape vectors is very encouraging and suggests that our theoretical findings may be extendable to more complicated scenarios.

**ACKNOWLEDGMENT**

The authors would like to thank Sean O’Connor and Prof. Jerome P. Lynch at the University of Michigan for helpful discussions on modal analysis and for providing us with real datasets to carry out the experiments presented in this paper.

**APPENDIX**

In this appendix we provide proofs of the main results. To do so, we take a perturbation theoretic viewpoint. We first describe how we can formulate our problem as a perturbation problem and then provide separate proofs for each theorem.

### A. Perturbation analysis

We start with the equation \([V] = [\Psi][\Gamma][S]\), and we allow the sample times \(t_1, \ldots, t_M\) to be arbitrary. To carry out perturbation analysis let us note that \([V][V]^* = [\Psi][\Gamma][S][S]^*[\Gamma]^*[\Psi]^*\), where \([S][S]^*\) is an \(N \times N\) matrix with entries
\[
([S][S]^*)_{l,n} = \begin{cases} 1, & l = n, \ v l \neq n. 
\end{cases}
\]

Thus, we can decompose this product as \([S][S]^* = [I] + [\Delta]\), where \([\Delta]\) contains the off-diagonal entries of \([S][S]^*\). Then
\[
[V][V]^* = ([\Psi][\Gamma][S][S]^*[\Gamma]^*[\Psi]^* + ([\Psi][\Gamma][\Delta][\Gamma]^*[\Psi]^* + [\Psi][\Gamma][S][S]^*[\Gamma]^*[\Psi]^*).
\]

The above expression allows us to view \([V][V]^*\) as the summation of a matrix \([H]\) and a matrix \([\delta H]\). We may view \([\delta H]\) as the perturbation matrix that is being added to \([H]\). Noting that the eigenvectors of \([H]\) are given by \([\Psi]\), our goal is to show that the eigenvectors of \([V][V]^*\) (which equal the left singular vectors of \([V]\)) are close to those of \([H]\). To do this, we follow the approach in [20] and employ [21, Theorem 1], which provides a perturbation bound on the eigenvectors as a function of the quantity
\[
\eta := \|[H]^{-\frac{1}{2}}[\delta H][H]^{-\frac{1}{2}}\|_2.
\]

In the second to last line we let \(y = ([\Gamma]^*[\Gamma]^*)^{-\frac{1}{2}}(\{x\}^*\{y\})\) and note that \(\{y\}^*\{y\} = \{x\}^*\{x\}\).

Our Theorems 2 and 3 follow by proving that \(\|[\delta H]\|_2 \leq \epsilon\). More complete details on how we can apply [21, Theorem 1] are provided in [20]. We do make two notes here concerning the application of [21, Theorem 1]. First, in order to apply this theorem, we require that \([H]\) be positive definite. This leads to the requirement that \(M \geq N\) in all of our results. Second, in order to obtain a final bound that depends only on the eigenvalues of \([H]\) and not also on those of \([V][V]^*\), it is necessary to prove that the the eigenvalues of \([V][V]^*\) approximate those of \([H]\). This fact also follows by proving that \(\|[\delta H]\|_2 \leq \epsilon\) and by applying [22, Lemma 2].

To compute an upper bound for \(\|[\delta H]\|_2\), let us note that \(\lambda_n([S][S]^*) = \lambda_n([I] + [\Delta]) = 1 + \lambda_n([\Delta])\), where we use \(\lambda_n([A])\) to denote the \(n\)th largest eigenvalue of \([A]\). If we can find upper and lower bounds on the eigenvalues of \([S][S]^*\) such that \(\lambda_n([S][S]^*) \leq \lambda_n([I] + [\Delta]) \leq \lambda_n([\Delta])\), then we can see that \(\lambda_n([I] + [\Delta]) \leq \lambda_n([\Delta]) = \lambda_n([\Delta])\), which in turn implies that \(\|[\delta H]\|_2 = \max_n |\lambda_n([\Delta])| \leq \max(|\lambda_n - 1|, |\lambda_n - 1|)\).

Therefore, our strategy is to bound \(\lambda_n([S][S]^*)\) from below and above in order to bound \(\|[\delta H]\|_2\). In the following sections
we establish this result for both the random and uniform sampling cases.

B. Proof of Theorem 3 (random sampling)

Let us first consider how to establish a bound on \( \|\Delta\|_2 \) if we were to sample \( t_1, \ldots, t_M \) uniformly at random in the time interval \([0, t_{\text{max}}]\). We can establish the following bound on the eigenvalues of \( [S][S]^* \).

**Theorem 6.** Given that we sample with \( t_{\text{max}} \) satisfying \([13]\) and \( M \) satisfying \([14]\), then with probability at least \( 1 - \tau \) we will have \( 1 - \epsilon < \lambda_\alpha([S][S]^*) < 1 + \epsilon \) for all \( n \).

**Proof:** To bound \( \| [S][S]^* \|_2 \) we use a slightly modified version of a theorem that appeared in \([18]\).

**Theorem 7.** ([18], Theorem 5.1) Consider a sequence \( \{Z_m\} \) of independent, d-dimensional, random, self-adjoint matrices that satisfy \( \|Z_m\|_2 \geq 0 \) and \( \lambda_{\text{max}}(Z_m) \leq c \) almost surely. Then for any \( \tilde{\mu}_\text{min} \) and \( \tilde{\mu}_\text{max} \) such that

\[
\tilde{\mu}_\text{min} \leq \mu_{\text{min}} \left( \frac{1}{M} \sum_{m=1}^{M} E[Z_m] \right) \quad \text{and} \quad \tilde{\mu}_\text{max} \geq \mu_{\text{max}} \left( \frac{1}{M} \sum_{m=1}^{M} E[Z_m] \right),
\]

we have

\[
P \left( \mu_{\text{min}} \left( \frac{1}{M} \sum_{m=1}^{M} Z_m \right) \leq \alpha \right) \leq d e^{-MD(\alpha/c)\|\mu_{\text{min}}/c|}
\]

for \( 0 \leq \alpha \leq \tilde{\mu}_\text{min} \) and

\[
P \left( \mu_{\text{max}} \left( \frac{1}{M} \sum_{m=1}^{M} Z_m \right) \geq \alpha \right) \leq d e^{-MD(\alpha/c)\|\mu_{\text{max}}/c|}
\]

for \( \tilde{\mu}_\text{max} \leq \alpha \leq c \).

In order to apply Theorem 7, let us write

\[
[S][S]^* = \frac{1}{M} \sum_{m=1}^{M} \left\{ e^{i\omega_{1} t_m}, \ldots, e^{i\omega_{N} t_m} \right\} \left\{ e^{-i\omega_{1} t_m}, \ldots, e^{-i\omega_{N} t_m} \right\}.
\]

We define the vector \( \{S_m\} = \{e^{-i\omega_{1} t_m}, \ldots, e^{-i\omega_{N} t_m}\}^* \), where \( \|\{S_m\}\|_2^2 = N \). Let us set the matrix \( [Z_m] \) that appears in the above theorem as \( [Z_m] = \{S_m\}[S_m]^* \). As a result, \( [Z_m] \) will be i.i.d. positive semi-definite matrices, i.e., \( [Z_m] \geq 0 \), of rank 1 with \( \lambda_{\text{max}}([Z_m]) = \|\{S_m\}\|_2^2 = N \). We wish to compute

\[
\lambda_{\text{min}} \left( \frac{1}{M} \sum_{m=1}^{M} E\{S_m\}[S_m]^* \right) = \lambda_{\text{min}} (E\{S_m\}[S_m]^*),
\]

\[
\lambda_{\text{max}} \left( \frac{1}{M} \sum_{m=1}^{M} E\{S_m\}[S_m]^* \right) = \lambda_{\text{max}} (E\{S_m\}[S_m]^*),
\]

or the appropriate lower and upper bound on the above quantities. Note that

\[
\{E\{S_m\}[S_m]^*\}_{l,n} = \begin{cases} 1, & l = n, \\
E e^{i(\omega_l - \omega_n) t_m}, & l \neq n \end{cases}
\]

The eigenvalues can also be written as \( \lambda_n (E\{S_m\}[S_m]^*) = \lambda_n (I + [\Delta_S]) = 1 + \lambda_n ([\Delta_S]) \), where \([\Delta_S] \) is the off-diagonal matrix of \( E\{S_m\}[S_m]^* \). According to Gershgorin’s circle theorem \([23]\) we know that every eigenvalue of \([\Delta_S] \) must lie within at least one Gershgorin disk. As \([\Delta_S] \) has zero diagonal entries, every Gershgorin disk must be centered at zero. Thus, the radius of the largest disk will provide a bound on all eigenvalues of \([\Delta_S] \). It follows that every eigenvalue of \([\Delta_S] \) will obey the following bound:

\[
|\lambda([\Delta_S])| \leq \max_l \sum_{n=1, n \neq l}^{N} \left| \frac{1}{2} \sin((\omega_l - \omega_n) t_{\text{max}}) \right|
\]

\[
\leq \max_l \sum_{n=1, n \neq l}^{N} \frac{2}{|\omega_l - \omega_n| t_{\text{max}}}
\]

\[
\leq \sum_{n=1, n \neq l}^{N} \frac{2}{|\omega_l - \omega_n| t_{\text{max}}}
\]

\[
\leq \frac{4}{\delta_{\text{min}} t_{\text{max}}} \sum_{n=1}^{N} \frac{1}{n},
\]

where \( \delta_{\text{min}} \) is the minimum non-zero singular value of \( [S][S]^* \).
where we have denoted the index of the middle row of $[\Delta_S]$ as $l'$ (when $N$ is even we can take either of $N/2$ or $N/2 + 1$ as the middle row), and we have used the fact that $\omega_n - \omega_{n+1} \geq \delta_{\text{min}}$ for $n = 1, \ldots, N - 1$. The summation term in the above bound is also known as the Harmonic number. We can simplify the above expression by using the following bound on the Harmonic number.

**Theorem 9.** Suppose we sample with a total time span $T_s = \frac{\pi}{\delta_{\text{max}}}$ and with sampling interval $T_s = \frac{\pi}{\max}$. Then we establish the following bound on the eigenvalues of $[\Delta_S]^*$: $1 - \epsilon \leq \lambda_n([\Delta_S]^*) \leq 1 + \epsilon$.

**Proof:** The off-diagonal matrix of $[S][S]^*$, denoted as $[\Delta]$, has the following entries: $[\Delta]_{l,n} = 0$ when $l = n$, and

$$[\Delta]_{l,n} = \frac{1}{M} \sum_{m=0}^{M-1} e^{i(\omega_l - \omega_m)MT_s} = \frac{1 - e^{i(\omega_l - \omega_m)MT_s}}{M - 1 - e^{i(\omega_l - \omega_m)T_s}}$$

$$= e^{i(\omega_l - \omega_m)T_s(M-1)/2} \sin((\omega_l - \omega_m)MT_s)$$

$$= e^{i(\omega_l - \omega_m)T_s(M-1)/2} \sin|\omega_l - \omega_m|MT_s/2$$

when $l \neq n$. The fraction of sinusoids in the above equation is known as the periodic sinc function or the Dirichlet function and is defined as $\text{psinc}(x) = \frac{\sin(Mx)}{M\sin(x)}$. More specifically,

$$\text{psinc}(x) = \begin{cases} \sin\left(\frac{Mx}{2}\right)/M\sin\left(\frac{x}{2}\right), & x \neq 2\pi k, \ k = 0, \pm 1, \pm 2, \ldots \end{cases}$$

As its name implies, the psinc function is a periodic function where the period is equal to $2\pi$ when $M$ is odd, and $4\pi$ when $M$ is even. Every time $x$ is equal to an integer multiple of $2\pi$, the psinc function will reach its maximum value.

Again, we bound the eigenvalues of $[\Delta]$ using Gershgorin’s disk theorem. Since every Gershgorin disk will again be centered at zero, every eigenvalue of $[\Delta]$ must obey the following bound:

$$|\lambda([\Delta])| \leq \max_{l \neq n} \sum_{n=1}^{N} |\text{psinc}(\omega_l - \omega_n)|T_s|T_s|$$.  

(19)

We can guarantee the evaluation of the psinc function to be small by restricting ourselves to only certain values of $T_s$ and $M$. We compute an upper bound on the psinc function by noting that the denominator can be lower bounded by a linear function for a certain range of $T_s$. In particular, for $T_s \leq \frac{\pi}{|\omega_l - \omega_n|}$ we have that $\text{psinc}(\omega_l - \omega_n|T_s \geq |\omega_l - \omega_n|^2\pi$. Applying this to (19), we have

$$|\lambda([\Delta])| \leq \max_{l \neq n} \sum_{n=1}^{N} \frac{|\pi \sin(\omega_l - \omega_n)|T_s}{|\omega_l - \omega_n|^2}$$

$$= \max_{l \neq n} \sum_{n=1}^{N} \frac{\pi}{2} \frac{\sin(\omega_l - \omega_n)^2MT_s}{|\omega_l - \omega_n|MT_s}$$

$$\leq \max_{l \neq n} \sum_{n=1}^{N} \frac{\pi}{|\omega_l - \omega_n|MT_s}$$

when $T_s \leq \frac{\pi}{|\omega_l - \omega_n|}$ for all $l$ and $n$. In the last line we have used the fact that $|\sin(x)| \leq 1/|x|$. To ensure that $T_s \leq \frac{\pi}{|\omega_l - \omega_n|}$ for all $l$ and $n$, we choose the sampling interval such that $\frac{T_s}{\delta_{\text{max}}} \leq \frac{\pi}{\delta_{\text{max}}}$. Following the same arguments we used in the proof of Theorem 3 this will give us

$$|\lambda([\Delta])| \leq \frac{\pi}{\delta_{\text{min}}MT_s} \sum_{n=1}^{N} \frac{2}{n} \frac{2(\log(N/2) + 1.01)}{\delta_{\text{min}}MT_s}$$

$$< 2\pi(\log(N/2) + 1.01)$$

$$< \frac{2\pi(\log(N/2) + 1.01)}{\delta_{\text{min}}(M-1)T_s}.$$
Now, note that \((M - 1)T_s = t_{\max}\) and if we choose \(t_{\max}\) so that
\[
t_{\max} \geq \frac{2\pi(\log(\lceil N/2 \rceil) + 1.01)}{\delta_{\min} \epsilon},
\]
we will have that \(|\lambda(\Delta)| < \epsilon\). In summary, when we sample within a sampling interval satisfying \(T_s \leq \frac{\pi}{\delta_{\max}}\) and a time span satisfying \(t_{\max} \geq \frac{2\pi(\log(\lceil N/2 \rceil) + 1.01)}{\delta_{\min} \epsilon}\), we will have
\[
1 - \epsilon < \lambda_n([S]|S|^*) = 1 + \lambda_n(\Delta) < 1 + \epsilon.
\]
Or, in other words, if we set \(T_s = \frac{\pi}{\delta_{\max}}\), and remembering that \(t_{\max} = (M - 1)T_s\), this means that we need to sample at least
\[
M \geq \frac{2(\log(\lceil N/2 \rceil) + 1.01)}{\epsilon} \delta_{\max} + 1,
\]
to achieve the above eigenvalue guarantee. Note that this is the smallest number of measurements we need since we set \(T_s\) as large as possible. If we were to reduce the sampling interval we would need to take more measurements to cover the same time span \(t_{\max}\).

**D. Proof of Theorem 5** (uniform sampling with matrix multiplication)

This result can be shown by simply using the triangle inequality. Let us write
\[
||\{\tilde{\psi}_n\} - \{\tilde{\psi}_n\}||_2 \leq ||\{\tilde{\psi}_n\} - \{\hat{\psi}_n\}||_2 + ||\{\hat{\psi}_n\} - \{\tilde{\psi}_n\}||_2,
\]
where \(\{\hat{\psi}_1\}, \ldots, \{\hat{\psi}_N\}\) denote the left singular vectors of \([Y]\). Note that the first term on the right hand side of the above inequality represents the error in the mode shape vectors due to the uniform sampling scheme as presented in Theorem 2. The second term represents the difference between the left singular vectors of \([V]\) and those of \([Y]\). To quantify the amount of this error we make use of [20, Theorem 1]. Substituting the upper bound for each term completes the proof.

**References**

[1] “Commercial buildings energy consumption survey (CBECs) – overview of commercial building characteristics.” [Online]. Available: http://www.eia.gov/consumption/commercial/data/2003/pdf/a5.pdf

[2] “American FactFinder.” [Online]. Available: http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_11_5YR_DP04

[3] “2013 report card on America’s infrastructure.” [Online]. Available: http://www.infrastructurereportcard.org/fact-sheet/bridges

[4] H. Sohn, C. R. Farrar, F. Hemez, and J. Czarnecki, “A review of structural health monitoring literature 1996–2001,” Los Alamos National Lab., NM, Tech. Rep., 2001.

[5] S. W. Doebling, C. R. Farrar, M. B. Prime, and D. W. Shevitz, “Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review,” Los Alamos National Lab., NM, Tech. Rep., 1996.

[6] J. P. Lynch, “An overview of wireless structural health monitoring for civil structures,” Philosophical Trans. of the Royal Society A, vol. 365, no. 1851, pp. 345–372, 2007.

[7] D. Donoho, “Compressed sensing,” IEEE Trans. Inf. Theory, vol. 52, no. 4, Apr. 2006.

[8] E. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[9] M. Wakin, S. Becker, E. Nakamura, M. Grant, E. Sovero, D. Ching, J. Yoo, J. Romberg, A. Emami-Neyestanak, and E. Candès, “A non-uniform sampler for wideband spectrally-sparse environments,” IEEE J. Emerging Sel. Topics Circuits Syst., vol. 2, no. 3, pp. 516–529, Sep. 2012.

[10] D. Mascarenas, D. Hash, J. Theiler, and C. Farrar, “The application of compressed sensing to detecting damage in structures,” in Proc. 8th Int. Workshop on Structural Health Monitoring, 2011.

[11] B. Yuequan, J. L. Beck, and L. Hui, “Compressive sampling for accelerometer signals in structural health monitoring,” Structural Health Monitoring, vol. 10, no. 3, pp. 235–246, 2011.

[12] R. Brincker, L. Zhang, and P. Andersen, “Modal identification from ambient responses using frequency domain decomposition,” Proc. of the 18th Int. Modal Anal. Conf., 2000.

[13] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” Appl. and Comput. Harm. Anal., vol. 26, no. 3, pp. 301–321, 2009.

[14] B. Boushah, Time frequency signal analysis and processing: A comprehensive reference. Elsevier, 2003.

[15] F. W. King, Hilbert Transforms. Cambridge University Press, 2009.

[16] S. R. Ibrahim and E. C. Mikułicz, “A method for the direct identification of vibration parameters from the free response,” Shock and Vibration Bulletin, 1977.

[17] J. N. Juang and R. S. Pappa, “An eigensystem realization algorithm for modal parameter identification and model reduction,” Journal of Guidance, Control, and Dynamics, vol. 8, no. 5, pp. 620–627, 1985.

[18] J. A. Tropp, “User-friendly tail bounds for sums of random matrices,” Found. Comput. Math., vol. 12, no. 4, pp. 389–434, 2012.

[19] M. Davenport, “Concentration of measure and sub-gaussian distributions,” 2009, available online at http://cnx.org/content/m32583/latest/

[20] A. C. Gilbert, J. Y. Park, and M. B. Wakin, “Sketched SVD: Recovering spectral features from compressive measurements,” Arxiv preprint arXiv:1211.0361, 2012.

[21] R. Mathias and K. Veselić, “A relative perturbation bound for positive definite matrices,” Linear Algebra and its Applications, vol. 270, pp. 315–321, 1998.

[22] J. Barlow and J. Demmel, “Computing accurate eigensystems of scaled diagonally dominant matrices,” Tech. Rep., 1980.

[23] S. Gershgorin, “Über die abgrenzung der eigenwerte einer matrix,” Bulletin de l’Academie des Sciences de l’URSS. Classe des sciences mathematiques et na, pp. 749–754, 1931.

[24] M. B. Villarino, “Sharp Bounds for the Harmonic Numbers,” Arxiv preprint arXiv:math/0510585, 2005.