Published article available online: Struct Control Health Monit. 2020;e2539.

PREPRINT

Sensitivity-based singular value decomposition 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. Reparametrization 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 $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$. 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

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

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* $\gamma = p - 2\alpha\,\text{tr}(\boldsymbol{H}_i^{-1} \boldsymbol{W}_\theta)$
  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 $\boldsymbol{C}$ 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].

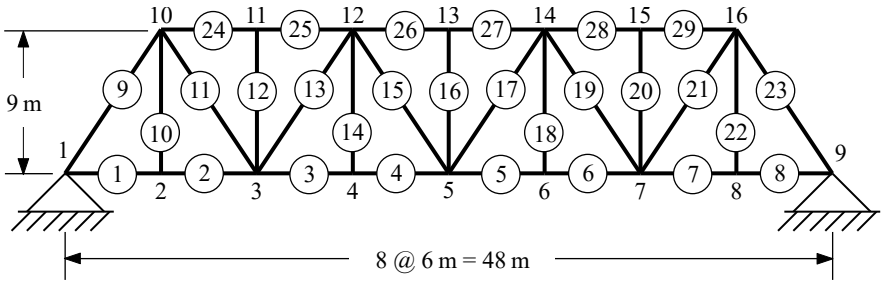


FIGURE 1 Truss structure layout with element and node numbers indicated

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

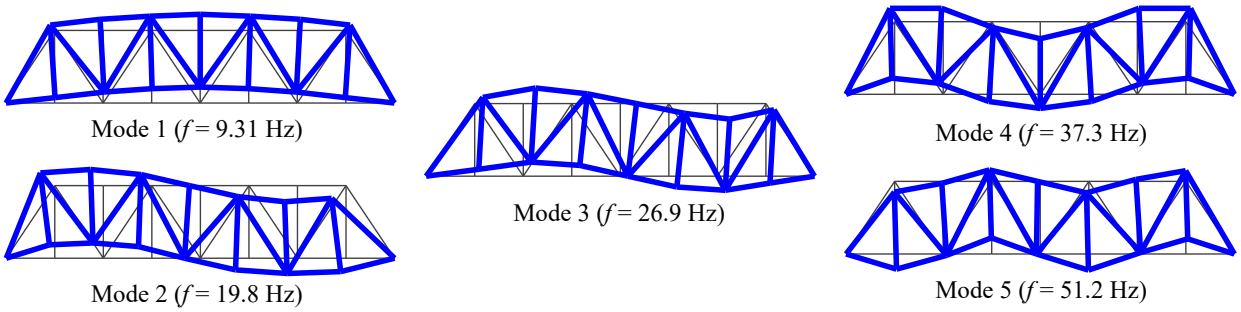


FIGURE 2 Truss mode shapes and natural frequencies

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.

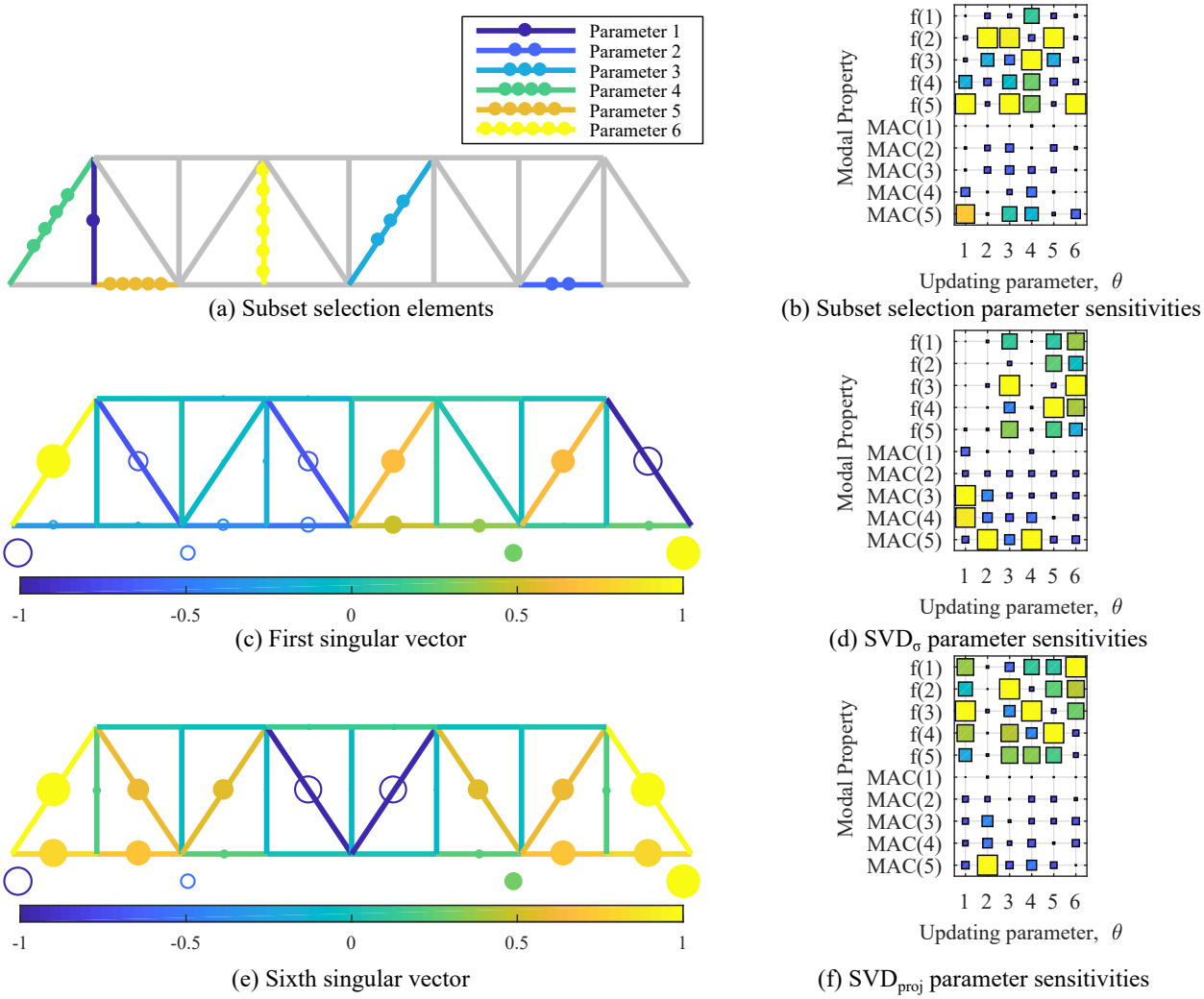
6.2 | Parametrization

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

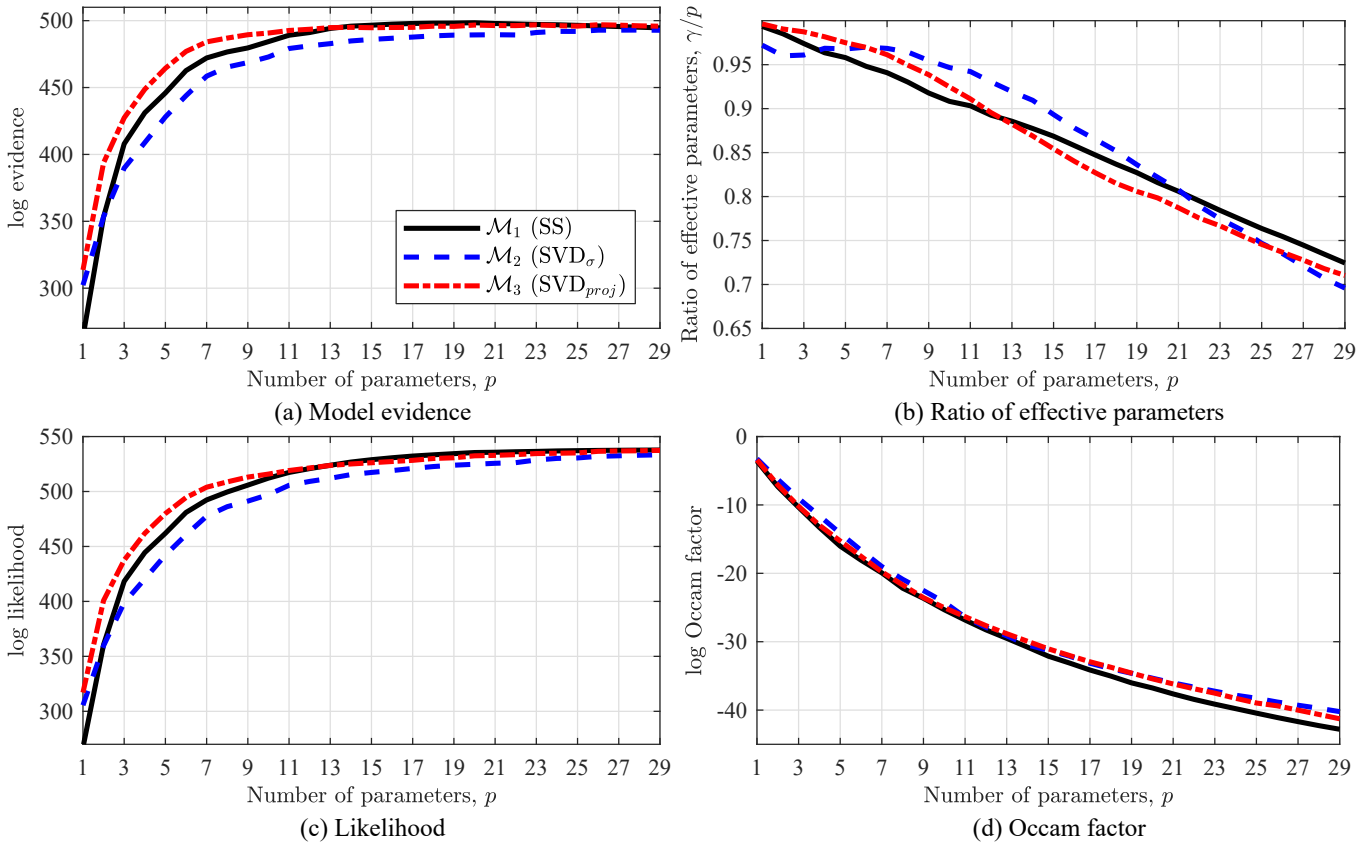
(f) SVD_{proj} parameter sensitivities

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

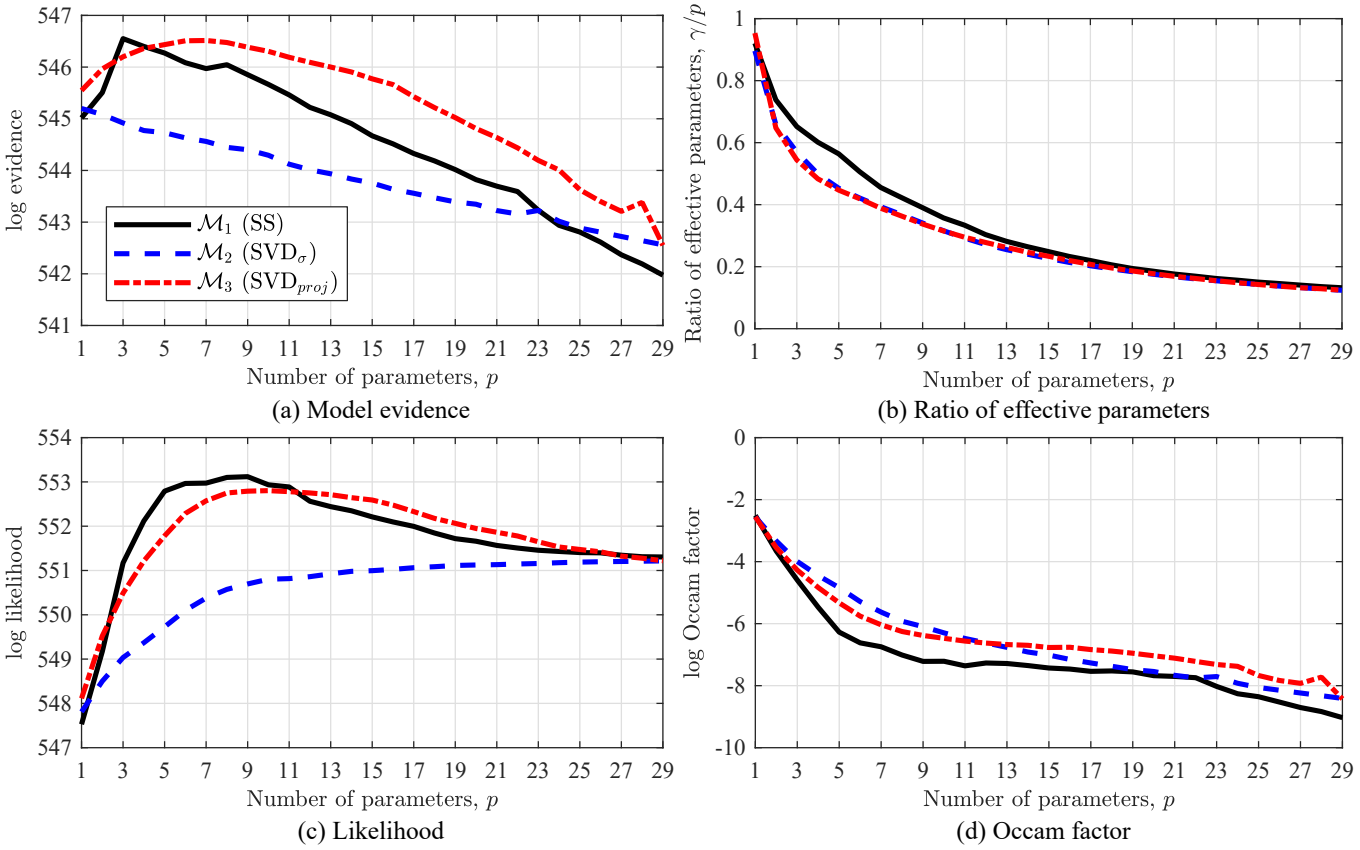
(c) Likelihood (d) Occam factor

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.



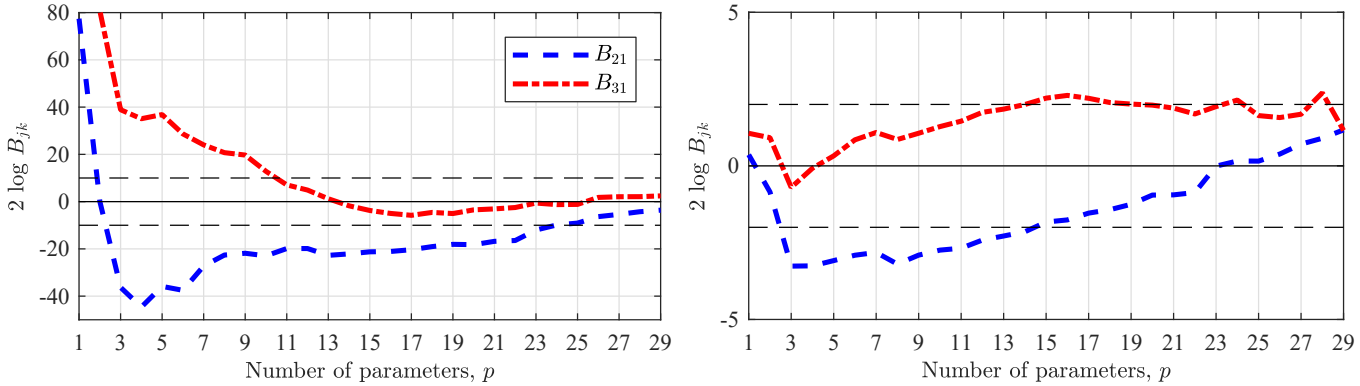
(a) Model evidence (b) Ratio of effective parameters

(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 $\boldsymbol{C}_{\tilde{\boldsymbol{z}}}$, 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 $w^s_{rj} = (0.05\ \text{std}(\tilde{\boldsymbol{\psi}}_j))^{-2}$, 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].

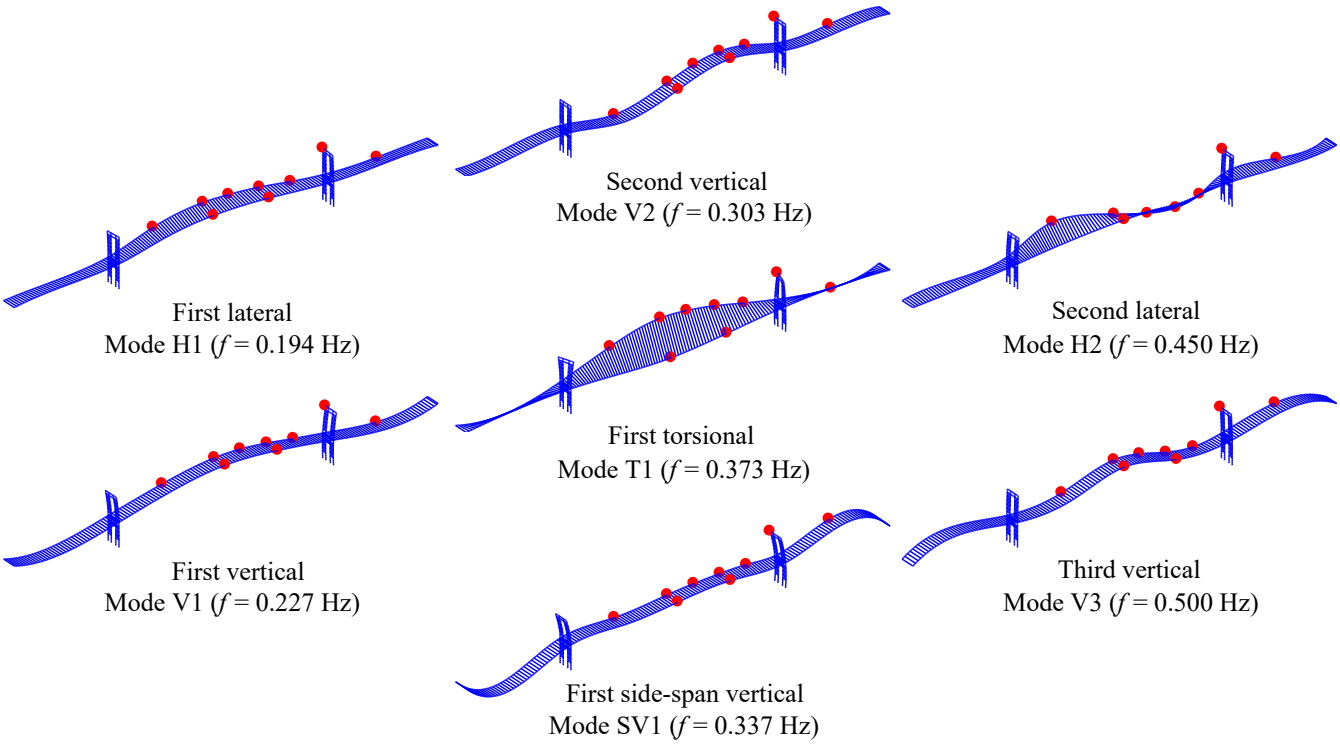


FIGURE 7 Suspension bridge measured modes (measurement locations indicated by red dots)

The natural frequencies 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 Figure 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 Figure 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.

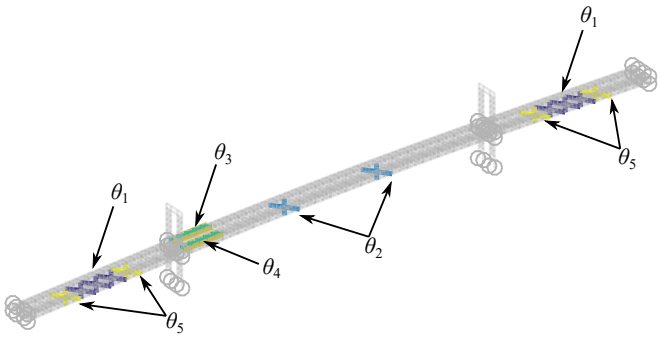
Both of the SVD-based parametrizations showed less specialization than SS, with parameters generally having significant effect on multiple modal properties. In particular, pretty much every parameter in SVD_σ affected multiple natural frequencies, generally with lower impact on mode shapes, while SVD_{proj} 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

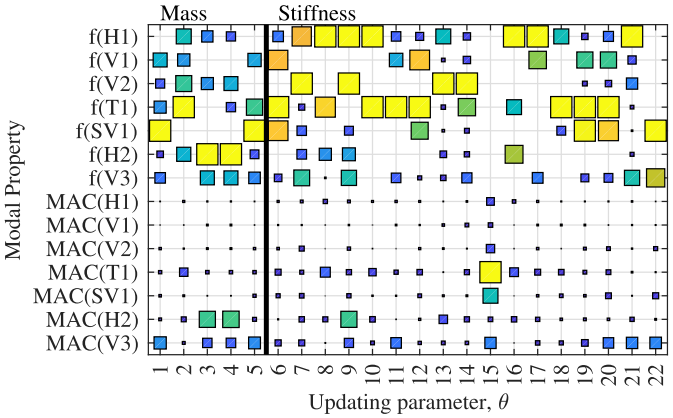
Model updating proceeded using the three parametrization methods, each with 5 mass and 17 stiffness updating parameters. The regularized objective function in Equation 14 was used with FE model parameter covariance matrix $\boldsymbol{W}_\delta = \boldsymbol{I}_d$. As noted in Section 6, all parametrizations 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).

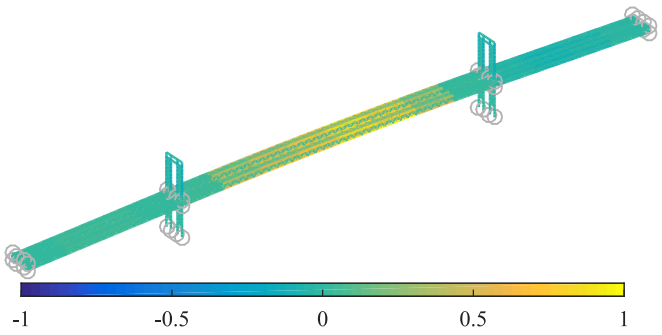
SS parametrization produced mediocre reductions in model error, giving 34% less total error than the initial model, splitting its total error almost equally between E_r^f and E_r^s. SS mildly improved natural frequency results across all modes. SS was relatively unsuccessful in reducing large natural 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.



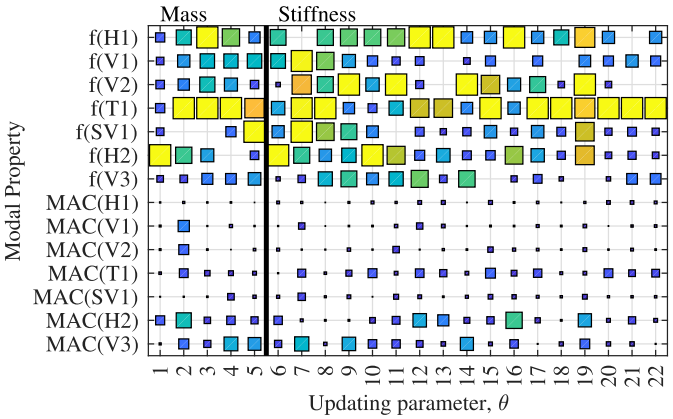
(a) Subset selection mass parameters



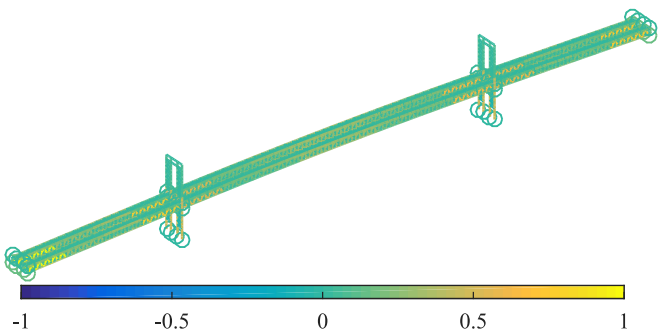
(b) Subset selection parameter sensitvities



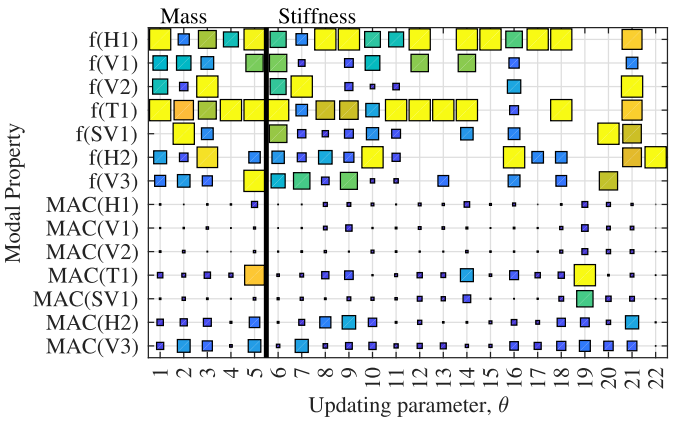
(c) Third mass singular vector



(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.