Model updating in structural dynamics: advanced parametrization, optimal regularization, and symmetry considerations

Daniel Thomas Bartilson

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

This page intentionally left blank.

This page intentionally left blank.

ABSTRACT

Model updating in structural dynamics: advanced parametrization, optimal regularization, and symmetry considerations

Daniel Thomas Bartilson

Numerical models are pervasive tools in science and engineering for simulation, design, and assessment of physical systems. In structural engineering, finite element (FE) models are extensively used to predict responses and estimate risk for built structures. While FE models attempt to exactly replicate the physics of their corresponding structures, discrepancies always exist between measured and model output responses. Discrepancies are related to aleatoric uncertainties, such as measurement noise, and epistemic uncertainties, such as modeling errors. Epistemic uncertainties indicate that the FE model may not fully represent the built structure, greatly limiting its utility for simulation and structural assessment. Model updating is used to reduce error between measurement and model-output responses through adjustment of uncertain FE model parameters, typically using data from structural vibration studies. However, the model updating problem is often ill-posed with more unknown parameters than available data, such that parameters cannot be uniquely inferred from the data.

This dissertation focuses on two approaches to remedy ill-posedness in FE model updating: parametrization and regularization. Parametrization produces a reduced set of updating parameters to estimate, thereby improving posedness. An ideal parametrization should incorporate model uncertainties, effectively reduce errors, and use as few parameters as possible. This is a challenging task since a large number of candidate parametrizations are available in any model updating problem. To ameliorate this, three new parametrization techniques are proposed: improved parameter clustering with residual-based weighting, singular vector decomposition-based parametrization, and incremental reparametrization. All

of these methods utilize local system sensitivity information, providing effective reduced-order parametrizations which incorporate FE model uncertainties.

The other focus of this dissertation is regularization, which improves posedness by providing additional constraints on the updating problem, such as a minimum-norm parameter solution constraint. Optimal regularization is proposed for use in model updating to provide an optimal balance between residual reduction and parameter change minimization. This approach links computationally-efficient deterministic model updating with asymptotic Bayesian inference to provide regularization based on maximal model evidence. Estimates are also provided for uncertainties and model evidence, along with an interesting measure of parameter efficiency.

Table of Contents

Dedicated to Fryderyk Chopin,

whose music accompanied me through this work

Preface

This dissertation consists of an extended summary of finite element model updating and four appended papers which function as independent chapters detailing specific advancements. The appended papers were prepared in collaboration with the co-authors. The author of this dissertation was responsible for a majority of the work, including development of theories, numerical implementation, paper planning, and writing.

The presented work in this dissertation was carried out at the Department of Civil Engineering and Engineering Mechanics at Columbia University between the years of 2014 and 2018. I would like to thank the university and department for supporting me during this time, with particular acknowledgment of Scott Kelly for his kind assistance. This research was partially funded by the National Science Foundation under Grant No. CMMI-1563364.

I am deeply indebted to my advisor, Professor Andrew W. Smyth, for his candid advice and mentoring over the past four years. His agreeable nature and enthusiasm created an ideal environment for me to develop as a researcher. I would like to recognize Professors Raimondo Betti and Ioannis Kougioumtzoglou for their insights and pleasant interactions. I wish to acknowledge Professor Tian Zheng for her valuable suggestions. I am fortunate that these four individuals, along with Professor Richard Longman, constitute my dissertation committee and I wish to thank them for their consideration.

Dr. Jinwoo Jang is deserving of special recognition for supporting me in all of my investigations in model updating. I can not thank him enough for his persistent positivity and our productive discussions.

Finally, I wish to sincerely thank my family and friends for their companionship in my life. You have each been a constant presence in my mind and source of comfort, whether near or far.

This page intentionally left blank.

PART I

EXTENDED SUMMARY

This page intentionally left blank.

1 Introduction

Scientists and engineers rely on numerical models to simulate the behavior of real systems in nearly all applications. Broadly stated, simulation is valuable when it is uneconomical, impractical, or impossible to directly test the real system. Numerical models may be based on physical understanding, as a simplification of reality, or may be a black box, only attempting to reproduce an input-output relation.

In the context of structural engineering, numerical modeling is typically accomplished through finite element (FE) models which physically represent designed or constructed works. Simulation via FE models is instrumental for the tasks of analysis, design, and assessment. Analysis hinges on simulation to estimate the response of a given structure (e.g. bridge vibrations) without needing to risk damage or interfere with ongoing operations (e.g. traffic). Design is closely related to analysis, comparing candidate configurations based on simulation to select an optimal system which satisfies given requirements. Assessment attempts to glean insight regarding the current configuration of the structure, perhaps for locating damage, measuring degradation, or estimating reliability.

For all of these tasks, it is essential that the FE model accurately represents the structure of interest. Unfortunately, discrepancies always exist between measured behavior and model-predicted behavior, indicating that the model does not fully represent the structure and has lower utility for simulation. This is the motivation for FE model updating, which is the process of adjusting uncertain FE model parameters to reduce discrepancies between measured data and model-output data (Friswell & Mottershead, 1995). Model updating is essentially a parameter estimation problem, with a wealth of deterministic and uncertainty quantification (UQ) methods available in the literature, described further in Section 2.1.

Regardless of the parameter estimation scheme, FE model updating problems are often

ill-posed, meaning that a unique solution does not exist. Even when the problem is well-posed, it may be ill-conditioned such that a small change in measured data or initial model state drastically changes the updating result. Ameliorating ill-posedness and ill-conditioning is non-trivial. One approach is to provide additional assumptions to constrain the data, such as by regularization in deterministic methods (Titurus & Friswell, 2008) or an informative prior probability distribution for the parameters in probabilistic UQ methods (Berger, Bernardo, & Sun, 2009). Construction of a reasonable regularizer or prior probability distribution function (PDF) is challenging since little is often known about the true variability of uncertain updating parameters, particularly when they do not represent physical quantities.

Another approach to improve posedness and condition is to develop a new, generally reduced parametrization. This is a challenging task since there are a multitude of possible parametrizations even for a simple FE model and the choice of parametrization will impact the efficacy and posedness of the model updating problem. A variety of parametrization methods exist in the literature, but there is still significant room for improvement, particularly among the wealth of machine learning approaches coming to the forefront.

The research in this dissertation focuses on these last two challenges, regularization and parametrization. Part I continues with a firm background for FE model updating in Chapter 2. Relevant difficulties in model updating relating to regularization and parametrization are detailed in Chapter 3, while Chapter 4 describes the contribution and importance of the author's research. This is reinforced in Chapter 5 through summaries of the four papers appended in Part II. Conclusions are drawn in Chapter 6.

2 Background

This chapter provides an overview of FE model updating, covering deterministic and UQ approaches in Section 2.1, parametrization methods in Section 2.2, and model comparison methods in Section 2.3.

FE model updating is best understood as one stage of a process, comprising model generation, model updating, and model comparison, depicted in Fig. 2.1.

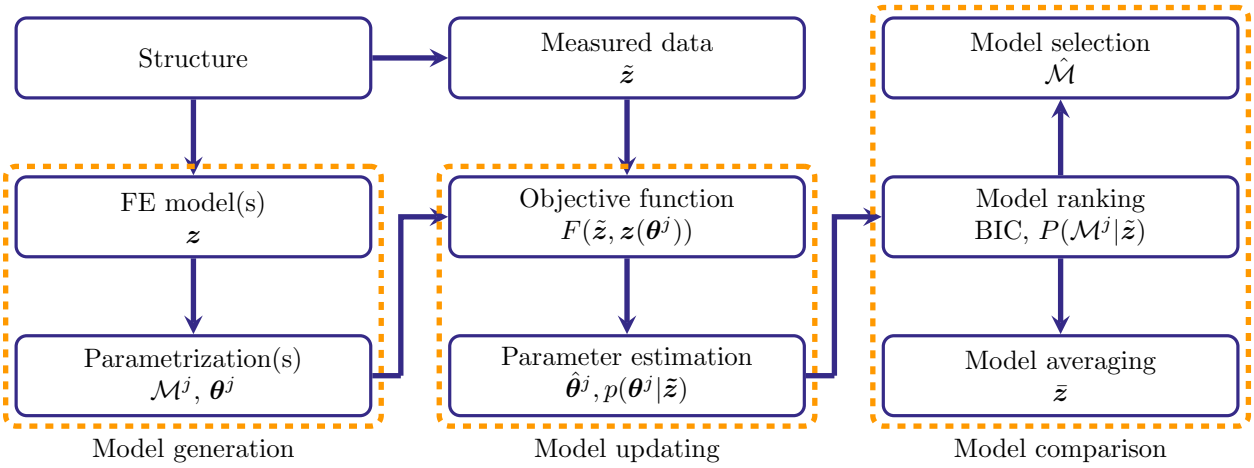


Figure 2.1: Flowchart of FE modeling, updating, and selection

– **Measured data** is one of the most critical items for model updating. It is often assembled into a vector of m components, denoted as $\tilde{\boldsymbol{z}}$, which can be homogeneous (comprising one data type) or heterogeneous (comprising several data types). When inputs (loadings) are measured, it may be practical to utilize time series of directly-measured quantities, such as strains, accelerations, and displacements, or inferred quantities such as response functions. However, inputs are typically unknown in practice, so inferred output quantities such as power spectral densities, or more commonly, natural frequencies and mode shapes, are used. Common methods in practice for estimating modal parameters from vi-

bration data are eigensystem realization algorithm (ERA) by Juang and Pappa (1985), observer/Kalman filter identification (OKID) by Juang, Phan, Horta, and Longman (1993), and frequency domain decomposition (FDD) by Brincker, Zhang, and Andersen (2000). A review of modern modal identification algorithms is available by Reynders (2012).

– **Model generation** comprises the two steps of FE modeling and parametrization. In general, there may be several candidate FE models with several tested parametrizations, but for now consider only one parametrized FE model denoted as the *model class* $\mathcal{M}$. This represents an input-output function between the vector of p updating parameters $\boldsymbol{\theta}$ and the model-outputs $\boldsymbol{z}(\boldsymbol{\theta})$. The term 'model class' is used since a given parametrization can be evaluated at varying parameter values. When the values of the parameters are fixed, say at the optimal value $\hat{\boldsymbol{\theta}}$, then the outputs $\boldsymbol{z}(\hat{\boldsymbol{\theta}})$ represent a *model* $\mathcal{M}(\hat{\boldsymbol{\theta}})$ for the measured data[1] (Beck & Yuen, 2004). Parametrization is covered in more depth in Section 2.2.

– **Model updating** refers to the process of parameter estimation for a specified model class, generally to produce point or distribution estimates of plausible parameter values. The choice of objective function is typically decided by the parameter estimation scheme, but almost always represents a goodness-of-fit measure between the measured data and model-outputs. The basic assumption of model updating is that the measured data can be reproduced by the FE model $\mathcal{M}$, but modeling errors and measurement errors create discrepancies:

$$\tilde{\boldsymbol{z}} = \boldsymbol{z}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon} + \boldsymbol{\eta} \tag{2.1}$$

where $\boldsymbol{\varepsilon}$ represents a vector of modeling errors and $\boldsymbol{\eta}$ represents a vector of measurement errors, such as noise. Measurement error is related to aleatoric uncertainty, meaning that it is inherent in the problem and cannot be reduced through improved modeling

[1]While the terms 'model' and 'model class' are more proper, it is typical to find model classes simply referred to as models within the literature and clarified by context.

(R. C. Smith, 2014). Modeling error relates to epistemic uncertainty[2], which comprises systematic errors arising from a lack of knowledge. For FE model updating, model errors can be categorized as (Mottershead, Link, & Friswell, 2011):

- idealization errors, related to assumptions and simplifications about system physics;
- discretization errors, introduced by poor construction of the FE model; and
- uncertainty in model parameters, such as material properties, section properties, geometry, and boundary conditions.

Unlike aleatoric uncertainties, epistemic uncertainties can be reduced through improved modeling. Therefore, the general goal of model updating is to estimate the parameters $\boldsymbol{\theta}$ which minimize the model error $\boldsymbol{\varepsilon}$ without fitting the measurement noise $\boldsymbol{\eta}$. FE model updating is typically formulated to correct FE model parameter errors, while idealization and discretization errors must be controlled by FE model construction. If FE model parameter uncertainties alone are minimized, the model is said to be *inconsistent*, but may still be valuable for response prediction. When all three error categories are minimized, the model is said to be *validated* and has greater utility for design and assessment (Mottershead et al., 2011). Model updating techniques are discussed in more detail in Section 2.1.

– The last stage, **model comparison**, is typically only performed when several candidate model classes are available. This relies on some measure of support for each model which can be used as a criterion for ranking the relative plausibility of each model $\mathcal{M}^j$ of a set $j = 1, \ldots, n_M$. A simple measure might be the model-output accuracy (maximum likelihood) criterion, but this tends to favor overfitted solutions and overly complex models (MacKay, 1992), so statistical inference is typically used. With an established ranking, it is possible to either perform *model selection* to choose a single best model $\hat{\mathcal{M}}$, or *model averaging* to produce a robust response estimate $\bar{\boldsymbol{z}}$. This is discussed more in Section 2.3.

[2]Aleatoric uncertainties represent intrinsic variability where the plausibility of a value may be described by a PDF, while epistemic uncertainties are assumed to have a single correct value which can be discovered through investigation.

2.1 Model updating

Modern model updating can be divided into deterministic and UQ approaches. Both of these methods are essentially parameter estimation schemes which attempt to solve an inverse problem by repeatedly running the forward problem. In this case, the forward problem provides FE model-output responses for a given set of parameter values, while the inverse problem asks for the point, range, or distribution of parameter values to reproduce the measured response. An excellent review is available by Simoen, De Roeck, and Lombaert (2015).

2.1.1 Deterministic approach

Deterministic model updating attempts to minimize the difference between the measured and model-output responses in Eq. (2.1) at a single optimal parameter estimate. This is achieved by minimizing an objective function F which utilizes the measurement and model-output data, written as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \; F\left(\tilde{\boldsymbol{z}}, \boldsymbol{z}(\boldsymbol{\theta})\right) \tag{2.2}$$

Typically the objective function is given as the weighted sum-of-square residual, where $\boldsymbol{r}(\boldsymbol{\theta})$ is the vector of m residuals:

$$F(\boldsymbol{\theta}) = \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} \qquad \boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \tag{2.3}$$

$\boldsymbol{W}_r$ is the $m \times m$ residual weighting matrix representing the uncertainty in the measurements. The optimal choice of $\boldsymbol{W}_r$, in a Bayesian sense, is the inverse of the measurement covariance matrix $\boldsymbol{C}_{\tilde{z}}$ (Collins, Hart, Haselman, & Kennedy, 1974). When covariance estimates are unavailable, $\boldsymbol{W}_r$ may instead reflect the relative importance of each residual term for reduction. Eq. (2.2) can also be written as the squared ℓ^2-norm of the weighted residual vector $\boldsymbol{q}(\boldsymbol{\theta})$:

$$F(\boldsymbol{\theta}) = \|\boldsymbol{q}(\boldsymbol{\theta})\|_2^2 \qquad \boldsymbol{q}(\boldsymbol{\theta}) = \boldsymbol{W}_r^{1/2} \boldsymbol{r}(\boldsymbol{\theta}) \tag{2.4}$$

Alternative objective functions in deterministic model updating include sums of individual ℓ^2-norms, e.g. Jang and Smyth (2017b), and the ℓ^1-norm, e.g. C. B. Smith and Hernandez (2018). In particular, use of the ℓ^1-norm may provide better detection of localized damage, but is typically more difficult to optimize.

With a given objective function, deterministic model updating then uses an optimization algorithm to produce the point estimate in Eq. (2.2). There are a wide variety of approaches available in the literature, naturally divided into global and local approaches. Global optimization methods attempt to find the best parameter estimate within the assumed range of parameter values and are typically used for non-convex problems or problems with multiple local minima. Common global approaches in model updating include genetic algorithms and simulated annealing (Levin & Lieven, 1998), coupled local minimizers (Teughels, De Roeck, & Suykens, 2003), and artificial bee colony algorithms (Kang, Li, & Xu, 2009). In most situations, global approaches are significantly more computationally-intensive than local approaches, which typically use numerical or analytical gradients to reach a local optimum. These approaches are relative fast, intuitive, and provide satisfactory convergence for FE model updating problems with 'small' initial errors.

– **The sensitivity method** (Mottershead et al., 2011) is one of the most common local approaches to deterministic model updating and is largely similar to Gauss–Newton minimization. This is an iterative approach where the parameter estimate is evolved from iteration i to $i+1$ until convergence by the update equation:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i \tag{2.5}$$

where $\boldsymbol{\theta}_i$ is the updating parameter vector at iteration i and $\Delta\boldsymbol{\theta}_i$ is the corresponding update term. While the residual vector $\boldsymbol{r}$ is usually non-linear with respect to the updating parameters $\boldsymbol{\theta}$, this is linearized at each iteration by truncating the Taylor series:

$$\boldsymbol{r}(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i) \approx \boldsymbol{r}(\boldsymbol{\theta}_i) + \boldsymbol{J}_i\Delta\boldsymbol{\theta}_i \tag{2.6}$$

$\boldsymbol{J}_i$ is the $m \times p$ Jacobian matrix of $\boldsymbol{r}$ with respect to $\boldsymbol{\theta}$ evaluated at $\boldsymbol{\theta}_i$:

$$\boldsymbol{J}_i = \left.\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_i} = -\left.\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_i} \tag{2.7}$$

which is also called the *sensitivity matrix* since it represents the local sensitivity of the residual (or equivalently, model-output) with respect to parameter changes. The parameter update is given as the Gauss–Newton solution

$$\Delta\boldsymbol{\theta}_i = -[\boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i]^{-1} \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i \tag{2.8}$$

where $\boldsymbol{r}_i$ is shorthand for $\boldsymbol{r}(\boldsymbol{\theta}_i)$. Obviously, the solution will not be unique when the number of measurements m is less than the number of updating parameters p (underdetermined), and $\boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i$ will not be invertible. Even when $m \geq p$, it is common for $\boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i$ to be ill-conditioned when the *effective* number of parameters exceeds the *effective* number of measurements, e.g. linear dependence in the columns of $\boldsymbol{J}_i$.

– **Regularization** is one common way to improve posedness and condition in deterministic FE model updating. This supplies additional constraints to the data, effectively increasing the number of measurements. An excellent review of regularization is provided by Titurus and Friswell (2008). A common choice is to penalize the sum-of-square parameter change, leading to a minimum-norm solution. For updating parameters that are originally zero (i.e. $\boldsymbol{\theta}_0 = \boldsymbol{0}$), this is accomplished by augmenting the original objective function with a penalty term:

$$F(\boldsymbol{\theta}) = \beta \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} + \alpha \boldsymbol{\theta}^T \boldsymbol{W}_\theta \boldsymbol{\theta} \tag{2.9}$$

where $\boldsymbol{W}_\theta$ is the parameter weighting term for parameters $\boldsymbol{\theta}$ and $\{\alpha, \beta\}$ are regularization terms. Similar to $\boldsymbol{W}_r$, the statistically-optimal choice of $\boldsymbol{W}_\theta$ is the inverse of the parameter covariance matrix $\boldsymbol{C}_\theta$ (Collins et al., 1974), but this is difficult to estimate in practice. Therefore, $\boldsymbol{W}_\theta$ often represents the relative importance (to the user) of reducing certain parameter changes. When $\boldsymbol{W}_\theta$ must be constructed without any information about the parameters, it is typical to use $\boldsymbol{W}_\theta = \boldsymbol{I}_p$, closely related to Tikhonov regular-

ization (Tikhonov & Arsenin, 1977). The solution to the regularized objective function in Eq. (2.9) is then given by

$$\Delta\boldsymbol{\theta}_i = -[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i + \alpha \boldsymbol{W}_\theta]^{-1} [\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i] \tag{2.10}$$

which essentially augments the solution to Eq. (2.8). The regularization parameters α and β control the relative importance of reducing residual against delivering a minimum parameter change solution. Popular approaches include L-curves and cross validation (Ahmadian, Mottershead, & Friswell, 1998; Titurus & Friswell, 2008) or simple heuristics (Mottershead et al., 2011). Ongoing challenges related to regularization are discussed in Chapter 3.

2.1.2 Uncertainty quantification approach

Where deterministic model updating seeks to produce a single solution to minimize the error between measurements and model-outputs in Eq. (2.1), UQ methods utilize uncertainty estimates to characterize a range or distribution of plausible parameter values. This is a key advantage over deterministic model updating since it enables users to evaluate the confidence in parameter estimates and response predictions, with further applications to model selection and reliability assessments. UQ model updating is comprised of two sub-classes: probabilistic and non-probabilistic methods (Simoen et al., 2015).

– **Probabilistic UQ methods** propagate uncertainties from input quantities (e.g. parameters and measurements) to output quantities (e.g. responses) by assigning PDFs to all quantities. Probabilistic methods include the popular Bayesian inference approach, first used in structural dynamics by Beck and Katafygiotis (1998), which uses Bayes' theorem to generate updated parameter estimates from initial parameter estimates and the measured data. In this approach, the measurement errors $\boldsymbol{\eta}$, modeling errors $\boldsymbol{\varepsilon}$ and uncertain parameters $\boldsymbol{\theta}$ are treated as random variables. The updated, or posterior, parameter

estimate is given by Bayes' expression:

$$p(\boldsymbol{\theta}|\tilde{\boldsymbol{z}}, \mathcal{M}) = \frac{p(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M})}{p(\tilde{\boldsymbol{z}}|\mathcal{M})} \tag{2.11}$$

The prior PDF for the parameters $p(\boldsymbol{\theta}|\mathcal{M})$ reflects the initial degree-of-belief in parameter values without knowledge of measured data. This is updated to the posterior PDF $p(\boldsymbol{\theta}|\tilde{\boldsymbol{z}}, \mathcal{M})$ by incorporating measured data $\tilde{\boldsymbol{z}}$ into parameter estimates. The transformation between prior and posterior estimates comes from the likelihood function $p(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \mathcal{M})$ which expresses the plausibility of observing the data $\tilde{\boldsymbol{z}}$ at a certain parameter value $\boldsymbol{\theta}$ within the parametrization $\mathcal{M}$. The denominator $p(\tilde{\boldsymbol{z}}|\mathcal{M})$ is termed the *evidence* for the model class $\mathcal{M}$, which is the integral of the numerator over all values:

$$p(\tilde{\boldsymbol{z}}|\mathcal{M}) = \int p(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M})d\boldsymbol{\theta} \tag{2.12}$$

Model evidence (technically model *class* evidence) represents the plausibility of the parametrization for describing the measured data, and is particularly important for Bayesian model comparison, discussed in Section 2.3.

The multidimensional integral for the model evidence can be challenging to compute. In situations where Gaussian assumptions can be made for measurement errors and the likelihood function, or in situations with a large amount of data relative to the number of parameters, it may be appropriate to assume the posterior will follow a Gaussian PDF. Thus, the evidence integral can be approximated using Laplace's method, considerably speeding up computation (Beck & Katafygiotis, 1998; MacKay, 1992). For more general posteriors (e.g. non-Gaussian, multimodal), it is typical to approximate the integral by a Monte Carlo (MC) method, typically Markov Chain Monte Carlo (MCMC), which attempts to draw samples from the posterior. Examples of MC methods used in model updating include Metropolis–Hastings MCMC (Beck & Au, 2002), transitional MCMC (Ching & Chen, 2007), hybrid MC (Cheung & Beck, 2009), nested sampling (Mthembu, Marwala, Friswell, & Adhikari, 2011), and approximate Bayesian computation (Vakilzadeh, 2016).

– **Non-probabilistic UQ methods** are a recent development for FE model updating and

generally express parameter, response, and measurement uncertainties as intervals corresponding to a certain confidence level. This is particularly appropriate when insufficient information is available to describe parameter uncertainties by a PDF and/or when few measurements are available to characterize measurement uncertainties (Khodaparast et al., 2017). Interval model updating was first applied by Khodaparast, Mottershead, and Badcock (2011) to establish bounds on the updating parameters corresponding to irreducible (aleatoric) uncertainty in measured data. Interval approaches have been extended to fuzzy modeling, which express the plausibility of a quantity (e.g. parameter value, response, measurement) through a membership function, performing uncertainty propagation from fuzzy measured data to fuzzy parameter estimates and responses. This has been applied to model updating by Haag, Herrmann, and Hanss (2010) and Erdogan, Gul, Catbas, and Bakir (2014), among others. These methods generally use large-scale optimization to find parameter bounds corresponding to plausible measurement uncertainty bounds, making them generally less computationally intensive than probabilistic UQ methods, but still orders-of-magnitude more intensive than deterministic model updating schemes.

2.2 Parametrization

While a parametrized FE model is a required quantity for model updating (Fig. 2.1), it is most natural to describe parametrization with an understanding of the limitations of parameter estimation schemes, namely posedness. Posedness is a property of the model updating problem, i.e. it occurs regardless of the parameter estimation scheme used. An ill-posed problem does not have a unique solution. This may be understood as a condition with insufficient measured information to constrain the parameter estimates, manifesting as an underdetermined system of equations in deterministic methods (Section 2.1.1) or unidentifiable parameters in UQ methods (Beck & Katafygiotis, 1998). Regularization is often used

for deterministic model updating, but generating an alternative parametrization is preferred (Simoen et al., 2015).

Even a relatively simple FE model may comprise a large number of parameters, such as element material and cross-sectional properties, connection and boundary condition models and properties, model geometry, etc., which may all be used as updating parameters. This may even be expanded if one considers the number of unique combinations of parameters which may be selected for updating. This must be contrasted against the effective number of measurements available in practice, which is limited by such factors as sampling frequency, sampling duration, measurement signal-to-noise ratio, excitation quality, and a finite number of sensors or practical arrangements. It is almost always possible to choose a parametrization with the number of parameters p much greater than the number of measurements m.

Parametrizing the FE model is non-trivial, with significant impact on the efficacy in reducing errors and posedness of the updating problem. A successful parametrization should satisfy three key conditions (Friswell & Mottershead, 1995):

- a limited number of parameters should be used to avoid ill-posedness,
- the parametrization should be chosen to correct model uncertainties, and
- model-outputs should be sensitive to the chosen parameters.

Fulfillment of these requirements generally requires expert knowledge of FE model sensitivities and uncertainties. Many approaches have been suggested in the literature. For example, Mottershead, Mares, Friswell, and James (2000) compared several parametrizations of an aluminum space frame, such as generic elements by Gladwell and Ahmadian (1995) and geometric joint offsets by Mottershead, Friswell, Ng, and Brandon (1996). A common parametrization is the set of FE model parameters $\boldsymbol{\delta}$ which linearly modify the element mass matrix $\boldsymbol{M}_e$ and element stiffness matrix $\boldsymbol{K}_e$ for each element e out of n_{el} elements. The global mass $\boldsymbol{M}$ and stiffness $\boldsymbol{K}$ matrices are thus functions of $\boldsymbol{\delta}$, which is a

column vector of $d = 2\,n_{\mathrm{el}}$ components:

$$\boldsymbol{M}(\boldsymbol{\delta}) = \sum_{e=1}^{n_{el}} \boldsymbol{M}_e(1-\delta_e^M) = \boldsymbol{M}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{M}_e\delta_e^M \tag{2.13}$$

$$\boldsymbol{K}(\boldsymbol{\delta}) = \sum_{e=1}^{n_{el}} \boldsymbol{K}_e(1-\delta_e^K) = \boldsymbol{K}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{K}_e\delta_e^K \tag{2.14}$$

where $\boldsymbol{M}_0$ and $\boldsymbol{K}_0$ are the unmodified mass and stiffness matrices, respectively. Obviously, this parametrization will still result in an ill-conditioned problem when $d > m$, so it is typical to select a reduced set of updating parameters based on this parametrization.

While many parametrizations may avail the user to lower residual, there is no guarantee that the model-outputs will be sensitive to the chosen parameters or that the problem will be well-posed. Parametrizations which incorporate model sensitivities therefore have a large advantage in that they are designed to provide sensitive parameters and counter ill-posedness. The two main sensitivity-based parametrization methods are subset selection and parameter clustering:

– **Subset selection** chooses a reduced number of FE model parameters to use as updating parameters, which is equivalent to fixing the unselected FE model parameters. This approach was originally used in regression analysis (Miller, 2002). For most problems, it is inappropriate or impossible to test all possible subsets of parameters, so heuristics are applied. The most common heuristic subset selection method is forward subset selection, first applied to FE model updating by Lallement and Piranda (1990) and used in further work by Friswell, Mottershead, and Ahmadian (1998, 2001). This method successively selects parameters based on the similarity of their sensitivity vectors to the residual vector $\boldsymbol{r}$ during an orthogonalization process. The sensitivity vector for each parameter δ_i corresponds to the columns of the sensitivity matrix, $\boldsymbol{J} = [\boldsymbol{j}_1 \cdots \boldsymbol{j}_d]$ from Eq. (2.7). The orthogonalization process ensures that each parameter has a different effect on reducing residual, which should improve posedness and condition by excluding parameters with linearly-dependent sensitivities. Silva, Maia, Link, and Mottershead (2016) provided an alternative method for selecting model updating parameters based on decomposition of

the sensitivity matrix, while Yuan, Liang, Silva, Yu, and Mottershead (2019) used a robust global sensitivity analysis for subset selection in model updating.

– **Parameter clustering** updates all of the FE model parameters $\boldsymbol{\delta}$, but creates a reduced parametrization by grouping FE model parameters with similar sensitivities into clusters which are each modified by a single updating parameter. This approach was proposed by Friswell et al. (1998) as an extension of subset selection, and was further developed by Shahverdi, Mares, Wang, and Mottershead (2009). Jang and Smyth (2017a, 2017b) used this approach to develop reduced parametrization for deterministic and UQ updating of a full-scale FE model. This approach typically uses hierarchical clustering to develop a dendrogram. At each step, the most similar two elements (clusters or single parameters), measured by cosine distance between sensitivity vectors, are combined to form a new cluster. This builds a dendrogram where 'branches' merge with increasing distance criterion. The user selects the number of clusters, perhaps based on a maximum distance criterion (e.g. ensure that parameters within the clusters have sensitivities within 0.01 cosine distance of each other). This defines groups of parameters which are then assigned a cluster updating parameter, reducing the number of updating parameters from d to the number of clusters. Since sensitivity-based parameter clustering combines FE model parameters with similar sensitivity vectors, linear dependence is reduced in the resulting Jacobian matrix, thereby improving posedness and condition. Alternative approaches for sensitivity-based clustering include information about the physical proximity of parameters, thereby creating physically contiguous clusters (Kim & Park, 2004, 2008).

2.2.1 Structural symmetry

For sensitivity-based model updating and sensitivity-based parametrization, it is important to understand structural properties to identify possible challenges in parametrization. In particular, symmetric structures have interesting vibration properties which may affect sensitivity-based parametrization methods. Many civil and mechanical structures have at

least one plane of symmetry for design and construction simplicity. A symmetric structure can be transformed to configurations which are identical to the original structure through a symmetry group, comprising a defined set of reflections and rotations (Zingoni, 2009). Glockner (1973) was the first to rigorously study of symmetry in structural dynamics, using group theory to form reduced representations of full structural systems. Notably, the mass and stiffness matrices of a symmetric structure can be decomposed into similar block-diagonal forms (Healey & Treacy, 1991). This greatly reduces the computational and memory requirements in vibration analysis, as the single, large eigendecomposition is transformed into several smaller, separable eigendecompositions. Group theory is the natural vehicle for establishing these transformations, and has seen extensive study in structural dynamics by Kaveh et al. (2003, 2007) and Zingoni (2009, 2014).

Structural symmetry discussed further in the context of challenges for parametrization in Chapter 3.

2.3 Model comparison

Model comparison (technically model *class* comparison) and model updating are best understood as two levels of statistical inference (MacKay, 1992). The first level of inference, model updating, implicitly assumes that the given model class $\mathcal{M}$ can correctly reproduce the measured data, i.e. Eq. (2.1), by inferring the correct parameter values in light of the measured data. The second level of inference, model comparison, seeks to assign a relative level of support for each model class $\mathcal{M}^j$ of a set of n_M candidate model classes $\{\mathcal{M}^1, \ldots, \mathcal{M}^{n_M}\}$. Within FE model updating, the candidate model classes correspond to one or more alternative FE models, each possibly having several parametrizations. Each parametrization of an FE model corresponds to a model class where the j^{th} parametrization uses a column vector of updating parameters $\boldsymbol{\theta}^j$ of size p^j.

In general, more complex models will provide better data fit than simpler models, but

they are more prone to *overfitting*. An overfitted model is less useful for future simulation because it is fitted to the specific measurement errors realized from aleatoric uncertainty, rather than capturing the general model errors associated with epistemic uncertainty (Beck & Yuen, 2004). To this end, model comparison should reflect *Occam's razor*, which states that a simpler description is preferable to an unnecessarily complicated one. For model selection, Occam's razor implies that the best model should exhibit strong agreement with the measured data with minimal complexity (Jeffreys, 1939).

Model comparison is typically performed using a Bayesian perspective, largely because this approach natively implements Occam's razor (Gull, 1988; MacKay, 1992), but also due to the prevalence of Bayesian UQ methods. The probability of the model class $\mathcal{M}^j$ given the data $\tilde{\boldsymbol{z}}$ is given by the Bayes expression:

$$P(\mathcal{M}^j|\tilde{\boldsymbol{z}}) = \frac{p(\tilde{\boldsymbol{z}}|\mathcal{M}^j)P(\mathcal{M}^j)}{\sum_{k=1}^{n_M} p(\tilde{\boldsymbol{z}}|\mathcal{M}^k)P(\mathcal{M}^k)} \tag{2.15}$$

where $P(\mathcal{M}^j)$ represents the prior probability of $\mathcal{M}^j$ out of the set of candidate models before observation of the data. The model evidence $p(\tilde{\boldsymbol{z}}|\mathcal{M}^j)$ was described in Eq. (2.12) and indicates the plausibility that the measured data was generated from the proposed model. Importantly, when all models are assumed to have equal prior probability (i.e. $P(\mathcal{M}^j) = 1/n_M$), then the probability of the model class is directly proportional to the model evidence:

$$P(\mathcal{M}^j|\tilde{\boldsymbol{z}}) \propto p(\tilde{\boldsymbol{z}}|\mathcal{M}^j) \tag{2.16}$$

Occam's razor can be discerned in the evidence expression by rewriting the log of Eq. (2.12) as the difference of two terms (Muto & Beck, 2008):

$$\begin{aligned} \log p(\tilde{\boldsymbol{z}}|\mathcal{M}^j) = &\int \log\left[p(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}^j, \mathcal{M}^j)\right] \, p(\boldsymbol{\theta}^j|\tilde{\boldsymbol{z}}, \mathcal{M}^j) d\boldsymbol{\theta}^j \\ &- \int \log\left[\frac{p(\boldsymbol{\theta}^j|\tilde{\boldsymbol{z}}, \mathcal{M}^j)}{p(\boldsymbol{\theta}^j|\mathcal{M}^j)}\right] p(\boldsymbol{\theta}^j|\tilde{\boldsymbol{z}}, \mathcal{M}^j) d\boldsymbol{\theta}^j \end{aligned} \tag{2.17}$$

The first term is the posterior mean of the log likelihood function, representing the average data fit provided by the model class $\mathcal{M}^j$. The second term is the relative entropy between the prior and posterior distributions for the parameters, measuring the information extracted

from the data during updating. This effectively functions as a penalty term which enforces Occam's razor since more complex parametrizations will extract additional information from the data to infer their larger number of parameters (Muto & Beck, 2008).

Bayesian model comparison naturally leads to model selection and model averaging (Beck & Yuen, 2004). In Bayesian model selection, the model class with the highest probability given the data $\hat{\mathcal{M}}$ is selected for further use. This is appropriate when one model class greatly outperforms the other candidates. In cases where several candidate models are assessed at similar posterior probability, it is recommended to use Bayesian model averaging to provide a robust aggregate response estimate. This is given by the theorem of total probability:

$$\bar{\boldsymbol{z}} = \mathbb{E}(\boldsymbol{z}|\tilde{\boldsymbol{z}}) = \sum_{j=1}^{n_M} \mathbb{E}(\boldsymbol{z}|\tilde{\boldsymbol{z}}, \mathcal{M}^j) P(\mathcal{M}^j|\tilde{\boldsymbol{z}}) \tag{2.18}$$

where $\mathbb{E}(\ )$ represents the expectation operator.

While Bayesian model comparison has many natural advantages and is commonly utilized, there are alternatives for use with deterministic model updating and non-probabilistic UQ methods, which generally cannot estimate model evidence. Deterministic model updating often makes use of Akiake's information criterion or AIC (Akaike, 1974), and the Bayesian information criterion or BIC (Schwarz, 1978). These measures assess the data fit of the each model class at the maximum likelihood result, but add penalty terms related to the number of model parameters, thereby attempting to enforce Occam's razor. Such methods avoid the costly computation of model evidence, but are generally only appropriate with large amounts of data compared to the number of model parameters (Kass & Raftery, 1995). While model selection in non-probabilistic UQ is underdeveloped, an example is available in Haag, González, and Hanss (2012), which compared models on the basis of combined uncertainty in model-outputs and identified model parameters.

This page intentionally left blank.

3 Research challenges

The work presented in this dissertation focuses on parametrization and regularization in FE model updating, both of which have significant ramifications on the efficiency, posedness, and condition of the resulting model updating problem. Ameliorating these difficulties enhances FE model fidelity, providing more reliable FE models for simulating structural responses, optimizing designs, and assessing structural changes. Contemporary challenges associated with parametrization and regularization are outlined below:

Parametrization

Model updating seeks to infer the best parameter values for a given model such that epistemic uncertainties (related to model errors) are reduced without fitting to the aleatoric uncertainties (such as measurement noise). In this process of statistical inference, it is often the case that there are more parameters to estimate than can be constrained by available measurements, leading to an ill-posed problem without a unique solution. At this stage, alternative parametrizations are desired which (1) give well-posed updating problems, (2) have some relevance to the uncertainties in the FE model, and (3) are capable of reducing model-output errors (Friswell & Mottershead, 1995). As discussed in Section 2.2, this is a challenging task with a large number of alternatives to assess.

It is generally possible to satisfy criterion (1) by setting the number of updating parameters p well below the number of measurements m. This may not be practical when a limited number of measurements are used (e.g. only natural frequencies), but schemes such as subset selection and parameter clustering can feasibly be used to generate extremely small parametrizations. Criterion (2) is generally enforced by user knowledge of existing uncertainties. In FE model updating, most kinds of epistemic uncertainty (e.g. structural damage

and deterioration, uncertain boundary conditions, uncertain joint behavior) are assumed to behave like errors in the global stiffness matrix $\boldsymbol{K}$. Therefore, the selected parametrization can sufficiently represent FE model uncertainties in most settings if the parametrization is based on $\boldsymbol{\delta}$, as in Eq. (2.14). Even derived parametrizations like subset selection and parameter clustering can be argued to reflect FE model uncertainties since their updating parameters are still based on $\boldsymbol{\delta}$.

Criterion (3) is related to the performance of a parametrization. In some settings, it may be appropriate to compare a large number of candidate parametrizations. However, in many practical cases, model updating is a very computationally-intensive process with long compute times for forward runs of full-scale FE models. Thus, it is desired to generate parametrizations which provide some *a priori* guarantee of effectiveness. The most natural vehicle for assessing the effectiveness of a parametrization prior to use is sensitivity analysis. This could be sensitivity analysis of candidate parametrizations, e.g. Mottershead et al. (2000). However, a better approach is to design parametrizations with respect to sensitivity, such as subset selection and parameter clustering. Such methods guarantee local, if not global, effectiveness of the selected parameters for reducing residual, and often have mechanisms for improving posedness and condition by eliminating linear dependence in the sensitivity matrix. Development of efficient parametrization schemes for improved performance and posedness is an ongoing challenge in model updating.

Symmetric structures (discussed in Section 2.2.1) are a special case deserving of further analysis for sensitivity-based parametrization. In one example, parameter clustering based on natural frequency sensitivity for a symmetric structure produced symmetric clusters (Jang & Smyth, 2017b). Recent work on the same structure using natural frequency and mode shape sensitivities led to asymmetric clusters (Bartilson, Jang, & Smyth, 2019), indicating that symmetric structures exhibit interesting sensitivity properties. Subset selection also stands to benefit from improved understanding of symmetry, since parameters are chosen based on their sensitivity vectors and an orthogonalization process will prevent two parameters from

being chosen which have the same sensitivities (Friswell et al., 1998).

Regularization

While probabilistic UQ methods are popular in the literature for their ability to assess confidence in estimates and perform model selection, they are often intractable for full-scale FE model updating due to previously-mentioned computational cost. Instead, it may be more attractive to use deterministic model updating, particularly approaches like the sensitivity method (Mottershead et al., 2011) with much lower computational demand. For deterministic methods, it is common to utilize some form of regularization for guaranteeing a well-posed and well-conditioned inverse problem. Even when condition is guaranteed through parametrization, it may still be advantageous to implement regularization since it can equalize the assumptions behind competing parametrizations.

As discussed in Section 2.1.1, a key difficulty is the determination of regularization parameters α and β in Eq. (2.9). These control the relative importance of reducing residual against the side constraint of providing a minimum norm solution. Ideally, regularization should have a minimal impact on the solution ($\beta \gg \alpha$) while still maintaining condition requirements. This has typically been done with L-curves by plotting the two objectives at various ratios of α/β to determine an optimal value, or by cross validation (Ahmadian et al., 1998; Titurus & Friswell, 2008). However, these processes can be very computationally-intensive since they require running a full model optimization at several test values for α/β. Other suggested methods are simple heuristics, e.g. $\alpha/\beta = 0.05$ (Mottershead et al., 2011), but such approaches will be sub-optimal for most problems since they fail to account for any differences in model uncertainty or measurement uncertainty.

As shown by MacKay (1992) and discussed in the context of model updating by Simoen et al. (2015), regularization in deterministic model updating can be related to the prior probability distribution for the parameters in Bayesian inference. Therefore, optimal regularization should reflect the underlying uncertainty in the parameters. This is not a simple

task, however, since there is typically little *a priori* information about the variability of FE model parameters. In fact, expressing the uncertainty in FE model parameters as a 'variability' is misleading, since parameters are taken to be epistemic with one 'true' value. This is one of the key critiques of Bayesian model updating (Simoen et al., 2015) and prompts the difficulty of assigning a PDF to FE model parameters which are non-probabilistic. Updating parameters are typically derived from FE model parameters, and assigning uncertainties to these quantities is challenging. When updating parameters are non-physical, e.g. generic elements by Gladwell and Ahmadian (1995), variability estimates must be made with essentially no prior knowledge.

4 Dissertation contributions

The presented work in this dissertation addresses the research challenges described in Chapter 3. By advancing the state-of-the-art within FE model updating, the proposed methods seek to provide updated FE models which accurately reproduce structural responses through well-posed model updating problems. The main contributions are outlined below:

Parametrization

An improved sensitivity-based parameter clustering scheme is described in **Paper A** which coherently combines heterogeneous data sources during cluster analysis. The proposed scheme builds upon previous sensitivity-based clustering methods which used natural frequency sensitivities, e.g. Shahverdi et al. (2009) and Jang and Smyth (2017a, 2017b), to also include mode shape sensitivities during cluster analysis. Furthermore, this method produces clusters which are consistent with the objective of reducing weighted residual, directly addressing a key requirement for a successful parametrization.

A novel parametrization approach is proposed in **Paper B** based on singular value decomposition (SVD) of the sensitivity matrix. This extends subset selection and parameter clustering methods to update linear combinations of FE model parameters by a condensed set of modifier terms. Two different paradigms are proposed for selecting the singular vectors, one which provides maximal improvement in condition, while another uses the logic of subset selection to best approximate the residual. This produces reduced parametrizations with enhanced local efficiency and improved posedness.

An extensible framework for extracting further utility from sensitivity-based parametrizations is presented in **Paper C**. This approach uses successive increments of sensitivity-based parametrization and model updating, enabling further reductions in error without increas-

ing the number of updating parameters. This non-intrusively improves efficiency without restricting posedness.

Paper D provides an in-depth study of the modal sensitivities for symmetric structures, addressing the nature of sensitivity-based clusters noted in Jang and Smyth (2017b), **Paper A**, and **Paper B**. This paper provides conclusive proof that natural frequency sensitivities are equal between symmetric parameter pairs, but mode shape sensitivities are unequal. This has significant ramifications for effective parametrization and improving posedness in updating symmetric structures.

Regularization

An optimal regularization scheme is proposed in **Paper A** which uses asymptotic Bayesian inference to relate regularization in deterministic model updating to the prior PDF in probabilistic methods (MacKay, 1992). The optimal regularizing constants are determined to maximize the model evidence, satisfying Occam's razor, while producing asymptotic estimates for measurement noise and parameter uncertainty. While FE model parameter uncertainties are still treated as aleatoric quantities, more informed estimates for the underlying uncertainties are developed through optimization. This approach estimates the model evidence, combining the computational efficiency of deterministic methods and the model comparison utility of UQ methods. Additionally, this approach naturally defines an 'effective number of parameters', which is an interesting quantity for assessing the efficiency of a parametrization. Regularization constants are determined during parameter identification, avoiding the computational cost of cross validation and L-curve methods. This approach is modified to provide more robust model evidence estimates in **Paper B** and **Paper C**.

5 Summary of appended papers

Paper A: Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization

Sensitivity-based parameter clustering presents an attractive reduced-order FE model parametrization which improves posedness and condition while retaining the physical relevance of the uncertain FE model parameters. This paper proposes an augmented parameter clustering method which weights the parameter sensitivities according to the objective function. This produces clusters which are consistent with the goal of reducing the weighted residual between measured and model-output data and can incorporate disparate data sources, such as natural frequency and mode shape sensitivity. This paper also implements optimal regularization in deterministic model updating which relates the regularizing constants to Bayesian prior PDFs and likelihood functions. The optimal regularizing constants are determined by maximizing the model evidence, simultaneously providing an estimate for the effective number of updating parameters. These two improvements are tested in model updating of a small-scale example with several sets of simulated data and a full-scale suspension bridge with real data. In both cases, the proposed parameter clustering scheme displayed moderate-to-strong efficiency improvements over existing parameter clustering methods.

Paper B: Sensitivity-based SVD parametrization and optimal regularization in finite element model updating

Analysis of the sensitivity matrix enables the development of FE model parametrizations which are (at least locally) effective in reducing residual with a minimal number of parameters. In this paper, SVD of the sensitivity matrix is used to develop a novel parametrization method for modifying linear combinations of FE model parameters. Two different SVD-

based parametrizations are described, one which gives a minimum-order representation of the original sensitivity matrix, focusing on condition, and another which best represents the measurement residual, focusing on efficiency. This paper also implements a modified form of Bayesian regularization to constrain parameter changes while estimating model evidence in a deterministic scheme. The proposed approach is tested against subset selection parametrization for model updating in an extensive study of a small-scale truss with analytical data and a full-scale study of a suspension bridge with real data. In both cases, the two SVD-based parametrizations display results on-par or better than subset selection when updating natural frequencies and mode shapes. The support for competing parametrizations is gauged by Bayes factor, providing robust inference.

Paper C: Incremental reparametrization in sensitivity-based subset selection and parameter clustering for finite element model updating

Sensitivity-based parametrization methods, such as subset selection and parameter clustering, are generally based on the local sensitivity matrix evaluated at the initial point. However, after updating this parametrization, a new parametrization may be more efficient due to changes in parameter sensitivities. In this paper, incremental reparametrization is proposed as a non-intrusive, extensible framework for updating an FE model using a new sensitivity-based parametrization at each increment. Importantly, this approach allows users to extract further improvement from an FE model without sacrificing posedness because each increment uses a reduced set of updating parameters. Deterministic model updating is implemented with a modified Bayesian regularization term to consistently penalize new parametrizations. The performance of the proposed reparametrization scheme is tested against subset selection and parameter clustering based only on the initial sensitivity matrix. Support for reparametrization is assessed via Bayes factor in model updating exercises of a small-scale truss with analytical data and a benchmark study with experimental data. Subset selection exhibits decisive support for reparametrization in both exercises, while the

benefits for parameter clustering are less pronounced.

Paper D: Symmetry properties of natural frequency and mode shape sensitivities in symmetric structures

The sensitivities of FE model-outputs to parameter changes are fundamental to sensitivity-based model updating and sensitivity-based parametrization methods. In particular, symmetric structures have been shown in previous work to produce symmetric sensitivity-based parameter clusters (Bartilson et al., 2019; Jang & Smyth, 2017b), indicating unique sensitivity behavior. This paper provides analytical proof that natural frequency sensitivities are equal between symmetric parameter pairs (e.g. the stiffnesses of symmetrically-placed elements), while mode shape sensitivities are necessarily unequal. The degree of inequality for mode shape sensitivities is explored in a small numerical study. The results are discussed with respect to parameter clustering, subset selection, and sensitivity-based model updating for symmetric structures.

This page intentionally left blank.

6 Conclusions

The work in this dissertation explored FE model updating by detailing challenges and proposing solutions in the topics of parametrization and regularization. FE model updating was approached from the perspective of statistical inference, attempting to estimate the 'correct' model parameters which improve correspondence between measured and model-output results. An updated FE model has enhanced simulation value, enabling users to accurately simulate responses, design changes, and damage scenarios.

A core challenge is that model updating problems are commonly ill-posed, such that a unique solution does not exist to the parameter inference problem, typically because there are more parameters to identify than data available to constrain the result. Parametrization and regularization are two approaches for improving posedness in model updating. A major challenge is producing alternative parametrizations and regularization which reflect underlying uncertainties and can efficiently improve simulation quality with a minimum number of updating parameters, thereby satisfying Occam's razor.

An effective method to evaluate the condition and effectiveness of a parametrization is through sensitivity analysis. In this vein, three new sensitivity-based parametrizations were proposed:

– The first method extends existing parameter clustering methods to reflect the weighted residual within the objective function. This approach also naturally incorporates heterogeneous data in cluster analysis, providing significantly better results than contemporary clustering methods when updating both natural frequencies and mode shapes.

– The second method utilizes SVD of the sensitivity matrix to parametrize using linear combinations of FE model parameters. This provides desirable properties, such as minimal representation of the original sensitivity matrix (to improve posedness) or maximal

effectiveness in reducing the residual (to improve effectiveness).

– The third method serves as a non-intrusive framework for delivering further error reductions with existing sensitivity-based parametrization methods. This approach utilizes the change in the sensitivity matrix after each increment of model updating to develop new, more efficient parametrizations without expanding the number of updating parameters.

These parametrizations were compared against contemporary methods on the basis of parameter efficiency, effectiveness in reducing residual, and relative support via model comparison. In all cases, the proposed methods were shown to increase effectiveness and efficiency without sacrificing posedness or physical meaning. In addition, analytical sensitivities were explored for symmetric structures, providing key insight into modal sensitivities for symmetric parameter pairs. The ramifications of the derived results were discussed for sensitivity-based parametrization and model updating of symmetric structures.

Optimal Bayesian regularization was also proposed in this dissertation. This approach relates regularization to Bayesian inference via an asymptotic assumption, delivering the computational efficiency of deterministic methods with the UQ benefits of probabilistic methods, such as model evidence estimates. Robust regularization is achieved by estimating model and measurement uncertainties. The proposed technique also assesses the local efficiency of the parametrization with applications to model comparison. The low computational cost ameliorates many of the difficulties associated with other regularization schemes, making this approach an attractive alternative for fast, approximate studies in practical FE model updating with applications to UQ and model comparison.

References

Ahmadian, H., Mottershead, J. E., & Friswell, M. I. (1998). Regularisation methods for finite element model updating. *Mechanical Systems and Signal Processing*, *12*(1), 47–64. https://doi.org/10.1006/mssp.1996.0133

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Bartilson, D. T., Jang, J., & Smyth, A. W. (2019). Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization. *Mechanical Systems and Signal Processing*, *114*, 328–345. https://doi.org/10.1016/j.ymssp.2018.05.024

Beck, J. L., & Au, S.-K. (2002). Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *Journal of Engineering Mechanics*, *128*(4), 380–391. https://doi.org/10.1061/(ASCE)0733-9399(2002)128:4(380)

Beck, J. L., & Katafygiotis, L. S. (1998). Updating models and their uncertainties I: Bayesian statistical framework. *Journal of Engineering Mechanics*, *124*(4), 455–461. https://doi.org/10.1061/(ASCE)0733-9399(1998)124:4(455)

Beck, J. L., & Yuen, K.-V. (2004). Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics*, *130*(2), 192–203. https://doi.org/10.1061/(ASCE)0733-9399(2004)130:2(192)

Berger, J. O., Bernardo, J. M., & Sun, D. (2009). The formal definition of reference priors. *The Annals of Statistics*, *37*(2), 905–938. https://doi.org/10.1214/07-AOS587

Brincker, R., Zhang, L., & Andersen, P. (2000). Modal identification from ambient responses using frequency domain decomposition. In *Proceedings of the 18th International Modal Analysis Conference.*

Cheung, S. H., & Beck, J. L. (2009). Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters. *Journal of Engineering Mechanics*, *135*(4), 243–255. https://doi.org/10.1061/(ASCE)0733-9399(2009)135:4(243)

Ching, J., & Chen, Y.-C. (2007). Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *Journal of Engineering Mechanics*, *133*(7), 816–832. https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816)

Collins, J. D., Hart, G. C., Haselman, T. K., & Kennedy, B. (1974). Statistical identification of structures. *AIAA Journal*, *12*(2), 185–190. https://doi.org/10.2514/3.49190

Erdogan, Y. S., Gul, M., Catbas, F. N., & Bakir, P. G. (2014). Investigation of uncertainty changes in model outputs for finite-element model updating using structural health monitoring data. *Journal of Structural Engineering*, *140*(11), 04014078. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001002

Friswell, M. I., & Mottershead, J. E. (1995). *Finite Element Model Updating in Structural Dynamics*. Dordrecht: Springer. https://doi.org/10.1007/978-94-015-8508-8

Friswell, M. I., Mottershead, J. E., & Ahmadian, H. (1998). Combining subset selection and parameter constraints in model updating. *Journal of Vibration and Acoustics*, *120*(4), 854–859. https://doi.org/10.1115/1.2893911

Friswell, M. I., Mottershead, J. E., & Ahmadian, H. (2001). Finite–element model updating using experimental test data: parametrization and regularization. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *359*(1778), 169–186. https://doi.org/10.1098/rsta.2000.0719

Gladwell, G. M. L., & Ahmadian, H. (1995). Generic element matrices suitable for finite element model updating. *Mechanical Systems and Signal Processing*, *9*(6), 601–614. https://doi.org/10.1006/mssp.1995.0045

Glockner, P. G. (1973). Symmetry in structural mechanics. *Journal of the Structural Division*, *99*(1), 71–89.

Gull, S. F. (1988). Bayesian inductive inference and maximum entropy. In *Maximum-entropy and Bayesian methods in science and engineering* (Vol. 1, pp. 53–74). Dordrecht: Kluwer.

Haag, T., González, S. C., & Hanss, M. (2012). Model validation and selection based on inverse fuzzy arithmetic. *Mechanical Systems and Signal Processing*, *32*, 116–134. https://doi.org/10.1016/j.ymssp.2011.09.028

Haag, T., Herrmann, J., & Hanss, M. (2010). Identification procedure for epistemic uncertainties using inverse fuzzy arithmetic. *Mechanical Systems and Signal Processing*, *24*(7), 2021–2034. https://doi.org/10.1016/j.ymssp.2010.05.010

Healey, T. J., & Treacy, J. A. (1991). Exact block diagonalization of large eigenvalue problems for structures with symmetry. *International Journal for Numerical Methods in Engineering*, *31*(2), 265–285. https://doi.org/10.1002/nme.1620310205

Jang, J., & Smyth, A. W. (2017a). Bayesian model updating of a full-scale finite element model with sensitivity-based clustering. *Structural Control and Health Monitoring*, *24*(11), e2004. https://doi.org/10.1002/stc.2004

Jang, J., & Smyth, A. W. (2017b). Model updating of a full-scale FE model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters. *Mechanical Systems and Signal Processing*, *83*, 337–355. https://doi.org/10.1016/j.ymssp.2016.06.018

Jeffreys, H. (1939). *Theory of Probability*. Oxford: Oxford University Press.

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

Juang, J.-N., Phan, M., Horta, L. G., & Longman, R. W. (1993). Identification of observer/Kalman filter Markov parameters: theory and experiments. *Journal of Guidance, Control, and Dynamics*, *16*(2), 320–329. https://doi.org/10.2514/3.21006

Kang, F., Li, J., & Xu, Q. (2009). Structural inverse analysis by hybrid simplex artificial bee colony algorithms. *Computers & Structures*, *87*(13-14), 861–870. https://doi.org/10.1016/j.compstruc.2009.03.001

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kaveh, A., & Nikbakht, M. (2007). Decomposition of symmetric mass–spring vibrating systems using groups, graphs and linear algebra. *Communications in Numerical Methods in Engineering*, *23*(7), 639–664. https://doi.org/10.1002/cnm.913

Kaveh, A., & Sayarinejad, M. A. (2003). Eigensolutions for matrices of special structures. *Communications in Numerical Methods in Engineering*, *19*, 125–136. https://doi.org/10.1002/cnm.576

Khodaparast, H. H., Govers, Y., Dayyani, I., Adhikari, S., Link, M., Friswell, M., ... Sienz, J. (2017). Fuzzy finite element model updating of the DLR AIRMOD test structure. *Applied Mathematical Modelling*, *52*, 512–526. https://doi.org/10.1016/j.apm.2017.08.001

Khodaparast, H. H., Mottershead, J. E., & Badcock, K. J. (2011). Interval model updating with irreducible uncertainty using the Kriging predictor. *Mechanical Systems and Signal Processing*, *25*(4), 1204–1226. https://doi.org/10.1016/j.ymssp.2010.10.009

Kim, G.-H., & Park, Y.-S. (2004). An improved updating parameter selection method and finite element model update using multiobjective optimisation technique. *Mechanical Systems and Signal Processing*, *18*(1), 59–78. https://doi.org/10.1016/S0888-3270(03)00042-6

Kim, G.-H., & Park, Y.-S. (2008). An automated parameter selection procedure for finite-element model updating and its applications. *Journal of Sound and Vibration*, *309*(3-5), 778–793. https://doi.org/10.1016/j.jsv.2007.07.076

Lallement, G., & Piranda, J. (1990). Localization methods for parametric updating of finite element models in elastodynamics. In *Proceedings of the 8th International Modal Analysis Conference* (pp. 579–585).

Levin, R. I., & Lieven, N. A. J. (1998). Dynamic finite element model updating using simulated annealing and genetic algorithms. *Mechanical Systems and Signal Processing*, *12*(1), 91–120. https://doi.org/10.1006/mssp.1996.0136

MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation*, *4*(3), 415–447. https://doi.org/10.1162/neco.1992.4.3.415

Miller, A. (2002). *Subset Selection in Regression*. London: Chapman & Hall.

Mottershead, J. E., Friswell, M. I., Ng, G. H. T., & Brandon, J. A. (1996). Geometric parameters for finite element model updating of joints and constraints. *Mechanical Systems and Signal Processing*, *10*(2), 171–182. https://doi.org/10.1006/mssp.1996.0012

Mottershead, J. E., Link, M., & Friswell, M. I. (2011). The sensitivity method in finite element model updating: a tutorial. *Mechanical Systems and Signal Processing*, *25*(7), 2275–2296. https://doi.org/10.1016/j.ymssp.2010.10.012

Mottershead, J. E., Mares, C., Friswell, M. I., & James, S. (2000). Selection and updating of parameters for an aluminium space-frame model. *Mechanical Systems and Signal Processing*, *14*(6), 923–944. https://doi.org/10.1006/mssp.2000.1303

Mthembu, L., Marwala, T., Friswell, M. I., & Adhikari, S. (2011). Model selection in finite element model updating using the Bayesian evidence statistic. *Mechanical Systems and Signal Processing*, *25*(7), 2399–2412. https://doi.org/10.1016/j.ymssp.2011.04.001

Muto, M., & Beck, J. L. (2008). Bayesian updating and model class selection for hysteretic structural models using stochastic simulation. *Journal of Vibration and Control*, *14*(1-2), 7–34. https://doi.org/10.1177/1077546307079400

Reynders, E. (2012). System identification methods for (operational) modal analysis: review and comparison. *Archives of Computational Methods in Engineering*, *19*(1), 51–124. https://doi.org/10.1007/s11831-012-9069-x

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shahverdi, H., Mares, C., Wang, W., & Mottershead, J. E. (2009). Clustering of parameter sensitivities: examples from a helicopter airframe model updating exercise. *Shock and Vibration*, *16*(1), 75–87. https://doi.org/10.3233/SAV-2009-0455

Silva, T. A. N., Maia, N. M. M., Link, M., & Mottershead, J. E. (2016). Parameter selection and covariance updating. *Mechanical Systems and Signal Processing*, *70*, 269–283. https://doi.org/10.1016/j.ymssp.2015.08.034

Simoen, E., De Roeck, G., & Lombaert, G. (2015). Dealing with uncertainty in model updating for damage assessment: a review. *Mechanical Systems and Signal Processing*, *56–57*, 123–149. https://doi.org/10.1016/j.ymssp.2014.11.001

Smith, C. B., & Hernandez, E. M. (2018). Detection of spatially sparse damage using impulse response sensitivity and LASSO regularization. *Inverse Problems in Science and Engineering*. https://doi.org/10.1080/17415977.2018.1434776

Smith, R. C. (2014). *Uncertainty Quantification: Theory, Implementation, and Applications*. Philadelphia, PA: SIAM.

Teughels, A., De Roeck, G., & Suykens, J. A. (2003). Global optimization by coupled local minimizers and its application to FE model updating. *Computers & Structures*, *81*(24-25), 2337–2351. https://doi.org/10.1016/S0045-7949(03)00313-4

Tikhonov, A. N., & Arsenin, V. I. (1977). *Solutions of ill-posed problems* (Vol. 14). V. H. Winston and Sons (distributed by Wiley, New York).

Titurus, B., & Friswell, M. I. (2008). Regularization in model updating. *International Journal for Numerical Methods in Engineering*, *75*(4), 440–478. https://doi.org/10.1002/nme.2257

Vakilzadeh, M. K. (2016). *Stochastic model updating and model selection: with application to structural dynamics* (PhD thesis). Chalmers University of Technology, Göteborg, Sweden.

Yuan, Z., Liang, P., Silva, T., Yu, K., & Mottershead, J. E. (2019). Parameter selection for model updating with global sensitivity analysis. *Mechanical Systems and Signal Processing*, *115*, 483–496. https://doi.org/10.1016/j.ymssp.2018.05.048

Zingoni, A. (2009). Group-theoretic exploitations of symmetry in computational solid and structural mechanics. *International Journal for Numerical Methods in Engineering*, *79*(3), 253–289. https://doi.org/10.1002/nme.2576

Zingoni, A. (2014). Group-theoretic insights on the vibration of symmetric structures in engineering. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *372*(2008). https://doi.org/10.1098/rsta.2012.0037

This page intentionally left blank.

PART II

APPENDED PAPERS A–D

This page intentionally left blank.

Paper A

Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization

---

Bibliographic information:

Bartilson, D. T., Jang, J., & Smyth, A. W. (2019). Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization. *Mechanical Systems and Signal Processing, 114*, 328–345. https://doi.org/10.1016/j.ymssp.2018.05.024

This page intentionally left blank.

Mechanical Systems and Signal Processing 114 (2019) 328–345

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization

Daniel T. Bartilson [a,*], Jinwoo Jang [b], Andrew W. Smyth [a]

[a] Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027, USA
[b] Department of Civil, Environmental & Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA

ARTICLE INFO

Article history:
Received 18 December 2017
Received in revised form 29 April 2018
Accepted 12 May 2018
Available online 25 May 2018

Keywords:
Finite element model updating
Sensitivity-based model updating
Sensitivity-based clustering
Parametrization
Regularization

ABSTRACT

Finite element model updating seeks to modify a structural model to reduce discrepancies between predicted and measured data, often from vibration studies. An updated model provides more accurate prediction of structural behavior in future analyses. Sensitivity-based parameter clustering and regularization are two techniques used to improve model updating solutions, particularly for high-dimensional parameter spaces and ill-posed updating problems. In this paper, a novel parameter clustering scheme is proposed which considers the structure of the objective function to facilitate simultaneous updating of disparate data, such as natural frequencies and mode shapes. Levenberg–Marquardt minimization with Bayesian regularization is also implemented, providing an optimal regularized solution and insight into parametrization efficiency. In a small-scale updating example with simulated data, the proposed clustering scheme is shown to provide moderate to excellent improvement over existing parameter clustering methods, depending on the accuracy of initial model. A full-scale updating example on a large suspension bridge shows similar improvement using the proposed parametrization scheme.

1. Introduction

Modern structural analysis generally depends on finite element (FE) models to predict dynamic behavior and understand the current state of a system. Though these models are often developed from detailed design drawings, discrepancies always exist between measured (observed) and model-output behavior [1]. Typical sources of discrepancy are model idealization, FE discretization errors, and uncertain model parameters such as material properties, section properties, geometry, and boundary conditions [1,2]. Discrepancies indicate that a model cannot reliably predict the behavior of its corresponding physical structure, limiting the utility of the model for future analysis.

Model updating is the process which seeks to reduce discrepancies between measured data and model-output data by adjusting parameters of an FE model [1–3]. Model updating has been successfully applied to a wide variety of aerospace, mechanical, and civil structures. Examples include a helicopter airframe [2,4], an aluminum space-frame [5], a prestressed

* Corresponding author.
E-mail addresses: dtb2121@columbia.edu (D.T. Bartilson), jangj@fau.edu (J. Jang), smyth@civil.columbia.edu (A.W. Smyth).
URL: http://www.columbia.edu/cu/civileng/smyth/ (A.W. Smyth).

https://doi.org/10.1016/j.ymssp.2018.05.024
0888-3270/

single-span highway bridge [6], a prestressed multi-span highway bridge [7], a concrete-filled steel tubular arch bridge [8], an actively-damped high-rise structure [9], and a residential reinforced concrete frame [10].

Model updating techniques may be divided into two categories: uncertainty quantification (UQ) methods and deterministic methods [11]. UQ methods incorporate measurement and model uncertainties in their solutions and can be grouped into probabilistic and non-probabilistic approaches. Probabilistic UQ methods estimate probability distributions functions for parameters and model outputs through repeated sampling in the parameter space. The most common non-probabilistic UQ method is fuzzy model updating, which uses optimization to estimate intervals for parameters and outputs corresponding to upper and lower bounds of measured data. However, both probabilistic and non-probabilistic UQ methods are orders-of-magnitude more computationally expensive than deterministic methods. An excellent review of UQ model updating can be found in [11].

Deterministic model updating produces a unique optimal solution and typically involves iterative adjustment of FE model parameters [3]. Of course, as these schemes generally involve minimizing a non-linear function, they are possibly subject to convergence problems. Among iterative methods, the sensitivity method [2] is one of the most intuitive and popular techniques for model updating. The sensitivity method approaches model updating as a non-linear least-squares minimization problem which is solved by iterations of linear approximations. The objective function is a sum of squared differences between measured and model-output data, making it easy to incorporate various data. The use of linear approximations also makes this method physically-intuitive and efficient, as the Jacobian matrix is directly relatable to model parameter sensitivities. However, the sensitivity method is often applied to ill-posed model updating problems, necessitating a reduction in the number of updating parameters and/or the inclusion of side-constraints in order to reach a unique, stable solution.

Shahverdi et al. [4] presented sensitivity-based parameter clustering as a viable method for reducing the number of updating parameters. By observing the sensitivities of model outputs to changes in model parameters, sensitivity-based parameter clustering generates clusters of model parameters which have similar effects on targeted model outputs. Then, each cluster of model parameters is updated by a single parameter. This gives a reduced-order model, generally with a better-conditioned Jacobian, while retaining the physical relevance of clustered model parameters. This technique was successfully applied to the updating of a helicopter airframe [4]. Jang and Smyth [12,13] applied this method for the updating of a large-scale suspension bridge.

Regularization is another technique used to solve ill-posed and noisy problems which often occur in FE model updating [2,14–16]. Generally, regularization adds equations which help constrain the updating solution. This can help produce a unique solution to an underdetermined problem (fewer measurements than parameters), though this situation should be avoided. Regularization is often used to give a minimum-norm solution, but it may also be used to enforce user-specified constraints between parameters [2].

While sensitivity-based parameter clustering is very promising, it is difficult to utilize disparate sources of data, such as natural frequencies and mode shapes, due to differences in scale. Previous work with parameter clustering only used one type of data [2,4], or used only natural frequency sensitivities for clustering despite the inclusion of mode shapes in the objective function [12,13]. To alleviate scaling issues during parameter clustering, it is necessary to develop a weighting technique which is efficient and reflective of the problem structure. The presented research details an objective-consistent weighting technique based on the residual. This paper also implements Bayesian regularization [17,18] in model updating, which gives a statistically optimal regularized solution. Bayesian regularization also provides insight into the effective number of updating parameters, which is used to explore the efficiency of competing parametrizations.

The paper begins with the definition of residual between measurements and corresponding model outputs, along with analytical sensitivities of model outputs to model parameters (Section 2). Model parametrization, clustering, and the objective-consistent weighting scheme are discussed in Section 3. The Levenberg–Marquardt minimization method, with the accompanying Bayesian regularization technique, are detailed in Section 4. Two model updating exercises are then performed to exhibit the efficiency of the objective-consistent clustering scheme for simultaneous updating of natural frequency and mode shape data. The first exercise uses a small-scale 2-dimensional truss with simulated measurements (Section 5), while the second uses a full-scale large suspension bridge with real data (Section 6). The findings are then discussed and concluding remarks are made in Section 7.

2. Residual definition and analytical sensitivity of model parameters

The sensitivity method for FE model updating [2] begins with the definition of a discrepancy, or residual, to be minimized by modifying a set of updating parameters. Traditionally, the residual $\boldsymbol{r}$ is defined as the difference between the column vector of m measured outputs $\tilde{\boldsymbol{z}}$ and the column vector of m analytical model outputs $\boldsymbol{z}(\boldsymbol{\theta})$ which is a function of the p updating parameters $\boldsymbol{\theta}$. The relationship between $\boldsymbol{r}$ and $\boldsymbol{\theta}$ is generally non-linear, but can be linearized by truncating the Taylor series after the linear term:

$$\boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \approx \boldsymbol{r}(\boldsymbol{\theta}_i) + \boldsymbol{J}_i(\boldsymbol{\theta} - \boldsymbol{\theta}_i) \tag{1}$$

At iteration i, $\boldsymbol{\theta}_i$ is the updating parameter vector and $\boldsymbol{J}_i \in \mathbb{R}^{m\times p}$ is the Jacobian matrix of $\boldsymbol{r}$ with respect to $\boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta}_i$:

$$\boldsymbol{J}_i = \left.\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_i} \tag{2}$$

The evaluation of $\boldsymbol{J}$, also known as the sensitivity matrix, forms the basis of Gauss–Newton updating methods [19]. Note that $\boldsymbol{J}$ is the sensitivity matrix for the residual, while other authors analyzed the sensitivity matrix for the model outputs, $\partial \boldsymbol{z}/\partial \boldsymbol{\theta}$ [1,4,12]. These sensitivity matrices are often equal but opposite in sign, which has no effect in sensitivity-based clustering methods. $\boldsymbol{J}$ is more general since the residuals, not the model-outputs, are being linearized.

In any model updating problem, one must choose data to set as a target for updating. Common choices are natural frequencies or eigenvalues, mode shapes, and frequency-response functions. For the purposes of this paper, the chosen residual is a concatenation of the natural frequency residual vector $\boldsymbol{r}_f$ (Section 2.1) and the mode shape residual vector $\boldsymbol{r}_s$ (Section 2.2), leading to $\boldsymbol{r} = [\boldsymbol{r}_f^T \;\; \boldsymbol{r}_s^T]^T$. The weighted sum-of-squared residual E_r can be written

$$E_r = \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_f^T & \boldsymbol{r}_s^T \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_r^f & \\ & \boldsymbol{W}_r^s \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_f \\ \boldsymbol{r}_s \end{bmatrix} \tag{3}$$

$$= \underbrace{\boldsymbol{r}_f^T \boldsymbol{W}_r^f \boldsymbol{r}_f}_{E_r^f} + \underbrace{\boldsymbol{r}_s^T \boldsymbol{W}_r^s \boldsymbol{r}_s}_{E_r^s} \tag{4}$$

which includes a term for the weighted sum-of-squared frequency residual (E_r^f) and a term for the weighted sum-of-squared mode shape residual (E_r^s). $\boldsymbol{W}_r$ is the residual weighting matrix, which is discussed in more detail in Section 4, and reflects the importance of each residual term.

The natural frequencies and mode shapes are assumed to be obtained from a model of an undamped structure, giving only real frequencies and mode shapes. A generic FE model of an undamped structure with N degrees-of-freedom (DoFs) consists of a stiffness matrix $\boldsymbol{K}$ and a mass matrix $\boldsymbol{M}$, where $\boldsymbol{K}, \boldsymbol{M} \in \mathbb{R}^{N \times N}$. These matrices are used to solve the generalized eigenvalue problem $\boldsymbol{K}\boldsymbol{\phi}_j = \omega_j^2 \boldsymbol{M}\boldsymbol{\phi}_j \quad \forall j = 1, \ldots, N$, where ω_j represents the j^{th} angular natural frequency (in units of rad/s) with an equivalent natural frequency f_j (in units of Hz). Each corresponding mode shape $\boldsymbol{\phi}_j$ is mass-normalized such that $\boldsymbol{\phi}_j^T \boldsymbol{M} \boldsymbol{\phi}_j = 1$.

For computing $\boldsymbol{J}$, note that each column is a concatenation of frequency residual gradients and mode shape residual gradients, i.e. $\partial \boldsymbol{r}/\partial \theta_k = [(\partial \boldsymbol{r}_f/\partial \theta_k)^T \; (\partial \boldsymbol{r}_s/\partial \theta_k)^T]^T$. Analytical methods for calculating the sensitivities are given in Sections 2.1 and 2.2. Alternatively (or for the purposes of verification), sensitivities can be estimated numerically using finite differences.

2.1. Undamped frequency residual

The frequency residual vector $\boldsymbol{r}_f$ is defined as the difference between the column vector of l measured natural frequencies $\tilde{\boldsymbol{f}}$ and the corresponding column vector of l model-output natural frequencies $\boldsymbol{f}(\boldsymbol{\theta})$:

$$\boldsymbol{r}_f = \tilde{\boldsymbol{z}}_f - \boldsymbol{z}_f(\boldsymbol{\theta}) = \tilde{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{\theta}) \tag{5}$$

It is essential to perform mode pairing [3] to ensure that measured and model-output data refer to the same modes when calculating residual. In this work, the mode pairing process proceeded in order through the measured modes, selecting the model-output mode which maximized the Modal Assurance Criterion (MAC) [20], or equivalently, minimized the angle between measured and model-output mode shapes. Model-output modes which were already paired were excluded from future pairing calculations.

When $\boldsymbol{W}_r^f$ is diagonal, the weighted sum-of-squared frequency residual E_r^f can be simplified using Eq. (5)

$$E_r^f = \boldsymbol{r}_f^T \boldsymbol{W}_r^f \boldsymbol{r}_f = \sum_{j=1}^{l} w_{rj}^f \left(\tilde{f}_j - f_j(\boldsymbol{\theta})\right)^2 \tag{6}$$

The frequency residual sensitivity vector (for use in $\boldsymbol{J}$) is

$$\frac{\partial \boldsymbol{r}_f}{\partial \theta_k} = -\frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \theta_k} \tag{7}$$

The sensitivity of model natural frequency f_j to a change in updating parameter θ_k can be calculated analytically using the results of [21,22]

$$2\pi \frac{\partial f_j}{\partial \theta_k} = \frac{\partial \omega_j}{\partial \theta_k} = \frac{1}{2\omega_j} \boldsymbol{\phi}_j^T \left[\frac{\partial \boldsymbol{K}}{\partial \theta_k} - \omega_j^2 \frac{\partial \boldsymbol{M}}{\partial \theta_k}\right] \boldsymbol{\phi}_j \tag{8}$$

where $\boldsymbol{\phi}_j$ is the j^{th} mass-normalized mode shape, i.e. $\boldsymbol{\phi}_j^T \boldsymbol{M} \boldsymbol{\phi}_j = 1$.

2.2. Undamped mode shape residual

The mode shape residual $\boldsymbol{r}_s$ is defined as the difference between the concatenated set of l measured mode shapes $\tilde{\boldsymbol{z}}_s$ and the corresponding concatenated model-output mode shapes $\boldsymbol{z}_s(\boldsymbol{\theta})$:

$$\boldsymbol{r}_s = \tilde{\boldsymbol{z}}_s - \boldsymbol{z}_s(\boldsymbol{\theta}) \tag{9}$$

The concatenated mode shape measurement vector is written as $\tilde{\boldsymbol{z}}_s = [\tilde{\boldsymbol{v}}_1^T \ \cdots \ \tilde{\boldsymbol{v}}_l^T]^T$ where $\tilde{\boldsymbol{v}}_j$ is the j^{th} unit-normalized measured mode shape, i.e. $\tilde{\boldsymbol{v}}_j = \tilde{\boldsymbol{\phi}}_j / (\tilde{\boldsymbol{\phi}}_j^T \tilde{\boldsymbol{\phi}}_j)^{1/2}$. Similarly, the concatenated model-output mode shapes are $\boldsymbol{z}_s(\boldsymbol{\theta}) = [\mu_1 \boldsymbol{v}_1^T(\boldsymbol{\theta}) \ \cdots \ \mu_l \boldsymbol{v}_l^T(\boldsymbol{\theta})]^T$, where $\boldsymbol{v}_j(\boldsymbol{\theta})$ is the j^{th} unit-normalized model-output mode shape. The model-output modes should be sorted according to the mode pairing results. μ_j is the modal scale factor between corresponding measured and model-output mode shapes, and minimizes the difference between $\tilde{\boldsymbol{v}}_j$ and $\boldsymbol{v}_j(\boldsymbol{\theta})$ in the least-squares sense [20]:

$$\mu_j = \tilde{\boldsymbol{v}}_j^T \boldsymbol{v}_j(\boldsymbol{\theta}) \tag{10}$$

Note that $\tilde{\boldsymbol{v}}_j$ and $\boldsymbol{v}_j(\boldsymbol{\theta})$ must be measured at the same n points on the physical and model structures. This means that $\boldsymbol{v}_j(\boldsymbol{\theta}), \tilde{\boldsymbol{v}}_j \in \mathbb{R}^n \quad \forall j = 1, \ldots, l$ and thus $\boldsymbol{r}_s \in \mathbb{R}^{nl}$.

Assuming that $\boldsymbol{W}_r^s$ is diagonal and can be decomposed into a scalar multiple of $\boldsymbol{I}_n$ for each mode ($[\boldsymbol{W}_r^s]_j = w_{rj}^s \boldsymbol{I}_n$), then the sum-of-squared mode shape residual E_r^s can be written

$$E_r^s = \boldsymbol{r}_s^T \boldsymbol{W}_r^s \boldsymbol{r}_s = \sum_{j=1}^{l} w_{rj}^s \boldsymbol{r}_{sj}^T \boldsymbol{r}_{sj} = \sum_{j=1}^{l} w_{rj}^s \| \tilde{\boldsymbol{v}}_j - \mu_j \boldsymbol{v}_j(\boldsymbol{\theta}) \|_2^2 \tag{11}$$

where $\boldsymbol{r}_{sj} = \tilde{\boldsymbol{v}}_j - \mu_j \boldsymbol{v}_j(\boldsymbol{\theta})$ is the residual for mode shape j and $\| \ \|_2$ is the L^2 norm. The MAC [20] between $\tilde{\boldsymbol{v}}_j$ and $\boldsymbol{v}_j(\boldsymbol{\theta})$ is defined as

$$\text{MAC}(\tilde{\boldsymbol{v}}_j, \boldsymbol{v}_j) = \frac{(\tilde{\boldsymbol{v}}_j^T \boldsymbol{v}_j)^2}{\tilde{\boldsymbol{v}}_j^T \tilde{\boldsymbol{v}}_j \cdot \boldsymbol{v}_j^T \boldsymbol{v}_j} \tag{12}$$

In the context of this paper, 'MAC' refers to 'crossMAC' between measured and model-output mode shapes, while 'auto-MAC' refers exclusively to MAC between the same set of modes (e.g. measured mode shapes). Use of Eq. (12) allows Eq. (11) to be rewritten as

$$E_r^s = \sum_{j=1}^{l} w_{rj}^s [1 - \text{MAC}(\tilde{\boldsymbol{v}}_j, \boldsymbol{v}_j(\boldsymbol{\theta}))] \tag{13}$$

The derivative of the residual for mode shape j, $\boldsymbol{r}_{sj}$ (for use in $\boldsymbol{J}$), is

$$\frac{\partial \boldsymbol{r}_{sj}}{\partial \theta_k} = -\frac{\partial (\mu_j \boldsymbol{v}_j)}{\partial \boldsymbol{v}_j} \frac{\partial \boldsymbol{v}_j}{\partial \theta_k} = -\left[\boldsymbol{v}_j \tilde{\boldsymbol{v}}_j^T + \mu_j \boldsymbol{I} \right] \frac{\partial \boldsymbol{v}_j}{\partial \theta_k} \tag{14}$$

The residual sensitivities for each mode are concatenated, such that $\partial \boldsymbol{r}_s / \partial \theta_k = [(\partial \boldsymbol{r}_{s1} / \partial \theta_k)^T \ \cdots \ (\partial \boldsymbol{r}_{sl} / \partial \theta_k)^T]^T$, where the sensitivity of the j^{th} mode shape can be analytically calculated by the results of [21,22]:

$$\frac{\partial \boldsymbol{v}_j}{\partial \theta_k} = \sum_{h=1}^{H} \frac{a_{jkh}}{\boldsymbol{v}_h^T \boldsymbol{M} \boldsymbol{v}_h} \boldsymbol{v}_h \tag{15}$$

The upper limit of summation H is the number of dynamic modes used to estimate the sensitivities, where $H \leqslant N$. When $H = N$, the results are exact. The factors a_{jkh} are given by

$$a_{jkh} = \frac{1}{\omega_j^2 - \omega_h^2} \boldsymbol{v}_h^T \left[\frac{\partial \boldsymbol{K}}{\partial \theta_k} - \omega_j^2 \frac{\partial \boldsymbol{M}}{\partial \theta_k} \right] \boldsymbol{v}_j; \quad j \neq h \tag{16}$$

$$a_{jkj} = -\frac{1}{2} \boldsymbol{v}_j^T \left[\frac{\partial \boldsymbol{M}}{\partial \theta_k} \right] \boldsymbol{v}_j \tag{17}$$

Combining the results of Eqs. (6) and (13), the weighted sum-of-squared residual from Eq. (3) is

$$E_r = E_r^f + E_r^s = \sum_{j=1}^{l} w_{rj}^f (\tilde{f}_j - f_j(\boldsymbol{\theta}))^2 + \sum_{j=1}^{l} w_{rj}^s [1 - \text{MAC}(\tilde{\boldsymbol{v}}_j, \boldsymbol{v}_j(\boldsymbol{\theta}))] \tag{18}$$

Note that when $\boldsymbol{W}_r$ is equal to the reciprocal of the covariance of $\tilde{\boldsymbol{z}}$ (for uncorrelated measurements), then Eq. (18) implies that each frequency has its own measurement variance, while the components of each mode shape vector have equal variance to other components of the same mode shape.

3. Selection of updating parameters

3.1. Model parametrization

When updating an FE model, the overarching approach is to modify a set of d model parameters $\boldsymbol{x}$ to reduce error. Following the parametrization in [12,13], the j^{th} element of the modified physical parameter vector $\boldsymbol{x}$ is defined by:

$$x_j = x_j^0(1-\delta_j) \tag{19}$$

where x_j^0 is the initial value of x_j and δ_j is the physical parameter modification term. Since $\boldsymbol{x}$ may contain parameters which differ by several orders of magnitude, such as stiffnesses and mass densities, updating $\boldsymbol{x}$ directly may result in a poorly-scaled Jacobian matrix. The use of $\boldsymbol{\delta}$ results in comparably-sized updating parameters and improved condition of $\boldsymbol{J}$. Perhaps the most natural implementation of this parametrization in an FE model is to modify each element or substructure e (out of a total number n_{el}) stiffness ($\boldsymbol{K}_e$) and mass ($\boldsymbol{M}_e$) matrices prior to summation into the global stiffness ($\boldsymbol{K}$) and mass ($\boldsymbol{M}$) matrices:

$$\boldsymbol{K} = \sum_{e=1}^{n_{el}} \boldsymbol{K}_e(1-\delta_e^k) = \boldsymbol{K}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{K}_e\delta_e^k \tag{20}$$

$$\boldsymbol{M} = \sum_{e=1}^{n_{el}} \boldsymbol{M}_e(1-\delta_e^m) = \boldsymbol{M}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{M}_e\delta_e^m \tag{21}$$

$\boldsymbol{K}_0$ and $\boldsymbol{M}_0$ are the initial global stiffness and mass matrices prior to modification or updating. δ_e^k and δ_e^m are the stiffness and mass physical parameter modifications, respectively, for element or substructure e. Note that $\boldsymbol{\delta}$ for this problem may be viewed as a concatenation of $\boldsymbol{\delta}^k$ and $\boldsymbol{\delta}^m$, therefore $\boldsymbol{\delta}$ has $d = 2n_{el}$ components. In this form, calculating partial derivatives for use in analytical sensitivity calculations is simple, as $\partial\boldsymbol{K}/\partial\delta_e^k = -\boldsymbol{K}_e$ and $\partial\boldsymbol{M}/\partial\delta_e^m = -\boldsymbol{M}_e$.

While FE model updating based on $\boldsymbol{\delta}$ is natural and straightforward, it can quickly become intractable or ill-posed in an FE model with thousands of elements, possibly with several parameters per element. Thus, a smaller set of updating parameters are sought through cluster analysis. The set of FE model parameters $\boldsymbol{\delta}$ are not directly updated, but instead $\boldsymbol{\theta}$ are updated and $\boldsymbol{\delta} = g(\boldsymbol{\theta})$. In general, it is desired that the clustering is "hard" (i.e. each physical parameter belongs to one and only one cluster). Thus, g may be written as a linear transformation:

$$\boldsymbol{\delta} = g(\boldsymbol{\theta}) = \boldsymbol{C}\boldsymbol{\theta} \tag{22}$$

where $\boldsymbol{C} \in \mathbb{R}^{d\times p}$ such that

$$C_{jk} = \mathbb{I}_k(j) = \begin{cases} 1 & \delta_j \text{ is included in cluster } k, \text{ updated by } \theta_k \\ 0 & \text{else} \end{cases} \tag{23}$$

in which $\mathbb{I}_k(j)$ is the indicator function.

The FE model elements can be summed into stiffness and mass substructures, $\boldsymbol{K}_j^\star$ and $\boldsymbol{M}_h^\star$, respectively, using the indicator function

$$\boldsymbol{K}_j^\star = \sum_{e=1}^{n_{el}} \mathbb{I}_j(e)\boldsymbol{K}_e; \quad \boldsymbol{M}_h^\star = \sum_{e=1}^{n_{el}} \mathbb{I}_h(e)\boldsymbol{M}_e \tag{24}$$

which are updated by a corresponding θ_j^k and θ_h^m:

$$\boldsymbol{K} = \sum_{j=1}^{p^k} \boldsymbol{K}_j^\star(1-\theta_j^k) = \boldsymbol{K}_0 - \sum_{j=1}^{p^k} \boldsymbol{K}_j^\star\theta_j^k \tag{25}$$

$$\boldsymbol{M} = \sum_{h=1}^{p^m} \boldsymbol{M}_h^\star(1-\theta_h^m) = \boldsymbol{M}_0 - \sum_{h=1}^{p^m} \boldsymbol{M}_h^\star\theta_h^m \tag{26}$$

Similar to $\boldsymbol{\delta}$, $\boldsymbol{\theta}$ is a concatenation of stiffness updating parameters $\boldsymbol{\theta}^k \in \mathbb{R}^{p^k}$ and mass updating parameters $\boldsymbol{\theta}^m \in \mathbb{R}^{p^m}$, with the total number of updating parameters $p = p^k + p^m$.

3.2. Sensitivity-based cluster analysis of updating parameters

Selection of updating parameters is an integral step in the model updating process, where three general conditions should be satisfied: (1) parameters should be chosen to avoid an ill-posed problem, (2) the choice of parameters should reflect the

objective of reducing modeling error, and (3) the model-outputs should be sensitive to the selected parameters [3]. Satisfying these three conditions is non-trivial. Ill-conditioning tends to occur in large problems when columns of the Jacobian are increasingly likely to exhibit linear dependence [4]. One solution is to use a subset of model parameters, a so-called 'subset selection' [23]. In this method, an orthogonal basis is created for the residual among a subset of parameter sensitivities. This reduces linear dependence in the Jacobian and ensures the parameters are effective in minimizing the residual.

An alternative approach retains all of the model parameters, but updates groups or clusters of model parameters, thereby reducing model order. This idea was presented in [24] as the 'best subspace approach', where clusters were chosen based on the angle between subspaces. This idea was improved upon and validated in Shahverdi et al.'s work [4], where clusters were selected based on hierarchical cluster analysis of parameter sensitivities. A major advantage of sensitivity-based clustering, as opposed to subset selection methods, is that clusters have a physical relevance: model parameters in the same cluster exhibit similar effect on model outputs. In the examples presented in [2,4,12,13], this meant that clustered parameters exhibited similar effects on the model natural frequencies.

Following the work of Shahverdi et al. [4] and Jang and Smyth [12,13], hierarchical clustering was selected for use as the grouping method in this paper. For a brief overview of hierarchical clustering, see [4], while a more thorough coverage is provided in [25]. When using hierarchical clustering, it is necessary to select a distance measure and a linkage method. The selected distance measure for this study is cosine distance. The cosine distance between the sensitivity vectors can be written

$$d_{cos}(\boldsymbol{j}_j,\boldsymbol{j}_k) = 1 - \frac{\boldsymbol{j}_j^T\boldsymbol{j}_k}{\sqrt{\boldsymbol{j}_j^T\boldsymbol{j}_j \cdot \boldsymbol{j}_k^T\boldsymbol{j}_k}} \tag{27}$$

where $\boldsymbol{j}_j$ and $\boldsymbol{j}_k$ are the sensitivities of δ_j and δ_k, respectively. The sensitivity vectors can be written as $\boldsymbol{j}_j = \partial\boldsymbol{r}/\partial\delta_j$, which are the columns of the Jacobian matrix $\boldsymbol{J} = [\boldsymbol{j}_1 \cdots \boldsymbol{j}_d]$.

Cosine distance is a measure of dissimilarity in shape between two vectors and does not consider the relative magnitudes. Furthermore, one can note that similar vectors are near-parallel. When near-parallel columns of the Jacobian are reduced through clustering, the condition of the problem tends to improve [4].

With a chosen clustering method and distance measure, the only remaining choice is the linkage method, which determines how vectors are combined to form clusters. The chosen method in this paper is the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [26]. This is an agglomerative method which combines the two nearest clusters at each step, evaluating the distance between the unweighted means of each cluster. This agglomerative process begins at the 'branches' of the dendrogram with single elements, then combines clusters at each step until it reaches the 'root' with only a single cluster encompassing all elements. When the dendrogram is built, the user is able to retrieve clusters by choosing to 'cut' the tree at a certain distance level, guaranteeing that each cluster is more than some chosen distance from any other. Alternatively, the user can input a desired number of clusters and the tree will be 'cut' to yield the desired number.

3.3. Objective-consistent scaling of cosine distance

While the use of cosine distance for sensitivity-based clustering was an excellent development, it is susceptible to becoming a skewed metric. When evaluating d_{cos} for two sensitivity vectors, skewing occurs when the sensitivity vectors comprise disparate sources, such as natural frequency and mode shape sensitivities. If the natural frequency sensitivities are considerably larger in magnitude than the mode shape sensitivities, the natural frequency sensitivities tend to dominate the resulting distance calculation. Indeed, there is no natural scale for mode shape sensitivities, as the mode shapes themselves can have arbitrary normalization. To mitigate the difference in magnitude for different sources of sensitivity, it is necessary to scale the residual sensitivity vectors.

To the authors' knowledge, no systematic method has been proposed for incorporating more than one type of data in cluster selection for FE model updating. Previous uses of parameter clustering in model updating only updated natural frequencies [2,4], or didn't incorporate mode shapes into cluster selection despite their inclusion in the objective function [12]. The proposed scaling for the residual sensitivity vectors is found by observing the gradient of the sum-of-squared residual from Eq. (3):

$$\frac{\partial E_r}{\partial\delta_j} = 2\boldsymbol{r}^T\boldsymbol{W}_r\frac{\partial\boldsymbol{r}}{\partial\delta_j} \tag{28}$$

where the sensitivity vector $\partial\boldsymbol{r}/\partial\delta_j$ is scaled by the weighted residual, $\boldsymbol{r}^T\boldsymbol{W}_r$. Given a diagonal $\boldsymbol{W}_r$, then each term of the sensitivity vector is multiplied by the corresponding residual and weighting term. Decomposing the inner product of $\boldsymbol{r}^T\boldsymbol{W}_r$ and $\partial\boldsymbol{r}/\partial\delta_j$ results in a sensitivity vector which is scaled by a matrix $\boldsymbol{W}_{\mathrm{OC}}$:

$$\boldsymbol{s}_j = \boldsymbol{W}_{\mathrm{OC}}\frac{\partial\boldsymbol{r}}{\partial\delta_j}; \quad \boldsymbol{W}_{\mathrm{OC}} = \mathrm{diag}(\boldsymbol{r}^T\boldsymbol{W}_r) \tag{29}$$

$\boldsymbol{W}_{\mathrm{OC}}$ is termed the 'objective-consistent' (OC) sensitivity scaling matrix, as it rescales the sensitivity vectors according to the residual objective function, E_r. The cosine distance between scaled sensitivities can be written as

$$d_{cos}(\boldsymbol{s}_j, \boldsymbol{s}_k) = 1 - \frac{\boldsymbol{j}_j^T \boldsymbol{W}_{OC}^2 \boldsymbol{j}_k}{\sqrt{\boldsymbol{j}_j^T \boldsymbol{W}_{OC}^2 \boldsymbol{j}_j \cdot \boldsymbol{j}_k^T \boldsymbol{W}_{OC}^2 \boldsymbol{j}_k}} \tag{30}$$

using the same notation as in Eq. (27), with $\boldsymbol{j}_j = \partial \boldsymbol{r} / \partial \delta_j$. This form clearly indicates the role of $\boldsymbol{W}_{OC}$.

This choice of scaling based on weighted residual confers several benefits. First, the cosine distance and corresponding clustering give more weight to sensitivity vector components which have a large corresponding residual, and less weight to sensitivity components with a small corresponding residual. This ameliorates scenarios where, say, a natural frequency sensitivity is large, but that particular natural frequency is already well-matched and has a residual component near zero. Second, the clustering reflects the scaling of the weighting matrix $\boldsymbol{W}_r$ and its effect on the objective value. Finally, this does not compromise the physical relevance of the clustering. Clusters are still selected based on the similarity of residual sensitivity vectors (which is related to model behavior), while the weighting only dictates the importance of each term in the cosine distance calculation.

4. Levenberg–Marquardt minimization with Bayesian regularization

With a well-defined residual and parametrized model, model updating proceeds with selection of an error minimization algorithm. The objective function is first modified to include regularization, then the minimization procedure is described following the approach in [17,18]. The regularized objective function, F, includes both a residual term E_r (see Eq. (3)) as well as an updating parameter size (penalty) term E_θ:

$$F(\boldsymbol{\theta}) = \beta E_r + \alpha E_\theta = \beta \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} + \alpha \boldsymbol{\theta}^T \boldsymbol{W}_\theta \boldsymbol{\theta} \tag{31}$$

α and β are scalar regularization parameters which influence the relative importance of E_r and E_θ during minimization. The process for calculating $\{\alpha, \beta\}$ is covered in Section 4.1.

The residual weighting matrix $\boldsymbol{W}_r$ and the parameter weighting matrix $\boldsymbol{W}_\theta$ must generally be symmetric positive semi-definite and should reflect the uncertainty in $\boldsymbol{r}$ and $\boldsymbol{\theta}$, respectively. The optimum choices (in a Bayesian sense) for $\boldsymbol{W}_r$ and $\boldsymbol{W}_\theta$ are the inverse of the measurement covariance matrix and the parameter covariance matrix, respectively [27–29]. The measurement covariance matrix is diagonal for the case of uncorrelated uncertainties, and is usually simple to estimate from data. The parameter covariance matrix is difficult to estimate, though [30] provides a method for relating $\boldsymbol{W}_\theta$ to the sensitivity matrix. The work presented in this paper uses a simple choice of $\boldsymbol{W}_\theta = \boldsymbol{I}$, similar to Tikhonov regularization [31].

The residual vector is a column of m elements given by

$$\boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \tag{32}$$

where $\tilde{\boldsymbol{z}}$ represents a vector of m measurements and $\boldsymbol{z}(\boldsymbol{\theta})$ represents a vector of m model outputs, given the column vector of p updating parameters $\boldsymbol{\theta}$. The parameter values and residual at iteration i are written $\boldsymbol{\theta}_i$ and $\boldsymbol{r}_i = \boldsymbol{r}(\boldsymbol{\theta}_i)$, respectively. At each iteration, the model parameters are updated such that

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i \tag{33}$$

To begin the solution algorithm, the residual at $\boldsymbol{\theta}_{i+1}$ is approximated using the truncated Taylor series of Eq. (1), giving

$$\boldsymbol{r}(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i) \approx \boldsymbol{r}_i + \boldsymbol{J}_i \Delta\boldsymbol{\theta}_i \tag{34}$$

where $\boldsymbol{J}_i$ is the Jacobian of $\boldsymbol{r}$ evaluated at $\boldsymbol{\theta}_i$, as in Eq. (2). Using this estimate for $\boldsymbol{r}_{i+1}$ in Eq. (31) and minimizing $F(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i)$ with respect to $\Delta\boldsymbol{\theta}_i$ yields

$$\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i] \tag{35}$$

which is the Gauss–Newton algorithm [19]. The Hessian at iteration i, $\boldsymbol{H}_i$, is estimated by the Gauss–Newton approximation

$$\boldsymbol{H}_i = \nabla\nabla F \approx 2[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i + \alpha \boldsymbol{W}_\theta] \tag{36}$$

Eq. (35) can be improved into the more robust Levenberg–Marquardt algorithm [32,33], which modifies the Hessian with a scalar damping term λ (unrelated to damping in mechanical vibrations), giving a trust-region solution:

$$\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i + 2\lambda \boldsymbol{I}]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i] \tag{37}$$

As $\lambda \to 0$, the Levenberg–Marquardt algorithm becomes the Gauss–Newton algorithm, while $\lambda \to \infty$ leads to the gradient-descent algorithm with infinitesimal step size. λ is adjusted using the multiplication scheme originally described by Marquardt [33], such that if the current value of λ results in $\Delta\boldsymbol{\theta}_i$ with $F_{i+1} < F_i$, λ is divided by a factor v for the next iteration. Otherwise λ is multiplied by a factor v, then $\Delta\boldsymbol{\theta}_i$, F_{i+1} are recomputed until $F_{i+1} < F_i$ or convergence. A reasonable set of initial values are $\lambda_0 = 0.01$ and $v = 10$, which fulfills the requirement that $\lambda \geqslant 0$.

4.1. Bayesian selection of regularization parameters

The regularization parameters α and β are used to produce a model updating solution that balances parameter values (E_θ) and residual (E_r). Within the Bayesian framework, α and β are treated as random variables [17]. If α and β are known, the posterior probability of the model parameters $\boldsymbol{\theta}$ is given by Bayes' rule:

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}},\alpha,\beta,\mathcal{M}) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta},\beta,\mathcal{M})P(\boldsymbol{\theta}|\alpha,\mathcal{M})}{P(\tilde{\boldsymbol{z}}|\alpha,\beta,\mathcal{M})} \tag{38}$$

where $\tilde{\boldsymbol{z}}$ are the measurements and $\mathcal{M}$ represents the chosen parametrization or model. $P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta},\beta,\mathcal{M})$ is the likelihood of the measurements given $\boldsymbol{\theta}$, $P(\boldsymbol{\theta}|\alpha,\mathcal{M})$ is the prior density of $\boldsymbol{\theta}$, and $P(\tilde{\boldsymbol{z}}|\alpha,\beta,\mathcal{M})$ is a normalization factor corresponding to the evidence for α and β. If it is assumed that noise in $\tilde{\boldsymbol{z}}$ is Gaussian, and that the prior distribution of $\boldsymbol{\theta}$ is also Gaussian, then the likelihood and the prior are

$$P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta},\beta,\mathcal{M}) = \frac{e^{-\beta E_r}}{Z_{\tilde{z}}}; \quad P(\boldsymbol{\theta}|\alpha,\mathcal{M}) = \frac{e^{-\alpha E_\theta}}{Z_\theta} \tag{39}$$

where the normalization factors are $Z_{\tilde{z}}(\beta) = (\pi/\beta)^{m/2}/\det(\boldsymbol{W}_r)^{1/2}$ and $Z_\theta(\alpha) = (\pi/\alpha)^{p/2}/\det(\boldsymbol{W}_\theta)^{1/2}$. Thus, the posterior is

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}},\alpha,\beta,\mathcal{M}) = \frac{e^{-F(\boldsymbol{\theta})}}{Z_F(\alpha,\beta)} \tag{40}$$

where $Z_F(\alpha,\beta)$ is a normalization factor which can be estimated by using a truncated Taylor series expansion of F from Eq. (31) about the local minimum point $\boldsymbol{\theta}_{\mathrm{MP}}$, which corresponds to the maximum probability point of the posterior distribution [17]. This approximation gives

$$Z_F(\alpha,\beta) \approx e^{-F(\boldsymbol{\theta}_{\mathrm{MP}})}(2\pi)^{p/2}/\det(\boldsymbol{H}_{\mathrm{MP}})^{1/2} \tag{41}$$

where $\boldsymbol{H}_{\mathrm{MP}} = \boldsymbol{H}(\boldsymbol{\theta}_{\mathrm{MP}})$ is the Hessian estimated from Eq. (36), as in [18]. Rewriting the posterior in Eq. (38) to find the evidence and substituting the results of Eqs. (39) and (40) gives

$$P(\tilde{\boldsymbol{z}}|\alpha,\beta,\mathcal{M}) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta},\beta,\mathcal{M})P(\boldsymbol{\theta}|\alpha,\mathcal{M})}{P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}},\alpha,\beta,\mathcal{M})} = \frac{Z_F(\alpha,\beta)}{Z_{\tilde{z}}(\beta)Z_\theta(\alpha)} \tag{42}$$

Substituting the expressions for the normalization terms and taking the logarithm of the evidence yields

$$\log P(\tilde{\boldsymbol{z}}|\alpha,\beta,\mathcal{M}) = -\beta E_r^{\mathrm{MP}} - \alpha E_\theta^{\mathrm{MP}} - \frac{1}{2}\log\det(\boldsymbol{H}_{\mathrm{MP}}) + \frac{m}{2}\log\beta + \frac{p}{2}\log\alpha + c \tag{43}$$

where c is not a function of α or β. Maximizing the log-evidence with respect to α and β yields

$$\alpha = \frac{\gamma}{2E_\theta^{\mathrm{MP}}} \quad \beta = \frac{m-\gamma}{2E_r^{\mathrm{MP}}} \tag{44}$$

where γ is the number of effective updating parameters [17], ranging from 0 to the number of parameters p:

$$\gamma = p - 2\alpha\,\mathrm{tr}(\boldsymbol{H}_{\mathrm{MP}}^{-1}\boldsymbol{W}_\theta) \tag{45}$$

γ provides an interesting insight into model efficiency and model selection [18]. If γ is close to p, then the problem might be 'saturated' and need more parameters to improve the solution. If increasing p results in the same γ as before, then the original p was sufficient (without changing the parameter selection scheme). $\gamma \ll p$ implies the solution is not sensitive to most of the chosen parameters, suggesting that p is too large or the selected parameters are inefficient. MacKay [17] suggests that $\gamma \approx p/2$ is a reasonable result, such that one 'unit' of noise is fitted for each well-estimated parameter.

Using knowledge of the relationships for α, β, and γ, the Levenberg–Marquardt minimization algorithm can then be modified to include iterative recalculation of the Bayesian regularization parameters, as done in [18]. The resulting pseudocode is provided in Table 1. Note that γ is recalculated in each iteration, but is only meaningful at a local minimum point (i.e. at a converged solution).

5. Model updating of a small-scale FE model: 29-element truss

5.1. Model description and cluster analysis

The first exercise of the proposed model updating scheme is the 29-element, 28-DoF, 2-dimensional truss depicted in Fig. 1. This truss was modified from the 29-DoF truss presented in Papadimitriou et al. [34]. The boundary conditions were changed to pin-pin (from pin-roller), thus making the structure symmetric and statically indeterminate. All truss elements utilized identical material properties, with Young's modulus (E) of 200 GPa, mass density (ρ) of 7800 kg/m^3, and area (A) 0.25 m^2. These values were essentially arbitrary, being chosen only to have a rough physical relevance.

Table 1
Pseudocode for Levenberg–Marquardt minimization with Bayesian regularization.

Input: Objective function $F(\boldsymbol{\theta}) = \beta E_r + \alpha E_\theta$ to be minimized
Output: Optimal parameters $\boldsymbol{\theta}_{\text{MP}}$, effective number of parameters γ
initialization: Set $\boldsymbol{\theta}_0$, $\alpha = 0.1$, $\beta = 1$, $\lambda = 0.01$, $v = 10$, $i = 0$;
while *not converged* **do**
  Compute $\boldsymbol{r}_i$, $\boldsymbol{J}_i$, and $\boldsymbol{H}_i = 2[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i + \alpha \boldsymbol{W}_\theta]$;
  Compute parameter update $\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i + 2\lambda \boldsymbol{I}]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i]$;
  Update $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i$;
  if $F(\boldsymbol{\theta}_{i+1}) > F(\boldsymbol{\theta}_i)$ **then**
    $\lambda \leftarrow \lambda \cdot v$;
    Go back to parameter update computation step ;
  else
    $\lambda \leftarrow \lambda / v$;
  end
  Compute effective number of parameters $\gamma = p - 2\alpha \,\text{tr}(\boldsymbol{H}_i^{-1} \boldsymbol{W}_\theta)$;
  Compute new estimates $\alpha = \gamma / (2E_\theta(\boldsymbol{\theta}_{i+1}))$ and $\beta = (m - \gamma)/(2E_r(\boldsymbol{\theta}_{i+1}))$;
  $i = i + 1$;
end

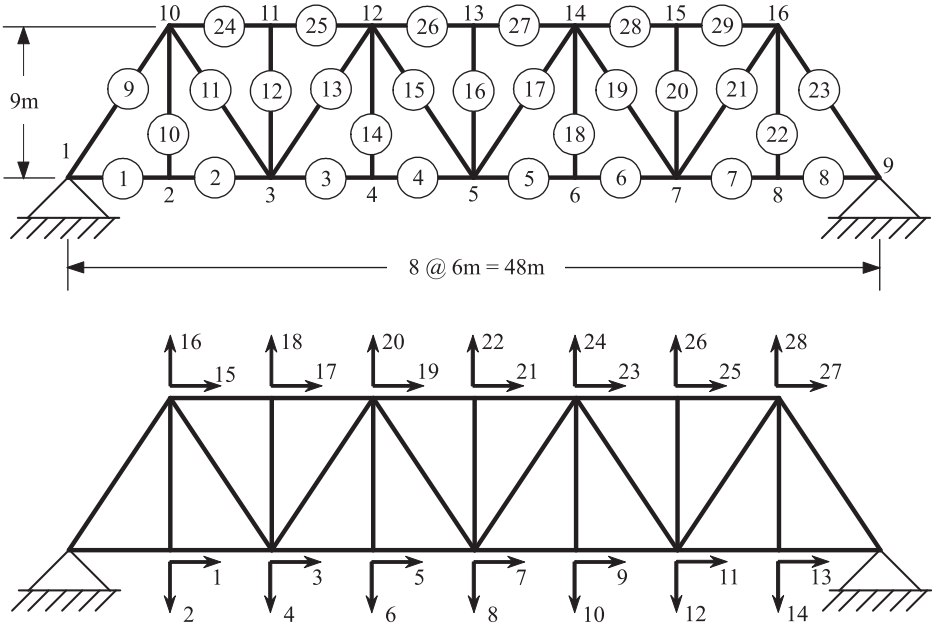


Fig. 1. 28-DoF truss structure, modified from Papadimitriou et al. [34].

The purpose of this small-scale model is not to demonstrate the power of parameter clustering for FE model updating, as a 29-element updating problem is easily tractable without any reduction in parameters. Instead, the small scale of the truss model allows a computationally-feasible comparison between different clustering methods across numerous structural states and realized measurements. A full-scale study only offers one structural state and realized measurement, limiting the ability to deduce if one updating method is generally superior to another method, or simply superior for that structural state. Furthermore, full-scale structural models are often impractical updating numerous randomized states due to prohibitive computation time.

The FE model of the truss was developed in MATLAB [35] using consistent (non-lumped) element mass matrices. For consistency with [34], the first five vibrational modes were analyzed, having natural frequencies of {9.31, 19.8, 26.9, 37.3, 51.2} Hz with no structural modification. It was assumed that the full field of 28 DoFs were measured.

Random realized states were generated by modifying the Young's modulus (E) and density (ρ) for each element, e, by an uncorrelated, uniform random number δ_e^k or δ_e^m, as in Eqs. (20) and (21). Since the truss element stiffness ($\boldsymbol{K}_e$) and element

mass ($\boldsymbol{M}_e$) are linearly dependent on the element E and ρ, respectively, this could also be seen as a direct modification of the element matrices. Two different modification levels were analyzed, $\delta \in [-0.5, 0.5]$ and $\delta \in [-0.05, 0.05]$, representing a poor and a good initial model state, respectively. 1000 random realized states were generated for each modification level.

Noise was then added to the natural frequency and mode shape measurements from each of the 2000 structural realizations. Each natural frequency measurement was corrupted with a Gaussian white noise sample with standard deviation equal to 0.5% of the measured natural frequency. Each mode shape was corrupted with a Gaussian white noise vector with standard deviation equal to 5% of the mode shape's standard deviation. This noise model was considered to be consistent with typical measurement conditions, in which natural frequencies are often reliable to within 1%, but mode shape measurements may exhibit an order-of-magnitude lower precision [1,36].

The objective function took the form of Eq. (18), with $\boldsymbol{W}_r$ equal to the inverse of measurement covariance matrix. Using the terminology in Eq. (18), this can be written as $w_{rj}^f = (0.005\tilde{f}_j)^{-2}$ and $w_{rj}^s = (0.05\ \mathrm{std}(\tilde{\boldsymbol{v}}_j))^{-2}$. Three different methods, or sets of sensitivity vectors, were used to generate clusters using the techniques described in Section 3.2. Six stiffness clusters were selected for each method, which avoided an overly complex depiction of stiffness clusters while still providing adequate similarity between element stiffness sensitivities in the same cluster. One mass cluster was selected (for a total $p = 7$), which was equivalent to updating the total structural mass. This was done to allow a limited ability to update mass discrepancy without requiring an additional description of mass clusters.

The first method (f cluster) used only natural frequency sensitivities $\partial \boldsymbol{f}/\partial \delta_k$, as in [2,4,12,13]. Of course, the authors in [2,4] only sought to update frequencies in their full-scale tests, so utilizing mode shape sensitivities would have been superfluous. The second method ($f + \phi$ cluster) used concatenated natural frequency and mode shape sensitivities $\partial \boldsymbol{f}/\partial \delta_k$ and $\partial \boldsymbol{v}_j/\partial \delta_k$, similar to $-\partial \boldsymbol{r}/\partial \delta_k$. The third method (OC cluster), which is the proposed method in this paper, utilized the objective-consistent weighting in Eq. (29) for each residual sensitivity vector.

Representative stiffness clusters for the three clustering methods, with the corresponding sensitivities, are shown in Fig. 2. Clustered elements are indicated by color and the number of dots on the element in Fig. 2e. Note that f clustering and $f + \phi$ clustering were based only on initial model sensitivities and were therefore constant for all realizations. OC clustering had weighting which depended on the realized residual and was not constant for all realizations. f clustering necessarily results in symmetric clusters for a symmetric structure, i.e. perturbing one element's stiffness or mass will have the same effect on natural frequencies as perturbing the symmetric element's stiffness or mass. Mode shape sensitivities were not observed to be symmetric or anti-symmetric across structural elements, even though the structure is symmetric. It is unknown if this is a general result and warrants further study. However, results do not seem trivial and are beyond the scope of the presented study.

Incorporating mode shape sensitivities into clustering resulted in asymmetric clusters, as in the $f + \phi$ and OC clustering in Fig. 2c and e. This asymmetry may not be a general result for OC clustering, e.g. a weighting matrix which strongly emphasizes natural frequency sensitivities may result in symmetric clustering. The absolute value of sensitivities are presented as 3-d bar plots in Fig. 2f. The cosine distance calculation utilized the full set of mode shape sensitivities (5 modes with 28 DoFs per mode), but are represented by MAC sensitivity in the bar plots for clarity.

5.2. Model updating results

Model updating was performed on each of the 2000 measurement realizations (i.e. for each randomized structural state, after the addition of noise). For each measurement realization, the three sets of clusters were generated, resulting in three different parametrizations of the same FE model in MATLAB. The parametrized models were then used to generate corresponding residual functions, $\boldsymbol{r}(\boldsymbol{\theta})$, relative to the realized measurements. The residual functions were regularized using Eq. (31) with the defined weighting matrices. The Levenberg–Marquardt method with Bayesian regularization was used to minimize the regularized objective function, as described in Section 4 and Table 1.

The average results of model updating with a poor initial model, i.e. $\delta \in [-0.5, 0.5]$, are shown in Table 2. The L^2 norm of the relative frequency error is used to provide a simple measure of natural frequency error, and is equal to a multiple of E_r^f since $w_{rj}^f = (0.005\tilde{f}_j)^{-2}$ in Eq. (18) under the implemented noise model. The L^2 norm of the relative frequency error showed significant improvement with each clustering method, with slightly better results for f clustering over OC clustering. This situation was reversed for average MAC (across all 5 mode shapes and 1000 realizations), where OC clustering slightly outperformed the other methods. However, OC clustering clearly outperformed the other methods in objective value (E_r), with approximately 17% lower objective function value relative to the initial value. OC clustering also resulted in slightly better γ, indicating that the objective function was more sensitive to those updating parameters. Each method had γ reasonably close to MacKay's suggestion of $\gamma = p/2 = 3.5$, suggesting that most of the parameters were effective in updating.

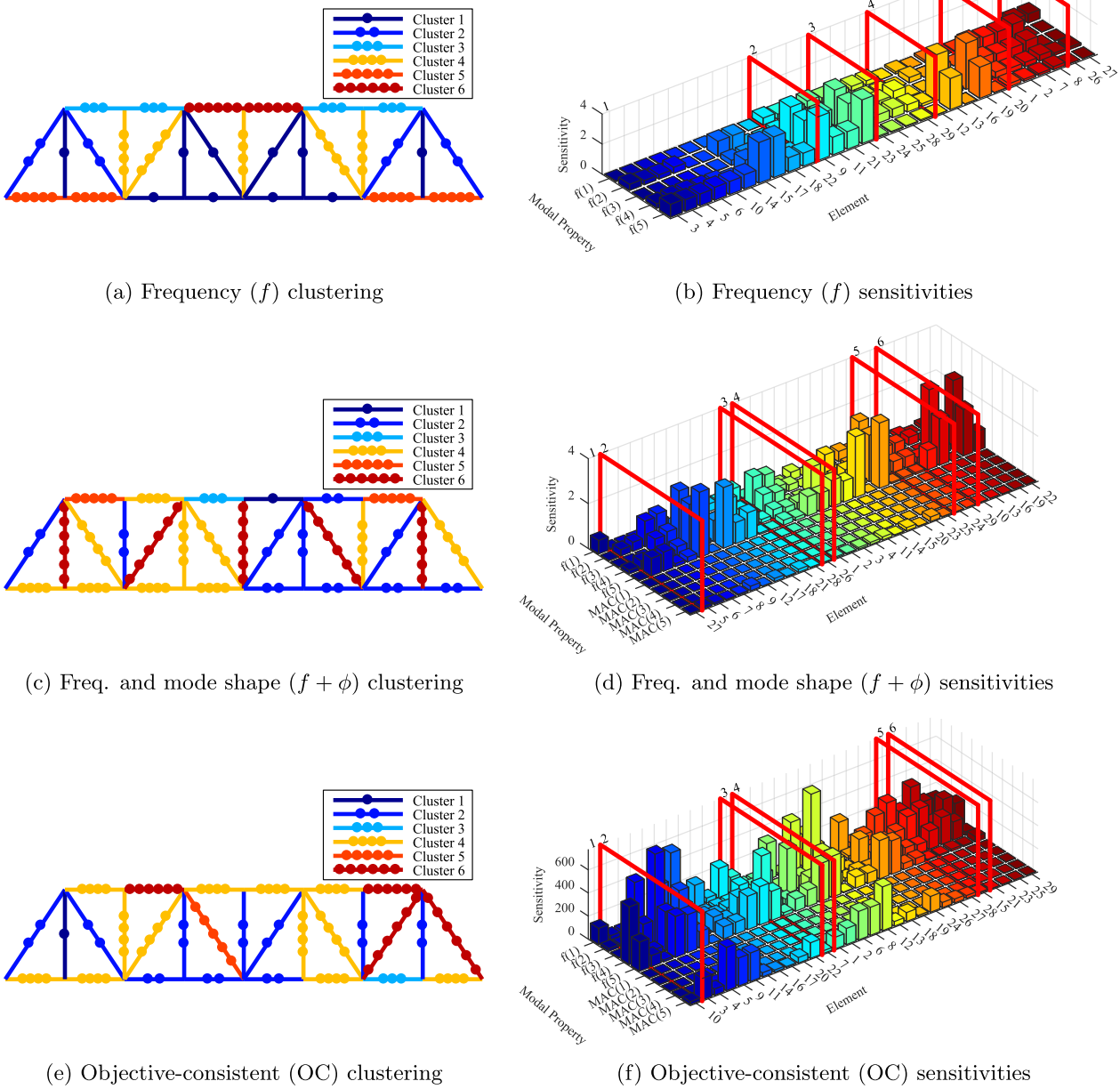
This set of observations is even more clear with a good initial model, $\delta \in [-0.05, 0.05]$ (Table 3). OC clustering was again outperformed by f and $f + \phi$ clustering for reducing frequency error, but it produced an average MAC value of 0.997 while f clustering essentially failed to improve the average MAC from 0.977. Ultimately, this means that f clustering was ineffective in reducing the objective function value E_r, while $f + \phi$ and OC clustering were extremely effective, with 80% and 86% reduction in E_r. The updating parameters selected by f clustering were largely ineffective, suggesting that fewer clusters could have

been used to achieve a similar result. OC clustering was considerably more effective, with approximately $\gamma = 5$ effective parameters.

6. Model updating of a full-scale FE model: large suspension bridge

6.1. System identification and model description

The second model updating exercise involves a full-scale FE model of a suspension bridge, as used in [12,13]. The bridge is a double-deck steel structure with two towers and four suspension cables, each 982 m long. The three spans (two side-spans, one main-span) total 2089 m of length, with the main-span comprising 451 m. In 2009, an ambient vibration study was performed to identify natural frequencies, damping ratios, and mode shapes of this bridge under normal operating conditions [13]. Dynamic responses were recorded at 9 locations using tri-axial force-balance accelerometers, totaling 27 measured DoFs at a sampling frequency of 200 Hz. The first seven dynamic modes were selected for model updating, totaling $m = 196$ modal measurements between 1 natural frequency and 27 DoFs for each mode.



(a) Frequency (f) clustering

(b) Frequency (f) sensitivities

(c) Freq. and mode shape ($f + \phi$) clustering

(d) Freq. and mode shape ($f + \phi$) sensitivities

(e) Objective-consistent (OC) clustering

(f) Objective-consistent (OC) sensitivities

Fig. 2. Selected truss clustering results and corresponding sensitivities.

Table 2
Truss model updating results, $\delta \in [-0.5, 0.5]$.

	Initial	Updated		
Average result		f cluster	$f+\phi$ cluster	OC cluster
$\|(\tilde{\boldsymbol{f}}-\boldsymbol{f})/\tilde{\boldsymbol{f}}\|_2$	15.5%	4.36%	9.20%	7.95%
MAC	0.920	0.933	0.944	0.956
Obj. value, E_r ($\times 10^3$)	6.47	4.43	5.05	3.33
Eff. no. params., γ		4.67	4.42	4.86

Table 3
Truss model updating results, $\delta \in [-0.05, 0.05]$.

	Initial	Updated		
Average result		f cluster	$f+\phi$ cluster	OC cluster
$\|(\tilde{\boldsymbol{f}}-\boldsymbol{f})/\tilde{\boldsymbol{f}}\|_2$	1.55%	1.01%	1.51%	1.25%
MAC	0.977	0.977	0.996	0.997
Obj. value, E_r ($\times 10^3$)	1.48	1.46	0.29	0.21
Eff. no. params., γ		1.62	4.52	4.81

Table 4
Suspension bridge measured and initial modal properties.

		Measured	Initial		
Mode	Description	$\tilde{f}$ (Hz)	f (Hz)	$(\tilde{f}-f)/\tilde{f}$	MAC
H1	First lateral	0.194	0.236	−21.7%	0.984
V1	First vertical	0.227	0.294	−29.8%	0.969
V2	Second vertical	0.303	0.356	−17.5%	0.986
T1	First torsional	0.373	0.384	−3.0%	0.741
SV1	First side-span vertical	0.337	0.453	−34.2%	0.879
H2	Second lateral	0.450	0.540	−19.8%	0.845
V3	Third vertical	0.500	0.596	−19.4%	0.743

Responses were recorded during four 1-h sessions in a single day, at 3 a.m., 8 a.m., 1 p.m., and 8 p.m. The responses during these time periods were assumed to be approximately stationary, as the environmental and operating conditions were unlikely to vary significantly. Output-only modal identification was performed on each of the 1-h data sets using the enhanced frequency domain decomposition (EFDD) method [37]. Power spectral density matrices were constructed using Welch's method with a Hamming window. The identified modal properties for all four 1-h periods were then averaged to give a set of modal properties representative of an average daily behavior. As shown in previous work, the measured modes had negligible imaginary components [12,13], so only the real components of the mode shapes were used. The average measured natural frequencies are given in Table 4.

Natural frequency measurements ranged in coefficient of variation from 0.25% (mode T1) to 5.0% (mode H1), estimated from the four measurements per mode. Mode shape measurement variation was gauged by the ratio of mode shape measurement standard deviation (averaged for all DoFs in the mode) to standard deviation of the average mode shape, which ranged from 3.9% (mode T1) to 12.0% (mode SV1). This data could have been used to estimate a measurement covariance matrix, but four sets of measurements was considered to be an inadequate sample. Instead, the measurement covariance matrix was formed based on an assumed noise model, equivalent to that in Section 5. It was assumed that the noise in each natural frequency measurement had a standard deviation equal to 0.5% of the measured frequency value. For each mode, it was assumed that every the noise in each component had a standard deviation equal to 5% of the measured mode shape's standard deviation. The residual weighting matrix, $\boldsymbol{W}_r$, was then the inverse of this assumed measurement covariance matrix. Using the terminology in Eq. (18), this noise model can be written as $w_{rj}^f = (0.005\tilde{f}_j)^{-2}$ and $w_{rj}^s = (0.05\ \mathrm{std}(\tilde{\boldsymbol{\nu}}_j))^{-2}$.

Fig. 3 depicts the measured mode shapes, with sensor locations indicated by dots. The mode shape magnitudes at the sensor locations are representative of the average measured modal data, while the unmeasured parts of the mode shape were interpolated to fit projected data and given boundary conditions. Note that the second lateral mode (H2) in Fig. 3f appears to be a torsional mode, but this is a visual illusion caused by the viewing angle which accentuates the deck torsion over lateral displacement. For more information on the interpolated mode shapes, please refer to [12,13]. Note that interpolation was only used for purposes of depiction, while any discussion of measured modes or MAC only utilizes data at the 27 measured DoFs. Fig. 4a shows the measured mode shape autoMAC matrix, indicating strong orthogonality.

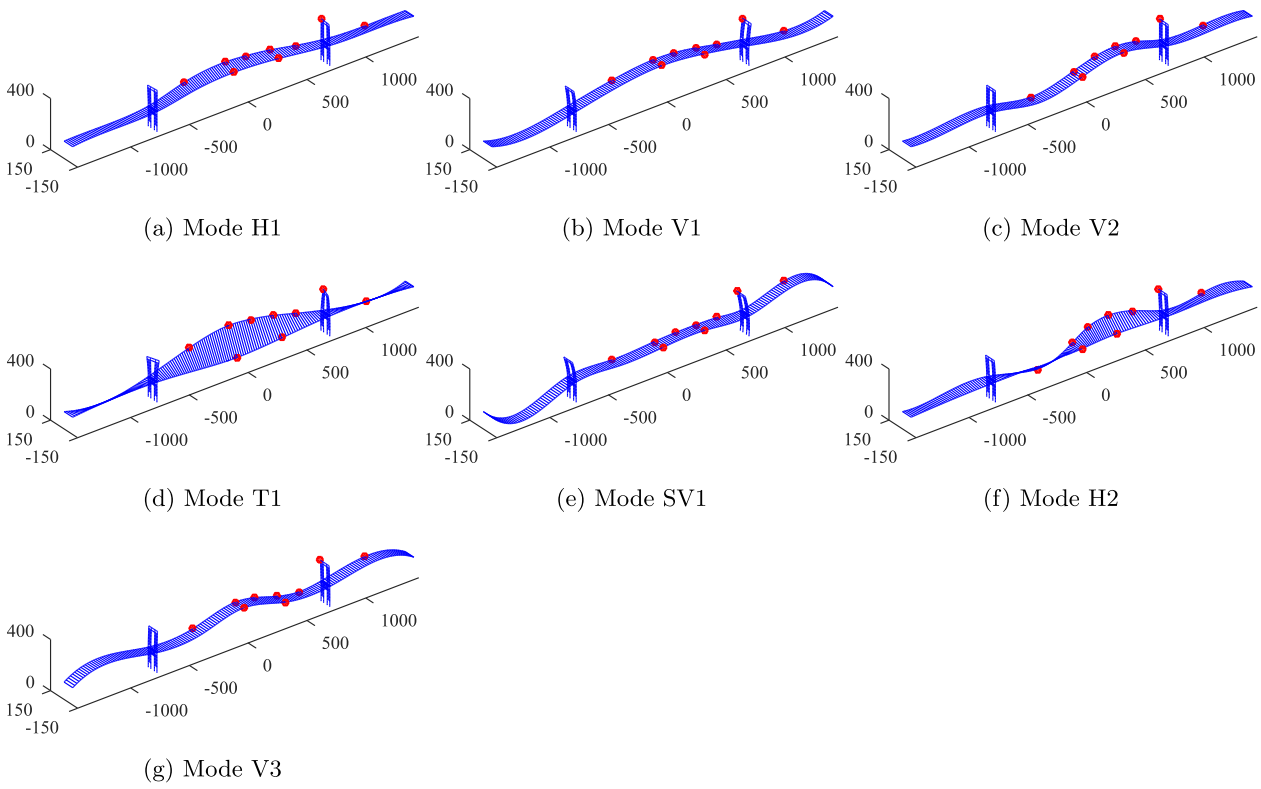


Fig. 3. Suspension bridge measured mode shapes (measurement locations indicated by red dots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

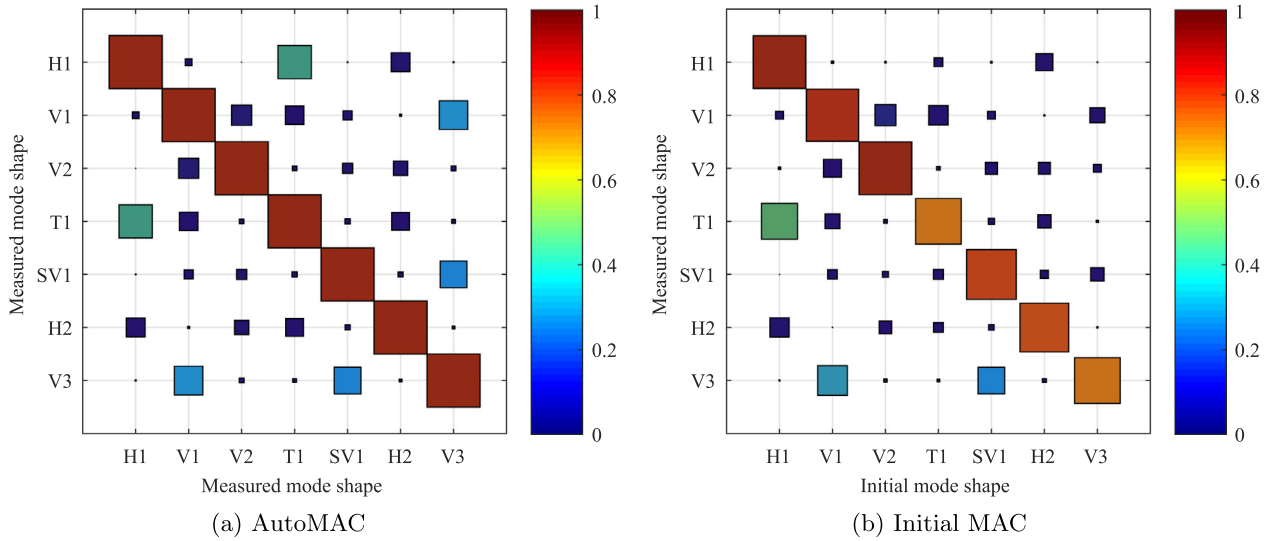


Fig. 4. Suspension bridge autoMAC for measured modes (a) and initial model MAC (b) matrices.

The FE model of the suspension bridge was implemented in ABAQUS [38] with 19,632 beam elements, 1464 truss elements, and 18,614 nodes. The node locations and connectivity were carefully defined based on partial technical drawings and photographs. Element properties were calculated from indicated cross-section data when available, while reasonable assumptions based on photographs were made for uninformed cases. Soil springs were included at the bottom of the pylons, the end of the deck, and the end of the suspension cables. To account for temperature expansion joints, hinge springs were implemented between the deck and towers. For a more thorough discussion of element types, boundary conditions, connections, and initial model material properties, please refer to Jang and Smyth's description of the same FE model [12].

The natural frequencies, relative frequency error, and MAC values of the initial FE model are given in Table 4. The initial MAC values are visually represented Fig. 4b. The initial MAC values are slightly different from those in Jang and Smyth's pre-

vious work [12,13] due to FE model modifications to account for more realistic structural behavior, particularly in the boundary conditions and interactions between cable and deck components. For each mode, the natural frequencies of the initial FE model were higher than their measured counterparts. The first torsional (T1) mode exhibited the lowest initial error at −3.0%, while the first side-span vertical (SV1) mode exhibited the highest initial error at −34.2%. Initial MAC values were excellent (greater than 0.95) for the first two vertical modes (V1 and V2) as well as the first lateral mode (H1). Conversely, the first torsional (T1) and the third vertical mode (V3) exhibited the most unsatisfactory MAC values at 0.741 and 0.743, respectively. Every mode exhibited a low MAC and/or a high frequency error, suggesting that every mode would be important in the updating process.

6.2. Parameter clustering

The selection of clusters and updating parameters for the full-scale suspension bridge model proceeded using similar methodology to the small-scale truss model in Section 5. The large scale of the suspension bridge eliminated the possibility of directly updating every element mass and stiffness. The FE model had well over 42,000 physical parameters to update, among more than 21,000 elements with separate Young's moduli and mass densities, plus several soil and hinge spring constants. Note that the geometry of the suspension bridge was not included within the updating parameters.

Prior to clustering, the structural components of the FE model were arranged into 132 substructures based on location and element type, which could be viewed as an expert-informed pre-clustering. The span was longitudinally divided into 8 main-span and 8 side-span segments, which were further partitioned based on element type. Each tower was divided into three vertical sections, which was further divided into bracing and pylon groups. The 132 resulting substructures were assigned mass density (δ_e^m) and Young's modulus (δ_e^k) modifications as model parameters. The soil and hinge springs, comprising 15 spring coefficients, were also assigned spring constant modification parameters (δ_e^k). This totaled 147 stiffness model parameters and 132 mass model parameters, for 279 total model parameters to be updated.

Two sets of clusters were selected using the methodology of Section 3.2, one based on frequency sensitivity (f cluster) and the other based on the objective-consistent weighted residual sensitivity (OC cluster). Previous work on this bridge selected clusters using a maximum cosine distance clustering criterion [12,13]. However, this criterion generally results in different numbers of clusters for different sensitivity vector sets. Models with more parameters are more likely to fit the data [17], making this clustering criterion inappropriate for comparing parametrizations. Therefore, the number of clusters was fixed at 5 mass clusters and 17 stiffness clusters for both the f clustering and OC clustering results, equivalent to prior work on this structure [12,13].

The sensitivities of the model outputs to the 22 OC cluster updating parameters are represented in Fig. 5. A more detailed representation of one mass cluster (3) and one stiffness cluster (19) are presented in Fig. 6. Similar to Section 5, OC clustering resulted in several asymmetric clusters, while frequency clustering, as done in previous work [12], resulted in symmetric

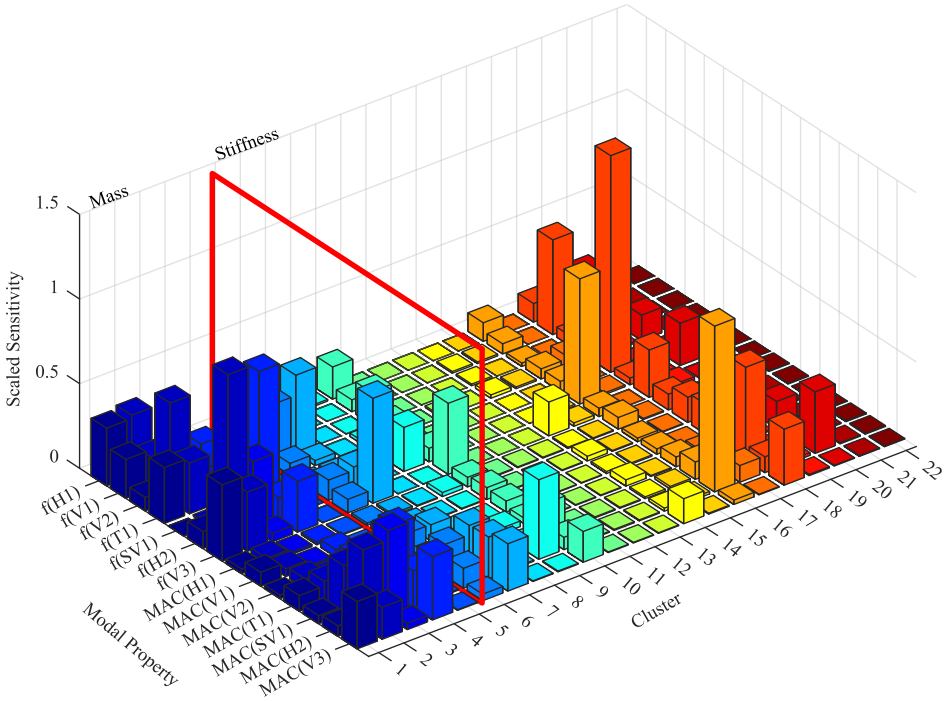
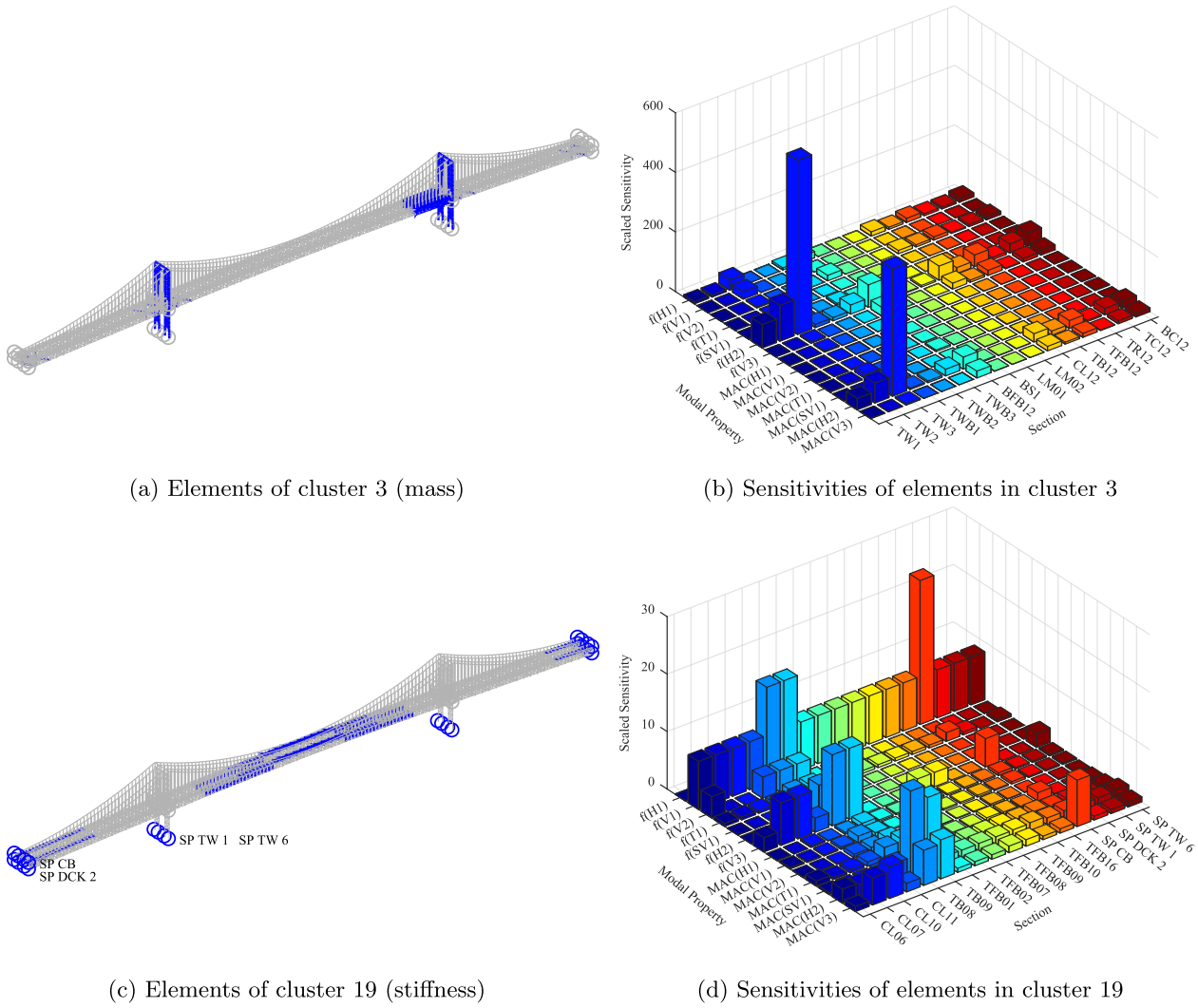


Fig. 5. Suspension bridge sensitivities to the 22 objective-consistent cluster updating parameters.



(a) Elements of cluster 3 (mass)

(b) Sensitivities of elements in cluster 3

(c) Elements of cluster 19 (stiffness)

(d) Sensitivities of elements in cluster 19

Fig. 6. Suspension bridge objective-consistent clustering results: selected clusters and sensitivities.

clusters. Cluster 3 is appreciably asymmetric with all clustered model parameters exhibiting a significant effect on the second horizontal mode (H2), and comprises mostly tower (TW) and tower brace (TWB) sections. Cluster 19 includes many top-deck braces (TB), top-deck floor beams (TFB), and columns (CL) in the mid-span area, TFBs at the ends of the side-spans, and cable-soil springs (SP TW). Physical intuition would indicate that these elements would be most impactful on the vertical modes (V1-V3), which is confirmed by the sensitivities in Fig. 6d.

It is important to consider the impact of sensor layout on clustering. Clustering schemes which only utilize model-outputs (e.g. frequency-based clustering) are not affected by the sensor layout. However, the proposed OC clustering scheme gives greater consideration to sensitivities associated with large weighted residuals, as in Eq. (29). Since there was only one sensor on the side-span (Fig. 3), it was expected that modes with significant side-span motion (e.g. SV1 and V3) would be under-represented in the clustering process. However, the low initial MAC values for modes SV1 and V3 indicates that most of the initial model mode shape outputs, not just the side-span output, had significant discrepancy. Therefore, modes SV1 and V3 were still highly weighted during clustering, as reflected in clusters 2 and 18 (Fig. 5).

6.3. Model updating results

The two clustering results were obtained using finite difference to estimate the sensitivity matrix for FE model parameters. With the models parametrized using the two clustering schemes, model updating proceeded by defining parametrized residual functions for each clustering result. These residual functions were used in the regularized objective in Eq. (31), which was minimized using Levenberg–Marquardt minimization with Bayesian regularization, detailed in Section 4 and Table 1. Note that clustering and updating were handled in MATLAB, while ABAQUS was used for modal analysis of the bridge. An application programming interface was built for communicating between MATLAB and ABAQUS.

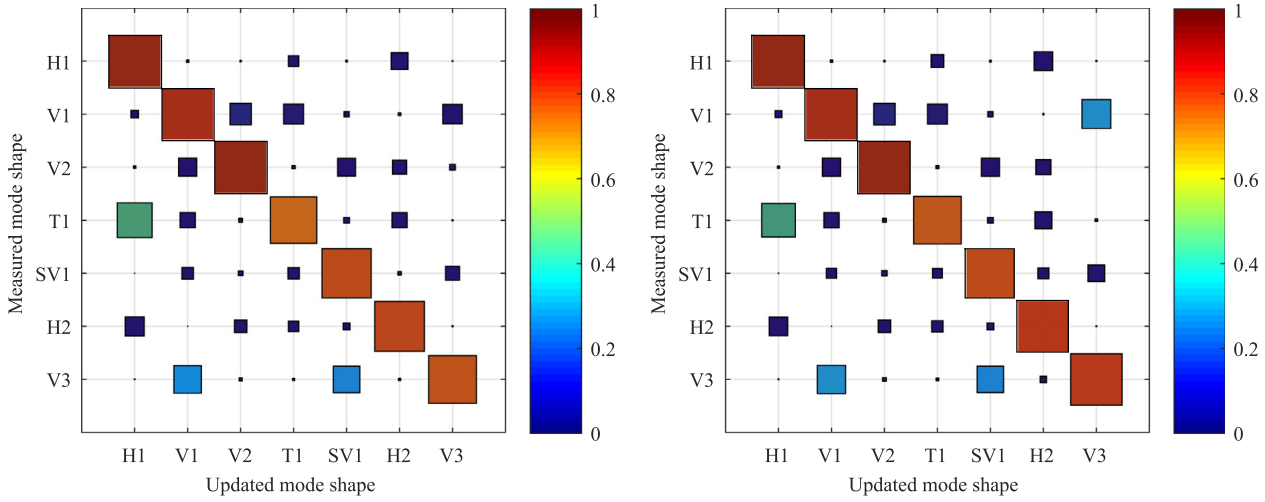
Table 5
Suspension bridge model updating results.

	Initial		Updated			
			f cluster		OC cluster	
Mode	$(\tilde{f}-f)/\tilde{f}$	MAC	$(\tilde{f}-f)/\tilde{f}$	MAC	$(\tilde{f}-f)/\tilde{f}$	MAC
H1	−21.7%	0.984	1.0%	0.989	4.4%	0.993
V1	−29.8%	0.969	−6.9%	0.966	−11.0%	0.973
V2	−17.5%	0.986	2.2%	0.977	2.2%	0.984
T1	−3.0%	0.741	12.9%	0.766	12.4%	0.832
SV1	−34.2%	0.879	−4.7%	0.861	−5.4%	0.869
H2	−19.8%	0.845	−3.6%	0.876	2.1%	0.970
V3	−19.4%	0.743	−5.3%	0.812	−5.6%	0.983
$\|(\tilde{\boldsymbol{f}}-\boldsymbol{f})/\tilde{\boldsymbol{f}}\|_2$	60.0%		16.8%		19.0%	
Av. MAC		0.878		0.893		0.928
Obj. value, E_r ($\times 10^4$)	2.39		0.95		0.70	
Eff. no. params., γ			6.4		10.1	

Due to the limited number of measurement locations, it was necessary to introduce an intermediate step during mode pairing. Initial FE model modes were paired with measured modes based on MAC using the 27 measured DoFs, which created an 'index' between initial FE model modes and measured modes. During the updating process, updated FE model modes were first paired with initial FE model modes using all model DoFs, producing high-fidelity pairing results. Then the previously-established mode index was used to correspond each updated mode to the correct measured mode. This approach was necessary to ensure consistent pairing between FE model modes and measured modes, since FE model mode shapes could exhibit large changes during the updating process and had relatively few DoFs for direct pairing.

The updating parameters of both clustering schemes converged in 3 iterations, with results shown in Table 5. As in Section 5.2, the L^2 norm of relative frequency error is utilized to give a summary of natural frequency error and is closely related to E_r^f.

Natural frequency updating results are similar between f and OC clustering for modes V2, T1, SV1, and V3. Both clustering schemes struggled to update the first torsional mode (T1), with neither achieving less than 12% relative frequency error. The difficulty with mode T1 was noted in previous work by Jang and Smyth [12], and may be due to non-linear geometry which was not included in the FE model. Three T1 modes were identified in the measured data sets, each with different in-phase or out-of-phase motions between the main cable and the deck, and different interactions between the main-span and the side-span. However, the FE model produced only one T1 mode because the geometrically non-linear deck-cable interaction was not included in the model. For the model updating study here, the measured T1 mode which had the highest MAC value with the mode shapes of the model was considered. Particular to the OC clustering algorithm, the low initial error of mode T1 tended to decrease the importance of this frequency measurement during clustering, perhaps explaining the difficulty for the proposed algorithm. Generally, f clustering achieved better natural frequency results, giving a 72% reduction in L^2 norm natural frequency error while OC clustering achieved 68% reduction.



(a) f cluster updated MAC (b) OC cluster updated MAC

Fig. 7. Suspension bridge frequency cluster (a) and OC cluster (b) updated MAC matrices.

However, the converse occurred for mode shape updating, as indicated by OC clustering increasing the average MAC to 0.928 from an initial value of 0.878 while *f* clustering only achieved a mild improvement to 0.893. OC clustering produced better MAC values than *f* clustering for every mode, with small increases noted for both methods on modes H1, V1, and V2, which had high initial MAC values. OC clustering had the most notable impact on the MAC of modes H2 and V3, which were increased from less than 0.850 to over 0.970, while *f* clustering showed only a minor improvement. Similarly, OC clustering was able to raise the MAC of mode T1 to 0.832 from 0.741 which is an appreciable increase and clearly outperformed *f* cluster's improvement to 0.766. Both clustering methods struggled to update the mode shape of the side-span vertical mode (SV1), actually reporting slight decreases in MAC value. During the bridge monitoring, various sensor configurations were tested. However, only one configuration was used for this study. The difficulty associated with the mode SV1 was likely related to the sensor configuration used for this study, which had only one sensor on the side-span (Fig. 3e), thereby limiting the importance of side span motion in MAC calculations. The updated MAC values are also presented diagrammatically in Fig. 7.

Despite mildly poorer results with natural frequencies, the enhanced improvement in mode shapes provided OC clustering with a significant edge over *f* clustering with respect to the objective function value E_r (Table 5). OC clustering reduced E_r by 70% from its initial value, while *f* clustering only reduced E_r by 60%. This result is partly explained by OC clustering having a higher effective number of parameters, γ, during the updating process. Out of $p = 22$ updating parameters, $\gamma = 10.1$ were effective for OC clustering which is quite close to MacKay's suggestion $\gamma = p/2$ [17]. Only $\gamma = 6.4$ parameters were effective for *f* clustering indicating that the updating solution was less sensitive to this parametrization.

7. Conclusions

Sensitivity-based parameter clustering presents a viable method for improving the condition and efficiency of FE model updating problems. The presented OC parametrization scheme allows for compatible use of disparate data (e.g. natural frequencies and mode shapes) in the selection of parameter clusters, thereby improving the efficiency and quality of results. The proposed clustering scheme retains the physical relevance of previous clustering schemes, which only used natural frequency sensitivities. Laid atop this foundation, OC clustering also considers the measurement residual and residual weighting inherent in the objective function. Indeed, the presented weighting scheme is extremely generic, as it helps produce clusters which have the most similar effect on the weighted residual. Bayesian regularization was also proposed for use in model updating. Built atop the Levenberg–Marquardt minimization algorithm, Bayesian regularization delivers an optimal set of regularization parameters with minimal computational overhead. Additionally, it provides key insight into the effective number of updating parameters, which can be used for model selection and model comparison.

The proposed OC clustering scheme was shown to be highly effective in both a small-scale exercise with simulated data and a full-scale exercise with real data. When compared to a clustering scheme based only on natural frequency sensitivity, OC clustering resulted in significantly better updating of mode shapes at the cost of slightly more frequency error. Incorporation of mode shape data produced asymmetric clusters, while clustering based on frequency produced symmetric clusters for the symmetric structures of study. For the small-scale model with a poor initial model (far from measured state), OC clustering achieved 17% better error reduction than frequency clustering. With a good initial model, OC clustering reduced the objective value by 84% more than frequency clustering. Observing the effective number of parameters further confirmed the improved efficiency of OC cluster analysis, which consistently showed a greater number of effective parameters compared to frequency clustering. These results were more tempered on the full-scale model, where OC clustering reduced the objective value by 10% more than frequency clustering, but it also showed significantly better agreement with measured mode shapes and greater number of effective parameters. For the model updating examples covered in this work, regularization was unnecessary for determination of a unique solution since parameter clustering resulted in overdetermined problems. Bayesian regularization was primarily useful for imposing equivalent prior probability distributions on competing parametrizations and estimating model efficiency.

The two contributions of this paper, OC clustering and Bayesian regularization, present simple and effective developments on existing methods. However, these contributions are not without drawbacks. While OC clustering is intuitive and showed improvement over existing schemes, it is not guaranteed to be the optimal technique for every structure or realization. Further work is required to analyze the effect of clustering on problem condition. OC clustering also suffers by not considering physical proximity of elements, resulting in clustered elements which are generally not physically adjacent. This lack of proximity poses problems for damage localization [12]. It may be noted that clusters are only selected once, starting at the initial model. As the model is updated, however, the model sensitivity matrix will change. Depending on the amount of change, it may be such that the initial clustering is no longer efficient. This problem was not considered, and warrants further research.

Regularization has the drawback of altering the objective, adding a side-constraint which may not be important to the user. While Bayesian regularization provides a key insight, it is not a refined tool for model selection. While it can be used to compare the efficiency of different parametrizations, it does not provide strong suggestions for alternative parametrizations other than to increase or decrease the number of parameters. Further work is needed to understand the limitations of this approach.

Acknowledgments

The authors gratefully acknowledge Columbia University's Graduate School of Arts and Sciences in support of the first author through the Guggenheim Fellowship and Presidential Fellowship. This work was partially supported by the U.S. National Science Foundation (Grant No. CMMI-1563364).

References

[1] J.E. Mottershead, M.I. Friswell, Model updating in structural dynamics: a survey, J. Sound Vib. 167 (2) (1993) 347–375.
[2] J.E. Mottershead, M. Link, M.I. Friswell, The sensitivity method in finite element model updating: a tutorial, Mech. Syst. Signal Process. 25 (7) (2011) 2275–2296.
[3] M.I. Friswell, J.E. Mottershead, Finite Element Model Updating in Structural Dynamics, vol. 38, Springer Science & Business Media, 1995.
[4] H. Shahverdi, C. Mares, W. Wang, J.E. Mottershead, Clustering of parameter sensitivities: examples from a helicopter airframe model updating exercise, Shock Vib. 16 (1) (2009) 75–87.
[5] J.E. Mottershead, C. Mares, M.I. Friswell, S. James, Selection and updating of parameters for an aluminium space-frame model, Mech. Syst. Signal Process. 14 (6) (2000) 923–944.
[6] J.M.W. Brownjohn, P. Moyo, P. Omenzetter, Y. Lu, Assessment of highway bridge upgrading by dynamic testing and finite-element model updating, J. Bridge Eng. 8 (3) (2003) 162–172.
[7] A. Teughels, G. De Roeck, Structural damage identification of the highway bridge Z24 by FE model updating, J. Sound Vib. 278 (3) (2004) 589–610.
[8] B. Jaishi, W.-X. Ren, Structural finite element model updating using ambient vibration test results, J. Struct. Eng. 131 (4) (2005) 617–628.
[9] J.R. Wu, Q.S. Li, Finite element model updating for a high-rise structure based on ambient vibration measurements, Eng. Struct. 26 (7) (2004) 979–990.
[10] P.G. Bakir, E. Reynders, G. De Roeck, Sensitivity-based finite element model updating using constrained optimization with a trust region algorithm, J. Sound Vib. 305 (1) (2007) 211–225.
[11] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: a review, Mech. Syst. Signal Process. 56 (2015) 123–149.
[12] J. Jang, A.W. Smyth, Model updating of a full-scale FE model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters, Mech. Syst. Signal Process. 83 (2017) 337–355.
[13] J. Jang, A.W. Smyth, Bayesian model updating of a full-scale finite element model with sensitivity-based clustering, Struct. Control Health Monitor. 24 (11) (2017) e2004.
[14] H. Ahmadian, J.E. Mottershead, M.I. Friswell, Regularisation methods for finite element model updating, Mech. Syst. Signal Process. 12 (1) (1998) 47–64.
[15] M.I. Friswell, J.E. Mottershead, H. Ahmadian, Finite-element model updating using experimental test data: parametrization and regularization, Philosoph. Trans. Roy. Soc. Lond. A: Math. Phys. Eng. Sci. 359 (1778) (2001) 169–186.
[16] B. Titurus, M.I. Friswell, Regularization in model updating, Int. J. Numer. Meth. Eng. 75 (4) (2008) 440–478.
[17] D.J.C. MacKay, Bayesian interpolation, Neural Comput. 4 (3) (1992) 415–447.
[18] F.D. Foresee, M.T. Hagan, Gauss-Newton approximation to Bayesian learning, in: International Conference on Neural Networks, vol. 3, IEEE, 1997, pp. 1930–1935.
[19] Å. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, PA, 1996.
[20] R.J. Allemang, D.L. Brown, A correlation coefficient for modal vector analysis, in: Proceedings of the 1st International Modal Analysis Conference, vol. 1, 1982, pp. 110–116.
[21] R.L. Fox, M.P. Kapoor, Rates of change of eigenvalues and eigenvectors, AIAA J. 6 (12) (1968) 2426–2429.
[22] S. Adhiakri, Rates of change of eigenvalues and eigenvectors in damped dynamic system, AIAA J. 37 (11) (1999) 1452–1458.
[23] G. Lallement, J. Piranda, Localization methods for parametric updating of finite element models in elastodynamics, in: International Modal Analysis Conference, 8th, 1990, pp. 579–585.
[24] M.I. Friswell, J.E. Mottershead, H. Ahmadian, Combining subset selection and parameter constraints in model updating, J. Vib. Acoust. 120 (4) (1998) 854–859.
[25] L. Rokach, O. Maimon, Clustering Methods, Springer, US, Boston, MA, 2005.
[26] R.R. Sokal, C.D. Michener, A statistical method for evaluating systematic relationships, Univ. Kansas Sci. Bull. 38 (1958) 1409–1438.
[27] J.D. Collins, G.C. Hart, T. Haselman, B. Kennedy, Statistical identification of structures, AIAA J. 12 (2) (1974) 185–190.
[28] M.I. Friswell, The adjustment of structural parameters using a minimum variance estimator, Mech. Syst. Signal Process. 3 (2) (1989) 143–155.
[29] T. Strutz, Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond, Vieweg and Teubner, Germany, 2010.
[30] M. Link, Updating of analytical models – procedures and experience, in: Proceedings of the Conference on Modern Practice in Stress and Vibration Analysis, Sheffield Academic Press, 1993, pp. 35–52.
[31] A.N. Tikhonov, V.I. Arsenin, Solutions of ill-posed problems, vol. 14, V.H. Winston and Sons (distributed by Wiley, New York), 1977.
[32] K. Levenberg, A method for the solution of certain problems in least squares, Q. Appl. Math. 2 (1944) 164–168.
[33] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Math. 11 (2) (1963) 431–441.
[34] C. Papadimitriou, J.L. Beck, S.-K. Au, Entropy-based optimal sensor location for structural model updating, J. Vib. Control 6 (5) (2000) 781–800.
[35] MATLAB, version 9.1.0 (R2016b), The MathWorks Inc., Natick, MA, 2016.
[36] S.V. Modak, T.K. Kundra, B.C. Nakra, Comparative study of model updating methods using simulated experimental data, Comput. Struct. 80 (5) (2002) 437–447.
[37] R. Brincker, C. Ventura, P. Andersen, Damping estimation by frequency domain decomposition, in: 19th International Modal Analysis Conference, 2001, pp. 698–703.
[38] ABAQUS/CAE, User's Guide: Version 6.14, Dassault Systèmes Simulia Corp., Providence, RI, 2014.

Paper B

Sensitivity-based SVD parametrization and optimal regularization in finite element model updating

---

Bibliographic information:

Bartilson, D. T., Jang, J., & Smyth, A. W. (2018). Sensitivity-based SVD parametrization and optimal regularization in finite element model updating. *Manuscript submitted for publication*.

This page intentionally left blank.

Received –; | Revised –; | Accepted –

DOI: xxx/xxxx

RESEARCH ARTICLE

Sensitivity-based SVD parametrization and optimal regularization in finite element model updating

Daniel T. Bartilson[1] | Jinwoo Jang[2] | Andrew W. Smyth[1]

[1]Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, USA
[2]Department of Civil, Environmental & Geomatics Engineering, Florida Atlantic University, Florida, USA

Correspondence
Daniel T. Bartilson, Department of Civil Engineering and Engineering Mechanics, 610 S.W. Mudd Building, New York, NY 10027, USA. Email: dtb2121@columbia.edu

Summary

Model updating is used to reduce error between measured structural responses and corresponding finite element (FE) model outputs, which allows accurate prediction of structural behavior in future analyses. In this work, reduced-order parametrizations of an underlying FE model are developed from singular value decomposition (SVD) of the sensitivity matrix, thereby improving efficiency and posedness in model updating. A deterministic error minimization scheme is combined with asymptotic Bayesian inference to provide optimal regularization with estimates for model evidence and parameter efficiency. Natural frequencies and mode shapes are targeted for updating in a small-scale example with simulated data and a full-scale example with real data. In both cases, SVD-based parametrization is shown to have as-good or better results than subset selection with very strong results on the full-scale model, as assessed by Bayes factor.

KEYWORDS:
finite element model updating, sensitivity-based parametrization, singular vector decomposition, Bayesian regularization, evidence-based model selection

1 | INTRODUCTION

Numerical models are essential tools for scientists and engineers to understand and predict the behavior of physical systems. In the context of structural engineering, finite element (FE) models are ubiquitously used to predict structural response and assess risk for existing structures under variable conditions and loadings. While numerical models should, ideally, provide exact predictions for their corresponding system, discrepancies always exist between measured behavior and model-predicted behavior. In FE modeling, these errors can be split into three categories[1]:

1. idealization errors, related to model simplification;
2. discretization errors, due to poor arrangement of the FE model; and
3. uncertainty in model parameters, such as mass densities, stiffnesses, and geometry.

The existence of modeling errors (and output discrepancies) indicates that the FE model is unreliable for predicting system behavior, diminishing its value for analysis.

The process of FE model updating seeks to correct the FE model, generally by modifying physical parameters to reduce discrepancy between measured and model-output data[1–3]. For structural applications, data often comes from vibration studies (which may provide natural frequencies, mode shapes, time histories, frequency-response functions, etc.) under forced or

ambient loadings. FE model updating has been successfully demonstrated on a multitude of civil structures[2,4], in addition to aerospace[1,5,6] and mechanical[7] structures.

It is important to note that FE model updating corrects model parameter errors (category 3), and generally cannot improve idealization or discretization errors (categories 1 and 2)[1]. When all three categories of FE modeling error are minimized, the model is said to be validated[1] and can give greater understanding of the current structural state, possibly for damage detection[8]. When the FE model exhibits idealization and/or discretization errors, the model is said to be inconsistent[1], but the updated model may still be valuable for response prediction within the measured frequency range.

FE model updating approaches can be divided into uncertainty quantification (UQ) methods and deterministic methods[9]. UQ methods naturally reflect measurement and model uncertainties in their results and can be further divided into probabilistic and non-probabilistic UQ methods. Probabilistic UQ methods estimate probability distributions functions for parameters and model outputs by drawing a large number of samples in the parameter space. Non-probabilistic UQ methods generally estimate intervals for parameters and outputs corresponding to upper and lower bounds of measured data. While non-probabilistic methods are generally less computationally-expensive than probabilistic methods, they are still orders-of-magnitude more expensive than deterministic methods and may be prohibitive for large models. Further detail on UQ methods in model updating is available by Simoen *et al.*[9].

Deterministic methods provide unique optimal solutions, generally by local or global minimization of a non-linear residual function. The sensitivity method[1] is a popular and intuitive local approach which iteratively minimizes a scalar objective function. The objective function is the sum of squared residual between measured and model-output data, making it easily extensible to many different sources or combinations of data. At each iteration, the non-linear residual function is linearized, forming the sensitivity matrix which intuitively captures the changes in model-outputs when modifying model parameters.

However, the sensitivity method is often applied to ill-posed model updating problems. Reparametization is one approach to improve posedness and efficiency by systematic selection of a new set of parameters to update the FE model. In this work, a novel parametrization technique is proposed based on the singular value decomposition (SVD) of the sensitivity matrix. Instead of selecting a reduced set of FE model parameters for updating, as in subset selection[10,11], linear combinations of FE model parameters are updated by single updating parameters. These linear combinations are defined by singular vectors. This is used to produce parametrizations which best represent the original sensitivity matrix with a reduced number of parameters. Alternatively, this can be closely related to subset selection by selecting singular vectors which best represent the residual.

Regularization is another approach to counter ill-posedness in model updating[1,12–14]. In general, regularization introduces additional equations to constrain the solution, such as equality constraints between nominally identical element material properties. More commonly, regularization is used to penalize large changes in updating parameters, representing a prior belief that parameter updates should be small. In this work, Bayesian regularization[15] is proposed for producing optimally regularized results through maximization of the model evidence. This method confers several benefits beyond parameter constraint, giving key insight into the support for competing models and parametrization efficiency, with strong ties to probabilistic methods.

In this work, the proposed SVD-based parametrization scheme is compared against subset selection in two FE model updating problems: a small-scale numerical example and a large-scale real example. In both cases, natural frequency and mode shape data for several dynamic modes are targeted for updating. Levenberg–Marquardt minimization with Bayesian regularization is implemented to provide deterministic model updating results, as well as estimates for model evidence and parameter efficiency. The paper is structured as follows. The objective function, comprising natural frequency and mode shape data, is detailed in Section 2. The objective function is then regularized using Bayesian inference in Section 3 with discussion of model evidence and parameter efficiency. In Section 4, the Levenberg–Marquardt minimization algorithm is briefly discussed in the context of regularization. Parametrization methods, including subset selection and the proposed SVD-based scheme, are detailed in Section 5. The proposed parametrization methods are first tested on a small-scale 2-dimensional truss with simulated data in Section 6, then on a full-scale large suspension bridge with real measurements in Section 7. Section 8 presents discussion of findings and conclusions.

2 | OBJECTIVE FUNCTION AND RESIDUAL DEFINITION

FE model updating begins with measured data from a structure, which can be written as a column vector of m components, $\tilde{\boldsymbol{z}}$. The corresponding column vector of m model outputs, $\boldsymbol{z}(\boldsymbol{\theta})$, is a function of the column vector of p updating parameters $\boldsymbol{\theta}$. A

common choice for the objective function is the weighted sum-of-square residual, E_r,

$$E_r = \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} \tag{1}$$

$$\boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \tag{2}$$

where $\boldsymbol{r}(\boldsymbol{\theta})$ is the residual vector and $\boldsymbol{W}_r$ is the residual weighting matrix. The residual weighting matrix should reflect the uncertainty in the measurements $\tilde{\boldsymbol{z}}$, giving the optimal weighting matrix as $\boldsymbol{W}_r = \boldsymbol{C}_{\tilde{z}}^{-1}$, where $\boldsymbol{C}_{\tilde{z}}$ is the covariance matrix of $\tilde{\boldsymbol{z}}$[16,17]. Since $\boldsymbol{C}_{\tilde{z}}$ is symmetric and positive semi-definite (SPSD), $\boldsymbol{W}_r$ is also SPSD. $\boldsymbol{W}_r$ and $\boldsymbol{C}_{\tilde{z}}$ are often diagonal or block-diagonal, representing statistical independence of measurements or sets of measurements, respectively.

When the measurement vector $\tilde{\boldsymbol{z}}$ contains disparate sources of data, it may be worthwhile to partition the problem. The examples studied in this paper utilize natural frequency and mode shape data, so $\boldsymbol{r}$ and E_r are partitioned into a natural frequency components ($\boldsymbol{r}_f$ and E_r^f) and mode shape components ($\boldsymbol{r}_s$ and E_r^s) with corresponding weighting matrices $\boldsymbol{W}_r^f$ and $\boldsymbol{W}_r^s$

$$E_r = \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_f^T & \boldsymbol{r}_s^T \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_r^f & \\ & \boldsymbol{W}_r^s \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_f \\ \boldsymbol{r}_s \end{bmatrix} \tag{3}$$

$$= \underbrace{\boldsymbol{r}_f^T \boldsymbol{W}_r^f \boldsymbol{r}_f}_{E_r^f} + \underbrace{\boldsymbol{r}_s^T \boldsymbol{W}_r^s \boldsymbol{r}_s}_{E_r^s} \tag{4}$$

The FE model-output natural frequencies and mode shapes are assumed to come from an undamped structural model, resulting in real-numbered outputs. The structural stiffness matrix $\boldsymbol{K}$ and mass matrix $\boldsymbol{M}$ are $N \times N$ symmetric real-valued matrices. For $j = 1, ..., N$, the j^{th} angular natural frequency ω_j (rad/s) and corresponding mass-normalized mode shape $\boldsymbol{\phi}_j$ satisfy the generalized eigenvalue problem $\boldsymbol{K}\boldsymbol{\phi}_j = \omega_j^2 \boldsymbol{M}\boldsymbol{\phi}_j$, where $\boldsymbol{\phi}_j^T \boldsymbol{M}\boldsymbol{\phi}_j = 1$. The equivalent natural frequency (Hz) is given by $f_j = \omega_j/(2\pi)$ and the unit-normalized mode shape is given $\boldsymbol{\psi}_j = \boldsymbol{\phi}_j/(\boldsymbol{\phi}_j^T \boldsymbol{\phi}_j)^{1/2}$.

2.1 | Natural frequency residual

The natural frequency residual column vector $\boldsymbol{r}_f$ is given by the difference between l measured natural frequencies $\tilde{\boldsymbol{f}}$ and corresponding model-output natural frequencies $\tilde{\boldsymbol{f}}(\boldsymbol{\theta})$

$$\boldsymbol{r}_f = \tilde{\boldsymbol{z}}_f - \boldsymbol{z}_f(\boldsymbol{\theta}) = \tilde{\boldsymbol{f}} - \boldsymbol{f}(\boldsymbol{\theta}) \tag{5}$$

It is essential to perform mode pairing[3] to ensure that measured and model-output modes are correctly correlated. Mode pairing generally pairs a model-output mode with the measured mode which maximizes the Modal Assurance Criterion (MAC)[18], or equivalently, minimizes the angle between their mode shapes.

When $\boldsymbol{W}_r^f$ is diagonal (i.e. natural frequency measurements are statistically independent) then the weighted sum-of-square natural frequency residual is

$$E_r^f = \boldsymbol{r}_f^T \boldsymbol{W}_r^f \boldsymbol{r}_f = \sum_{j=1}^{l} w_{rj}^f (\tilde{f}_j - f_j(\boldsymbol{\theta}))^2 \tag{6}$$

2.2 | Mode shape residual

The mode shape residual column vector $\boldsymbol{r}_s$ is given by the difference between the concatenated set of l measured mode shapes $\tilde{\boldsymbol{z}}_s = [\tilde{\boldsymbol{\psi}}_1^T \ \cdots \ \tilde{\boldsymbol{\psi}}_l^T]^T$ and the corresponding concatenated model-output (unit-normalized) mode shapes $\boldsymbol{z}_s(\boldsymbol{\theta}) = [\mu_1 \boldsymbol{\psi}_1^T(\boldsymbol{\theta}) \ \cdots \ \mu_l \boldsymbol{\psi}_l^T(\boldsymbol{\theta})]^T$. The modal scaling factor, μ_j, is used to minimize the difference between corresponding measured and model-output mode shapes, $\tilde{\boldsymbol{\psi}}$ and $\boldsymbol{\psi}_j(\boldsymbol{\theta})$ in the least-squares sense[18]

$$\mu_j = \tilde{\boldsymbol{\psi}}_j^T \boldsymbol{\psi}_j(\boldsymbol{\theta}) \tag{7}$$

Each measured mode shape $\tilde{\boldsymbol{\psi}}_j$ and model-output mode shape $\boldsymbol{\psi}_j(\boldsymbol{\theta})$ must have measurements corresponding to the same n degrees of freedom (DoFs), making $\boldsymbol{r}_s$ a column vector of nl elements

$$\boldsymbol{r}_s = \tilde{\boldsymbol{z}}_s - \boldsymbol{z}_s(\boldsymbol{\theta}) \tag{8}$$

If $\boldsymbol{W}_r^s$ is diagonal and decomposable into a scalar multiple of $\boldsymbol{I}_n$ for each mode ($[\boldsymbol{W}_r^s]_j = w_{rj}^s \boldsymbol{I}_n$), then the sum-of-squared mode shape residual E_r^s can be written

$$E_r^s = \boldsymbol{r}_s^T \boldsymbol{W}_r^s \boldsymbol{r}_s = \sum_{j=1}^{l} w_{rj}^s \boldsymbol{r}_{sj}^T \boldsymbol{r}_{sj} = \sum_{j=1}^{l} w_{rj}^s \|\tilde{\boldsymbol{\psi}}_j - \mu_j \boldsymbol{\psi}_j(\boldsymbol{\theta})\|_2^2 \tag{9}$$

where $\boldsymbol{r}_{sj} = \tilde{\boldsymbol{\psi}}_j - \mu_j \boldsymbol{\psi}_j(\boldsymbol{\theta})$ is the residual for mode shape j and $\| \ \|_2$ is the ℓ^2 norm. Equation 9 can be rewritten in a more familiar form as

$$E_r^s = \sum_{j=1}^{l} w_{rj}^s \left[1 - \mathrm{MAC}(\tilde{\boldsymbol{\psi}}_j, \boldsymbol{\psi}_j(\boldsymbol{\theta}))\right] \tag{10}$$

where MAC is defined as[18]

$$\mathrm{MAC}(\tilde{\boldsymbol{\psi}}_j, \boldsymbol{\psi}_j) = \frac{(\tilde{\boldsymbol{\psi}}_j^T \boldsymbol{\psi}_j)^2}{\tilde{\boldsymbol{\psi}}_j^T \tilde{\boldsymbol{\psi}}_j \cdot \boldsymbol{\psi}_j^T \boldsymbol{\psi}_j} \tag{11}$$

2.3 | Partitioned objective function

Equations 6 and 10 can be combined into Equation 3 to give

$$E_r = E_r^f + E_r^s = \sum_{j=1}^{l} w_{rj}^f (\tilde{f}_j - f_j(\boldsymbol{\theta}))^2 + \sum_{j=1}^{l} w_{rj}^s \left[1 - \mathrm{MAC}(\tilde{\boldsymbol{\psi}}_j, \boldsymbol{\psi}_j(\boldsymbol{\theta}))\right] \tag{12}$$

As noted before, the residual weighting matrix $\boldsymbol{W}_r$ should be equal to the inverse of the measurement covariance matrix $\boldsymbol{C}_{\tilde{z}}$. The measurement covariance model used in the included examples uses a diagonal covariance matrix. The standard deviation of each natural frequency measurement j is assumed to be a scalar (c_f) multiple of the measured natural frequency $\tilde{f}_j$, giving $w_{rj}^f = (c_f \tilde{f}_j)^{-2}$. Similarly, the standard deviation for each component of measured mode j is assumed to be equivalent to a scalar (c_s) multiplied by the standard deviation of measured mode shape j, giving $w_{rj}^s = (c_s \ \mathrm{std}(\tilde{\boldsymbol{\psi}}_j))^{-2}$. Inserting these results into Equation 12 gives

$$E_r = \frac{1}{c_f^2} \sum_{j=1}^{l} (1 - f_j(\boldsymbol{\theta})/\tilde{f}_j)^2 + \frac{1}{c_s^2} \sum_{j=1}^{l} \frac{1}{\mathrm{var}(\tilde{\boldsymbol{\psi}}_j)} \left[1 - \mathrm{MAC}(\tilde{\boldsymbol{\psi}}_j, \boldsymbol{\psi}_j(\boldsymbol{\theta}))\right] \tag{13}$$

which provides further insight into relative weighting of the natural frequency and mode shape error components.

3 | MODEL EVIDENCE ESTIMATION AND BAYESIAN REGULARIZATION

While the general goal of FE model updating is to optimize the objective function, such as E_r in Equation 1, this often results in an ill-posed problem and/or an overfitted solution[1,12–14]. Ill-posedness may develop when there are more updating parameters than measurements (underdetermined), leading to non-unique solutions. Overfitting occurs when the model updating solution fits to the measurement noise at the expense of generality, reducing its utility for prediction.

Both of these problems can be ameliorated through regularization, which adds an additional term to the objective function. This increases the number of equations, reducing ill-posedness, and penalizes overly large updating parameter values, reducing overfitting. Equation 1 is modified to include the regularization term $E_\theta = \boldsymbol{\theta}^T \boldsymbol{W}_\theta \boldsymbol{\theta}$.

$$F(\boldsymbol{\theta}) = \beta E_r + \alpha E_\theta = \beta \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} + \alpha \boldsymbol{\theta}^T \boldsymbol{W}_\theta \boldsymbol{\theta} \tag{14}$$

The regularization parameters α and β control the relative importance of reducing residual against reducing the amount of parameter modification. The Bayesian approach to regularization[15] treats α and β as random variables. The optimal values for the regularizing parameters maximize the model evidence, which is a key component of Bayesian analysis.

3.1 | Model evidence estimation

Given a model $\mathcal{M}_j$ (parametrization of an FE model) and values of α and β, the posterior probability of the updating parameters can be written using Bayes' rule:

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}_j) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \beta, \mathcal{M}_j) P(\boldsymbol{\theta}|\alpha, \mathcal{M}_j)}{P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_j)} \tag{15}$$

in which $P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \beta, \mathcal{M}_j)$ is the likelihood function of the measured data $\tilde{\boldsymbol{z}}$, $P(\boldsymbol{\theta}|\alpha, \mathcal{M}_j)$ is the prior probability density function (PDF) of $\boldsymbol{\theta}$, and $P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_j)$ is a normalization term also known as the evidence for model $\mathcal{M}_j$.

The likelihood function is proportional to the probability of the data $\tilde{\boldsymbol{z}}$ given $\boldsymbol{\theta}$ for a model $\mathcal{M}_j$. If the noise in $\tilde{\boldsymbol{z}}$ is assumed to be additive, zero-mean, and Gaussian, with covariance $\boldsymbol{C}_{\tilde{z}} = [2\beta \boldsymbol{W}_r]^{-1}$ then the likelihood is written

$$P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \beta, \mathcal{M}_j) = \frac{e^{-\beta E_r}}{Z_{\tilde{z}}(\beta)}; \qquad Z_{\tilde{z}}(\beta) = \pi^{m/2} \det(\beta \boldsymbol{W}_r)^{-1/2} \tag{16}$$

However, the likelihood is not a PDF and $Z_{\tilde{z}}(\beta)$ should not be viewed as the integral of $e^{-\beta E_r}$ over $\boldsymbol{\theta}$. The prior distribution for $\boldsymbol{\theta}$ is assumed to be a zero-mean Gaussian with covariance $\boldsymbol{C}_\theta = [2\alpha \boldsymbol{W}_\theta]^{-1}$, giving

$$P(\boldsymbol{\theta}|\alpha, \mathcal{M}_j) = \frac{e^{-\alpha E_\theta}}{Z_\theta(\alpha)}; \qquad Z_\theta(\alpha) = \pi^{p/2} \det(\alpha \boldsymbol{W}_\theta)^{-1/2} \tag{17}$$

Substituting Equations 16 and 17 into Equation 15 simplifies to

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}_j) = \frac{e^{-F(\boldsymbol{\theta})}}{Z_F(\alpha, \beta)} \tag{18}$$

where $Z_F(\alpha, \beta)$ is a normalization term. This can be estimated by expanding the regularized objective function $F(\boldsymbol{\theta})$ (Equation 14) using a Taylor series truncated after the quadratic term[15]. $F(\boldsymbol{\theta})$ is estimated as

$$F(\boldsymbol{\theta}) \approx F(\boldsymbol{\theta}_{\text{MP}}) + (\boldsymbol{\theta} - \boldsymbol{\theta}_{\text{MP}})^T \boldsymbol{H}_{\text{MP}} (\boldsymbol{\theta} - \boldsymbol{\theta}_{\text{MP}}) \tag{19}$$

The expansion is performed about the minimum point of F, $\boldsymbol{\theta}_{\text{MP}}$, which is the maximum of the posterior probability. Therefore the evaluated gradient $\{\nabla F\}(\boldsymbol{\theta}_{\text{MP}})$ is zero, where $\nabla = \partial/\partial\boldsymbol{\theta}$. $\boldsymbol{H}_{\text{MP}}$ is the Hessian of $F(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}_{\text{MP}}$, $\boldsymbol{H}_{\text{MP}} = \{\nabla\nabla F\}(\boldsymbol{\theta}_{\text{MP}})$. $Z_F(\alpha, \beta)$ is then evaluated as the Gaussian integral, using Laplace's method[15]

$$Z_F(\alpha, \beta) = \int e^{-F(\boldsymbol{\theta})} d\boldsymbol{\theta} \approx e^{-F(\boldsymbol{\theta}_{\text{MP}})} (2\pi)^{p/2} \det(\boldsymbol{H}_{\text{MP}})^{-1/2} \tag{20}$$

Rewriting Equation 15 to find the evidence and substituting in Equations 16-18 gives

$$P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_j) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}, \beta, \mathcal{M}_j) P(\boldsymbol{\theta}|\alpha, \mathcal{M}_j)}{P(\boldsymbol{\theta}|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}_j)} = \frac{Z_F(\alpha, \beta)}{Z_{\tilde{z}}(\beta) Z_\theta(\alpha)} \tag{21}$$

Evaluating the log-evidence using the normalization terms yields[15]

$$\log P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_j) = \underbrace{-\beta E_r^{\text{MP}} + \frac{1}{2} \log\det\left((\beta/\pi)\boldsymbol{W}_r\right)}_{\text{log likelihood}} \underbrace{-\alpha E_\theta^{\text{MP}} + \frac{1}{2} \log\det\left(\boldsymbol{H}_{\text{MP}}^{-1}[2\alpha \boldsymbol{W}_\theta]\right)}_{\text{log Occam factor}} \tag{22}$$

which can be separated into terms related to the log likelihood and the log Occam factor. The likelihood is maximized by reducing the sum-of-square residual, E_r, which favors complex models that may overfit the data. The Occam factor penalizes overly complex models, representing Occam's principle that simpler models are preferable[9,15]. The first Occam term penalizes overly large parameter values, while the second term is the ratio of the prior curvature or volume relative to the posterior curvature or volume, which penalizes overly large prior parameter spaces. The second Occam term also reflects the robustness of the model[19], penalizing highly peaked posteriors which imply poor model generalization.

3.2 | Optimal regularization

Estimating the log evidence using Laplace's method, generally referred to as an "asymptotic approach", is well-known in model updating[9,20,21]. However, within these works, the prior PDF was fixed. In general, the prior PDF of updating parameters for a given model $P(\boldsymbol{\theta}|\mathcal{M}_j)$ is mostly unknown and uninformed assumptions are made. The work done by MacKay[15] provides a method for determining the "width" of a Gaussian prior, α, to maximize the log evidence, simultaneously delivering optimal regularization. To the authors' knowledge, this approach has not been previously used for evidence estimation in FE model updating and presents a step forward for deterministic model updating. Previous work by the authors[22] implemented Bayesian regularization, but also optimized β, which is inappropriate for evidence estimation and model evidence comparison, as will be discussed below.

The optimal regularizing constant α is determined by maximizing the log evidence in Equation 22 with respect to α, giving

$$\alpha = \frac{\gamma}{2E_\theta^{\text{MP}}} \tag{23}$$

where γ is called the "effective number of parameters"[15]. The Hessian, $\boldsymbol{H}(\boldsymbol{\theta})$, can be separated as

$$\boldsymbol{H}(\boldsymbol{\theta}) = \{\nabla\nabla F\}(\boldsymbol{\theta}) = \beta \boldsymbol{B}(\boldsymbol{\theta}) + \alpha \boldsymbol{A} \tag{24}$$

where $\boldsymbol{B}(\boldsymbol{\theta}) = \{\nabla\nabla E_r\}(\boldsymbol{\theta})$ and $\boldsymbol{A} = \nabla\nabla E_\theta = 2\boldsymbol{W}_\theta$. This allows γ to be written using the trace operator or a sum of eigenvalues

$$\gamma = p - 2\alpha \operatorname{tr}(\boldsymbol{H}_{\mathrm{MP}}^{-1}\boldsymbol{W}_\theta) = \sum_{j=1}^{p} \frac{\beta\lambda_j}{\beta\lambda_j + \alpha} \tag{25}$$

where λ_j is the j^{th} eigenvalue of $[\boldsymbol{W}_\theta^{-1}\boldsymbol{B}_{\mathrm{MP}}]$.

$\gamma = 0$ implies the estimated posterior curvature (Hessian) is identical to the prior curvature, representing a null updating result and ineffective parametrization. $\gamma \to p$ implies the posterior curvature is infinitely greater than the prior curvature, such that the prior has no impact relative to the likelihood during updating and the result is the maximum likelihood estimate. MacKay[15] suggests that $\gamma = p/2$ is a reasonable result for many updating problems, but $\gamma \to p$ is desirable because it suggests that the updating result is controlled by the data rather than by regularization.

At this stage, the optimal value of β could be found by evidence maximization, as in $\beta = (m - \gamma)/(2E_r^{\mathrm{MP}})$[15], as in previous work[22], but this has several disadvantages. Foremost, the likelihood function will no longer be model-independent since β will depend on the optimized model error E_r^{MP}. Additionally, this will fix the regularized objective function at $F(\boldsymbol{\theta}_{\mathrm{MP}}) = m/2$, which causes difficulty for model evidence comparison. The approach adopted in this work evaluates $\boldsymbol{W}_r$ as the inverse of the measurement covariance matrix, $\boldsymbol{W}_r = \boldsymbol{C}_{\tilde{z}}^{-1}$ with β fixed to the value of $1/2$. If the measurement covariance is unknown, $\boldsymbol{W}_r$ can reflect the relative importance of each residual term for reduction, but evidence estimates should be analyzed cautiously since β will be arbitrary.

3.3 | Model comparison via Bayes factor

The relative evidence, or Bayes factor[23], can be used to evaluate the strength of support for competing models. Given the evidence (approximate or exact) for two models $\mathcal{M}_j$ and $\mathcal{M}_k$ with equally likely prior probabilities, the Bayes factor is

$$B_{jk} = \frac{P(\tilde{\boldsymbol{z}}|\mathcal{M}_j)}{P(\tilde{\boldsymbol{z}}|\mathcal{M}_k)} \tag{26}$$

which gives the support for using $\mathcal{M}_j$ instead of $\mathcal{M}_k$. Note that this form drops dependence on the regularizing constants α and β. These parameters can either be marginalized (integrating over all values), or more reasonably, the model comparison can be performed using the optimal regularizing constants[15]. Kass and Raftery[23] provided a widely used set of criteria for interpreting the Bayes factor, given in Table 1. Note that $\log B_{jk} = -\log B_{kj}$, so negative results can be interpreted as support for $\mathcal{M}_k$.

TABLE 1 Interpretation of Bayes factors, adapted from Kass and Raftery[23]

$2\log B_{jk}$	B_{jk}	Evidence against $\mathcal{M}_k$
0-2	1-3	Not worth more than a bare mention
2-6	3-20	Positive
6-10	20-150	Strong
>10	>150	Very strong

4 | LEVENBERG–MARQUARDT MINIMIZATION ALGORITHM

With a well-defined objective function (Section 2) and regularization (Section 3), FE model updating can proceed by determining an optimal model through minimization of the regularized objective function. The parameter values at iteration i, $\boldsymbol{\theta}_i$, are updated by $\Delta\boldsymbol{\theta}_i$ to give the values at the next iteration:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i \tag{27}$$

The goal is then to find the parameter update $\Delta\boldsymbol{\theta}_i$ such that the objective function $F(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i)$ is minimized. Using the notation of $\boldsymbol{r}_i = \boldsymbol{r}(\boldsymbol{\theta}_i)$, then the updated residual can be estimated by its linearization

$$\boldsymbol{r}(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i) \approx \boldsymbol{r}_i + \boldsymbol{J}_i \Delta\boldsymbol{\theta}_i \tag{28}$$

where $\boldsymbol{J}_i$ is the Jacobian of $\boldsymbol{r}$ evaluated at $\boldsymbol{\theta}_i$, $\boldsymbol{J}_i = \{\nabla \boldsymbol{r}\}(\boldsymbol{\theta}_i)$. Linearization of the residual forms the basis of the sensitivity method[1], where $\boldsymbol{J}_i$ is also called the sensitivity matrix and its columns represent the sensitivity of the residual to changes in each parameter. The Jacobian can be estimated numerically using a finite-difference scheme, or it can be calculated analytically. For FE model updating of natural frequencies and mode shapes, the analytical Jacobian can be assembled column-wise[24]. When the linearized residual is used and $F(\boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i)$ is minimized with respect to $\Delta\boldsymbol{\theta}_i$, then the parameter update is

$$\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i] \tag{29}$$

which represents the Gauss–Newton algorithm. The terms $2(\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i)$ are the gradient of F at $\boldsymbol{\theta}_i$, $\{\nabla F\}(\boldsymbol{\theta}_i)$. $\boldsymbol{H}_i$ is the Hessian of F at $\boldsymbol{\theta}_i$, which is approximated as

$$\boldsymbol{H}_i = \{\nabla\nabla F\}(\boldsymbol{\theta}_i) \approx 2[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i + \alpha \boldsymbol{W}_\theta] \tag{30}$$

Comparing this to Equation 24 indicates that the approximate Hessian of E_r is $\{\nabla\nabla E_r\}(\boldsymbol{\theta}_i) = \boldsymbol{B}(\boldsymbol{\theta}_i) \approx \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i$.

The Gauss–Newton algorithm is transformed into the more robust Levenberg–Marquardt algorithm[25,26] by adding a damping term λ to the diagonal of $\boldsymbol{H}$, giving the Levenberg–Marquardt parameter update

$$\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i + 2\lambda \boldsymbol{I}]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i] \tag{31}$$

This trust region approach collapses to the Gauss–Newton algorithm when $\lambda \to 0$, and to the gradient-descent algorithm (with infinitesimal step size) when $\lambda \to \infty$. λ is controlled by the multiplicative process given by Marquardt[26]. The utilized Levenberg–Marquardt algorithm is described in Algorithm 1, including the scheme for iteratively evaluating the hyperparameters λ, α, and γ.

Algorithm 1 Pseudocode for Levenberg–Marquardt minimization with Bayesian regularization

Input: Regularized objective function $F(\boldsymbol{\theta}) = \beta E_r + \alpha E_\theta$ to be minimized, model $\mathcal{M}_j$

Output: Optimal parameters $\boldsymbol{\theta}_{\text{MP}}$, effective number of parameters γ, log evidence estimate $\log P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_j)$

initialization Set $\boldsymbol{\theta}_0$, $\alpha = 0.5$, $\beta = 0.5$, $\lambda = 0.01$, $\upsilon = 10$, $i = 0$
while not converged **do**
  Compute residual $\boldsymbol{r}_i$, Jacobian $\boldsymbol{J}_i$, and approximate Hessian $\boldsymbol{H}_i = 2[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{J}_i + \alpha \boldsymbol{W}_\theta]$
  Compute parameter update $\Delta\boldsymbol{\theta}_i = -2[\boldsymbol{H}_i + 2\lambda \boldsymbol{I}]^{-1}[\beta \boldsymbol{J}_i^T \boldsymbol{W}_r \boldsymbol{r}_i + \alpha \boldsymbol{W}_\theta \boldsymbol{\theta}_i]$
  Evaluate trial parameters $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \Delta\boldsymbol{\theta}_i$
  if Objective value increased $F(\boldsymbol{\theta}_{i+1}) > F(\boldsymbol{\theta}_i)$ **then**
    Increase damping term $\lambda \leftarrow \lambda \cdot \upsilon$
    Go back to parameter update computation step (4)
  else Decrease damping term $\lambda \leftarrow \lambda / \upsilon$
  end if
  Compute effective number of parameters* $\gamma = p - 2\alpha \operatorname{tr}(\boldsymbol{H}_i^{-1} \boldsymbol{W}_\theta)$
  Reestimate regularization parameter $\alpha = \gamma / (2E_\theta(\boldsymbol{\theta}_{i+1}))$
  $i \leftarrow i + 1$
end while
Estimate evidence for updated model $\mathcal{M}_j$ by $\log P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_j)$ (Equation 22)

* Note: γ is only meaningful at a converged solution, $\boldsymbol{\theta}_{\text{MP}}$

5 | MODEL PARAMETRIZATION

Parametrization is a crucial part of FE model updating. Even a small model can easily have thousands of possible parametrizations among combinations of material properties, geometry, and external conditions. In general, parametrizations should satisfy three requirements[3]:

1. ill-posedness should be avoided by limiting the number of parameters,
2. parameters should reflect model uncertainty, and
3. FE model-outputs should be sensitive to chosen parameters.

Fulfilling these requirements generally requires physical understanding of the FE model. Mottershead *et al.*[7] studied several parametrizations of a frame joint, including geometric and element-eigenvalue modifications. While these parametrizations may be more effective for reducing modeling error, they are often difficult to justify physically. Other methods directly use FE model parameters, but select a reduced number of updating parameters to alleviate ill-posedness. This has been accomplished through subset selection[10,11], which is described in Section 5.1. Smith and Hernandez[27] recently proposed LASSO for combined subset selection and ℓ^1 regularization which is appropriate for sparse model errors. Alternative methods include parameter clustering, in which all FE model parameters are retained and grouped into clusters (substructures) based on sensitivity considerations[1,5,22,28,29]. Each cluster is then updated by a single parameter, giving a reduced parametrization.

The simplest parametrization is the vector of uncertain FE physical properties $\boldsymbol{x}$, such as mass densities, Young's moduli, geometry, cross-sectional properties, etc. Since $\boldsymbol{x}$ may contain parameters which differ by several orders of magnitude, updating $\boldsymbol{x}$ directly may result in a poorly-scaled Jacobian matrix. The use of physical parameter modification parameters $\boldsymbol{\delta}$ results in comparably-sized updating parameters and improved condition of $\boldsymbol{J}$. Then the e^{th} updated FE physical properties can be written

$$x_e = x_e^0(1 - \delta_e) \tag{32}$$

where x_e^0 is the initial value of x_e. The FE model physical properties utilized in this work include the Young's modulus and mass density for each element (or substructure), e, out of a total number n_{el}. Thus, each element mass matrix ($\boldsymbol{M}_e$) and stiffness matrix ($\boldsymbol{K}_e$) is modified prior to summation into the global mass ($\boldsymbol{M}$) and stiffness ($\boldsymbol{K}$) matrices, similar to other work[1,5,22,28,29]:

$$\boldsymbol{M}(\boldsymbol{\delta}) = \sum_{e=1}^{n_{el}} \boldsymbol{M}_e(1 - \delta_e^m) = \boldsymbol{M}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{M}_e\delta_e^m \tag{33}$$

$$\boldsymbol{K}(\boldsymbol{\delta}) = \sum_{e=1}^{n_{el}} \boldsymbol{K}_e(1 - \delta_e^k) = \boldsymbol{K}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{K}_e\delta_e^k \tag{34}$$

where $\boldsymbol{M}_0$ and $\boldsymbol{K}_0$ are the initial global stiffness and mass matrices, respectively. δ_e^m and δ_e^k are the stiffness and mass physical parameter modifications for element e, making $\boldsymbol{\delta}$ a vector comprising these $d = 2n_{el}$ components.

Parameterizing the model updating problem using $\boldsymbol{\theta} = \boldsymbol{\delta}$ is simple, but it is intractable for FE models with thousands of uncertain physical parameters. Not only does the complexity of computing the Jacobian increase linearly with the number of updating parameters, but the Jacobian is increasingly likely to exhibit ill-conditioning[5]. These problems can be resolved by intelligently reparameterizing.

For the methods covered here, it is possible to write a linear transformation between the selected parametrization $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$, which will be called the natural parametrization:

$$\boldsymbol{\delta} = \boldsymbol{T}\boldsymbol{\theta} \tag{35}$$

where $\boldsymbol{T}$ is the $d \times p$ transformation matrix from $\boldsymbol{\theta}$ to $\boldsymbol{\delta}$. This notation confers several insights. Namely, the Jacobian with respect to $\boldsymbol{\theta}$, $\boldsymbol{J}'$, can be related to the Jacobian with respect to $\boldsymbol{\delta}$, $\boldsymbol{J}$:

$$\boldsymbol{J}' = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\delta}}\frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{\theta}} = \boldsymbol{J}\boldsymbol{T} \tag{36}$$

Using the condition that E_θ should be invariant under reparametrization, then $\boldsymbol{W}_\theta$ can be related to $\boldsymbol{W}_\delta$ by

$$\boldsymbol{W}_\theta = \boldsymbol{T}^T\boldsymbol{W}_\delta\boldsymbol{T} \tag{37}$$

This relation is useful for generating consistently-defined $\boldsymbol{W}_\theta$ when performing model comparison, such that each $\boldsymbol{W}_\theta$ reflects the uncertainty in the underlying FE model parameters.

5.1 | Parameter subset selection

The parameter subset selection method[30] chooses a subset of parameters by testing candidate parameter groups of fixed size. The parameter subset which results in minimum residual is chosen. Since testing all possible parameter subsets is intractable for practical problems, greedy methods are typically used. One such approach, forward selection, was applied to FE model updating by Lallement and Piranda[10] and by Friswell *et al.*[11].

Consider a natural FE model parametrization using the column vector $\boldsymbol{\delta}$ with d components. Model updating using this parametrization would find $\Delta\boldsymbol{\delta}$ which minimizes E_r. This can be written as the ℓ^2 norm of the weighted residual:

$$E_r(\boldsymbol{\delta} + \Delta\boldsymbol{\delta}) = \|\boldsymbol{q} + \boldsymbol{G}\Delta\boldsymbol{\delta}\|_2^2 \tag{38}$$

where $\boldsymbol{G}$ and $\boldsymbol{q}$ are the weighted Jacobian and residual, respectively.

$$\boldsymbol{G} = \boldsymbol{W}_r^{1/2}\boldsymbol{J}; \qquad \boldsymbol{q} = \boldsymbol{W}_r^{1/2}\boldsymbol{r} \tag{39}$$

Forward subset selection chooses $p < d$ elements of $\boldsymbol{\delta}$ with corresponding columns of $\boldsymbol{G} = [\boldsymbol{g}_1 \ \cdots \ \boldsymbol{g}_d]$ which minimize E_r[10,11,30]. The iterative process begins by identifying the parameter δ_a (and corresponding column of $\boldsymbol{G}$) which minimizes E_r at the initial state $\boldsymbol{\delta}_0$

$$\boldsymbol{g}_a = \arg\min_{\boldsymbol{g}_j \in \boldsymbol{G}} \|\boldsymbol{q} + \boldsymbol{g}_j\widehat{\Delta\delta_j}\|_2^2 \tag{40}$$

where $\widehat{\Delta\delta_j}$ is the least-squares estimate of the j^{th} parameter, $\widehat{\Delta\delta_j} = -\boldsymbol{g}_j^T\boldsymbol{q}/\boldsymbol{g}_j^T\boldsymbol{g}_j$. This is equivalent to identifying the parameter sensitivity which has the minimum angle with the weighted residual at the initial state. Then the columns of $\boldsymbol{G}$ and the weighted residual $\boldsymbol{q}$ are replaced by

$$\boldsymbol{g}_j \leftarrow \boldsymbol{g}_j - \boldsymbol{g}_a(\boldsymbol{g}_a^T\boldsymbol{g}_j/\boldsymbol{g}_j^T\boldsymbol{g}_j); \qquad \boldsymbol{q} \leftarrow \boldsymbol{q} + \boldsymbol{g}_a\widehat{\Delta\delta_a} \tag{41}$$

Thus $\boldsymbol{q}$ and the remaining columns of $\boldsymbol{G}$ are orthogonal to $\boldsymbol{g}_a$, and the process is iterated until p parameters are selected. The transformation matrix can be written

$$T_{ak} = \begin{cases} 1 & \delta_a \text{ selected in iteration } k;\ \delta_a \text{ updated by } \theta_k \\ 0 & \text{else} \end{cases} \tag{42}$$

Each of the p columns of $\boldsymbol{T}$ is a unique member of the standard basis of $\mathbb{R}^d$, thus $\boldsymbol{T}$ is orthogonal ($\boldsymbol{T}^T\boldsymbol{T} = \boldsymbol{I}_p$).

Low-sensitivity parameters may be excluded from subset selection[1,3,5]. If the degree-of-sensitivity of a parameter is given by the ℓ^2 norm of its sensitivity vector, $\|\boldsymbol{g}_j\|_2^2 = \boldsymbol{g}_j^T\boldsymbol{g}_j$, it is clear that low-sensitivity parameters will tend to require large update terms $\widehat{\Delta\delta_j}$. This poses difficulties, since deterministic model updating generally depends on initial model parameters being close to global optimal values. Additionally, if model parameters are related to physical quantities, then large parameter updates may go beyond physically-plausible bounds (e.g. negative mass).

Unfortunately, there is no consensus for what constitutes low sensitivity. This could be taken as a relative term, i.e. $\|\boldsymbol{g}_b\|^2 \ll \|\boldsymbol{g}_j\|^2\ \forall j \neq b$, but this doesn't guarantee a limit on the parameter update. Low sensitivity could instead be tied to the estimate for the parameter update, $\widehat{\Delta\delta_j}$, at the first iteration, but this may be a poor estimate for the parameter update since $\widehat{\Delta\delta_j}$ comes from a one-dimensional optimization instead of the true multi-dimensional optimization. Alternatively, any parameters which result in unacceptably-sized parameter updates could be removed (and possibly replaced) after-the-fact, but there is no guarantee that the new parametrization will result in a properly bounded set of parameter updates. Since there isn't a clear method for removing low sensitivity parameters, no parameters are excluded from analysis in this work.

5.2 | SVD-based parametrization

The parametrization method proposed in this paper shares many similarities with subset selection and SVD. The subset selection method seeks a reduced set of parameters from $\boldsymbol{\delta}$ with residual gradients (columns of $\boldsymbol{G}$) which best represent the residual using an orthogonalization process. The proposed method also uses an orthogonalization process, but instead of selecting a subset from $\boldsymbol{\delta}$, it forms linear combinations of parameters using SVD and updates $\boldsymbol{\delta}$ along these vectors. While SVD is not new to FE model updating, it has usually been used for regularization[12]. Recently, Silva *et al.*[31] selected parameters based on contribution to the output covariance matrix, closely related to the SVD of the Jacobian matrix. Note that this was used for subset selection, while the proposed approach forms linear combinations of existing parameters and uses very different logic.

The proposed parametrization method begins by considering the SVD of the weighted Jacobian, $\boldsymbol{G}$, which is $m \times d$ with rank $r \leq \min(m, d)$. In general, the process of reparametrization is used for underdetermined problems, such that $m < d$. The SVD

of $\boldsymbol{G}$ is[32]

$$\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \sum_{j=1}^{r} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^T \tag{43}$$

where $\boldsymbol{U}$ is a $m \times m$ orthogonal matrix of the left singular vectors, $\boldsymbol{U} = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_m]$, and $\boldsymbol{V}$ is a $d \times d$ orthogonal matrix of the right singular vectors, $\boldsymbol{V} = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_d]$. $\boldsymbol{\Sigma}$ is an $m \times d$ matrix with singular values $[\sigma_1 \cdots \sigma_r]$ along its main diagonal, arranged in descending order. There are at most $r = \min(m, d)$ non-zero singular values and associated left and right singular vectors. Any singular vectors which correspond to zero singular values are outside of the column or row space of $\boldsymbol{G}$.

Using SVD, the parameter update which minimizes Equation 38 is given by the sum

$$\Delta\boldsymbol{\delta} = \sum_{j=1}^{r} \frac{\boldsymbol{u}_j^T \boldsymbol{q}}{\sigma_j} \boldsymbol{v}_j \tag{44}$$

An approximate solution (or, equivalently, approximation to $\boldsymbol{G}$) can be obtained by truncating the sums in Equations 43 and 44 using only $p < r$ singular vectors. Using this logic, a set C of p singular values and right singular vectors is retained, and the model is updated along the right singular vectors in this set, such that the transformation matrix is $\boldsymbol{T} = \boldsymbol{V}_C$. In other words, each chosen singular vector defines a linear combination of model parameters in $\boldsymbol{\delta}$ which is updated by a single updating parameter in $\boldsymbol{\theta}$, such that

$$\boldsymbol{\delta} = \sum_{j \in C} \theta_j \boldsymbol{v}_j \tag{45}$$

Since each of the p right-singular vectors in $\boldsymbol{V}_C$ is a column vector of size d, then $\boldsymbol{T}$ is $d \times p$ orthogonal matrix (i.e. $\boldsymbol{T}^T\boldsymbol{T} = \boldsymbol{I}_p$).

The resulting parametrization $\boldsymbol{\theta}$ has a weighted Jacobian matrix given by $\boldsymbol{G}'$, as in Equation 36. Since the columns of $\boldsymbol{V}$ are orthogonal and $\boldsymbol{V}_C$ is a subset of these columns, this is equivalent to writing

$$\boldsymbol{G}' = \boldsymbol{G}\boldsymbol{V}_C = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{V}_C = \sum_{j \in C} \sigma_j \boldsymbol{u}_j \boldsymbol{e}_j^T \tag{46}$$

where $\boldsymbol{e}_j$ is the j^{th} standard basis vector of d-space. Therefore, the singular values of $\boldsymbol{G}'$ are a subset of the singular values of $\boldsymbol{G}$.

5.2.1 | Parametrization to maximize singular values

The most critical issue, then, is to choose the set of p singular vectors to retain, C. The first proposed parametrization delivers the best approximation to $\boldsymbol{G}$ by retaining the largest singular values, which correspond to the first p singular vectors[32]. This set can be defined recursively such that the set of p retained singular values σ_C is greater than all other singular values σ_b not in σ_C:

$$\boldsymbol{T} = \boldsymbol{V}_C; \qquad C = \{j \mid \sigma_j > \sigma_b;\ \forall b \notin C\} \tag{47}$$

This method improves the condition of $\boldsymbol{G}'$ (when using the reduced parametrization) since the range of singular values is reduced. Under a set of conditions, it can be shown that this choice will maximize the effective number of parameters, γ. Starting with a natural FE model parametrization $\boldsymbol{\delta}$ with d parameters, then γ is given by Equation 25:

$$\gamma = \sum_{j=1}^{d} \frac{\beta\lambda_j}{\beta\lambda_j + \alpha} \tag{48}$$

where α and β are regularization parameters which are assumed to be constant. λ_j is the j^{th} eigenvalue of $[\boldsymbol{W}_\delta^{-1}\boldsymbol{B}_{\text{MP}}]$. The Hessian of the residual objective function is $\boldsymbol{B}(\boldsymbol{\delta}) = \{\nabla\nabla E_r\}(\boldsymbol{\delta})$, and $\boldsymbol{\delta}_{\text{MP}}$ are the updating parameters at the minimum value of F. $\boldsymbol{B}_{\text{MP}}$ may be estimated by $\boldsymbol{B} = \boldsymbol{J}^T\boldsymbol{W}_r\boldsymbol{J} = \boldsymbol{G}^T\boldsymbol{G}$ (Equations 30 and 39), where the Jacobian is evaluated at the minimum point.

Since $\boldsymbol{V}_C$ is a subset of these columns corresponding to the largest σ_j, this is equivalent to writing

$$\boldsymbol{G}' = \boldsymbol{U}\boldsymbol{\Sigma}\begin{bmatrix} \boldsymbol{I}_p \\ \boldsymbol{0} \end{bmatrix} = \sum_{j=1}^{p} \sigma_j \boldsymbol{u}_j \boldsymbol{e}_j^T \tag{49}$$

where $\boldsymbol{0}$ is a $(d - p) \times p$ matrix of zeros. Therefore, the singular values of $\boldsymbol{G}'$ are only the maximal p singular values of $\boldsymbol{G}$. The singular values of $\boldsymbol{G}$ are equal to the square root of the eigenvalues of $\boldsymbol{B} = \boldsymbol{G}^T\boldsymbol{G}$. If the parameter values $\boldsymbol{\delta}$ are sufficiently close to $\boldsymbol{\delta}_{\text{MP}}$, then $\boldsymbol{B} \approx \boldsymbol{B}_{\text{MP}}$. Using a uniform prior distribution for the FE model parameters $\boldsymbol{W}_\delta \propto \boldsymbol{I}_d$, then $\boldsymbol{W}_\theta = \boldsymbol{T}^T\boldsymbol{W}_\delta\boldsymbol{T} \propto \boldsymbol{I}_p$. Therefore $\lambda_j = \sigma_j^2$ are both the first p eigenvalues of $[\boldsymbol{W}_\delta^{-1}\boldsymbol{G}^T\boldsymbol{G}]$ and all the eigenvalues of $[\boldsymbol{W}_\theta^{-1}\boldsymbol{G}'^T\boldsymbol{G}']$. Thus, the new

parametrization represented by the transformation $\boldsymbol{T} = \boldsymbol{V}_C$ results in the maximal number of effective updating parameters since

$$\gamma' = \sum_{j=1}^{p} \frac{\beta\sigma_j^2}{\beta\sigma_j^2 + \alpha} \geq \sum_{D} \frac{\beta\sigma_j^2}{\beta\sigma_j^2 + \alpha} \; \forall D \neq \{1, ..., p\} \tag{50}$$

This comes from the fact that $x/(x+1)$ is maximized when x is maximized and $\{\sigma_1, .., \sigma_p\} \geq \{\sigma_j | j \notin \{1, ..., p\}\}$.

5.2.2 | Parametrization to maximize projection onto residual

While Equation 47 provides the best representation of the sensitivity matrix, there is no guarantee that the selected 'directions' (singular vectors) for updating will be effective in reducing the residual, and may even be orthogonal to $\boldsymbol{q}$. To avoid this, the logic of subset selection is used, with the set of right singular vectors $\boldsymbol{V}_C$ chosen such that the resulting sensitivity matrix $\boldsymbol{G'} = \boldsymbol{G}\boldsymbol{V}_C$ has maximum projection onto the residual $\boldsymbol{q}$. The new sensitivity matrix can be decomposed using Equation 46 and therefore the j^{th} column of $\boldsymbol{G'}$ is equal to $\boldsymbol{g}'_j = \sigma_j \boldsymbol{u}_j$. The projection of $\boldsymbol{g}'_j$ onto $\boldsymbol{q}$ is then $\sigma_j \boldsymbol{u}_j^T \boldsymbol{q}$, but can be normalized to $\boldsymbol{u}_j^T \boldsymbol{q}/\boldsymbol{q}^T \boldsymbol{q}$, which is the cosine of the angle between $\boldsymbol{u}_j$ and $\boldsymbol{q}$.

Therefore, the following parametrization is proposed: the set of singular values C is chosen which correspond to the p largest projections of left singular vectors $\boldsymbol{u}_j$ onto the weighted residual vector $\boldsymbol{q}$:

$$\boldsymbol{T} = \boldsymbol{V}_C; \qquad C = \{j \,|\, \boldsymbol{u}_j^T \boldsymbol{q} > \boldsymbol{u}_b^T \boldsymbol{q}; \; \forall b \notin C\} \tag{51}$$

This ensures that the updating parameters will be (at least locally) effective in reducing the residual. However, this approach may result in amplification of noise when $\boldsymbol{u}_j^T \boldsymbol{q} > \sigma_j$ [13] and may require very large updating parameter values when $\boldsymbol{u}_j^T \boldsymbol{q} \gg \sigma_j$. It may be practical to exclude singular values which are too small, but this requires definition of a threshold, evoking many of the difficulties previously noted for subset selection.

6 | UPDATING A SMALL-SCALE TRUSS MODEL WITH SIMULATED DATA

6.1 | Model description

The efficiency of the proposed model parametrization scheme was first tested on a 29-element, 28-DoF, 2-dimensional truss shown in Figure 1. This structure was modified from Papadimitriou *et al.*'s work [33] to have symmetric (pin-pin) boundary conditions and slightly different scale. This truss was also used in previous work by the authors [22]. Each truss element had identical section properties, with mass density of 7800 kg/m^3, Young's modulus of 200 GPa, and area 0.25 m^2. Therefore, this structure was statically indeterminate and symmetric. The truss FE model was implemented in MATLAB [34].

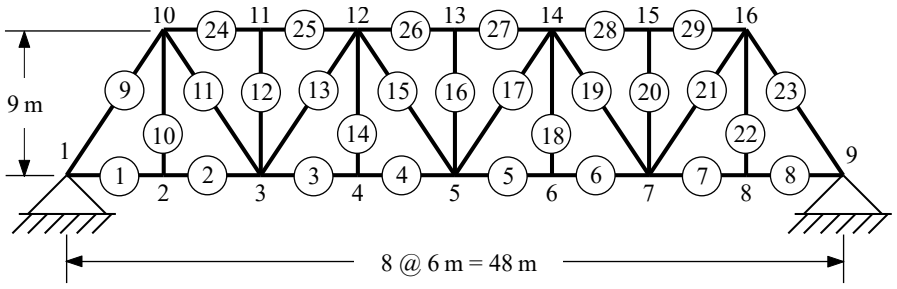


FIGURE 1 Truss structure layout with element and node numbers indicated

The first five vibrational modes were selected for analysis, with natural frequencies and mode shapes depicted in Figure 2. All 28 free DoFs were assumed to be measured, giving $m = 145$ measurements across the 5 natural frequencies and mode shapes.

In order to capture a full comparison between the different FE model parametrization schemes, the presented truss was used to generate a large number of related FE model updating problems. This was accomplished in two stages: random structural modification and addition of random measurement noise. Beginning with the unmodified truss, the mass density and Young's modulus of each element were modified using uncorrelated Gaussian random variables, δ_m^e and δ_k^e, as in Equations 33 and

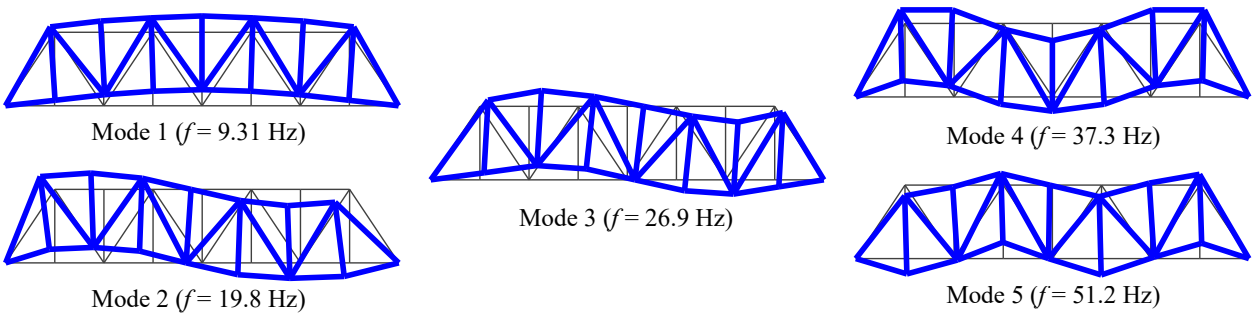


FIGURE 2 Truss mode shapes and natural frequencies

34. Two different levels of variability were analyzed, with case I corresponding to high uncertainty in physical parameters, $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, (0.1)^2 \boldsymbol{I}_d)$ and case II corresponding to low uncertainty in physical parameters, $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, (0.01)^2 \boldsymbol{I}_d)$. 100 random realized states were generated for each case.

The second stage was adding noise to the natural frequency and mode shape measurements for each of the 200 realized states. Natural frequency measurement noise was sampled from a Gaussian random variable with standard deviation equal to 0.5% of the measured natural frequency, i.e. $\tilde{f}_j \sim \mathcal{N}(f_j, (0.005 f_j)^2)$. Mode shape measurement noise was sampled from a Gaussian random variable with standard deviation equal to 5% of the corresponding mode shape (vector) standard deviation, i.e. $\tilde{\boldsymbol{\psi}}_j \sim \mathcal{N}(\boldsymbol{\psi}_j, (0.05\ \mathrm{std}(\boldsymbol{\psi}_j))^2 \boldsymbol{I}_l)$. This measurement noise model was intended to reflect typical conditions, in which natural frequency measurements are reliable within 1% of true values, while mode shape measurements exhibit an order-of-magnitude greater variability[2,35]. Note that all measurement noise was uncorrelated.

The residual weighting matrix $\boldsymbol{W}_r$ was evaluated as the inverse of the measurement covariance matrix, using measured quantities. Therefore, $\boldsymbol{W}_r$ was diagonal with $w^f_{rj} = (0.005 \tilde{f}_j)^{-2}$ and $w^s_{rj} = (0.05\ \mathrm{std}(\tilde{\boldsymbol{\psi}}_j))^{-2}$ in Equation 12. Obviously, $\boldsymbol{W}_r$ would vary slightly depending on the realization. The regularization parameter β was fixed at a value of 1/2, as discussed in Section 3, to ensure that the likelihood function was independent of the parametrization. Thus, $\beta \boldsymbol{W}_r$ was fixed for all parameterizations.

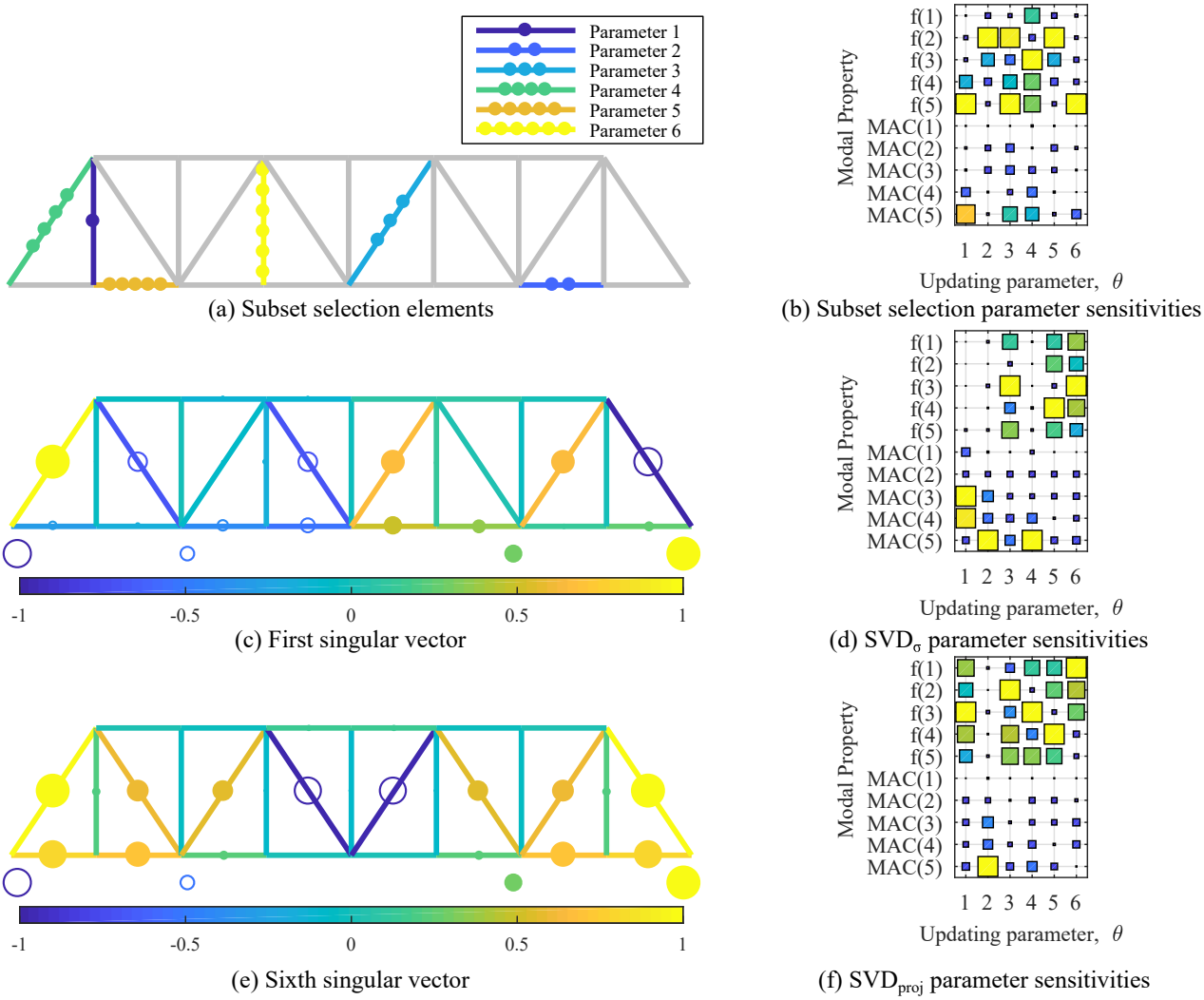
6.2 | Parametrization

For each realization, the FE model was parametrized using the methods described in Section 5. For a given number of parameters, p, each parametrization could also be called a "model class" or model, denoted as $\mathcal{M}_j$ in Equation 15. The first parametrization method was subset selection, described in Section 5.1, which is denoted as SS or $\mathcal{M}_1$. The second parametrization method was SVD-based parametrization with maximal singular values, defined by Equation 47, which is denoted as SVD_σ or $\mathcal{M}_2$. The third parametrization was SVD-based parametrization with maximal singular vector projection, defined by Equation 51, denoted as $\mathrm{SVD}_{\mathrm{proj}}$ or $\mathcal{M}_3$. Each parametrization was tested with the number of updating parameters ranging from $p = 1$ to $p = n_{el}$, only updating the Young's modulus of each element. Element mass densities were not used as FE model updating parameters to aid in depiction of parametrizations.

A sample set of model parametrizations is depicted in Figures 3a, c, and e using $p = 6$ updating parameters. Corresponding parameter sensitivities are shown in Figures 3b, d, and f, where larger squares indicate larger (magnitude) of sensitivity. For clarity of depiction, MAC sensitivity is shown in lieu of mode shape sensitivities. The sample SS parametrization in Figure 3a didn't show any preference for symmetry, but tended to avoid grouping all selected elements near one area of the truss. The sensitivities of the SS parameters showed good coverage of all natural frequencies, but none of the selected parameters were particularly impactful on the MAC of mode 1.

The first and sixth (right) singular vectors are shown in Figures 3c and e, respectively, with positive components denoted by filled circles and negative components denoted by hollow circles. Each singular vector encodes the change in FE physical parameters as a result of a change in the corresponding updating parameter. Each singular vector is generally non-zero for all FE Young's moduli, defining a relative amount of stiffness increase or decrease in each element. Thus the first singular vector depicts the stiffness changes corresponding to θ_1 in SVD_σ, which mostly affected the stiffness of the end diagonal elements. The sixth singular vector in Figure 3e corresponds to θ_6 in SVD_σ as well as θ_1 in $\mathrm{SVD}_{\mathrm{proj}}$. The sixth singular vector had a large effect on diagonal elements and bottom chord elements near the supports.

It is interesting to note that both SVD-based parametrization methods had parameters which were mostly effective on natural frequencies (e.g. θ_3 in SVD_σ and θ_1 in SVD_{proj}) and separate parameters which were impactful on mode shapes (e.g. θ_1 in SVD_σ and θ_2 in SVD_{proj}). This may be explained by phenomena noticed in previous work[22], in which natural frequency sensitivities were symmetric for symmetric parameters, but mode shape sensitivities were asymmetric for symmetric parameters. Intuitively, this implies that symmetric singular vectors (e.g. Figure 3e) will be more impactful on natural frequencies while anti-symmetric singular vectors (e.g. Figure 3c) will contribute more to mode shape sensitivities.



(a) Subset selection elements

(b) Subset selection parameter sensitivities

(c) First singular vector

(d) SVD_σ parameter sensitivities

(e) Sixth singular vector

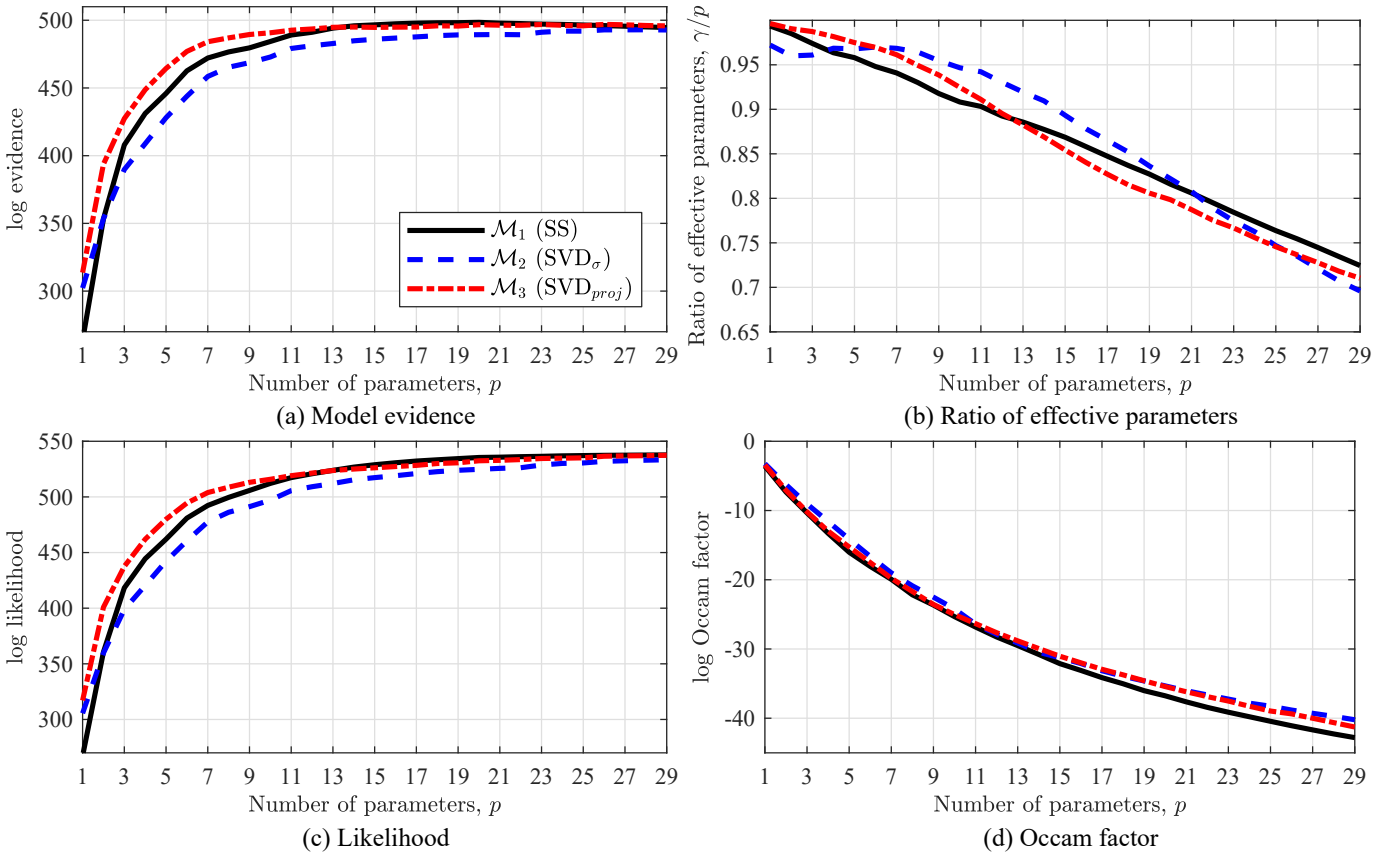
(f) SVD_{proj} parameter sensitivities

FIGURE 3 Sample truss parametrizations and parameter sensitivities for $p = 6$ updating parameters

6.3 | Model updating results

FE model updating was performed for each of the 200 realizations, using the three described parametrization methods and model sizes from 1 to n_{el}. The regularized objective function in Equation 14 was used with the FE model parameter covariance matrix $\boldsymbol{W}_\delta = \boldsymbol{I}_d$. It is interesting to note that all parametrization methods used in this work used orthogonal transformation matrices, and therefore $\boldsymbol{W}_\theta = \boldsymbol{I}_p$ from Equation 37. The regularization parameter α was allowed to vary as determined by the estimation algorithm to provide an optimal parameter weighting matrix. The Levenberg–Marquardt minimization with Bayesian regularization scheme (Section 4 and Algorithm 1) was used to optimize the objective function. The evidence, likelihood, and Occam factor were evaluated using Equation 22 at $\boldsymbol{\theta}_{\text{MP}}$.

The average posterior results for model evidence, likelihood, and Occam factor are shown for high FE model parameter uncertainty (case I) in Figures 4a, c, and d, respectively. In this case, the model evidence for each parametrization peaked between $p = 12$ and $p = 24$, favoring larger parametrizations. Each parametrization method resulted in similar evidence, likelihood, and Occam factor curves, with likelihood increasing for larger parametrizations as expected. The Occam factor was not linear with respect to p, indicating that adding parameters gave diminishing returns in terms of information extraction. Significant differences in average evidence are visible at $p < 12$, with all parametrizations performing similarly for larger models. SVD_{proj} outperformed SS at small p, indicating that it was more efficient with highly reduced parametrizations. SVD_{σ} was outperformed by both SVD_{proj} and SS at all values of p, suggesting that it was far more important to incorporate the weighted residual $\boldsymbol{q}$ during parametrization, rather than trying to extract as much information from the Jacobian matrix. SS had the highest (magnitude) Occam factor, which implies that it extracted mildly more information from the data during updating.



(a) Model evidence (b) Ratio of effective parameters

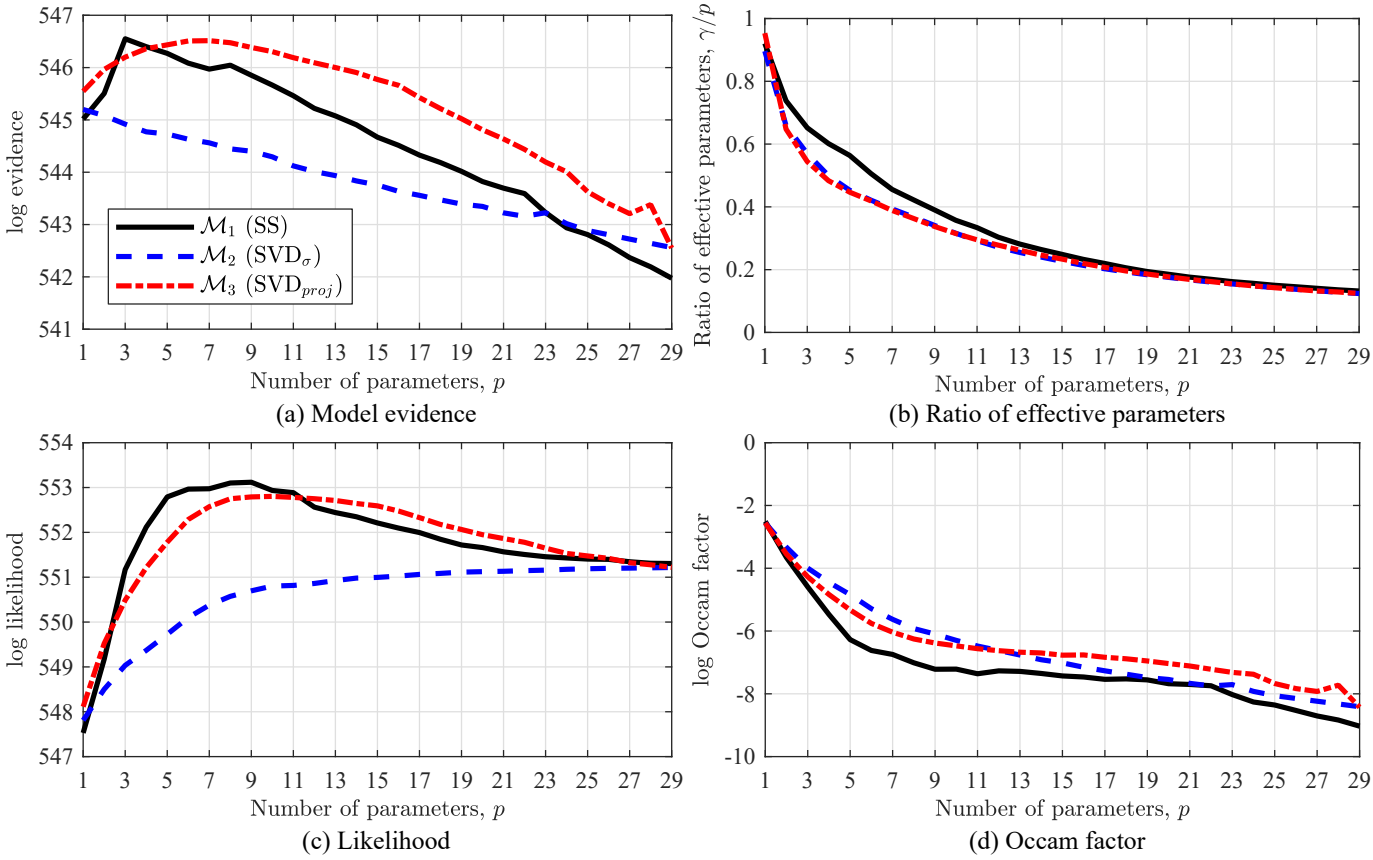
(c) Likelihood (d) Occam factor

FIGURE 4 Average truss model updating results, high uncertainty in FE model parameters (case I)

Instead of plotting the effective number of parameters γ from Equation 25, the ratio of effective parameters γ/p is plotted in Figure 4b. All parametrizations began with nearly full saturation of parameters, $\gamma/p = 1$, and decayed to $\gamma/p \approx 0.70$ even when using $p = d$. SS parametrization resulted in a nearly linear decay in efficiency, while the SVD-based parametrizations were non-linear. Despite the fact that SVD_{σ} theoretically should have had maximal γ for a set number of parameters p (Section 5.2.2), this was only true between $p = 6$ and $p = 20$. This may be due to the large discrepancy between initial model parameters and updated model parameters, which renders some assumptions in Section 5.2.2 inappropriate.

Figure 5 presents the average truss model updating results using low FE model parameter uncertainty (case II). Evidence results favored small models, peaking between $p = 1$ and $p = 7$ for each of the three parametrizations. Unlike the results in Figure 4, the behavior of the evidence and likelihood curves was markedly different for each parametrization. SS showed mildly greater average evidence at its peak of $p = 3$ and decreased nearly linearly after its peak. SVD_{proj} showed a flatter evidence curve that outperformed the other parametrizations at for nearly all p. SVD_{σ} didn't display a peak evidence, decreasing monotonically from $p = 1$ and with lower evidence than the other parametrizations. SS actually outperformed the other parametrizations in

terms of data fit (likelihood) between $p = 2$ and $p = 11$, but this was offset by its greater Occam factor. It may be concerning to note that the likelihood curves (Figure 5c) weren't monotonically increasing with more parameters for SS and SVD_{proj}, but this is explained by the fact that the posterior probability (F) was maximized in Algorithm 1 rather than the likelihood (E_r). The parametrization efficiency (Figure 5b) again started with near-saturation, but decayed non-linearly to the much lower level of 0.15 for all models. Despite the fact that the initial FE model parameters were much closer (on average) to their true values, SVD_σ didn't provide the maximal value for γ at each given p.



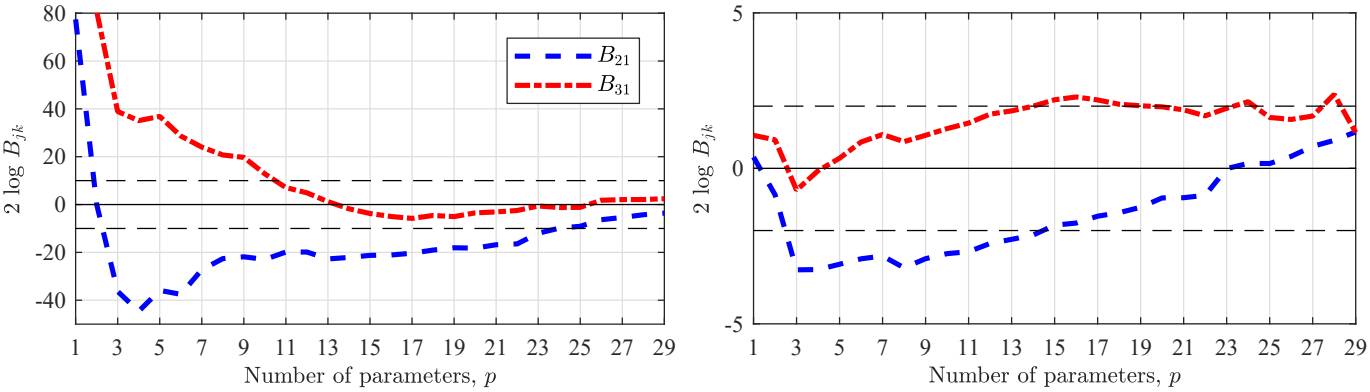
(a) Model evidence (b) Ratio of effective parameters

(c) Likelihood (d) Occam factor

FIGURE 5 Average truss model updating results, low uncertainty in FE model parameters (case II)

Interpretation of the support for each parametrization ($\mathcal{M}_j$) is performed most naturally using Bayes factors[23], as defined in Equation 26. The support for SVD_σ over SS ($\mathcal{M}_2$ over $\mathcal{M}_1$: B_{21}) and for SVD_{proj} over SS ($\mathcal{M}_3$ over $\mathcal{M}_1$: B_{31}) are plotted in Figure 6 for each level of FE model parameter uncertainty, including dashed lines for interpreting the significance of the support from Table 1. Note that evidence comparison between SVD_σ and SVD_{proj} could be inferred by the log difference between B_{21} and B_{31} since $\log B_{23} = \log B_{21} - \log B_{31}$. For high parameter uncertainty, $\log B_{21}$ was almost entirely less than 0, indicating support for SS over SVD_σ with very strong significance for small p and diminishing to strong or less for $p > 24$. At $p = 1$ however, SVD_σ was actually supported over SS. Conversely, SVD_{proj} was very strongly supported over SS for $p < 10$ (B_{31}), then decreasing into weak support for SS with $p > 14$.

Bayes factors were much lower in significance for low uncertainty in FE model parameters (Figure 6b). Again, SS was supported over SVD_σ with $\log B_{21} < 0$ for most p, varying from positive for $2 < p < 14$ and weak results for larger models. Support for SVD_{proj} over SS (B_{31}) was inconclusive for $p < 14$, but was generally positive for $p > 14$, indicating that these parametrizations are equally supported by the data with a slight edge toward SVD_{proj} for larger parametrizations.



(a) High uncertainty in FE model parameters (case I)

(b) Low uncertainty in FE model parameters (case II)

FIGURE 6 Bayes factors for competing truss parametrizations

7 | UPDATING A LARGE-SCALE SUSPENSION BRIDGE MODEL WITH REAL DATA

7.1 | System identification

The second test of the proposed parametrization schemes was FE model updating of a full-scale suspension bridge with measured data. The studied structure is a double-deck steel bridge with four suspension cables and two towers, as used in previous work[22,28,29]. The structure is symmetric with a 2089 m total length among two side-spans and a 451 m mid-span. A series of ambient vibration studies were performed in 2009 to identify modal properties such as natural frequencies, mode shapes, and damping ratios under typical operating conditions[28,29]. Vibrational responses were captured using tri-axial accelerometers at 9 locations on the spans and towers, giving 27 measured DoFs for each mode shape. The data from one day was used in this study, using measurements during four 1-hour periods. The measured modal data (natural frequencies and mode shapes) was then averaged across the four measurement periods to provide an estimate for average daily modal properties. Note that more detailed data about identification techniques and hourly data can be found in Jang and Smyth's work[28,29].

The first 7 vibrational modes were chosen for use in model updating, giving $m = 196$ measurements for 7 natural frequencies and mode shapes. The average mode shapes, including mode labels and average natural frequencies, are given in Figure 7. The depicted mode shapes indicate the mode shape amplitude at the 9 measured locations (indicated by dots), while the unmeasured modal displacements were interpolated with reasonable boundary conditions. This interpolation was only used for the purposes of depiction; any use of measured data only utilizes the 27 directly-measured DoFs. The suspension cables and suspenders are omitted from Figure 7 for clarity.

The measured data could have been used to estimate the measurement covariance matrix $\boldsymbol{C}_{\tilde{z}}$, but four observations was considered to be inadequate. Thus, the measurement covariance matrix was formed based on an assumed noise model. The noise in each natural frequency measurement was assumed to have a standard deviation equal to 0.5% of the measured natural frequency value. The noise in each mode shape component was assumed to have a standard deviation equal to 5% of the measured mode shape's standard deviation. The residual weighting matrix, $\boldsymbol{W}_r$, was then the inverse of this assumed measurement covariance matrix. Using Equation 12, this can be written as $w^f_{rj} = (0.005\tilde{f}_j)^{-2}$ and $w^s_{rj} = (0.05\ \mathrm{std}(\tilde{\boldsymbol{\psi}}_j))^{-2}$, or $c_f = 0.005$ and $c_s = 0.05$ in Equation 13. As in Section 6, the regularization parameter related to the residual, β, was fixed at a value of 1/2 to provide more accurate evidence comparisons between parametrizations.

7.2 | FE model description

The suspension bridge FE model was developed in ABAQUS[36] using partially-available technical drawings to define the geometry and element properties. In cases where technical drawings were uninformative, reasonable guesses for element properties were made using photography. The FE model comprised approximately 21,000 elements and 18,000 nodes. Soil interaction and thermal expansion joints were incorporated through boundary condition springs and hinge springs, respectively. For a thorough discussion of element types, boundary conditions, connections, and initial model material properties, please refer to Jang and Smyth's description of the same FE model[28].

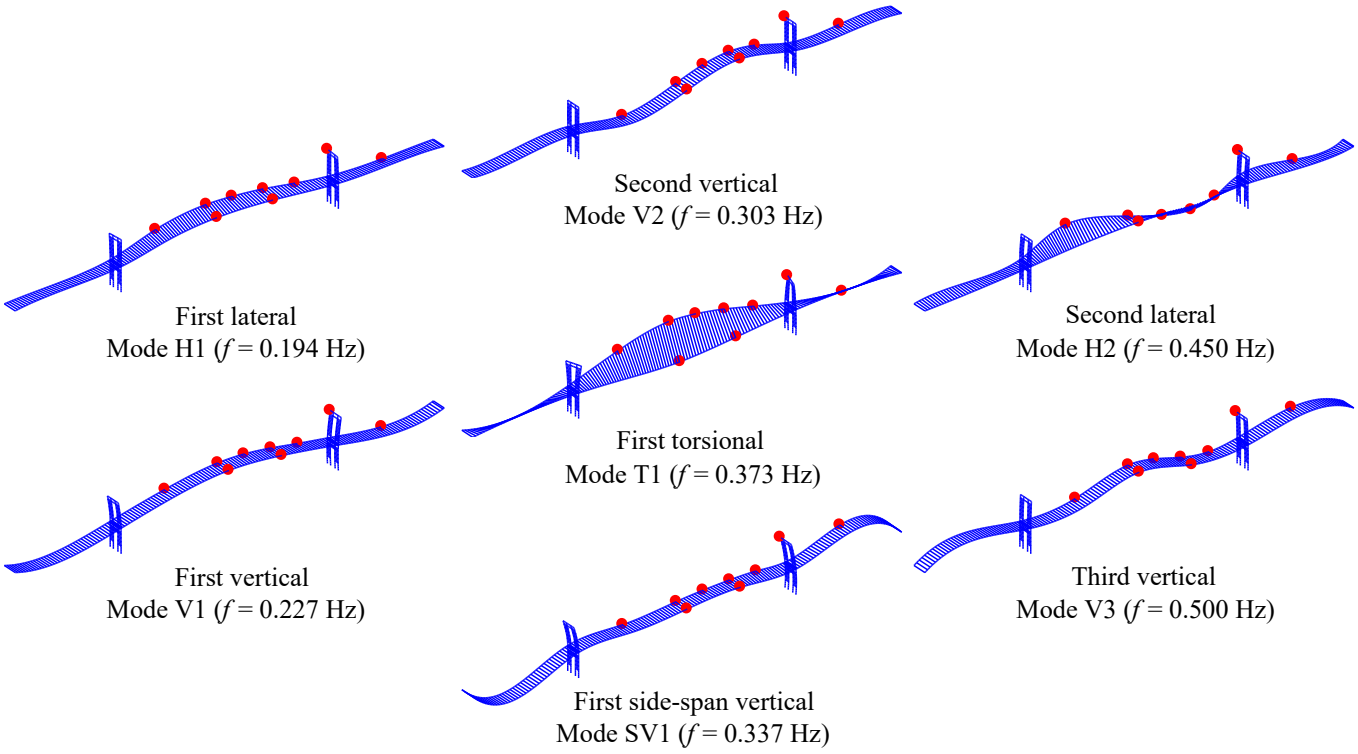


FIGURE 7 Suspension bridge measured modes (measurement locations indicated by red dots)

The MAC values of the initial FE model are given in Table 2, along with the relative frequency error, $f_{\text{err}} = (\tilde{f} - f)/\tilde{f}$. Due to FE model modifications to account for more realistic structural behavior, particularly in the boundary conditions and interactions between cable and deck components, the initial f_{err} and MAC values are slightly different from those in Jang and Smyth's previous work[28,29]. The natural frequencies of the initial FE model were higher than their measured counterparts. The first torsional mode (T1) had the lowest initial natural frequency error at -3.0%, while the first side-span vertical mode (SV1) exhibited the highest initial natural frequency error at -34.2%. The first two vertical modes (V1 and V2) and the first lateral mode (H1) were already very close to their measured counterparts, with MAC values above 0.950. The first torsional mode (T1) and the third vertical mode (V3) exhibited the lowest initial MAC values at 0.741 and 0.743, respectively. Every mode exhibited a high frequency error and/or a low MAC, indicating that every mode would be important in model updating. The total initial FE model error E_r was a summation of the natural frequency error E_r^f, comprising 60% of the total, and the mode shape error E_r^s, comprising the other 40% of the total (Equation 12).

Due to the relatively low number of mode shape measurements, an intermediate step was implemented during mode pairing. Initial FE model modes and measured modes were paired using MAC from the 27 measured DoFs, creating an index between the modes of the initial FE model and measured modes. During the model updating, FE model modes were first paired with the initial FE model modes using all FE model DoFs to increase pairing fidelity. Then the index between initial FE model-measured modes was used to relate each updated mode to the correct measured mode. This approach ensured consistent pairing between FE model modes and measured modes, since FE model mode shapes could change significantly and had relatively few DoFs for direct pairing.

7.3 | Parametrization

Due to the large number of FE model physical parameters (approximately 42,000 physical parameters, among mass densities, Young's moduli, and spring coefficients), it was necessary to first arrange elements into substructures to avoid an intractably large set of sensitivity calculations. Structural components were decomposed into 132 substructures based on element type and location. The main span was divided into 8 longitudinal groups and the two side-spans were each divided into four longitudinal groups. The towers were divided into three vertical groups. These groups were then further divided based on element type.

The properties of these substructures were used as the FE model physical parameters, giving 132 mass densities, 132 Young's moduli, 15 spring coefficients to update. Thus, the natural FE parametrization $\boldsymbol{\delta}$ had $d = 279$ components.

To be consistent with Jang and Smyth's prior work, each parametrization used 5 mass parameters and 17 stiffness (and spring) parameters, giving $p = 22$ total updating parameters. This was achieved by separately selecting mass and stiffness parameters based on $\partial \boldsymbol{q}/\partial \boldsymbol{\delta}^m$ and $\partial \boldsymbol{q}/\partial \boldsymbol{\delta}^k$ sensitivities, respectively. The FE model was parametrized according to subset selection ($\mathcal{M}_1$ or SS), SVD-based parametrization with maximal singular values ($\mathcal{M}_2$ or SVD_σ), and SVD-based parametrization with maximal singular vector projection ($\mathcal{M}_3$ or SVD_{proj}) methods described in Section 5.

Figure 8 depicts the parametrizations of the bridge, as well as the sensitivities for the 22 updating parameters of each parametrization in subfigures b, d, and f. Note that these plots reflect the absolute value of the sensitivity, so visually similar parameters (e.g. θ_{20} to θ_{22} in Figure 8d) may have components which differ significantly in sign. The sensitivity plots are separated into mass parameters (θ_1 to θ_5) and stiffness parameters (θ_6 to θ_{22}).

Figure 8a depicts the mass substructures chosen by SS, with θ_1 and θ_5 affecting the mass density of the side-span lateral bracing. As expected, these parameters have the largest impact on the side-span vertical mode, SV1, as confirmed by the sensitivities in Figure 8b. θ_2 comprises mid-span lateral bracing and mainly affects natural frequencies for main-span modes. θ_3 and θ_4 comprise bracing and chord main-span elements near one tower, mainly affecting modes H2 and V3. It is interesting to note that SS parameters seem to be specialized in the sense that each parameter mainly affects one or two modal properties with very little influence on other properties. This is exemplified by θ_{15} which only has significant effect on the T1 and SV1 mode shapes.

The third mass singular vector (i.e. $\boldsymbol{v}_3$ of $\partial \boldsymbol{q}/\partial \boldsymbol{\delta}^m$) is shown in Figures 8c, which corresponds to θ_3 in SVD_σ and θ_1 in SVD_{proj}. This can be viewed as the change in element masses when either of those two updating parameters are perturbed, which essentially adds mass to the midspan area. Unsurprisingly, these parameters mostly affect the natural frequency of main-span modes (Figures 8d and f). The third stiffness singular vector (i.e. $\boldsymbol{v}_3$ of $\partial \boldsymbol{q}/\partial \boldsymbol{\delta}^k$) is shown in Figure 8e, which mainly influences the truss element stiffnesses near the span-ends and the tower elements. This corresponds to θ_8 in SVD_σ and θ_5 in SVD_{proj}. These updating parameters have large influence on natural frequencies, mainly for mode T1, with little impact on mode shapes. As noted in Section 6, since the singular vectors depicted in Figures 8c and e are approximately symmetric, they are mostly impactful on natural frequencies. In general, the SVD-based parametrizations had low mode shape sensitivity compared to natural frequency sensitivity, except for θ_{19} of SVD_{proj}, which also exhibited separation between the two kinds of sensitivity.

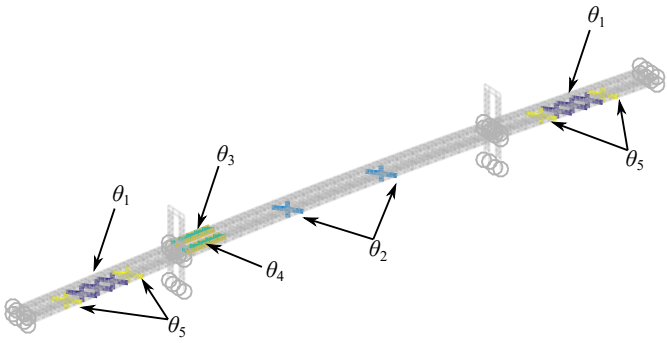
Both of the SVD-based parametrizations showed less specialization than SS, with parameters generally having significant effect on multiple modal properties. In particular, pretty much every parameter in SVD_σ affected multiple natural frequencies, generally with lower impact on mode shapes, while SVD_{proj} had much some parameters which were specialized (e.g. θ_4 and θ_{13}) and more impact on mode shapes. This suggests that parametrizations which incorporate the measurement residual $\boldsymbol{q}$ show a tendency towards specialized parameters, perhaps reflecting non-uniform distribution of measurement error.

7.4 | Model updating results

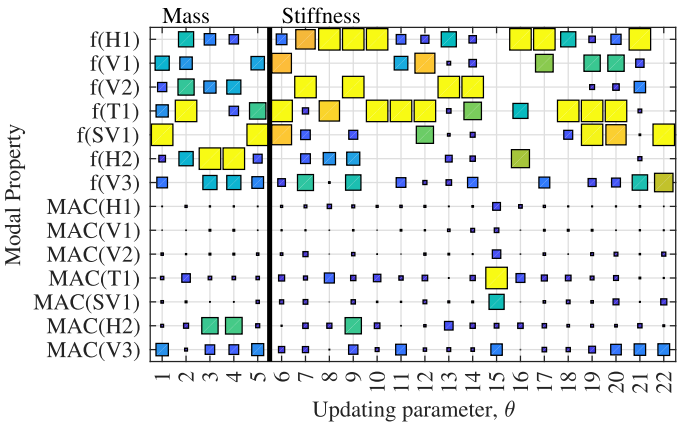
Model updating proceeded using the three parametrization methods, each with 5 mass and 17 stiffness updating parameters. The regularized objective function in Equation 14 was used with FE model parameter covariance matrix $\boldsymbol{W}_\delta = \boldsymbol{I}_d$. As noted in Section 6, all parametrizations used orthogonal transformation matrices giving $\boldsymbol{W}_\theta = \boldsymbol{I}_p$. The Levenberg–Marquardt minimization with Bayesian regularization scheme (Section 4 and Algorithm 1) was used to optimize the objective function, with α free to be determined by the algorithm. Parametrization and optimization were performed in MATLAB [34], while modal analysis was performed in ABAQUS [36]. Communication between MATLAB and ABAQUS was controlled by an application programming interface. The evidence, likelihood, and Occam factor were evaluated using Equation 22 at the minimum point $\boldsymbol{\theta}_{\text{MP}}$.

The converged results are shown in Table 2, including the relative frequency error f_{err} and MAC for each mode in both the initial and updated states. The total sum-of-square residual E_r, composed of natural frequency residual E_r^f and mode shape residual E_r^s (as in Equation 12) is also included, along with the parameter efficiency γ/p (Equation 25).

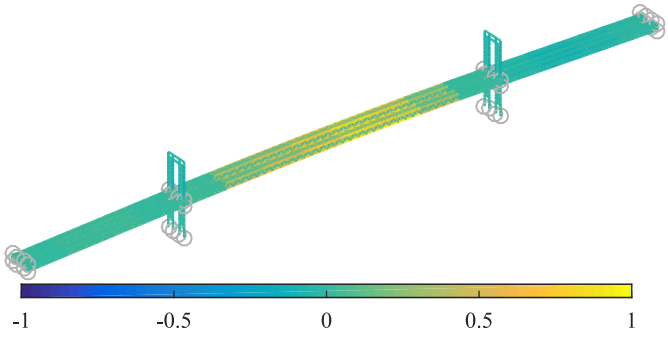
SS parametrization produced mediocre reductions in model error, giving 34% less total error than the initial model, splitting its total error almost equally between E_r^f and E_r^s. SS mildly improved natural frequency results across all modes. SS was relatively unsuccessful in reducing large natural frequencies errors, but did show promising results for H1 and V2. Mode shape results were underwhelming, with very slight gains for most modes and slight loss in MAC for mode SV1. Parameter efficiency was also quite low at 0.31, indicating that the prior distribution (regularization) played a larger role than the likelihood (data) in determining the parameter values.



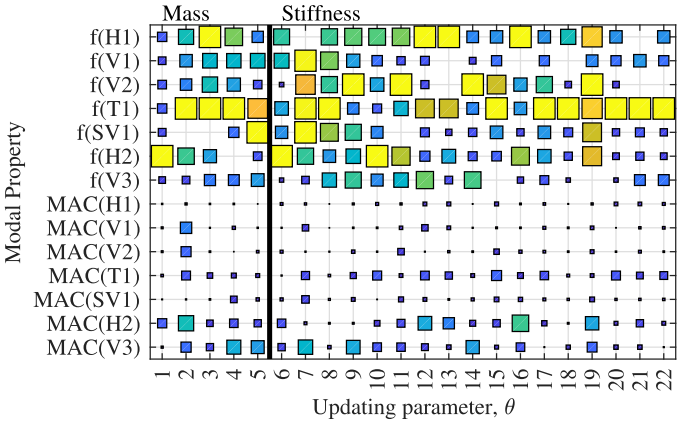
(a) Subset selection mass parameters



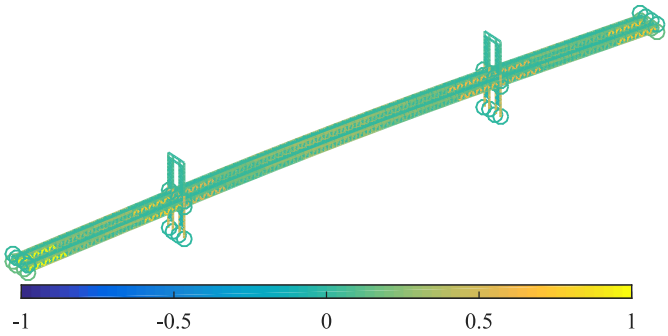
(b) Subset selection parameter sensitvities



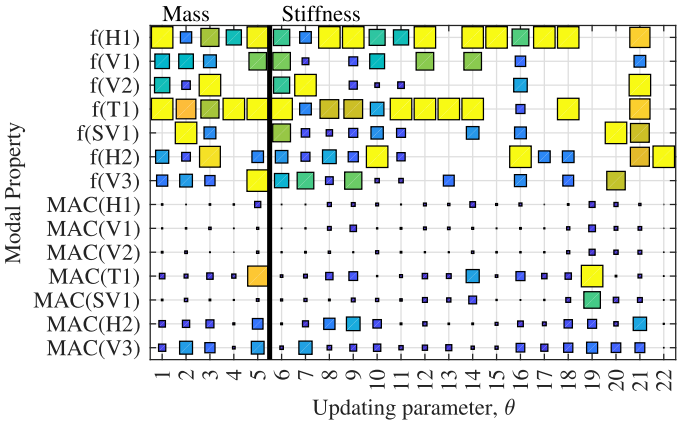
(c) Third mass singular vector



(d) SVD_σ parameter sensitvities



(e) Third stiffness singular vector



(f) SVD_{proj} parameter sensitvities

FIGURE 8 Suspension bridge parametrizations and parameter sensitivities

The SVD-based parametrizations both showed very strong results with 74% and 76% reduction in E_r for SVD_σ and SVD_{proj}, respectively. SVD_σ focused more on reducing more shape error, with E_r^s comprising 77% of its total error, while it comprised 82% of the total error for SVD_{proj}. Both SVD-based parametrizations showed similar natural frequency results, with SVD_{proj} slightly outperforming SVD_σ for all modes except V2. These parametrizations were extremely successful in reducing natural frequency error, cutting E_r^f by approximately 90% from its initial value. Both methods struggled with mode T1, increasing the relative error from -3% to approximately 10%. Problems with mode T1 were noted in previous work[22,28], and may be related to unmodeled non-linear geometry. Three T1 modes were identified in the measured data sets, each with different in-phase or out-of-phase motions between the main cable and deck, and different interactions between the main-span and side-spans. However, only one T1 mode was produced in the FE model because the geometrically non-linear deck-cable interaction was not modeled.

TABLE 2 Suspension bridge model updating results

	Initial FE model		Updated FE model					
			$\mathcal{M}_1$ (SS)		$\mathcal{M}_2$ (SVD$_\sigma$)		$\mathcal{M}_3$ (SVD$_{\text{proj}}$)	
Mode	f_{err}	MAC	f_{err}	MAC	f_{err}	MAC	f_{err}	MAC
H1	-21.7%	0.984	-10.5%	0.990	5.4%	0.992	2.4%	0.991
V1	-29.8%	0.969	-25.9%	0.971	-9.4%	0.970	-8.9%	0.963
V2	-17.5%	0.986	-8.8%	0.984	2.8%	0.980	5.1%	0.979
T1	-3.0%	0.741	1.7%	0.795	11.3%	0.820	9.7%	0.846
SV1	-34.2%	0.879	-26.0%	0.876	0.4%	0.872	0.3%	0.889
H2	-19.8%	0.845	-11.3%	0.903	-1.1%	0.962	-0.3%	0.939
V3	-19.4%	0.743	-16.8%	0.774	-10.3%	0.974	-7.4%	0.968
Total error, E_r (10^4)	2.39		1.57		0.61		0.57	
Nat. freq. error, E_r^f (10^4)	1.44		0.78		0.14		0.10	
Mode shape error, E_r^s (10^4)	0.95		0.79		0.47		0.47	
Parameter efficiency, γ/p	–		0.31		0.94		0.92	

TABLE 3 Suspension bridge posterior results and Bayes factors (all results $\times 10^3$)

	Posterior results (log)			Bayes factor, $2 \log B_{jk}$		
				Model $\mathcal{M}_k$		
Model $\mathcal{M}_j$	Evidence	Likelihood	Occam factor	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$
$\mathcal{M}_1$ (SS)	-7.11	-7.08	-0.027	–	-9.41	-9.87
$\mathcal{M}_2$ (SVD$_\sigma$)	-2.41	-2.32	-0.085	9.41	–	-0.46
$\mathcal{M}_3$ (SVD$_{\text{proj}}$)	-2.18	-2.11	-0.069	9.87	0.46	–

Similar mode shape updating results were also noted for the two SVD-based parametrizations, decreasing the mode shape error E_r^s by about 50% from its initial value. Both parametrizations were highly successful in improving the MAC of modes T1 and V3, beginning near 0.750 and ending around 0.830 and 0.970, respectively. Similarly strong improvement was noted for mode H2. Parametrization efficiencies were excellent, at 0.94 for SVD$_\sigma$ and 0.92 for SVD$_{\text{proj}}$, respectively. This indicates that the parameters were very efficient in utilizing the data for updating, with minimal influence of regularization. The slightly higher values for SVD$_\sigma$ was expected due to the results of Section 5.2.2.

Table 3 displays the posterior results, including the (log) evidence, likelihood, and Occam factor for each parametrization. Since β and $\boldsymbol{W}_r$ were fixed for all models, the log likelihood was controlled by the total error E_r. Therefore, it was expected that SVD$_{\text{proj}}$ would have higher likelihood and also greater evidence since the Occam factors were small in magnitude. The evidence for SS was very low compared to the SVD-based parametrizations, due the high total error and therefore low likelihood. The Occam factors for the SVD-based methods were similar, with SS having a much lower Occam factor, reflecting the statements made for γ/p (i.e. low data utilization). The relative evidence for pairs of models were examined using Bayes factors (see Equation 26 and Table 1). There was extremely strong evidence for the SVD-based parametrizations over SS (>9000) with order-of-magnitude lower, but still decisive, evidence for SVD$_{\text{proj}}$ over SVD$_\sigma$ in this model updating exercise.

8 | CONCLUSIONS

The approach proposed in this work utilizes SVD of the sensitivity matrix to develop robust, reduced-order parametrizations which improve posedness and efficiency in FE model updating. Singular vectors are used to define linear combinations of underlying FE model parameters which are each updated by a single updating parameter. SVD-based parametrization can form an optimal, reduced representation of the sensitivity matrix, or singular vectors can be selected to best represent the measurement residual, similar to subset selection. Model error is minimized using a deterministic scheme which incorporates Bayesian

inference to perform regularization and estimate both parametrization efficiency and model evidence. This is closely related to Laplace's method with minimization via the Levenberg–Marquardt algorithm. The main novelty of this approach is in optimally selecting the regularization parameter, which corresponds to estimating an improved prior PDF. The proposed approach to model updating combines the low computational cost of regularized deterministic methods with strong ties probabilistic methods.

The proposed SVD-based parametrization schemes were tested against subset selection on two vibration-based model updating problems: a small-scale 2-dimensional truss with simulated measurements and a large-scale suspension bridge with real data. In both cases, natural frequencies and mode shapes were targeted for updating. The truss example provided an efficient testbed for comparing a range of model sizes and parameter uncertainty levels across a significant number of randomized realizations. Parametrization using only the largest singular vectors was generally not supported by the data compared to subset selection, while incorporating the residual into choice of singular vectors resulted in as-good or better support compared to subset selection. Support was measured by the relative evidence (Bayes factor). Model updating of the large-scale suspension bridge produced mediocre results when parametrized using subset selection, while the SVD-based methods provided excellent reductions in error. In this example, incorporating data into choice of singular vectors was again shown to be effective.

While the proposed parametrizations showed excellent results on the presented examples, there was significant variation in the strength of model support in the two model updating exercises. Further work is required to understand if subset selection is inherently ineffective on large-scale models or only for the presented suspension bridge. Regularization was generally unnecessary for the included examples which were always overdetermined, but it provided a consistent set of prior beliefs and allowed estimates of model evidence. The posterior estimates may be inaccurate for non-Gaussian prior distributions and likelihoods, or for small data sets.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Columbia University's Graduate School of Arts and Sciences in support of the first author through the Guggenheim Fellowship and Presidential Fellowship. This work was partially supported by the U.S. National Science Foundation (Grant No. CMMI-1563364).

References

1. Mottershead JE, Link M, Friswell MI. The sensitivity method in finite element model updating: a tutorial. *Mech Syst Signal Pr.* 2011;25(7):2275–2296.

2. Mottershead JE, Friswell MI. Model updating in structural dynamics: a survey. *J Sound Vib.* 1993;167(2):347–375.

3. Friswell MI, Mottershead JE. *Finite Element Model Updating in Structural Dynamics*. Dordrecht: Springer; 1995.

4. Brownjohn JMW, De Stefano A, Xu Y-L, Wenzel H, Aktan AE. Vibration-based monitoring of civil infrastructure: challenges and successes. *J Civil Struct Hlth Monitor.* 2011;1(3–4):79–95.

5. Shahverdi H, Mares C, Wang W, Mottershead JE. Clustering of parameter sensitivities: examples from a helicopter airframe model updating exercise. *Shock Vib.* 2009;16(1):75–87.

6. Goller B, Broggi M, Calvi A, Schueller GI. A stochastic model updating technique for complex aerospace structures. *Finite Elem Anal Des.* 2011;47(7):739–752.

7. Mottershead JE, Mares C, Friswell MI, James S. Selection and updating of parameters for an aluminium space-frame model. *Mech Syst Signal Pr.* 2000;14(6):923–944.

8. Teughels A, De Roeck G. Structural damage identification of the highway bridge Z24 by FE model updating. *J Sound Vib.* 2004;278(3):589–610.

9. Simoen E, De Roeck G, Lombaert G. Dealing with uncertainty in model updating for damage assessment: a review. *Mech Syst Signal Pr.* 2015;56–57:123–149.

10. Lallement G, Piranda J. Localization methods for parametric updating of finite element models in elastodynamics. In: Proc 8th Int Modal Anal Conf; Jan 29–Feb 1, 1990; Kissimmee.

11. Friswell MI, Mottershead JE, Ahmadian H. Combining subset selection and parameter constraints in model updating. *J Vib Acoust.* 1998;120(4):854–859.

12. Ahmadian H, Mottershead JE, Friswell MI. Regularisation methods for finite element model updating. *Mech Syst Signal Pr.* 1998;12(1):47–64.

13. Friswell MI, Mottershead JE, Ahmadian H. Finite–element model updating using experimental test data: parametrization and regularization. *Philos T Roy Soc A.* 2001;359(1778):169–186.

14. Titurus B, Friswell MI. Regularization in model updating. *Int J Numer Meth Eng.* 2008;75(4):440–478.

15. MacKay DJC. Bayesian interpolation. *Neural Comput.* 1992;4(3):415–447.

16. Collins JD, Hart GC, Haselman TK, Kennedy B. Statistical identification of structures. *AIAA J.* 1974;12(2):185–190.

17. Friswell MI. The adjustment of structural parameters using a minimum variance estimator. *Mech Syst Signal Pr.* 1989;3(2):143–155.

18. Allemang RJ, Brown DL. A correlation coefficient for modal vector analysis. In: Proc 1st Int Modal Anal Conf; Nov 8–10, 1982; Orlando.

19. Beck JL, Yuen K-V. Model selection using response measurements: Bayesian probabilistic approach. *J Eng Mech.* 2004;130(2):192–203.

20. Papadimitriou C, Beck JL, Katafygiotis LS. Asymptotic expansions for reliability and moments of uncertain systems. *J Eng Mech.* 1997;123(12):1219–1229.

21. Beck JL, Katafygiotis LS. Updating models and their uncertainties I: Bayesian statistical framework. *J Eng Mech.* 1998;124(4):455–461.

22. Bartilson DT, Jang J, Smyth AW. Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization. *Mech Syst Signal Pr.* 2019;114:328–345.

23. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* 1995;90(430):773–795.

24. Fox RL, Kapoor MP. Rates of change of eigenvalues and eigenvectors. *AIAA J.* 1968;6(12):2426–2429.

25. Levenberg K. A method for the solution of certain non-linear problems in least squares. *Q Appl Math.* 1944;2(2):164–168.

26. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Ind Appl Math.* 1963;11(2):431–441.

27. Smith CB, Hernandez EM. Detection of spatially sparse damage using impulse response sensitivity and LASSO regularization. *Inverse Probl Sci Eng.* 2018;.

28. Jang J, Smyth AW. Model updating of a full-scale FE model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters. *Mech Syst Signal Pr.* 2017;83:337–355.

29. Jang J, Smyth AW. Bayesian model updating of a full-scale finite element model with sensitivity-based clustering. *Struct Control Health Monit.* 2017;24(11):e2004.

30. Miller A. *Subset Selection in Regression*. London: Chapman & Hall; 2002.

31. Silva TAN, Maia NMM, Link M, Mottershead JE. Parameter selection and covariance updating. *Mech Syst Signal Pr.* 2016;70:269–283.

32. Golub GH, Reinsch C. Singular value decomposition and least squares solutions. *Numer Math.* 1970;14(5):403–420.

33. Papadimitriou C, Beck JL, Au S-K. Entropy-based optimal sensor location for structural model updating. *J Vib Control.* 2000;6(5):781–800.

34. *MATLAB* [computer program]. Version 9.1.0 (R2016b). Natick, MA: The MathWorks Inc; 2016.

35. Modak SV, Kundra TK, Nakra BC. Comparative study of model updating methods using simulated experimental data. *Comput Struct.* 2002;80(5):437–447.

36. *ABAQUS/CAE* [computer program]. Version 6.14. Providence, RI: Dassault Systèmes; 2014.

This page intentionally left blank.

Paper C

Incremental reparametrization in sensitivity-based subset selection and parameter clustering for finite element model updating

---

Bibliographic information:

Bartilson, D. T., Jang, J., & Smyth, A. W. (2018). Incremental reparametrization in sensitivity-based subset selection and parameter clustering for finite element model updating. *Manuscript in preparation.*

This page intentionally left blank.

Incremental reparametrization in sensitivity-based subset selection and parameter clustering for finite element model updating

Daniel T. Bartilson[a,*], Jinwoo Jang[b], Andrew W. Smyth[a]

[a] *Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027, USA*
[b] *Department of Civil, Environmental & Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA*

Abstract

Finite element (FE) model updating improves the prediction value of an FE model by adjusting parameters to enhance correspondence with measured responses. Parametrizing the FE model is non-trivial, with significant impact on the efficacy in reducing errors and posedness of the updating problem. The concept of incremental reparametrization is proposed and explored as an extensible framework for extracting further utility from an FE model with a reduced set of updating parameters. This non-intrusive approach successively creates new parametrizations of an FE model by examining the sensitivity matrix after each model updating increment. A deterministic parameter estimation algorithm is proposed which ties asymptotic Bayes inference to regularization such that new parametrizations are consistently penalized. The proposed reparametrization scheme is tested on natural frequency and mode shape updating exercises, including a small-scale model with analytical data and a benchmark problem with experimental data. In both examples, the proposed reparametrization scheme is strongly supported for subset selection, while it exhibits weaker returns for parameter clustering.

Keywords: Finite element model updating, Parametrization, Sensitivity analysis, Subset selection, Parameter clustering

1. Introduction

Finite element (FE) model updating is a well-developed and common technique for adjusting the parameters of an FE model to reduce discrepancy between measured and model-output data [1, 2]. An updated model has greater utility for response prediction, and may provide further insight into the structural state when model idealization and discretization errors are controlled [2]. For structural applications, the measurements are often obtained from vibration studies, giving such data as natural frequencies, mode shapes, time series, and response functions.

Generally, there are a large number of possible FE model parameters which may be adjusted to update an FE model, such as element mass and stiffness properties, connection and boundary condition spring coefficients, or even geometry. It is typically impractical to individually update all parameters, even with a moderately-sized FE model, because the problem rapidly becomes ill-posed. Ill-posedness is a property of the parameter estimation problem, in which there is insufficient information to constrain the parameter estimates. For deterministic model updating, such as Gauss–Newton minimization, this is understood as an underdetermined system of equations, with more parameters than effective measurements [2, 3]. For uncertainty quantification (UQ) methods of model updating, such as Monte Carlo-based sampling, this is typically referred to as unidentifiability [4]. For an excellent review of deterministic and UQ methods of model updating, the reader is referred to Simoen *et al.* [3].

There are two principal remedies for ill-posedness: regularization and reparametrization. Regularization has seen significant study in FE model updating [5–7] and creates a side-constraint, usually to minimize the norm of parameter changes. This introduces additional equations to overcome ill-posedness, which must be appropriately balanced against the main objective of reducing the residual between measured and model-output data. In any form

*Corresponding author
Email addresses: dtb2121@columbia.edu (Daniel T. Bartilson), jangj@fau.edu (Jinwoo Jang), smyth@civil.columbia.edu (Andrew W. Smyth)
URL: http://faculty.eng.fau.edu/jangj/ (Jinwoo Jang), http://www.columbia.edu/cu/civileng/smyth/ (Andrew W. Smyth)

of regularization, however, the additional constraints are unrelated to the original objective and tend to pull the solution away from its maximum likelihood result [8].

The other common counter for ill-posedness is reparametrization, which is any systematic method for selecting a new (usually reduced) set of parameters to update the underlying FE model. Many parametrization methods, such as subset selection and parameter clustering, use sensitivity analysis to choose an effective set of elements or substructures to modify. Nearly all sensitivity-based parametrization methods are based on the local sensitivity matrix, or Jacobian, evaluated at the initial model parameter state. This is seemingly appropriate when the initial parameter estimates are close to their final estimates, and the sensitivity matrix is unlikely to vary significantly during updating. However, it is difficult to guarantee this condition and a more efficient parametrization might be found at the end of one updating increment.

In this work, incremental reparametrization is explored in the context of forward subset selection and sensitivity-based parameter clustering for FE model updating. This is phrased as a model selection problem in which the relative evidence is evaluated for reparametrization. In Section 2, an objective function is defined which, for the purposes of this work, uses both natural frequency and mode shape data. A review of subset selection and parameter clustering methods is given in Section 3. The incremental reparametrization scheme is detailed along with a novel completion criterion for detecting permutations of the current parametrization. The chosen parameter identification algorithm is detailed in Section 4, which is based on the asymptotic Bayesian inference scheme implemented in previous work by the authors [7]. This approach is adapted to accommodate reparametrization and provide model evidence estimates. The proposed reparametrization scheme is then tested on two model updating exercises: a small-scale truss with analytical data (Section 5) and a Structural Health Monitoring (SHM) benchmark structure with experimental data (Section 6). Conclusions are drawn in Section 7.

2. Objective definition

FE model updating generally requires the following items:

1. Measured data from a structure, $\tilde{\boldsymbol{z}}$;
2. A parametrization of an FE model, $\mathcal{M}$, with parameters $\boldsymbol{\theta}$ and which outputs data similar to the measured data $\boldsymbol{z}(\boldsymbol{\theta})$;
3. An objective function which penalizes discrepancies between measured and model-output data;
4. A parameter estimation algorithm which produces a point, range, or distribution of optimal parameter values.

It is common to write the measured data $\tilde{\boldsymbol{z}}$ (item 1) as a column vector with m components. While this approach naturally extends to many kinds of measured data, in this work $\tilde{\boldsymbol{z}}$ is written as a concatenation of natural frequency and mode shape data:

$$\tilde{\boldsymbol{z}} = [\tilde{f}_1 \ \cdots \ \tilde{f}_l \quad \tilde{\boldsymbol{\psi}}_1^T \ \cdots \ \tilde{\boldsymbol{\psi}}_l^T]^T \tag{1}$$

This comprises l measured modes, each with a natural frequency $\tilde{f}_j$ and mode shape vector $\tilde{\boldsymbol{\psi}}_j$ which is unit-normalized (i.e. $\tilde{\boldsymbol{\psi}}_j^T \tilde{\boldsymbol{\psi}}_j = 1$). Since each mode shape vector has n components corresponding to n measured degrees-of-freedom (DoFs), the total number of measurements is $m = l(n+1)$.

FE model parametrization (item 2) is covered in greater depth in Section 3. For now, it is sufficient to assume a parametrization $\mathcal{M}$ exists which takes in a column vector of p parameters, $\boldsymbol{\theta}$, and outputs a column vector of m components, $\boldsymbol{z}(\boldsymbol{\theta})$. The FE model output data should estimate similar quantities to the measured data, so for the purposes of this work, it is given as

$$\boldsymbol{z}(\boldsymbol{\theta}) = [f_1(\boldsymbol{\theta}) \ \cdots \ f_l(\boldsymbol{\theta}) \quad \mu_1 \boldsymbol{\psi}_1^T(\boldsymbol{\theta}) \ \cdots \ \mu_l \boldsymbol{\psi}_l^T(\boldsymbol{\theta})]^T \tag{2}$$

where there are l corresponding natural frequency, $f_j(\boldsymbol{\theta})$, and unit-normalized mode shape, $\boldsymbol{\psi}_j(\boldsymbol{\theta})$, outputs. μ_j is the j^{th} modal scale factor [9] which minimizes the square difference between measured and model-output modes

$$\mu_j = \tilde{\boldsymbol{\psi}}_j^T \boldsymbol{\psi}_j(\boldsymbol{\theta}) \tag{3}$$

and is also a function of the parameters $\boldsymbol{\theta}$. Correlation between measured and model-output modes should be established by mode pairing [1], which generally matches modes on the basis of Modal Assurance Criterion (MAC) [9]. The MAC between measured mode j and model-output mode k is given as

$$\text{MAC}(\tilde{\boldsymbol{\psi}}_j, \boldsymbol{\psi}_k) = \frac{(\tilde{\boldsymbol{\psi}}_j^T \boldsymbol{\psi}_k)^2}{\tilde{\boldsymbol{\psi}}_j^T \tilde{\boldsymbol{\psi}}_j \cdot \boldsymbol{\psi}_k^T \boldsymbol{\psi}_k} \tag{4}$$

and ranges between 0 (orthogonal mode shapes) and 1 (parallel mode shapes). Pairs are established which maximize the MAC, and the model-output modes are sorted to reflect their corresponding order in the measured data.

A common choice for objective function (item 3) is the weighted sum-of-square residual, E_r, where the residual vector $\boldsymbol{r}(\boldsymbol{\theta})$ is simply the difference between measured and model-output data vectors:

$$E_r = \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} \tag{5}$$

$$\boldsymbol{r}(\boldsymbol{\theta}) = \tilde{\boldsymbol{z}} - \boldsymbol{z}(\boldsymbol{\theta}) \tag{6}$$

$\boldsymbol{W}_r$ is the $m \times m$ residual weighting matrix, which should represent the uncertainty in the measurement vector $\tilde{\boldsymbol{z}}$. The statistically-optimal (in a Bayesian sense) choice of $\boldsymbol{W}_r$ is the inverse of the measurement covariance matrix $\boldsymbol{C}_{\tilde{z}}$ [10], where both $\boldsymbol{C}_{\tilde{z}}$ and $\boldsymbol{W}_r$ are symmetric and positive semi-definite. It is typical to assume statistical independence between measurements, making $\boldsymbol{C}_{\tilde{z}}$ and $\boldsymbol{W}_r$ both diagonal. A diagonal $\boldsymbol{W}_r$ can be written as

$$\text{diag}[\boldsymbol{W}_r] = \left\{ w_{r1}^f \ \cdots \ w_{rl}^f \quad \text{diag}[\boldsymbol{W}_{r1}^\psi] \ \cdots \ \text{diag}[\boldsymbol{W}_{rl}^\psi] \right\} \tag{7}$$

where w_{rj}^f is the j^{th} natural frequency weighting term, and $\boldsymbol{W}_{rj}^\psi$ is the corresponding diagonal mode shape weighting matrix. When a weighting matrix like that in Eq. (7) is used in Eq. (5), then the following separation can be made between natural frequency and mode shape errors.

$$E_r = \sum_{j=1}^{l} w_{rj}^f (\tilde{f}_j - f_j)^2 + \sum_{j=1}^{l} \{\tilde{\boldsymbol{\psi}}_j - \mu_j \boldsymbol{\psi}_j\}^T \boldsymbol{W}_{rj}^\psi \{\tilde{\boldsymbol{\psi}}_j - \mu_j \boldsymbol{\psi}_j\} \tag{8}$$

For further discussion of this objective function, including partitioning and analytical sensitivity calculations, please refer to previous work by the authors [7].

3. Parametrization

With a defined set of measured data $\tilde{\boldsymbol{z}}$, and an objective function E_r to minimize, the next requirement for FE model updating is a parametrized FE model to generate corresponding data. There are many parametrization schemes available in the literature, and each parametrization scheme may be able to generate a large number of possible parametrizations for a given model. Model parametrizations should satisfy three conditions [1]:

1. a limited number of parameters should be used to avoid ill-posedness,
2. the parametrization should be chosen to correct model uncertainties, and
3. model-outputs should be sensitive to the chosen parameters.

Fulfillment of these requirements generally requires expert knowledge of FE model sensitivities and uncertainties. Mottershead *et al.* [11] studied several parametrizations of an aluminum frame, some based on consideration of model uncertainties and others that were mostly numerical. While these methods may reduce the residual, there is no guarantee that the model-outputs will be sensitive to the selected parameters or that posedness will be improved. Parametrizations which incorporate model sensitivities therefore have a large advantage in that they are designed to provide sensitive parameters and counter ill-posedness. These methods commonly select a reduced subset of FE model parameters [5, 12, 13], or retain all FE model parameters but group them into clusters (substructures) of 'similar' elements which are updated by single parameters [2, 7, 14–16].

The common starting place for sensitivity-based parametrizations is the base FE model parametrization, denoted as $\boldsymbol{\delta}$. A typical choice of $\boldsymbol{\delta}$ is the set of parameters which linearly modify the element mass $\boldsymbol{M}_e$ and/or stiffness $\boldsymbol{K}_e$ matrices for each element e out of a total n_{el}. Therefore, the global mass $\boldsymbol{M}$ and stiffness $\boldsymbol{K}$ matrices (which are $N \times N$ for N FE model DoFs) are given as functions of $\boldsymbol{\delta}$:

$$\boldsymbol{M}(\boldsymbol{\delta}) = \sum_{e=1}^{n_{el}} \boldsymbol{M}_e (1 - \delta_e^M) = \boldsymbol{M}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{M}_e \delta_e^M \tag{9}$$

$$\boldsymbol{K}(\boldsymbol{\delta}) = \sum_{e=1}^{n_{el}} \boldsymbol{K}_e (1 - \delta_e^K) = \boldsymbol{K}_0 - \sum_{e=1}^{n_{el}} \boldsymbol{K}_e \delta_e^K \tag{10}$$

where $\boldsymbol{M}_0$ and $\boldsymbol{K}_0$ are the unmodified mass and stiffness matrices, respectively. $\boldsymbol{\delta}$ is thus a column vector of $d = 2\,n_{\text{el}}$ components when both mass and stiffness are modified, or $d = n_{\text{el}}$ when only one matrix is modified.

Within this work, an incremental reparametrization scheme is defined which uses a different set of updating parameters in each increment i. Therefore, it is convenient to define the updating parameters for the i^{th} increment as $\boldsymbol{\theta}^i$, while the FE model parameters at the beginning of the increment are denoted as $\boldsymbol{\delta}^i$. $\boldsymbol{\delta}^i$ can simply be thought of as initial values for the FE model parameters in a typical model updating scheme. For the parametrizations in this work, the relationship between $\boldsymbol{\delta}$ and $\boldsymbol{\theta}^i$ can be defined by a linear transformation matrix $\boldsymbol{T}^i$, such that

$$\boldsymbol{\delta} = \boldsymbol{\delta}^i + \boldsymbol{T}^i\boldsymbol{\theta}^i \tag{11}$$

This parametrization is denoted as $\mathcal{M}^i$, and encompasses the initial values of the FE model parameters $\boldsymbol{\delta}^i$, the column vector of $p \leq d$ updating parameters $\boldsymbol{\theta}^i$, and the $d \times p$ transformation matrix $\boldsymbol{T}^i$. Importantly, since $\boldsymbol{\delta}^i$ is a fixed quantity for the parametrization $\mathcal{M}^i$, then the Jacobian of $\boldsymbol{\delta}$ with respect to $\boldsymbol{\theta}^i$ is simply $\boldsymbol{T}^i$. This can be exploited to show that the Jacobian of the residual $\boldsymbol{r}$ with respect to $\boldsymbol{\theta}^i$, $\boldsymbol{J}^i_j$, is related to the Jacobian of the residual with respect to $\boldsymbol{\delta}$, $\boldsymbol{S}^i_j$:

$$\boldsymbol{J}^i_j = \left.\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\theta}^i}\right|_{\boldsymbol{\theta}^i = \boldsymbol{\theta}^i_j} = \left.\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\delta}}\right|_{\boldsymbol{\delta} = \boldsymbol{\delta}^i + \boldsymbol{T}^i\boldsymbol{\theta}^i_j} \frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{\theta}^i} = \boldsymbol{S}^i_j\boldsymbol{T}^i \tag{12}$$

where the Jacobians are evaluated at a certain value $\boldsymbol{\theta}^i_j$.

Two parametrization methods, namely forward subset selection (Section 3.1) and sensitivity-based parameter clustering (Section 3.2), are described for producing the transformation matrix $\boldsymbol{T}^i$. The incremental reparametrization scheme is described in Section 3.3.

3.1. Parameter subset selection

Subset selection attempts to find a reduced number of model parameters which produce a successful updating result according to an objective function such as minimum residual or maximum evidence, and was originally applied in regression [17]. An exhaustive search of all subsets may be practical for small problems, but heuristic methods are essential for most practical problems. A variety of selection methods exist in FE model updating, but the most common heuristic subset selection method is forward subset selection, first applied to FE model updating by Lallement and Piranda [12] and used in further work [5, 13]. This method, simply called subset selection from here-on, successively selects parameters based on their similarity of their sensitivity to the residual vector during an orthogonalization process.

First, consider updating using an FE model parametrization $\boldsymbol{\delta}$ with d parameters. This can be stated as "find the parameter update $\Delta\boldsymbol{\delta}^i$ which minimizes the objective function $E_r(\boldsymbol{\delta}^i + \Delta\boldsymbol{\delta}^i)$":

$$E_r(\boldsymbol{\delta}^i + \Delta\boldsymbol{\delta}^i) = \|\boldsymbol{q}^i + \boldsymbol{G}^i\Delta\boldsymbol{\delta}^i\|_2^2 \tag{13}$$

where $\|\ \|_2$ is the L^2 norm. $\boldsymbol{q}^i$ is the weighted residual and $\boldsymbol{G}^i$ is the weighted Jacobian, or equivalently, the Jacobian of $\boldsymbol{q}^i$:

$$\boldsymbol{q}^i = \boldsymbol{W}_r^{1/2}\boldsymbol{r}^i; \qquad \boldsymbol{G}^i = \boldsymbol{W}_r^{1/2}\boldsymbol{S}^i = \left.\frac{\partial \boldsymbol{q}^i}{\partial \boldsymbol{\delta}}\right|_{\boldsymbol{\delta}=\boldsymbol{\delta}^i}; \qquad \boldsymbol{S}^i = \left.\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\delta}}\right|_{\boldsymbol{\delta}=\boldsymbol{\delta}^i}; \tag{14}$$

Subset selection seeks a reduced set of $p < d$ parameters from $\boldsymbol{\delta}$ which minimize E_r, using an iterative process to greedily select one parameter at a time.

The process begins at selection iteration $a = 1$ by identifying and selecting the parameter δ_b which (by itself) provides the greatest reduction in E_r:

$$b = \arg\min_{j \in 1,..,d} \|\boldsymbol{q}^i + \boldsymbol{g}^i_j\widehat{\Delta\delta^i_j}\|_2^2 \tag{15}$$

where $\boldsymbol{g}^i_j$ is the weighted sensitivity vector of parameter j, or equivalently the j^{th} column of $\boldsymbol{G}^i = [\boldsymbol{g}^i_1\ \cdots\ \boldsymbol{g}^i_d]$ and $\widehat{\Delta\delta^i_j} = -\boldsymbol{g}^{iT}_j\boldsymbol{q}^i/\boldsymbol{g}^{iT}_j\boldsymbol{g}^i_j$ is the least-squares estimate for the j^{th} parameter. This is equivalent to choosing the

parameter which has the least angle between its sensitivity vector and the residual [13]. The process then continues by orthogonalizing the sensitivity vectors and residual with respect to the the sensitivity for δ_b, $\boldsymbol{g}_b^i$:

$$\boldsymbol{g}_j^i \leftarrow \boldsymbol{g}_j^i - \boldsymbol{g}_b^i(\boldsymbol{g}_b^{iT}\boldsymbol{g}_j/\boldsymbol{g}_j^{iT}\boldsymbol{g}_j^i); \qquad \boldsymbol{q}^i \leftarrow \boldsymbol{q}^i + \boldsymbol{g}_b^i\widehat{\Delta\delta_b^i} \tag{16}$$

A new iteration is started using these orthogonalized quantities, identifying the parameter which minimizes the cosine distance between the modified $\boldsymbol{q}^i$ and the modified columns of $\boldsymbol{G}^i$. This process is performed for all iterations a until p parameters are selected. This orthogonalization process ensures that each parameter has a different effect on reducing residual (i.e. sensitivity), which should improve posedness by excluding parameters with linearly-dependent sensitivities.

The updating parameters correspond to individual FE model parameters, so this is equivalent to directly updating p FE model parameters while keeping the other $d-p$ fixed. The $d \times p$ transformation matrix $\boldsymbol{T}^i$ can be written as

$$T_{ba}^i = \begin{cases} 1 & \delta_b \text{ chosen in selection iteration } a \text{ (i.e. } \delta_b \text{ updated by } \theta_a^i) \\ 0 & \text{else} \end{cases} \tag{17}$$

This matrix is orthogonal ($\boldsymbol{T}^{iT}\boldsymbol{T}^i = \boldsymbol{I}_p$) since each of the p columns is a unique member of the standard basis of d-space.

3.2. Sensitivity-based parameter clustering

Where subset selection only updates a reduced set of FE model parameters, parameter clustering updates all of the FE model parameters but groups model parameters into clusters which are updated together, giving a reduced parametrization. This approach was first proposed by Friswell *et al.* [13] as an extension of subset selection, and was further developed by Shahverdi *et al.* [14] to use hierarchical clustering based on the cosine distance between parameter sensitivity vectors. This approach was also used by the authors in previous work [7, 15, 16].

Hierarchical clustering [18] requires a distance measure to evaluate similarity between parameters and a linkage method for determining the distance between candidate clusters. For sensitivity-based parameter clustering, cosine distance is usually selected as the distance measure. The cosine distance between the weighted sensitivity vectors for parameters δ_j and δ_k at the beginning of increment i is

$$d_{\cos}(\boldsymbol{g}_j^i, \boldsymbol{g}_k^i) = 1 - \boldsymbol{g}_j^{iT}\boldsymbol{g}_k^i/(\boldsymbol{g}_j^{iT}\boldsymbol{g}_j^i \cdot \boldsymbol{g}_k^{iT}\boldsymbol{g}_k^i)^{1/2} \tag{18}$$

which ranges between 0 (parallel) and 2 (anti-parallel). The chosen linkage method is the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [18], which combines clusters based on the distance between unweighted cluster means. At each step, the most similar two elements (clusters or single parameters), measured by distance between unweighted cluster means, are combined to form a new cluster. This builds a dendrogram where 'branches' combine with increasing distance criterion. Thus, for a specified maximum linkage distance, the dendrogram can be 'cut', resulting in a number of clusters. For the purposes of this work, the number of clusters p may also be specified, requiring that the dendrogram is cut at such a level to give the prescribed number of clusters. Since sensitivity-based parameter clustering combines FE model parameters with similar sensitivity vectors, linear dependence is reduced in the resulting Jacobian matrix, $\boldsymbol{J}^i$, thereby improving posedness.

The $d \times p$ transformation matrix for the i^{th} parametrization, $\boldsymbol{T}^i$, can be written as

$$T_{ba}^i = \begin{cases} |\mathcal{A}|^{-1/2} & \delta_b \text{ is part of cluster } a \text{ (i.e. } \delta_b \text{ updated by } \theta_a^i) \\ 0 & \text{else} \end{cases} \tag{19}$$

where $|\mathcal{A}|$ is the cardinality of cluster a. When defined in this way, $\boldsymbol{T}^i$ is orthogonal ($\boldsymbol{T}^{iT}\boldsymbol{T}^i = \boldsymbol{I}_p$) because each parameter is assigned to exactly one cluster and each column is unit-normalized.

3.3. Incremental reparametrization for sensitivity-based parametrization methods

Subset selection and parameter clustering methods, as part of the broader class of sensitivity-based parametrization techniques, are generally based on the sensitivity of residual $\boldsymbol{r}$ to FE model parameters $\boldsymbol{\delta}$ at the initial FE model state, $\boldsymbol{S}^0$ (see Eq. (14)). This is used to generate a parametrization $\mathcal{M}^i$ for the $i = 0$ increment, which uses the parameters $\boldsymbol{\theta}^i$ and transformation matrix $\boldsymbol{T}^i$. This is then optimized using an algorithm to give a point estimate for the optimal $\boldsymbol{\theta}^i$, denoted as $\hat{\boldsymbol{\theta}}^i$. This may be a deterministic parameter solution, as will be covered in Section 4, or it

could be the maximum likelihood or maximum a posteriori estimate from a probabilistic scheme. This provides the FE model parameter update as

$$\Delta\boldsymbol{\delta}^i = \boldsymbol{T}^i\hat{\boldsymbol{\theta}}^i \tag{20}$$

which is then used to increment the FE model parameters:

$$\boldsymbol{\delta}^{i+1} = \boldsymbol{\delta}^i + \Delta\boldsymbol{\delta}^i \tag{21}$$

This is commonly the final result of a FE model updating exercise. However, it is possible that the sensitivity matrix at the end of one increment, $\boldsymbol{S}^1$, is considerably different to $\boldsymbol{S}^0$, leading to new possible parametrizations and further error reduction. This is particularly applicable to forward subset selection methods because parameters are chosen based on similarity to the residual vector. Since both the parameter sensitivities and residual vector will change during updating, it is very likely that a new parametrization will be more appropriate after one updating increment.

Therefore, a technique called incremental reparametrization is proposed. First, it is necessary to define a parametrization method $\mathcal{M}_k$, which can be used at each increment i, giving increment parametrizations $\mathcal{M}_k^i$. Then, starting at increment $i = 0$ with a defined initial FE model parameter vector $\boldsymbol{\delta}^0$, incremental reparametrization proceeds as follows:

1. Evaluate the FE model parameter sensitivities at the current increment $\boldsymbol{S}^i = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\delta}}|_{\boldsymbol{\delta}=\boldsymbol{\delta}^i}$.
2. Generate a parametrization for current increment $\mathcal{M}_k^i$ such that $\Delta\boldsymbol{\delta}^i = \boldsymbol{T}^i\boldsymbol{\theta}^i$.
3. Check if the current parametrization is equal to the last parametrization. If so, exit.
4. Estimate optimal parameters $\hat{\boldsymbol{\theta}}^i$ and update FE model parameters $\boldsymbol{\delta}^{i+1} = \boldsymbol{\delta}^i + \boldsymbol{T}^i\hat{\boldsymbol{\theta}}^i$.
5. Increase the increment index $i \leftarrow i + 1$ and return to step 1.

For a deterministic scheme such as the sensitivity method, the optimal parameters are estimated by an iterative method within each increment. A depiction of this process is shown in Fig. 1 with an illustration of the decrease in an objective value, such as the L^2 norm of residual. Since each increment is treated as a separate non-linear problem, the number of iterations in each increment can be variable.

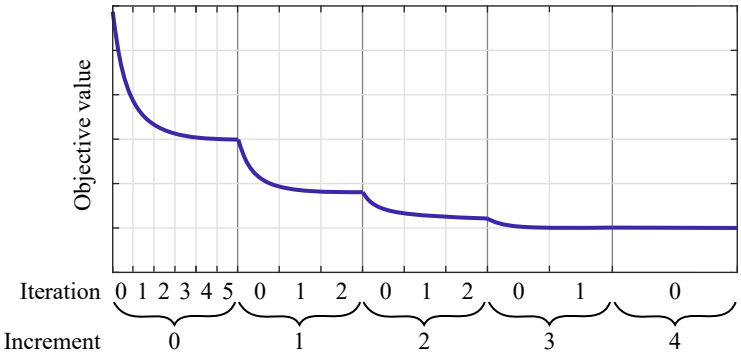


Figure 1: Illustration of proposed incrementation scheme

The proposed algorithm is non-intrusive. Steps 1, 2, and 4 comprise a typical sensitivity-based model updating scheme, while steps 3 and 5 are simple control logic. In fact, this process may be thought of as performing several steps of model updating, where each new step is initialized at the results of the previous step. A large benefit of this approach is that the number of updating parameters p can be kept constant across all increments, which is a critical part of ensuring posedness in model updating.

The only step which is not obvious or previously covered in this or other work is step 3. While it may seem possible to simply test if $\boldsymbol{T}^i = \boldsymbol{T}^{i-1}$, this does not work for situations in which parameters are simply relabeled or reordered. For example, in parameter clustering, cluster indices are essentially arbitrary. Therefore, a test for parametrization equivalence is proposed which recognizes cases where the columns of $\boldsymbol{T}^i$ are a permutation of the columns of $\boldsymbol{T}^{i-1}$. Define $\boldsymbol{Q}^{hi}$ between the parametrizations at increments h and i as

$$\boldsymbol{Q}^{hi} = \boldsymbol{T}^{hT}\boldsymbol{T}^i \tag{22}$$

where $\boldsymbol{T}^h$ and $\boldsymbol{T}^i$ are not necessarily orthogonal, but each column is normalized such that diag$[\boldsymbol{T}^T\boldsymbol{T}]$ is a vector of 1s for both parametrizations. If $\boldsymbol{T}^h$ and $\boldsymbol{T}^i$ are column-permutations is readily shown that $\boldsymbol{Q}^{hi}$ is a 'quasi-permutation' matrix with exactly one entry of 1 in each row and column and values less than 1 elsewhere. If $\boldsymbol{T}^h$ and $\boldsymbol{T}^i$ are not column-permutations, then one or more columns and rows of $\boldsymbol{Q}^{hi}$ will not have exactly one entry of 1, making this simple to check algorithmically. Furthermore, if $\boldsymbol{T}^h$ and $\boldsymbol{T}^i$ are each orthogonal, as established for subset selection and parameter clustering, then $\boldsymbol{Q}^{hi}$ will be a true permutation matrix when $\boldsymbol{T}^h$ and $\boldsymbol{T}^i$ are column-permutations, with exactly one entry of 1 in each row and column and 0 in all other entries.

Convergence is an important consideration in this incremental scheme. Within each increment, most parameter estimation algorithms attempt to solve a (locally) convex optimization problem to find the optimal $\boldsymbol{\theta}^i$ to minimize E_r or some related objective, which is a convergent problem. Thus, the question is whether or not the parametrizations (increments) will necessarily converge. Under the requirement that the reducible error in E_r is finite, then there can only be a finite number of increments which reduce E_r (within the limits of computer precision). Therefore, it is important to select a parameter estimation algorithm which only performs updates which reduce E_r, otherwise it may be possible to select an infinite number of new parametrizations without affecting the objective, creating an infinite loop. Within this work, it was generally noted by the authors that the amount of parameter change $\Delta\boldsymbol{\delta}^i$ tended to reduce with successive increments. This resulted in decreasing changes in parameter sensitivities, leading toward parametrization convergence as illustrated in Fig. 1.

4. Levenberg–Marquardt minimization scheme with asymptotic Bayesian inference

With a defined set of measured data, an objective function (Section 2), and a parametrized FE model (Section 3), the final step in model updating is to select a parameter estimation algorithm. UQ methods are growing in popularity because they naturally incorporate measurement and model uncertainties into their results, giving a range or distribution of parameter estimates [3]. However, they are typically orders-of-magnitude more computationally expensive than deterministic methods and may be prohibitive for full-scale model updating. Deterministic model updating provides a computationally inexpensive method for obtaining a local or global optimum parameter estimate. The sensitivity method [2] is an intuitive local approach for deterministic model updating, and includes a wide-class of methods for performing Gauss–Newton-type optimization of a quadratic residual function, like in Eq. (5).

While the sensitivity method isn't typically formulated to provide UQ, regularization can be used to provide a natural link to Bayesian inference [3, 8]. This approach was used previously by the authors [7] to provide optimal regularization based on the algorithm from Foresee and Hagan [19]. This approach hinges on the use of Gaussian distributions, or alternately on asymptotic assumptions, which has previously seen use in FE model updating [3, 20].

The regularized objective function F is written as a weighted sum of the quadratic residual term E_r from Eq. (5) and a quadratic penalty term E_δ:

$$F = \beta E_r + \alpha E_\delta = \beta \boldsymbol{r}^T \boldsymbol{W}_r \boldsymbol{r} + \alpha \boldsymbol{\delta}^T \boldsymbol{W}_\delta \boldsymbol{\delta} \tag{23}$$

These two objectives of minimizing residual and limiting FE model parameter changes are balanced by the regularization parameters α and β. In this work, the regularization parameter α is determined to maximize model evidence, as detailed in Section 4.1. Similar to $\boldsymbol{W}_r$, the FE model parameter weighting matrix $\boldsymbol{W}_\delta$ should reflect the uncertainty in the FE model parameters, optimally given as the inverse of the parameter covariance matrix [10]. One of the main contributions of this work is regularization with respect to the FE model parameters $\boldsymbol{\delta}$ rather than the updating parameters at each increment, $\boldsymbol{\theta}^i$. This allows for consistent penalization even after reparametrization. Additionally, regularizing with respect to $\boldsymbol{\delta}$ should theoretically improve the convexity of the incremental updating problem, since new parametrizations and further updates to $\boldsymbol{\delta}$ are penalized as part of the objective function. This helps create a stable minimum at which point further parametrization is sufficiently penalized, leading toward convergence.

In the deterministic Levenberg–Marquardt minimization scheme proposed in this work, FE model parameters are evolved from one increment to the next using the previous values $\boldsymbol{\delta}^i$ which are updated by the point estimate for the optimal increment updating parameters, $\hat{\boldsymbol{\theta}}^i$, and the corresponding transformation matrix $\boldsymbol{T}^i$:

$$\boldsymbol{\delta}^{i+1} = \boldsymbol{\delta}^i + \boldsymbol{T}^i \hat{\boldsymbol{\theta}}^i \tag{24}$$

Within each increment i, the updating parameters are evolved from iteration j to $j+1$ until convergence:

$$\boldsymbol{\theta}^i_{j+1} = \boldsymbol{\theta}^i_j + \Delta\boldsymbol{\theta}^i_j \tag{25}$$

It is notationally convenient to refer to the FE model parameters at increment i and iteration j as

$$\boldsymbol{\delta}_j^i = \boldsymbol{\delta}^i + \boldsymbol{T}^i \boldsymbol{\theta}_j^i \tag{26}$$

and equivalently, any quantity such as the residual $\boldsymbol{r}$ evaluated at $\boldsymbol{\delta}_j^i$ may be indicated as

$$\boldsymbol{r}(\boldsymbol{\delta}_j^i) = \boldsymbol{r}_j^i \tag{27}$$

The residual at iteration $j+1$ is approximated by the truncated Taylor series:

$$\boldsymbol{r}_{j+1}^i \approx \boldsymbol{r}_j^i + \boldsymbol{J}_j^i \Delta\boldsymbol{\theta}_j^i; \tag{28}$$

which uses the Jacobian with respect to $\boldsymbol{\theta}^i$ evaluated at $\boldsymbol{\delta}_j^i$, indicated as $\boldsymbol{J}_j^i$. This is used in Eq. (23) to approximate the updated regularized objective value:

$$F_{j+1}^i = \beta \boldsymbol{r}_{j+1}^{iT} \boldsymbol{W}_r \boldsymbol{r}_{j+1}^i + \alpha \boldsymbol{\delta}_{j+1}^{iT} \boldsymbol{W}_\delta \boldsymbol{\delta}_{j+1}^i \tag{29}$$

where the expression for the updated parameter values is exact: $\boldsymbol{\delta}_{j+1}^i = \boldsymbol{\delta}^i + \boldsymbol{T}^i \boldsymbol{\theta}_{j+1}^i$. The approximated objective F is minimized with respect to the parameter update, such that

$$\frac{\partial F_{j+1}^i}{\partial(\Delta\boldsymbol{\theta}_j^i)} = 0 \tag{30}$$

The resulting parameter update is then given by the Gauss–Newton solution

$$\Delta\boldsymbol{\theta}_j^i = -2[\bar{\boldsymbol{H}}_j^i]^{-1}[\beta \boldsymbol{J}_j^{iT} \boldsymbol{W}_r \boldsymbol{r}_j^i + \alpha \boldsymbol{T}^{iT} \boldsymbol{W}_\delta \boldsymbol{\delta}_j^i] \tag{31}$$

where $\bar{\boldsymbol{H}}_j^i$ is the approximate Hessian of F with respect to $\boldsymbol{\theta}^i$ evaluated at $\boldsymbol{\delta}_j^i$:

$$\bar{\boldsymbol{H}}_j^i = 2[\beta \boldsymbol{J}_j^{iT} \boldsymbol{W}_r \boldsymbol{J}_j^i + \alpha \boldsymbol{W}_\theta^i] \tag{32}$$

and $\boldsymbol{W}_\theta^i$ is the weighting matrix for the i^{th} increment updating parameters:

$$\boldsymbol{W}_\theta^i = \boldsymbol{T}^{iT} \boldsymbol{W}_\delta \boldsymbol{T}^i \tag{33}$$

The Gauss–Newton solution can be improved into the more robust Levenberg–Marquardt [21, 22] solution by augmenting the Hessian with a damping term λ along the diagonal (unrelated to damping in structural vibration):

$$\Delta\boldsymbol{\theta}_j^i = -2[\bar{\boldsymbol{H}}_j^i + 2\lambda \boldsymbol{I}_p]^{-1}[\beta \boldsymbol{J}_j^{iT} \boldsymbol{W}_r \boldsymbol{r}_j^i + \alpha \boldsymbol{T}^{iT} \boldsymbol{W}_\delta \boldsymbol{\delta}_j^i] \tag{34}$$

Notably, Eq. (34) becomes the Gauss–Newton update as $\lambda \to 0$ and to the infinitesimal gradient descent update as $\lambda \to \infty$. λ is controlled by the multiplicative scheme developed by Marquardt [22], with further description given in Table 1.

4.1. Asymptotic Bayesian inference for optimal regularization and evidence estimation

The determination of regularization parameters in FE model updating has typically been done by simple heuristics or by expensive L-curve methods, with an excellent review by Titurus and Friswell [6]. The approach used in this paper relates regularization to Bayesian inference, as done by MacKay [8]. The authors proposed a similar approach in previous work [7], which adapted the algorithm by Foresee and Hagan [19]. This technique is now expanded to handle the proposed change in regularization from $\boldsymbol{\theta}$ to $\boldsymbol{\delta}$, and β is fixed to improve model evidence estimates.

Given a model or parametrization $\mathcal{M}_k^i$, data $\tilde{\boldsymbol{z}}$, and values for α and β, the posterior probability of the updating parameters is given by the Bayes expression

$$P(\boldsymbol{\theta}^i|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}_k^i) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}^i, \beta, \mathcal{M}_k^i) P(\boldsymbol{\theta}^i|\alpha, \mathcal{M}_k^i)}{P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_k^i)} \tag{35}$$

where $P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}^i, \beta, \mathcal{M}_k^i)$ is the likelihood of the data at a specified parameter value, $P(\boldsymbol{\theta}^i|\alpha, \mathcal{M}_k^i)$ is the prior probability of updating parameters, and $P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_k^i)$ is the evidence for the model $\mathcal{M}_k^i$. If the measurements are assumed to come from a Gaussian distribution with mean vector $\tilde{\boldsymbol{z}}$ and covariance matrix $\boldsymbol{C}_{\tilde{z}} = [2\beta \boldsymbol{W}_r]^{-1}$, then the likelihood can be expressed as

$$P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}^i, \beta, \mathcal{M}_k^i) = e^{-\beta E_r}/Z_r; \qquad \log Z_r = \frac{m}{2}\log \pi - \frac{1}{2}\log|\beta \boldsymbol{W}_r| \tag{36}$$

where $|\ |$ indicates the determinant for a matrix quantity. Similarly, if the uncertain FE model parameters $\boldsymbol{\delta}$ are assumed to come from a Gaussian distribution with zero mean vector and covariance matrix $\boldsymbol{C}_\delta = [2\alpha \boldsymbol{W}_\delta]^{-1}$, then the prior probability can be stated as

$$P(\boldsymbol{\theta}^i|\alpha, \mathcal{M}_k^i) = e^{-\alpha E_\delta}/Z_\delta; \qquad \log Z_\delta = -\alpha E_\delta^{i\star} + \frac{p}{2}\log \pi - \frac{1}{2}\log|\alpha \boldsymbol{W}_\theta^i| \tag{37}$$

which has a special term related to the parameter penalty term at increment i:

$$E_\delta^{i\star} = \boldsymbol{\delta}^{iT}\boldsymbol{W}_\delta \boldsymbol{\delta}^i - \boldsymbol{\delta}^{iT}\boldsymbol{W}_\delta \boldsymbol{T}^{iT}[\boldsymbol{W}_\theta^i]^{-1}\boldsymbol{T}^i \boldsymbol{W}_\delta \boldsymbol{\delta}^i \tag{38}$$

This term vanishes when $\boldsymbol{T}^i$ is invertible, but this is not generally the case since $\boldsymbol{T}^i$ is $d \times p$ with $p < d$ in general. Substituting Eqs. (36) and (37) into Eq. (35) yields the following expression for the posterior probability:

$$P(\boldsymbol{\theta}^i|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}_k^i) = e^{-F}/Z_F; \qquad Z_F = \int e^{-F} d\boldsymbol{\theta}^i \tag{39}$$

The normalization factor Z_F can be approximated using Laplace's method [8], which is appropriate for Gaussian posteriors as well as large data sets that result in highly peaked posteriors [3, 20]. This approximation is

$$\log Z_F \approx -\hat{F}^i + \frac{p}{2}\log 2\pi - \frac{1}{2}\log|\hat{\boldsymbol{H}}^i| \tag{40}$$

where $\hat{\boldsymbol{\theta}}^i$ is the parameter estimate for increment i, which is equivalently occurs at the minimum point of F ($\hat{F}^i$) and the maximum point of the assumed posterior distribution. $\hat{\boldsymbol{H}}^i$ is the Hessian evaluated at $\hat{\boldsymbol{\theta}}^i$, which may be the approximate Hessian as in Eq. (32). These expressions can be combined into Eq. (35) to give the following expression for the model evidence:

$$P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_k^i) = \frac{P(\tilde{\boldsymbol{z}}|\boldsymbol{\theta}^i, \beta, \mathcal{M}_k^i)P(\boldsymbol{\theta}^i|\alpha, \mathcal{M}_k^i)}{P(\boldsymbol{\theta}^i|\tilde{\boldsymbol{z}}, \alpha, \beta, \mathcal{M}_k^i)} = \frac{Z_F}{Z_r Z_\delta} \tag{41}$$

Taking the logarithm and substituting the normalization results from Eqs. (36), (37), and (40) yields the following expression for the log evidence

$$\log P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_k^i) = \underbrace{-\beta \hat{E}_r^i + \frac{1}{2}\log|(\beta/\pi)\boldsymbol{W}_r|}_{\text{log likelihood}} \underbrace{-\alpha(\hat{E}_\delta^i - E_\delta^{i\star}) + \frac{1}{2}\log\left|[\hat{\boldsymbol{H}}^i]^{-1}[2\alpha \boldsymbol{W}_\theta^i]\right|}_{\text{log Occam factor}} \tag{42}$$

where $\hat{E}_r^i$ is the quadratic residual and $\hat{E}_\delta^i$ is the quadratic parameter penalty term, both evaluated at $\hat{\boldsymbol{\theta}}^i$. This separates in into a log likelihood expression and an Occam factor expression. Without regularization, minimizing E_r would simply maximize the likelihood, generally preferring overly complex models and overfitted solutions. The Occam factor is closely related to the regularization and penalizes complexity, enforcing Occam's principle that simpler models are preferred [3, 4, 8]. The first term in the Occam factor penalizes large FE model parameter changes and the second term penalizes parametrization complexity, evaluated as the ratio between the prior and posterior volumes.

It is important to note that the model evidence in Eq. (42) is only an approximation in cases where the true posterior is non-Gaussian. In these cases, it is more appropriate to sample the posterior distribution using a Monte Carlo scheme [3, 4, 16]. However, Eq. (42) is still more informative than likelihood alone and may be used in a similar fashion to the closely-related Schwarz criterion [23, 24]. Therefore, the approximation in Eq. (42) should be applied carefully, but can provide a meaningful balance between data fit and model complexity in most cases.

This approach for determining model evidence has seen use in model updating, often referred to as an asymptotic approach [3, 20]. In most works, however, the prior distribution for the parameters had to be based on uninformed assumptions since little is typically known *a priori*. The main improvement suggested by MacKay [8] is to determine the regularization parameter α to maximize the model evidence. This scales the width of the Gaussian prior, which simultaneously provides optimal regularization. Previous work by the authors [7] also optimized β, but this is inappropriate for model evidence estimation since likelihood (controlled by β) becomes model-dependent and the optimal objective value $F(\hat{\boldsymbol{\theta}}^i)$ becomes fixed at a value of $m/2$ [8].

In this work, β is fixed at a value of $1/2$, such that $2\beta \boldsymbol{W}_r = \boldsymbol{C}_{\tilde{z}}^{-1}$. The parameter α is chosen to maximize the model evidence in Eq. (42), giving

$$\alpha = \frac{\gamma}{2(\hat{E}_\delta^i - E_\delta^{i\star})} \tag{43}$$

with $E_\delta^{i\star}$ from Eq. (38) and γ as the effective number of parameters [8]:

$$\gamma = p - 2\alpha \operatorname{tr}([\hat{\boldsymbol{H}}^i]^{-1} \boldsymbol{W}_\theta^i) \tag{44}$$

Even though this is referred to as a 'number', it should not be taken as an integer quantity. It is most naturally understood as a measure of the relative influence of the data (rather than the regularization) for determining parameter values [8]. γ ranges between 0 and p, or alternatively, the 'parameter efficiency' γ/p ranges between 0 and 1. If γ/p is close to 0, then few of the parameters are updated by the data, such that the posterior is close to the prior. In this case, the solution is dominated by the regularization. When γ/p is close to 1, then the data is much more influential than the regularization in updating the parameters, and the posterior is much more peaked than the prior. MacKay [8] suggests that $\gamma/p \approx 1/2$ is a reasonable result, but for the purposes of this work, a higher γ for the same number of parameters is desirable, implying less influence of regularization.

The proposed solution algorithm, including the elements of incremental reparametrization from Section 3.3, Levenberg–Marquardt minimiziation within each increment, and Bayesian determination of the regularization parameter α, is given in Table 1.

Table 1: Pseudocode for incremental reparametrization with Bayesian-regularized Levenberg–Marquardt minimization

Input: Regularized obj. function $F(\boldsymbol{\delta}) = \beta E_r + \alpha E_\delta$, parametrization method $\mathcal{M}_k$
Output: Opt. param. $\boldsymbol{\delta}^{\mathrm{MP}}$, eff. no. params. γ, log evidence est. $\log P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_k^i)$

initialization: Set $\boldsymbol{\delta}^0$, $\alpha = 0.5$, $\beta = 0.5$, $\lambda = 0.01$, $v = 10$, $i = 0$, $j = 0$;
while *Parametrization* $\mathcal{M}_k^i$ *not converged* **do**
  Generate parametrization for current increment i, $\mathcal{M}_k^i : \boldsymbol{\delta}^{i+1} = \boldsymbol{\delta}^i + \boldsymbol{T}^i \boldsymbol{\theta}^i$;
  Set iteration count to $j = 0$ and initialize increment parameters, $\boldsymbol{\theta}_0^i = \boldsymbol{0}$;
  while *Parameters* $\boldsymbol{\theta}_j^i$ *not converged* **do**
    Compute residual $\boldsymbol{r}_j^i$, Jacobian $\boldsymbol{J}_j^i$, and approx. Hessian $\bar{\boldsymbol{H}}_j^i = 2[\beta \boldsymbol{J}_j^{iT} \boldsymbol{W}_r \boldsymbol{J}^{iT} + \alpha \boldsymbol{W}_\theta^i]$;
    Compute parameter update $\Delta\boldsymbol{\theta}_j^i = -2[\bar{\boldsymbol{H}}_j^i + 2\lambda \boldsymbol{I}_p]^{-1}[\beta \boldsymbol{J}_j^{iT} \boldsymbol{W}_r \boldsymbol{r}_j^i + \alpha \boldsymbol{T}^{iT} \boldsymbol{W}_\delta \boldsymbol{\delta}_j^i]$;
    Update parameters $\boldsymbol{\theta}_{j+1}^i = \boldsymbol{\theta}_j^i + \Delta\boldsymbol{\theta}_j^i$;
    if *Objective value increased* $F_{j+1}^i > F_j^i$ **then**
      Increase damping term $\lambda \leftarrow \lambda \cdot v$;
      Go back to parameter update computation step (8) ;
    else
      Accept parameter update and reduce damping term $\lambda \leftarrow \lambda / v$;
    Compute effective number of parameters $\gamma = p - 2\alpha \operatorname{tr}([\bar{\boldsymbol{H}}_j^i]^{-1} \boldsymbol{W}_\theta^i)$;
    Reestimate regularization parameter $\alpha = \gamma / (2(E_{\delta,j}^i - E_\delta^{i\star}))$;
    Increase iteration count $j \leftarrow j + 1$;
  Estimate log evidence for current model: $\log P(\tilde{\boldsymbol{z}}|\alpha, \beta, \mathcal{M}_k^i)$ using Eq. (42);
  Increase increment count $i \leftarrow i + 1$;

4.2. Model comparison by Bayes factor

Model comparison, for the purposes of comparing different parametrizations, is most naturally accomplished using a measure of relative evidence. This can be expressed via the Bayes factor [23], which gauges the strength of support for two competing models with equal prior probability. For the purposes of this work, model comparisons will be performed within a set generated from a single parametrization method $\mathcal{M}_k$, say at two increments i and h. Then the support for $\mathcal{M}_k^i$ over $\mathcal{M}_k^h$ is given by the Bayes factor:

$$\log B_k^{ih} = \log P(\tilde{\boldsymbol{z}}|\mathcal{M}_k^i) - \log P(\tilde{\boldsymbol{z}}|\mathcal{M}_k^h) \tag{45}$$

This can be interpreted using the common set of criteria by Kass and Raftery [23], reproduced in Table 2. Since $\log B_k^{ih} = -\log B_k^{hi}$, negative results can simply be interpreted as support for $\mathcal{M}_k^h$ over $\mathcal{M}_k^i$.

Table 2: Interpretation of Bayes factors, adapted from Kass and Raftery [23]

$2 \log B_k^{ih}$	Evidence against $\mathcal{M}_k^h$
0-2	Not worth more than a bare mention
2-6	Positive
6-10	Strong
>10	Very strong

Note that the model evidence estimate from Eq. (42) retains dependence on α, so α can either be marginalized, or as used in this work, model evidence can be estimated at the optimal value (Eq. (43)) [8]. This is equivalent to fixing the prior at the maximum evidence width.

5. Exercise I: Small-scale truss with analytical data

5.1. Model description and data generation

The proposed reparametrization scheme was first tested in model updating of a truss with simulated data. This truss was used in previous work in FE model updating by the authors [7] and was adapted from Papadimitriou *et al.*'s work [25]. The truss is 2-dimensional, statically indeterminate, and symmetric, with 29 elements, 28 DoFs, and pinned boundary conditions. Each element has a mass density of 7800 kg/m^3, Young's modulus of 200 GPa, and cross-sectional area of 0.25 m^2. The FE model for the truss was developed in MATLAB [26].

The first five vibrational modes were selected for analysis and updating ($l = 5$), with natural frequencies and mode shapes (in the unmodified structural state) indicated in Fig. 2. Measured data was assumed to consist of one natural frequency measurement and a full field of $n = 28$ mode shape measurements per mode, for a total of $m = 145$ measurements.

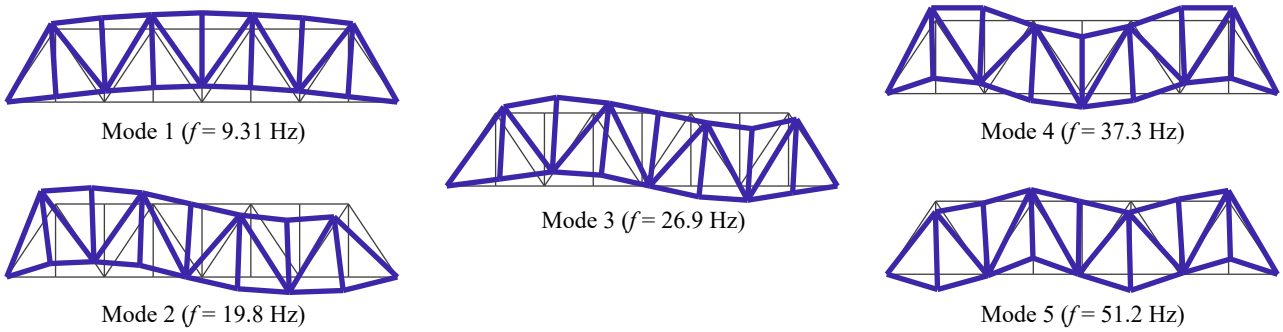


Figure 2: Truss mode shapes and natural frequencies

Several realized data sets were created by first randomizing structural modification and then corrupting the measurement vectors with random noise to produce a more generalized and reproducible ensemble of model updating exercises. Beginning with the described model in its unmodified state, randomized structural modifications were generated by modifying the mass density and Young's modulus of each element by an independent Gaussian random

variable, as in Eqs. (9) and (10). Two levels of structural modification were utilized, where case I covers high uncertainty in FE model parameters, $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, (0.1)^2 \boldsymbol{I}_d)$, and case II covers low uncertainty in FE model parameters, $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, (0.01)^2 \boldsymbol{I}_d)$, with 100 random modified structural states realized for each case.

Random noise was then added to each measurement vector produced from the 200 realized structural states. The measurement noise was generated identically for both cases I and II. Natural frequency measurements were subject to an additive Gaussian noise with zero mean and standard deviation equal to 0.5% of the true natural frequency of the modified structure, $\tilde{f}_j \sim \mathcal{N}(0, (0.005 f_j)^2)$. Mode shape measurements were subject to an additive Gaussian noise vector with zero mean. For each mode shape, the measurement noise vector had standard deviation equal to 5% of the standard deviation of true the mode shape of the modified structure, $\tilde{\boldsymbol{\psi}}_j \sim \mathcal{N}(\mathbf{0}, (0.05 \text{ std}(\boldsymbol{\psi}_j))^2 \boldsymbol{I}_n)$. This noise model was intended to simulate realistic measurement conditions and was previously used by the authors with this structure [7].

The residual weighting matrix $\boldsymbol{W}_r$ was evaluated for each generated FE model updating problem as the inverse of the assumed measurement covariance matrix. Using the notation in Eq. (7), this can be written with $w^f_{rj} = (0.005 \tilde{f}_j)^{-2}$ and $\boldsymbol{W}^{\psi}_{rj} = (0.05 \text{ std}(\tilde{\boldsymbol{\psi}}_j))^{-2} \boldsymbol{I}_n$. $\boldsymbol{W}_r$ would vary slightly depending on the realized data, while β was fixed at a value of $1/2$, as discussed in Section 4.

5.2. Parametrization

Each realized measurement vector was treated as an individual FE model updating problem, and aggregate results were analyzed for both cases. The FE model parameters were chosen to be the $d = n_{\text{el}} = 29$ element Young's moduli from Eq. (10). Element mass densities were not chosen for updating to simplify depiction of parametrizations. Each realized state was parametrized using the incremental reparametrization scheme described in Section 3, where the subset selection parametrization method is indicated as SS or $\mathcal{M}_1$ and parameter clustering is indicated as Cl. or $\mathcal{M}_2$. A forward finite-difference scheme was used to estimate sensitivities at all stages, though analytical results are available in previous work [7].

A sample set of truss parametrizations is depicted in Fig. 3 using $p = 6$ stiffness updating parameters. This includes initial parametrizations ($\mathcal{M}^0_k$) in Figs. 3a and c, and final parametrizations ($\mathcal{M}^f_k$, after the incremental reparametrization schemes converged) in Figs. 3b and d. These depictions aid in understanding how the parametrizations evolve through the incrementing process. For instance, the subset selection parametrizations began with $\mathcal{M}^0_1$ in Fig. 3a and converged to $\mathcal{M}^f_1$ in Fig. 3b. It is notable that two FE model parameters were selected in both the initial and final parametrizations (parameters 1 and 4 in $\mathcal{M}^0_1$ and parameter 6 and 3 in $\mathcal{M}^f_1$, respectively). This demonstrates how the same underlying parameter can change labels, as discussed in the motivation for a reparametrization stopping criterion which can detect parameter label changes (Section 3.3). $\mathcal{M}^0_1$ and $\mathcal{M}^f_1$ are noticeably different, indicating significant changes in parametrization during the incremental process.

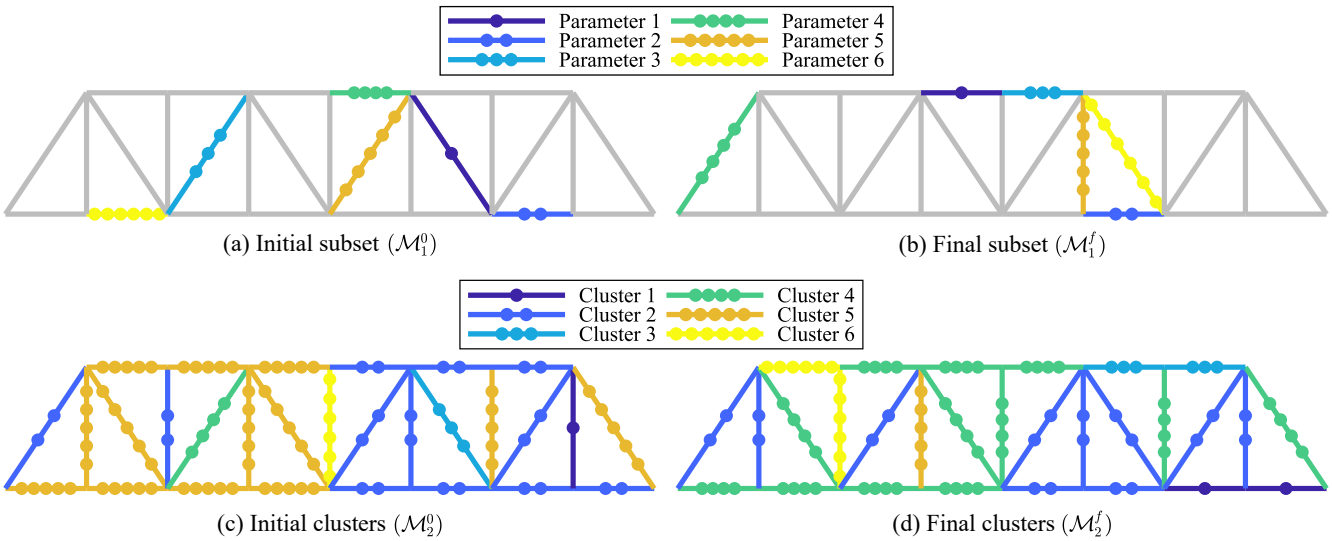


Figure 3: Sample truss parametrizations based on initial and final parameter sensitivities ($p = 6$)

Conversely, the initial parameter clustering parametrization ($\mathcal{M}_2^0$) in Fig. 3c is quite similar to the final clustering parametrization ($\mathcal{M}_2^f$) in Fig. 3d. While the initial clusters 1, 3, 4, and 6 do not have a similar complement in the final clusters, these are single element clusters and comprise only 4 of the 29 FE model parameters. 6 of the 12 parameters in cluster 2 remain the same between initial and final parametrizations, with both sets comprising many elements on the right-central structure and a few web elements on the left side of the structure. Cluster 5 in $\mathcal{M}_2^0$ shares 10 of its 13 parameters with cluster 4 in $\mathcal{M}_2^f$, again indicating the importance of detecting cluster label changes in stopping criteria. With 16 of the 29 FE model parameters contained within similar clusters between initial and final parametrizations, this sample parametrization indicates a recurring observation within this work: reparametrization in sensitivity-based clustering produces less significant changes in parametrization, converging more quickly, while reparametrization in subset selection produces significant variation in parametrization and converges more slowly.

5.3. Model updating results

Model updating was performed on each of the 200 realizations using iterative reparametrization for subset selection and sensitivity-based clustering, with model sizes ranging from $p = 1$ to $p = d$. This used natural frequency and mode shape data in the objective function from Eq. (8), which was regularized and solved using the Levenberg–Marquardt algorithm described in Table 1. The weighting matrix for the FE model parameters, $\boldsymbol{W}_\delta$, was taken as the identity matrix $\boldsymbol{I}_d$ and the regularization parameter α was allowed to vary to maximize model evidence. Incidentally, since both parametrization methods used orthogonal transformation matrices $\boldsymbol{T}^i$, then the updating parameter matrix $\boldsymbol{W}_\theta^i$ from Eq. (33) was equal to $\boldsymbol{I}_d$ for all increments i. The model evidence was estimated using Eq. (42) at the converged solution for each increment i, $\hat{\boldsymbol{\theta}}^i$. FE modeling, modal analysis, parametrization, and data processing were all performed in MATLAB [26].

The average (mean) posterior results, along with parameter efficiency, Bayes factor, and number of increments for convergence in case I (high FE model parameter uncertainty) are shown in Fig. 4. The initial parametrizations ($\mathcal{M}_k^0$) represent previous approaches which only parametrized at the initial structural state and are compared against the results after convergence of the incremental reparametrization schemes ($\mathcal{M}_k^f$). The initial subset selection parametrization, on average, had an optimal model size of $p = 18$, with relatively similar evidence for larger models in Fig. 4a. Reparametrization increased the model evidence across all model sizes for subset selection, moving the optimal model size to approximately $p = 5$. Direct comparison of evidence, via Bayes factor, is given in Fig. 4e, which indicates strong support (see Table 2) for reparametrization in subset selection for all $p < 24$, with increasingly strong support for smaller models. The log likelihood curves in Fig. 4c are similar to the model evidence curve since the penalty term in Fig. 4d only became significant at higher p. Interestingly, the Occam factor was considerably lower in magnitude after reparametrization in subset selection. Since it is unlikely that the total amount of parameter change E_δ decreased with more increments, this likely reflects a more robust (less peaked) posterior.

The parameter efficiency γ/p curves in Fig. 4b are also a measure of the relative influence of the Occam factor and the likelihood. For subset selection, these curves showed a slight preference for the initial subset selection parametrization over reparametrization, though both initial and final parametrizations show good parametrization efficiencies (>70%) at all model sizes. Note that α is selected to provide an optimal model evidence, and therefore tries to avoid the large Occam factor penalties associated with sharply-peaked posteriors, even though this may result in a lower γ/p. Fig. 4f reveals the computational cost (convergence rates) of reparametrization in subset selection. Small parametrizations required, on average, over 15 increments to converge, reducing gradually for larger parametrizations. So while the greatest gains in model evidence were obtained with smaller p, this was offset by requiring several increments to converge.

Parameter clustering showed markedly less improvement from reparametrization, with moderate increases in model evidence and no change in the optimal model size. However, assessing via Bayes factor in Fig. 4e shows conclusive support for reparametrization at most reasonable model sizes ($3 < p < 25$). Parameter clustering showed slightly lower Occam factor after reparametrization. The initial and final parameter efficiency curves showed generally similar behavior, which is explained by the fast convergence and small changes in parametrization, noted in the discussion of Fig. 3. Clustering required between 2 and 4 increments to converge for most model sizes, which was a relatively small price to pay for decisive increases in model evidence.

The average results for case II (low FE model parameter uncertainty) are shown in Fig. 5. Using subset selection, the optimal model size for both initial and final parametrizations occurred between $p = 3$ and $p = 5$, favoring small models. Bayes factor results were more tempered than in case I, exhibiting very strong support for reparametrization with p, decaying to equivocal support for $p \geq 17$. This is related to the much larger effect of the Occam factor than in

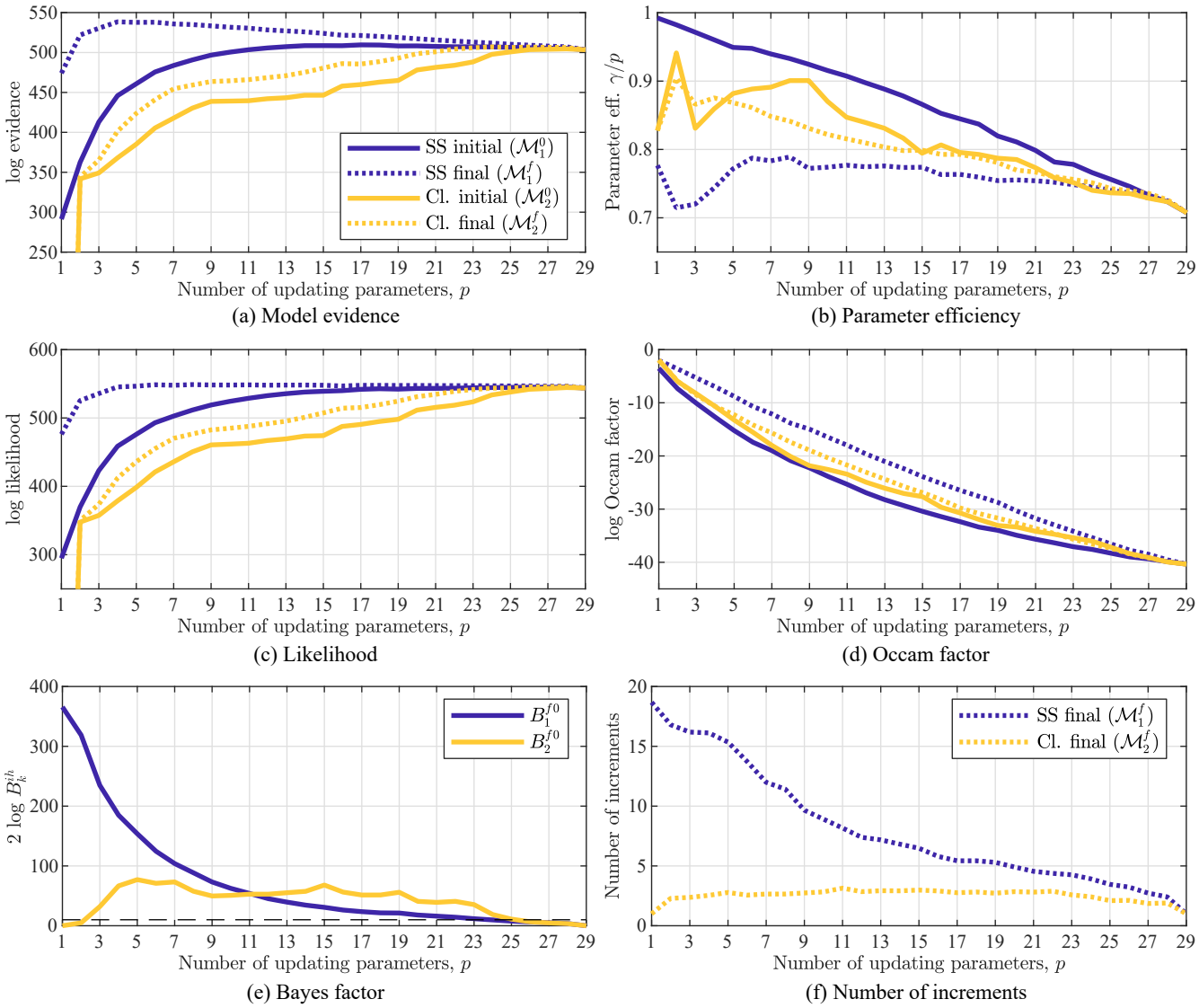


Figure 4: Average truss model updating results, high uncertainty in FE model parameters (case I)

case I, since changes in likelihood were much more limited in case II. Interestingly, the likelihood actually decreased beyond $p = 5$. While this may seem counterintuitive, this is a result of maximizing the regularized objective function (i.e. the posterior) rather than the data fit (i.e. the likelihood). Parametrization efficiency declined much more quickly with increasing p in case II, with $\mathcal{M}_1^f$ providing significantly lower efficiency at all p. However, this lower γ/p was expected due to a greater influence of measurement noise in case II than in case I, giving less information to be extracted from the measurements. This increased the effect of the prior, decreasing γ/p. Subset selection took far fewer increments to converge in case II compared to case I, averaging between 8 increments for a small p to 3 increments for a large p.

Parameter clustering showed similar optimal model size to subset selection in case II, with an optimal p between 2 and 5 for both initial and final parametrizations. The support for reparametrization never reached a significant level in Fig. 5e, giving little basis for reparametrization in parameter clustering for small FE model uncertainties. Reparametrization even failed to provide meaningful increases in data fit (likelihood). The computational cost stayed low, with only 1-2 increments needed for convergence on average. Parameter efficiency curves were large unaffected by reparametrization, peaking around 60% for optimal model sizes and decaying to 20% for large models.

This work was not intended to comment on the relative performance of subset selection and clustering methods, as both approaches have differing merits. However, it may be noted that for this structure, realization method,

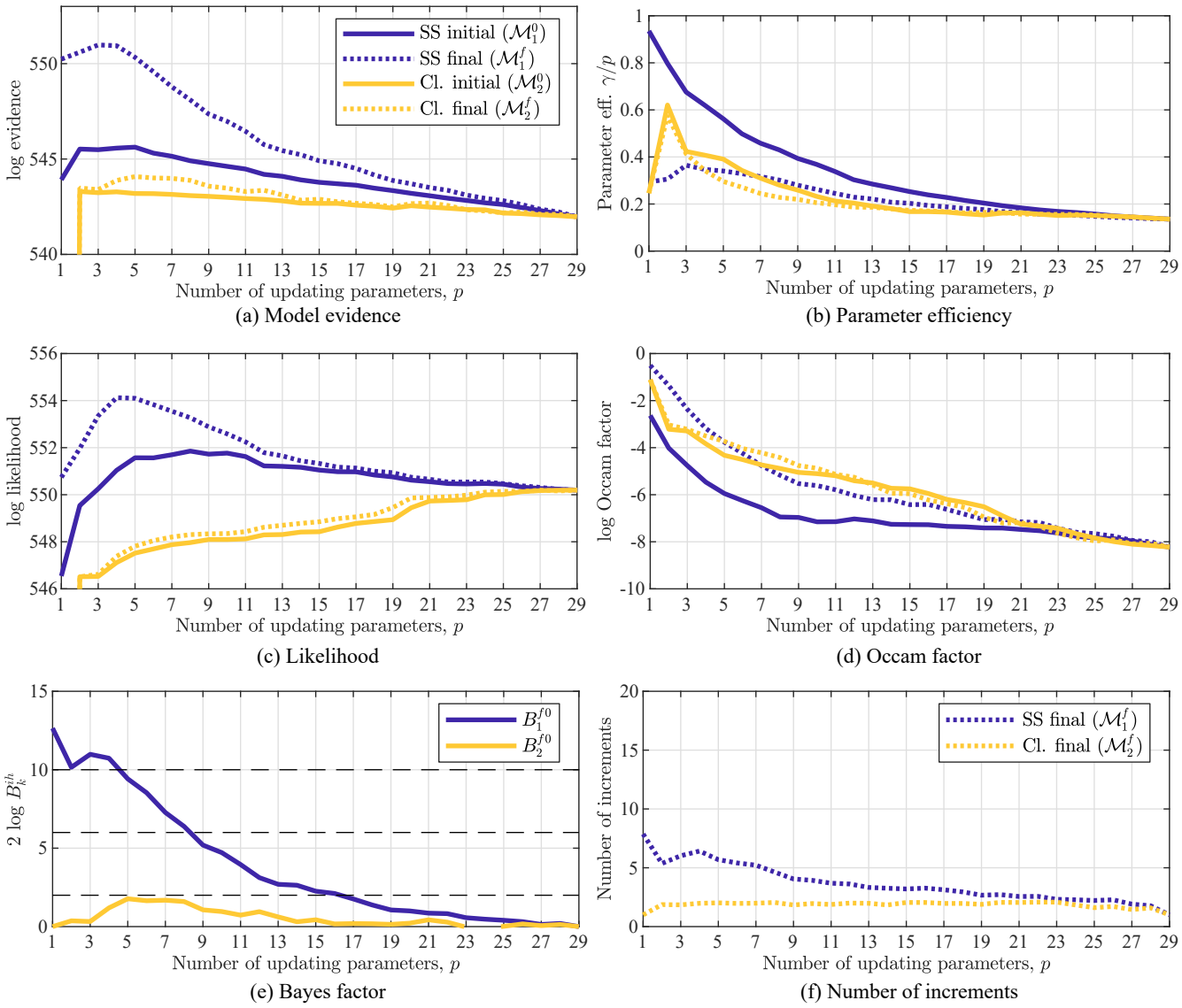


Figure 5: Average truss model updating results, low uncertainty in FE model parameters (case II)

and model updating technique, subset selection outperformed parameter clustering (in terms of evidence) by a considerable margin at all model sizes in both case I and case II. The model evidence, parameter efficiency, and number of increment curves for all parametrizations converged at $p = d$, since there was only one choice of parametrization at this point (i.e. $\boldsymbol{\theta} = \boldsymbol{\delta}$).

6. Exercise II: IASC–ASCE SHM benchmark structure with experimental data

6.1. System identification

The second test of the proposed reparametrization scheme was a more realistic model updating exercise of the benchmark structure created by a task group of the International Association for Structural Control (IASC) and American Society of Civil Engineers (ASCE) for SHM. The structure and results of the Phase II experimental studies [27] are used herein, with data obtained from the Network for Earthquake Engineering Simulation (NEES) database [28]. The structure is a four-story, two-bay by two-bay steel frame structure which is 2.5 m by 2.5 m in plan and 3.6 m height. The benchmark structure was instrumented using three uniaxial accelerometers on each of the four

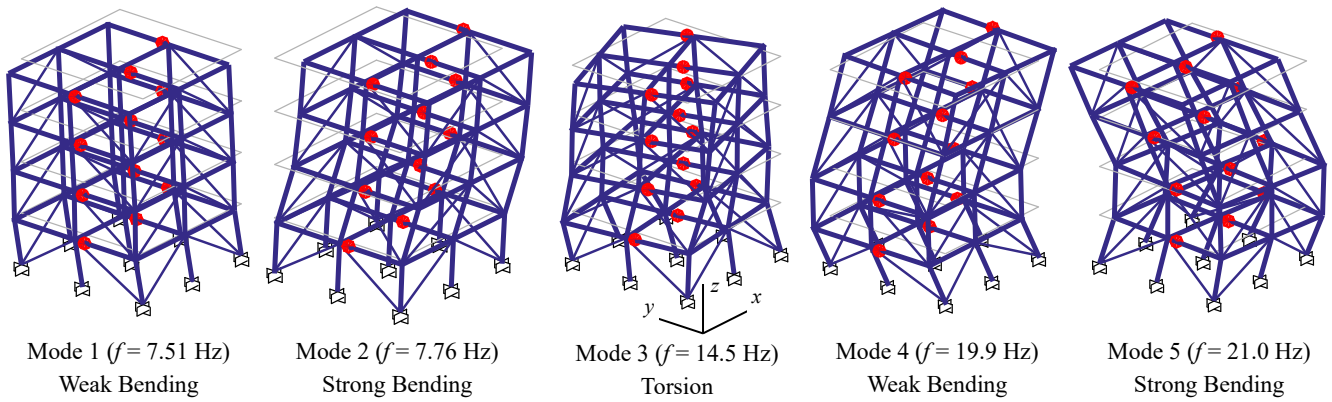


Figure 6: IASC–ASCE SHM benchmark structure measured modes and natural frequencies (measurement locations indicated by red dots)

floors, with the outer two sensors measuring in the y-direction (strong) and the inner sensor measuring in the x-direction (weak) for 12 total measured DoFs. The structure was subject to an ambient excitation for 300 s. Greater detail about the test structure and experimental setup is available online [28] and in [27]. While several structural configurations were available, only the undamaged configuration (case 1, fully braced) is used in this work.

The measured data was cleaned to remove the mean then split into 6 equal-length, non-overlapping segments. Output-only modal identification was performed on each segment separately using the Enhanced Frequency Domain Decomposition (EFDD) algorithm [29]. Similar to Ching and Beck's work with this structure and data [30], five dominant modes were identified from the data, with one natural frequency and 12 mode shape measurements per mode for $m = 65$ total measurements. The 6 separate identification results were used to estimate the mean measurement vector $\tilde{\boldsymbol{z}}$ and the diagonal measurement covariance matrix $\boldsymbol{C}_{\tilde{z}}$. Mode shape results were unit-normalized and set with common orientation prior to averaging. The average natural frequencies and mode shapes are presented in Fig. 6. The depicted modal deformations were created by assuming that the floors acted as rigid diaphragms while the columns and diagonals were made to comply with the floor displacements. This was only done for the purposes of depiction, while any comparison between measured and model-output mode shapes only utilizes the measured locations.

Natural frequency measurements had coefficients of variation of $\{0.66,\ 0.53,\ 0.18,\ 2.0(10)^{-4},\ 2.5(10)^{-4}\}\%$, signifying low variation across the board but particularly precise estimates for modes 4 and 5. Mode shape measurement variation was gauged by the ratio of mode shape measurement standard deviation (averaged for all DoFs in the mode) to the standard deviation of the average mode shape, which gives $\{11,\ 9.7,\ 3.9,\ 1.2,\ 2.5\}\%$. This is approximately an order of magnitude greater variation than for natural frequencies, again with the strongest estimates for modes 4 and 5.

The residual weighting matrix $\boldsymbol{W}_r$ was evaluated as the inverse of the estimated measurement covariance matrix $\boldsymbol{C}_{\tilde{z}}$, which were both diagonal since independence of measurement errors was assumed. The regularization factor β was fixed at a value of $1/2$ to ensure consistent estimates of model evidence, as discussed in Section 4.

6.2. Model description

The FE model for the IASC–ASCE SHM benchmark structure was developed in SAP2000 [31] using technical descriptions available in Phase II literature [27], the NEES database [28], and properties from Phase I literature [32]. The structure was modeled using 9 vertical columns extending the full height, bolted to a foundation, all oriented with the section strong axis along the y-direction. This defined the structure with its global weak axis in the x-direction (about the y-axis) and its global strong axis in the y-direction (about the x-axis). The column bases were assumed to be perfectly restrained against translation and torsion, but strong- and weak-axis base rotations were treated by individual torsion springs. The initial values for the torsion stiffnesses were taken as a multiplier times the nominal column moment stiffnesses (EI/L). The multiplier was set at 10, near the middle of the range for partially-restrained (PR) connections in AISC code [33]. Using the section and material properties in Phase I literature [32], this gave initial torsion stiffnesses for strong-axis column springs as $4.38(10)^6$ N-m/rad and weak-axis column springs as $1.48(10)^6$ N-m/rad. These 18 torsion spring stiffnesses were treated as uncertain stiffness

parameters during updating.

Each of the 4 floors had 12 floor beams which were connected to columns at both ends. These connections were assumed to provide full restraint against relative translation, torsion, and weak-axis rotation, while strong-axis rotation was modeled with PR connections. A PR connection parameter was assigned to each floor beam, giving 48 uncertain floor beam PR connection stiffnesses to be updated. As with the column spring stiffnesses, the initial values of the PR spring stiffnesses were set at 10 times the nominal strong axis moment stiffness, giving $1.95(10)^6$ N-m/rad. Similar to the 120-DoF Phase I analytical model [32], horizontal translation and rotation about the vertical axis was constrained for all nodes on the same floor. Therefore, the floors behaved as rigid diaphragms for in-plane motions, but allowed global and differential vertical motion within each floor. This reduced the number of FE model DoFs from 216 (6 per node) to 120. Two diagonal braces were modeled on each external face of each floor, for a total of 32 braces. The stiffness of each brace was treated as an uncertain parameter, with an initial value set at the nominal stiffness (EA/L), $3.28(10)^7$ N/m.

Floor masses were modeled with four 1000 kg slabs on floors 1-3 and four 750 kg slabs on floor 4 [27, 28]. Each slab was modeled as a uniform distributed line mass added to the two floor beams that supported it. Each additive mass was treated as an uncertain parameter, giving two uncertain parameters per slab, for a total of 32 mass updating parameters. This choice of updating parameters was equivalent to updating the total mass and a single ordinate (along the long axis of the slab) of each slab. Ultimately, there were $d = 130$ uncertain parameters to update between 32 mass parameters and 98 stiffness parameters.

The relative frequency error ($f_{\text{err}} = (\tilde{f} - f)/\tilde{f}$) and MAC values of corresponding measured and initial FE model modes are given in Table 3. The FE model bending modes exhibited 10%-20% greater natural frequencies, with greater error for higher modes. However, the torsional mode (3) exhibited approximately 15% lower natural frequency in the FE model than in the measured structure, suggesting that the model errors were more complex than simply overestimating total stiffness or underestimating total mass. The MAC values for modes 2-4 were very good (>0.900), while modes 1 and 5 exhibited poor and mediocre correspondence, respectively. There was no mode with low natural frequency and mode shape error, suggesting that significant model-structure discrepancies were present.

Due to the relatively low number of mode shape measurements, an intermediate step was implemented during mode pairing. An index was created by pairing initial FE model modes and measured modes using MAC from the 12 measured DoFs. During model updating, FE model modes were first paired with the initial FE model modes using all FE model DoFs to increase pairing fidelity. Then the index between initial FE model-measured modes was used to relate each updated mode to the correct measured mode. This approach ensured consistent pairing between FE model modes and measured modes, since FE model mode shapes could change significantly and had relatively few DoFs for direct pairing.

6.3. Parametrization

Parametrization of the benchmark structure used the incremental reparametrization methodology detailed in Section 3, separately treating the 32 mass and 98 stiffness FE model parameters. As in Section 5, subset selection parametrization method is indicated as SS or $\mathcal{M}_1$ and parameter clustering is indicated as Cl. or $\mathcal{M}_2$. The number of mass and stiffness updating parameters was determined by analyzing dendrograms of weighted initial FE model parameter sensitivities, depicted in Fig. 7. This was equivalent to the weighted cosine distance from Eq. (18). Based on this analysis, a cutoff distance of 10^{-3} (in measure of average cosine distance) was chosen, giving 8 mass clusters and 10 stiffness clusters. This was applied to subset selection as well, giving $p^M = 8$ mass updating parameters and $p^K = 10$ stiffness updating parameters, for a total of $p = 18$ updating parameters.

Fig. 7 also depicts the weighted sensitivity dendrogram of the FE model parameters at the solution for $\mathcal{M}_2^f$. FE model parameter sensitivities became more distinct during updating, indicated by a general increase in the height of corresponding 'branches'. This was more pronounced with the mass sensitivities, where cutting the final dendrogram at 10^{-3} would result in 9 clusters rather than the initial 8. Stiffness sensitivities changed less than mass sensitivities as a whole. Changes in cluster membership are indicated by line crossings in the final dendrograms, with more membership changes noted in the mass parameters than in the stiffness parameters.

Partial depictions of the initial and final parametrizations are presented in Fig. 8. The initial and final subset selection stiffness parametrizations are presented in Figs. 8a and b, respectively. Possible FE model stiffness parameters included the brace axial stiffnesses, floor beam PR spring stiffnesses, and column boundary condition torsion spring stiffnesses. Incidentally, there was no overlap in membership between the initial and final stiffness subsets, in direct contrast to the behavior of the truss in Fig. 3. This may mean that the parameter sensitivities remained

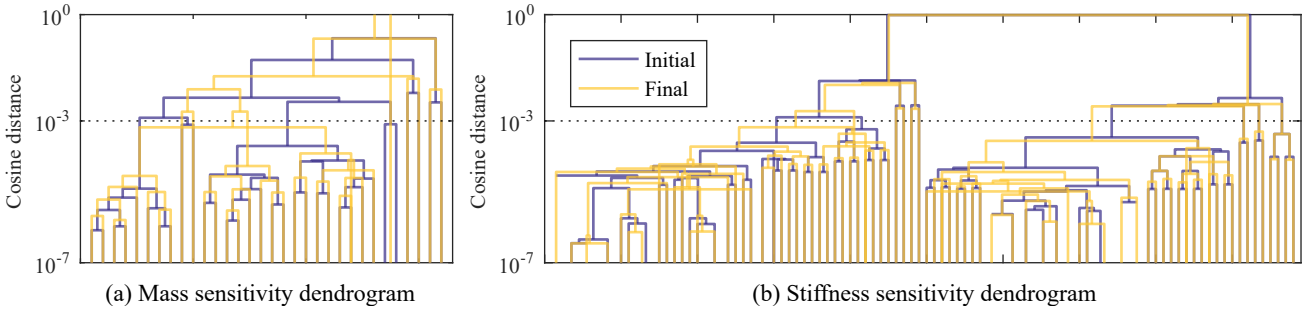


Figure 7: Mass and stiffness parameter sensitivity dendrograms for IASC–ASCE SHM benchmark structure (initial cutoff indicated by dashed line at 10^{-3})

quite distinct during updating for this structure, such that previous use of a parameter greatly reduced its ability to further reduce residual in later increments (i.e. parameter sensitivities varied little during updating).

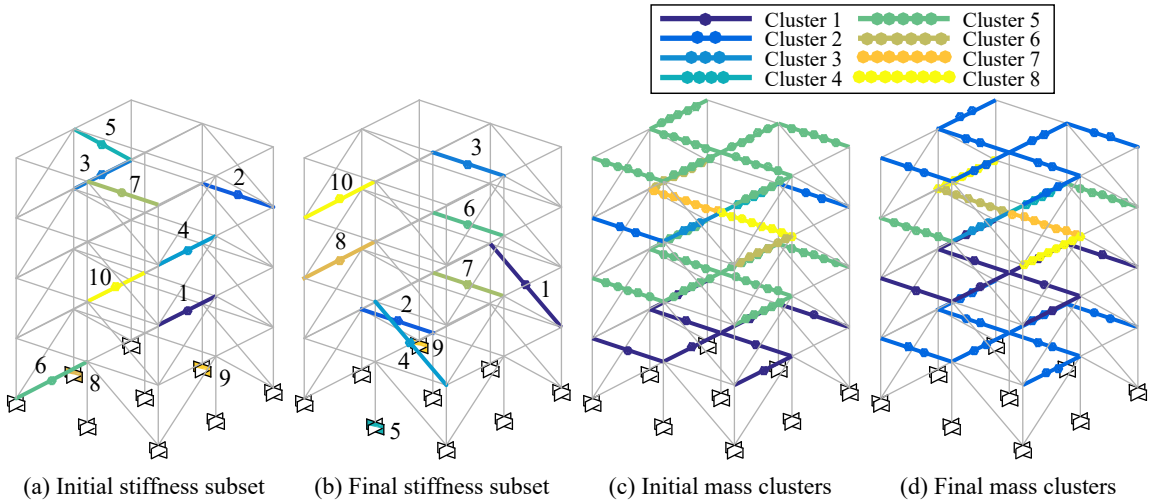


Figure 8: IASC–ASCE SHM benchmark structure sample parametrizations

Figs. 8c and d depict the initial and final parameter clusters based on mass sensitivities, respectively. These clusters were developed from the 32 floor beams having uncertain additive mass from the applied slabs. The initial and final clusters show broad similarity. The initial cluster 6 corresponds with the final cluster 2, both containing the beams in floor 4 and one other floor. In both initial and final parametrizations, only the third floor is broken up between different clusters. This may have to do with the relatively small motion of the third floor in modes 4 and 5 (Fig. 6), such that detailed changes in the masses of floor 3 could affect modes 1-3 without significantly affecting modes 4 and 5. Careful inspection reveals that clusters 2, 6, 7, and 8 in the initial parametrization directly correspond to 5, 8, 6, and 7 in the final parametrization and clusters 3 and 4 are unchanged. In other words, the only change (beyond relabeling) was in the membership of beams on the second floor. Again, this emphasizes the importance of detecting label changes (cluster index permutations) in a successful algorithm stopping criterion.

6.4. Model updating results

Model updating was performed using the defined incremental reparametrization schemes, with parameter estimation and incrementation via the scheme described in Section 4 and Table 1. Natural frequency and mode shape data were incorporated in the objective function from Eq. (8). The covariance matrix for FE model parameters $\boldsymbol{W}_\delta$ was initialized at the identity matrix $\boldsymbol{I}_d$, and the regularization/prior scaling parameter α was allowed to vary to maximize model evidence. As in Section 5, the updating parameter matrix $\boldsymbol{W}_\theta^i$ from Eq. (33) was $\boldsymbol{I}_p$ since the transformation matrices were orthogonal. FE modeling and modal analysis were performed with SAP2000 [31],

while parametrization and data processing were handled in MATLAB [26], with a user-defined API to communicate between the applications.

The model updating results of the initial and final (converged) parametrizations are shown in Table 3, with results for f_{err} and MAC for each of the five identified modes as well as the total quadratic residual E_r and the parameter efficiency γ/p. The initial subset selection parametrization, representing previous approaches of only using one increment, is indicated by $\mathcal{M}_1^0$. This single increment showed a very respectable decrease in E_r, with strong improvements in f_2, f_4, f_5, and MAC_1, where MAC_j is short-hand for $\text{MAC}(\tilde{\boldsymbol{\psi}}_j, \boldsymbol{\psi}_j)$ from Eq. (4). Mild improvements can be noted in all quantities except the natural frequency and MAC of mode 3, which showed slight decreases in correspondence. Between the initial and final subset selection parametrization ($\mathcal{M}_1^f$), E_r was reduced by approximately 23% from its previous value while most modal quantities changed measurably. This required three increments to converge, far fewer than what was needed in the truss example of Section 5. The heavily-weighted f_4 and f_5 saw no change, but slight improvements were made in the MAC_4 and MAC_5, while the MAC_1 reduced significantly and MAC_2 reduced slightly. This trade-off reflects the higher weighting of the mode shape measurements for modes 4 and 5, noted in Section 6.1. Mode 3 saw limited improvements in its correspondence. The parameter efficiency γ/p improved slightly from about 66% in the initial parametrization to 77% in the final parametrization, likely to do with a more robust (wider posterior) choice of final parameters.

Table 3: IASC–ASCE SHM benchmark structure model updating results

	Init. FE model		SS (init.)		SS (final)		Cl. (init. and final)	
			$\mathcal{M}_1^0$		$\mathcal{M}_1^f$		$\mathcal{M}_2^0$ and $\mathcal{M}_2^f$	
Mode	f_{err}	MAC	f_{err}	MAC	f_{err}	MAC	f_{err}	MAC
1 (Weak bending)	-11.7%	0.787	-5.1%	0.992	-7.5%	0.898	-6.0%	0.789
2 (Strong bending)	-11.0%	0.914	-2.1%	0.961	-5.8%	0.944	-6.3%	0.913
3 (Torsion)	15.7%	0.935	21.8%	0.898	17.8%	0.890	18.7%	0.940
4 (Weak bending)	-18.4%	0.925	0.0%	0.941	0.0%	0.951	-0.0%	0.933
5 (Strong bending)	-17.3%	0.832	-0.0%	0.914	-0.0%	0.923	0.0%	0.835
Total error, E_r	$1.32(10)^{10}$		$3.84(10)^4$		$2.96(10)^4$		$6.76(10)^4$	
Parameter eff., γ/p	–		0.66		0.77		0.74	

Where subset selection saw significant change in output quantities between initial and final parametrizations, parameter clustering showed no reportable change as a result of incrementation in Table 3. Both initial ($\mathcal{M}_2^0$) and final ($\mathcal{M}_2^f$) parametrizations showed the same modal output results. Again, the heavily-weighted f_4 and f_5 were excellently matched, but very little improvement was made in the MAC of modes 1, 2, and 5 compared to subset selection. This was a deciding factor in its noticeably lower reduction in E_r compared to subset selection, with more than double E_r compared to $\mathcal{M}_1^f$. Interestingly, while little improvement was made in f_3, parameter clustering was able to measurably increase MAC_3 to 0.940, suggesting that it had greater effect on the problematic torsional mode. As well, parameter clustering exhibited good parameter efficiency at 74%, suggesting that the regularization was effective, but didn't overpower the data during parameter estimation.

All of the parametrizations were able to reduce E_r by 6 orders of magnitude. This was probably related to the near-elimination of natural frequency error from modes 4 and 5 by all parametrizations. Since these measurements had very low variance (Section 6.1), they were very highly weighted by $\boldsymbol{W}_r$. This is the main similarity in results for all parametrizations, along with roughly halving the natural frequency error for modes 1 and 2 while failing to improve the natural frequency error of mode 3. Mode 3 was likely problematic due to the rigid diaphragm assumption made for each of the 4 floors, since the in-plane rigidity of the floors would be most impactful on torsion. Imagery of the structure indicates that diagonal braces were used in the four bays of each floor, but this may have been insufficient to provide the assumed shear stiffness.

The posterior results are given in Table 4, comprising the (log) evidence, likelihood, and Occam factor for initial and final parametrizations. These quantities were estimated at the end of the corresponding increment using asymptotic estimates from Eq. (42). It is important to note β and $\boldsymbol{W}_r$ were fixed for all parametrizations, so log likelihood was only related to the total error E_r. Thus, the additional insight provided by the evidence estimate comes from the Occam factor, which is a measure of model complexity. Interestingly, despite significant changes in parametrization noted in Fig. 8, the Occam factor for $\mathcal{M}_1^0$ and $\mathcal{M}_1^f$ exhibited little change. Thus, the change

in evidence was mostly controlled by E_r, giving the final subset selection parametrization overwhelming support compared to the initial parametrization, as suggested by Bayes factor from Eq. (45) and Table 2. This suggests that for this realistic model updating problem, incremental reparametrization was extremely valuable for subset selection, despite only taking 3 increments to converge with $p = 18$ total updating parameters.

Conversely, reparametrization with parameter clustering saw a slight *decrease* in evidence compared to only performing one parametrization. While the log likelihood was increased by approximately 10 from $\mathcal{M}_2^0$ to $\mathcal{M}_2^f$, the log Occam factor saw a decrease of nearly 20. This decrease in Occam factor represents greater complexity (more total parameter change, more peaked posterior) in the final parametrization. This was enough to provide decisive support for $\mathcal{M}_2^0$ over $\mathcal{M}_2^f$, as assessed by Bayes factor criteria in Table 2. While many good notes can be made about the results of parameter clustering on this structure, incremental reparametrization was not valuable in this scenario. The increased complexity of the reparametrization scheme did not offer any further error reductions nor improvement in parameter efficiency in Table 3, and it did not provide any improvement in model evidence in Table 2.

Table 4: IASC–ASCE SHM benchmark structure posterior results and Bayes factors

	Posterior results (log)			Bayes factor, $2\log B_k^{ih}$	
				Model $\mathcal{M}_k^h$	
Model $\mathcal{M}_k^i$	Evidence	Likelihood	Occam factor	$\mathcal{M}_k^0$	$\mathcal{M}_k^f$
$\mathcal{M}_1^0$ (SS init.)	-1.901(10)^4	-1.895(10)^4	-56.95	–	-8.87(10)^3
$\mathcal{M}_1^f$ (SS final)	-1.457(10)^4	-1.452(10)^4	-55.87	8.87(10)^3	–
$\mathcal{M}_2^0$ (Cl. init.)	-3.361(10)^4	-3.356(10)^4	-50.91	–	20.0
$\mathcal{M}_2^f$ (Cl. final)	-3.362(10)^4	-3.355(10)^4	-69.01	-20.0	–

7. Conclusions

Incremental reparametrization presents a novel, natural approach to extract further FE model error reduction from an existing sensitivity-based parametrization, such as subset selection and parameter clustering. This is accomplished non-intrusively by initializing successive increments at the final state of the previous increment. The primary cost of this proposed approach is the computation of the residual sensitivity matrix (with respect to the FE model parameters) at the start of each increment, but this is offset by limited reprogramming requirements and relatively quick parametrization convergence noted on the included exercises. Importantly, the proposed approach allows further error reduction without increasing the number of updating parameters, which is critical for maintaining posedness of the model updating problem.

While the proposed approach is extensible to any parameter estimation scheme, a deterministic scheme with ties to asymptotic Bayesian inference is proposed in this work. This approach provided robust, fast parameter estimation and approximations for model evidence, which was key in assessing the relative support for reparametrization by Bayes factor. A novel penalty term was implemented which produced consistent regularization despite an incrementally changing parametrization. While regularization is typically used to improve posedness, this was unnecessary in this work since all of the examples had posedness controlled by parametrization. Instead, regularization was used to implement a prior probability for the FE model parameters as part of the model evidence estimation scheme.

Incremental reparametrization was explored in two model updating exercises: a small-scale 2-dimensional truss with analytical data and the IASC–ASCE SHM benchmark structure with experimental data. In both examples, natural frequency and mode shape data were targeted for updating. The truss example allowed for effective exploration of a range of model sizes and parameter uncertainty levels using a significant number of randomized trials. In this example, incremental reparametrization was decisively supported for model updating problems with high parameter uncertainties, in which parameter sensitivities were likely to exhibit a significant change during updating, creating opportunities for more effective parametrizations. Subset selection benefited more from reparametrization, as assessed by Bayes factor, than parameter clustering. This probably relates to the use of both parameter sensitivities and the residual vector in subset selection which greatly influences the choice of parameters. Broadly similar results were noted on the realistic model updating exercise of the IASC–ASCE SHM benchmark structure, where reparametrization was extremely effective for subset selection while parameter clustering did not exhibit an improvement from further parametrization.

Acknowledgments

The authors gratefully acknowledge Columbia University's Graduate School of Arts and Sciences in support of the first author through the Guggenheim and Presidential Fellowships. This work was partially supported by the U.S. National Science Foundation (Grant No. CMMI-1563364).

References

[1] M. I. Friswell, J. E. Mottershead, Finite Element Model Updating in Structural Dynamics, Springer, Dordrecht, 1995. doi:10.1007/978-94-015-8508-8.

[2] J. E. Mottershead, M. Link, M. I. Friswell, The sensitivity method in finite element model updating: a tutorial, Mech. Syst. Signal Process. 25 (7) (2011) 2275–2296. doi:10.1016/j.ymssp.2010.10.012.

[3] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: a review, Mech. Syst. Signal Process. 56–57 (2015) 123–149. doi:10.1016/j.ymssp.2014.11.001.

[4] J. L. Beck, K.-V. Yuen, Model selection using response measurements: Bayesian probabilistic approach, J. Eng. Mech. 130 (2) (2004) 192–203. doi:10.1061/(ASCE)0733-9399(2004)130:2(192).

[5] M. I. Friswell, J. E. Mottershead, H. Ahmadian, Finite–element model updating using experimental test data: parametrization and regularization, Philos. T. Roy. Soc. A 359 (1778) (2001) 169–186. doi:10.1098/rsta.2000.0719.

[6] B. Titurus, M. I. Friswell, Regularization in model updating, Int. J. Numer. Meth. Eng. 75 (4) (2008) 440–478. doi:10.1002/nme.2257.

[7] D. T. Bartilson, J. Jang, A. W. Smyth, Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization, Mech. Syst. Signal Process. 114 (2019) 328–345. doi:10.1016/j.ymssp.2018.05.024.

[8] D. J. C. MacKay, Bayesian interpolation, Neural Comput. 4 (3) (1992) 415–447. doi:10.1162/neco.1992.4.3.415.

[9] R. J. Allemang, D. L. Brown, A correlation coefficient for modal vector analysis, in: 1st International Modal Analysis Conference, 1982.

[10] J. D. Collins, G. C. Hart, T. K. Haselman, B. Kennedy, Statistical identification of structures, AIAA J. 12 (2) (1974) 185–190. doi:10.2514/3.49190.

[11] J. E. Mottershead, C. Mares, M. I. Friswell, S. James, Selection and updating of parameters for an aluminium space-frame model, Mech. Syst. Signal Process. 14 (6) (2000) 923–944. doi:10.1006/mssp.2000.1303.

[12] G. Lallement, J. Piranda, Localization methods for parametric updating of finite element models in elastodynamics, in: 8th International Modal Analysis Conference, 1990.

[13] M. I. Friswell, J. E. Mottershead, H. Ahmadian, Combining subset selection and parameter constraints in model updating, J. Vib. Acoust. 120 (4) (1998) 854–859. doi:10.1115/1.2893911.

[14] H. Shahverdi, C. Mares, W. Wang, J. E. Mottershead, Clustering of parameter sensitivities: examples from a helicopter airframe model updating exercise, Shock Vib. 16 (1) (2009) 75–87. doi:10.3233/SAV-2009-0455.

[15] J. Jang, A. W. Smyth, Model updating of a full-scale FE model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters, Mech. Syst. Signal Process. 83 (2017) 337–355. doi:10.1016/j.ymssp.2016.06.018.

[16] J. Jang, A. W. Smyth, Bayesian model updating of a full-scale finite element model with sensitivity-based clustering, Struct. Control Health Monit. 24 (11) (2017) e2004. doi:10.1002/stc.2004.

[17] A. Miller, Subset Selection in Regression, Chapman & Hall, London, 2002.

[18] L. Rokach, O. Maimon, Clustering Methods, Springer, Boston, MA, 2005, Ch. 15, pp. 321–352.

[19] F. D. Foresee, M. T. Hagan, Gauss–Newton approximation to Bayesian learning, in: IEEE International Conference on Neural Networks, 1997. doi:10.1109/ICNN.1997.614194.

[20] J. L. Beck, L. S. Katafygiotis, Updating models and their uncertainties I: Bayesian statistical framework, J. Eng. Mech. 124 (4) (1998) 455–461. doi:10.1061/(ASCE)0733-9399(1998)124:4(455).

[21] K. Levenberg, A method for the solution of certain non-linear problems in least squares, Q. Appl. Math. 2 (2) (1944) 164–168. doi:10.1090/qam/10666.

[22] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Math. 11 (2) (1963) 431–441. doi:10.1137/0111030.

[23] R. E. Kass, A. E. Raftery, Bayes factors, J. Am. Stat. Assoc. 90 (430) (1995) 773–795. doi:10.1080/01621459.1995.10476572.

[24] G. Schwarz, Estimating the dimension of a model, Ann. Stat. 6 (2) (1978) 461–464. doi:10.1214/aos/1176344136.

[25] C. Papadimitriou, J. L. Beck, S.-K. Au, Entropy-based optimal sensor location for structural model updating, J Vib. Control 6 (5) (2000) 781–800. doi:10.1177/107754630000600508.

[26] MATLAB, Version 9.1.0 (R2016b), The MathWorks Inc., Natick, MA, 2016.

[27] S. J. Dyke, D. Bernal, J. Beck, C. Ventura, Experimental phase II of the structural health monitoring benchmark problem, in: 16th ASCE Engineering Mechanics Conference, 2003.

[28] Network for Earthquake Engineering Simulation (NEES), NEES Database: Structural Control and Monitoring Benchmark Problems, https://datacenterhub.org/dataviewer/view/neesdatabases:db/structural_control_and_monitoring_benchmark_problems/ (accessed 6 September 2018).

[29] R. Brincker, C. E. Ventura, P. Andersen, Damping estimation by frequency domain decomposition, in: 19th International Modal Analysis Conference, 2001, pp. 698–703.

[30] J. Ching, J. L. Beck, Bayesian analysis of the phase II IASC–ASCE structural health monitoring experimental benchmark data, J. Eng. Mech. 130 (10) (2004) 1233–1244. doi:10.1061/(ASCE)0733-9399(2004)130:10(1233).

[31] SAP2000, Version 20.0.0, Computers and Structures, Inc., Berkeley, CA, 2017.

[32] E. A. Johnson, H. F. Lam, L. S. Katafygiotis, J. L. Beck, Phase I IASC–ASCE structural health monitoring benchmark problem using simulated data, J Eng. Mech. 130 (1) (2004) 3–15. doi:10.1061/(ASCE)0733-9399(2004)130:1(3).

[33] AISC, Specification for Structural Steel Buildings (ANSI/AISC 360-16), American Institute of Steel Construction, Chicago, IL, 2016, pp. 275–279.

Paper D

Symmetry properties of natural frequency and mode shape sensitivities in symmetric structures

---

Bibliographic information:

Bartilson, D. T., Jang, J., & Smyth, A. W. (2018). Symmetry properties of natural frequency and mode shape sensitivities in symmetric structures. *Manuscript submitted for publication.*

This page intentionally left blank.

Symmetry properties of natural frequency and mode shape sensitivities in symmetric structures

Daniel T. Bartilson[a,*], Jinwoo Jang[b], Andrew W. Smyth[a]

[a] *Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027, USA*
[b] *Department of Civil, Environmental & Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA*

Abstract

When updating a finite element (FE) model to match the measured properties of its corresponding structure, the sensitivities of FE model outputs to parameter changes are of significant interest. These sensitivities form the core of sensitivity-based model updating algorithms, but they are also used for developing reduced parametrizations, such as in subset selection and clustering. In this work, the sensitivities of natural frequencies and mode shapes are studied for structures having at least one plane of reflectional symmetry. It is first shown that the mode shapes of these structures are either symmetric and anti-symmetric, which is used to prove that natural frequency sensitivities are equal for symmetric parameters. Conversely, mode shape sensitivities are shown to be unequal for symmetric parameters, as measured by cosine distance. These topics are explored with a small numerical example, where it is noted that mode shape sensitivities for symmetric parameters exhibit similar properties to asymmetric parameters.

Keywords: Modal analysis, Sensitivity-based model updating, Eigenvalue problem, Symmetry, Sensitivity analysis

1. Introduction

Finite element (FE) models are of great importance in science and engineering for modeling of physical systems. In structural engineering, an FE model generally corresponds to a unique physical structure but there may exist significant discrepancy between measured and model-predicted behavior. The presence of discrepancy limits the predictive value of an FE model. FE model updating seeks to reduce discrepancy between measured and model-output behavior, often modal properties, by adjusting parameters of an FE model, such as element masses and stiffnesses [1]. An excellent review of model updating, encompassing contemporary methods in uncertainty quantification, is available by Simoen *et al.* [2].

One of the most popular and intuitive methods for FE model updating is the sensitivity method [1] which uses a series of linear approximations to minimize a non-linear sum-of-square error function between measured and model-output data vectors. In the linear approximation step, the sensitivity method directly utilizes the Jacobian matrix, also called the sensitivity matrix, which captures the derivatives of model outputs with respect to parameter changes [1]. The condition of the sensitivity matrix is of paramount importance to the sensitivity method, since parameters are updated using the pseudo-inverse of the sensitivity matrix. Poor conditioning occurs when there are more unknowns (parameters to update) than equations (measurements), but can also result from noisy data or poor selection of updating parameters [1].

Two general approaches have been adopted for improving the condition of the sensitivity matrix in FE model updating: regularization and reduced parametrization. Regularization has been widely studied and utilized in FE model updating [1, 3–5], and involves adding equations to constrain the amount of parameter modification or enforce user-specified equalities. Reduced parametrization, on the other hand, seeks to decrease the number of updating parameters through an intelligent, often automated process. Subset selection [5, 6] produces a reduced set of updating parameters which are (locally) most effective in reducing the modeling errors. Parameter clustering [1, 3, 7–9] has been established as a viable method for grouping model parameters into clusters which are each

*Corresponding author
Email addresses: dtb2121@columbia.edu (Daniel T. Bartilson), jangj@fau.edu (Jinwoo Jang), smyth@civil.columbia.edu (Andrew W. Smyth)
URL: http://faculty.eng.fau.edu/jangj/ (Jinwoo Jang), http://www.columbia.edu/cu/civileng/smyth/ (Andrew W. Smyth)

updated with a single parameter. This relies on machine learning [10] to group model parameters which have similar effects on model-outputs, as measured by sensitivity. In these contexts, it is important to understand the properties of the sensitivity vectors for various model parameters. Parameter clustering has typically been based on natural frequency sensitivities. When this approach is used on symmetric structures, it has produced symmetric clusters [3, 8, 9]. Recently, the authors incorporated both natural frequency and mode shape sensitivities which notably lead to asymmetric clusters [3]. This also has ramifications for subset selection, since parameters are chosen based on their sensitivity vectors and an orthogonalization process will prevent two parameters from being chosen which have the same sensitivities [5].

Beyond sensitivity-based parametrization methods, other applications also benefit from an in-depth understanding of natural frequency and mode shape sensitivities for symmetric structures. Natural frequencies and mode shapes are widely used damage-sensitive features, not only for damage detection, but also for damage localization. Since natural frequencies are not enough to indicate structural damage locations, they are typically used alongside mode shape data to localize structural damage [11–13]. For the same reason, when model updating is used for a structural damage detection application, it usually includes a combination of natural frequencies and mode shapes in an objective function. Since many civil and mechanical structures have at least one plane of symmetry (for design and construction simplicity), it is fundamentally important to understand the effects of parameter changes in symmetric structures. To this end, developing an understanding of natural frequency and mode shape sensitivities will contribute to robust damage detection algorithms and parametrization methods for these special, yet ubiquitous structures.

In this work, natural frequency and mode shape sensitivities are thoroughly explored in the context of structures having at least one plane of symmetry. In Section 2, prior work is reviewed to establish the symmetry and anti-symmetry of mode shapes in symmetric structures. This is then combined with analytical natural frequency and mode shape sensitivity results to show that natural frequency sensitivities are equal (Section 3) and mode shape sensitivities are unequal (Section 4) between symmetric parameters. These techniques are applied to an example symmetric truss in Section 5 with further analysis of mode shape sensitivities for symmetric parameters. Section 6 presents discussion of findings and conclusions.

2. Eigensolution for symmetric structures

The mathematical study of symmetry in structural dynamics began with Glockner [14] who applied group theory to form reduced representations of full structural systems. A symmetric structure can be transformed through a defined set of reflections and rotations, called the symmetry group, to configurations which are identical to the original structure [15]. When a structure exhibits symmetry, its mass and stiffness matrices can be decomposed into similar block-diagonal forms [16]. In vibration analyses of linear structures, this greatly reduces the computational and memory requirements, as the single, large eigendecomposition is transformed into several smaller, separable eigendecompositions [16]. Group theory is the natural vehicle for establishing these transformations, and has seen extensive study in structural dynamics by Kaveh *et al.* [17–20] and Zingoni [15, 21, 22].

In this work, a single type of symmetry is analyzed, namely a single reflection or bilateral symmetry. This corresponds to the C_{1v} symmetry group which comprises two operations: the identity operation e and a single reflection about a vertical plane σ_v. The identity operation is part of any symmetry group, even for asymmetric structures, so it is often disregarded as a trivial transformation [20]. An example of a structure exhibiting C_{1v} symmetry is given in Fig. 1.

Consider a structure which is divided by a reflection plane into left and right substructures, with link elements cut by the plane. At this point, the structure can be general, with different mass, stiffness, and connectivity properties within the left and right substructures. The only requirement which will be put on the structure is that the left and right substructures contain an equal number of degrees-of-freedom (DoFs), at n. The global stiffness matrix $\boldsymbol{K}$ and mass matrix $\boldsymbol{M}$ for the full structure are thus symmetric positive-definite with dimension $N \times N$ where $N = 2n$. The N-element column vector $\boldsymbol{u}$ represents the displacements of the system. The displacement and force vectors can be partitioned into $\boldsymbol{u}^T = \{\boldsymbol{u}_1^T \ \boldsymbol{u}_2^T\}$ and $\boldsymbol{f}^T = \{\boldsymbol{f}_1^T \ \boldsymbol{f}_2^T\}$, respectively. $\boldsymbol{u}_1$ and $\boldsymbol{f}_1$ correspond to displacements and forces on the left substructure, while $\boldsymbol{u}_2$ and $\boldsymbol{f}_2$ correspond to displacements and forces on the right substructure. $\boldsymbol{u}_1$, $\boldsymbol{f}_1$, $\boldsymbol{u}_2$ and $\boldsymbol{f}_2$ are thus equally-sized column vectors with n components.

One way of defining C_{1v} symmetry in a structure is that the displacements and accelerations are symmetric given a symmetric force vector. Note that "symmetry" in this context refers to being unaffected by reflection, not to the typical transpose-symmetry of matrices. If the displacements are symmetric with reflection, then $\boldsymbol{u}^T = \{\boldsymbol{u}_1^T \ \boldsymbol{u}_2^T\} =$

$\{\boldsymbol{u}_2^T \ \boldsymbol{u}_1^T\}$, or $\boldsymbol{u}_1 = \boldsymbol{u}_2$. A more systematic way of representing this operation is through a linear transformation representing reflection, defined as

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \tag{1}$$

which is symmetric and involutory (i.e. $\boldsymbol{T}^T = \boldsymbol{T}^{-1} = \boldsymbol{T}$). Therefore, stiffness-symmetry is defined such that $\boldsymbol{u} = \boldsymbol{T}\boldsymbol{u}$ given $\boldsymbol{f} = \boldsymbol{T}\boldsymbol{f}$. Substituting these transformations into the stiffness equation gives

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{f} \quad \rightarrow \quad \boldsymbol{K}\boldsymbol{T}\boldsymbol{u} = \boldsymbol{T}\boldsymbol{f} \quad \rightarrow \quad \boldsymbol{T}\boldsymbol{K}\boldsymbol{T} = \boldsymbol{K} \tag{2}$$

In other words, the stiffness matrix of a symmetric structure is unaffected by reflection and is indistinguishable when measuring the stiffness of the left or right substructure DoFs. A similar approach can be applied to the mass matrix, where a mass-symmetry is defined such that the acceleration vector is unaffected by reflection: $\ddot{\boldsymbol{u}} = \boldsymbol{T}\ddot{\boldsymbol{u}}$. This is used to satisfy the relation $\boldsymbol{M}\ddot{\boldsymbol{u}} = \boldsymbol{f}$, where $\boldsymbol{f}$ is again symmetric. This leads to $\boldsymbol{T}\boldsymbol{M}\boldsymbol{T} = \boldsymbol{M}$.

The stiffness matrix for this general structure can be written in block-form as

$$\boldsymbol{K} = \begin{bmatrix} \boldsymbol{K}^{(1)} & \boldsymbol{K}^{(2)} \\ \boldsymbol{K}^{(2)T} & \boldsymbol{K}^{(3)} \end{bmatrix} \tag{3}$$

where $\boldsymbol{K}^{(1)}$ and $\boldsymbol{K}^{(3)}$ represent the connectivity within the left and right side DoFs, respectively (intraconnectivity). $\boldsymbol{K}^{(2)}$ and $\boldsymbol{K}^{(2)T}$ represent the connectivity between the left and right side DoFs, respectively (interconnectivity). As with the displacement vector, assume that the block matrices are equal in size with dimension $n \times n$. If the structure is symmetric such that $\boldsymbol{T}\boldsymbol{K}\boldsymbol{T} = \boldsymbol{K}$, then

$$\boldsymbol{T}\boldsymbol{K}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{K}^{(3)} & \boldsymbol{K}^{(2)T} \\ \boldsymbol{K}^{(2)} & \boldsymbol{K}^{(1)} \end{bmatrix} \qquad \rightarrow \qquad \boldsymbol{K}^{(1)} = \boldsymbol{K}^{(3)}; \qquad \boldsymbol{K}^{(2)} = \boldsymbol{K}^{(2)T} \tag{4}$$

A similar result holds for the mass matrix $\boldsymbol{M}$. Intuitively, this means that the stiffness and mass properties are identical between the left and right substructures which are connected by link elements cut by the plane of symmetry. In other words, for a structure to exhibit symmetrical displacement (and acceleration) given a symmetric input force, it must exhibit stiffness (and mass) symmetry. Therefore, the stiffness and mass matrices can be written as

$$\boldsymbol{K} = \begin{bmatrix} \boldsymbol{K}^{(1)} & \boldsymbol{K}^{(2)} \\ \boldsymbol{K}^{(2)} & \boldsymbol{K}^{(1)} \end{bmatrix} \qquad \boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}^{(1)} & \boldsymbol{M}^{(2)} \\ \boldsymbol{M}^{(2)} & \boldsymbol{M}^{(1)} \end{bmatrix} \tag{5}$$

These matrices have a special construction, with two identical blocks along the diagonal and two separate identical blocks along the off-diagonal. This is defined as a Form II matrix, studied by Kaveh and Sayarinejad [17]. Since $\boldsymbol{K}$ and $\boldsymbol{M}$ are symmetric, then so are the block matrices. $\boldsymbol{K}^{(1)}$ corresponds to the stiffness of the substructures on either side of the symmetry plane, while $\boldsymbol{K}^{(2)}$ corresponds to the stiffness of elements linking together the symmetric substructures. Note that this requires a consistent selection of DoFs. In the case of Fig. 1, this required nodal numbering and DoF selection which was isomorphic with reflection.

$\boldsymbol{K}$ and $\boldsymbol{M}$ are used in a generalized eigenvalue problem for finding natural frequencies $\omega = \sqrt{\lambda}$ and mode shapes $\boldsymbol{\phi}$ which satisfy:

$$[\boldsymbol{K} - \lambda \boldsymbol{M}]\boldsymbol{\phi} = \boldsymbol{0} \tag{6}$$

As noted before, $\boldsymbol{K}$ and $\boldsymbol{M}$ can be block-diagonalized via a transformation. Defining an orthogonal matrix $\boldsymbol{P}$ that rotates the space of eigenvectors, then the eigenvalue problem can be rewritten using transformed stiffness ($\bar{\boldsymbol{K}}$) and mass ($\bar{\boldsymbol{M}}$) matrices. Note that the eigenvalues are invariant to this transformation.

$$\boldsymbol{\phi} = \boldsymbol{P}\bar{\boldsymbol{\phi}} \quad \rightarrow \quad [\underbrace{\boldsymbol{P}^T\boldsymbol{K}\boldsymbol{P}}_{\bar{\boldsymbol{K}}} - \lambda \underbrace{\boldsymbol{P}^T\boldsymbol{M}\boldsymbol{P}}_{\bar{\boldsymbol{M}}}]\bar{\boldsymbol{\phi}} = \boldsymbol{0} \tag{7}$$

If the matrix $\boldsymbol{P}$ is defined as follows, then this block-diagonalizes $\bar{\boldsymbol{K}}$ and $\bar{\boldsymbol{M}}$

$$\boldsymbol{P} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{I} \end{bmatrix} \quad \rightarrow \quad \bar{\boldsymbol{K}} = \begin{bmatrix} \bar{\boldsymbol{K}}^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{K}}^{(2)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}^{(1)} + \boldsymbol{K}^{(2)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}^{(1)} - \boldsymbol{K}^{(2)} \end{bmatrix} \tag{8}$$

where $\bar{\boldsymbol{K}}^{(1)}$ and $\bar{\boldsymbol{K}}^{(2)}$ are called the condensed submatrices of $\boldsymbol{K}$ and similar for $\boldsymbol{M}$. Substituting these results into Eq. (7) yields the separable eigenvalue problem:

$$\left[\begin{array}{cc} \bar{\boldsymbol{K}}^{(1)} - \lambda \bar{\boldsymbol{M}}^{(1)} & \mathbf{0} \\ \mathbf{0} & \bar{\boldsymbol{K}}^{(2)} - \lambda \bar{\boldsymbol{M}}^{(2)} \end{array} \right] \bar{\boldsymbol{\phi}} = \mathbf{0} \tag{9}$$

The eigenvalues satisfy the following relation:

$$\det(\bar{\boldsymbol{K}} - \lambda \bar{\boldsymbol{M}}) = \det(\bar{\boldsymbol{K}}^{(1)} - \lambda \bar{\boldsymbol{M}}^{(1)}) \det(\bar{\boldsymbol{K}}^{(2)} - \lambda \bar{\boldsymbol{M}}^{(2)}) = 0 \tag{10}$$

This gives two separable eigendecompositions, where the first set of eigenvalues is denoted as $\{\lambda^{(1)}\}$ with corresponding eigenvectors $\{\bar{\boldsymbol{\phi}}^{(1)}\}$ which satisfy $[\bar{\boldsymbol{K}}^{(1)} - \lambda^{(1)} \bar{\boldsymbol{M}}^{(1)}] \bar{\boldsymbol{\phi}}^{(1)} = \mathbf{0}$. Similarly, the second set of eigenvalues is denoted as $\{\lambda^{(2)}\}$ with corresponding eigenvectors $\{\bar{\boldsymbol{\phi}}^{(2)}\}$ which satisfy $[\bar{\boldsymbol{K}}^{(2)} - \lambda^{(2)} \bar{\boldsymbol{M}}^{(2)}] \bar{\boldsymbol{\phi}}^{(2)} = \mathbf{0}$. The generalized eigenvalues of $\boldsymbol{K}$, $\boldsymbol{M}$, denoted as $\{\lambda(\boldsymbol{K}, \boldsymbol{M})\}$, are then given by the set union of the eigenvalues of the condensed problems:

$$\{\lambda(\boldsymbol{K}, \boldsymbol{M})\} = \{\lambda(\bar{\boldsymbol{K}}, \bar{\boldsymbol{M}})\} = \{\lambda^{(1)}\} \cup \{\lambda^{(2)}\} \tag{11}$$

The eigenvectors of $\boldsymbol{K}$, $\boldsymbol{M}$, denoted as $\{\boldsymbol{\phi}(\boldsymbol{K}, \boldsymbol{M})\}$ can be evaluated from the eigenvectors of the condensed problems after rotation back through $\boldsymbol{P}$. Inserting the i^{th} eigenvalue of the first condensed problem $\lambda_i^{(1)}$ into Eq. (9) yields

$$\bar{\boldsymbol{\phi}}_i = \left\{ \begin{array}{c} \bar{\boldsymbol{\phi}}_i^{(1)} \\ \mathbf{0} \end{array} \right\} \quad \rightarrow \quad \boldsymbol{\phi}_i = \boldsymbol{P} \bar{\boldsymbol{\phi}}_i = \left\{ \begin{array}{c} \bar{\boldsymbol{\phi}}_i^{(1)} \\ \bar{\boldsymbol{\phi}}_i^{(1)} \end{array} \right\} \tag{12}$$

where the scalar was dropped because mode shape scaling is arbitrary. This set of eigenvectors is assembled by duplicating the eigenvectors of the first condensed problem. These eigenvectors are thus symmetric, since $\boldsymbol{\phi}_i = \boldsymbol{T} \boldsymbol{\phi}_i$ for all i from the first condensed problem. Since the eigenvectors are symmetric (i.e. unaffected by reflection) the corresponding modes are said to be symmetric modes.

The eigenvector associated with the j^{th} eigenvalue of the second condensed problem, $\lambda_j^{(2)}$ is similarly evaluated as

$$\bar{\boldsymbol{\phi}}_j = \left\{ \begin{array}{c} \mathbf{0} \\ \bar{\boldsymbol{\phi}}_j^{(2)} \end{array} \right\} \quad \rightarrow \quad \boldsymbol{\phi}_j = \boldsymbol{P} \bar{\boldsymbol{\phi}}_j = \left\{ \begin{array}{c} \bar{\boldsymbol{\phi}}_j^{(2)} \\ -\bar{\boldsymbol{\phi}}_j^{(2)} \end{array} \right\} \tag{13}$$

where the scalar was again omitted. The eigenvectors are assembled by duplication from the eigenvectors of the condensed problem, but now the two blocks are opposite. These eigenvectors are thus anti-symmetric since $\boldsymbol{\phi}_j = -\boldsymbol{T} \boldsymbol{\phi}_j$ for all j from the second condensed problem. Similar to the symmetric mode shapes, the modes corresponding to anti-symmetric eigenvectors are said to be anti-symmetric modes.

The total set of generalized eigenvectors, $\{\boldsymbol{\phi}(\boldsymbol{K}, \boldsymbol{M})\}$, is then given by the union of those two sets of eigenvectors

$$\{\boldsymbol{\phi}(\boldsymbol{K}, \boldsymbol{M})\} = \left\{ \begin{array}{c} \bar{\boldsymbol{\phi}}^{(1)} \\ \bar{\boldsymbol{\phi}}^{(1)} \end{array} \right\} \cup \left\{ \begin{array}{c} \bar{\boldsymbol{\phi}}^{(2)} \\ -\bar{\boldsymbol{\phi}}^{(2)} \end{array} \right\} \tag{14}$$

which takes significant liberty with notation to indicate that it comprises only symmetric and anti-symmetric forms of the condensed problem eigenvectors.

These results are applicable to structures which may exhibit further symmetry, as long as they are decomposable by (at least) one C_{1v} symmetry group which produces Form II $\boldsymbol{K}$ and $\boldsymbol{M}$ matrices, as defined in Eq. (5). Furthermore, this approach can be extended to structures which have DoFs along the line of symmetry. Using the notation of Kaveh and Nikbakht [20], this corresponds to a Form III matrix which is essentially an augmentation of a Form II matrix. The symmetry and anti-symmetry of mode shapes (excluding the DoFs along the line of symmetry) are preserved for Form III matrices. The reflection matrix $\boldsymbol{T}$ becomes more complex in these situations, and further work is warranted to derive the solutions for these structures.

3. Natural frequency sensitivity

Previous applications of group theory to structural dynamics sought to reduce computational expense or memory requirements for the analysis of symmetric structures. In this work, the special properties of mode shapes in

symmetric structures are combined with analytical sensitivity results to explore natural frequency and mode shape sensitivities. To simplify notation in the remaining work, λ refers to $\lambda(\boldsymbol{K}, \boldsymbol{M})$ and $\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{K}, \boldsymbol{M})$. The resulting natural frequency sensitivities can be calculated analytically using the result of Fox and Kapoor [23], which is also derived in Adhikari [24] for arbitrary mode shape normalization and for systems with damping:

$$\lambda_{i,k} = \boldsymbol{\phi}_i^T \boldsymbol{G}_{i,k} \boldsymbol{\phi}_i \tag{15}$$

Note that $\boldsymbol{G}_{i,k} = \boldsymbol{K}_{,k} - \lambda_i \boldsymbol{M}_{,k}$ where $\square_{,k}$ represents the partial derivative with respect to parameter k, $\partial\square/\partial\theta_k$. The term $\boldsymbol{G}_{i,k}$ is evaluated for two parameters, θ_k and θ_l which are symmetric, such that

$$\boldsymbol{K}_{,k} = \begin{bmatrix} \boldsymbol{K}_{,k}^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \qquad \boldsymbol{K}_{,l} = \boldsymbol{T}\boldsymbol{K}_{,k}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_{,k}^{(1)} \end{bmatrix} \tag{16}$$

and similar for $\boldsymbol{M}_{,k}$ and $\boldsymbol{M}_{,l}$. This uses the reflection matrix $\boldsymbol{T}$ defined in Eq. (1). Symmetric parameters may be, for example, the mass densities of two symmetrically-located (mirrored) elements. However, the definition is generic and may include groups of parametrized elements which have a mirror group. The only requirement is that the symmetric parameters satisfy $\boldsymbol{K}_{,l} = \boldsymbol{T}\boldsymbol{K}_{,k}\boldsymbol{T}$. Note that the off-diagonal block matrices of $\boldsymbol{K}_{,k}$ and $\boldsymbol{M}_{,k}$ are necessarily zero since they correspond to the stiffness and mass properties, respectively, of link elements. Since link elements are cut by the reflection plane, they are not part of either substructure and are never going to give symmetric parameters. $\boldsymbol{G}_{i,k}$ and $\boldsymbol{G}_{i,l}$ are then formed

$$\boldsymbol{G}_{i,k} = \begin{bmatrix} \boldsymbol{K}_{,k}^{(1)} - \lambda_i \boldsymbol{M}_{,k}^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \qquad \boldsymbol{G}_{i,l} = \boldsymbol{T}\boldsymbol{G}_{i,k}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_{,k}^{(1)} - \lambda_i \boldsymbol{M}_{,k}^{(1)} \end{bmatrix} \tag{17}$$

3.1. Symmetric mode shape

As established in Section 2, the mode shapes of symmetric structures are either symmetric or anti-symmetric. Analysis begins with an arbitrary symmetric mode i, such that

$$\boldsymbol{\phi}_i = \begin{Bmatrix} \boldsymbol{v}_i \\ \boldsymbol{v}_i \end{Bmatrix} = \boldsymbol{T}\boldsymbol{\phi}_i \tag{18}$$

for some non-zero vector $\boldsymbol{v}_i$. $\boldsymbol{\phi}_i$ is unaffected by the reflection matrix $\boldsymbol{T}$ in Eq. (1). Writing the eigenvalue sensitivity with respect to parameter θ_l and using the relation in Eq. (17):

$$\lambda_{i,l} = \boldsymbol{\phi}_i^T \boldsymbol{G}_{i,l} \boldsymbol{\phi}_i = \boldsymbol{\phi}_i^T [\boldsymbol{T}\boldsymbol{G}_{i,k}\boldsymbol{T}] \boldsymbol{\phi}_i = [\boldsymbol{\phi}_i^T \boldsymbol{T}] \boldsymbol{G}_{i,k} [\boldsymbol{T}\boldsymbol{\phi}_i] = \boldsymbol{\phi}_i^T \boldsymbol{G}_{i,k} \boldsymbol{\phi}_i = \lambda_{i,k} \tag{19}$$

then equivalence is satisfied, and eigenvalue sensitivities are equal between symmetric parameters for symmetric modes.

3.2. Anti-symmetric mode shape

This examination is repeated for an arbitrary anti-symmetric mode j,

$$\boldsymbol{\phi}_j = \begin{Bmatrix} \boldsymbol{v}_j \\ -\boldsymbol{v}_j \end{Bmatrix} = -\boldsymbol{T}\boldsymbol{\phi}_j \tag{20}$$

which is negated when transformed by the reflection matrix $\boldsymbol{T}$ in Eq. (1). The eigenvalue sensitivity with respect to θ_l can again be equated to the sensitivity with respect to θ_k:

$$\lambda_{j,l} = \boldsymbol{\phi}_j^T \boldsymbol{G}_{j,l} \boldsymbol{\phi}_j = \boldsymbol{\phi}_j^T [\boldsymbol{T}\boldsymbol{G}_{j,k}\boldsymbol{T}] \boldsymbol{\phi}_j = [-\boldsymbol{\phi}_j^T \boldsymbol{T}] \boldsymbol{G}_{j,k} [-\boldsymbol{T}\boldsymbol{\phi}_j] = \boldsymbol{\phi}_j^T \boldsymbol{G}_{j,k} \boldsymbol{\phi}_j = \lambda_{j,k} \tag{21}$$

and thus eigenvalue sensitivities are equal between symmetric parameters also for anti-symmetric modes.

Therefore eigenvalue (and correspondingly, natural frequency) sensitivities are equal between symmetric parameters for all modes, comprising symmetric and anti-symmetric modes, in symmetric structures. This is a significant result and gives an analytical explanation for the symmetry of clusters based on natural frequency sensitivity, as noted in [3, 8, 9]. Since natural frequency sensitivity vectors will always be equal between symmetric parameters, they will be clustered together using any distance metric or clustering technique. Furthermore, this means that

subset selection based on natural frequency sensitivity will necessarily select against a symmetric parameter after its complement has been chosen. This occurs because the selection process orthogonalizes after each selection, and the identical sensitivity of the parameter pair will be removed.

Finally, this has large ramifications for the sensitivity matrix in sensitivity-based model updating. If an updating problem is set to update only natural frequencies, then two columns corresponding to symmetric parameters will be linearly dependent and the rank of the sensitivity matrix will be less than the number of columns. Therefore, special care must be taken when updating the natural frequencies of symmetric structures to correctly determine the rank of the sensitivity matrix prior to updating and take the correct ameliorating steps of regularization or reparametrization.

4. Mode shape sensitivity

This approach is now extended to mode shape sensitivities. The sensitivity, or derivative, of mode shape $\boldsymbol{\phi}_i$ with respect to parameter θ_k is given by [23]:

$$\boldsymbol{\phi}_{i,k} = -\underbrace{[\boldsymbol{F}_i\boldsymbol{F}_i + 2\boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T\boldsymbol{M}]}_{\boldsymbol{L}_i}^{-1} \underbrace{[\boldsymbol{F}_i\boldsymbol{F}_{i,k} + \boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T\boldsymbol{M}_{,k}]}_{\boldsymbol{R}_{i,k}}\boldsymbol{\phi}_i \tag{22}$$

This can be split into the inverse of a symmetric matrix $\boldsymbol{L}_i$ which does not depend on the parameter θ_k, a matrix $\boldsymbol{R}_{i,k}$, and the mode shape $\boldsymbol{\phi}_i$. $\boldsymbol{L}_i$ is Form II, meaning that $\boldsymbol{T}\boldsymbol{L}_i\boldsymbol{T} = \boldsymbol{L}_i$. Thus $\boldsymbol{L}_i^{-1}$ is also Form II ($\boldsymbol{L}_i^{-1} = [\boldsymbol{T}\boldsymbol{L}_i\boldsymbol{T}]^{-1} = \boldsymbol{T}\boldsymbol{L}_i^{-1}\boldsymbol{T}$), giving properties that will be exploited later. $\boldsymbol{F}_i$ represents the eigenvalue problem with eigenvalue λ_i, which has the same form as the $\boldsymbol{K}$ and $\boldsymbol{M}$ matrices:

$$\boldsymbol{F}_i = \boldsymbol{K} - \lambda_i\boldsymbol{M} = \begin{bmatrix} \boldsymbol{F}_i^{(1)} & \boldsymbol{F}_i^{(2)} \\ \boldsymbol{F}_i^{(2)} & \boldsymbol{F}_i^{(1)} \end{bmatrix} \tag{23}$$

The derivative of $\boldsymbol{F}_i$ with respect to parameter θ_k is given as $\boldsymbol{F}_{i,k} = \boldsymbol{K}_{,k} - \lambda_i\boldsymbol{M}_{,k} - \lambda_{i,k}\boldsymbol{M}$, which is closely related to $\boldsymbol{G}_{i,k}$ from Eq. (17):

$$\boldsymbol{F}_{i,k} = \begin{bmatrix} \boldsymbol{F}_{i,k}^{(1)} & \boldsymbol{F}_{i,k}^{(2)} \\ \boldsymbol{F}_{i,k}^{(2)} & \boldsymbol{F}_{i,k}^{(3)} \end{bmatrix} = \boldsymbol{G}_{i,k} - \lambda_{i,k}\boldsymbol{M} \tag{24}$$

The derivative with respect to a symmetric parameter $\boldsymbol{\theta}_l$ is

$$\boldsymbol{F}_{i,l} = \begin{bmatrix} \boldsymbol{F}_{i,l}^{(1)} & \boldsymbol{F}_{i,l}^{(2)} \\ \boldsymbol{F}_{i,l}^{(2)} & \boldsymbol{F}_{i,l}^{(3)} \end{bmatrix} = \boldsymbol{G}_{i,l} - \lambda_{i,k}\boldsymbol{M} \tag{25}$$

where the result $\lambda_{i,l} = \lambda_{i,k}$ was used from Section 3. This means that $\boldsymbol{F}_{i,l}$ is the reflection of $\boldsymbol{F}_{i,k}$, using the reflection matrix in Eq. (1):

$$\boldsymbol{F}_{i,l} = \boldsymbol{T}\boldsymbol{F}_{i,k}\boldsymbol{T} \tag{26}$$

Using these definitions, the goal is to describe the similarity between mode shape sensitivities for two symmetric parameters. Generally, subset selection methods and parameter clustering methods phrase similarity in terms of the cosine distance between vectors [3, 5, 7–9]. Between two mode shape sensitivity vectors, this can be written:

$$d_{\cos}(\boldsymbol{\phi}_{i,k}, \boldsymbol{\phi}_{i,l}) = 1 - \frac{\boldsymbol{\phi}_{i,k}^T\boldsymbol{\phi}_{i,l}}{\sqrt{\boldsymbol{\phi}_{i,k}^T\boldsymbol{\phi}_{i,k} \cdot \boldsymbol{\phi}_{i,l}^T\boldsymbol{\phi}_{i,l}}} \tag{27}$$

This is equal to $1 - \cos(\psi)$ where ψ is the angle between the two vectors. $d_{\cos}$ ranges between 0 (vectors are parallel) and 2 (vectors are anti-parallel). When $\boldsymbol{\phi}_{i,k}$ and $\boldsymbol{\phi}_{i,l}$ have the same magnitude (i.e. $\boldsymbol{\phi}_{i,k}^T\boldsymbol{\phi}_{i,k} = \boldsymbol{\phi}_{i,l}^T\boldsymbol{\phi}_{i,l}$), then this can be written as

$$d_{\cos}(\boldsymbol{\phi}_{i,k}, \boldsymbol{\phi}_{i,l}) = \{\boldsymbol{\phi}_{i,k} - \boldsymbol{\phi}_{i,l}\}^T\{\boldsymbol{\phi}_{i,k} - \boldsymbol{\phi}_{i,l}\}/(2\boldsymbol{\phi}_{i,k}^T\boldsymbol{\phi}_{i,k}) \tag{28}$$

In order for two symmetric parameters, θ_k and θ_l, to have $d_{\cos} = 0$ when using Eq. (28) then the following condition must hold:

$$\boldsymbol{\phi}_{i,k} - \boldsymbol{\phi}_{i,l} = -\boldsymbol{L}_i^{-1}[\boldsymbol{R}_{i,k}\boldsymbol{\phi}_i - \boldsymbol{R}_{i,l}\boldsymbol{\phi}_i] = \boldsymbol{0} \quad \rightarrow \quad \boldsymbol{R}_{i,k}\boldsymbol{\phi}_i - \boldsymbol{R}_{i,l}\boldsymbol{\phi}_i = \boldsymbol{0} \tag{29}$$

where $\boldsymbol{L}_i^{-1}$ can be reduced because it is invertible.

4.1. Symmetric mode shape

Eq. (29) is analyzed in the context of a symmetric mode shape, $\boldsymbol{\phi}_i$, with form given by Eq. (18). However, we will first show that the symmetric parameter sensitivities are equal in magnitude to allow use of Eq. (28). To do this, Eq. (22) is examined for mode i and parameter l:

$$\begin{aligned}\boldsymbol{R}_{i,l}\boldsymbol{\phi}_i &= [\boldsymbol{F}_i\boldsymbol{F}_{i,l} + \boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T\boldsymbol{M}_{,l}]\boldsymbol{\phi}_i = [\boldsymbol{F}_i[\boldsymbol{T}\boldsymbol{F}_{i,k}\boldsymbol{T}] + \boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T[\boldsymbol{T}\boldsymbol{M}_{,k}\boldsymbol{T}]]\boldsymbol{\phi}_i \\ &= [\boldsymbol{F}_i\boldsymbol{T}\boldsymbol{F}_{i,k} + \boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T\boldsymbol{T}\boldsymbol{M}_{,k}]\boldsymbol{T}\boldsymbol{\phi}_i = [\boldsymbol{F}_i\boldsymbol{T}\boldsymbol{F}_{i,k} + \boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T\boldsymbol{M}_{,k}]\boldsymbol{\phi}_i\end{aligned} \tag{30}$$

This uses the reflection matrix $\boldsymbol{T}$ in Eq. (1) to describe the transformations of $\boldsymbol{F}_{i,l}$ in Eq. (26) and $\boldsymbol{M}_{,l}$, similar to Eq. (16). This is further simplified using the fact that $\boldsymbol{T}\boldsymbol{\phi}_i = \boldsymbol{\phi}_i$ for a symmetric mode shape. Given that $\boldsymbol{L}_i^{-1}$ is a Form II matrix, then $\boldsymbol{T}\boldsymbol{L}_i^{-1} = \boldsymbol{L}_i^{-1}\boldsymbol{T}$. Thus, we can show that $\boldsymbol{T}\boldsymbol{\phi}_{i,k} = \boldsymbol{\phi}_{i,l}$ using a part of Eq. (22) and Eq. (30):

$$\boldsymbol{T}\boldsymbol{\phi}_{i,k} = -\boldsymbol{L}_i^{-1}\boldsymbol{T}\boldsymbol{R}_{i,k}\boldsymbol{\phi}_i = -\boldsymbol{L}_i^{-1}[\boldsymbol{F}_i\boldsymbol{T}\boldsymbol{F}_{i,k} + \boldsymbol{M}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T\boldsymbol{M}_{,k}]\boldsymbol{\phi}_i = \boldsymbol{\phi}_{i,l} \tag{31}$$

This utilizes the facts that $\boldsymbol{F}_i$ and $\boldsymbol{M}$ are Form II to give $\boldsymbol{T}\boldsymbol{F}_i = \boldsymbol{F}_i\boldsymbol{T}$ and $\boldsymbol{T}\boldsymbol{M} = \boldsymbol{M}\boldsymbol{T}$. Therefore the two sensitivity vectors have equal magnitude, $\boldsymbol{\phi}_{i,l}^T\boldsymbol{\phi}_{i,l} = \boldsymbol{\phi}_{i,k}^T\boldsymbol{T}\boldsymbol{T}\boldsymbol{\phi}_{i,k} = \boldsymbol{\phi}_{i,k}^T\boldsymbol{\phi}_{i,k}$, so Eq. (28) can be used.

Comparing Eq. (30) to $\boldsymbol{R}_{i,k}\boldsymbol{\phi}_i$ in Eq. (22) gives the following relation to satisfy Eq. (29):

$$[\boldsymbol{F}_i - \boldsymbol{F}_i\boldsymbol{T}]\boldsymbol{F}_{i,k}\boldsymbol{\phi}_i = \boldsymbol{0} \tag{32}$$

Note that $\boldsymbol{F}_i$ can not be reduced from this equation since it is not invertible. For a particular mode shape $\boldsymbol{\phi}_i$ and parameter θ_k, this equation may be satisfied by several $\boldsymbol{F}_i$ because $\boldsymbol{F}_{i,k}$ isn't necessarily full-rank. However, this equation must hold for any arbitrary parameter θ_k. Therefore, $\boldsymbol{F}_{i,k}$ can be treated as an arbitrary matrix because the $\boldsymbol{K}_{,k}$ and $\boldsymbol{M}_{,k}$ are arbitrary based on parametrization. Since $\boldsymbol{\phi}_i$ is necessarily non-zero by definition of an eigenvector, the only general solution is $\boldsymbol{F}_i = \boldsymbol{F}_i\boldsymbol{T}$. This requirement can be written as $\boldsymbol{F}_i^{(1)} = \boldsymbol{F}_i^{(2)}$ from Eq. (23):

$$\boldsymbol{K}^{(1)} - \lambda_i\boldsymbol{M}^{(1)} = \boldsymbol{K}^{(2)} - \lambda_i\boldsymbol{M}^{(2)} \tag{33}$$

This may have infinitely many solutions for a particular eigenvalue λ_i, but it is desired to find a condition such that mode shape sensitivities are equal for symmetric parameters *for all symmetric modes*. Therefore, λ_i is treated as arbitrary, giving the requirements that

$$\boldsymbol{K}^{(1)} = \boldsymbol{K}^{(2)} \qquad \boldsymbol{M}^{(1)} = \boldsymbol{M}^{(2)} \tag{34}$$

While generally difficult to attain, this condition doesn't violate any properties of the stiffness and mass matrices (i.e. symmetry, non-zero).

4.2. Anti-symmetric mode shape

Eq. (29) is now analyzed in the context of an anti-symmetric mode shape, $\boldsymbol{\phi}_j$, with form given by Eq. (20). Again, it is first shown that the symmetric parameters have equal magnitude sensitivity vectors by examining part of Eq. (22) This is simplified using $\boldsymbol{T}\boldsymbol{\phi}_j = -\boldsymbol{\phi}_j$ for an anti-symmetric mode shape:

$$\begin{aligned}\boldsymbol{R}_{j,l}\boldsymbol{\phi}_j &= [\boldsymbol{F}_j\boldsymbol{F}_{j,l} + \boldsymbol{M}\boldsymbol{\phi}_j\boldsymbol{\phi}_j^T\boldsymbol{M}_{,l}]\boldsymbol{\phi}_j = [\boldsymbol{F}_j[\boldsymbol{T}\boldsymbol{F}_{j,k}\boldsymbol{T}] + \boldsymbol{M}\boldsymbol{\phi}_j\boldsymbol{\phi}_j^T[\boldsymbol{T}\boldsymbol{M}_{,k}\boldsymbol{T}]]\boldsymbol{\phi}_j \\ &= [\boldsymbol{F}_j\boldsymbol{T}\boldsymbol{F}_{j,k} + \boldsymbol{M}\boldsymbol{\phi}_j\boldsymbol{\phi}_j^T\boldsymbol{T}\boldsymbol{M}_{,k}]\boldsymbol{T}\boldsymbol{\phi}_j = -[\boldsymbol{F}_j\boldsymbol{T}\boldsymbol{F}_{j,k} - \boldsymbol{M}\boldsymbol{\phi}_j\boldsymbol{\phi}_j^T\boldsymbol{M}_{,k}]\boldsymbol{\phi}_j\end{aligned} \tag{35}$$

This can be used to show that $\boldsymbol{\phi}_{j,l} = -\boldsymbol{T}\boldsymbol{\phi}_{j,k}$ in a similar process to Eq. (31). Thus, two sensitivity vectors have equal magnitude, $\boldsymbol{\phi}_{j,l}^T\boldsymbol{\phi}_{j,l} = \boldsymbol{\phi}_{j,k}^T\boldsymbol{\phi}_{j,k}$, so Eq. (28) can be used. Comparing Eq. (35) to Eq. (29) for mode j gives the following condition such that Eq. (29) is satisfied:

$$[\boldsymbol{F}_j + \boldsymbol{F}_j\boldsymbol{T}]\boldsymbol{F}_{j,k}\boldsymbol{\phi}_j = \boldsymbol{0} \tag{36}$$

As discussed previously, this equation has to hold for an arbitrary parameter θ_k and non-zero $\boldsymbol{\phi}_j$, making $\boldsymbol{F}_{j,k}\boldsymbol{\phi}_j$ an arbitrary vector. The only solution, then, is $\boldsymbol{F}_j = -\boldsymbol{F}_j\boldsymbol{T}$, or equivalently, $\boldsymbol{F}_j^{(1)} = -\boldsymbol{F}_j^{(2)}$. This can be written as

$$\boldsymbol{K}^{(1)} - \lambda_j\boldsymbol{M}^{(1)} = -[\boldsymbol{K}^{(2)} - \lambda_j\boldsymbol{M}^{(2)}] \tag{37}$$

Since this must hold for various values of λ_j, corresponding to different anti-symmetric modes, the only solution which satisfies $\boldsymbol{\phi}_{j,k} = \boldsymbol{\phi}_{j,l}$ for arbitrary (symmetric) parameters θ_k and θ_l and arbitrary anti-symmetric mode j is

$$\boldsymbol{K}^{(1)} = -\boldsymbol{K}^{(2)} \qquad \boldsymbol{M}^{(1)} = -\boldsymbol{M}^{(2)} \tag{38}$$

Comparing the equivalence criterion for symmetric mode shapes in Eq. (34) and anti-symmetric Eq. (38) yields the only solution to both sets: $\boldsymbol{K}^{(1)} = \boldsymbol{K}^{(2)} = \boldsymbol{0}$ and $\boldsymbol{M}^{(1)} = \boldsymbol{M}^{(2)} = \boldsymbol{0}$, which is a null result. Therefore, we have analytically shown that there is no symmetric structure for which all mode shape sensitivities will be equal between symmetric parameters, in direct contrast to natural frequency sensitivities, which will be equal. In other words, the presence of structural symmetry is never going to let symmetric parameters have the same mode shape sensitivities. Intuitively, one would expect that the more closely mode shape sensitivity equivalence is satisfied for symmetric modes, the greater discrepancy would exist for anti-symmetric modes, and vice versa. This comes from observation of Eqs. (32) and (36), but $\boldsymbol{L}_i^{-1}$, $\boldsymbol{F}_{i,k}$, and $\boldsymbol{\phi}_i$ will vary for each mode, so it is not guaranteed that this condition will exist.

However, the observation that $d_{\cos}(\boldsymbol{\phi}_{i,k}, \boldsymbol{\phi}_{i,l}) > 0$ is limited in utility. While the cosine distance between mode shape sensitivities for symmetric parameters is necessarily greater than zero, the cosine distance may still be close enough to zero such that symmetric parameters are still clustered together (in cluster analysis) or nearly orthogonalized (in subset selection) or essentially linearly dependent (in the sensitivity matrix). Since the equations do not readily admit a bound on $d_{\cos}$, this will be explored in a small numerical study.

5. Example structure with C_{1v} symmetry

The structure of study, in Fig. 1, was modified from Kaveh and Nikbakht [20]. This is a symmetric 10 element truss formed from two equilateral triangles (elements $\{1, 2, 3\}$ and elements $\{8, 9, 10\}$) with link elements $\{4, 5, 6, 7\}$. The element properties are uniform, with dimensionless Young's modulus of 1, area 1, and mass density 1. The structure is pinned at the symmetric nodes 1 and 4, giving 8 free DoFs. Note that DoFs 5 and 7 are reversed relative to the direction of 1 and 3 such that DoFs 5-8 are isomorphic to DoFs 1-4 after reflection, giving the structure C_{1v} symmetry. The displacements of the structure are thus partitioned as

$$\boldsymbol{u}^T = \left\{\boldsymbol{u}_1^T \mid \boldsymbol{u}_2^T\right\} = \{u_1\ u_2\ u_3\ u_4 \mid u_5\ u_6\ u_7\ u_8\} \tag{39}$$

corresponding to left and right substructures. The stiffness $\boldsymbol{K}$ and mass $\boldsymbol{M}$ matrices are Form II, as described in

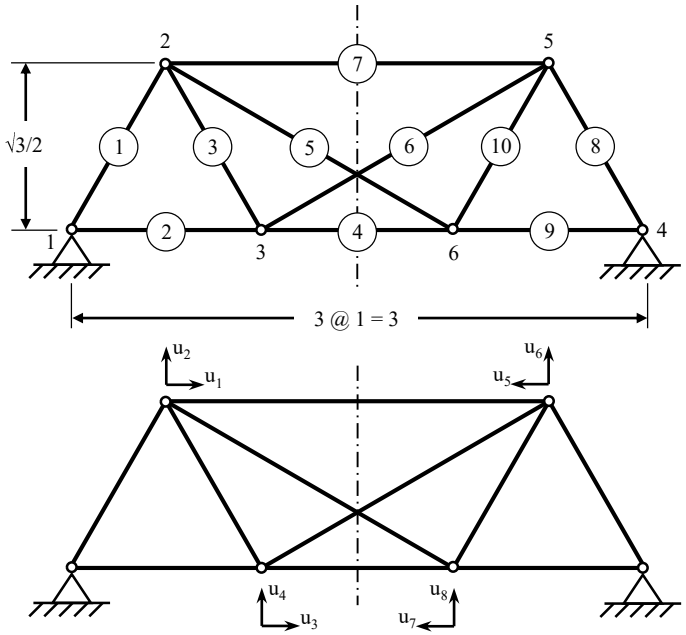


Figure 1: Symmetric 8-DoF truss structure adapted from Kaveh and Nikbakht [20]

Eq. (5) with block matrices given by

$$\boldsymbol{K} = \left[\begin{array}{c|c} \boldsymbol{K}^{(1)} & \boldsymbol{K}^{(2)} \\ \hline \boldsymbol{K}^{(2)} & \boldsymbol{K}^{(1)} \end{array}\right] = \left[\begin{array}{rrrr|rrrr} 1.43 & -0.25 & -0.25 & 0.43 & 0.50 & 0 & 0.43 & 0.25 \\ -0.25 & 1.64 & 0.43 & -0.75 & 0 & 0 & -0.25 & -0.14 \\ -0.25 & 0.43 & 2.68 & -0.18 & 0.43 & -0.25 & 1 & 0 \\ 0.43 & -0.75 & -0.18 & 0.89 & 0.25 & -0.14 & 0 & 0 \\ \hline 0.50 & 0 & 0.43 & 0.25 & 1.43 & -0.25 & -0.25 & 0.43 \\ 0 & 0 & -0.25 & -0.14 & -0.25 & 1.64 & 0.43 & -0.75 \\ 0.43 & -0.25 & 1 & 0 & -0.25 & 0.43 & 2.68 & -0.18 \\ 0.25 & -0.14 & 0 & 0 & 0.43 & -0.75 & -0.18 & 0.89 \end{array}\right] \tag{40}$$

$$\boldsymbol{M} = \left[\begin{array}{c|c} \boldsymbol{M}^{(1)} & \boldsymbol{M}^{(2)} \\ \hline \boldsymbol{M}^{(2)} & \boldsymbol{M}^{(1)} \end{array}\right] = \left[\begin{array}{rrrr|rrrr} 1.91 & 0 & 0.17 & 0 & -0.33 & 0 & -0.29 & 0 \\ 0 & 1.91 & 0 & 0.17 & 0 & 0.33 & 0 & 0.29 \\ 0.17 & 0 & 1.58 & 0 & -0.29 & 0 & -0.17 & 0 \\ 0 & 0.17 & 0 & 1.58 & 0 & 0.29 & 0 & 0.17 \\ \hline -0.33 & 0 & -0.29 & 0 & 1.91 & 0 & 0.17 & 0 \\ 0 & 0.33 & 0 & 0.29 & 0 & 1.91 & 0 & 0.17 \\ -0.29 & 0 & -0.17 & 0 & 0.17 & 0 & 1.58 & 0 \\ 0 & 0.29 & 0 & 0.17 & 0 & 0.17 & 0 & 1.58 \end{array}\right] \tag{41}$$

Transforming these matrices by $\boldsymbol{P}$ in Eq. (8) to be block-diagonal yields the condensed matrices $\bar{\boldsymbol{K}}^{(1)}$, $\bar{\boldsymbol{K}}^{(2)}$, $\bar{\boldsymbol{M}}^{(1)}$ and $\bar{\boldsymbol{M}}^{(2)}$. This leads to the separable eigenvalue problem in Eq. (9), which give the eigenvalues of $\boldsymbol{K}$ and $\boldsymbol{M}$ as the union of the eigenvalues for the condensed problems:

$$\{\lambda^{(1)}\} = \{0.08,\ 0.88,\ 1.76,\ 2.70\} \qquad \{\lambda^{(2)}\} = \{0.19,\ 0.42,\ 0.83,\ 1.81\} \tag{42}$$

Therefore, the eigenvalues are $\{\lambda\} = \{0.08,\ 0.19,\ 0.42,\ 0.83,\ 0.88,\ 1.76,\ 1.81,\ 2.70\}$ with modes 1, 5, 6, and 8 coming from the first condensed problem (symmetric) and modes 2, 3, 4, and 7 from the second condensed problem (anti-symmetric). The associated eigenvectors for the separable problems are

$$\bar{\boldsymbol{\phi}}^{(1)} = \left[\begin{array}{rrrr} -0.30 & 1 & 1 & 0.20 \\ 0.58 & 0.77 & -0.82 & 0.05 \\ 0.04 & -0.21 & -0.06 & 1 \\ 1 & -0.42 & 0.96 & -0.04 \end{array}\right] \qquad \bar{\boldsymbol{\phi}}^{(2)} = \left[\begin{array}{rrrr} 1 & 0.31 & 0.59 & -0.57 \\ -0.33 & 0.63 & 1 & 0.80 \\ 0.68 & 0.02 & -0.90 & 1 \\ -0.40 & 1 & -0.84 & -0.36 \end{array}\right] \tag{43}$$

These can be transformed into the eigenvectors of the full problem using Eq. (7), which are separated into

$$\boldsymbol{\phi} = \left[\begin{array}{cc} \bar{\boldsymbol{\phi}}^{(1)} & \bar{\boldsymbol{\phi}}^{(2)} \\ \bar{\boldsymbol{\phi}}^{(1)} & -\bar{\boldsymbol{\phi}}^{(2)} \end{array}\right] \tag{44}$$

Note that this doesn't reflect the typical mode ordering (based on ascending eigenvalue), but is used to show that the mode shapes are symmetric and anti-symmetric. The mode shapes are depicted in Fig. 2.

The model was parametrized to modify the stiffness (or equivalently, Young's modulus) of each element l out of a total $p = 10$

$$\boldsymbol{K}(\boldsymbol{\theta}) = \sum_{l=1}^{p} \boldsymbol{K}_l(1 - \theta_l) \tag{45}$$

where $\boldsymbol{\theta}$ is the vector of parameters and $\boldsymbol{K}_l$ is the element stiffness matrix of element l. Therefore, the derivatives are simple, with $\boldsymbol{K}_{,l} = -\boldsymbol{K}_l$ and $\boldsymbol{M}_{,l} = \boldsymbol{0}$. The parameter sets $\{1, 8\}$, $\{2, 9\}$, and $\{3, 10\}$ were symmetric.

The natural frequency sensitivities and mode shape sensitivities were computed numerically for the purposes of verification. A selection of these mode shape sensitivities for symmetric parameters are shown in Fig. 3. The parameter (element) is indicated with a bold line while the sensitivity is indicated by a red arrow for each DoF.

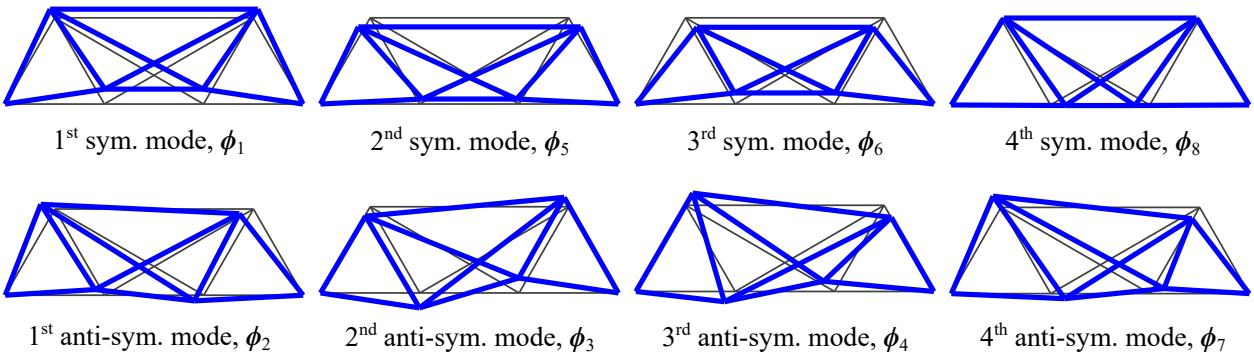


Figure 2: Truss mode shapes

This represents how the mode shape changes as a result of perturbing the indicated parameter. The numerically-computed sensitivities obeyed the reflection properties noted in Section 4, with $\boldsymbol{\phi}_{1,1} = \boldsymbol{T}\boldsymbol{\phi}_{1,8}$ and $\boldsymbol{\phi}_{2,3} = -\boldsymbol{T}\boldsymbol{\phi}_{2,10}$. In words, mode shape sensitivities for symmetric parameters are reflections for a symmetric mode, while they are negative reflections for anti-symmetric modes. Visually, $\boldsymbol{\phi}_{1,1}$ and $\boldsymbol{\phi}_{1,8}$ appear to be near-opposite, with most of the sensitivities having opposite direction between the two parameters. Conversely, $\boldsymbol{\phi}_{2,3}$ and $\boldsymbol{\phi}_{2,10}$ appear quite similar, especially for the large magnitude terms on nodes 3 and 6.

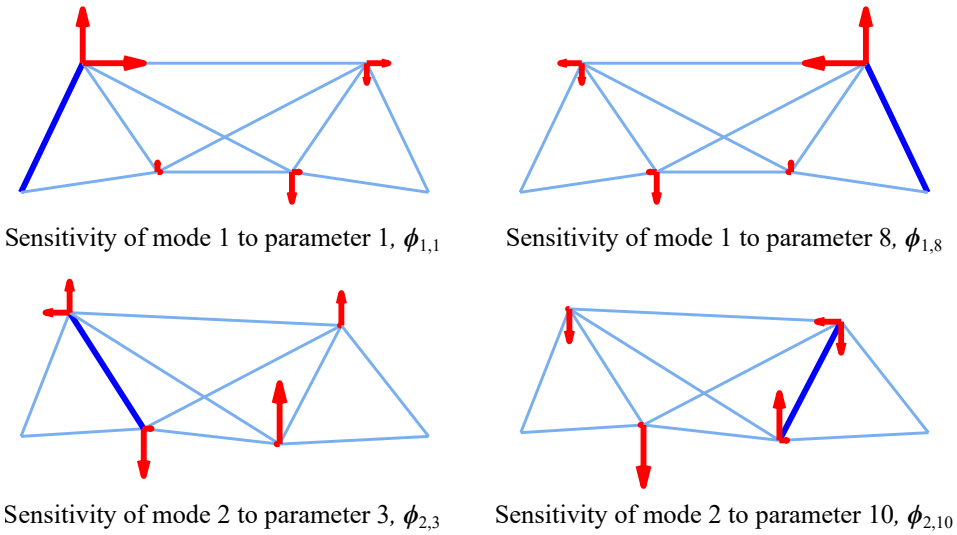


Figure 3: Selected mode shape sensitivities for symmetric parameters

To quantify the similarity between sensitivity vectors, the cosine distance between mode shape sensitivities was computed between all parameter pairs using Eq. (27). This comprised $p(p-1)/2 = 45$ unique pairs of parameters, of which 3 represented the symmetric parameter pairs. This was performed for all 8 modes, with results shown in Fig. 4. This allowed for a limited comparison of cosine distance between mode shape sensitivities for symmetric parameter pairs and for all other pairs.

The cosine distance between mode shape sensitivities for the three symmetric parameter pairs was largely similar to the behavior for any other parameter pair. For all symmetric parameter pairs, $d_{\cos}$ was near its maximum for modes 1 and 4-8, showing little preference between symmetric and anti-symmetric mode shapes. A maximal value of $d_{\cos} = 2$ indicated that the mode shapes sensitivities were opposites, as discussed in $\boldsymbol{\phi}_{1,1}$ and $\boldsymbol{\phi}_{1,8}$ of Fig. 3. Conversely, lower values of $d_{\cos}$ can be noted for the symmetric parameter pairs on modes 2 and 3. In particular, $d_{\cos}(\boldsymbol{\phi}_{2,3}, \boldsymbol{\phi}_{2,10})$ was quite small (≈ 0.14), as suggested in the discussion of Fig. 3.

The other 42 parameter pairs showed largely uniform behavior for modes 1-3 and 8, with few values below 1. Conversely, modes 4-7 showed a large proportion of extreme values, having many pairs with $d_{\cos}$ near zero and also near 2. This doesn't seem to indicate a difference between symmetric and anti-symmetric mode shape sensitivity

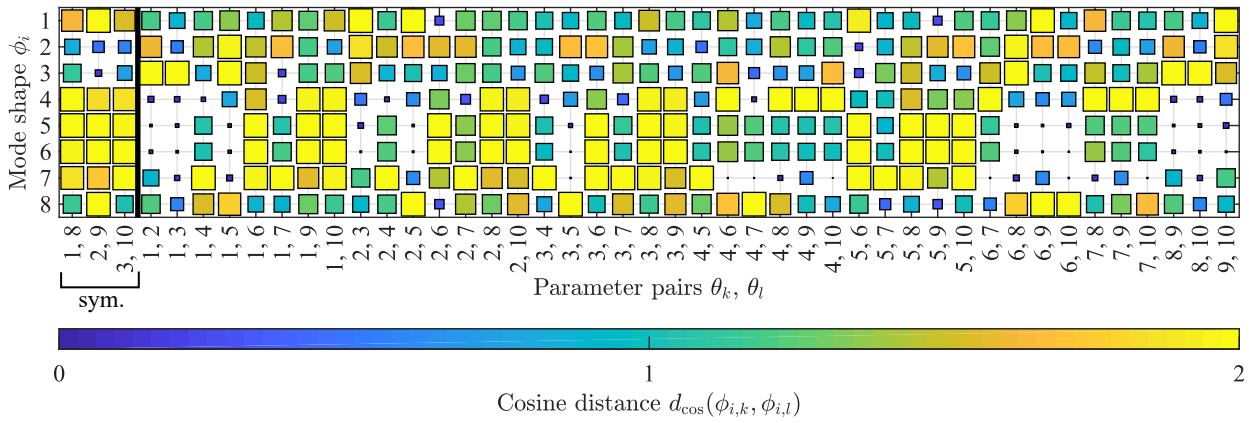


Figure 4: Cosine distances between parameter pairs, separated into symmetric pairs and all other pairs

behavior, as these distinct behaviors encompassed both symmetric and anti-symmetric modes.

The mean of $d_{\cos}$ (across all modes) for the symmetric parameter pairs was approximately 1.5, while it was near 1.1 for the other parameter pairs. While this indicates that symmetric parameter pairs are more likely to exhibit significantly different mode shape sensitivities, this is not a guarantee. Symmetric parameter pairs still exhibited significant variability, with $d_{\cos}$ values noted near zero as well as near 2. This is an important result, as it indicates that there are no special properties (e.g. equality, negation) of mode shape sensitivities between symmetric parameters. For the purposes of clustering and subset selection, symmetric parameters aren't expected to exhibit different behavior than any arbitrary pair of parameters. Therefore, they are expected to contribute to asymmetry in clustering and prevent orthogonalization of symmetric parameters in subset selection. For sensitivity-based updating, they are expected to behave like any other parameter pair, with no reason to suspect that they won't contribute to the column-space (rank) of the sensitivity matrix.

6. Conclusions

The properties of natural frequency and mode shape sensitivities were explored in the context of structures possessing at least one plane of reflectional symmetry. For these structures, it was shown that the stiffness and mass matrices can be partitioned into a special form, which reduces the eigenvalue problems into two smaller problems. The reduced problems provide exclusively symmetric or anti-symmetric mode shapes for the symmetric structure. The special symmetry properties of the mode shapes were then used to explore the derivatives (or sensitivities) of natural frequencies and mode shapes to changes in symmetric parameters. It was analytically proved that natural frequency sensitivities are necessarily equal between symmetric parameters, while mode shape sensitivity vectors are always unequal. These topics were applied to a small example truss with symmetry to quantify the difference between mode shape sensitivity vectors, as measured by cosine distance. It was noted that symmetric parameters generally had greater cosine distance between their mode shape sensitivity vectors, as compared to other (asymmetric) parameter pairs, but still exhibited significant variability.

For sensitivity-based clustering, this proves that using natural frequency sensitivities will lead to symmetric clusters, while incorporating mode shape sensitivities will tend to create asymmetric clusters, as observed previously by the authors [3]. This stems from mode shape sensitivities being necessarily unequal between symmetric parameters, while natural frequency sensitivities are always equal between symmetric parameters. For parameter subset selection (which uses the sensitivity matrix), if natural frequency sensitivities are used exclusively, then only one of a pair of symmetric parameters can possibly be chosen, as its counterpart will be orthogonalized. However, incorporating mode shape sensitivity will differentiate the sensitivities of the symmetric parameters, possibly allowing both parameters to be selected. This has similar ramifications for sensitivity-based updating, where use of natural frequency sensitivities will necessarily lead to linearly dependent columns of the sensitivity vector, since the sensitivity vectors for symmetric parameters will be equal. Using mode shape data, with or without natural frequency data, is expected to ameliorate this problem and improve the rank of the sensitivity matrix.

The derived results were validated on the example problem, but the degree of variability between mode shape sensitivities corresponding to symmetric parameters deserves further investigation on a variety of structures.

Acknowledgments

The authors gratefully acknowledge Columbia University's Graduate School of Arts and Sciences in support of the first author through the Guggenheim and Presidential Fellowships. This work was partially supported by the U.S. National Science Foundation (Grant No. CMMI-1563364).

References

[1] J. E. Mottershead, M. Link, M. I. Friswell, The sensitivity method in finite element model updating: a tutorial, Mech. Syst. Signal Process. 25 (7) (2011) 2275–2296. doi:10.1016/j.ymssp.2010.10.012.

[2] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: a review, Mech. Syst. Signal Process. 56–57 (2015) 123–149. doi:10.1016/j.ymssp.2014.11.001.

[3] D. T. Bartilson, J. Jang, A. W. Smyth, Finite element model updating using objective-consistent sensitivity-based parameter clustering and Bayesian regularization, Mech. Syst. Signal Process. 114 (2019) 328–345. doi:10.1016/j.ymssp.2018.05.024.

[4] H. Ahmadian, J. E. Mottershead, M. I. Friswell, Regularisation methods for finite element model updating, Mech. Syst. Signal Process. 12 (1) (1998) 47–64. doi:10.1006/mssp.1996.0133.

[5] M. I. Friswell, J. E. Mottershead, H. Ahmadian, Combining subset selection and parameter constraints in model updating, J. Vib. Acoust. 120 (4) (1998) 854–859. doi:10.1115/1.2893911.

[6] G. Lallement, J. Piranda, Localization methods for parametric updating of finite element models in elastodynamics, in: International Modal Analysis Conference, 8th, 1990, pp. 579–585.

[7] H. Shahverdi, C. Mares, W. Wang, J. E. Mottershead, Clustering of parameter sensitivities: examples from a helicopter airframe model updating exercise, Shock Vib. 16 (1) (2009) 75–87. doi:10.3233/SAV-2009-0455.

[8] J. Jang, A. W. Smyth, Model updating of a full-scale FE model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters, Mech. Syst. Signal Process. 83 (2017) 337–355. doi:10.1016/j.ymssp.2016.06.018.

[9] J. Jang, A. W. Smyth, Bayesian model updating of a full-scale finite element model with sensitivity-based clustering, Struct. Control Health Monit. 24 (11) (2017) e2004. doi:10.1002/stc.2004.

[10] L. Rokach, O. Maimon, Clustering Methods, Springer, Boston, 2005, Ch. 15, pp. 321–352. doi:10.1007/0-387-25465-X_15.

[11] O. S. Salawu, Detection of structural damage through changes in frequency: a review, Eng. Struct. 19 (9) (1997) 718–723. doi:10.1016/S0141-0296(96)00149-6.

[12] N. Bicanic, H.-P. Chen, Damage identification in framed structures using natural frequencies, Int. J. Numer. Meth. Eng. 40 (23) (1997) 4451–4468. doi:10.1002/(SICI)1097-0207(19971215)40:23<4451::AID-NME269>3.0.CO;2-L.

[13] H.-P. Chen, Structural damage identification from measured vibration modal data, Ph.D. thesis, Department of Civil Engineering, University of Glasgow, UK (1998).

[14] P. G. Glockner, Symmetry in structural mechanics, J. Struct. Div. ASCE 99 (1) (1973) 71–89.

[15] A. Zingoni, Group-theoretic exploitations of symmetry in computational solid and structural mechanics, Int. J. Numer. Meth. Eng. 79 (3) (2009) 253–289. doi:10.1002/nme.2576.

[16] T. J. Healey, J. A. Treacy, Exact block diagonalization of large eigenvalue problems for structures with symmetry, Int. J. Numer. Meth. Eng. 31 (2) (1991) 265–285. doi:10.1002/nme.1620310205.

[17] A. Kaveh, M. A. Sayarinejad, Eigensolutions for matrices of special structures, Commun. Numer. Meth. Eng. 19 (2003) 125–136. doi:10.1002/cnm.576.

[18] A. Kaveh, B. Salimbahrami, Eigensolution of symmetric frames using graph factorization, Commun. Numer. Meth. Eng. 20 (12) (2004) 889–910. doi:10.1002/cnm.711.

[19] A. Kaveh, M. A. Sayarinejad, Graph symmetry and dynamic systems, Comput. Struct. 82 (23-26) (2004) 2229–2240. doi:10.1016/j.compstruc.2004.03.066.

[20] A. Kaveh, M. Nikbakht, Decomposition of symmetric mass–spring vibrating systems using groups, graphs and linear algebra, Commun. Numer. Meth. Eng. 23 (7) (2007) 639–664. doi:10.1002/cnm.913.

[21] A. Zingoni, On group-theoretic computation of natural frequencies for spring–mass dynamic systems with rectilinear motion, Commun. Numer. Meth. Eng. 24 (11) (2008) 973–987. doi:10.1002/cnm.1003.

[22] A. Zingoni, Group-theoretic insights on the vibration of symmetric structures in engineering, Phil. Trans. R. Soc. A 372 (2008) (2014) 20120037. doi:10.1098/rsta.2012.0037.

[23] R. L. Fox, M. P. Kapoor, Rates of change of eigenvalues and eigenvectors, AIAA J. 6 (12) (1968) 2426–2429. doi:10.2514/3.5008.

[24] S. Adhikari, Rates of change of eigenvalues and eigenvectors in damped dynamic system, AIAA J. 37 (11) (1999) 1452–1458. doi:10.2514/2.622.