# Modal Identification of Bridges Using Mobile Sensors with Sparse Vibration Data

Soheil Sadeghi Eshkevari, S.M.ASCE[1]; Shamim N. Pakzad, A.M.ASCE[2]; Martin Takáč[3]; and Thomas J. Matarazzo, A.M.ASCE[4]

**Abstract:** Dynamic sensor networks have the potential to significantly increase the speed and scale of infrastructure monitoring. Structural health monitoring (SHM) methods have been long developed under the premise of utilizing fixed sensor networks for data acquisition. Over the past decade, applications of mobile sensor networks have emerged for bridge health monitoring. Yet, when it comes to modal identification, there remain gaps in knowledge that have ultimately prevented implementations on large structural systems. This paper presents a structural modal identification methodology based on sensors in a network of moving vehicles: a large-scale data collection mechanism that is already in place. Vehicular sensor networks scan the bridge's vibrations in space and time to build a sparse representation of the full response, i.e., an incomplete data matrix with a low rank. This paper introduces modal identification using matrix completion (MIMC) methods to extract dynamic properties (frequencies, damping, and mode shapes) from data collected by a large number of mobile sensors. A dense matrix is first constructed from sparse observations using alternating least-square (ALS) then decomposed for structural modal identification. This paper shows that the completed data matrix is the product of a spatial matrix and a temporal matrix from which modal properties can be extracted via methods such as principal component analysis (PCA). Alternatively, an impulse-response structure can be embedded into the temporal matrix and then natural frequencies and damping ratios are determined using Newton's method with an inverse Hessian approximation. For the case of ambient vibrations, the natural excitation technique (NExT) is applied and then structured optimization (Newton's method) is performed. Both approaches are evaluated numerically, and results are compared in terms of data sparsity, modal property accuracy, and postprocessing complexity. Results show that both techniques extract accurate modal properties, including high-resolution mode shapes from sparse dynamic sensor network data; they are the first to provide a complete modal identification using data from a large-scale dynamic sensor network. **DOI: 10.1061/(ASCE)EM.1943-7889.0001733.** 



## Introduction

Dynamic sensor networks are an element of modern society. Humans are increasingly relying on ubiquitous smartphones, internet-of-things devices, and data-driven services in their daily lives. This phenomenon has given life to massive amounts of data, which have driven an entire field of studies (Barabasi 2005; Wang et al. 2010, 2012; Alexander et al. 2015; Tachet et al. 2017) on human activity in the urban environment and the development of smart city applications (Tachet et al. 2016; Anjomshoaa et al. 2018; Vazifeh et al. 2018). Sensors carried by humans create an inexpensive, large-scale mobile sensor network—and as smart and self-driving cars continue to emerge, vehicles will become a growing source for sensory data on the built environment (Gurney et al. 2015; Massaro et al. 2017). Throughout hundreds of millions of trajectories each day, humans capture data on bridges and other infrastructure routinely and comprehensively.

This has sparked a large interest in the use of mobile sensors for structural health monitoring (SHM), a field that has exclusively relied on data sets collected by networks of fixed sensors (Sony et al. 2019; Pakzad et al. 2008; Lynch and Loh 2006; Kurata et al. 2011; Matarazzo and Pakzad 2016b). Mobile sensor networks have low setup costs and address major shortcomings of fixed sensor networks. Mobile sensor data contain a denser spatial resolution when compared to data collected by the same number of distributed fixed sensors. Overall, mobile sensor networks are scalable and produce significantly more spatial information per sensor. The development of SHM approaches that are able to incorporate crowdsensed mobile sensor data (Matarazzo and Pakzad 2018; Matarazzo et al. 2018; Mei et al. 2019; Mei and Gül 2018) accelerates the rate at which engineers acquire knowledge on the true conditions of infrastructure; the corresponding boom in information would prove invaluable to bridge monitoring and management (Kleywegt and Sinha 1994; Smith 2016). Early work on mobile sensors in SHM demonstrated that a sensor within a moving vehicle can be used to detect the fundamental frequency of a bridge (Yang et al. 2004; Lin and Yang 2005; Yang and Chang 2009). Subsequent studies built on the theory and application of vehicle-bridge interaction in order to estimate damping using various *drive-by* setups (González et al. 2012; Keenahan et al. 2014; McGetrick et al. 2015) or stiffness information (Li et al. 2014) whose variations may be indicative of structural damage. Very recent studies have theoretically and experimentally developed methodologies for damage detection using the aggregation of passing-by vehicles on bridges (Mei et al. 2019; Mei and Gül 2018).

[1]Ph.D. Candidate, Dept. of Civil and Environmental Engineering, Lehigh Univ., Bethlehem, PA 18015 (corresponding author). ORCID: https://orcid.org/0000-0001-9285-6911. Email: soheil.sadeghi69@gmail.com

[2]Associate Professor, Dept. of Civil and Environmental Engineering, Lehigh Univ., Bethlehem, PA 18015.

[3]Assistant Professor, Dept. of Industrial and Systems Engineering, Lehigh Univ., Bethlehem, PA 18015.

[4]Postdoctoral Researcher, Senseable City Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139; Researcher, Cornell Tech, Cornell Univ., New York, NY 10044. ORCID: https://orcid.org/0000-0001-8978-1357

Note. This manuscript was submitted on March 4, 2019; approved on August 27, 2019; published online on January 28, 2020. Discussion period open until June 28, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Engineering Mechanics***, © ASCE, ISSN 0733-9399.

Modern system identification (SID) algorithms for civil structures are reliable, repeatable, and often have a formidable mathematical foundation, e.g., frequency domain decomposition (FDD) (Brincker et al. 2001), eigensystem realization analysis (ERA) (Juang and Pappa 1985; James et al. 1995), stochastic system identification (SSI) (Peeters and De Roeck 2001), fast Bayesian FFT method (Au 2011), distributed modal identification (Pakzad et al. 2011), Kalman filter-based SID (Chang and Pakzad 2014), and stochastic iterative SID (Dorvash and Pakzad 2013). As with SHM, classical SID methods were formulated under the expectation of fixed sensor network data and are incompatible with mobile sensor data. Recently, there has been progress in the development of SID methods that are designed for mobile sensor data. Matarazzo and Pakzad (2016a) proposed an updated version of the structural modal identification using expectation maximization (STRIDE) algorithm to accept incomplete data sets and mobile sensor network data. A method proposed by Marulanda et al. (2017) enabled the identification of spatially dense mode shapes using a hybrid sensor network: one mobile and one fixed. Matarazzo and Pakzad (2016b) derived three state-space models that expect data from dynamic sensor networks and recommended the truncated physical model based on its configurable model complexity and ability to distinguish between sensing nodes and model degree of freedom (DOF). Matarazzo and Pakzad (2018) introduced a new version of the STRIDE algorithm called STRIDEX, which is able to identify the parameters of the truncated physical state-space model and, therefore, produce comprehensive estimates of structural modal properties using data from a mobile sensor network. In an initial experimental application, STRIDEX produced an accurate mode shape with 248 points using data from two mobile sensors. Subsequent experimental applications demonstrated accurate identification of higher modes, complete with dense mode shapes, based exclusively on mobile sensor data (Matarazzo et al. 2018).

This paper focuses on bridge SID based on vehicular sensor networks, an especially widespread dynamic sensor network with increasing sensing capabilities. In everyday traffic, connected vehicles with embedded sensors scan the bridge's vibrations in space and time. The vehicular sensor network data is a sparse representation of the full response—more specifically, an incomplete data matrix with a low rank. The sparsity of the data matrix depends on attributes of the vehicular sensor network, such as the number of mobile sensors, position time-series, sampling rates, etc. This paper presents modal identification using matrix completion (MIMC) methods for SID of bridges based on data collected by a large number of mobile sensors.

## Background Theory and Scope

### Mathematical Model for Dynamic Sensor Networks

Dynamic sensor network data are produced when some aspect of the sensor network varies during measurement. For instance, mobile sensor data are classified as dynamic sensor network data because the positions of the sensors change throughout discrete-time sampling. Dynamic sensor networks were introduced in Matarazzo and Pakzad (2016b) along with state-space models that were designed to analyze the resulting data sets for structural dynamics applications; the stochastic truncated physical state-space model is the primary model for the case of ambient vibrations [for complete mathematical details, see Matarazzo and Pakzad (2016b, 2018)]. The model is given in Eq. (1)

$$\begin{aligned} x_k &= \boldsymbol{A}\boldsymbol{x}_{\boldsymbol{k-1}} + \boldsymbol{\eta}_{\boldsymbol{k}} \\ y_k &= \boldsymbol{G}_{\boldsymbol{k}}\boldsymbol{C}\boldsymbol{x}_{\boldsymbol{k}} + \boldsymbol{\nu}_{\boldsymbol{k}} \end{aligned} \tag{1}$$

where $\boldsymbol{x_1} \sim N(\hat{\mu}, \hat{V})$, $\boldsymbol{\eta_k} \sim N(0, Q)$, and $\boldsymbol{\nu_k} \sim N(0, R)$; and $\boldsymbol{A} \in \mathbb{R}^{pN_\alpha \times pN_\alpha}$ is the state matrix, $\boldsymbol{C} \in \mathbb{R}^{N_0 \times pN_\alpha}$ is the observation matrix, $\boldsymbol{x_k} \in \mathbb{R}^{pN_\alpha}$ and $\boldsymbol{y_k} \in \mathbb{R}^{N_0}$ are the state and observation vectors, respectively; $\boldsymbol{\eta_k}$ and $\boldsymbol{\nu_k}$ are the input excitation and measurement noises, which are both modeled as Gaussian white noise with covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively; $N_\alpha$ is the number of virtual probing locations; $N_0$ is the size of the observation vectors; $p$ is the model order; and $\boldsymbol{G_k}$ is the mode shape regression (MSR) function [in this study, $\boldsymbol{G_k}$ is identical to $\boldsymbol{\Omega_k}$ in Matarazzo and Pakzad (2016b)], which maps structural responses at the locations of the mobile sensors to responses at virtual probing locations (VPLs). In one definition, VPLs are the structural DOF of interest and are the points at which mode shapes are computed. In theory, the MSR function is a product of two matrices, which are based on the exact structural mode shapes. Matarazzo and Pakzad (2016b) showed that the MSR function could be approximated accurately using sinc basis functions with $\hat{\boldsymbol{G}}_{\boldsymbol{k}}$, as shown in Eq. (2)

$$\hat{\boldsymbol{G}}_{\boldsymbol{k}} = \left[\operatorname{sinc}\left(\frac{1}{\Delta s^\alpha}(\boldsymbol{s}_k^O - s_1^\alpha)\right), \operatorname{sinc}\left(\frac{1}{\Delta s^\alpha}(\boldsymbol{s}_k^O - s_2^\alpha)\right), \ldots, \operatorname{sinc}\left(\frac{1}{\Delta s^\alpha}(\boldsymbol{s}_k^O - s_N^\alpha)\right)\right] \tag{2}$$

In this equation $\boldsymbol{s}_k^O$ is a vector of sensors' positions at time step $k$; $s_i^\alpha$ is the location of the $i$th VPL; and $\Delta s^\alpha$ is the uniform distance between the VPLs. In spatial signal reconstruction, the role of $\operatorname{sinc}(x) = \sin(x)/x$ is to map measurements from an observed location to a different unobserved location based on the ideal low-pass filter (Gensun 1996; Schanze 1995; Moheimani et al. 2003).

In this mathematical framework, the STRIDEX method has been successful in both numerical and experimental applications where the VPLs were set to be uniformly spaced [as recommended by Matarazzo and Pakzad (2016b)]. Yet, the applications to date have considered a relatively small number of mobile sensors (less than ten). In addition, signal reconstruction literature shows the errors associated with nonuniform sampling (Jerri 1977; Maymon and Oppenheim 2011); yet, further research is needed to quantify how the spacing of the VPLs, or the mobile sensors, influences the accuracy of the sinc MSR approximations used in STRIDEX, which ultimately impact modal property estimates. When relying on vehicular sensor networks, it is prudent to consider the following: (1) a very large number of sensors collecting data simultaneously; (2) a variable number of participating sensors during data collection; and (3) vehicle-bridge interaction and pavement roughness effects. The approach described in this paper addresses the first two considerations (throughout the paper, the dynamic effect of the vehicles are ignored, and road surface is assumed to be smooth); the preceding topic (3) is studied in Eshkevari and Pakzad (2019) and Eshkevari et al. (2020). Despite the significance of topic (3) in the overall formulation of the problem, it is outside of the scope of the current study. The effect of vehicle

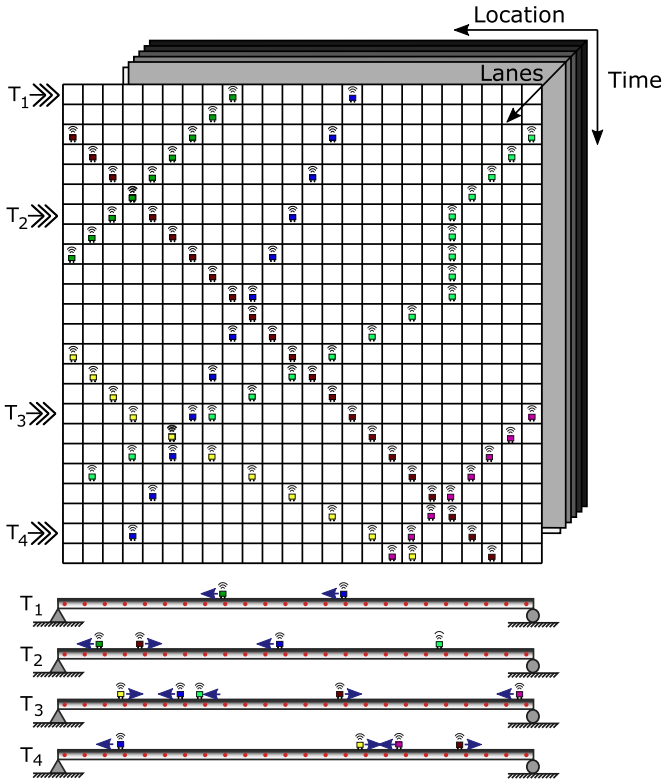


**Fig. 1.** Illustration of a sparse matrix produced by a large vehicular sensor network. The full bridge response is represented by a three-dimensional data matrix: two spatial dimensions and one temporal dimension. A two-dimensional (2D) matrix shows the response along the bridge length for a given traffic lane. The methods in this paper consider the problem of completing such a sparse 2D data matrix and extracting bridge modal properties. The instantaneous vehicle configurations corresponding to time samples $T_1$, $T_2$, $T_3$, and $T_4$ are depicted.

dynamics and roughness-caused vibrations has been broadly studied in recent years (Lin and Yang 2005; Malekjafarian et al. 2015; Malekjafarian and OBrien 2014).

### Dynamic Sensor Networks with Sparse Data

The structural responses recorded by a vehicular sensor network are both sparse and dynamic. At each instant, the vast majority of the bridge is not sampled in space (sparse), and the sensing configuration is time variant (dynamic). Consider a large response matrix with temporal and spatial dimensions. Because the data collected by a moving sensor is a function of both time and space, the entries build in two directions simultaneously, i.e., not just in a single row or column. In the case of a constant speed, the mobile sensor data is a diagonal band within the response matrix.

Fig. 1 shows an illustration of a generic sparse data matrix produced by a large vehicular sensor network. The response matrix is three-dimensional (3D) with two spatial dimensions on the bridge (longitudinal and transverse directions) and one temporal dimension (time). A two-dimensional (2D) slice of this matrix removes the dimension along the cross section (transverse) of the bridge and shows the data collected for a given traffic lane. There are two key observations: (1) this data matrix is mostly empty; and (2) while there are some visible patterns among the available entries in the data matrix, ultimately there is an unknown stochastic structure that governs. The methods presented in this paper are evaluated based on this 2D slice of the full 3D response matrix, both of which are sparse tensors.

The general approach is to *complete* the unobserved parts of this matrix given the sparse entries, i.e., the matrix completion problem. For example, in Eshkevari and Pakzad (2020) and Eshkevari et al. (2019), the alternating least-squares (ALS) method (Jain et al. 2013; Zachariah et al. 2012) was implemented to complete a sparse structural response matrix, which resulted in responses at all considered DOFs. Using ALS, the sparse matrix is estimated as the product of two matrices with rank $K$, which is much smaller than the dimensions of the original matrix. Further details of this method are discussed later in this paper.

### Contributions

This paper presents a set of modal identification using matrix completion methods, which enables a comprehensive bridge SID based solely on data collected by a large-scale vehicular sensor network. As mentioned previously, vehicular sensor network data are unique for SHM because they are both sparse and dynamic. Two approaches, the principal component analysis (PCA) and structured optimization analysis (SOA), are proposed to transform a completed response matrix into structural modal properties. The MIMC methods have the following novelties:

- The MIMC methods are the first SID algorithms designed to process data collected by a large network of mobile sensors and are immediately applicable to vehicular sensor networks.
- MIMC methods are able to process unstructured mobile sensor data, which include records from vehicles that have random and independent trajectories, e.g., different speeds, sampling rates, etc.
- MIMC methods are able to estimate very high-resolution mode shapes after one computational run (no iterations are necessary).
- Within the MIMC methods, a new optimization technique for the structured matrix completion problem that uses an approximate Newton's method is proposed and validated.
- MIMC methods adapt to the availability of the observed data. The procedure is successful in identifying structural modes even when the original data matrix is very sparse, e.g., 0.5% completeness.

In the next section, the matrix completion problem is presented with respect to incomplete structural response data. Then three MIMC methods are proposed and described in detail. In Simulation and Results, three mobile sensing simulations based on finite-element models are presented to evaluate the performance of the MIMC methods. In the Discussion, the simulation results are discussed. An overview of the methods and primary remarks are presented in the Conclusion.

## Completion of a Sparse Structural Response Matrix

In this matrix completion problem, the full structural response matrix $\boldsymbol{Y}$ is represented as the product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ [Eq. (3)]. Within alternating least squares (ALS), the problem reduces to determining the optimal values of decomposition matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (Jain et al. 2013; Zachariah et al. 2012). Note that $\boldsymbol{Y}$ is the true complete response matrix, which is not available; only a subset of it is observed ($\boldsymbol{Y}_{obs}$) during data collection. As a result, the mathematical expression of the objective function is described in Eq. (5)

$$\boldsymbol{Y} = \boldsymbol{AB} \tag{3}$$

$$\boldsymbol{Y}_{obs} = \boldsymbol{\Phi}(\boldsymbol{Y}) \tag{4}$$

$$\min_{\boldsymbol{A},\boldsymbol{B}} \frac{1}{2}\|\boldsymbol{Y}_{\text{obs}} - \Phi(\boldsymbol{AB})\|^2 + \frac{\lambda}{2}(\|\boldsymbol{A}\|^2 + \|\boldsymbol{B}\|^2) \tag{5}$$

where $\boldsymbol{Y}_{obs}$ is a sparse matrix, which is a subset of the original time-space matrix $\boldsymbol{Y}$. The matrix $\boldsymbol{\Phi}$ is a mapping function that selects

observed entries from the full matrix $\boldsymbol{Y}$. In the ALS method, the objective function shown in Eq. (5) is minimized using alternating gradient descent steps on matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (the goal is to find matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ that minimize the objective function). Conventionally, a regularization term is added to the objective function to prohibit overfitting.

So far, the algorithm provides a solution to complete the response matrix using sparse observations. The next step is to process this matrix in order to extract modal properties of the bridge. The first approach could be to apply ERA (Juang and Pappa 1985) for modal identification. Despite its broad application, ERA is computationally impractical when the number of output channels is large (Kramer and Gugercin 2016; Krishnan et al. 2011; Kramer and Gorodetsky 2018). In addition, no distributed implementation of ERA-based algorithms is known to the authors. To address this concern, and because the focus of this paper is not on the system identification methods, alternative algorithms have to be proposed that are computationally affordable and can adequately demonstrate the efficiency and accuracy of this approach.

ALS requires that the original matrix (in this study, $\boldsymbol{Y}$) has a low-rank. From structural dynamics, it is known that in a multi degree of freedom (MDOF) system, the response at each location can be represented in modal coordinates as shown in Eq. (6). In this equation, $\boldsymbol{\Omega}$ is a matrix of stacked mode shapes, and $\boldsymbol{Q}$ is a matrix of stacked single degree of freedom (SDOF) responses of modal coordinates

$$Y = \boldsymbol{\Omega Q} \approx \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{Q}} \tag{6}$$

The dimensionality of this equation can be reduced through modal truncation. For many structures, the dynamic responses can be estimated accurately using only the most significant modes, e.g., the first $K$ modes. The truncated matrices of the mode shapes and modal coordinates are presented as $\tilde{\boldsymbol{\Omega}}$ and $\tilde{\boldsymbol{Q}}$ with the ranks equal to $K$ (model order). From Eqs. (3) and (6), $\boldsymbol{AB} \approx \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{Q}}$. This suggests that matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are the transformed versions of the matrix of mode shapes ($\tilde{\boldsymbol{\Omega}}$) and modal coordinates ($\tilde{\boldsymbol{Q}}$), respectively.

From modal analysis, it is known that the modal coordinates are orthogonal with respect to the mass and stiffness matrices, while there is no orthogonality condition when matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are optimized using ALS (unconditional optimization). The normalized mass matrix of a large and homogeneous bridge converges to an identity matrix. This fact simplifies the mass orthogonality condition to a simple orthogonality condition. Therefore, to mass-orthogonalize components of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, a simple approach is to transform them using the principal component analysis (PCA). PCA extract orthogonal principal components from a set of data (note that it is nearly equivalent to apply PCA over the reconstructed matrix with significantly more computational cost). It is expected that after orthogonalization of $\boldsymbol{A}$ and $\boldsymbol{B}$, these new matrices are better estimations of the actual modal components. However, it is known that the mass and stiffness orthogonality between two modes does not guarantee that these modes are natural mode shapes. For instance, Ritz modes (Wilson et al. 1982) are orthogonal but not natural (Ritz modes may contain an exclusive subset of natural modes). To address this problem, an alternative approach is proposed to estimate modal components from matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in which the structure of matrix $\boldsymbol{B}$ is prefixed.

### Impulse Response Analysis using Structured Optimization

The impulse response of each mode in a MDOF system has a certain structure. In fact, in this response, $\tilde{\boldsymbol{Q}}$ contains the free vibration responses of the modal coordinates as presented in Eq. (7) for undamped and Eq. (8) for damped problems

$$\begin{aligned}\boldsymbol{Y}(x,t) &= \sum_{n=1}^{K} \boldsymbol{\Omega}_n(x)\sin(\omega_n t + \psi_n)\\ &= \boldsymbol{\Omega}_1(x)\sin(\omega_1 t + \psi_1) + \cdots + \boldsymbol{\Omega}_K(x)\sin(\omega_K t + \psi_K)\\ &= [\,\boldsymbol{\Omega}_1(x) \quad \boldsymbol{\Omega}_2(x) \quad \cdots \quad \boldsymbol{\Omega}_K(x)\,]\begin{bmatrix}\sin(\omega_1 t + \psi_1)\\ \sin(\omega_2 t + \psi_2)\\ \vdots\\ \sin(\omega_K t + \psi_K)\end{bmatrix}\\ &= \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{Q}}\end{aligned} \tag{7}$$

$$\begin{aligned}\boldsymbol{Y}(x,t) &= \sum_{n=1}^{K} \boldsymbol{\Omega}_n(x)e^{-\xi_n\omega_n t}\sin(\omega_n t + \psi_n)\\ &= [\,\boldsymbol{\Omega}_1(x) \quad \boldsymbol{\Omega}_2(x) \quad \cdots \quad \boldsymbol{\Omega}_K(x)\,]\begin{bmatrix}e^{-\xi_1\omega_1 t}\sin(\omega_1 t + \psi_1)\\ e^{-\xi_2\omega_2 t}\sin(\omega_2 t + \psi_2)\\ \vdots\\ e^{-\xi_K\omega_K t}\sin(\omega_K t + \psi_K)\end{bmatrix}\\ &= \tilde{\boldsymbol{\Omega}}\tilde{\boldsymbol{Q}}\end{aligned} \tag{8}$$

where $\boldsymbol{\Omega}_i(x) = i$th mode shape magnitude at location $x$; and $\xi_i$, $\omega_i$, and $\phi_i$ are modal damping ratio, frequency, and phase angle for mode $i$. This reiterates the fact that the dense time and space response matrix $\boldsymbol{Y}$ can be decomposed into two submatrices that include modal information. The rows of the matrix $\tilde{\boldsymbol{Q}}$ are decaying harmonics.

In this study, as an alternative for the PCA method for extracting modal components from ALS results, the following optimization problem is considered. Say matrix $\boldsymbol{Y}$ is estimated from $\boldsymbol{Y}_{obs}$ using the ALS method; the new objective function is expressed as Eq. (9), in which optimization variables are shown in Eq. (10)

$$\min_{\hat{\boldsymbol{\Omega}},\hat{\boldsymbol{Q}}} \frac{1}{2}\|\boldsymbol{Y} - \hat{\boldsymbol{\Omega}}\hat{\boldsymbol{Q}}\|^2 \tag{9}$$

$$\hat{\boldsymbol{Q}} = \begin{bmatrix}e^{-w_{11}w_{12}t}\sin(w_{12}t + w_{13})\\ e^{-w_{21}w_{22}t}\sin(w_{22}t + w_{23})\\ \vdots\\ e^{-w_{K1}w_{K2}t}\sin(w_{K2}t + w_{K3})\end{bmatrix}_{K\times N}$$

$$\hat{\boldsymbol{\Omega}} = \begin{bmatrix}\vdots & \vdots & & \vdots\\ \boldsymbol{V}_1 & \boldsymbol{V}_2 & \cdots & \boldsymbol{V}_K\\ \vdots & \vdots & & \vdots\end{bmatrix}_{M\times K} \tag{10}$$

Matrix $\hat{w}$ is unknown in all components, while $\hat{\boldsymbol{Q}}$ has only three unknowns per row ($w_{i1}$, $w_{i2}$, and $w_{i3}$). This optimization problem aims to find the best parameters for $w_{ij}$ and $\boldsymbol{V}_k$ in order to minimize the objective function. A visualization of this decomposition is given in Fig. 2.

By imposing the structure, it is guaranteed that if the optimal parameters are found, they will provide modal property estimates, i.e., frequency, damping, and mode shapes. The ALS method, like other common linear optimization methods for matrix completion

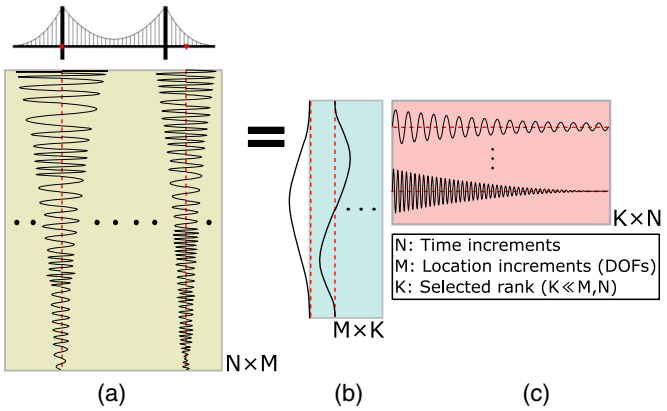


**Fig. 2.** Concept of matrix decomposition from free vibration response: in this case, a bridge is subjected to impulsive loading, and its vibration responses at all DOFs are stacked and form the matrix (a). This response matrix can be accurately approximated as a multiplication of two matrices (b and c), both with rank $K$ (model order). The vertical matrix contains natural mode shapes, while the horizontal matrix includes natural modal fluctuations.

(Cai et al. 2010; Candès and Recht 2009), is inapplicable to such a highly-constrained optimization problem. Therefore, in this study, Newton's optimization method is incorporated for estimating parameters. The main challenge in applying Newton's method to high dimensional data is the numerical calculation of the Hessian inverse (Zhu et al. 1997). An approximate approach is utilized in this study to do this task more efficiently.

### *Ambient Response Analysis using NExT*

The structured optimization technique is suitable only when the impulse responses of an MDOF system are available. In order to generalize this method for random structural vibrations, the natural excitation technique (NExT) is embedded (James et al. 1995) into the proposed method. In short, this technique converts the response of a structure under ambient loading into impulse responses in cross-correlational coordinates. The equation of motion for a damped system under ambient loading is shown in Eq. (11)

$$\boldsymbol{M}\ddot{\boldsymbol{y}}(t) + \boldsymbol{C}\dot{\boldsymbol{y}}(t) + \boldsymbol{K}\boldsymbol{y} = \boldsymbol{F}(t) \tag{11}$$

where $\boldsymbol{M}$, $\boldsymbol{C}$, and $\boldsymbol{K}$ = mass, damping, and stiffness matrices of the system, respectively; $\boldsymbol{y}(t)$ = response vector at time $t$; and $\boldsymbol{F}(t)$ = random loading vector. In case of a random ambient load (e.g., white noise), the cross-correlation function of the response vector with an arbitrary reference response channel results in a free vibration equation of motion, as in Eq. (12)

$$\boldsymbol{M}\ddot{\boldsymbol{R}}_{\ddot{\boldsymbol{y}},\ddot{\boldsymbol{y}}_{ref}}(t) + \boldsymbol{C}\dot{\boldsymbol{R}}_{\ddot{\boldsymbol{y}},\ddot{\boldsymbol{y}}_{\mathrm{ref}}}(t) + \boldsymbol{K}\boldsymbol{R}_{\ddot{\boldsymbol{y}},\ddot{\boldsymbol{y}}_{ref}}(t) = \boldsymbol{0} \tag{12}$$

where $\boldsymbol{R}_{\ddot{\boldsymbol{y}}_i,\ddot{\boldsymbol{y}}_{\mathrm{ref}}(t)}$ = cross-correlation between the response signal $\boldsymbol{y}_i$ at DOF $i$ and a reference response signal $\ddot{\boldsymbol{y}}_{ref}$ (could be any chosen DOF, with minimal constrains). Note that this technique was developed to use with measurements from a fixed sensor network. Yet, in this case, after matrix completion, the columns of the response matrix $\boldsymbol{Y}$, i.e., time histories at the DOFs, are available. Therefore, it is possible to apply NExT on the columns of $\boldsymbol{Y}$ to produce a response matrix with the required structure, i.e., free vibration of a damped SDOF system. A demonstration of how to incorporate NExT with structured optimization analysis (SOA) is given in Fig. 3.

## Modal Identification using Matrix Completion Methods

In this section, the proposed procedure (MIMC) is explained in its three possible forms: (1) matrix completion with PCA; (2) matrix completion with structured optimization analysis; and (3) matrix completion with structured optimization analysis integrated with the NExT method. In all three methods, the preliminary step is to complete the bridge response matrix using the ALS method. The first two methods are designated for bridges subjected to impulsive loads (free vibration response), while the third method is generalized for structures responding to ambient white noise excitations.

### *Method 1: Matrix Completion with PCA*

This method includes two main steps: (1) use ALS to complete the response matrix based on the observed mobile sensor data; and (2) apply PCA to the decomposition matrices to produce

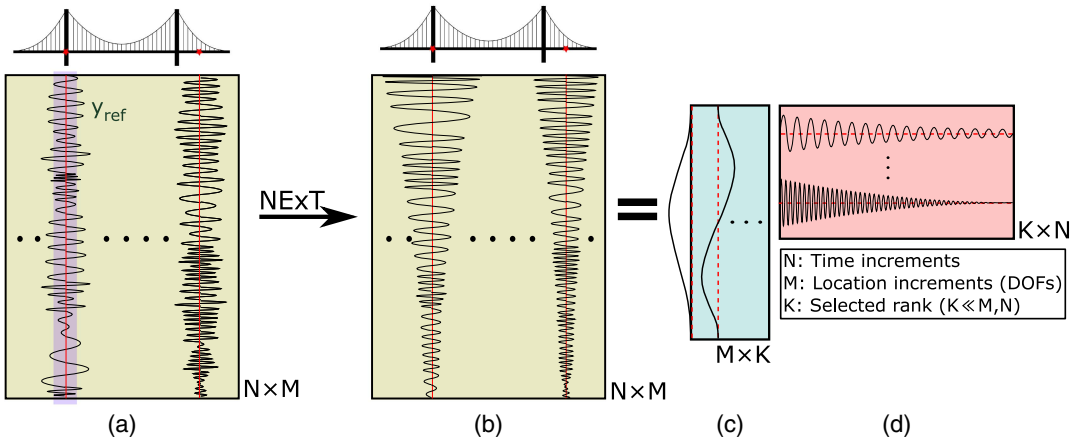


**Fig. 3.** Concept of matrix decomposition from ambient loading response: in this case, a bridge is subjected to ambient loading, and its vibration response is shown as a matrix (a). The vibration responses at each DOF is not decaying, therefore, these random responses are mapped into decaying signals, matrix (b), using NExT. Once this matrix is formed, the rest of the process is identical to the free vibration case (Fig. 2).

orthogonal modes. In order to complete the highly sparse response matrix, the objective function shown in Eq. (5) should be minimized by tuning variable matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. This task can be done by taking gradient steps toward optimality as expressed in Algorithm 1 (Jain et al. 2013).

**Algorithm 1.** Alternating minimization for matrix completion

1: **Input**: $\boldsymbol{Y}_{obs}$, $\boldsymbol{\Phi}(\cdot)$, $T$
2: Initialize $\boldsymbol{A}^1_{M,K}, \boldsymbol{B}^1_{K,N}$ to be random matrices
3: **for** t = 1, ..., T **do**
4: $\quad \boldsymbol{A}^{t+1} = \text{argmin}_{\boldsymbol{A}}(\|\boldsymbol{Y}_{\text{obs}} - \boldsymbol{\Phi}(\boldsymbol{A}^t\boldsymbol{B}^t)\|_2^2)$ ▷ Gradient Descent
5: $\quad \boldsymbol{B}^{t+1} = \text{argmin}_{\boldsymbol{B}}(\|\boldsymbol{Y}_{\text{obs}} - \boldsymbol{\Phi}(\boldsymbol{A}^{t+1}\boldsymbol{B}^t)\|_2^2)$ ▷ Gradient Descent
6: **Return** $\boldsymbol{A}^T, \boldsymbol{B}^T$

In this algorithm, $\boldsymbol{Y}^{M\times N}_{obs}$ = sparse matrix with an assumed model order $K$; $\boldsymbol{\Phi}(\cdot)$ = binary location matrix with ones at observed coordinates and zeros elsewhere; and $T$ = desired number of iterations. Within the argmin(.) function, the variable matrices are updated using a magnitude proportional to their gradients and with a certain step size. For brevity, these common steps are summarized into the argmin(.) function.

After this step, the PCA algorithm is applied on $K$ columns (or rows) of matrix $\boldsymbol{A}$ (or $\boldsymbol{B}$) to decouple its components into estimations of modal coordinates. A similar task was done by Poncelet et al. (2007) using PCA, independent component analysis (ICA), and second-order blind identification (SOBI). These tools are all able to extract uncoupled sources from mixed data by assuming different characteristics for separated sources (e.g., orthogonality, statistical independence, and uncorrelatedness for the three mentioned methods). In this study, the simplest method (PCA) is incorporated for the separation task. In PCA, a matrix is decomposed into its singular vectors and values and then, based on a desired level of accuracy, multiple singular terms with the most participation are used to repopulate an estimation of the original matrix. We say that the estimated matrix $\boldsymbol{Y} = \boldsymbol{A}^T \times B^T$ has a rank equal to the model order $K$. Therefore, by applying PCA, $K$ orthogonal modes are expected that are estimations of the natural mode shapes.

### *Method 2: Matrix Completion with Structured Optimization Analysis*

The second MIMC method can be viewed as a variation of Method 1. In the second step of Method 1, PCA may not produce the expected results, i.e., natural modal coordinates. There are multiple modal coordinates, among which each pair satisfies the orthogonality condition; however, they are not necessarily true structural modes. In fact, PCA performs ideally on regular and symmetric structures with a uniform mass distribution. In these cases, the mass matrix is approximated as a scaled identity matrix and mass orthogonality condition is simplified to the regular orthogonality condition, which is guaranteed in PCA. To force this step of the algorithm to produce the natural modes, structured optimization analysis (SOA) is proposed, as shown in Eq. (10).

To estimate parameters of the structured decomposition matrices [$\hat{\boldsymbol{Q}}$ and $\hat{\boldsymbol{\Omega}}$ in Eq. (10)], an approximate Newton's optimization method is adopted from Eisen et al. (2017). In principle, Newton's optimization method takes the step shown in Eq. (13) toward the optimal point, using second order information from the Hessian

$$w_{n+1} = w_n - \lambda \boldsymbol{H}[f(w_n)]^{-1}\nabla f(w_n) \tag{13}$$

where $\boldsymbol{H}(.)$ = Hessian of a function $f(w)$; and $\lambda$ = step size that ensures the step satisfies Wolfe conditions (Wright and Nocedal 1999) (a requirement for line search algorithms). The inverse calculation for the Hessian is a computationally expensive task for high-dimensional data. In this problem, the dimension includes all entries of matrix $\hat{\boldsymbol{Q}}$ and three parameters for each columns of matrix $\hat{\boldsymbol{\Omega}}$, in total $(M+3)\times K$. Despite its lower dimension with respect to the matrix completion problem using ALS, the problem is still high dimensional, and the inverse Hessian calculation is a bottleneck. To circumvent this issue, k-truncated adoptive Newton's method (k-TAN) is implemented as shown in Algorithms 2 and 3:

**Algorithm 2.** Newton algorithm using approximated Hessian inverse

1: **Input**: $\boldsymbol{Y}_{obs}$, $\boldsymbol{\Phi}(.), eps, threshold$
2: $d = 1$, $\boldsymbol{w}_1 =$ random;
3: **procedure** LOSSFUNCTION(W)
4: $\quad \xi, \omega, \phi, \boldsymbol{B} \coloneqq w$ ▷ $\boldsymbol{w}$ is flattened vector of variables
5: $\quad A = f(\xi, \omega, \phi)$
6: $\quad$ **return** $loss(\boldsymbol{w}) \coloneqq |\boldsymbol{\Phi}(\boldsymbol{AB}) - \boldsymbol{Y}_{obs}|$
7: **while** $d > threshold$ **do**
8: $\quad a_{new}, \boldsymbol{P}_k = LineSearch(eps)$ ▷ Based on strong Wolfe conditions
9: $\quad \boldsymbol{w}_2 = \boldsymbol{w}_1 + a_{new}\boldsymbol{P}_k$
10: $\quad d = |loss(\boldsymbol{w}_2) - \text{loss}(\boldsymbol{w}_1)|$
11: $\quad \boldsymbol{w}_1 = \boldsymbol{w}_2$
12: **Return** $\boldsymbol{w}_1$

**Algorithm 3.** Hessian inverse approximation using truncated absolute eigenvalues

**procedure** INVERSEHESSIAN
  **Input**: $\boldsymbol{H}, eps$
  $\boldsymbol{U}^T\boldsymbol{V}\boldsymbol{U} \coloneqq \boldsymbol{H}$ ▷ Eigenvalue decomposition of Hessian
  $\lambda_i = diag(\boldsymbol{V})$
  **If** $|\lambda_i| > eps$: ▷ Eigenvalue truncation
    $\gamma_i \coloneqq 1/|\lambda_i|$
  **Else:**
    $\gamma_i \coloneqq 0$
  $\boldsymbol{V}_{new} \coloneqq diag(\gamma_i)$
  $\boldsymbol{H}_{inv} \coloneqq \boldsymbol{U}^T\boldsymbol{V}_{new}\boldsymbol{U}$ ▷ Inverse Hessian approximation
  **Return** $\boldsymbol{H}_{inv}$

In this algorithm, *eps* is a threshold for eigenvalue truncation; *threshold* is the acceptable accuracy indicator; and $f(.)$ = function that converts $\xi, \omega$, and $\phi$ into decaying oscillations and stack them to form matrix $\boldsymbol{A}$. The LineSearch(.) function is adopted from Wright and Nocedal (1999) to calculate valid step sizes $a_{\text{new}}$ and produce $\boldsymbol{H}(\text{loss})^{-1}\nabla\text{loss}$ for each time step. Note that $\boldsymbol{\Phi}(.)$ is the binary matrix that activates available entries of $\boldsymbol{Y}_{obs}$ in $\boldsymbol{AB}$. However, because this step (Newton's optimizer) is implemented following ALS, after all entries of $\boldsymbol{Y}_{obs}$ have been estimated, $\boldsymbol{\Phi}(\cdot)$ is equal to a matrix of all ones. Most importantly, $\boldsymbol{H}(.)^{-1}$ is calculated according to Algorithm 3. This algorithm uses truncated absolute values of the Hessian eigenvalues to approximate its inverse. The proposed method using ALS and SOA is illustrated visually in Figs. 4 and 5.

SOA is a nonconvex optimization problem and is sensitive to the initial values of the variables. Accordingly, an appropriate initialization of the frequencies is possible by detecting peaks in the PSD estimate of a random scan. This *warm start* of the frequency variables proved to be sufficient for the algorithm to find the optimal parameters.

### Method 2b: Structured Optimization Only

Method 2 can proceed without the matrix completion step in advance. In Algorithm 2, the objective function still includes $\boldsymbol{\Phi}(.)$, which is an element-wise multiplication of a binary matrix.

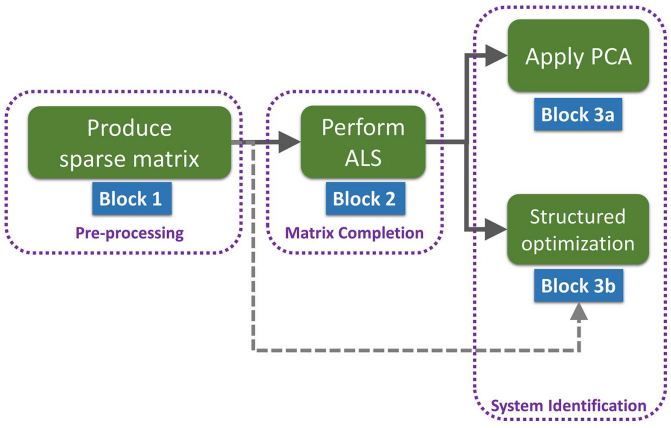


**Fig. 4.** Process of the methods for free vibration-based identification. The method consists of three main phases: Block 1: preprocessing; Block 2: matrix completion; and Block 3: system identification. The third phases can be performed using PCA (Block 3a) or structured optimization (Block 3b).

In Method 2, the binary matrix $\boldsymbol{\Phi}(.)$ is all ones because the unobserved entries of the response matrix are estimated. Alternatively, instead of minimizing the objective function calculated over the entire matrix, $\boldsymbol{Y}_{obs}$ (sparse matrix) can be fed into the algorithm along with its corresponding binary matrix $\boldsymbol{\Phi}(.)$ to tailor this algorithm for an SOA-only method. In other words, SOA has this possibility to be applied directly on the original sparse matrix of the observations, in contrast to PCA in which a preliminary matrix completion step is necessary. This approach is successful in SID when impulse responses are considered; however, for a generalized, random structural response, it is beneficial to integrate NExT into the process. For that case, ALS as a preliminary step is required.

### Method 3: Matrix Completion with Structured Optimization Analysis Integrated with NExT

The third method combines Method 2 with the natural excitation technique (NExT) in order to process response data from a randomly excited structure, which is the most commonly considered scenario in SID. Fig. 6 shows that in this approach, after completing the partially observed response matrix $\boldsymbol{Y}_{obs}$, one DOF is selected as the reference signal ($\boldsymbol{y}_{ref}$ in Fig. 3), and a new matrix of cross-correlated signals with respect to the reference signal is produced. Then, this matrix is estimated using the SOA technique to produce modal property estimates.

The critical hyper-parameter in this method is the index of the reference signal (DOF). According to literature on NExT, the best selection for $\ddot{\boldsymbol{y}}_{ref}$ is a location whose response is influenced by the structural modes of interest (i.e., it is not at or near a modal node for modes of interest). This method needs the matrix completion step to be performed before SOA because the completed matrix after Block 2 (Fig. 6) does not have the presumed decaying structure in its columns. A better illustration of this proposed method is given in Fig. 7.

To evaluate the proposed methods, in the next section, they are validated in multiple numerical simulations.

## Simulation and Results

In this section, three simulations of the MIMC methods are presented to demonstrate their performances for SID: (1) an undamped structure subjected to an impulsive load; (2) a damped structure subjected to an impulsive load; and (3) a damped structure subjected to a random ambient load. The goal of the simulations is to accurately identify the first four modes of a bridge using mobile sensors scanning data. In the first two simulations, Method 1 (ALS+PCA) and Method 2 (ALS+SOA) are implemented and their performances are compared for three different levels of data

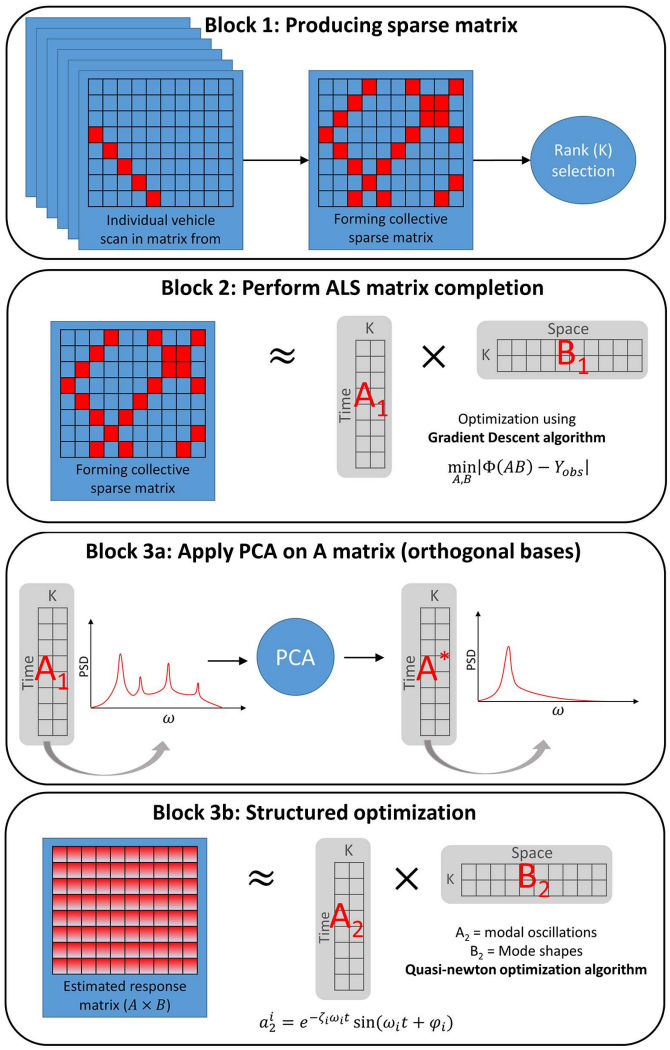


**Fig. 5.** Components of the free vibration-based methods: Block 1: superposing dynamic mobile scans into a global response matrix; Block 2: applying ALS to complete matrix from the sparse matrix; Block 3a: applying PCA to extract uncoupled components from matrices $\boldsymbol{A}$ and $\boldsymbol{B}$; and Block 3b: applying structured optimization to fit free vibration signal parameters for each mode from the completed matrix.

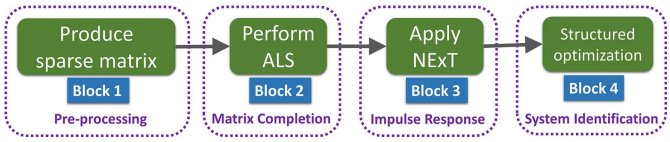


**Fig. 6.** Process of the method for ambient vibration-based identification. The identification phase can be performed by structured optimization after applying NExT to the completed matrix.

availability, which are directly related to the total number of vehicles that contribute to the mobile sensor data set: 75 scans, 100 scans, and 125 scans (cases A, B, and C, respectively). The completeness of the corresponding data matrices in these cases are 0.75%, 1.00%, and 1.25%, respectively. These cases are presented first as a reference to better understand the methods in order to present comprehensively in the third simulation case. Moreover,

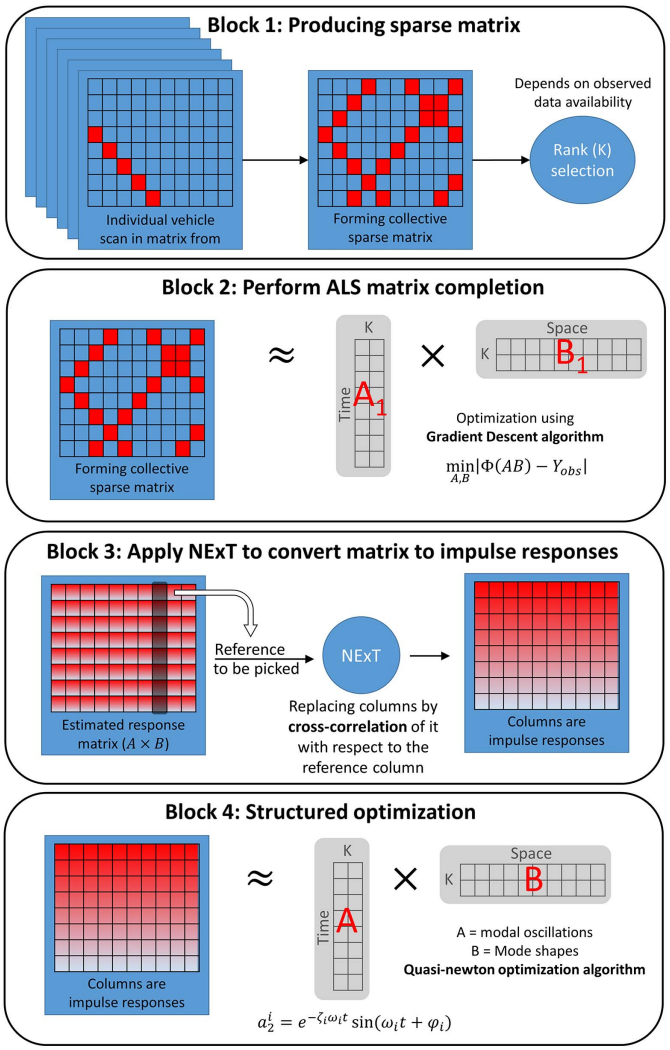


**Fig. 7.** Components of the ambient vibration-based method: Block 1: superposing dynamic mobile scans into a global response matrix; Block 2: applying ALS to complete matrix from the sparse matrix; Block 3: applying NExT to convert nondecaying signals to decaying signals; and Block 4: applying structured optimization for modal properties extraction.

in bridges with potholes or construction/expansion joints, impulsive loads are a proper way of modeling large trucks and buses passing through these bumps and holes. In Simulation III, ambient vibrations are considered along with four data completion levels (0.50%, 0.75%, 1.00%, and 1.25%) for which Method 3 is implemented.

### *Finite-Element Model and Mobile Sensing Setup*

The simulations are based on a 500-m single-span bridge, whose linear responses are modeled using OpenSees. The responses at 5,000 equally-spaced DOFs of the bridge are considered, which are sampled at 50 Hz and have a duration of 100 s. Forces are applied at nine evenly-spaced nodes, along the bridge with magnitudes that vary in each simulation. The response data at all DOFs form a $5{,}000 \times 5{,}000$ dense response matrix from which a sparse response matrix $\boldsymbol{Y}_{obs}$ can be subsampled based on the paths of the simulated mobile sensors. Rayleigh damping is assigned by setting damping ratios equal to 2% for modes one and eight. A modal



J. Eng. Mech., 2020, 146(4): 04020011

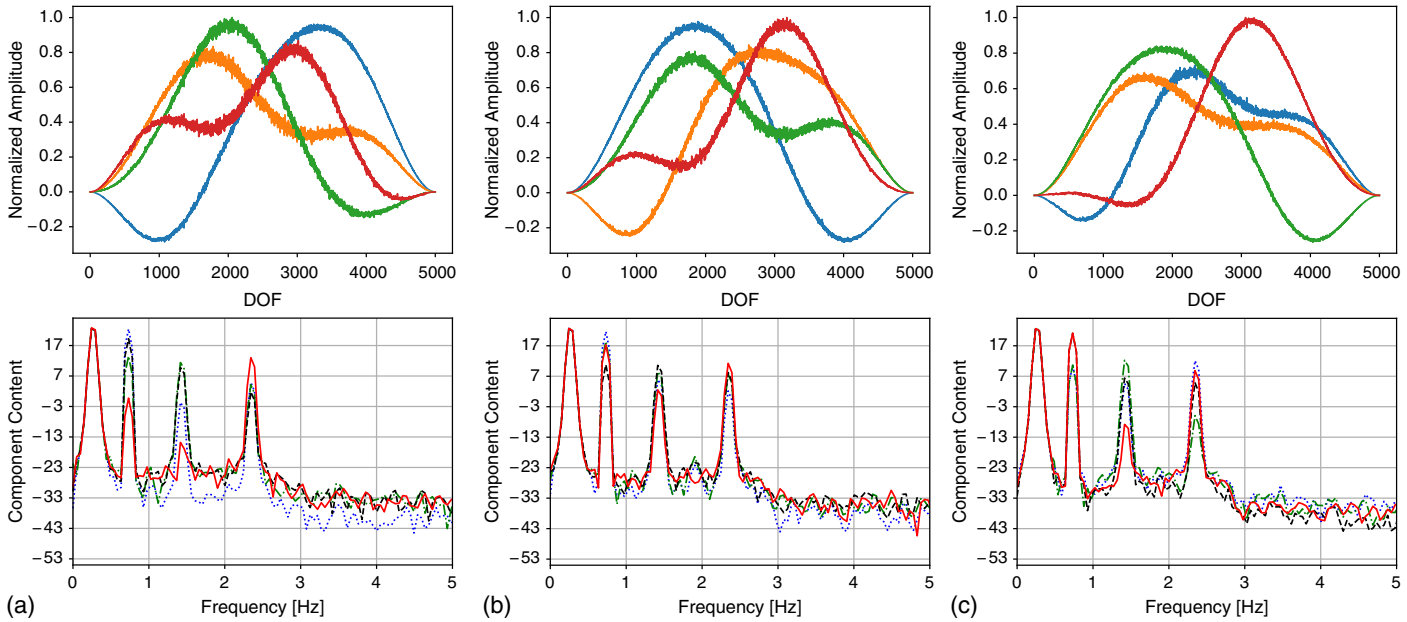


**Fig. 8.** Simulation I: ALS results for three cases: (a) Case A; (b) Case B; and (c) Case C. DOF versus normalized amplitude plots show matrix $\boldsymbol{A}$ components in space and frequency versus component content plots show Welch PSD estimates of matrix $\boldsymbol{B}$ components.

analysis of the structure shows the natural frequencies of the first four modes: 0.2655, 0.7322, 1.4356, and 2.3731 Hz. Because the bridge is modeled in 2D, all modes are vertical. After data generation in OpenSees, the matrix database is exported to Python for further analyses.

The mobile scanning data are generated by selecting appropriate entries from the dense response matrix—a process that is illustrated in Block 1 of Fig. 5. To simulate the random nature of vehicles scanning the bridge, the starting times and locations of each vehicle are selected randomly. This process is repeated $n$ times to mimic a sensing scenario, including $n$ vehicular sensors, which scan the bridge's response within a 100 s period. Note the lengths of the individual mobile scans (diagonals) vary among the vehicles because their trajectories begin at a random point in the data matrix. In other words, the mobile sensors collect data independently. Finally, for simplicity, the speeds of all the sensors are set to 5.0 m/s to allow just enough time (100 s) for a vehicle starting at one end to complete one full-length bridge scan, meaning that most scans only cover a portion of the bridge.

### *Simulation I: Undamped Bridge Subjected to Impulsive Load*

The first step in all the proposed methods is to complete the matrix using the ALS method, for which Algorithm 1 is implemented in Python using TensorFlow framework and performed on the sparse matrix. For applying the algorithm, a value for $K$ (desired model order of the decomposition matrices) must be selected. In all three cases, a $K$ value between four and six yields the best performance—further details about this selection are discussed later. Fig. 8 shows the ALS results for cases A, B, and C: top plots display the components of the $\boldsymbol{A}$ matrix and bottom plots show the power spectral density (PSD) estimates of the components of the $\boldsymbol{B}$ matrix. The components of the matrix $\boldsymbol{A}$ are similar to the expected mode shapes of the bridge; however, they are not exactly the natural modes. The peaks of the PSD estimates of the components of the $\boldsymbol{B}$ matrices correspond to the modal frequencies. Note these components are coupled, i.e., each component does not have a distinct peak. PCA is used to make these components orthogonal to one another.

PCA is performed on columns of matrix $\boldsymbol{A}$ as well as rows of matrix $\boldsymbol{B}$ to produce orthogonal modes. Fig. 9 shows the mode shapes resulting from PCA for cases A, B, and C, respectively. The estimated mode shapes for the first four modes are consistent with the true mode shapes of the bridge as indicated by the modal assurance criteria (MAC) (Allemang and Brown 1982) values in Table 1, which are all greater than 0.95. The first and second mode shapes were perfectly identified in all cases. The PSD estimates of these extracted modes for all three cases are presented in Fig. 10. In this figure, each mode shows only one peak in its frequency representation, which means PCA is successful in decoupling the modes.

It is important to notice that the identified mode shapes contain 5,000 points; a very high resolution—in fact, these are the densest mode shapes that are extracted from accelerometer data in the literature of the field. The implication is that, in case A, on average, each mobile sensor scan produced about 66 points.

The same problem is approached using Method 2 (matrix completion with structured optimization). In this method, after the matrix completion step, the response matrix is reformed into a structured matrix to approximate modal properties of each mode. Likewise, Algorithm 2 is implemented in Python to tune unknown variables. The estimated mode shapes for cases A, B, and C are shown in Fig. 11. In Method 2, the number of points in the identified mode shapes is reduced to 100, a user-defined value. This particular selection enables a quick solution to the optimization problem while maintaining a dense mode shape.

The quality of the identified mode shapes is quantified using MAC values and displayed in Table 1. The MAC values for all four modes are above 0.94 in all three cases, which indicates the identified shapes are consistent with the true mode shapes. In addition, the mode shape estimates are insensitive to data completion levels (from case A to case C). MAC values for two first modes using both PCA and SOA are above 0.99. This accuracy is especially significant considering the high resolution of the identified mode shapes. Comparing two methods in natural frequency estimation shows that PCA has a 2% error in the worst case, while SOA estimates are within 1% of the true values. Regarding the frequency

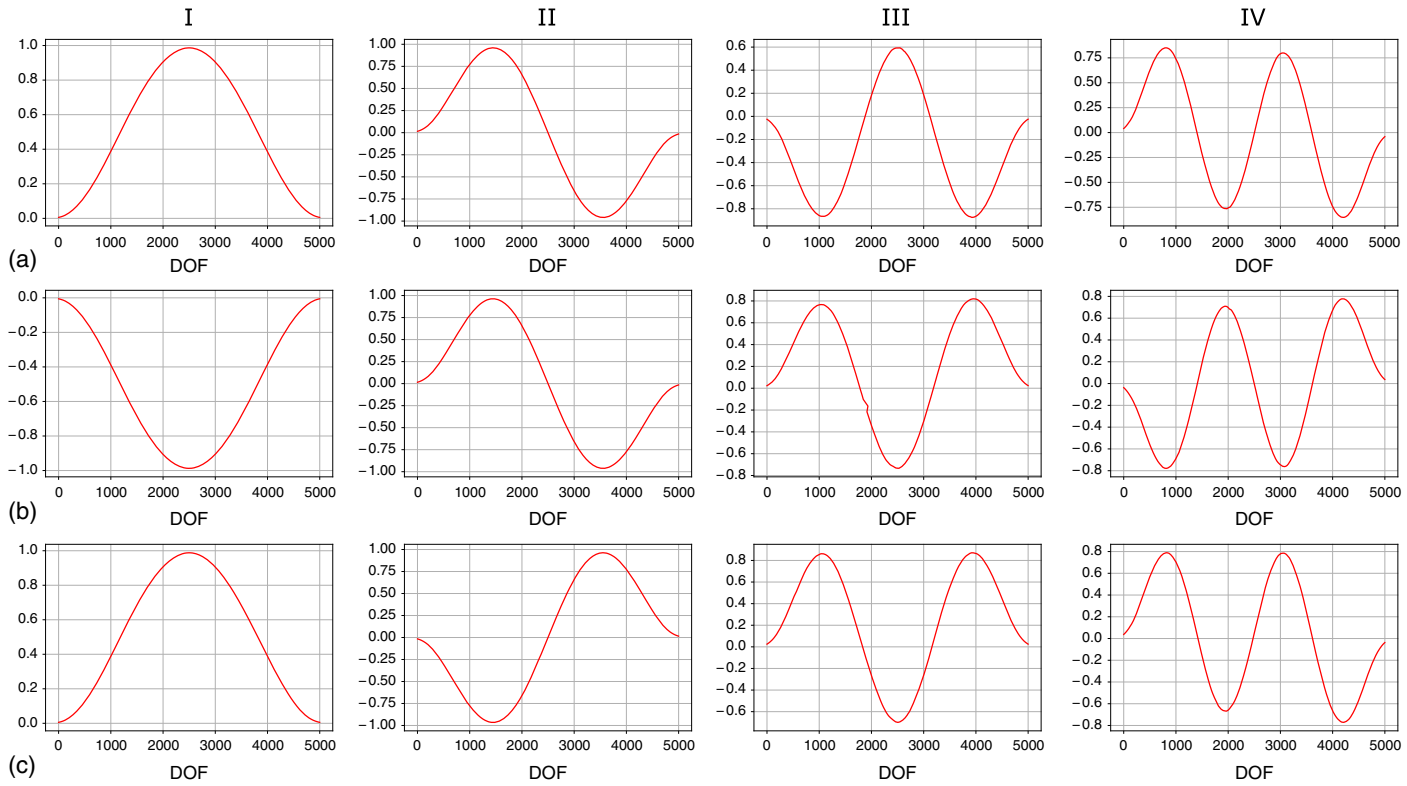


**Fig. 9.** Simulation I: mode shapes resulted from Method 1: (a) Case A; (b) Case B; and (c) Case C.

**Table 1.** Identification results for simulation I

| Case | Method | Mode 1 | Mode 2 | Mode 3 | Mode 4 |
|---|---|---|---|---|---|
| | | MAC values | | | |
| A | PCA | 0.9999 | 0.9993 | 0.9408 | 0.9585 |
| | SOA | 0.9995 | 0.9987 | 0.9575 | 0.9511 |
| B | PCA | 0.9999 | 0.9996 | 0.9577 | 0.963 |
| | SOA | 0.9995 | 0.9985 | 0.9418 | 0.9543 |
| C | PCA | 0.9999 | 0.9993 | 0.9533 | 0.9566 |
| | SOA | 0.9995 | 0.9988 | 0.9485 | 0.9519 |
| | | Natural frequencies (Hz) | | | |
| A | PCA | 0.2562 | 0.7291 | 1.4187 | 2.3498 |
| | SOA | 0.2656 | 0.7317 | 1.4066 | 2.3554 |
| B | PCA | 0.2704 | 0.7375 | 1.4356 | 2.3451 |
| | SOA | 0.2656 | 0.7316 | 1.4322 | 2.3553 |
| C | PCA | 0.2653 | 0.7274 | 1.4305 | 2.3499 |
| | SOA | 0.2656 | 0.7317 | 1.4313 | 2.3553 |
| — | Actual | 0.2655 | 0.7322 | 1.4356 | 2.3731 |

representation of the modes identified with SOA, Fig. 12 shows that in all three cases, modes are independent and without spectral leakage. By comparing these results with Fig. 10, the advantage of SOA is noticeable; SOA includes a constraint that guarantees natural mode extraction, which is different from PCA. This special feature is demonstrated again in the following simulation.

Table 1 summarizes the estimated modal frequencies from both methods and all three cases; the results confirm that both methods produce accurate modal property estimates. Overall, SOA provided more accurate frequency estimates than PCA. It is important to keep in mind, that while successful, this simulation was basic; a damped structure is considered in Simulation II.

### *Simulation II: Damped Bridge Subjected to Impulsive Load*

Simulation II is similar to simulation I except that damping is considered. After forming a sparse matrix of the damped bridge

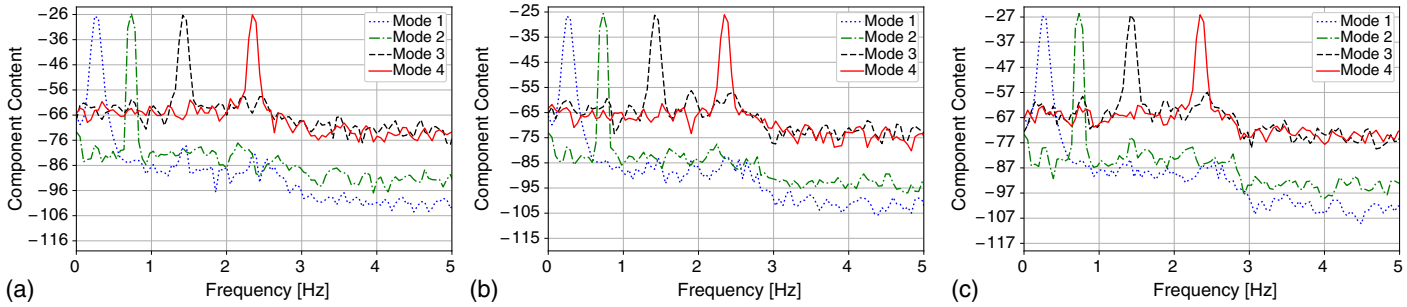


**Fig. 10.** Simulation I: Welch PSD estimates of identified modes for (a) Case A; (b) Case B; and (c) Case C using Method 1.

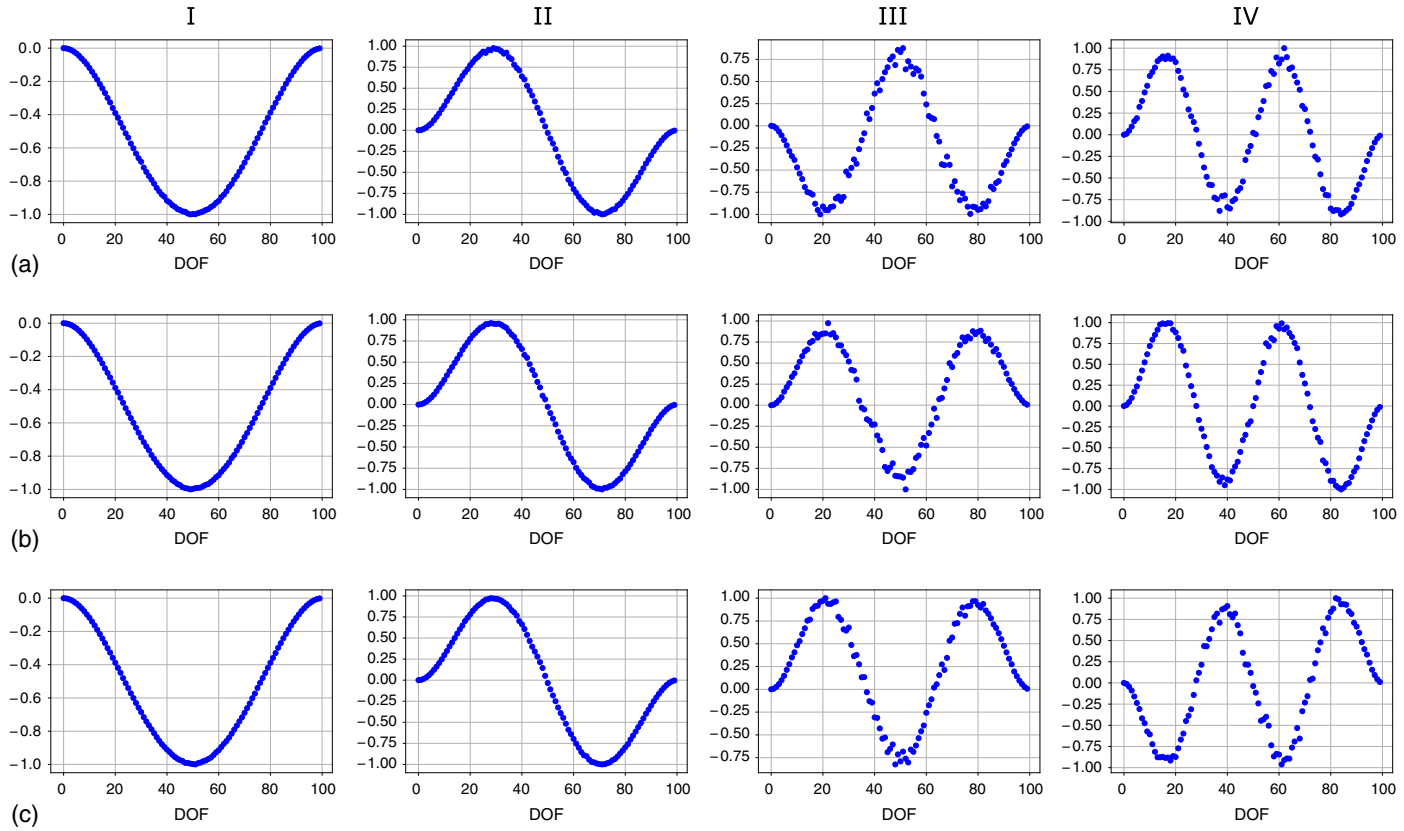


**Fig. 11.** Simulation I: mode shapes resulted from Method 2: (a) Case A; (b) Case B; and (c) Case C.

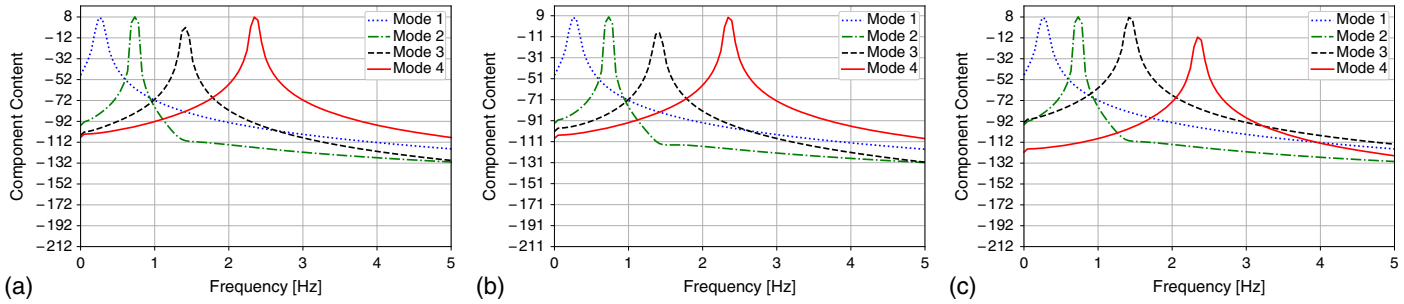


**Fig. 12.** Simulation I: Welch PSD estimates of identified modes for (a) Case A; (b) Case B; and (c) Case C using Method 2.

response subjected to an impulsive load from finite-element analysis, the matrix is completed using the ALS algorithm; the results are shown in Fig. 13. These plots are different when compared with ALS results of the undamped case; for instance, the components of the $\boldsymbol{A}$ matrix are noisy and contain outliers. In addition, the peaks in the PSD estimates are less prominent. One explanation for this difference is that the matrix completion algorithm performs more desirably when the values of the matrix do not decay significantly. In the damped case, the rapid decays of the signal interfere with the algorithm's ability to properly estimate the tails. Despite this, Table 2 shows that Method 1 is still successful in identifying highly dense mode shapes. A comparison between Fig. 10 (undamped) and Fig. 14 (damped) highlights the spectral leakage among the modes in the frequency domain in the damped case, which influences the amplitudes of each peak.

Regarding three sparsity cases, it is observed that as the amount of available data increases, the estimated components of matrix $\boldsymbol{A}$ are less noisy. However, to address this noise (unwanted high frequency content), a moving average window is applied to the mode shapes resulting from PCA to filter out high frequency noises.

In Method 1 (matrix completion with PCA), the damping ratios for each mode are calculated based on free vibration decay (Chopra 2017). After applying PCA, the corresponding modal coordinates are plotted in time, the amplitude decay within a certain number of cycles is measured, and the modal damping is estimated. The results shown in Table 2 indicate that the damping ratios are estimated properly.

Next, Method 2 is used, in which SOA is implemented after matrix completion. This algorithm is very fast and converges in

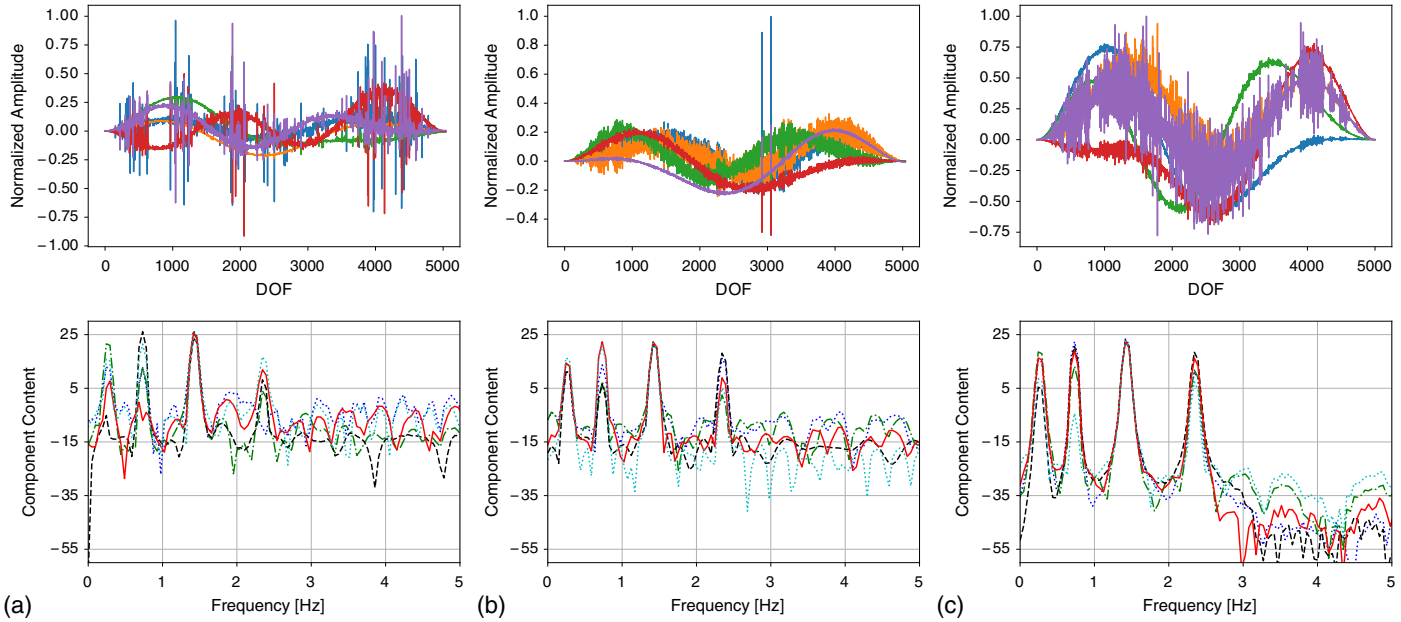


**Fig. 13.** Simulation II: ALS results for three cases: (a) Case A; (b) Case B; and (c) Case C. DOF versus normalized amplitude plots show matrix $\boldsymbol{A}$ components in space and frequency versus component content plots show Welch PSD estimates of matrix $\boldsymbol{B}$ components.

**Table 2.** Identification results for simulation II

| Case | Method | Mode 1 | Mode 2 | Mode 3 | Mode 4 |
|---|---|---|---|---|---|
| | | MAC | | | |
| A | PCA | 0.9993 | 0.9986 | 0.9578 | 0.9582 |
| | SOA | 0.9902 | 0.9948 | 0.9555 | 0.7996 |
| B | PCA | 0.9959 | 0.9260 | 0.9579 | 0.9426 |
| | SOA | 0.9875 | 0.9994 | 0.9554 | 0.9246 |
| C | PCA | 0.9933 | 0.9929 | 0.9579 | 0.9700 |
| | SOA | 0.9831 | 0.9988 | 0.9557 | 0.9572 |
| | | Natural frequencies (Hz) | | | |
| A | PCA | 0.2655 | 0.7325 | 1.4258 | 2.3500 |
| | SOA | 0.2657 | 0.7318 | 1.4313 | 2.3495 |
| B | PCA | 0.2606 | 0.7325 | 1.4307 | 2.3599 |
| | SOA | 0.2655 | 0.7319 | 1.4313 | 2.3539 |
| C | PCA | 0.2706 | 0.7429 | 1.4315 | 2.3653 |
| | SOA | 0.2657 | 0.7319 | 1.4316 | 2.3550 |
| — | Actual | 0.2655 | 0.7322 | 1.4356 | 2.3731 |
| | | Damping ratios (%) | | | |
| A | PCA | 2.048 | 1.004 | 0.707 | 0.473 |
| | SOA | 1.886 | 0.777 | 0.576 | 0.450 |
| B | PCA | 2.101 | 0.976 | 0.619 | 0.521 |
| | SOA | 1.783 | 0.751 | 0.574 | 0.535 |
| C | PCA | 2.155 | 0.923 | 0.735 | 0.933 |
| | SOA | 1.882 | 0.792 | 0.576 | 0.539 |
| — | Actual | 2.000 | 0.800 | 0.600 | 0.700 |

fewer than 20 iterations. In contrast to Method 1, after convergence, the mode shapes, damping ratios, and frequencies are all calculated separately, and there is no need for post processing. The identified modes are not presented for brevity; however, the MAC values, damping ratios, and frequencies are presented in Table 2.

Overall, the mode shapes are estimated accurately; MAC values in Table 2 exceed 0.92 in all cases (except for the fourth mode in case A). Note that the identified modes from Method 1 consist of 5,000 points, while Method 2 results in 100 points. In terms of natural frequency estimates, Table 2 shows that both methods were equally successful. Estimated damping ratios shown in Table 2 indicate that both algorithms are successful in this aspect. SOA results in almost exact estimations for the first three modes while PCA estimates are not as accurate. The PSD estimates in Fig. 15 display four independent and smooth frequency plots for the structural modes.

Until now, we showed that the proposed methods are successful for estimating modal characteristics of a bridge (damped or undamped) subjected to the impulsive load and using mobile sensors data. In the next simulation, a more realistic loading case (ambient random load) is investigated.

### *Simulation III: Damped Bridge Subjected to Ambient Load (Operational Condition)*

In this section, Method 3 is used for SID using ambient vibration data. Based on the results in the previous two simulations, it is concluded that SOA produces more consistent SID results. Simultaneously, PCA excels at determining modes that are orthogonal to one another; yet, it does not automatically produce true structural modes as the condition for these is orthogonality with respect to the mass and stiffness matrices. As a result, PCA performs ideally on single-span bridges with a uniform mass because in these cases, the condition for component orthogonality is equivalent to that for mass-stiffness orthogonality. SOA, on the other hand, is not built on this assumption and is widely applicable to more generic structures. Based on these considerations, SOA is recommended to be used with NExT for ambient vibrations and is selected for this simulation. In this example, to further test the limits of sparse data, a lower level of data completion case is added. The four cases, 50, 75, 100, and 125 scans, correspond to data completion levels 0.50%, 0.75%, 1.00%, and 1.25%, respectively.

The bridge is excited randomly at nine locations while the mobile sensors scan the response. This case is simulating the passage of a network of moving sensors over evenly spaced potholes or expansion joints. However, other random loading patterns with

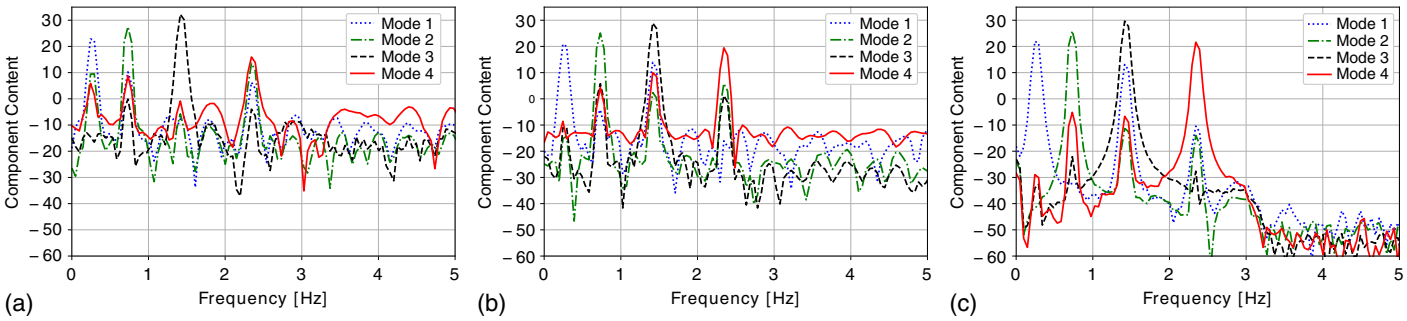


**Fig. 14.** Simulation II: Welch PSD estimates of identified modes for (a) Case A; (b) Case B; and (c) Case C using Method 1.

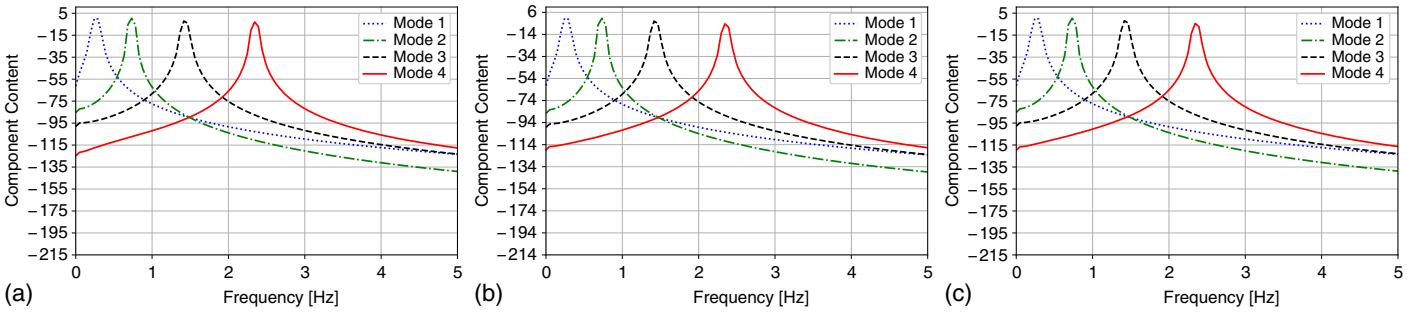


**Fig. 15.** Simulation II: Welch PSD estimates of identified modes for (a) Case A; (b) Case B; and (c) Case C using Method 2.

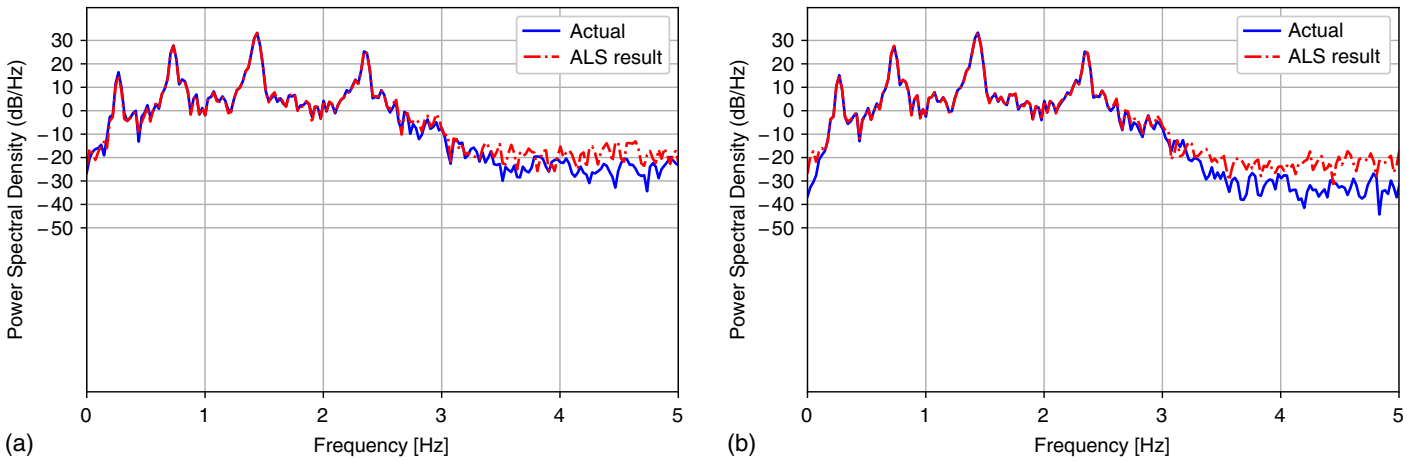


**Fig. 16.** Simulation III: Welch PSD estimates of completed versus actual bridge response signals: (a) 50 scans; and (b) 125 scans.

more loading points can be assigned with no loss of generality in the performance. In addition, 5% white measurement noise is added to the mobile sensing data. The first step is to apply ALS to complete the response matrix. To demonstrate this process, Fig. 16 compares the PSD estimates of the true response at the 1,000th DOF with the ALS approximations for two extreme sparsity cases. Overall, the ALS results contain accurate frequency content—especially below 3 Hz. These plots do not indicate any significant differences among the data completion levels.

After the matrix completion step, NExT is applied to the estimated signals (according to the procedure in Fig. 6). Out of the 100 DOFs available (reduced from 5,000 to improve the CPU speed), the 40th is selected as the reference signal because it includes contributions from all four modes. Then, a response matrix with decaying signals in each column is constructed for SOA. Fig. 17 shows the estimated modes for each of the data availability levels. In general, the modal property estimates from this simulation are more accurate than those from Simulation II.

The estimated mode shapes are evaluated using MAC values in Table 3. The MAC values for the identified mode shapes, for all modes, and all sparsity cases are greater than 0.95, which indicates a strong agreement with the true mode shapes. In particular, the shapes for the first and second modes were perfectly identified in all incomplete data cases, as measured by MAC values over 0.99. Furthermore, it is worth reiterating that Method 3 is successful in estimating high resolution mode shapes

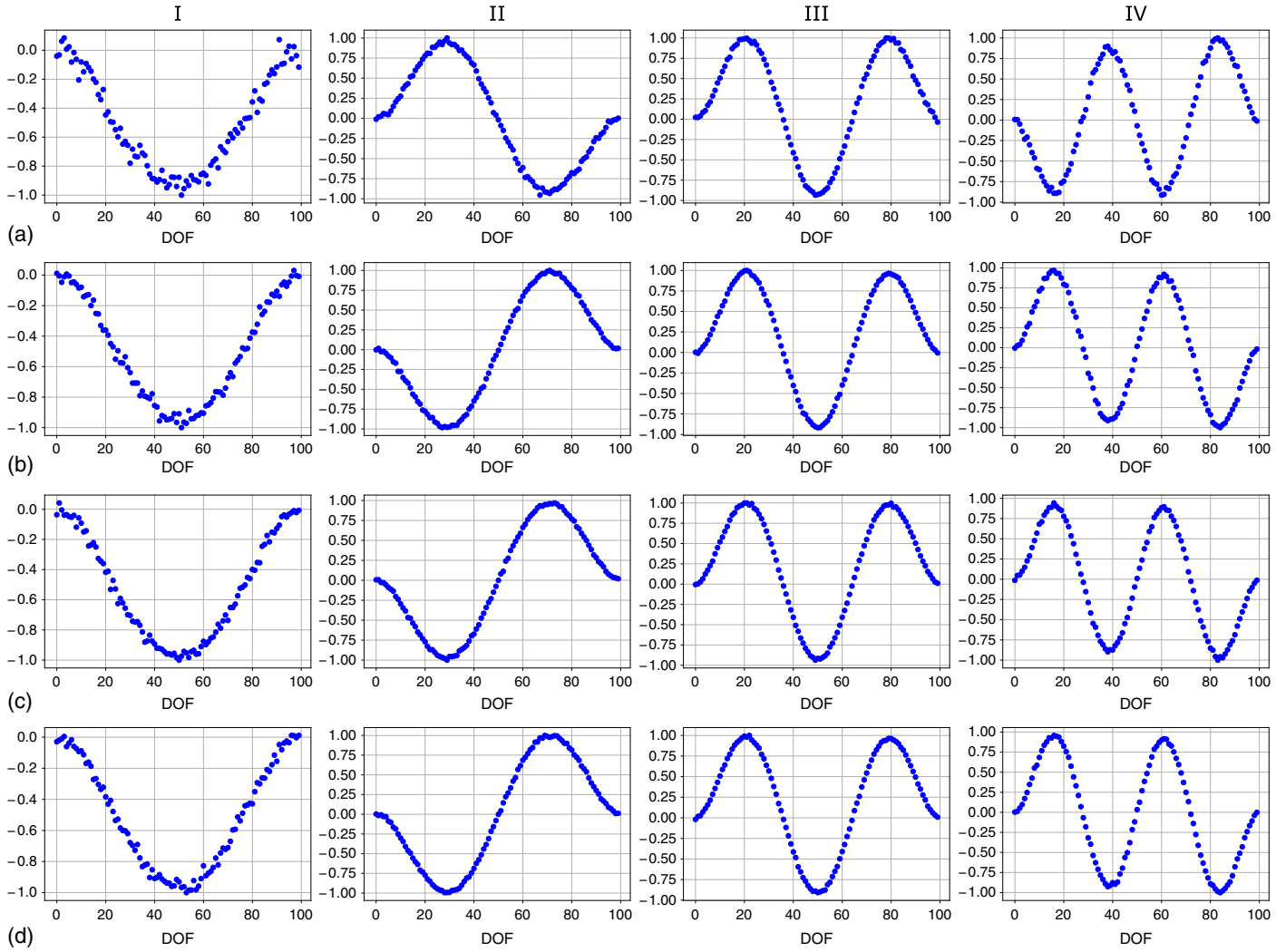


**Fig. 17.** Simulation III: mode shapes resulted from Method 3: (a) 50 scans; (b) 75 scans; (c) 100 scans; and (d) 125 scans.

(100 points each). Fig. 18 illustrates the method's accuracy in extracting clear and independent modal frequencies without spectral leakage. This result can be attributed to the structure imposed by SOA.

Lastly, by superposing identified modes and comparing with the completed response matrix results, Fig. 19 is created. This figure demonstrates that the combined reconstructed signal from SOA ideally resembles the underlying natural frequency contents of the bridge. In addition, by comparing figures, different cases (data availability variations) result in the same level of modal estimation accuracy. Table 3 shows that the natural frequency estimations are all within 0.7% error range, which is promising. Damping ratio estimations from Table 3 agree closely with the actual values, especially in the first three modes. The same level of accuracy is also evident from MAC values of the estimated mode shapes presented in Table 3 (two first modes are identified with MAC values higher than 0.99). By comparing the accuracy of the estimated parameters in different data availability cases, it is evident that MIMC performs consistently well within the tested range (50–125 scans). This is important because it shows the robustness of the proposed method for different mobile data availability. This robustness was also observed in two former simulations and in SOA results. In fact, in SOA, the number of unknown variables reduces from $(M + N) \times K$ for ALS step to $(M + 3) \times K$ ($M \times K$ for mode shape estimations and $3 \times K$ for modal coordinate estimations), and this lowers the sensitivity of the algorithm to the observed data availability.

## Discussion

In the previous sections, the proposed MIMC methods were applied to a set of examples for demonstration and validation. This section provides further information to assist users with proper implementations of MIMC methods.

Operations on high dimensional data are vulnerable to computational instabilities. ALS and SOA are two optimization techniques based on sparse data whose performances depend on a number of parameters. In the matrix completion with ALS, a decaying step size, tuned for the gradient descent optimization, provided an adequate performance. A decaying step size is generally suitable for smooth and convex objective functions. The SOA objective function is nonconvex; thus, a line search algorithm with strong Wolfe conditions (Wright and Nocedal 1999) was implemented, as presented in Algorithm 2.

In the optimization process (ALS step), the sparse matrix is more accurately estimated if the model order $K$ is slightly more than the expected rank of the response matrix. In general, the rank of the decomposition matrix is close to the target number of modes

**Table 3.** Identification results for simulation III

| Scans | Mode 1 | Mode 2 | Mode 3 | Mode 4 |
|---|---|---|---|---|
| | | MAC | | |
| 50 | 0.9932 | 0.9981 | 0.9546 | 0.9527 |
| 75 | 0.9965 | 0.9978 | 0.9552 | 0.9542 |
| 100 | 0.9978 | 0.9980 | 0.9552 | 0.9549 |
| 125 | 0.9976 | 0.9972 | 0.9555 | 0.9573 |
| | | Natural frequencies (Hz) | | |
| 50 | 0.2654 | 0.7316 | 1.4311 | 2.3549 |
| 75 | 0.2654 | 0.7333 | 1.4312 | 2.3555 |
| 100 | 0.2654 | 0.7331 | 1.4311 | 2.3554 |
| 125 | 0.2654 | 0.7330 | 1.4311 | 2.3553 |
| Actual | 0.2655 | 0.7322 | 1.4356 | 2.3731 |
| | | Damping ratios (%) | | |
| 50 | 1.855 | 0.894 | 0.421 | 0.214 |
| 75 | 1.857 | 1.004 | 0.426 | 0.213 |
| 100 | 1.867 | 0.996 | 0.422 | 0.213 |
| 125 | 1.867 | 0.993 | 0.349 | 0.213 |
| Actual | 2.000 | 0.800 | 0.600 | 0.700 |

to be identified (e.g., four in this study). However, increasing the rank (model order) marginally can improve ALS optimization based on the sparse data and produce beneficial SID results (Eshkevari and Pakzad 2020). In all simulations, $K$ is tuned to produce the best performance.

In Simulation II, Fig. 14 shows that despite the fair accuracy of the identified modes (shown in Table 2), the modes from PCA are not as well separated as the first simulation (modal interactions are present). Fig. 15 shows that this problem does not exist in the modes identified with SOA. In SOA, modal interactions or spectral leakage are unlikely, because the structure only accepts one set of frequency, damping ratio, and phase shift parameters for each component; this constraint guarantees the extraction of distinct and independent modes.

For a matrix completion problem, a range of data completion between 0.5% and 2.0% is reasonable (Candès and Recht 2009). This ratio is convertible to different numbers of mobile scans according to the dimension of the response matrix. For instance, in this study, the response matrix is $5{,}000 \times 5{,}000$, which needs nearly 100 mobile scans to suffice a 1.0% data completion rate. These 100 scans are divided in the number of lanes over the bridge (e.g., for the numerical case study, we can assume three lanes in each direction). Therefore for each lane, 17 scans are needed within the monitoring duration (or 25 scans for a four-lane bridge). Considering the sampling rate, the vehicle speed, and the length of the bridge, this number is easily practical. The important note in this study is that these 17 scans do not need to visit all DOFs within the time frame, e.g., partial scans are acceptable. For crowded bridges, the number of mobile scans is significantly more than this range; however, the operator can randomly pick sufficient measurements and use for the matrix completion and modal identification tasks.

To better understand the effect of model order $K$ with respect to the data availability levels, various $K$ values and scan levels were considered by Method 3. Table 4 shows the corresponding final objective function values of ALS and SOA and documents the runtimes and number of identified modes. In this table, ALS objective function values are reported per observed matrix entry and SOA values are reported over the entire response matrix after completion with normalization. A smaller ALS objective function value indicates a more successful fit based on the observed entries and suggests a more accurate signal reconstruction. Generally, as $K$ increases, the ALS objective function decreases. At the lowest data availability (50 scans), when $K$ grew from 5 to 6, the ALS objective function value increased slightly—which implies that a limit had been reached. For the more complete data cases, the ALS objective function always decreased as $K$ increased. Similarly, the SOA objective function value measures the success in imposing a structure over the response matrix; however, this value was much less sensitive to changes in $K$ or data availability.

The runtimes of each technique are compared in Table 4; while these are subject to the processor's type, the relative metrics are informative. At the lowest data level (50 scans) and with $K$ equal to 6, five modes were identified; however, the runtime for ALS was significant. Alternatively, five modes can be identified ten times faster, and about four times more accurately, when 100 or 150 scans were available (Table 4). Generally, when the number of scans increased from 100 to 150, there was no improvement in SID. These points demonstrate some of the trade-offs between $K$, data availability, and the target number of identified modes. In certain circumstances, more data does not improve the SID results.

There are numerous variables that influence the data collected by a vehicular sensor network. In this paper, the proposed MIMC methods were shown to produce accurate SID results in the case of a random monitoring process. The speed of the mobile sensors is an influential parameter that was not explicitly studied in this paper—instead it was linked to a particular data duration and the bridge

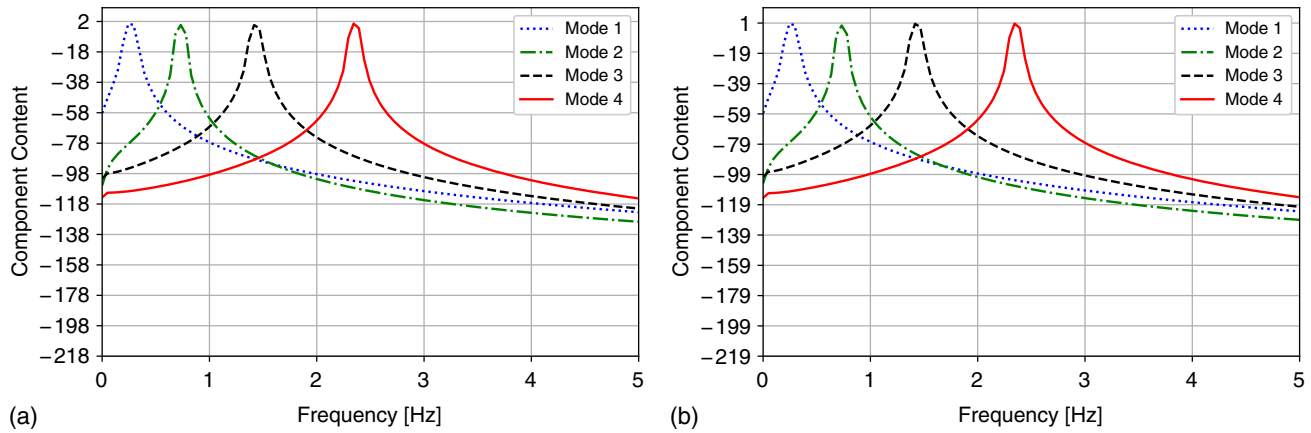


**Fig. 18.** Simulation III: Welch PSD estimates of identified modes using Method 3: (a) 50 scans; and (b) 125 scans.



J. Eng. Mech., 2020, 146(4): 04020011

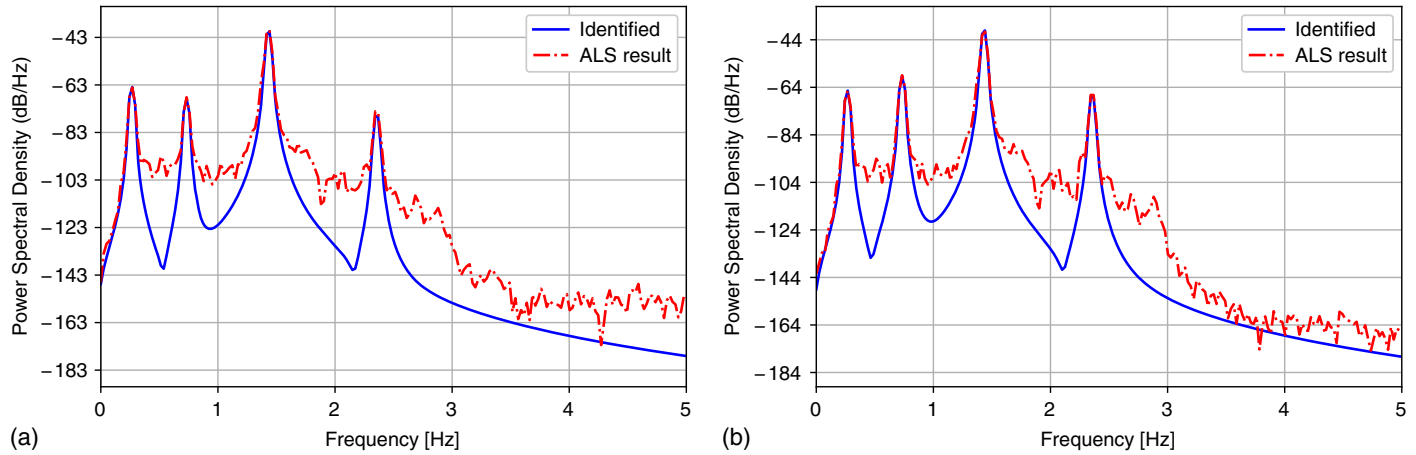


**Fig. 19.** Simulation III: Welch PSD estimates of optimal structured response versus ALS estimated signal: (a) 50 scans; and (b) 125 scans.

**Table 4.** Runtimes, number of identified modes, and objective functions for various K and data availability

| Measure | Scans | K = 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Runtime (s) | 50 | 155 | 161 | >1,000 | >1,000 | >1,000 |
| | 100 | 31 | 30 | 55 | 93 | 162 |
| | 150 | 46 | 31 | 57 | 31 | 233 |
| Identified modes | 50 | 2 | 3 | 3 | 4 | 5 |
| | 100 | 0 | 3 | 4 | 5 | 5 |
| | 150 | 2 | 3 | 4 | 5 | 5 |
| | | | Objective function values | | | |
| ALS step | 50 | 77.1751 | 40.6841 | 23.9679 | 11.0315 | 12.7113 |
| | 100 | 90.8833 | 51.8795 | 29.3165 | 7.3299 | 3.6893 |
| | 150 | 93.5409 | 51.0089 | 26.5513 | 9.0379 | 4.6717 |
| SOA step | 50 | 0.0221 | 0.0479 | 0.0535 | 0.0025 | 0.0159 |
| | 100 | 0.1313 | 0.0483 | 0.0476 | 0.0446 | 0.0404 |
| | 150 | 0.0378 | 0.0549 | 0.0807 | 0.0469 | 0.0461 |

length. It is important to clarify that this example was presented for demonstration purposes and to show that there are no aspects of the MIMC methods algorithms that restrict speed variations of the mobile sensors. The velocities of the individual vehicles affect the locations of the observed entries in the response matrix and, in turn, the shape of the available data (see the first block of the proposed methods shown in Figs. 5 and 7). Matrix completion literature (Candès and Recht 2009) discusses how sparse matrices with randomly arranged entries provide an ideal starting point for reaching the global optimum, i.e., yielding the most accurate full response matrix. Thus, a stochastic vehicular sensor network with various speeds, sampling rates, scanning intervals, etc., is better suited to achieve this condition. In addition, speed variations can help to better facilitate the proposed methods for shorter span bridges. For instance, if the bridge span is 300 m and the vehicle network crosses the bridge with a speed of 10 km/h, the same level of spatial discretization could be achieved. This means that the response matrix will have the same dimensions, and consequently, the conclusions of the current case studies (e.g., high resolution natural mode shape identification) are achievable.

## Conclusion

In this study, novel methods were proposed for a comprehensive modal identification of a bridge based on data collected by a large number of moving sensors (i.e., vehicles). Bridge response data collected by a vehicular sensor network are both sparse and dynamic. The full (unobserved) bridge response is viewed as a very large data matrix of which the aggregate mobile sensing data provide a sparse representation. The modal identification using matrix completion methods proposed utilizing matrix completion, i.e., alternating least squares (Jain et al. 2013), to complete the full matrix based on sparse entries. Then, the completed matrix was analyzed to extract a complete set of modal properties, e.g., frequencies, damping ratios, and high-resolution mode shapes. For this, two algorithms, principal component analysis (Jolliffe 2011) and structured optimization analysis were proposed (the latter was developed by the authors) and applied on the completed matrix (Methods 1 and 2, respectively). To extend the applicability of this technique to ambient structural vibrations, a third method was proposed based on the natural excitation technique (James et al. 1995) (Method 3).

The proposed methods were evaluated numerically in three different simulation studies, and results were presented. Method 1 was able to extract natural mode shapes of the bridge under impulsive loading with 5,000 points, which are the densest identified mode shapes in the existing literature. Alternatively, Method 2 can extract mode shapes with a user-defined number of points. In Simulation III, Method 3 showed that it is a robust and accurate SID solution for bridges using mobile sensor network with minimal sensitivity to the data completion rate. However, the methods generally improve in terms of computational costs and SID results when more data

completion rates are available. It was also shown that Method 2 and Method 3 could extract fully decoupled modal components while Method 1 suffered from modal leakage in some modes. A sensitivity study on the user-defined model order $K$ (rank of the decomposition matrices) was done. The study showed that while larger $K$'s usually lead to better signal reconstruction results, it will increase the computational costs. A balanced configuration of $K$ and data completion rate yielded the best performance.

The accuracy of estimated modal properties is promising in all three simulations and methods (within 2% error in most cases). Method 1 was able to identify a very high-resolution fundamental mode with MAC value equals to 0.9999. The frequency estimated in all three simulations are very accurate (e.g., Method 3 resulted in 0.07% estimation error in the worst case on Simulation III). Method 2 and 3 showed a desired performance in damping ratio estimation, especially on the first three modes. This work further supports the practice of dynamic sensor networks for SHM applications, especially system identification. MIMC methods are applicable to data collected by vehicular sensor networks, which present new opportunities to monitor bridge vibrations at unprecedented rates and scales. This proposed methodology paves the way toward a fully autonomous and real-time bridge health monitoring platform using crowdsourced data provided by smart devices.

## Data Availability Statement

Models and codes generated and used during the study are available from the corresponding author by request.

## Acknowledgments

Research funding is partially provided by the National Science Foundation through Grant CMMI-1351537 by the Hazard Mitigation and Structural Engineering program, Grants CCF-1618717 and CCF-1740796, and by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA). The authors would like to thank Anas S.p.A, Allianz, Brose, Cisco, Dover Corporation, Ford, the Amsterdam Institute for Advanced Metropolitan Solutions, the Fraunhofer Institute, the Kuwait-MIT Center for Natural Resources and the Environment, LabCampus, RATP, Singapore-MIT Alliance for Research and Technology (SMART), SNCF Gares & Connexions, UBER, and all the members of the MIT Senseable City Lab Consortium for supporting this research.

## References

Alexander, L., S. Jiang, M. Murga, and M. C. González. 2015. "Origin–destination trips by purpose and time of day inferred from mobile phone data." *Transp. Res. Part C: Emerging Technol.* 58 (Part B): 240–250. https://doi.org/10.1016/j.trc.2015.02.018.

Allemang, R. J., and D. L. Brown. 1982. "A correlation coefficient for modal vector analysis." In Vol. 1 of *Proc., 1st Int. Modal Analysis Conf.*, 110–116. Orlando, FL: SEM.

Anjomshoaa, A., F. Duarte, D. Rennings, T. Matarazzo, P. de Souza, and C. Ratti. 2018. "City scanner: Building and scheduling a mobile sensing platform for smart city services." *IEEE Internet Things J.* 5 (6): 4567–4579. https://doi.org/10.1109/JIOT.2018.2839058.

Au, S.-K. 2011. "Fast Bayesian FFT method for ambient modal identification with separated modes." *J. Eng. Mech.* 137 (3): 214–226. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000213.

Barabasi, A.-L. 2005. "The origin of bursts and heavy tails in human dynamics." *Nature* 435 (7039): 207. https://doi.org/10.1038/nature03459.

Brincker, R., L. Zhang, and P. Andersen. 2001. "Modal identification of output-only systems using frequency domain decomposition." *Smart Mater. Struct.* 10 (3): 441. https://doi.org/10.1088/0964-1726/10/3/303.

Cai, J.-F., E. J. Candès, and Z. Shen. 2010. "A singular value thresholding algorithm for matrix completion." *SIAM J. Optim.* 20 (4): 1956–1982. https://doi.org/10.1137/080738970.

Candès, E. J., and B. Recht. 2009. "Exact matrix completion via convex optimization." *Found. Comput. Math.* 9 (6): 717. https://doi.org/10.1007/s10208-009-9045-5.

Chang, M., and S. N. Pakzad. 2014. "Observer Kalman filter identification for output-only systems using interactive structural modal identification toolsuite." *J. Bridge Eng.* 19 (5): 04014002. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000530.

Chopra, A. K. 2017. *Dynamics of structures theory and applications to earthquake engineering*. Hoboken, NJ: Pearson.

Dorvash, S., and S. Pakzad. 2013. "Stochastic iterative modal identification algorithm and application in wireless sensor networks." *Struct. Control Health Monit.* 20 (8): 1121–1137. https://doi.org/10.1002/stc.1521.

Eisen, M., A. Mokhtari, and A. Ribeiro. 2017. "Large scale empirical risk minimization via truncated adaptive newton method." Preprint, submitted May 22, 2017. http://arxiv.org/abs/1705.07957.

Eshkevari, S. S., and S. Pakzad. 2019. "Bridge structural identification using moving vehicle acceleration measurements." In Vol. 2 of *Dynamics of civil structures*, 251–261. Cham, Switzerland: Springer.

Eshkevari, S. S., and S. N. Pakzad. 2020. "Signal reconstruction from mobile sensors network using matrix completion approach." In Vol. 8 of *Topics in modal analysis and testing*, 61–75. Cham, Switzerland: Springer.

Eshkevari, S. S., T. J. Matarazzo, and S. N. Pakzad. 2020. "Bridge modal identification using acceleration measurements within moving vehicles." Preprint, submitted January 6, 2020. http://arxiv.org/abs/2001.01797.

Eshkevari, S. S., M. Takác, S. N. Pakzad, and S. S. Eshkevari. 2019. "High resolution bridge mode shape identification via matrix completion approach." *Struct. Health Monit.* https://doi.org/10.12783/shm2019/32499.

Gensun, F. 1996. "Whittaker–Kotelnikov–Shannon sampling theorem and aliasing error." *J. Approximation Theory* 85 (2): 115–131. https://doi.org/10.1006/jath.1996.0033.

González, A., E. J. OBrien, and P. McGetrick. 2012. "Identification of damping in a bridge using a moving instrumented vehicle." *J. Sound Vib.* 331 (18): 4115–4131. https://doi.org/10.1016/j.jsv.2012.04.019.

Gurney, K. R., et al. 2015. "Climate change: Track urban emissions on a human scale." *Nat. News* 525 (7568): 179. https://doi.org/10.1038/525179a.

Jain, P., P. Netrapalli, and S. Sanghavi. 2013. "Low-rank matrix completion using alternating minimization." In *Proc., 45th Annual ACM Symp. on Theory of Computing*, 665–674. New York: Association for Computing Machinery.

James, G., T. G. Carne, and J. P. Lauffer. 1995. "The natural excitation technique (NExT) for modal parameter extraction from operating structures." *Modal Anal. Int. J. Anal. Exp. Modal Anal.* 10 (4): 260.

Jerri, A. J. 1977. "The Shannon sampling theorem—Its various extensions and applications: A tutorial review." *Proc. IEEE.* 65 (11): 1565–1596. https://doi.org/10.1109/PROC.1977.10771.

Jolliffe, I. 2011. "Principal component analysis." In *International encyclopedia of statistical science*, 1094–1096. Cham, Switzerland: Springer.

Juang, J.-N., and R. S. Pappa. 1985. "An eigensystem realization algorithm for modal parameter identification and model reduction." *J. Guidance Control Dyn.* 8 (5): 620–627. https://doi.org/10.2514/3.20031.

Keenahan, J., E. J. OBrien, P. J. McGetrick, and A. Gonzalez. 2014. "The use of a dynamic truck–trailer drive-by system to monitor bridge damping." *Struct. Health Monit.* 13 (2): 143–157. https://doi.org/10.1177/1475921713513974.

Kleywegt, A., and K. Sinha. 1994. *Tools for bridge management data analysis*. Transportation Research Circular 423. Washington, DC: Transportation Research Board.

Kramer, B., and A. A. Gorodetsky. 2018. "System identification via cur-factored Hankel approximation." *SIAM J. Sci. Comput.* 40 (2): A848–A866. https://doi.org/10.1137/17M1137632.

Kramer, B., and S. Gugercin. 2016. "Tangential interpolation-based eigensystem realization algorithm for MIMO systems." *Math. Comput. Modell. Dyn. Syst.* 22 (4): 282–306. https://doi.org/10.1080/13873954.2016.1198389.

Krishnan, S. S., Z. Sun, A. Irfanoglu, S. J. Dyke, and G. Yan. 2011. "Evaluating the performance of distributed approaches for modal identification." In Vol. 7981 of *Sensors and smart structures technologies for civil, mechanical, and aerospace systems 2011*, 79814M. Bellingham, WA: International Society for Optics and Photonics.

Kurata, M., J. Kim, Y. Zhang, J. P. Lynch, G. Van Der Linden, V. Jacob, E. Thometz, P. Hipley, and L.-H. Sheng. 2011. "Long-term assessment of an autonomous wireless structural health monitoring system at the new Carquinez suspension bridge." In Vol. 7983 of *Nondestructive characterization for composite materials, aerospace engineering, civil infrastructure, and homeland security 2011*, 798312. Bellingham, WA: International Society for Optics and Photonics.

Li, W.-M., Z.-H. Jiang, T.-L. Wang, and H.-P. Zhu. 2014. "Optimization method based on generalized pattern search algorithm to identify bridge parameters indirectly by a passing vehicle." *J. Sound Vib.* 333 (2): 364–380. https://doi.org/10.1016/j.jsv.2013.08.021.

Lin, C., and Y. Yang. 2005. "Use of a passing vehicle to scan the fundamental bridge frequencies: An experimental verification." *Eng. Struct.* 27 (13): 1865–1878. https://doi.org/10.1016/j.engstruct.2005.06.016.

Lynch, J. P., and K. J. Loh. 2006. "A summary review of wireless sensors and sensor networks for structural health monitoring." *Shock Vib. Dig.* 38 (2): 91–130. https://doi.org/10.1177/0583102406061499.

Malekjafarian, A., P. J. McGetrick, and E. J. OBrien. 2015. "A review of indirect bridge monitoring using passing vehicles." *Shock Vib.* 2015: 286139. https://doi.org/10.1155/2015/286139.

Malekjafarian, A., and E. J. OBrien. 2014. "Identification of bridge mode shapes using short time frequency domain decomposition of the responses measured in a passing vehicle." *Eng. Struct.* 81 (Dec): 386–397. https://doi.org/10.1016/j.engstruct.2014.10.007.

Marulanda, J., J. M. Caicedo, and P. Thomson. 2017. "Modal identification using mobile sensors under ambient excitation." *J. Comput. Civ. Eng.* 31 (2): 04016051. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000619.

Massaro, E., C. Ahn, C. Ratti, P. Santi, R. Stahlmann, A. Lamprecht, M. Roehder, and M. Huber. 2017. "The car as an ambient sensing platform [point of view]." *Proc. IEEE.* 105 (1): 3–7. https://doi.org/10.1109/JPROC.2016.2634938.

Matarazzo, T. J., and S. N. Pakzad. 2016a. "Structural identification for mobile sensing with missing observations." *J. Eng. Mech.* 142 (5): 04016021. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001046.

Matarazzo, T. J., and S. N. Pakzad. 2016b. "Truncated physical model for dynamic sensor networks with applications in high-resolution mobile sensing and bigdata." *J. Eng. Mech.* 142 (5): 04016019. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001022.

Matarazzo, T. J., and S. N. Pakzad. 2018. "Scalable structural modal identification using dynamic sensor network data with STRIDEX." *Comput.-Aided Civ. Infrastruct. Eng.* 33 (1): 4–20. https://doi.org/10.1111/mice.12298.

Matarazzo, T. J., P. Santi, S. N. Pakzad, K. Carter, C. Ratti, B. Moaveni, C. Osgood, and N. Jacob. 2018. "Crowdsensing framework for monitoring bridge vibrations using moving smartphones." *Proc. IEEE.* 106 (4): 577–593. https://doi.org/10.1109/JPROC.2018.2808759.

Maymon, S., and A. V. Oppenheim. 2011. "Sinc interpolation of nonuniform samples." *IEEE Trans. Signal Process.* 59 (10): 4745–4758. https://doi.org/10.1109/TSP.2011.2160054.

McGetrick, P. J., C.-W. Kim, A. González, and E. J. Brien. 2015. "Experimental validation of a drive-by stiffness identification method for bridge monitoring." *Struct. Health Monit.* 14 (4): 317–331. https://doi.org/10.1177/1475921715578314.

Mei, Q., and M. Gül. 2018. "A crowdsourcing-based methodology using smartphones for bridge health monitoring." *Struct. Health Monit.* 18 (5–6): 1602–1619. https://doi.org/10.1177/1475921718815457.

Mei, Q., M. Gül, and M. Boay. 2019. "Indirect health monitoring of bridges using Mel-frequency cepstral coefficients and principal component analysis." *Mech. Syst. Sig. Process.* 119 (Mar): 523–546. https://doi.org/10.1016/j.ymssp.2018.10.006.

Moheimani, S. R., D. Halim, and A. J. Fleming. 2003. Vol. 10 of *Spatial control of vibration: Theory and experiments*. Singapore: World Scientific.

OBrien, E. J., and A. Malekjafarian. 2016. "A mode shape-based damage detection approach using laser measurement from a vehicle crossing a simply supported bridge." *Struct. Control Health Monit.* 23 (10): 1273–1286. https://doi.org/10.1002/stc.1841.

Pakzad, S. N., G. L. Fenves, S. Kim, and D. E. Culler. 2008. "Design and implementation of scalable wireless sensor network for structural monitoring." *J. Infrastruct. Syst.* 14 (1): 89–101. https://doi.org/10.1061/(ASCE)1076-0342(2008)14:1(89).

Pakzad, S. N., G. V. Rocha, and B. Yu. 2011. "Distributed modal identification using restricted auto regressive models." *Int. J. Syst. Sci.* 42 (9): 1473–1489. https://doi.org/10.1080/00207721.2011.563875.

Peeters, B., and G. De Roeck. 2001. "Stochastic system identification for operational modal analysis: A review." *J. Dyn. Syst. Meas. Contr.* 123 (4): 659–667. https://doi.org/10.1115/1.1410370.

Poncelet, F., G. Kerschen, J.-C. Golinval, and D. Verhelst. 2007. "Output-only modal analysis using blind source separation techniques." *Mech. Syst. Sig. Process.* 21 (6): 2335–2358. https://doi.org/10.1016/j.ymssp.2006.12.005.

Schanze, T. 1995. "Sinc interpolation of discrete periodic signals." *IEEE Trans. Signal Process.* 43 (6): 1502–1503. https://doi.org/10.1109/78.388863.

Smith, I. F. 2016. "Studies of sensor data interpretation for asset management of the built environment." *Front. Built Environ.* 2: 8. https://doi.org/10.3389/fbuil.2016.00008.

Sony, S., S. Laventure, and A. Sadhu. 2019. "A literature review of next-generation smart sensing technology in structural health monitoring." *Struct. Control Health Monit.* 26 (3): e2321. https://doi.org/10.1002/stc.2321.

Tachet, R., O. Sagarra, P. Santi, G. Resta, M. Szell, S. Strogatz, and C. Ratti. 2017. "Scaling law of urban ride sharing." *Sci. Rep.* 7: 42868. https://doi.org/10.1038/srep42868.

Tachet, R., P. Santi, S. Sobolevsky, L. I. Reyes-Castro, E. Frazzoli, D. Helbing, and C. Ratti. 2016. "Revisiting street intersections using slot-based systems." *PLoS One.* 11 (3): e0149607. https://doi.org/10.1371/journal.pone.0149607.

Vazifeh, M. M., P. Santi, G. Resta, S. Strogatz, and C. Ratti. 2018. "Addressing the minimum fleet problem in on-demand urban mobility." *Nature* 557 (7706): 534. https://doi.org/10.1038/s41586-018-0095-1.

Wang, H., F. Calabrese, G. Di Lorenzo, and C. Ratti. 2010. "Transportation mode inference from anonymized and aggregated mobile phone call detail records." In *Proc., 13th Int. IEEE Conf. on Intelligent Transportation Systems*, 318–323. New York: IEEE.

Wang, P., T. Hunter, A. M. Bayen, K. Schechtner, and M. C. González. 2012. "Understanding road usage patterns in urban areas." *Sci. Rep.* 2: 1001. https://doi.org/10.1038/srep01001.

Wilson, E. L., M.-W. Yuan, and J. M. Dickens. 1982. "Dynamic analysis by direct superposition of Ritz vectors." *Earthquake Eng. Struct. Dyn.* 10 (6): 813–821. https://doi.org/10.1002/eqe.4290100606.

Wright, S., and J. Nocedal. 1999. *Numerical optimization*. New York: Springer.

Yang, Y., and K. Chang. 2009. "Extracting the bridge frequencies indirectly from a passing vehicle: Parametric study." *Eng. Struct.* 31 (10): 2448–2459. https://doi.org/10.1016/j.engstruct.2009.06.001.

Yang, Y.-B., C. Lin, and J. Yau. 2004. "Extracting bridge frequencies from the dynamic response of a passing vehicle." *J. Sound Vib.* 272 (3–5): 471–493. https://doi.org/10.1016/S0022-460X(03)00378-X.

Zachariah, D., M. Sundin, M. Jansson, and S. Chatterjee. 2012. "Alternating least-squares for low-rank matrix reconstruction." *IEEE Signal Process Lett.* 19 (4): 231–234. https://doi.org/10.1109/LSP.2012.2188026.

Zhu, C., R. H. Byrd, P. Lu, and J. Nocedal. 1997. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization." *ACM Trans. Math. Software (TOMS).* 23 (4): 550–560. https://doi.org/10.1145/279232.279236.