

University of Thessaly
School of Engineering
Department of Mechanical Engineering

PhD Thesis

**Bayesian Tools for Uncertainty Quantification and Propagation (UQP) in
Structural Dynamics Simulations**

by
Xinyu Jia

A thesis submitted
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

November, 2021

© 2021 Xinyu Jia

The approval of this thesis by the Mechanical Engineering Department of the School of Engineering of University of Thessaly does not imply the acceptance of the personal views of the author (L. 5342/32 ar. 202 par. 2).

dyvirt-etn.com

Dynamic virtualisation: modelling performance of engineering structures



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 764547.

Approved by the members of the seven-member examination committee:

First Examiner (Supervisor)	Dr. Costas Papadimitriou Professor, Department of Mechanical Engineering, University of Thessaly
Second Examiner	Dr. David Wagg Professor, Department of Mechanical Engineering, The University of Sheffield
Third Examiner	Dr. Eleni Chatzi Associate Professor, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich
Fourth Examiner	Dr. Sotirios Natsiavas Professor, School of Mechanical Engineering, Aristotle University of Thessaloniki
Fifth Examiner	Dr. Dimitrios Giagopoulos Associate Professor, Department of Mechanical Engineering, University of Western Macedonia
Sixth Examiner	Dr. Lambros Katafygiotis Professor, Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology
Seventh Examiner	Dr. Ilias Dimitrakopoulos Associate Professor, Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology

Acknowledgements

I would like to express my deepest and most sincere gratitude to my advisor Prof. Costas Papadimitriou for his insightful guidance, sustained encouragement and extensive support throughout my Ph.D. study. I much appreciate for his understanding, patience and great help at the beginning of my research. His guidance and a lot of further discussions between us have been the key to my academic growth over the past three years. He taught me how to be a good learner, a good thinker and a good researcher. I benefited a lot from him not only the knowledge itself but also the way of critical thinking, the rigorous attitude, and diligence towards research. His passion and enthusiasm for research inspire me to pursue a career in academia, and I am looking forward to continuing our collaborations.

I would also like to thank Prof. David Wagg and Dr. Andrew Halfpenny, for their invaluable help on my life and research during my first secondment in The University of Sheffield and HBM Prencscia Company in the UK. Special gratitude goes to my cross institutional supervisor, Prof. Eleni Chatzi in ETH Zurich, for her great support and of course the resources for my second secondment. My appreciation also extends to other examination committee members of my Ph.D. thesis including Prof. Lambros Katafygiotis, Prof. Ilias Dimitrakopoulos, Prof. Sotirios Natsiavas and Prof. Dimitrios Giagopoulos. Special thanks go to my collaborator Dr. Omid Sedehi from HKUST for his guidance, suggestions and discussions on my research. Special thanks also go to the collaborator Prof. Babak Moaveni from Tufts University.

Many thanks to my lab mates and friends in System Dynamics Laboratory including Tulay Ercan, George Trouzas, Matina Moraiti, Menghao Ping (visitor from Hunan University) and Victor Flores Terrazas (visitor from HKUST) for the enjoyable discussions in research and indeed the wonderful time that we had during my Ph.D. Special thanks to my friend Tulay for a lot of discussions and arguments on the life which enables me to learn, to think and to face problems from different perspectives. Many thanks to the staff Nicoletta Sachinidou in the department for her great help on my life in the past three years.

The DyVirt project (Grant No. 764547) under the Marie Skłodowska-Curie Innovative Training Network is also gratefully acknowledged. Much appreciate for offering me with financial support for my research, living and traveling expenses over the past three years. Special thanks also go to the other thirteen early stage researchers within the project. Thanks for the great moments that we had during the project workshops, training weeks, as well as summer schools.

Lastly, I would like to express my gratitude to my parents for their endless love, continuous support and encouragement to allow me to explore the beauty of the world. I am definitely in debt to them for not being able to celebrate the Chinese New Year together with them for so many years.

Xinyu Jia

Bayesian Tools for Uncertainty Quantification and Propagation (UQP) in Structural Dynamics Simulations

Xinyu Jia

Department of Mechanical Engineering, University of Thessaly, 2021

Supervisor: Prof. Dr. Costas Papadimitriou

Abstract

This thesis is focused on the development of Bayesian techniques for uncertainty quantification and propagation (UQP) in engineering simulations based on physics-based models of dynamic structures and measurements collected during system operation. It introduces a hierarchical Bayesian modeling (HBM) framework to account for the uncertainty due to the variabilities that arise from model error, experimental data, manufacturing process, assembling process as well as nonlinear mechanisms activated under different loading conditions. First, it extends the HBM framework developing it further for model inference based on modal properties. Then, it generalizes the framework to nonlinear dynamical systems for calibration and quantification of uncertainties of parameters of nonlinear models. Finally, it advances the framework to account for multi-level physics-based modeling of systems and multi-level models of uncertainties using multi-level test data. Asymptotic approximations are introduced, developed and incorporated into the HBM framework in order to gain more insights on the interpretation of diverse sources of uncertainties. Introducing such approximations can substantially reduce the computational burden of the HBM framework compared to the high computational effort required in a full sampling procedure. Simulated and experimental studies are conducted to verify the effectiveness of the proposed methodologies. It is shown that the proposed HBM framework provides a better account for the parameter uncertainties, distinguishing between irreducible and reducible uncertainties, while the conventional Bayesian modeling (CBM) framework often underestimates the uncertainties for the parameters and aggregates such uncertainties into model error. Moreover, this thesis revisits the underestimation of uncertainties issues within a classical Bayesian point of view developing further data features-based models and presenting novel formulations for the constructions of the likelihood function. It is illustrated that the proposed methods offer consistent parameter uncertainties which is independent of

the sampling rate used for the accurate representation of the same time history responses. Furthermore, for the issues where the PDFs or statistics of the measurements are available, this thesis also presents a methodology under a hierarchical modeling setting for the model parameters to account for the uncertainty due to variability. As an alternative to HBM method, the presented approach is successfully applied to structural dynamical example and fatigue S-N curve analysis for the parameter estimation and predictions given the available statistics of the measured quantities. The proposed methodologies in this thesis have the great potential to be applied to other disciplines of engineering and science.

Contents

Chapter 1. Introduction	1
1.1 Research Background and Motivation	1
1.2 Research Objectives	6
1.3 Contributions and Organization of the Thesis	7
1.4 References	12
Chapter 2. Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Features.....	20
2.1 Introduction	21
2.2 Proposed HBM Framework.....	22
2.3 Computational Algorithm.....	35
2.4 Illustrative Examples	36
2.5 Conclusion.....	52
2.6 References	56
Chapter 3. Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework	61
3.1 Introduction	62
3.2 Hierarchical Bayesian Nonlinear Model Updating.....	64
3.3 Application to Nonlinear Systems using Bouc-Wen hysteresis	75
3.4 Illustrative Examples	77
3.5 Conclusion.....	92
3.6 References	94
Chapter 4. Hierarchical Bayesian Learning Framework for Multi-Level Modeling using Multi-level Data	100
4.1 Introduction	101
4.2 Hierarchical Bayesian Modeling Framework for Multi-level Models	103
4.3 Numerical Example.....	121
4.4 Concluding Remarks	139
4.5 References	143
Chapter 5. Data Features-based Bayesian Learning for Time-domain Model Updating and Robust Predictions in Structural Dynamics	148
5.1 Introduction	149
5.2 Classical Formulation for Bayesian Model Updating	150
5.3 Data Features Models for Bayesian Model Updating and Robust Predictions.....	152
5.4 Illustrative Example	158
5.5 Conclusion.....	171
5.6 References	171

Chapter 6. Statistics-based Bayesian Modeling Framework for Uncertainty Quantification and Propagation.....	175
6.1 Introduction	176
6.2 Proposed Bayesian Modeling Framework.....	177
6.3 Application to Structural Dynamics using Measured Modal Properties	184
6.4 Simulated Example	185
6.5 Estimation of S-N Curve Model Parameters using Fatigue Data	193
6.6 Conclusion.....	198
6.7 References	200
Chapter 7. Conclusions and Future Work	204
7.1 Concluding Remarks	204
7.2 Future Work	207

List of Figures

Fig. 2-1 Graphical representation of the proposed HBM framework for probabilistic quantification and propagation	24
Fig. 2-2 (a) 3-DOF shear model of a building and (b) base earthquake excitation	37
Fig. 2-3 Posterior distribution of hyper parameters and prediction error parameter for case 2. The diagonal sub-figures depict the marginal distributions for each parameter, and sub-figures below the diagonal show the pairwise contour plots.	41
Fig. 2-4 Identification uncertainty and parameter uncertainty as a function of the number of datasets N_D , with N_D ranging from 1 to 100 in linear scale for all cases 1-4 (top-to-bottom) .	41
Fig. 2-5 Prediction of the ratio of modal frequencies over the nominal values obtained by propagating the uncertainty of (a) the structural parameters for case 2 (b) all parameters for Case 2 (c) the structural parameters for case 3 (d) all parameters for Case 3.....	43
Fig. 2-6 Prediction of the mode shapes by propagating the uncertainty of (a) the structural parameters for case 2 (b) all parameters for case 2 (c) the structural parameters for case 3 (d) all parameters for case 3	43
Fig. 2-7 Prediction of the response time histories of the third story for case 2 considering (a) only structural parameters uncertainty (b) total uncertainty	44
Fig. 2-8 Prediction of the response time histories of the third story for case 3 considering (a) only structural parameters uncertainty (b) total uncertainty	45
Fig. 2-9 Posterior predictive distributions of the modal frequencies normalized over the nominal values for the simulated Case 2 by using (a) HBM (b) CBM.....	46
Fig. 2-10 (a) Posterior predictive distribution of the maximum drift in the simulated case 2 (b) the exceedance probability of the maximum drift	46
Fig. 2-11 (a) 3-Story prototype structure tested on a shaking table (b) 3-DOF linear numerical structural model	47
Fig. 2-12 Posterior distribution of hyper-parameters and prediction error parameter	48
Fig. 2-13 (a) Uncertainty quantification of the stiffness parameters obtained using the approximation A-2 (b) magnified views of the UB on a logarithmic scale	49
Fig. 2-14 Predictions of modal frequencies normalized over the nominal values considering (a) only structural parameters uncertainty, (b) both the structural parameters and prediction error uncertainties	50

Fig. 2-15 Predicted mode shapes considering (a) only structural parameters uncertainty, (b) both structural and model error parameter uncertainties	50
Fig. 2-16 Predicted acceleration response history of the third story considering (a) only structural parameters uncertainty, (b) both structural and prediction error uncertainties	51
Fig. 3-1 Graphical representation of the proposed hierarchical Bayesian modeling framework for nonlinear model updating	66
Fig. 3-2 (a) SDOF system with BW hysteresis (b) SDOF linear model of the system in (a) ..	78
Fig. 3-3 Posterior distribution of the hyper-parameters	79
Fig. 3-4 Estimates of identification uncertainty and ensemble uncertainty	81
Fig. 3-5 Predictions of acceleration using CBM (top) and HBM (bottom)	82
Fig. 3-6 Predictions of displacement using CBM (top) and HBM (bottom)	83
Fig. 3-7 (a) 5-DOF shear model of a building system (b) Base earthquake excitation	84
Fig. 3-8 (a) Comparisons between the measured acceleration response and prediction and (b) the actual hysteresis loop and prediction from optima model of the first floor	84
Fig. 3-9 Posterior distributions of hyper parameters computed based on Eq. (3.23)	86
Fig. 3-10 Posterior distributions of hyper parameters computed based on the analytical solutions	87
Fig. 3-11 Estimates of identification uncertainty in each dataset and uncertainty due to variability overall datasets	87
Fig. 3-12 Mean estimates of the hyper parameters and prediction error parameters by using the proposed method, FS and CBM	89
Fig. 3-13 Predictions of accelerations of the first floor using CBM (top) and HBM (bottom)	90
Fig. 3-14 Predictions of displacement of the first floor CBM (top) and HBM (bottom)	90
Fig. 3-15 Predictions of nonlinear displacements of the first floor CBM (top) and HBM (bottom)	91
Fig. 3-16 Probabilistic distributions of the model and prediction error parameters	92
Fig. 4-1 Graphical representation of (a) decomposition of a physical structure into components, subsystems and system and (b) parameterized models and datasets available	104
Fig. 4-2 Uncertainty propagation for response predictions	119
Fig. 4-3 (a) Six-DOF model of a mechanical system, (b) test configuration of subsystem SS1, (c,d) test configurations of the components C1 and C2, (e) component C3	122
Fig. 4-4 The MLEs and the identification uncertainties of model parameters computed based on the individual dataset from components and sub-system	126

Fig. 4-5 Posterior distributions of the hyper parameters considering the datasets from components only (top), sub-system only (middle) and both the components and sub-system (bottom).....	126
Fig. 4-6 Posterior distributions of model error for C1, C2 and SS1	128
Fig. 4-7 Posterior distribution of model parameters considering the datasets from components only, sub-system only and both the components and sub-system	130
Fig. 4-8 Posterior distributions of model parameters.....	131
Fig. 4-9 Base input.....	132
Fig. 4-10 Predictions of displacement of (a) the second and (b) the fourth DOFs by considering the uncertainty of parameter $\theta^{(1)}$ inferred from C1 only and also inferred from all information.....	133
Fig. 4-11 Predictions of displacement of (a) the second and (b) the fourth DOFs by considering the uncertainty of parameter $\theta^{(2)}$ inferred from C2 only and inferred from all information.....	135
Fig. 4-12 Prediction of the ratio of modal frequencies to their nominal values obtained by propagating the uncertainty of the model parameter $\theta^{(2)}$ (a) inferred from C2 level datasets only (b) inferred from all available datasets	135
Fig. 4-13 Prediction of the mode shapes by propagating the uncertainty of the model parameter $\theta^{(2)}$ (a) inferred from C2 level datasets only, and (b) inferred from all available datasets.....	136
Fig. 4-14 Predictions of acceleration of the fifth DOF by propagating the uncertainties of (a) model parameters inferred from SS only and inferred from all information and (b) model parameters and prediction error	138
Fig. 4-15 Displacement and acceleration of the sixth DOF by propagating the parameter uncertainties through the system model.....	138
Fig. 5-1 (a) 10-story physical model (b) parameterized model (c) base excitation	160
Fig. 5-2 Posterior distributions of the model and prediction error parameters using (a) NORM (b) TN (c) EXP models	162
Fig. 5-3 Parameter estimation using different sampling rates of the measurements	164
Fig. 5-4 Parameter estimation using TN and ABC models	165
Fig. 5-5 PDFs of model parameters corresponding to the sampling rate 100Hz of measurements.....	165
Fig. 5-6 Parameter estimation using CONORM and MTN	167

Fig. 5-7 Posterior distributions of model and prediction error parameters using (a) CONORM and (b) MTN models.....	167
Fig. 5-8 Predictions of acceleration of the first story based on the results from NORM model (top) and from EXP model (bottom).....	170
Fig. 5-9 Predictions of displacement of the second story based on the results from NORM model (top) and from EXP model (bottom).....	171
Fig. 6-1 Proposed probabilistic modeling framework	178
Fig. 6-2 3-DOF spring mass chain system.....	186
Fig. 6-3 Samples of joint posterior distribution of hyper parameters	188
Fig. 6-4 Samples of posterior distribution of $\boldsymbol{\theta}$, $e = \{e_r, e_{r,j};, r = 1, 2, \dots, R; , j = 1, 2, \dots, n_0\}$ and ε	189
Fig. 6-5 Comparison between measured and predicted PDFs for the modal properties.....	191
Fig. 6-6 Base excitation	191
Fig. 6-7 Comparison between measured and predicted results for the displacement time history at DOF 3.....	192
Fig. 6-8 Comparison between measured and predicted results for the displacement time history at DOF 3 without consideration of error terms in prediction	192
Fig. 6-9 Samples of joint posterior distribution of hyper parameters	195
Fig. 6-10 Samples of posterior distribution of α , β and ε	196
Fig. 6-11 Predicted 90% credible intervals and comparison with measured fatigue data	197
Fig. 6-12 Comparison of PDF of experimental data and model predicted PDFs for the four stress levels	198

List of Tables

Table 1-1 Thesis chapters	7
Table 2-1 Information for 4 simulated cases	38
Table 2-2 Estimates of mean and standard deviation of the model parameters and prediction	39
Table 2-3 Estimates of mean and standard deviation of the model parameters and prediction	48
Table 3-1 Estimates of mean and standard deviation of the model parameters and prediction error parameters	80
Table 3-2 Estimates of means of hyper parameters	86
Table 3-3 Computational effort of the proposed method and FS method	89
Table 3-4 Statistical information of model parameters and prediction error parameter corresponding to Fig. 3-16.....	92
Table 4-1 Information for computational models.....	122
Table 4-2 Estimates of the means of hyper parameters and model error.....	129
Table 4-3 Computational effort of A-1, A-2 and FS methods.....	129
Table 5-1 Mean and 5%, 95% quantiles of model and prediction error parameters using NORM, TN and EXP models	163
Table 5-2 Mean and 5%, 95% quantiles of model and prediction error parameters using CONORM and MTN models.....	168
Table 5-3 Logarithm evidence values with different sampling rates.....	168
Table 6-1 Modal properties.....	186
Table 6-2 Mean and variance of simulated modal properties.....	187
Table 6-3 Upper and lower boundaries of hyper parameters.....	188
Table 6-4 Estimates of mean and standard deviation of hyper parameters	188
Table 6-5 Estimates of mean and standard deviation of model parameters and error terms .	189
Table 6-6 Predicted mean and variance of the modal properties.....	190
Table 6-7 Fatigue life test data of aluminum alloy 2524-T3	195
Table 6-8 Upper and lower bounds of hyper parameters.....	195
Table 6-9 Most probable values of hyper parameters.....	196

Chapter 1. Introduction

1.1 Research Background and Motivation

1.1.1 Structural modeling and model updating

The concept of mathematical or numerical modeling is essentially the fundamental of aspects of science and engineering [1]. Computational models are thus extensively employed to simulate the behavior of physical structures and perform engineering analyses. Finite element (FE) modeling and simulation is the most flexible and widely used numerical analysis technique in the field of structural dynamics [2–4]. Conceptually, physical dynamical structures can be described by mathematical models constituting a set of equations based on mechanics principles and empirical observations. Undoubtedly, FE modeling and simulation provides a powerful numerical tool to characterize and perceive the dynamical behavior as well as to predict the output response under future loading [5–7] without physically testing the system. This process can significantly reduce the experimental time and costs in the research and the process of product developments. However, it requires the detailed knowledge of a structure, including mechanical characteristics, physical boundary/initial conditions, nonlinearity, material properties, etc. Therefore, it poses difficulties due to a lack of knowledge of physical information or other restrictions [8]. Assumptions and idealizations are then required and implicitly applied for the representation of physical behavior. Such assumptions may detract from the quality and accuracy of the computational models [9]. As a result, achieving highly accurate of FE models for complex dynamical structures is an imperative yet challenging task.

To this end, FE model updating has been received considerable attentions in recent decades [10–13]. Essentially, FE model updating is an inverse process aiming at diminishing the discrepancies between the numerical model predictions and the experiential tests [14]. In other words, given a parameterize model, FE model updating seeks to determine a collection of most plausible/probable values of model parameters to replicate, as close as possible, the experimental test data. The experimental data used in the process of model updating are

classified into time-domain and frequency-domain tests. Response time histories data including accelerations, displacements or strains can be directly utilized to quantify the misfit between numerical models and experimental systems, while the frequency-domain data consisting of modal frequencies, mode shapes as well as damping ratios needs to be identified initially based on a procedure of operational modal analysis [15]. Such an inverse updating process belongs to the category of system identification [16] or structural health monitoring [17], producing more reliable analytical models capable of providing more accurate response predictions under future loadings.

1.1.2 Uncertainty quantification and propagation

Uncertainty inevitably exist in practically all areas of science and engineering. It can be broadly categorized into two classes: epistemic uncertainty and aleatory uncertainty [18,19]. The epistemic uncertainty is due to the lack of knowledge or poor quality of information, and thus can be reducible by given additional information, while the aleatory uncertainty is caused by the inherent randomness or unpredictability of a physic system and it is thus irreducible by collecting more information.

In the process of FE model updating, uncertainties can arise from several sources [9]. For example, parametric uncertainty (epistemic), also referred as identification uncertainty, is attributed to the incomplete knowledge of model parameters or insufficient experimental data. Model structure uncertainty (epistemic), also referred as model form uncertainty or model error, is due to the assumptions or approximations in the computational model that are used to connect the input variables and output responses. Uncertainty in the experimental data (epistemic), related to the experimental error, is owing to the uncertainty in the function relationship between the output response and measurement noise. Parameter variability (aleatory), also referred as natural variability, can be arisen from model error, environmental conditions, material properties, and component manufacturing process. Such uncertainties play an important role and need to be quantified during the process of updating models.

On the other hand, quantifying the uncertainty of output response is beneficial especially for estimating robust predictions of system performance, reliability and safety [20,21]. This leads to the development of uncertainty propagation techniques. They aim at propagating the uncertainties through the updated model under future loadings to the output quantities of interest (QoI), providing reasonable confidence interval or probability density function of output responses. Consequently, quantifying and propagating the uncertainties is indispensable for improving the accuracy of computational FE models.

1.1.3 Bayesian inference in structural dynamics

A bunch of research efforts have been contributed to deterministic model updating approaches [11,13]. However, structure unidentifiability and the presence of uncertainty impel the development of probabilistic means. Due to a rigorous probabilistic framework, Bayesian inference can rationally quantify parameter and modeling uncertainties as well as perform system identification by integrating experimental data and physics-based models. Bayesian statistical framework in structural dynamics was initially developed by Beck and Katafygiotis [22]. The framework requires the initial knowledge of uncertainties in the model parameters including it into the prior probability density function (PDF), and updates the uncertainties by estimating the posterior PDF in light of the experimental data. According to this framework, the probability of a model is interpreted as a measure of how plausible the model is within a set of models, conditional on the available information [23]. This interpretation is consistent with the Bayesian point of view and thus it provides a possibility to extend the axioms of probability theory to the fields. This pushes forward a great deal of discoveries and breakthroughs following up previous pioneering developments. Katafygiotis and Beck defined then global system identifiability, local system identifiability and system unidentifiability based on data in [24]. Vanik et al. applied the Bayesian framework to structural health monitoring based on observations of structural behavior [25]. Papadimitriou et al. implemented Bayesian inference to structural reliability analysis [26]. Beck and Au introduced Markov chain Monte Carlo (MCMC) sampling methods into the framework to handle model identification issues [27]. Yuen et al. presented a Bayesian damage detection approach and successfully applied to IASC-ASCE benchmark problem [28]. Ching and Beck proposed a new Bayesian model updating method utilizing the Expectation-Maximization algorithm to determine the most probable values (MPV) of the model parameters [29], and further applied it to the IASC-ASCE benchmark problem [30]. Lam et al. developed a new Bayesian artificial neural network approach for structural health monitoring [31]. Muto and Beck developed a Bayesian updating method for hysteretic structural models using stochastic simulation [32]. Taflanidis and Beck presented an efficient framework for further robust design using stochastic simulation [20]. More investigations from the Bayesian School can be found for selecting the most appropriate models among alternative competing ones [33–37], detecting the possible structural damages [38–41], calibrating and updating the parameters along with their uncertainties [42–48] as well as propagating uncertainties to predict important QoI in operation and safety of structural systems [21,49–51].

Significant progress has been achieved in the past two decades based on the mainstream Bayesian approach. However, some challenges still remain unresolved. For example, the deficiency of the Bayesian framework proposed in [22] is the underestimation of uncertainties due to the inherent reduction of the parameter uncertainty as the number of data points increases. This reduction in uncertainty caused, for example, when higher sampling rate is used in measured time histories, is not justified when the information contained in the inherent dynamics is already captured by lower sampling rates. More importantly, this framework focus on epistemic uncertainty, ignoring the uncertainty due to variability in experimental data, environmental conditions, material properties and manufacturing process. Such uncertainty is basically performed by test-to-test variability in reality, which results in different identified values of the model parameters from dataset to dataset. To properly account for such variability in the modeling of a structure one needs to embed uncertainties within the model by introducing a hierarchy in the model parameters.

This advanced the development of hierarchical Bayesian modeling (HBM) framework in recent years. This framework develops the next generation of Bayesian methodologies and opens up new horizons in data-driven uncertainty quantification offering reasonable and realistic uncertainties. It was first developed by mathematicians to estimate the unknown parameters at higher levels [52]. Subsequently, it has been applied to several scientific disciplines, such as molecular dynamics [53–55], mechanical models [56,57], fatigue predictions [58,59] as well as structural dynamics [60]. Particularly in the field of structural dynamics, it was first applied by Ballesteros et al. [60] to infer the parameters of a structural model introduced to represent a population of identically manufactured structural elements, using modal properties as experimental data. In civil engineering, the HBM was first introduced by Behmanesh et al. for structural identification using a simulated case where the effects of modal data incompleteness and model error were investigated in the framework [61]. It is clearly demonstrated in this work that the uncertainty of the model parameters is underestimated and incorporated it into the prediction error function when using the classical Bayesian strategy, while the HBM framework can properly consider the parameters uncertainties and the uncertainty from model error. The HBM framework is then applied to a real footbridge to account for environmental variability, model error as well as parameter estimation uncertainties [62]. Later on, Song et al. presented the framework to a numerical 10-story building model, demonstrating the effect of significant model error [63]. The framework was also implemented to a two-story reinforced concrete building by accounting for the variation in the amplitude of excitation [64]. A review of the applications of this

framework can be found in literature [65]. Additionally, Patsialis et al. applied the HBM framework for reduced order structural models in earthquake engineering [66] and Akhlaghi et al. developed the framework for site characterization and site response using surface wave measurements and H/V spectral data [67]. The above-mentioned efforts have made significant progress on the development of HBM framework. However, the computational burden is substantially higher than the conventional Bayesian framework, since a relatively large number of model runs is required. For improving the computational cost of this framework, Sedehi et al. recently proposed new asymptotic approximations into the framework for time-domain model updating [68,69] and operational modal analysis [70]. Such improvements analytically show the HBM framework can characterize the ensemble variability of structural parameters observed over multiple datasets together with the identification uncertainty obtained based on the discrepancy between the measured and model outputs within a dataset. It is shown that the HBM framework embeds a substantial amount of uncertainty within the structural parameters rather than the prediction error parameters. Therefore, the response QoI can be predicted considering only the posterior uncertainty of structural parameters. This feature is activated through an explicit marginalization of the prediction error variance from the joint posterior distribution. It is helpful for the response predictions when there is no information about the prediction error variance.

Even though the source of the uncertainty underestimation is well-understood within the HBM framework, the developments and applications of this powerful tool in structural dynamics are still in early stage. Therefore, there is still plenty of space to be explored for improvements. For example, the asymptotic approximations developed in [68,69] are mainly based on time histories responses where a large number of data points are available, the developments relied on modal properties are still lacking. Also, the prediction error term is treated as a nuisance variable which poses difficulties for propagating the overall uncertainties to the observed QoI. Additionally, the framework has been successfully applied to linear model updating. However, most physical structures are inherently characterized by nonlinear behaviours with higher uncertainties due to material and geometric nonlinearities, and the implementation of HBM on updating nonlinear models needs to be explored. More importantly, a non-hierarchical model structure for the underlying system was employed in the HBM framework, however, physical structure often consists of complex models and nonlinearities at all hierarchy levels such as components, sub-systems as well as system levels. Developing a HBM technique for such complex systems and deciding on how to

allocate test resources to get the most out of the resulting data to validate the predictive accuracy of the models are unresolved and challenging issues. Moreover, given only the PDFs or statistical characteristics for measurements, the uncertainty due to variability in the model parameters needs to be considered, yet this issue has not been solved within the HBM framework. Furthermore, although the HBM framework has been introduced into the fatigue prediction problem, this area is still open for new contributions to reasonably quantify the uncertain parameters and accurately predict the uncertainty bounds of $S-N$ curves.

1.2 Research Objectives

The overall goal of the research work presented in this thesis is to develop a comprehensive Bayesian probabilistic framework for uncertainty quantification and propagation (UQP) in engineering simulations on the basis of complex physics-based models of dynamical systems, and measurements collected during system operation. The framework should account for uncertainties arising from model and measurement errors, as well as uncertainties arising from experimental, operational, environmental and manufacturing variabilities. Individual objectives addressed in this thesis to accomplish the overall goal of the research are as follows.

- Develop a framework integrating hierarchical modeling techniques and physics-based models with information provided from multiple datasets
- Provide a more realistic description of parameter uncertainties as well as properly accounting for measurements uncertainties under different environment and operational conditions during operation of a structure, as well as manufacturing variabilities
- Address the theoretical and computational issues in systems involving various types of modeling complexities such as nonlinearities from hysteresis, damping, stiffness, etc
- Extend the framework of a full-scale structure and further develop a new UQP framework for multi-level modeling of a structure considering different levels of models and data hierarchies
- Gain more insightful expressions for understanding the sources of uncertainties within a multi-level modeling approach

- Resolve the uncertainty underestimation issue due to redundant information in the response data consisting usually of long time history measurements
- Integrate into the framework the case where only statistical information is available for the measurements
- Demonstrate and validate the framework using data from structural dynamics and fatigue predictions.

1.3 Contributions and Organization of the Thesis

This dissertation contributes to the development of the next generation of Bayesian modeling framework for tackling the challenging UQP issues in structural dynamic simulations. This framework provides a comprehensive computation tool for reasonably quantifying the uncertainties, accurately predicting the output quantities, as well as efficiently updating the safety and system reliability. Efforts made on the framework have led to four papers submitted for review and possible publications [71–74], one paper to be submitted to international scientific journals [75], and a number of peer-reviewed conference papers [76–79]. Five journal papers in total, devoted to the objectives, are listed as the individual chapters in this thesis, and summarized in Table 1-1. The specific topics and the corresponding contributions of these chapters are described as follows. A summary of contributions made in this dissertation and the research directions in the future are also given in the last Chapter.

Table 1-1 Thesis chapters

Chapter	Title
2	Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Features
3	Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework
4	Hierarchical Bayesian Learning Framework for Multi-Level Modeling using Multi-level Data
5	Data Features-based Bayesian Learning for Time-domain Model Updating and Robust Predictions in Structural Dynamics
6	Statistics-based Bayesian Modeling Framework for Uncertainty Quantification and Propagation

Chapter 2. Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Features

(<https://doi.org/10.5281/zenodo.5078051>, Submitted for review and possible publication to Mechanical Systems and Signal Processing (MSSP))

This chapter focus on the development of Bayesian learning tools for model calibration and response predictions in linear dynamical systems using multiple datasets consisting of modal properties identified from measured response time histories. Uncertainties are embedded into the structural model parameters by introducing a hierarchical modeling level for which the model parameters are assigned a parameterized prior distribution with the hyperparameters to be inferred from the multiple datasets. The unmodelled dynamics are quantified by introducing a zero-mean Gaussian prediction error term to account for the difference between the measured and model predicted modal properties. The proposed HBM framework can properly account for the variation in the parameter estimates due to experimental, operational, environmental and manufacturing variabilities. Two asymptotic approximations, valid for large number of measured data within each dataset and differing in the derivations of Taylor expansion in terms of the prediction error variance, are integrated into the proposed framework to reduce the high computational cost associated with full sampling approaches for quantifying and propagating uncertainties. It provides reasonable uncertainty bounds that are irreducible as the number of modal data and the number of datasets increases. In contrast, conventional Bayesian inference techniques result in unrealistically small parameter estimation or identification uncertainties as the number of data increased, failing to provide a reasonable account of uncertainty bounds. Uncertainties of model parameters and prediction error, inferred from modal properties data, are considered for predictions of modal frequencies, mode shapes as well as time histories response. Reliability analysis is also performed within the framework considering both the parameter uncertainties and input uncertainties.

Main contributions of this chapter are as follows.

- Develop a hierarchical Bayesian modeling framework for UQP based on modal features
- Handle uncertainty due to variability over multiple datasets
- Reduce the high computational effort involved in conventional sampling approaches required in the framework by introducing accurate asymptotic approximations

- Exploit the asymptotic results so as to provide insightful information on uncertainty management
- Offer robust data-informed predictions.

Chapter 3. Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework (<https://doi.org/10.5281/zenodo.5520607>, Submitted for review and possible publication to Computer Methods in Applied Mechanics and Engineering (CMAME))

This chapter extends the HBM framework outlined in Chapter 3, developing it further for updating physics-based models with nonlinearities using measured response time histories measurements. Following the framework in Chapter 2, an extension involves embedding uncertainties not only to structural model parameters but also to prediction error model parameters, introducing respectively a parameterized Gaussian probabilistic distribution and an inverse Gamma distribution to models the structural and prediction error parameters, with the hyperparameters inferred by the available datasets. Based on asymptotic approximations developed, analytical and insightful expressions of the posterior distribution of hyperparameters are derived, given sufficient data points within a dataset. Examples conducted on structural dynamics benchmark models with nonlinearities based on a nonlinear Bouc-Wen model demonstrate that one dominant source of the parameter uncertainty is due to the variability that arise from model error. The uncertainty in the response predictions are explored by propagating the parameter uncertainties and the uncertainties from model error through the nonlinear model.

Main contributions of this chapter are as follows.

- Develop a HBM framework for nonlinear model updating
- Tackle the uncertainty due to variability that arise from model error
- Investigate the effect of sensor locations/numbers on parameter uncertainty
- Simplify the HBM process of nonlinear model updating using asymptotic approximations
- Validate the HBM framework using hysteresis type nonlinearities.

Chapter 4. Hierarchical Bayesian Learning Framework for Multi-Level Modeling using Multi-level Data (<https://doi.org/10.5281/zenodo.5702385>, Submitted for review and possible publication to MSSP)

This chapter contains the core methodological contribution of the Thesis. It exploits the developments in Chapters 2 and 3 to address the challenging issues of validating a hierarchy of models arising when assembling a system from subsystems and components and the tests are carried out at different levels of model hierarchy. It advances the HBM framework proposed in Chapter 2 and generalizes it further for a multi-level modeling approach followed for a structural system. Parameterized physics-based models are introduced at different levels of hierarchy, uncertainties are embedded into the model parameters by introducing parameterized prior distributions of model parameters, and the hyperparameters in these distributions are informed from the datasets available from the multiple tests performed at different levels of model hierarchy (component, subsystem and system levels). Novel computing techniques based on asymptotic approximations are developed for parameter inference and uncertainty propagation in the multi-level modeling approach.

Main contributions of this chapter are as follows.

- Develop a comprehensive hierarchical Bayesian learning framework for model calibration and robust predictions in a multi-level modeling approach
- Capture the parameter variability that arises from environmental/operational conditions manufacturing process, as well as assembling process
- Integrate the data hierarchies to account for parameter variability arisen from test-to-test
- Reduce the computation burden of the framework by introducing accurate asymptotic approximations
- Accurately propagate the uncertainties through the modeling hierarchy levels for making robust predictions at the system level
- Exploit the asymptotic approximations to develop analytical and insightful expressions for the uncertainty in the hyperparameters, model parameters and output QoI
- Demonstrate the accuracy and computational efficiency of the whole framework with an application on simplified hierarchical model structure encountered in structural dynamics.

Chapter 5. Data Features-based Bayesian Learning for Time-domain Model Updating and Robust Predictions in Structural Dynamics

This chapter points out an issue of parameter uncertainty underestimation inferred that arises from the application of conventional Bayesian inference framework when single dataset of response time histories is used to infer the model parameters. It promotes solutions to this problem by proposing new likelihood formulations within the classical Bayesian framework. For this, three probabilistic prediction models are introduced based on the features between measurements and model predicted outputs. Formulations are derived which reveal the relation between the proposed likelihood-informed Bayesian framework and established likelihood-free approximate Bayesian computation (ABC) framework. Spatially and temporally correlation is also considered for the proposed prediction model within the framework. Dynamical example is used to demonstrate the effectiveness of the proposed framework.

Main contributions of this chapter are as follows.

- Propose data-features likelihood-based Bayesian methodology which can reasonably account for the parameter uncertainty making it independent of the sampling rate of the measured time histories response for large enough sampling rates
- Reveal the relation between the proposed likelihood-informed Bayesian framework and likelihood-free ABC framework
- Provide reasonable uncertainty bounds of QoI, consistent with the ones obtained from ABC formulation
- Ensure the correlation case of prediction error where the uncertainty is also independent of different sampling rates of the time histories response.

Chapter 6. Statistics-based Bayesian Modeling Framework for Uncertainty Quantification and Propagation (<https://doi.org/10.5281/zenodo.5545922>, Submitted for review and possible publication to MSSP)

Motivated by the idea of HBM framework proposed in Chapters 2 and 3, this chapter develops further a statistics-based Bayesian inference strategy to account for the parameter variability due to model error, measurement noise as well as environmental, operational and manufacturing variabilities. The proposed method formulates the likelihood function based on the Kullback Leibler divergence (KL-div) used to quantify the discrepancy between the PDFs of the model predictions and measurements. Introducing such formulation is beneficial especially for the case where the full measurements are not available and only the statistics or PDFs are provided. For output QoI that depend linearly on the inferred model parameters, the

posterior distribution of the hyperparameters is analytically derived with respect to the first two moments of the measured model predicted PDFs. For nonlinear dependence of the output QoI on the inferred model parameters, the analytical expressions developed for the linear case offer a convenient framework to approximate the posterior PDF of the hyperparameters. Numerical methods to estimate the resulting multidimensional integrals for the lowest two moments of the model predicted PDF are explored. Examples from structural dynamics and fatigue *S-N* curves are employed to verify the effectiveness of the proposed approach.

Main contributions of this chapter are as follows.

- Propose an alternative HBM framework to quantify the parameter variability that is especially convenient to be used when only the statistics or PDFs of the measurements are available
- Use KL-divergence to quantify discrepancy between model prediction and measurement PDFs
- Develop analytical posterior distribution of model parameters based on lower two moments of PDFs
- Provide an alternative way to identify the hyper parameters in HBM proposed
- Demonstrate framework for structural dynamics and S-N fatigue curve applications

1.4 References

- [1] S.A. Billings, Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains, John Wiley & Sons, 2013.
- [2] O.C. Zienkiewicz, R.L. Taylor, P. Nithiarasu, J.Z. Zhu, The finite element method, McGraw-hill London, 1977.
- [3] O.C. Zienkiewicz, R.L. Taylor, The finite element method for solid and structural mechanics, Elsevier, 2005.
- [4] T.J.R. Hughes, The finite element method: linear static and dynamic finite element analysis, Courier Corporation, 2012.
- [5] J. Guggenberger, H. Grundmann, Stochastic response of large FEM models with hysteretic behaviour in beam elements, Computer Methods in Applied Mechanics and Engineering. 194 (2005) 1739–1756.

- [6] G. Kerschen, K. Worden, A.F. Vakakis, J.C. Golinval, Past, present and future of nonlinear system identification in structural dynamics, *Mechanical Systems and Signal Processing*. 20 (2006) 505–592.
- [7] V. Papadopoulos, M. Papadrakakis, G. Deodatis, Analysis of mean and mean square response of general linear stochastic finite element systems, *Computer Methods in Applied Mechanics and Engineering*. 195 (2006) 5454–5471.
- [8] Z. Lai, *Sparse Structural System Identification and Damage Detection*, Rice University, 2018.
- [9] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, *Mechanical Systems and Signal Processing*. 56 (2015) 123–149.
- [10] J.E. Mottershead, M.I. Friswell, Model updating in structural dynamics: a survey, *Journal of Sound and Vibration*. 167 (1993) 347–375.
- [11] J. Mottershead, M. Friswell, *Finite Element Model Updating in Structural Dynamics*, (1995).
- [12] A. Teughels, G. De Roeck, Damage detection and parameter identification by finite element model updating, *Revue Européenne de Génie Civil*. 9 (2005) 109–158.
- [13] M. Friswell, J.E. Mottershead, *Finite element model updating in structural dynamics*, Springer Science & Business Media, 2013.
- [14] J.E. Mottershead, M. Link, M.I. Friswell, The sensitivity method in finite element model updating: A tutorial, *Mechanical Systems and Signal Processing*. 25 (2011) 2275–2296.
- [15] S.-K. Au, *Operational Modal Analysis*, 2017.
- [16] E. Reynders, System Identification Methods for (Operational) Modal Analysis: Review and Comparison, *Archives of Computational Methods in Engineering*. 19 (2012) 51–124.
- [17] K. Worden, C.R. Farrar, G. Manson, G. Park, The fundamental axioms of structural health monitoring, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 463 (2007) 1639–1664.
- [18] W.E. Walker, P. Harremoës, J. Rotmans, J.P. van der Sluijs, M.B.A. van Asselt, P. Janssen, M.P. Krayen von Krauss, Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support, *Integrated Assessment*. 4 (2003) 5–17.

- [19] A. Der Kiureghian, O. Ditlevsen, Aleatory or epistemic? Does it matter?, *Structural Safety*. 31 (2009) 105–112.
- [20] A.A. Taflanidis, J.L. Beck, An efficient framework for optimal robust stochastic system design using stochastic simulation, *Computer Methods in Applied Mechanics and Engineering*. 198 (2008) 88–101.
- [21] J.L. Beck, A.A. Taflanidis, Prior and Posterior Robust Stochastic Predictions for Dynamical Systems Using Probability Logic, *International Journal for Uncertainty Quantification*. 3 (2013) 271–288.
- [22] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, *Journal of Engineering Mechanics*. 124 (1998) 455–461.
- [23] J.L. Beck, Bayesian system identification based on probability logic, *Structural Control and Health Monitoring*. 17 (2010) 825–847.
- [24] L.S. Katafygiotis, J.L. Beck, Updating Models and Their Uncertainties. II: Model Identifiability, *Journal of Engineering Mechanics*. 124 (1998) 463–467.
- [25] M.W. Vanik, J.L. Beck, S.-K. Au, Bayesian probabilistic approach to structural health monitoring, *Journal of Engineering Mechanics*. 126 (2000) 738–745.
- [26] C. Papadimitriou, J.L. Beck, L.S. Katafygiotis, Updating robust reliability using structural test data, *Probabilistic Engineering Mechanics*. 16 (2001).
- [27] J.L. Beck, S.-K. Au, Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation, *Journal of Engineering Mechanics*. 128 (2002) 380–391.
- [28] K.-V. Yuen, S.K. Au, J.L. Beck, Two-Stage Structural Health Monitoring Approach for Phase I Benchmark Studies, *Journal of Engineering Mechanics*. 130 (2004) 16–33.
- [29] J. Ching, J.L. Beck, New Bayesian model updating algorithm applied to a structural health monitoring benchmark, *Structural Health Monitoring*. 3 (2004) 313–332.
- [30] J. Ching, J.L. Beck, Bayesian Analysis of the Phase II IASC–ASCE Structural Health Monitoring Experimental Benchmark Data, *Journal of Engineering Mechanics*. 130 (2004) 1233–1244.
- [31] H.F. Lam, K.V. Yuen, J.L. Beck, Structural health monitoring via measured ritz vectors utilizing artificial neural networks, *Computer-Aided Civil and Infrastructure Engineering*. 21 (2006) 232–241.
- [32] M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, *JVC/Journal of Vibration and Control*. 14 (2008) 7–34.

- [33] J.L. Beck, K.V. Yuen, Model Selection Using Response Measurements: Bayesian Probabilistic Approach, *Journal of Engineering Mechanics*. 130. No. 2 (2004) 192–203.
- [34] K.V. Yuen, Recent developments of Bayesian model class selection and applications in civil engineering, *Structural Safety*. 32 (2010) 338–346.
- [35] K.V. Yuen, *Bayesian Methods for Structural Dynamics and Civil Engineering*, 2010.
- [36] A. Ben Abdesslem, N. Dervilis, D. Wagg, K. Worden, Model selection and parameter estimation in structural dynamics using approximate Bayesian computation, *Mechanical Systems and Signal Processing*. 99 (2018) 306–325.
- [37] A. Ben Abdesslem, N. Dervilis, D. Wagg, K. Worden, Model selection and parameter estimation of dynamical systems using a novel variant of approximate Bayesian computation, *Mechanical Systems and Signal Processing*. 122 (2019) 364–386.
- [38] K.V. Yuen, J.L. Beck, L.S. Katafyglotis, Unified probabilistic approach for model updating and damage detection, *Journal of Applied Mechanics, Transactions ASME*. 73 (2006) 555–564.
- [39] E. Ntotsios, C. Papadimitriou, P. Panetsos, G. Karaiskos, K. Perros, P.C. Perdikaris, Bridge health monitoring system based on vibration measurements, *Bulletin of Earthquake Engineering*. 7 (2009).
- [40] T. Yin, H.F. Lam, H.M. Chow, A Bayesian probabilistic approach for crack characterization in plate structures, *Computer-Aided Civil and Infrastructure Engineering*. 25 (2010) 375–386.
- [41] Y. Huang, J.L. Beck, H. Li, Hierarchical sparse Bayesian learning for structural damage detection: Theory, computation and application, *Structural Safety*. 64 (2017) 37–53.
- [42] K.V. Yuen, S.C. Kuok, Bayesian methods for updating dynamic models, *Applied Mechanics Reviews*. 64 (2011).
- [43] W.J. Yan, L.S. Katafygiotis, A novel Bayesian approach for structural model updating utilizing statistical modal information from multiple setups, *Structural Safety*. 52 (2015) 260–271.
- [44] W.J. Yan, D. Chronopoulos, C. Papadimitriou, S. Cantero-Chinchilla, G.S. Zhu, Bayesian inference for damage identification based on analytical probabilistic model of scattering coefficient estimators and ultrafast wave scattering simulation scheme, *Journal of Sound and Vibration*. 468 (2020) 115083.

- [45] Y.-C. Zhu, S.-K. Au, Bayesian data driven model for uncertain modal properties identified from operational modal analysis, *Mechanical Systems and Signal Processing*. 136 (2020) 106511.
- [46] A. Kamariotis, E. Chatzi, D. Straub, Value of information from vibration-based structural health monitoring extracted via Bayesian model updating, 166 (2021).
- [47] F. Schneider, I. Papaioannou, D. Straub, C. Winter, G. Müller, Bayesian parameter updating in linear structural dynamics with frequency transformed data using rational surrogate models, 166 (2022).
- [48] M. Song, L. Renson, B. Moaveni, G. Kerschen, Bayesian model updating and class selection of a wing-engine structure with nonlinear connections using nonlinear normal modes, *Mechanical Systems and Signal Processing*. 165 (2022).
- [49] A.A. Taflanidis, J.L. Beck, Life-cycle cost optimal design of passive dissipative devices, *Structural Safety*. 31 (2009) 508–522.
- [50] P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian uncertainty quantification and propagation in molecular dynamics simulations, *ECCOMAS 2012 - European Congress on Computational Methods in Applied Sciences and Engineering, e-Book Full Papers*. 144103 (2012) 3595–3608.
- [51] P.E. Hadjidoukas, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, $\Pi 4U$: A high performance computing framework for Bayesian uncertainty quantification of complex models, *Journal of Computational Physics*. 284 (2015).
- [52] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, Aki Vehtari, D.B. Rubin, *Bayesian data analysis*, Chapman & Hall/CRC, 2013.
- [53] S. Wu, P. Angelikopoulos, C. Papadimitriou, R. Moser, P. Koumoutsakos, A hierarchical Bayesian framework for force field selection in molecular dynamics simulations, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 374 (2016).
- [54] S. Wu, P. Angelikopoulos, G. Tauriello, C. Papadimitriou, P. Koumoutsakos, Fusing heterogeneous data for the calibration of molecular dynamics force fields using hierarchical Bayesian models, *Journal of Chemical Physics*. 145 (2016).
- [55] A. Economides, G. Arampatzis, D. Alexeev, S. Litvinov, L. Amoudruz, L. Kulakova, C. Papadimitriou, P. Koumoutsakos, Hierarchical Bayesian Uncertainty Quantification for a Model of the Red Blood Cell, *Physical Review Applied*. 15 (2021) 34062.

- [56] J.B. Nagel, N. Mojsilovic, B. Sudret, Bayesian assessment of the compressive strength of structural masonry, 12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP 2015. (2015).
- [57] J.B. Nagel, B. Sudret, A unified framework for multilevel uncertainty quantification in Bayesian inverse problems, *Probabilistic Engineering Mechanics*. 43 (2016) 68–84.
- [58] X.W. Liu, D.G. Lu, P.C.J. Hoogenboom, Hierarchical Bayesian fatigue data analysis, *International Journal of Fatigue*. 100 (2017) 418–428.
- [59] X.W. Liu, D.G. Lu, Survival analysis of fatigue data: Application of generalized linear models and hierarchical Bayesian model, *International Journal of Fatigue*. 117 (2018) 39–46.
- [60] G.C. Ballesteros, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian hierarchical models for uncertainty quantification in structural dynamics, in: *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, 2014: pp. 1615–1624.
- [61] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, *Mechanical Systems and Signal Processing*. 64–65 (2015) 360–376.
- [62] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, *Journal of Sound and Vibration*. 374 (2016) 92–110.
- [63] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Modeling error estimation and response prediction of a 10-story building model through a hierarchical bayesian model updating framework, *Frontiers in Built Environment*. 5 (2019).
- [64] M. Song, B. Moaveni, C. Papadimitriou, A. Stavridis, Accounting for amplitude of excitation in model updating through a hierarchical Bayesian approach: Application to a two-story reinforced concrete building, *Mechanical Systems and Signal Processing*. 123 (2019) 68–83.
- [65] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Accounting for Modeling Errors and Inherent Structural Variability through a Hierarchical Bayesian Model Updating Approach: An Overview, *Sensors*. 20 (2020) 3874.
- [66] D. Patsialis, A.P. Kyprioti, A.A. Taflanidis, Bayesian calibration of hysteretic reduced order structural models for earthquake engineering applications, *Engineering Structures*. 224 (2020) 111204.

- [67] M.M. Akhlaghi, M. Song, M. Pontrelli, B. Moaveni, L.G. Baise, Site Characterization Through Hierarchical Bayesian Model Updating Using Dispersion and H/V Data, in: *Model Validation and Uncertainty Quantification, Volume 3*, Springer, 2020: pp. 333–335.
- [68] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Probabilistic hierarchical Bayesian framework for time-domain model updating and robust predictions, *Mechanical Systems and Signal Processing*. 123 (2019) 648–673.
- [69] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Data-driven uncertainty quantification and propagation in structural dynamics through a hierarchical Bayesian framework, *Probabilistic Engineering Mechanics*. 60 (2020).
- [70] O. Sedehi, L.S. Katafygiotis, C. Papadimitriou, Hierarchical Bayesian operational modal analysis: Theory and computations, *Mechanical Systems and Signal Processing*. 140 (2020).
- [71] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Features, Submitted for review and possible publication to MSSP. (2021). <https://zenodo.org/record/5078051>.
- [72] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework, Submitted for review and possible publication to CMAME. (2021). <https://doi.org/10.5281/zenodo.5520607>.
- [73] X. Jia, C. Papadimitriou, Hierarchical Bayesian Learning Framework for Multi-Level Modeling using Multi-level Data, Submitted for review and possible publication to MSSP. (2021). <https://doi.org/10.5281/zenodo.5702385>
- [74] M. Ping, X. Jia, C. Papadimitriou, X. Han, C. Jiang, Statistics-based Bayesian modeling framework for uncertainty quantification and propagation, Submitted for review and possible publication to MSSP. (2021). <https://doi.org/10.5281/zenodo.5545922>.
- [75] X. Jia, C. Papadimitriou, Data Features-based Bayesian Learning for Time-domain Model Updating and Robust Predictions in Structural Dynamics, In Preparation to Submit. (2021).
- [76] X. Jia, C. Papadimitriou, Data features-based likelihood-informed Bayesian finite element model updating, in: *Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering*, 2019: pp. 103–113. <https://doi.org/10.7712/120219.6328.18902>.

- [77] X. Jia, O. Sedehi, C. Papadimitriou, L. Katafygiotis, B. Moaveni, Two-Stage Hierarchical Bayesian Framework for Finite Element Model Updating, in: *Model Validation and Uncertainty Quantification*, Springer, Cham., 2020: pp. 383–387. https://doi.org/10.1007/978-3-030-47638-0_42
- [78] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Computationally efficient hierarchical Bayesian modeling framework for learning embedded model uncertainties, *Proceedings of the International Conference on Structural Dynamic , EURODYN. 2* (2020) 3886 – 3893. <https://doi.org/10.47964/1120.9317.19651>
- [79] X. Jia, O. Sedehi, L.S. Katafygiotis, B. Moaveni, C. Papadimitriou, Hierarchical Bayesian Model Updating for Nonlinear Structures Using Response Time Histories, *3* (2021). https://doi.org/10.1007/978-3-030-77348-9_14.

Chapter 2. Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Features

Original Paper:

X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Features, Mechanical Systems and Signal Processing. Submitted. (2021). <https://zenodo.org/record/5078051>

ABSTRACT

The hierarchical Bayesian modeling (HBM) framework has recently been developed to tackle the uncertainty quantification and propagation in structural dynamics inverse problems. This new framework characterizes the ensemble variability of structural parameters observed over multiple datasets together with the identification uncertainty obtained based on the discrepancy between the measured and model outputs. The present paper expands on this framework, developing it further for model inference based on modal features. It generalizes the HBM framework by considering an additional hyper distribution to characterize the uncertainty of prediction error variances across different datasets. Moreover, computationally efficient approximations are developed to simplify the computation of the posterior distribution of hyper-parameters. Conditions are presented under which the approximations are expected to be accurate. The asymptotic approximations provide insightful information on the relation of the estimates of the hyper-parameters and their uncertainties with the variability of the estimations and identification uncertainties. Introducing the HBM formulation is beneficial, particularly for the propagation of uncertainty based on both structural and prediction error parameters providing reasonable uncertainty bounds. The posterior uncertainty of the structural and prediction error parameters is propagated to

estimate data-informed output quantities of interests, including failure probabilities, which offers robustness to the variability over datasets. The proposed approximations are tested and verified using simulated and experimental examples. The effects of the uncertainty due to dataset variability and the prediction error uncertainty are illustrated in these examples.

2.1 Introduction

Updating models and predicting responses using data-driven approaches has been substantially investigated in structural dynamics using deterministic [1] and probabilistic approaches [2-5]. Due to a rigorous probabilistic framework, Bayesian tools can rationally integrate data and physics-based models in order to select the most appropriate models among alternative competing ones [6-8], estimate the parameters of these models and their uncertainties [5,9,10], as well as propagating uncertainties to predict important quantities of interest (QoI) in operation and safety of structural systems [11-14]. Undoubtedly, the Bayesian framework offers an indispensable and powerful mathematical tool for quantifying and propagating uncertainties in simulations. However, challenges still remain. For example, owing to the redundant information carried in the data, the conventional Bayesian approach often underestimates the uncertainty, resulting in an inherent reduction of the parameter uncertainty as the number of data increases [2]. Furthermore, standard Bayesian procedures do not properly take into account the uncertainty in the parameters attributed to the variability in experimental data, environmental conditions, material properties, manufacturing process, assembling process, and nonlinear mechanisms activated under different loading conditions [15-19].

In light of above, a comprehensive hierarchical Bayesian modeling (HBM) framework has been further developed in various scientific disciplines [20-24] to properly quantify the uncertainties within the model parameters. Specifically in the field of structural dynamics, the HBM approach was used to quantify uncertainties due to environmental variabilities [25], as well as amplitude of excitation variabilities [26]. Sampling-based techniques, for instance Markov chain Monte Carlo (MCMC) and Gibbs sampler [27], are employed within the HBM framework to update models, calibrate uncertainties, and propagate the uncertainties to response QoI in structural dynamics using multiple modal datasets [20,25,28]. A review of the applications of this framework can be found in [29].

Recently, for improving the computational cost of this framework, new asymptotic approximations have been proposed for time-domain model updating [30,31] and operational

modal analysis [32]. It is shown that the HBM framework embeds a substantial amount of uncertainty within the structural parameters rather than the prediction error parameters. Thus, the response QoI can be predicted considering only the posterior uncertainty of structural parameters [30]. This feature is activated through an explicit marginalization of the prediction error variance from the joint posterior distribution. It is helpful for the response predictions when there is no information about the prediction error variance. However, predicting unobserved quantities would require subjective assumptions about the prediction error variance.

Motivated by the above considerations, this paper proposes a new HBM formulation for model updating and robust predictions based on modal test data. It treats the prediction error variance as a hyper-parameter updated based on the data. It also proposes two novel asymptotic approximations to reduce the computational cost. When this framework is used for uncertainty propagation, parameter and prediction error uncertainties are both incorporated automatically. An analytical formulation is developed to provide insights as to the interpretation of uncertainties in the HBM framework. Finally, the applications of this framework are illustrated using both numerical and experimental examples.

This study is structured as follows. Section 2.2 describes the proposed HBM formulation for model updating and prediction, including the asymptotic approximations. Section 2.3 presents the computational algorithm of the proposed approach. In Sections 2.4, a simulated example with four different cases and one experimental example are employed to verify the effectiveness of the proposed formulations. Section 2.5 reports the conclusions of this study.

2.2 Proposed HBM Framework

2.2.1 Probabilistic quantification and propagation model

Let $\mathbf{D} = \{D_i, i = 1, 2, \dots, N_D\}$ be the measured datasets from the structure comprising N_D independent experimental datasets. Each individual experimental data $D_i = \{\hat{\lambda}_{r,i}, \hat{\phi}_{r,i} \in R^{N_0}, r = 1, 2, \dots, m\}$ consists of the square of the modal frequencies $\hat{\omega}_{r,i}$ in rad/s, $\hat{\lambda}_{r,i} = \hat{\omega}_{r,i}^2$, and the mode shapes $\hat{\phi}_{r,i}$ at $N_{0,i}$ measured degrees of freedoms (DOF) for mode r , where m is the number of the observed modes. For each experimental dataset, the modal properties have been obtained by modal estimation techniques [33-35] using available input-output or output-only vibration measurements. Consider a parameterized class of structural model M that is used to predict the square of the r -th modal frequency $\lambda_r(\boldsymbol{\theta}_i) = \omega_r^2(\boldsymbol{\theta}_i)$ and

the r -th mode shapes $\phi_r(\theta_i)$ for a particular value of the model parameter set $\theta_i \in R^{N_\theta}$ in the i -th dataset, where N_θ is the total number of the parameter vector θ_i . The discrepancy between the i -th predicted modal properties and the i -th measured dataset is quantified based on the prediction error equations, given as follows [28]:

$$\begin{aligned}\hat{\lambda}_{r,i} &= \lambda_r(\theta_i) + e_r, \\ \frac{\hat{\phi}_{r,i}}{\|\hat{\phi}_{r,i}\|} &= \beta_r(\theta_i) \frac{\phi_r(\theta_i)}{\|\phi_r(\theta_i)\|} + \varepsilon_r,\end{aligned}\tag{2.1}$$

where e_r and ε_r are respectively the prediction errors for the square of the modal frequency and the mode shape components of the r -th mode, and $\beta_r(\theta_i) = (\hat{\phi}_{r,i}^T \phi_r(\theta_i)) / (\|\hat{\phi}_{r,i}\| \|\phi_r(\theta_i)\|)$ is a normalization constant that accounts for the different scaling between the measured and the predicted mode shape for given parameter set θ_i . The prediction errors are then modeled by zero-mean Gaussian variables $e_r \sim N(e_r | 0, \tilde{\lambda}_r^2 \sigma^2)$ and $\varepsilon_r \sim N(\varepsilon_r | \mathbf{0}, \Sigma)$, where $\tilde{\lambda}_r$ is a factor that scales the standard deviation $\tilde{\lambda}_r \sigma$ of the error for the r -th square of the modal frequency to be proportional to an expected value $\tilde{\lambda}_r$ of the square of the modal frequency. The modal error factors $\tilde{\lambda}_r$ can be chosen to correspond to the values $\tilde{\lambda}_r = \lambda_r(\theta_{ref})$ predicted by a reference or nominal model of the structure corresponding to parameter values θ_{ref} .

Alternatively, they can be chosen to be $\tilde{\lambda}_r = \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\lambda}_{r,i}$, the average value of the identified modal frequencies over all datasets. The covariance matrix Σ is defined as $\Sigma = \sigma^2 \mathbf{I}$, and the parameter σ^2 is to be estimated using multiple datasets.

A HBM framework is used to infer the values of the model parameters [20,30]. According to the HBM framework, the uncertainties due to model error and environmental, manufacturing and experimental variabilities are embedded into the model parameters θ . Realization of θ varies across the datasets, where θ_i corresponds to the i -th dataset. As a common way, herein we assign the Gaussian prior probability density function (PDF):

$$p(\theta_i) = N(\theta_i | \mu_\theta, \Sigma_\theta)\tag{2.2}$$

for θ_i with hyper mean μ_θ and hyper covariance matrix Σ_θ to be estimated using the multiple datasets. Fig. 2-1 shows the graphical representation of the proposed HBM framework. As seen, the unknown hyper-parameters can be incorporated as a set of

2.2.2 Likelihood function for each dataset

$$p(D_i | \boldsymbol{\theta}_i, \sigma^2) \propto \frac{1}{\sigma^{m(N_0+1)}} \exp \left[-\frac{m(N_0+1)}{2\sigma^2} J(\boldsymbol{\theta}_i) \right] \quad (2.3)$$
$$J(\boldsymbol{\theta}_i) = \frac{1}{N_0 + 1} [J_1(\boldsymbol{\theta}_i) + J_2(\boldsymbol{\theta}_i)] \quad (2.4)$$
$$J_1(\boldsymbol{\theta}_i) = \frac{1}{m} \sum_{r=1}^m \frac{(\hat{\lambda}_{r,i} - \lambda_r(\boldsymbol{\theta}_i))^2}{(\tilde{\lambda}_r)^2} \quad (2.5)$$
$$J_2(\boldsymbol{\theta}_i) = \frac{1}{m} \sum_{r=1}^m \left\| \frac{\hat{\boldsymbol{\phi}}_{r,i}}{\|\hat{\boldsymbol{\phi}}_{r,i}\|} - \beta_{r,i}(\boldsymbol{\theta}_i) \frac{\boldsymbol{\phi}_r(\boldsymbol{\theta}_i)}{\|\boldsymbol{\phi}_r(\boldsymbol{\theta}_i)\|} \right\|^2 \quad (2.6)$$

Institutional Repository - Library & Information Centre - University of Thessaly
21/05/2022 09:13:57 EEST - 137.108.70.13

2.2.2.1 Asymptotic approximation 1 (A-1) for likelihood function

Introducing the function $L(\boldsymbol{\theta}_i, \sigma^2)$ defined as the negative of the logarithm of the likelihood function

$$L(\boldsymbol{\theta}_i, \sigma^2) = -\ln p(D_i | \boldsymbol{\theta}_i, \sigma^2) = \frac{m(N_0 + 1)}{2} \ln \sigma^2 + \frac{m(N_0 + 1)}{2\sigma^2} J(\boldsymbol{\theta}_i) \quad (2.7)$$

and using Taylor expansion, the likelihood function can be approximated by a normal distribution (valid for larger number of data):

$$p(D_i | \boldsymbol{\theta}_i, \sigma^2) = \exp(-L(\boldsymbol{\theta}_i, \sigma^2)) \propto \exp \left[-\frac{1}{2} \left[(\boldsymbol{\theta}_i^T, \sigma^2) - (\hat{\boldsymbol{\theta}}_i^T, \hat{\sigma}_i^2) \right] \mathbf{H}(\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2) \left[\begin{pmatrix} \boldsymbol{\theta}_i \\ \sigma^2 \end{pmatrix} - \begin{pmatrix} \hat{\boldsymbol{\theta}}_i \\ \hat{\sigma}_i^2 \end{pmatrix} \right] \right] \quad (2.8)$$

with the mean equal to the most probable values (MPV) of the parameters $\hat{\boldsymbol{\theta}}_i$ and $\hat{\sigma}_i^2$ obtained by minimizing the objective function $L(\boldsymbol{\theta}_i, \sigma^2)$, and covariance matrix $\hat{\Sigma}_i(\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2) = \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2)$, where $\mathbf{H}(\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2)$ is the Hessian matrix of the function $L(\boldsymbol{\theta}_i, \sigma^2)$ with respect to the parameters $\{\boldsymbol{\theta}_i, \sigma^2\}$, evaluated at the MPV $\{\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2\}$. Using Eq. (2.8), the optimal values $\hat{\boldsymbol{\theta}}_i$ and $\hat{\sigma}_i^2$ are given by:

$$\hat{\boldsymbol{\theta}}_i = \text{Arg min}_{\boldsymbol{\theta}_i} (J(\boldsymbol{\theta}_i)) \quad (2.9)$$

$$\hat{\sigma}_i^2 = J(\hat{\boldsymbol{\theta}}_i) \quad (2.10)$$

Basically, the MPV can be directly computed using an optimization tool, while the hessian matrix can be computed numerically or analytically [36,37]. The Hessian matrix of $L(\boldsymbol{\theta}_i, \sigma^2)$ evaluated at $\hat{\boldsymbol{\theta}}_i$ and $\hat{\sigma}_i^2$ is readily obtained in the form

$$\mathbf{H}(\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2) = \frac{m(N_0 + 1)}{2} \begin{bmatrix} \frac{1}{\hat{\sigma}_i^2} \mathbf{H}_J(\hat{\boldsymbol{\theta}}_i) & \mathbf{0} \\ \mathbf{0} & \frac{1}{\hat{\sigma}_i^4} \end{bmatrix} \quad (2.11)$$

where $\mathbf{H}_J(\hat{\boldsymbol{\theta}}_i) = \left. \frac{\partial^2 J(\boldsymbol{\theta}_i)}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_i^T} \right|_{\boldsymbol{\theta}_i = \hat{\boldsymbol{\theta}}_i}$ is the Hessian matrix of the averaged measure of fit $J(\boldsymbol{\theta}_i)$. One

can directly obtain the MPV $\hat{\boldsymbol{\theta}}_i$ by minimizing the function $J(\boldsymbol{\theta}_i)$ with respect to the model parameters $\boldsymbol{\theta}_i$, independently of the value of the parameter σ^2 . Once the MPV of the model parameter set $\hat{\boldsymbol{\theta}}_i$ is available, the optimal prediction error parameter $\hat{\sigma}_i^2$ can then be

calculated through Eq. (2.10) as the measure of fit function (2.4), evaluated at the MPV $\hat{\boldsymbol{\theta}}_i$ of the model parameters. The Hessian matrix $\mathbf{H}(\hat{\boldsymbol{\theta}}_i, \hat{\sigma}_i^2)$ is a block diagonal matrix which results in uncorrelated $\boldsymbol{\theta}_i$ and σ^2 . It can be estimated from the hessian matrix $\mathbf{H}_J(\hat{\boldsymbol{\theta}}_i)$ of the measure-of-fit $J(\boldsymbol{\theta}_i)$ evaluated at $\hat{\boldsymbol{\theta}}_i$. The block diagonal hessian matrix in Eq. (2.8) allows simplifying the likelihood $p(D_i | \boldsymbol{\theta}_i, \sigma^2)$ in the form:

$$p(D_i | \boldsymbol{\theta}_i, \sigma^2) \propto N(\boldsymbol{\theta}_i | \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}) N(\sigma^2 | \hat{\sigma}_i^2, \hat{\Sigma}_{\sigma^2}) \quad (2.12)$$

where

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i} = \frac{2}{m(N_0 + 1)} J(\hat{\boldsymbol{\theta}}_i) \mathbf{H}_J^{-1}(\hat{\boldsymbol{\theta}}_i) = \frac{2\hat{\sigma}_i^2}{m(N_0 + 1)} \mathbf{H}_J^{-1}(\hat{\boldsymbol{\theta}}_i) \quad (2.13)$$

and

$$\hat{\Sigma}_{\sigma^2} = \frac{2\hat{\sigma}_i^4}{m(N_0 + 1)} \quad (2.14)$$

Therefore, the likelihood function can be approximated as the multiplication of two Gaussian distributions. Moreover, from the definition of $J(\boldsymbol{\theta}_i)$ in Eq. (2.4), it can be observed that as the number of data or modes increases, the value of $J(\hat{\boldsymbol{\theta}}_i)$ tends to a finite value, which represents the average error between model predicted and measured (identified) modal properties. Similarly, the entries of the Hessian matrix $\mathbf{H}(\hat{\boldsymbol{\theta}}_i)$ stabilize to finite values as the number of data increases. As a result, the expressions in Eq. (2.13) and Eq. (2.14) suggest that uncertainties in the estimates of $\boldsymbol{\theta}_i$ and σ^2 are inversely proportional to the square root of the number of data $m(N_0 + 1)$, guarantying that the identification uncertainties are decreased as the number of data within a dataset increases. Also, based on Eq. (2.13), the identification uncertainty depends on the magnitude of the prediction error $\hat{\sigma}_i^2$. It is notable that small values of prediction error variance are desirable as it implies good fitting accuracy.

2.2.2.2 Asymptotic approximation 2 (A-2) for likelihood function

An alternative approximation is to expand the function $L(\boldsymbol{\theta}_i, \sigma^2)$ with respect to the structural model parameters $\boldsymbol{\theta}_i$ only, about the MPV $\hat{\boldsymbol{\theta}}_i$ which is also given by Eq. (2.9) and is independent of σ^2 , and following a similar procedure, the likelihood can be also simplified in the form:

$$p(D_i | \boldsymbol{\theta}_i, \sigma^2) \propto (\sigma^2)^{-\frac{m(N_0+1)}{2}} \exp\left(-\frac{m(N_0+1)}{2\sigma^2} J(\hat{\boldsymbol{\theta}}_i)\right) \sqrt{|\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2)|} N(\boldsymbol{\theta}_i | \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2)) \quad (2.15)$$

where

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2) = \frac{2\sigma^2}{m(N_0+1)} \mathbf{H}_J^{-1}(\hat{\boldsymbol{\theta}}_i) \quad (2.16)$$

is the inverse of the Hessian matrix evaluated at the MPV $\hat{\boldsymbol{\theta}}_i$. Details for A-2 can be found in Appendix A. The difference between the two approximations is that in the second one there is no approximation introduced with respect to the prediction error parameters σ^2 , resulting in a covariance matrix $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2)$ of the normal distribution that depends on the value of σ^2 .

2.2.3 Joint distribution for a set of hyper-parameters $\{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma^2\}$

2.2.3.1 Joint posterior distribution of all parameters

According to the Bayes' rule, one can find the joint posterior distribution of all parameters:

$$p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | \mathbf{D}) \propto p(\mathbf{D} | \{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \quad (2.17)$$

where $p(\mathbf{D} | \{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ denotes the likelihood function for all datasets and $p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ denotes the prior distribution of all parameters. The datasets are statistically independent, and the probability of observing the i -th dataset D_i depends only on $\boldsymbol{\theta}_i$ and σ^2 , i.e. $p(D_i | \{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = p(D_i | \boldsymbol{\theta}_i, \sigma^2)$. Therefore, the likelihood function can be simplified as the product of the individual likelihood functions $p(D_i | \boldsymbol{\theta}_i, \sigma^2)$, given as follows:

$$p(\mathbf{D} | \{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = \prod_{i=1}^{N_D} p(D_i | \boldsymbol{\theta}_i, \sigma^2) \quad (2.18)$$

Substituting the approximation 1 given by the expression in Eq. (2.12) into Eq. (2.18), the likelihood function can be then written as:

$$p(\mathbf{D} | \{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \propto \prod_{i=1}^{N_D} N(\boldsymbol{\theta}_i | \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}) N(\sigma^2 | \hat{\sigma}_i^2, \hat{\Sigma}_{\sigma_i^2}) \quad (2.19)$$

The joint prior distribution of all the parameters $p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ is given as:

$$p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D} | \sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) p(\sigma^2 | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \quad (2.20)$$

Due to the independence of θ_i and σ^2 , and the independence of σ^2 and μ_0, Σ_0 , the joint prior distribution is expressed as:

$$p(\{\theta_i\}_{i=1}^{N_D}, \sigma^2, \mu_0, \Sigma_0) = p(\mu_0, \Sigma_0) p(\sigma^2) \prod_{i=1}^{N_D} p(\theta_i | \mu_0, \Sigma_0) \quad (2.21)$$

where $p(\theta_i | \mu_0, \Sigma_0)$ follows the Gaussian distribution assumed in Eq. (2.2). $p(\mu_0, \Sigma_0)$ and $p(\sigma^2)$ denote the prior distributions for the hyper-parameters and prediction errors, respectively. Substituting Eqs. (2.19) and (2.21) into Eq. (2.17), the joint posterior distribution takes the final form:

$$\begin{aligned} & p(\{\theta_i\}_{i=1}^{N_D}, \sigma^2, \mu_0, \Sigma_0 | \mathbf{D}) \\ & \propto p(\mu_0, \Sigma_0) p(\sigma^2) \prod_{i=1}^{N_D} [N(\sigma^2 | \hat{\sigma}_i^2, \hat{\Sigma}_{\sigma_i^2})] \prod_{i=1}^{N_D} N(\theta_i | \hat{\theta}_i, \hat{\Sigma}_{\theta_i}) N(\theta_i | \mu_0, \Sigma_0) \end{aligned} \quad (2.22)$$

2.2.3.2 Marginalization over the model parameter θ_i

The joint distribution of a set of parameters $\{\mu_0, \Sigma_0, \sigma^2\}$ can be calculated by integrating the posterior distribution in Eq. (2.22) over the model parameter space $\{\theta_i\}_{i=1}^{N_D}$, resulting in:

$$\begin{aligned} p(\sigma^2, \mu_0, \Sigma_0 | \mathbf{D}) &= \int_{\theta_i} p(\{\theta_i\}_{i=1}^{N_D}, \sigma^2, \mu_0, \Sigma_0 | \mathbf{D}) d\{\theta_i\}_{i=1}^{N_D} \\ &\propto p(\mu_0, \Sigma_0) p(\sigma^2) \prod_{i=1}^{N_D} \left[N(\sigma^2 | \hat{\sigma}_i^2, \hat{\Sigma}_{\sigma_i^2}) \right] \prod_{i=1}^{N_D} \int_{\theta_i} N(\theta_i | \mu_0, \Sigma_0) N(\theta_i | \hat{\theta}_i, \hat{\Sigma}_{\theta_i}) d\theta_i \end{aligned} \quad (2.23)$$

The integrals of the product of two normal distributions involved in Eq. (2.23) can be simplified as follows [32,38]:

$$\int_{\theta_i} N(\theta_i | \mu_0, \Sigma_0) N(\theta_i | \hat{\theta}_i, \hat{\Sigma}_{\theta_i}) d\theta_i = N(\mu_0 | \hat{\theta}_i, \Sigma_0 + \hat{\Sigma}_{\theta_i}) \quad (2.24)$$

while the third factor in Eq. (2.23), involving the product of N_D normal distribution of scalar variable σ^2 , is also given by a normal distribution $N(\sigma^2 | \hat{\mu}_{\sigma^2}, \hat{\sigma}_{\sigma^2}^2)$ with mean and variance [39]

$$\begin{aligned} \hat{\mu}_{\sigma^2} &= [(\hat{\sigma}_i^2)^{-2}]_{ave}^{-1} [(\hat{\sigma}_i^2)^{-1}]_{ave} = [J^{-2}(\hat{\theta}_i)]_{ave}^{-1} [J^{-1}(\hat{\theta}_i)]_{ave} \\ \hat{\sigma}_{\sigma^2}^2 &= \frac{2}{m(N_0 + 1)N_D} [(\hat{\sigma}_i^2)^{-2}]_{ave}^{-1} = \frac{2}{m(N_0 + 1)N_D} [J^{-2}(\hat{\theta}_i)]_{ave}^{-1} \end{aligned} \quad (2.25)$$

where the symbol $[a_i]_{ave} = \frac{1}{N_D} \sum_{i=1}^{N_D} a_i$ denotes the arithmetic mean of the values

$a_i, i = 1, \dots, N_D$. Since $J(\hat{\theta}_i)$ stabilizes for a large number of data existing within a dataset, the mean value $\hat{\mu}_{\sigma^2}$ is independent of the number of datasets N_D , while the variance $\hat{\sigma}_{\sigma^2}^2$ is

inversely proportional to $m(N_0+1)N_D$, that is, it is inversely proportional to the number of data points within a dataset and the number of dataset. The mean of the prediction error parameter σ^2 generally depends on an average of the individual prediction errors $J(\hat{\theta}_i)$ for each dataset, arising from model and measurement errors. The uncertainty $\hat{\sigma}_{\sigma^2}^2$ given in Eq. (2.25) reduces as the number of data within a given dataset and the number of datasets N_D increases. Thus, the joint distribution of a set of parameters $\{\mu_0, \Sigma_0, \sigma^2\}$ is expressed analytically as:

$$p(\sigma^2, \mu_0, \Sigma_0 | \mathbf{D}) \propto p(\mu_0, \Sigma_0) p(\sigma^2) N(\sigma^2 | \hat{\mu}_{\sigma^2}, \hat{\sigma}_{\sigma^2}^2) \prod_{i=1}^{N_D} N(\mu_0 | \hat{\theta}_i, \Sigma_0 + \hat{\Sigma}_{\theta_i}) \quad (2.26)$$

Similarly, using the approximation A-2 of the likelihood function, the posterior distribution of a set of parameters $\{\mu_0, \Sigma_0, \sigma^2\}$ is given by (see derivation in Appendix A)

$$p(\sigma^2, \mu_0, \Sigma_0 | \mathbf{D}) \propto p(\mu_0, \Sigma_0) p(\sigma^2) T(\sigma^2) \prod_{i=1}^{N_D} N(\mu_0 | \hat{\theta}_i, \Sigma_0 + \hat{\Sigma}_{\theta_i}(\sigma^2)) \quad (2.27)$$

where $T(\sigma^2)$ is defined as:

$$T(\sigma^2) = (\sigma^2)^{\frac{[m(N_0+1)-N_\theta]N_D}{2}} \exp\left(-\frac{m(N_0+1)N_D}{2\sigma^2} J_D\right) \quad (2.28)$$

and $J_D = [J(\hat{\theta}_i)]_{ave} = \frac{1}{N_D} \sum_{i=1}^{N_D} J(\hat{\theta}_i)$ is the average of the measures of fit over all datasets.

Therefore, based on Eq. (2.27), one can also get the joint posterior distribution of parameters $\{\mu_0, \Sigma_0, \sigma^2\}$. It should be noted that the structure of Eq. (2.27) is exactly the same as the structure of Eq. (2.26) with the only difference that the covariance $\hat{\Sigma}_{\theta_i}(\sigma^2)$ in Eq. (2.27) is replaced by $\hat{\Sigma}_{\theta_i}$ and the expression $T(\sigma^2)$ in Eq. (2.27) is replaced by the normal distribution $N(\sigma^2 | \hat{\mu}_{\sigma^2}, \hat{\sigma}_{\sigma^2}^2)$. In the numerical example, the accuracy of both approximations will be investigated.

According to Eq. (2.26) or (2.27), samples of prediction error parameter and the hyper-parameters can be generated using any Markov Chain Monte Carlo (MCMC) algorithm [40] such as the Transitional MCMC (TMCMC) algorithm [41,42] or the Nested algorithm [43]. It should be noted that the sampling approaches no longer involve model runs once the optimal values are obtained. These samples obtained here will be then used to estimate uncertainties in the model parameter θ in section 2.2.5 and to propagate the uncertainty for predicting the QoI in section 2.2.6.

2.2.4 Calculation of the MPV of the hyper-parameters and predictions error variance

The MPV of the hyper-parameters $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\Sigma}_\theta$ and the prediction error parameter σ^2 are obtained by minimizing the negative logarithm of the joint distribution $p(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \sigma^2 | \mathbf{D})$, given by:

$$L_p(\sigma^2, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) = -\ln p(\sigma^2, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta | \mathbf{D}) \quad (2.29)$$

For uniform prior distributions $p(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ and $p(\sigma^2)$ assigned for the hyper-parameters, Eq. (2.29) simplifies to

$$\begin{aligned} L_p(\sigma^2, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) \\ = \frac{1}{2\hat{\sigma}_{\sigma^2}^2} (\sigma^2 - \hat{\mu}_{\sigma^2})^2 + \frac{1}{2} \sum_{i=1}^{N_D} \ln(|\boldsymbol{\Sigma}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i}|) + \frac{1}{2} \sum_{i=1}^{N_D} (\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i)^T (\boldsymbol{\Sigma}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} (\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i) + c \end{aligned} \quad (2.30)$$

where c is constant. The first derivatives of the objective function $L_p(\sigma^2, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ with respect to the respective parameters $\boldsymbol{\mu}_\theta$, $\boldsymbol{\Sigma}_\theta$ and σ^2 can be derived, based on Eqs. (A.5), (A.7) and (A.8) in Appendix A, in the form [32]:

$$\frac{\partial L_p}{\partial \boldsymbol{\mu}_\theta} = \sum_{i=1}^{N_D} (\boldsymbol{\Sigma}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} (\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i) \quad (2.31)$$

$$\frac{\partial L_p}{\partial \boldsymbol{\Sigma}_\theta} = \frac{1}{2} \sum_{i=1}^{N_D} (\boldsymbol{\Sigma}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} - \frac{1}{2} \sum_{i=1}^{N_D} (\boldsymbol{\Sigma}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} (\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i) (\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i)^T (\boldsymbol{\Sigma}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} \quad (2.32)$$

$$\frac{\partial L_p}{\partial \sigma^2} = \frac{1}{\hat{\sigma}_{\sigma^2}^2} (\sigma^2 - \hat{\mu}_{\sigma^2}) \quad (2.33)$$

The MPV of the hyper-parameters ($\bar{\boldsymbol{\mu}}_\theta$ and $\bar{\boldsymbol{\Sigma}}_\theta$) are computed by setting the first derivatives in Eqs. (2.31) and (2.32) equal to zero and solving the resulting equations. This results in [32]:

$$\bar{\boldsymbol{\mu}}_\theta = \sum_{i=1}^{N_D} \mathbf{w}_i \hat{\boldsymbol{\theta}}_i \quad (2.34)$$

$$\sum_{i=1}^{N_D} \mathbf{w}_i (\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i) (\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i)^T \mathbf{w}_i^T = \left[\sum_{i=1}^{N_D} (\bar{\boldsymbol{\Sigma}}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} \right]^{-1} \quad (2.35)$$

where \mathbf{w}_i is the weighting matrix, given by:

$$\mathbf{w}_i = \left[\sum_{i=1}^{N_D} (\bar{\boldsymbol{\Sigma}}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} \right]^{-1} (\bar{\boldsymbol{\Sigma}}_\theta + \hat{\boldsymbol{\Sigma}}_{\theta_i})^{-1} \quad (2.36)$$

Thus the optimizations can be carried out according to Eqs. (2.34) and (2.35). The MPV of the hyper-parameter $\bar{\boldsymbol{\mu}}_0$ is a weighted sum of the MPVs of the model parameters for each dataset ($\hat{\boldsymbol{\theta}}_i$).

To gain more insight into these expressions, the identification uncertainty $\hat{\boldsymbol{\Sigma}}_{\theta_i}$ over all datasets are assumed to be equal to a covariance matrix $\boldsymbol{\Sigma}$ taken as

$$\boldsymbol{\Sigma} = \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\boldsymbol{\Sigma}}_{\theta_i} \quad (2.37)$$

the average of the identification uncertainty of each dataset. This assumption is used to simplify the expressions and one can derive the MPV estimations in the form [32]:

$$\bar{\boldsymbol{\mu}}_0 = \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\boldsymbol{\theta}}_i \quad (2.38)$$

$$\bar{\boldsymbol{\Sigma}}_{\theta} = \boldsymbol{\Sigma}_v - \boldsymbol{\Sigma} \quad (2.39)$$

$$\bar{\sigma}^2 = \hat{\mu}_{\sigma^2} \quad (2.40)$$

where

$$\boldsymbol{\Sigma}_v = \frac{1}{N_D} \sum_{i=1}^{N_D} (\bar{\boldsymbol{\mu}}_0 - \hat{\boldsymbol{\theta}}_i)(\bar{\boldsymbol{\mu}}_0 - \hat{\boldsymbol{\theta}}_i)^T \quad (2.41)$$

is the uncertainty due to the variability of the MPV from each dataset around their mean value, while $\boldsymbol{\Sigma}$ is the identification uncertainty assumed the same for each dataset.

Eq. (2.39) provides valuable insight into the posterior uncertainty of the structural parameters. In particular, the following points can be outlined:

- The posterior uncertainty of the structural parameters is obtained as the difference between the variability in the parameter estimates and the identification uncertainty.
- The variability of the structural parameters can be due to environmental changes, material/geometric nonlinearities, manufacturing variability, etc. This variability might increase or decrease depending on the testing conditions.
- The identification uncertainty arises from the parameter estimation process when a specific dataset is used. It is expected to decrease as additional data points are incorporated.
- When the number of data points in each dataset increases substantially, the hyper covariance can be generally approximated as the variability since the identification uncertainty $\boldsymbol{\Sigma}$ becomes almost negligible.

- Eq. (2.39) holds true only if the covariance matrix $\bar{\Sigma}_\theta = \Sigma_v - \Sigma$ is a positive definite matrix. A negative definite matrix $\Sigma_v - \Sigma$ might be encountered when the variability is smaller than the identification uncertainty. In this case, the optimal value lies at the boundary of the domain of definition of the hyper-parameters $\{\mu_\theta, \Sigma_\theta, \sigma^2\}$.
- For a sufficiently large number of data within each dataset, the identification uncertainty is expected to be extremely small, and the positive definiteness of the covariance matrix $\bar{\Sigma}_\theta = \Sigma_v - \Sigma$ is guaranteed.

For the second approximation, by following the same procedure from Eq. (2.29) to Eq. (2.40), the MPV of the hyper-parameters and prediction error parameter can be computed as:

$$\begin{aligned}\bar{\mu}_\theta &= \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\theta}_i \\ \bar{\Sigma}_\theta &= \Sigma_v - \Sigma(\bar{\sigma}^2) \\ \bar{\sigma}^2 &= \frac{m(N_0 + 1)}{m(N_0 + 1) - N_\theta} J_D\end{aligned}\tag{2.42}$$

where

$$\Sigma(\bar{\sigma}^2) = \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\Sigma}_{\theta_i}(\bar{\sigma}^2)\tag{2.43}$$

and N_θ is the number of the model parameters. More details for the derivations can be found in Appendix A.

2.2.5 Posterior predictive distribution of structural parameters

Based on the total probability theorem, the posterior predictive distribution of the model parameters can be calculated as:

$$p(\theta_{new} | \mathbf{D}) = \int \int \int p(\theta_{new} | \mu_\theta, \Sigma_\theta) p(\sigma^2, \mu_\theta, \Sigma_\theta | \mathbf{D}) d\mu_\theta d\Sigma_\theta d\sigma^2\tag{2.44}$$

where θ_{new} is a new sample from the hyper distribution $N(\theta_{new} | \mu_\theta, \Sigma_\theta)$, which does not depend on σ^2 . Thus, one can replace $p(\theta_{new} | \sigma^2, \mu_\theta, \Sigma_\theta)$ by $N(\theta_{new} | \mu_\theta, \Sigma_\theta)$. Using samples $(\mu_\theta^{(i)}, \Sigma_\theta^{(i)}, \sigma^{2(i)})$ from the PDF given in Eq. (2.26) or (2.27), a sample estimate of the multi-dimensional integral is obtained as:

$$p(\theta_{new} | \mathbf{D}) \approx \frac{1}{N} \sum_{i=1}^N N(\theta_{new} | \mu_\theta^{(i)}, \Sigma_\theta^{(i)})\tag{2.45}$$

where N is the number of samples.

2.2.6 Propagation of uncertainties in output QoI

The model parameter and prediction error uncertainties are next propagated to predict the uncertainties in the modal properties and time histories. Once the structural model $M(\boldsymbol{\theta}_{new})$ is updated based on the multiple datasets, the uncertainties can be propagated to predict the output QoI. Herein, in addition to the uncertainty of structural parameters, the uncertainty of prediction error parameter σ^2 is considered for predicting an output QoI. The samples of prediction error parameter and structural parameters $\{\sigma^{2(q)}, \boldsymbol{\theta}_{new}^{(q)}\}$ are obtained according to Eqs. (2.26) and (2.45), respectively using the Nested sampling technique [43].

2.2.6.1 Propagation for estimating modal properties

Samples of the modal properties (the square of modal frequencies $\tilde{\lambda}_{r,q}$ and mode shapes $\tilde{\phi}_{r,q}$) can be then computed as:

$$\begin{aligned} \lambda_r^{(q)} &= \lambda_r(\boldsymbol{\theta}_{new}^{(q)}) + e_r^{(q)}, & e_r^{(q)} &\sim N(e_r | 0, \tilde{\lambda}_r^2 \sigma^{2(q)}) \\ \phi_r^{(q)} &= \frac{\phi_r(\boldsymbol{\theta}_{new}^{(q)})}{\|\phi_r(\boldsymbol{\theta}_{new}^{(q)})\|} + \boldsymbol{\varepsilon}_r^{(q)}, & \boldsymbol{\varepsilon}_r^{(q)} &\sim N(\boldsymbol{\varepsilon}_r | \mathbf{0}, \sigma^{2(q)} \mathbf{I}) \end{aligned} \quad (2.46)$$

where the symbol $\boldsymbol{\theta}_{new}^{(q)}$ denotes the q -th sample of model parameter $\boldsymbol{\theta}_{new}$ generated from its distribution $p(\boldsymbol{\theta}_{new} | \mathbf{D})$. $\lambda_r(\boldsymbol{\theta}_{new}^{(q)})$ and $\phi_r(\boldsymbol{\theta}_{new}^{(q)})$ are respectively the r -th predicted square of the modal frequency and mode shape referring to the sample $\boldsymbol{\theta}_{new}^{(q)}$. The prediction error parameters $e_r^{(q)}$ and $\boldsymbol{\varepsilon}_r^{(q)}$ corresponding to the r -th modal properties follow zero-mean Gaussian distribution with variance $\tilde{\lambda}_r^2 \sigma^{2(q)}$ and covariance $\sigma^{2(q)} \mathbf{I}$. Using the samples in Eq. (2.46), the estimates of mean, standard deviation and $\alpha\%$ to $1-\alpha\%$ quantiles can be obtained for the modal properties.

2.2.6.2 Propagation for estimating response time histories

Once the samples of modal properties are obtained, one can predict response time histories (e.g. displacement, velocity, acceleration, inter-story drift and stresses) by using modal superposition method [28]:

$$g^{(q)}(t) = C(\boldsymbol{\theta}_{new}^{(q)}) \sum_{r=1}^m \phi_r^{(q)} \xi_r^{(q)}(t) \quad (2.47)$$

where $\phi_r^{(q)}$ is the r -th predicted mode shapes computed by Eq. (2.46), $g(t)$ denotes either the displacement, inter-story drift or stress histories. The velocity or acceleration time histories can also be readily calculated by replacing the modal coordinates $\xi_r^{(q)}(t)$ in Eq. (2.47) with $\dot{\xi}_r^{(q)}(t)$ or $\ddot{\xi}_r^{(q)}(t)$. $C(\theta_{new}^{(q)})$ is a mapping matrix for calculating other quantities, e.g. inter-story drift time histories. The samples $\xi_r^{(q)}(t)$ of the modal responses satisfy the following modal equation:

$$\ddot{\xi}_r^{(q)}(t) + 2\zeta_r^{(q)}\omega_r^{(q)}\dot{\xi}_r^{(q)}(t) + \omega_r^{2(q)}\xi_r^{(q)}(t) = \frac{1}{(\phi_r^{T(q)}M_m\phi_r^{(q)})}\phi_r^{T(q)}L\mathbf{F}(t) \quad (2.48)$$

where $\omega_r^{(q)}$ is the sample of natural frequency in rad/s given by $\omega_r^{(q)} = \sqrt{\lambda_r^{(q)}}$, and $\zeta_r^{(q)}$ is the sample of damping ratio, generated by assigning a PDF for ζ_r . $\mathbf{F}(t)$ is the input force vector, and L is a matrix that associates the independent forces to the DOF of the model. Using the samples $g^{(q)}(t)$, $q = 1, \dots, N$, the mean, standard deviation and quantiles, quantifying the response uncertainties can be estimated.

2.2.6.3 Propagation for estimating failure probability

The uncertainties in the model parameters and prediction error can also be used to obtain data-driven and robust-to-parameters estimates of the probability of failure of the structure subjected to deterministic or stochastic excitations. The probability of failure is defined in this work as the probability of exceeding a threshold level associated with the system performance. For example, failure may be associated with the inter-story drift exceeding a threshold level. Let $g(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{e}) < 0$ be the failure domain identified by the limit state function $g(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{e})$, where $\mathbf{e}^T = [e_1, \dots, e_m, \boldsymbol{\epsilon}_1^T, \dots, \boldsymbol{\epsilon}_m^T]$ includes the normally distributed prediction error terms, and $\boldsymbol{\phi}$ is the set of parameters associated with other uncertainties (e.g. input related uncertainties) that were not identified by the HBM framework. It is assumed that $\boldsymbol{\phi}$ is uncorrelated to $\{\boldsymbol{\theta}, \mathbf{e}\}$ and a PDF $p(\boldsymbol{\phi})$ is assigned to quantify uncertainties in $\boldsymbol{\phi}$. The probability of failure $\Pr(F | \mathbf{D})$ given the data is defined by [44,45]

$$\begin{aligned} \Pr(F | \mathbf{D}) &= \Pr(g(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{e}) < 0 | \mathbf{D}) \\ &= \int_{g(\boldsymbol{\phi}, \boldsymbol{\theta}, \mathbf{e}) < 0} p(\boldsymbol{\phi})p(\mathbf{e} | \sigma^2) p(\boldsymbol{\theta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) p(\sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | \mathbf{D}) d\boldsymbol{\phi} d\boldsymbol{\theta} d\mathbf{e} d\sigma^2 d\boldsymbol{\mu}_0 d\boldsymbol{\Sigma}_0 \end{aligned} \quad (2.49)$$

Assuming a narrow distribution for the hyper-parameters so that the probability distribution can be approximated by a very narrow distribution in the hyper-parameter space, the failure probability integral can be approximated in the form

$$\begin{aligned} \Pr(F | \mathbf{D}) &= \Pr(g(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mathbf{e}) < 0 | \mathbf{D}, \bar{\boldsymbol{\mu}}_{\boldsymbol{\theta}}, \bar{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}, \bar{\sigma}^2) \\ &= \int_{g(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mathbf{e}) < 0} p(\boldsymbol{\varphi}) p(\mathbf{e} | \bar{\sigma}^2) p(\boldsymbol{\theta} | \bar{\boldsymbol{\mu}}_{\boldsymbol{\theta}}, \bar{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}) d\boldsymbol{\varphi} d\boldsymbol{\theta} d\mathbf{e} \end{aligned} \quad (2.50)$$

The integration in Eq. (2.50) can be performed using subset simulation (SubSim) algorithm [46,47]. One can exploit the fact that the probability distributions are Normal and use available software [48] to carry out sample estimates of the probability integrals.

2.3 Computational Algorithm

The procedure for applying the proposed HBM framework for parameter estimation and robust predictions is summarized in Algorithm 1. Two steps are presented for the parameter estimation. The first step requires model runs to approximate the likelihood function for each dataset, and thus it is computationally the most expensive step. In the second step, no model runs are required. Only the MPV and the uncertainties of the model and prediction error parameters generated from the first step are needed to obtain the samples from the joint distribution $p(\boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \sigma^2 | \mathbf{D})$. Note that an alternative full sampling (FS) approach can be employed to acquire the samples of the hyper-parameters. The sampling procedure provides an accurate solution for a large number of samples. It can be accomplished by drawing the samples of model and prediction error parameters from the likelihood function for each dataset in the first step, and thus it is computationally much more expensive than the asymptotic approximations. The available samples from the first step are subsequently utilized for computing the posterior distributions of the hyper-parameters in the second step. More details for the sampling approach can be found in [23,49]. Comparisons of the sampling approach with asymptotic techniques are presented in the next section.

The second stage in the Algorithm propagates both parameter and prediction error uncertainties to compute any QoI. For this purpose, the uncertainty of modal parameters can be propagated for predicting (un-)observed modal or response quantities, considering both parameter and prediction error uncertainties. This is in contrast to the parameter inference done directly based on response time histories, where the model prediction error is evaluated only for the measured QoI (e.g. acceleration) and cannot be transferred to other QoI, such as drift, or stresses that are important when safety and operational issues are considered.

Algorithm 1:

Proposed HBM framework for parameter estimation and robust predictions

1. Parameter estimation

Step 1: Model runs are required

1a) Find the MPV of θ_i 's and the hessian matrix of $L(\theta_i, \sigma^2)$ for each dataset D_i , $i = 1, \dots, N_D$

1.1) Minimize $J(\theta_i)$ to compute the MPV $\hat{\theta}_i$ using Eq. (2.9)

1.1.1) Only for A-1: Compute $J(\hat{\theta}_i)$ and obtain $\hat{\mu}_{\sigma^2}$ and $\hat{\sigma}_{\sigma^2}^2$ using Eq. (2.25)

1.2) Evaluate the Hessian matrix $\mathbf{H}_J(\hat{\theta}_i)$ and obtain $\hat{\Sigma}_{\theta_i}$ using Eq. (2.13) for A-1 or $\hat{\Sigma}_{\theta_i}(\sigma^2)$ using Eq. (2.16) for A-2

Step 2: No model runs are required

1b) Sample from the joint distribution $p(\mu_0, \Sigma_0, \sigma^2 | \mathbf{D})$ of the set of parameters $\{\mu_0, \Sigma_0, \sigma^2\}$ using Eq. (2.26) for A-1 or Eq. (2.27) for A-2

2. Robust predictions

- 2a) Draw samples $\theta_{new}^{(q)}$, $q = 1, \dots, N_q$, from $p(\theta_{new}^{(q)} | \mathbf{D})$ in Eq. (2.45), and samples $e_r^{(q)}$ and $\epsilon_r^{(q)}$ from $N(e_r | 0, \tilde{\lambda}_r^2 \sigma^{2(q)})$ and $N(\epsilon_r | \mathbf{0}, \sigma^{2(q)} \mathbf{I})$, respectively, where N_q is the number of samples from the last stage of nested sampling
- 2b) Compute samples of modal properties and time histories responses using Eqs. (2.46), (2.47) and (2.48)
- 2c) Use the samples to estimate the statistical properties, such as the mean, standard deviation and quantiles
- 2d) Calculate the mean values of $\bar{\mu}_0$, $\bar{\Sigma}_0$ and $\bar{\sigma}^2$ from step 1b and calculate the failure probability using Eq. (2.50)
-

2.4 Illustrative Examples

2.4.1 Simulated case study: 3 degrees-of-freedom (3-DOF) system

A 3-DOF shear model of a building system is employed as a simulated example to investigate the performance of the proposed HBM framework. The structure is shown in Fig. 2-2(a), where the system is fixed at the base. The nominal masses ($m_{0,i}$) are 1kg, and the stiffness ($k_{0,i}$) of each story is 1800N/m. Given these properties, the natural frequencies of the three modes of the system are 3.0Hz, 8.42Hz and 12.17Hz. The damping ratio of each mode is assumed to be 0.02. The system is excited at the base with an earthquake excitation (\ddot{y}_g) shown in Fig. 2-2(b).

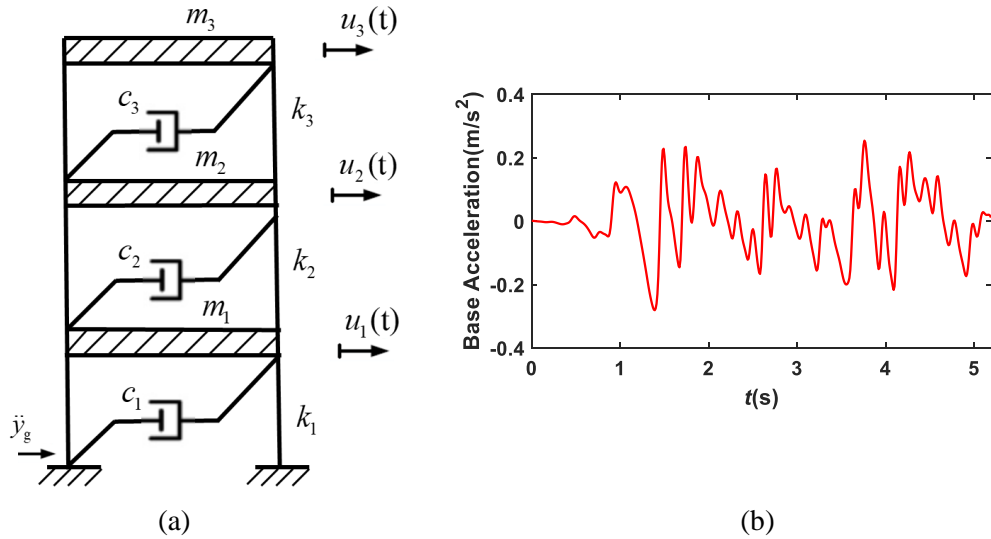


Fig. 2-2 (a) 3-DOF shear model of a building and (b) base earthquake excitation

The purpose of the identification is to estimate the stiffness of each story. The model is parameterized using three model parameters $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \theta_3]^T$ associated with the stiffness of each story, defined as $\theta_i = k_i / k_{0,i}$ and representing the ratio of the stiffness to the nominal stiffness value. A Gaussian prior distribution is introduced for the three parameters with hyper mean $\boldsymbol{\mu}_\theta = [\mu_{\theta_1} \ \mu_{\theta_2} \ \mu_{\theta_3}]^T$ and hyper covariance $\boldsymbol{\Sigma}_\theta$ assumed to be a diagonal matrix with diagonal element $\sigma_{\theta_i}^2$.

Measurement data for the modal properties of each dataset is simulated from the system for fixed values of $\boldsymbol{\mu}_\theta$, $\boldsymbol{\Sigma}_\theta$, and σ^2 . Specifically, to simulate a dataset D_i , a sample $\boldsymbol{\theta}_i$ is drawn from the Gaussian prior distribution $N(\boldsymbol{\theta}_i | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ and the modal properties $\omega_{r,i}$ and $\phi_{r,i}$ are computed. These modal properties are perturbed according to Eq. (2.1) by adding an independent Gaussian error term to the modal frequencies and mode shape components for all modes to derive the measured modal frequencies $\hat{\omega}_{r,i}$ and mode shapes $\hat{\phi}_{r,i}$. The intensity of the error is controlled by the value of the standard deviation σ . The procedure is repeated for generating all datasets $i = 1, \dots, N_D$. In total, $N_D = 100$ datasets of modal frequencies and mode shapes are simulated.

Four cases are explored for creating the datasets, summarized in Table 2-1. The first one simulates no variability over datasets; the second case considers a large variability in $\boldsymbol{\theta}_i$'s while having small measurement noise; the third case simulates the opposite of the case 2; the last case considers large variability and measurement noise. In all four cases, the hyper mean

vectors are selected to be $\mu_{0_i} = \mathbf{1}$. The proposed HBM framework is then applied for parameter estimation and model predictions. The Nested sampling algorithm [43] is employed in this work to generate samples from the joint PDF given by Eq. (2.26) or (2.27). In the sampling algorithm, the number of the initial samples is chosen as 500, and the tolerance is set to 0.01.

Table 2-1 Information for 4 simulated cases

Case	Variability σ_{θ_i} (%)	Noise σ (%)	Remarks
1	0.00	2.00	No variability, i.e. $D_i = D_1$ for $i = 1, \dots, N_D$
2	2.00	0.05	Large variability, small measurement noise
3	0.05	2.00	Small variability, large measurement noise
4	2.00	2.00	Large variability, large measurement noise

Table 2-2 summarizes the MPV and mean estimates of the hyper-parameters and the prediction error variance for the four simulation cases. Algorithm 1 is used for the model updating and predictions. The full sampling (FS) approach is also applied to investigate the accuracy of asymptotic approximations. Moreover, the results from the classical Bayesian modeling (CBM) framework [2] that treats the multiple datasets as a single dataset are presented in these Tables. Compared to the FS results, the A-2 gives sufficiently accurate estimations in all four cases. The A-1 provides accurate results for the cases 1 and 2 but it fails to provide sufficiently accurate results for the cases 3 and 4. These results suggest the A-1 is not as accurate as the A-2. This is due to the fact that for each dataset the normal distribution is not an adequate approximation of the individual likelihood functions when the prediction error variance is included. In contrast, the normal distribution is an adequate approximation of the likelihood function for each dataset, seen as a function of the structural parameters for a given value of the prediction error variance. The approximation seems to be accurate even in this example, wherein we have a small number of data points (three modal frequencies and nine mode shape components).

Results in Table 2-2 for $N_D=100$ reveal that the HBM framework recovers the actual values of the hyper mean, hyper variance and prediction error variance in almost all cases. However, this is not the case for the hyper standard deviations in case 3. Although they are estimated to be small (0.5%), indicating small uncertainty in the structural parameters, their actual values are up to one order of magnitude larger than the identified ones (0.05%). This is

due to fact that a small part of the prediction error uncertainty is assigned as the uncertainty in the structural parameters.

Based on the results of the CBM presented in cases 2-4, it can be seen that the parameter uncertainty is severely underestimated. The mean estimate of the prediction error variance is one order of magnitude higher than the one used to simulate the measurements. In these cases, the misfit between the model predictions and the measurements for each dataset is accounted for by the prediction error term. In case 1, where no variability is present, the HBM framework and the CBM frameworks both yield very close results.

Table 2-2 Estimates of mean and standard deviation of the model parameters and prediction error parameters

Parameters		$\hat{\mu}_{\theta_1}$	$\hat{\sigma}_{\theta_1}$	$\hat{\mu}_{\theta_2}$	$\hat{\sigma}_{\theta_2}$	$\hat{\mu}_{\theta_3}$	$\hat{\sigma}_{\theta_3}$	$\hat{\sigma}$	$\hat{\Sigma}_{\sigma}$
Case 1	A-1	0.9911	0.0031	1.0076	0.0030	1.0470	0.0024	0.0198	0.00037
	A-2	0.9919	0.0026	1.0076	0.0027	1.0472	0.0022	0.0199	0.00038
	FS	0.9906	0.0028	1.0068	0.0024	1.0472	0.0023	0.0200	0.00040
	CBM	0.9928	0.0030	1.0069	0.0030	1.0469	0.0022	0.0199	0.00039
Case 2	A-1	1.0020	0.0192	0.9955	0.0213	1.0015	0.0188	0.00032	0.74×10^{-5}
	A-2	1.0021	0.0190	0.9954	0.0213	1.0013	0.0189	0.00049	1.1×10^{-5}
	FS	1.0019	0.0191	0.9956	0.0216	1.0013	0.0191	0.00065	1.6×10^{-5}
	CBM	1.0015	0.0013	0.9954	0.0014	1.0012	0.0008	0.0087	0.00016
Case 3	A-1	0.9978	0.0122	1.0014	0.0134	1.0056	0.0107	0.0132	0.00032
	A-2	0.9973	0.0070	1.0013	0.0056	1.0060	0.0047	0.0195	0.00041
	FS	0.9956	0.0073	1.0044	0.0079	1.0045	0.0060	0.0195	0.00042
	CBM	0.9988	0.0032	1.0011	0.0032	1.0059	0.0019	0.0196	0.00042
Case 4	A-1	1.0024	0.0280	1.0008	0.0254	0.9947	0.0255	0.0134	0.00031
	A-2	1.0007	0.0224	1.0004	0.0233	0.9955	0.0171	0.0198	0.00046
	FS	1.0023	0.0217	1.0013	0.0234	0.9963	0.0177	0.0201	0.00047
	CBM	1.0023	0.0031	1.0011	0.0031	0.9947	0.0022	0.0221	0.00051

Fig. 2-3 shows the posterior distribution of the hyper-parameters and prediction error variance for the simulated case 2. The posterior uncertainty is very small for the hyper mean values and relatively small for the hyper standard deviations.

Based on the estimated $\hat{\theta}_i$'s and $\hat{\Sigma}_{\theta_i}$'s, the MPV and the identification uncertainty is plotted in Fig. 2-4 for each dataset (error bars). The MPV of the hyper standard deviation is

also displayed as a function of the number of datasets, ranging from $N_D=1$ to 100. The results are demonstrated only for θ_2 , but similar trends are observed for θ_1 and θ_3 , not shown in Fig. 2-4.

In case 1, due to the fact that the data in each dataset is chosen to be identical, the MPV $\hat{\theta}_i$ and identification uncertainty $\hat{\Sigma}_{\theta_i}$ are identical for all datasets. As a result, using Eq. (2.38), the hyper mean $\hat{\mu}_\theta$ is independent of the number of the datasets, and $\Sigma_v = \mathbf{0}$ (see Eq. (2.41)), making the covariance matrix $\bar{\Sigma}_\theta$ in Eq. (2.39) non-positive definite. Thus, the analysis in Section 2.2.4 is not applicable. The uncertainty of σ_{θ_2} is a decreasing function of the number of datasets, tending to zero for a large number of datasets due to the fact that all datasets are identical and no variability is expected. In this case, the HBM framework provides the same result as the CBM.

In cases 2 and 4 (Fig. 2-4), it is seen that the variability of the parameters is significant and approaches a constant non-zero value for a large number of datasets. The identification uncertainties are small in case 2 but they are considerably larger in case 4. For a small number of datasets, the parameter uncertainty bounds (UB) vary across the datasets with a tendency to decrease as a function of the number of datasets. After including approximately 10 datasets, the uncertainty bound stabilizes and tends to a constant value.

In case 3 (Fig. 2-4), it can be seen that for a sufficiently large number of datasets the identification uncertainty for a single dataset is significantly larger than the parameter uncertainty due to the variability in the datasets. Thus, only a small portion of the uncertainty is embedded in the structural parameters while the largest portion is attributed to the prediction errors. As the number of datasets increases, the uncertainty in the structural parameters decreases, tending to a constant non-zero value representing the variability in the parameter estimates across datasets.

Based on these results, it can be concluded that the proposed method is capable of providing realistic UB as the number of datasets increases. In contrast, the conventional Bayesian method underestimates the parameter uncertainty and for a large number of datasets they give unrealistically small uncertainties. Moreover, the proposed method can identify the uncertainty in the structural parameters due to environmental, operational and manufacturing variability, whereas the conventional Bayesian method fails to account for such uncertainties and embeds it in the prediction error term.

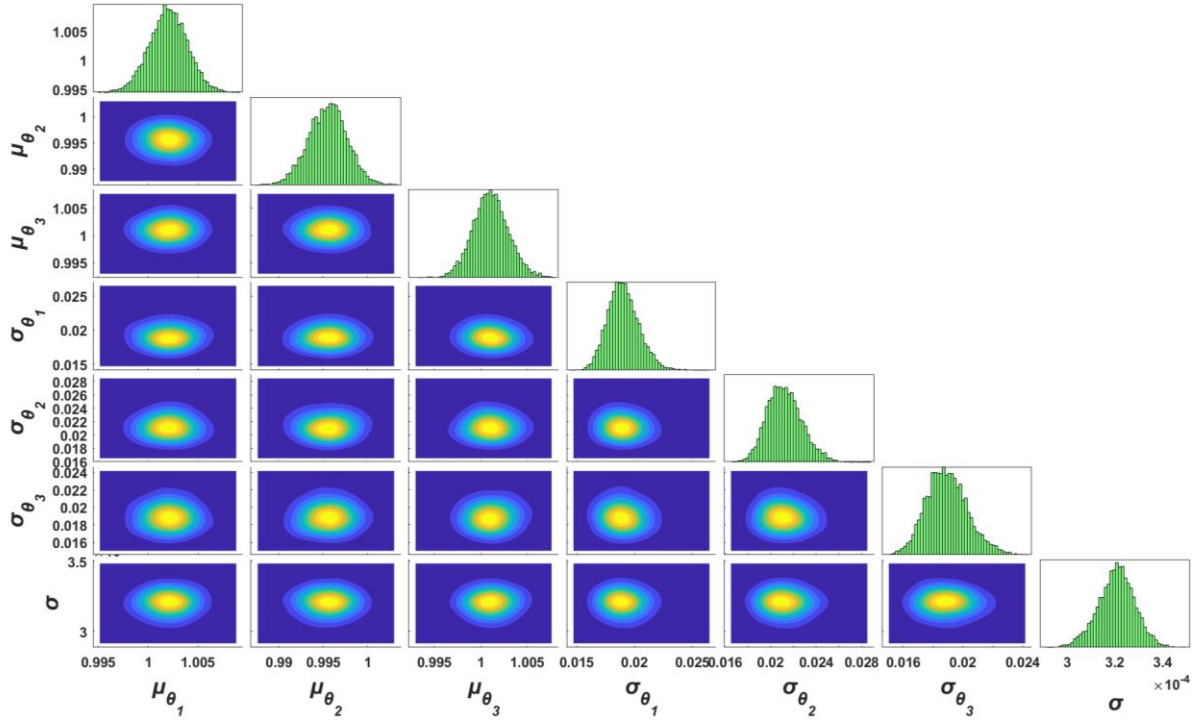


Fig. 2-3 Posterior distribution of hyper parameters and prediction error parameter for case 2.

The diagonal sub-figures depict the marginal distributions for each parameter, and sub-figures below the diagonal show the pairwise contour plots.

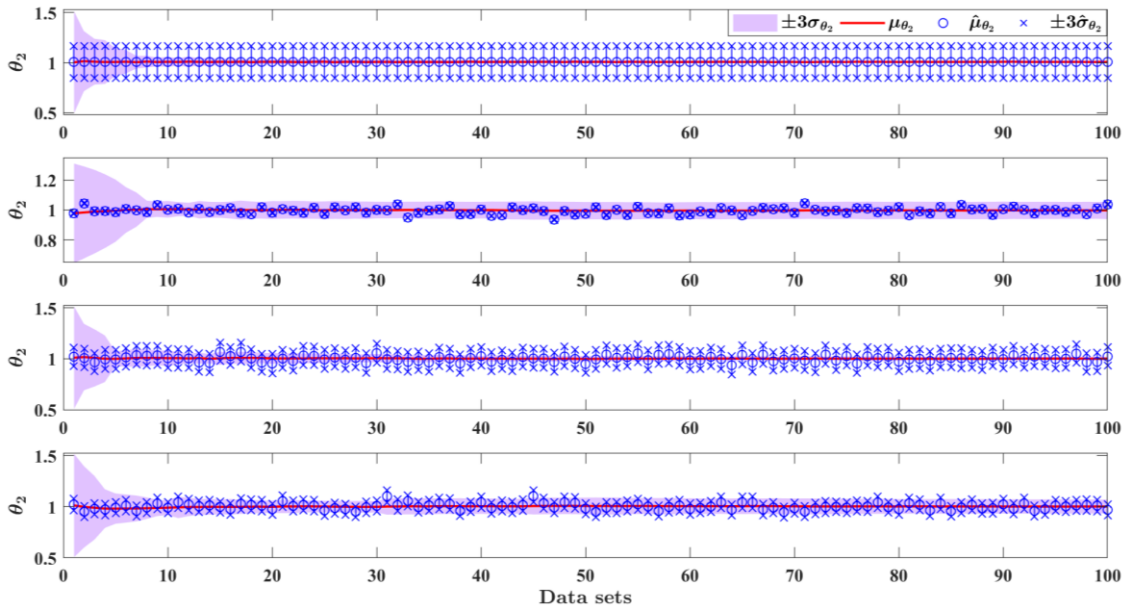


Fig. 2-4 Identification uncertainty and parameter uncertainty as a function of the number of datasets N_D , with N_D ranging from 1 to 100 in linear scale for all cases 1-4 (top-to-bottom)

Based on the hyper-parameter samples, one can get the samples of structural parameters using Eq. (2.45). The prediction error uncertainty and the uncertainty of the structural

parameters are propagated to predict response QoI. Representative results are presented for the simulated cases 2 and 3 assuming that $N_D=100$. The predicted modal frequencies are shown in Fig. 2-5, and the predicted mode shape of the first mode is shown in Fig. 2-6. For ease of comparison, the modal frequencies are normalized by the nominal values of the modal frequencies. The results provide 95% UB in the modal properties. Two uncertainty propagation cases are explored. In the first one only structural parameter uncertainty is considered, while in the second one the uncertainty in the model prediction error in Eq. (2.46) is also considered. The results are also compared with the 95% UB estimated from the measured modal properties for all datasets, as well as with the measured modal properties corresponding to a new dataset, not included in the inference.

In the simulated case 2, it can be seen that the results with and without the prediction error terms are almost identical (compare Fig. 2-5(a) and (b)), attributed to the small prediction error uncertainty assumed in this case. The predicted UB, is significant, with the UB containing the values of the experimental modal properties for all datasets (compare with 95% UB for measurements). Note that the model predicted 95% UB of the measurements are close to model predicted ones for the modal frequencies. This is due to the fact that the variability in the multiple datasets can be adequately captured by varying the structural parameters, with the prediction error to be insignificant. This finding is consistent with the very small prediction error assumed when simulating the modal properties. Referring to Fig. 2-5(a-b), similar results are obtained for the mode shape components although the uncertainty at the component level seems to be smaller than the uncertainty observed for modal frequencies.

In the simulated case 3, Fig. 2-5(c) and Fig. 2-6 (c) indicate that the 95% UB fails to contain the 95% UB of the measured modal properties. Thus, propagating only the small uncertainties in the model parameters is inadequate and could result in underestimation of uncertainties of the modal properties. As shown in Fig. 2-5(d) and Fig. 2-6(d), when the uncertainties from prediction error terms are included, the 95% UB are close to 95% UB for the measurements.

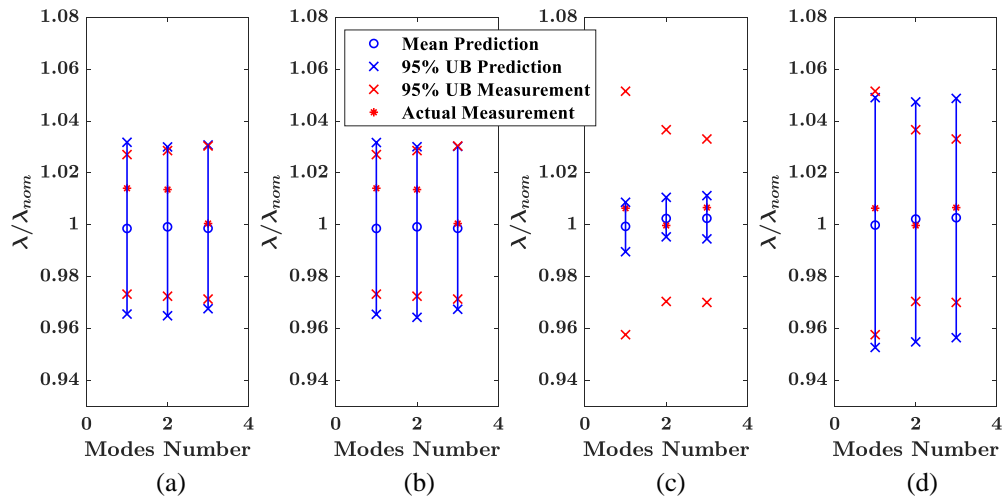


Fig. 2-5 Prediction of the ratio of modal frequencies over the nominal values obtained by propagating the uncertainty of (a) the structural parameters for case 2 (b) all parameters for Case 2 (c) the structural parameters for case 3 (d) all parameters for Case 3

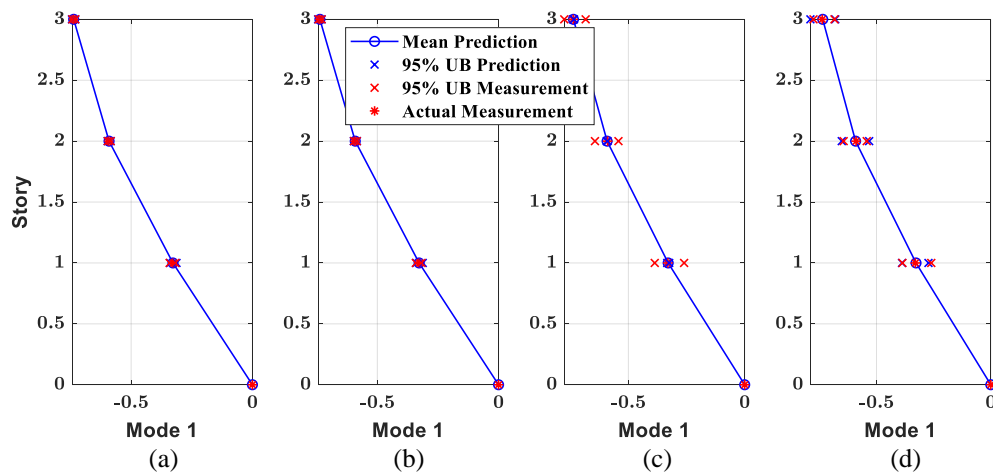


Fig. 2-6 Prediction of the mode shapes by propagating the uncertainty of (a) the structural parameters for case 2 (b) all parameters for case 2 (c) the structural parameters for case 3 (d) all parameters for case 3

The uncertainties are also propagated to predict response time histories. Model predictions of the acceleration and inter-story drift time histories are shown in Fig. 2-7 (case 2) and Fig. 2-8 (case 3). For simulated case 2, it can be seen that the UB are wide and contain the response predictions from the nominal model, as well as the measured acceleration corresponding to a new dataset generated from the measurement simulation procedure. It is also noted that the mean acceleration and the mean inter-story drifts are quite close to the

corresponding nominal responses since the hyper mean values of the structural parameters are well estimated. Neglecting the prediction error uncertainty in this case, the parameter uncertainty can be propagated to obtain reasonable predictions of modal properties and response time histories. Regarding the predicted 95% UB of the response time histories for the simulated case 3, the UB predicted for the case of no prediction error uncertainty are thin, as shown in Fig. 2-8(a), failing to contain the simulated response. On the contrary, as seen in Fig. 2-8(b), the larger 95% UB of the response time histories are obtained when the prediction error uncertainties are included in the propagation. These 95% UB contain the simulated response.

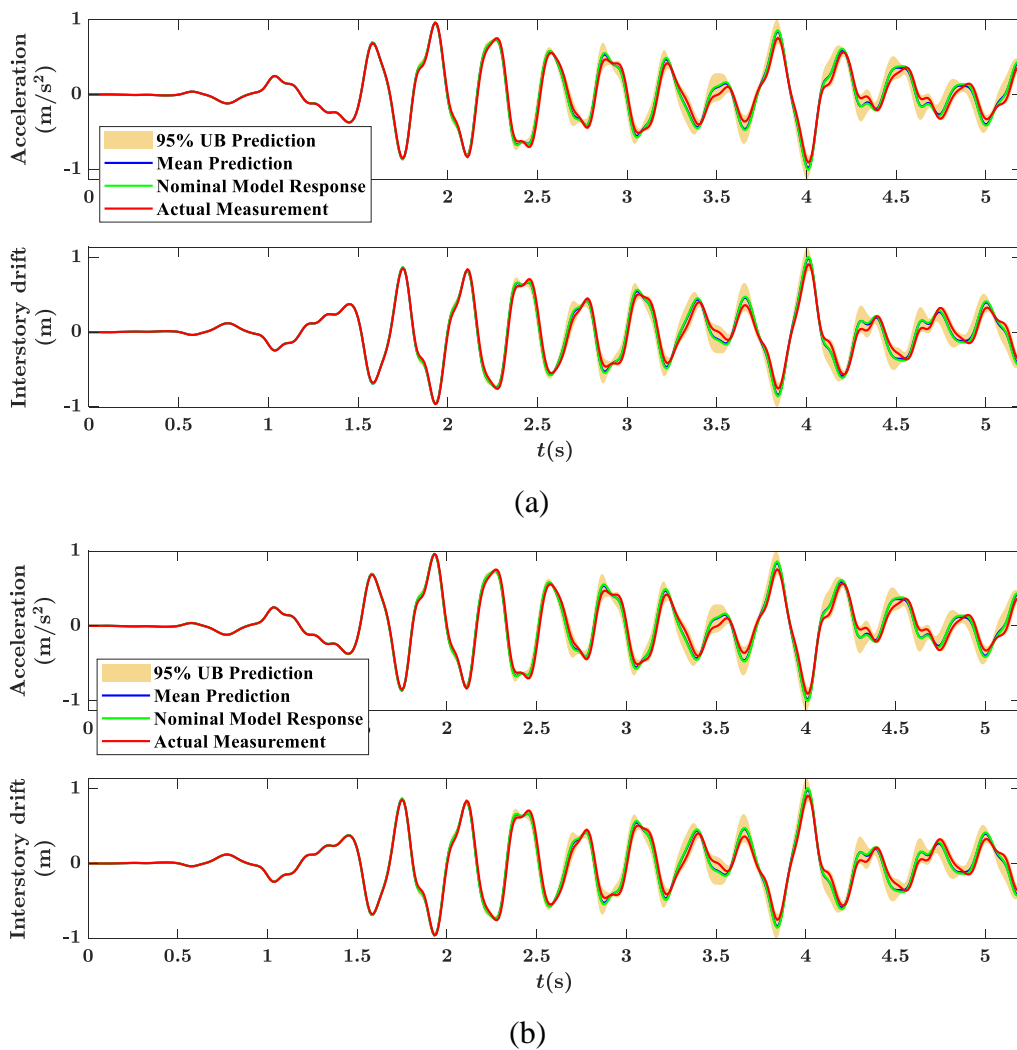


Fig. 2-7 Prediction of the response time histories of the third story for case 2 considering (a) only structural parameters uncertainty (b) total uncertainty

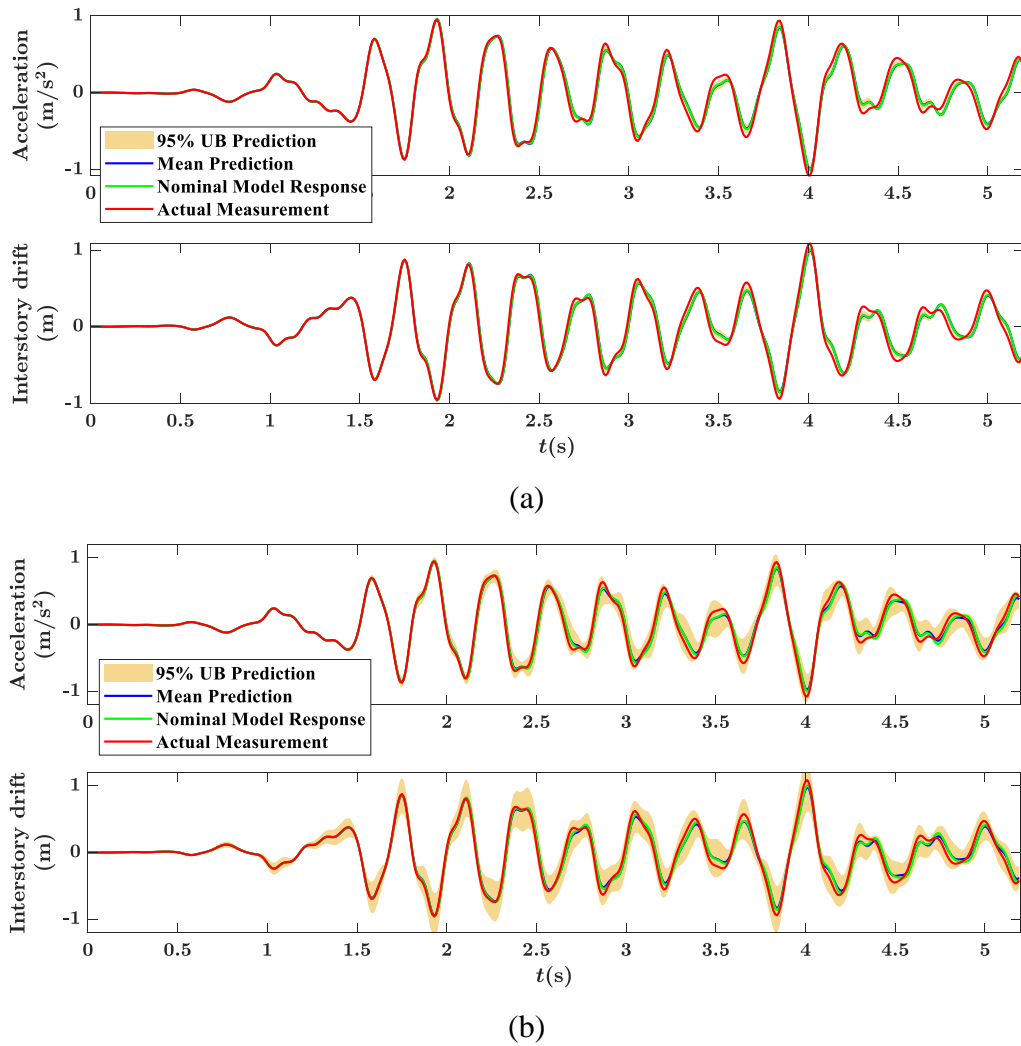


Fig. 2-8 Prediction of the response time histories of the third story for case 3 considering (a) only structural parameters uncertainty (b) total uncertainty

Next, the HBM and the CBM framework are compared for the simulated case 2. The uncertainties in the modal frequencies are displayed in Fig. 2-9 for the HBM and CBM. The posterior distributions obtained by these methods differ qualitatively. Uncertainties calculated by the HBM are correlated since they arise from the parameter uncertainties. In contrast, the uncertainties obtained by the CBM are uncorrelated due to the statistical independence of the prediction error terms.

In the context of earthquake engineering, the maximum inter-story drift is an important performance index. In this example, the parameter and prediction error uncertainties are propagated to compute the posterior predictive distribution of the drift based on Eq. (2.50). To this end, the Subset Simulation [48] is used to integrate over the multi-dimensional parameter space. Fig. 2-10(a) shows the posterior distributions obtained for the HBM and

CBM frameworks. Although the mean values of the maximum drift computed by both methods are approximately the same, the uncertainty estimated by the HBM is significantly greater than the uncertainty obtained by the CBM framework.

Additionally, the probability that the maximum drift exceeds a predefined threshold is calculated. For this purpose, the limit state function is defined as $g(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mathbf{e}) = \max_i \{d_i(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mathbf{e})\} - d_0$, where $d_i(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mathbf{e})$ is the i -th story drift and d_0 is the exceedance threshold. Fig. 2-10(b) shows the failure probability for both the HBM and CBM methods. The predictions given by the two methods differ substantially by one-to-three orders of magnitude, depending on the value of the exceedance levels. These large differences can be justified by the different predictions of the PDF for the maximum floor drift shown in Fig. 2-10(a).

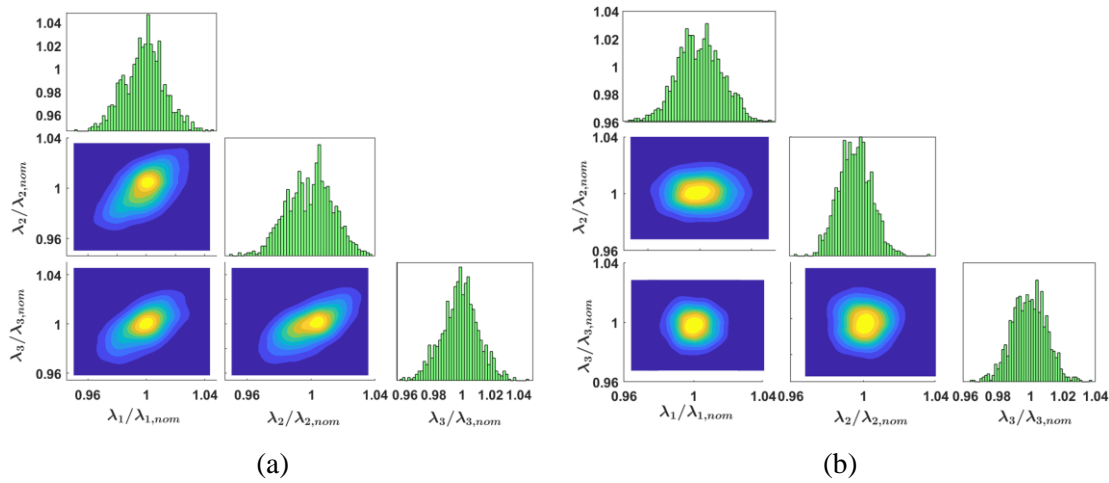


Fig. 2-9 Posterior predictive distributions of the modal frequencies normalized over the nominal values for the simulated Case 2 by using (a) HBM (b) CBM

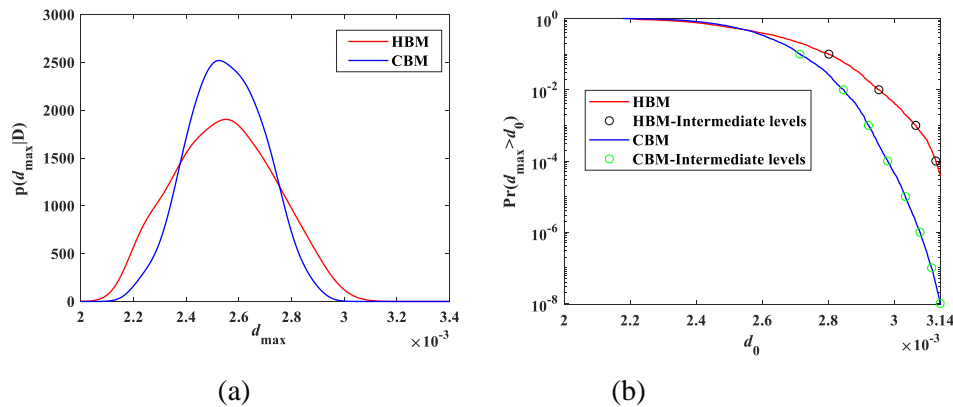


Fig. 2-10 (a) Posterior predictive distribution of the maximum drift in the simulated case 2 (b) the exceedance probability of the maximum drift

2.4.2 Experimental case study: 3-story prototype structure

An experimental 3-story prototype structure tested on a shaking table is employed to validate the proposed method, as shown in Fig. 2-11(a). The structure is excited with $N_D=19$ Gaussian white noise base excitations with the same intensity so that the excitation (operational) characteristics be identical. However, the white noise excitations differ in the time history details. Three accelerometers are placed on each floor to record the acceleration responses. The measurements are 120s long sampled at 0.005s intervals. Operational modal analysis methods are used [33,50] here to obtain the 19 datasets of modal properties. The structure is modeled as a 3-DOF shear building depicted in Fig. 2-11(b). The nominal values of the mass in each floor equals to $m_1=5.63\text{kg}$, $m_2=6.03\text{kg}$, $m_3=4.66\text{kg}$, respectively. The nominal values of the three springs are $k_1=20.88\text{kN/m}$, $k_2=22.37\text{kN/m}$, $k_3=24.21\text{kN/m}$. These nominal values are taken from the design plans and used to construct a simplified shear model of the structure. All the details of the structure properties, experimental setup, as well as the model assumptions can be found in [30].

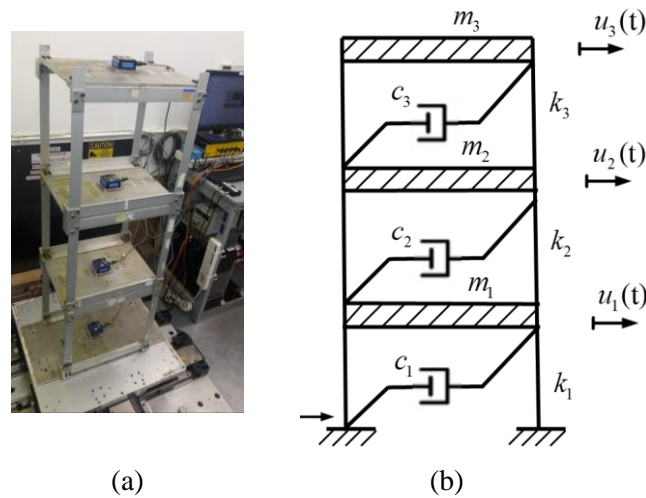


Fig. 2-11 (a) 3-Story prototype structure tested on a shaking table (b) 3-DOF linear numerical structural model

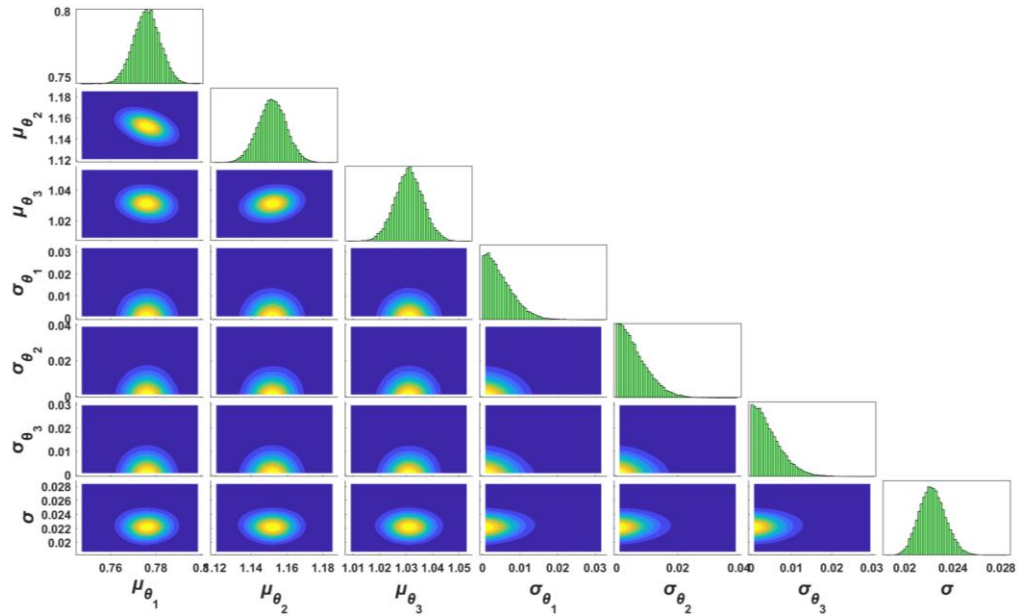


Fig. 2-12 Posterior distribution of hyper-parameters and prediction error parameter

The structural model is exactly the same as the one used for the simulated case, and so the same parameterization is employed. It should be noted that all 19 experiments are conducted under very similar environmental and operational conditions. So one would not expect variabilities in the structural model parameters due to changes in the environmental and operational conditions. The variabilities are mostly due to model error.

The mean estimates of the hyper-parameters and the prediction error variance are presented in Table 2-3 using approximations A-1 and A-2. The results are also compared with the corresponding mean estimates obtained using the full sampling technique. Additionally, the results are compared with those obtained using the CBM method. Based on this table, it can be seen that the A-2 estimates the hyper mean and covariance accurately while the A-1 fails to provide adequate accuracy for the covariance hyper-parameters. As expected, the CBM framework underestimates the uncertainties in the model parameters and the model and measurement errors are mostly captured by the prediction error term.

Table 2-3 Estimates of mean and standard deviation of the model parameters and prediction error parameters

Parameters	$\hat{\mu}_{\theta_1}$	$\hat{\sigma}_{\theta_1}$	$\hat{\mu}_{\theta_2}$	$\hat{\sigma}_{\theta_2}$	$\hat{\mu}_{\theta_3}$	$\hat{\sigma}_{\theta_3}$	$\hat{\sigma}$	$\hat{\Sigma}_{\sigma}$
A-1	0.7735	0.0068	1.1548	0.0084	1.0301	0.0063	0.0215	0.0010
A-2	0.7757	0.0040	1.1523	0.0041	1.0317	0.0041	0.0223	0.0012
FS	0.7769	0.0043	1.1524	0.0044	1.0309	0.0042	0.0224	0.0012
CBM	0.7755	0.0032	1.1526	0.0033	1.0317	0.0034	0.0226	0.0012

The joint posterior distribution of the hyper-parameters and the prediction error variance $\{\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \sigma^2\}$ is shown in Fig. 2-12 using the approximation A-2. It is observed that the MPV of the hyper mean are considerably different from the nominal values, indicating that the simplified structural model contains large modeling errors. Specifically, the first two stiffness parameters are close to 0.7757 and 1.1523, respectively. For the hyper standard deviations, the mean of the samples is small, which shows that the uncertainty in the model arising from the variability in the datasets is small. The identification uncertainty for each dataset is large due to large modeling errors. A significant amount of uncertainty, originated from the mismatch between the predicted and measured data, is captured through the prediction error term. This can be realized from the high values of the prediction error parameter ($\sigma \approx 0.022$). The results are similar to the ones obtained for the simulated case 3.

Fig. 2-13(a) shows the identification of the stiffness parameters as more datasets are included. The error bars indicate the MPV and the identification uncertainty obtained for each data set; the red line shows the estimated mean, and the colored area shows 99% UB. As shown, the UB initially decreases as the number of datasets increases, and this finding is consistent with the small variability of the MPV across datasets. However, as shown in Fig. 2-13(b), the UB tends to stabilize for larger values of the number of data sets.

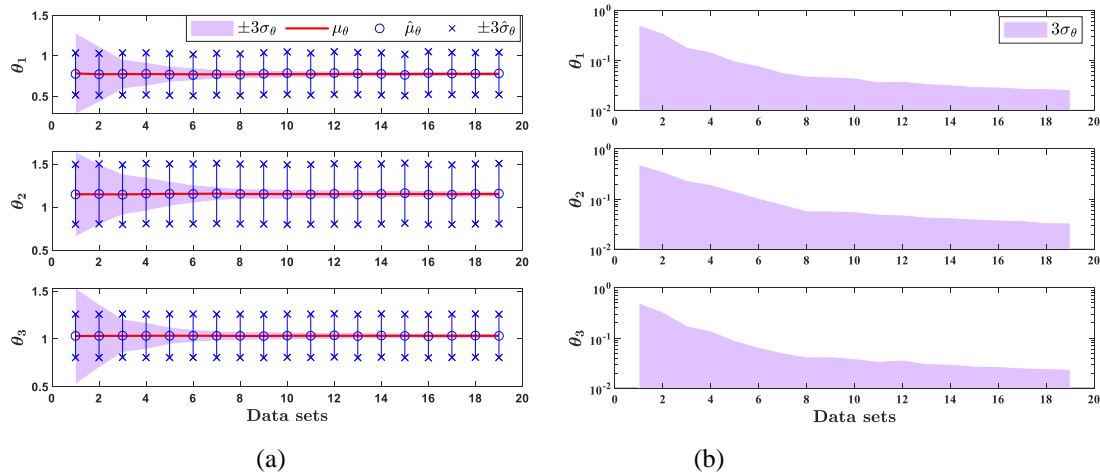


Fig. 2-13 (a) Uncertainty quantification of the stiffness parameters obtained using the approximation A-2 (b) magnified views of the UB on a logarithmic scale

Next, the uncertainties are propagated to estimate uncertainties in modal properties and response histories. The propagation with and without prediction error are presented separately.

Fig. 2-14 and Fig. 2-15 show the predicted modal properties. When the prediction error uncertainty is ignored, the 95% UB of the first and second modal frequencies fail to contain the UB obtained based on the variability across 19 datasets. On the contrary, when the prediction error uncertainty is considered, the UB well covers the variability of modal frequencies over datasets. Regarding the mode shapes, almost all components are not contained in the UB of predictions unless the prediction error uncertainty is incorporated.

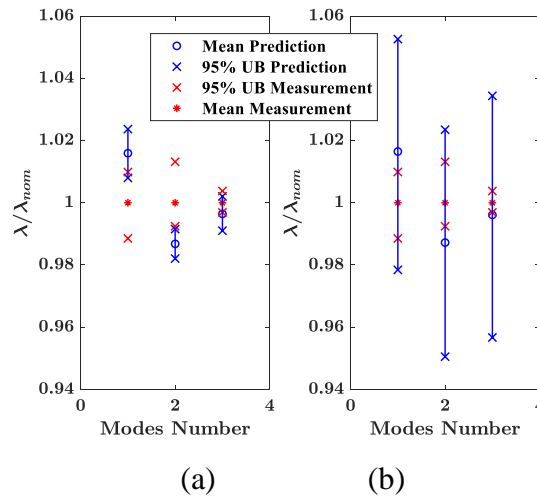


Fig. 2-14 Predictions of modal frequencies normalized over the nominal values considering (a) only structural parameters uncertainty, (b) both the structural parameters and prediction error uncertainties

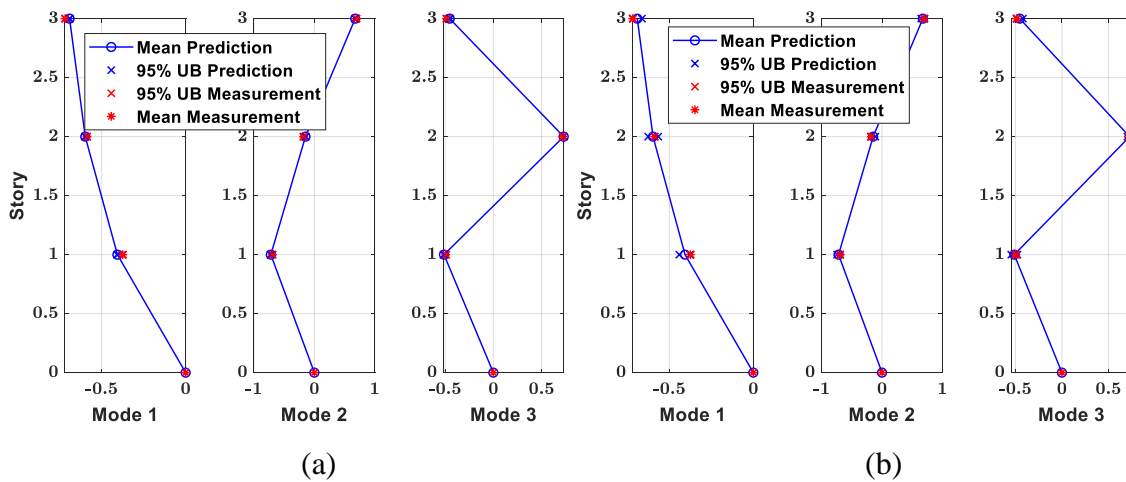


Fig. 2-15 Predicted mode shapes considering (a) only structural parameters uncertainty, (b) both structural and model error parameter uncertainties

Fig. 2-16 shows the predictions of acceleration response history when the structure is subjected to new base excitation. For this, the mean values of damping ratios corresponding to the three mode shapes are obtained as 2.39%, 0.87%, and 0.65%. The cases with and without prediction error uncertainty are shown separately in Fig. 2-16(a) and (b). The Zoomed views are included for a better comparison. Ignoring the prediction error terms, the 95% UB of the predicted acceleration is thin, failing to contain the actual response from one of the 19 datasets. However, the inclusion of the prediction error uncertainty results in a much larger UB, which almost contains the entire measured response. This result provides evidence as to the validity and effectiveness of the proposed approach.

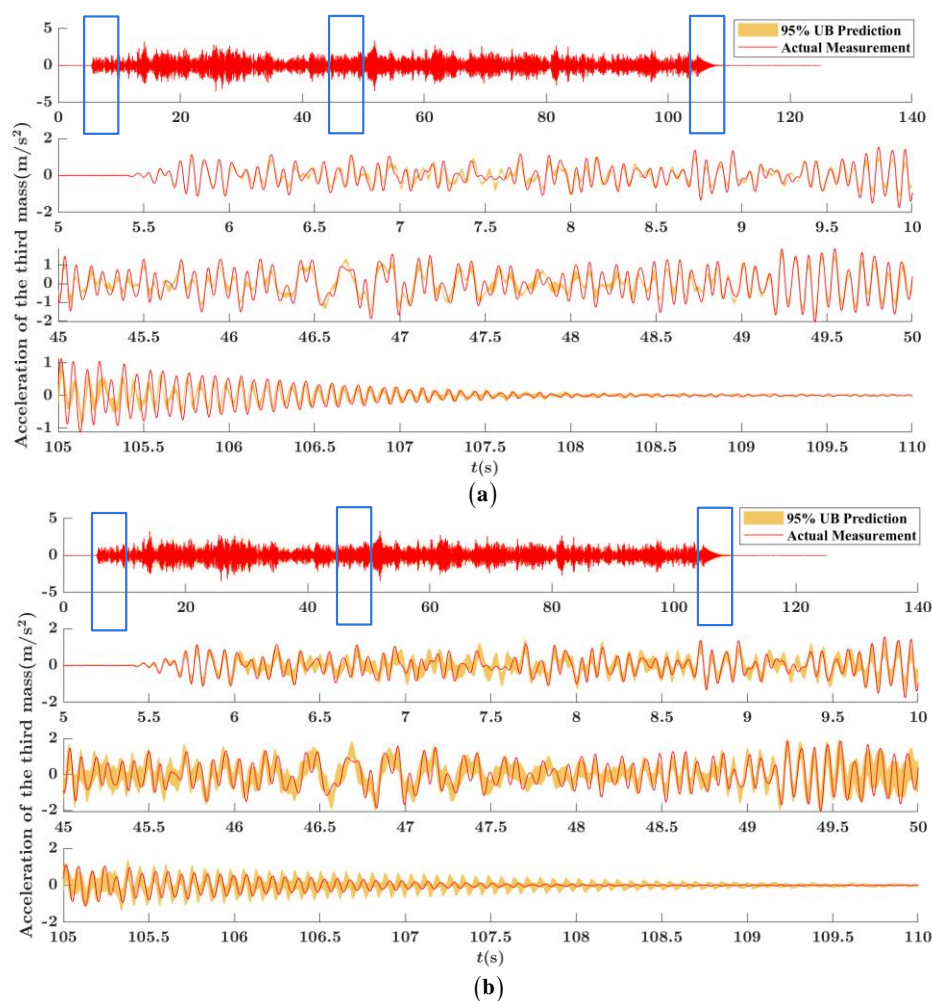


Fig. 2-16 Predicted acceleration response history of the third story considering (a) only structural parameters uncertainty, (b) both structural and prediction error uncertainties

2.5 Conclusion

A new HBM framework is developed for updating structural models based on modal test data. A remarkable feature of the proposed method lies in the capability to account for both the uncertainty of the structural and prediction error parameters, considering modeling errors, environmental effects, operational variabilities, and measurement noise. Two asymptotic approximations are proposed to simplify the likelihood function, differing in the formulation of Taylor series expansion with respect to the prediction error variance. These approximations are next used for investigating the asymptotic behavior of the HBM framework analytically, providing explicit formulations for the MPV of the hyper-parameters. Subsequently, the uncertainties are propagated for predicting modal frequencies, mode shape vectors, structural parameters, response histories, and engineering demand parameters. Ultimately, a reliability analysis perspective is integrated within the proposed framework to provide a robust estimation of the failure probability, considering both parameter and input uncertainties.

Two illustrative examples are employed to validate the proposed framework. The asymptotic approximations are examined by comparing the results with the actual values obtained from the full simulation approach. The HBM framework is also compared with the CBM framework to investigate the treatment of uncertainties. From the results of these examples, the following conclusions are drawn:

- The asymptotic approximation expanded based on the full likelihood function (A-2) is superior to the approximation built upon the individual likelihood function (A-1). The A-2 also demonstrates good accuracy and stability when compared with the sampling strategy. Thus, the A-2 is prescribed as an efficient and accurate methodology.
- The proposed framework reliably accounts for the variability of parameters across datasets, providing reasonable parameter and response uncertainties, whereas the CBM framework underestimates the uncertainties.
- The prediction of engineering demand parameters (e.g., inter-story drift) and their exceedance probabilities based on the HBM framework is more accurate and robust when compared with the CBM framework. This achievement is crucial to avoid misleading damage alarms.

Appendix A. Asymptotic approximation A-2

An alternative approximation for the likelihood function $p(D_i | \boldsymbol{\theta}_i, \sigma^2)$ is derived by viewing $L(\boldsymbol{\theta}_i, \sigma^2)$ in Eq. (2.7) as a function of $\boldsymbol{\theta}_i$ only and expanding the function with respect to the structural model parameters $\boldsymbol{\theta}_i$ about the MPV $\hat{\boldsymbol{\theta}}_i$, resulting in the likelihood function

$$\begin{aligned} p(D_i | \boldsymbol{\theta}_i, \sigma^2) &= \exp(-L(\boldsymbol{\theta}_i, \sigma^2)) \\ &= \exp(-L(\hat{\boldsymbol{\theta}}_i, \sigma^2) - \frac{1}{2}(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)^T \mathbf{H}_{\hat{\boldsymbol{\theta}}_i}(\sigma^2)(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i) + \dots) \\ &\approx \exp(-\frac{m(N_0+1)}{2} \ln \sigma^2 - \frac{m(N_0+1)}{2\sigma^2} J(\hat{\boldsymbol{\theta}}_i) - \frac{1}{2}(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)^T \mathbf{H}_{\hat{\boldsymbol{\theta}}_i}(\sigma^2)(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i) \quad (\text{A.1}) \\ &\propto (\sigma^2)^{-\frac{m(N_0+1)}{2}} \exp(-\frac{m(N_0+1)}{2\sigma^2} J(\hat{\boldsymbol{\theta}}_i)) \sqrt{|\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2)|} N(\boldsymbol{\theta}_i | \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2)) \end{aligned}$$

where $\hat{\boldsymbol{\theta}}_i$, obtained by minimizing the objective function $L(\boldsymbol{\theta}_i, \sigma^2)$ with respect to $\boldsymbol{\theta}_i$, is given by Eq. (2.9), independently of the value of σ^2 , and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2) = \mathbf{H}_{\hat{\boldsymbol{\theta}}_i}^{-1}(\sigma^2)$ is the inverse of the Hessian matrix $\mathbf{H}_{\hat{\boldsymbol{\theta}}_i}(\sigma^2)$ of $L(\boldsymbol{\theta}_i, \sigma^2)$ evaluated at the MPV $\hat{\boldsymbol{\theta}}_i$, given by

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2) &= \mathbf{H}_{\hat{\boldsymbol{\theta}}_i}^{-1}(\sigma^2) \equiv \left[\frac{\partial^2 L(\boldsymbol{\theta}_i, \sigma^2)}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_i^T} \bigg|_{\boldsymbol{\theta}_i = \hat{\boldsymbol{\theta}}_i} \right]^{-1} \\ &= \frac{2\sigma^2}{m(N_0+1)} \left[\frac{\partial^2 J(\boldsymbol{\theta}_i)}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_i^T} \bigg|_{\boldsymbol{\theta}_i = \hat{\boldsymbol{\theta}}_i} \right]^{-1} = \frac{2\sigma^2}{m(N_0+1)} \mathbf{H}_J^{-1}(\hat{\boldsymbol{\theta}}_i) \quad (\text{A.2}) \end{aligned}$$

Following the same procedure regarding the calculation of distribution $p(\sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | \mathbf{D})$ in the main text and replacing the likelihood function by Eq. (A.1), one can readily obtain the joint posterior distribution $p(\sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | \mathbf{D})$ in the form (2.27) where $T(\sigma^2)$ is defined in (2.28).

The MPV of parameters $\{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma^2\}$ can be derived by minimizing the negative logarithm of the joint distribution $p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma^2 | \mathbf{D})$, defined as $L_p = -\ln p(\sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. Assuming that the prior distributions $p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and $p(\sigma^2)$ are assigned uniform distributions, L_p can be expressed as:

$$\begin{aligned} L_p &= -\ln p(\sigma^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | \mathbf{D}) = -\ln T(\sigma^2) - \sum_{i=1}^{N_D} N(\boldsymbol{\mu}_0 | \hat{\boldsymbol{\theta}}_i, \boldsymbol{\Sigma}_0 + \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}(\sigma^2)) + c \\ &= \frac{m(N_0+1)N_D}{2} \ln \sigma^2 + \frac{m(N_0+1)N_D}{2\sigma^2} J_D + L_{p1} + L_{p2} + L_{p3} + c \quad (\text{A.3}) \end{aligned}$$

where L_{p1} , L_{p2} and L_{p3} are defined as:

$$\begin{aligned}
L_{p1} &= L_{p1}(\sigma^2) = -\frac{1}{2} \sum_{i=1}^{N_D} \ln |\hat{\Sigma}_{\theta_i}(\sigma^2)| \\
L_{p2} &= L_{p2}(\Sigma_{\theta}, \sigma^2) = \frac{1}{2} \sum_{i=1}^{N_D} \ln |\Sigma_{\theta} + \hat{\Sigma}_{\theta_i}(\sigma^2)| \\
L_{p3} &= L_{p3}(\mu_{\theta}, \Sigma_{\theta}, \sigma^2) = \frac{1}{2} \sum_{i=1}^{N_D} (\mu_{\theta} - \hat{\theta}_i)^T (\Sigma_{\theta} + \hat{\Sigma}_{\theta_i}(\sigma^2))^{-1} (\mu_{\theta} - \hat{\theta}_i)
\end{aligned} \tag{A.4}$$

Using the fact that

$$\frac{\partial \mathbf{a}^T (\mathbf{X} + \mathbf{Y})^{-1} \mathbf{a}}{\partial \mathbf{a}} = 2(\mathbf{X} + \mathbf{Y})^{-1} \mathbf{a} \tag{A.5}$$

the first derivative of the objective function L_p with respect to the parameters μ_{θ} can be computed as:

$$\frac{\partial L_p}{\partial \mu_{\theta}} = \frac{\partial L_{p3}}{\partial \mu_{\theta}} = \sum_{i=1}^{N_D} (\Sigma_{\theta} + \hat{\Sigma}_{\theta_i}(\sigma^2))^{-1} (\mu_{\theta} - \hat{\theta}_i) \tag{A.6}$$

Using the following identities for two square matrices \mathbf{X} and \mathbf{Y} ,

$$\frac{\partial \ln(|\mathbf{X} + \mathbf{Y}|)}{\partial \mathbf{X}} = ((\mathbf{X} + \mathbf{Y})^{-1})^T \tag{A.7}$$

$$\frac{\partial \mathbf{a}^T (\mathbf{X} + \mathbf{Y})^{-1} \mathbf{a}}{\partial \mathbf{X}} = -((\mathbf{X} + \mathbf{Y})^{-1} \mathbf{a} \mathbf{a}^T (\mathbf{X} + \mathbf{Y})^{-1})^T \tag{A.8}$$

the first derivative of the objective function L_p with respect to the parameters Σ_{θ} can be computed as

$$\begin{aligned}
\frac{\partial L_p}{\partial \Sigma_{\theta}} &= \frac{\partial L_{p2}}{\partial \Sigma_{\theta}} + \frac{\partial L_{p3}}{\partial \Sigma_{\theta}} \\
&= \frac{1}{2} \sum_{i=1}^{N_D} (\Sigma_{\theta} + \hat{\Sigma}_{\theta_i}(\sigma^2))^{-1} - \frac{1}{2} \sum_{i=1}^{N_D} (\Sigma_{\theta} + \hat{\Sigma}_{\theta_i}(\sigma^2))^{-1} (\mu_{\theta} - \hat{\theta}_i) (\mu_{\theta} - \hat{\theta}_i)^T (\Sigma_{\theta} + \hat{\Sigma}_{\theta_i}(\sigma^2))^{-1}
\end{aligned} \tag{A.9}$$

The first derivative of the objective function L_p with respect to the parameters σ^2 can be computed as:

$$\frac{\partial L_p}{\partial \sigma^2} = \frac{m(N_0 + 1)N_D}{2\sigma^2} - \frac{m(N_0 + 1)N_D}{2\sigma^4} J_D + \frac{\partial L_{p1}}{\partial \sigma^2} + \frac{\partial L_{p2}}{\partial \sigma^2} + \frac{\partial L_{p3}}{\partial \sigma^2} \tag{A.10}$$

Using the following identities for two square matrices $\mathbf{X}(t)$ and \mathbf{Y} and a vector \mathbf{a}

$$\frac{\partial \ln(|\mathbf{X}(t)|)}{\partial t} = \text{tr} \left((\mathbf{X}(t))^{-1} \frac{\partial \mathbf{X}(t)}{\partial t} \right) \tag{A.11}$$

$$\frac{\partial \ln(|\mathbf{X}(t) + \mathbf{Y}|)}{\partial t} = \text{tr} \left((\mathbf{X}(t) + \mathbf{Y})^{-1} \frac{\partial \mathbf{X}(t) + \mathbf{Y}}{\partial t} \right) \tag{A.12}$$

$$\frac{\partial \mathbf{a}^T (\mathbf{X}(t) + \mathbf{Y})^{-1} \mathbf{a}}{\partial t} = -tr \left((\mathbf{X}(t) + \mathbf{Y})^{-1} \frac{\partial \mathbf{X}(t) + \mathbf{Y}}{\partial t} (\mathbf{X}(t) + \mathbf{Y})^{-1} \mathbf{a} \mathbf{a}^T \right) \quad (\text{A.13})$$

the derivatives of L_{p1} , L_{p2} and L_{p3} with respect to σ^2 can be computed as:

$$\begin{aligned} \frac{\partial L_{p1}}{\partial \sigma^2} &= -\frac{1}{2\sigma^2} N_\theta \\ \frac{\partial L_{p2}}{\partial \sigma^2} &= \frac{1}{m(N_0 + 1)} \sum_{i=1}^{N_D} tr \left((\mathbf{\Sigma}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2))^{-1} (\mathbf{H}_J(\hat{\boldsymbol{\theta}}_i))^{-1} \right) \\ &= \frac{1}{2\sigma^2} \sum_{i=1}^{N_D} tr \left((\mathbf{\Sigma}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2))^{-1} \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2) \right) \\ \frac{\partial L_{p3}}{\partial \sigma^2} &= -\frac{1}{2\sigma^2} \sum_{i=1}^{N_D} tr \left((\mathbf{\Sigma}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2))^{-1} \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2) (\mathbf{\Sigma}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2))^{-1} (\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i)(\boldsymbol{\mu}_\theta - \hat{\boldsymbol{\theta}}_i)^T \right) \end{aligned} \quad (\text{A.14})$$

where N_θ is the number of the model parameters $\boldsymbol{\theta}$. Therefore, the MPVs $\bar{\boldsymbol{\mu}}_\theta$, $\bar{\mathbf{\Sigma}}_\theta$ and $\bar{\sigma}^2$ of parameters $\boldsymbol{\mu}_\theta$, $\mathbf{\Sigma}_\theta$, and σ^2 satisfy the following equations:

$$\begin{cases} \sum_{i=1}^{N_D} (\bar{\mathbf{\Sigma}}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\bar{\sigma}^2))^{-1} (\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i) = 0 \\ \sum_{i=1}^{N_D} (\bar{\mathbf{\Sigma}}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\bar{\sigma}^2))^{-1} - \sum_{i=1}^{N_D} (\bar{\mathbf{\Sigma}}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\bar{\sigma}^2))^{-1} (\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i)(\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i)^T (\bar{\mathbf{\Sigma}}_\theta + \hat{\mathbf{\Sigma}}_{\theta_i}(\bar{\sigma}^2))^{-1} = 0 \\ \frac{m(N_0 + 1)N_D}{2\bar{\sigma}^2} - \frac{m(N_0 + 1)N_D}{2\bar{\sigma}^4} J_D + \left. \frac{\partial L_{p1}}{\partial \sigma^2} \right|_{\bar{\sigma}^2} + \left. \frac{\partial L_{p2}}{\partial \sigma^2} \right|_{\bar{\mathbf{\Sigma}}_\theta, \bar{\sigma}^2} + \left. \frac{\partial L_{p3}}{\partial \sigma^2} \right|_{\bar{\boldsymbol{\mu}}_\theta, \bar{\mathbf{\Sigma}}_\theta, \bar{\sigma}^2} = 0 \end{cases} \quad (\text{A.15})$$

Similarly, when the identification uncertainty $\hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2)$ over all datasets are assumed to be equal to a covariance matrix $\mathbf{\Sigma}$ taken to be $\frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\mathbf{\Sigma}}_{\theta_i}(\sigma^2)$, the average of the identification uncertainty of each dataset, the MPV of parameters $\boldsymbol{\mu}_\theta$ and $\mathbf{\Sigma}_\theta$ can be solved based on the first two formulas in Eq. (A.15), given as:

$$\begin{aligned} \bar{\boldsymbol{\mu}}_\theta &= \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\boldsymbol{\theta}}_i \\ \bar{\mathbf{\Sigma}}_\theta &= \mathbf{\Sigma}_v - \mathbf{\Sigma} \end{aligned} \quad (\text{A.16})$$

where the variability uncertainty and the identification uncertainty are defined as:

$$\begin{aligned} \mathbf{\Sigma}_v &= \frac{1}{N_D} \sum_{i=1}^{N_D} (\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i)(\bar{\boldsymbol{\mu}}_\theta - \hat{\boldsymbol{\theta}}_i)^T \\ \mathbf{\Sigma} &= \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\mathbf{\Sigma}}_{\theta_i}(\bar{\sigma}^2) \end{aligned} \quad (\text{A.17})$$

Substituting Eq. (A.17) to $\left. \frac{\partial L_{p2}}{\partial \sigma^2} \right|_{\bar{\Sigma}_0, \bar{\sigma}^2}$ and $\left. \frac{\partial L_{p3}}{\partial \sigma^2} \right|_{\bar{\mu}_0, \bar{\Sigma}_0, \bar{\sigma}^2}$ yields:

$$\begin{aligned} \left. \frac{\partial L_{p2}}{\partial \sigma^2} \right|_{\bar{\Sigma}_0, \bar{\sigma}^2} &= \frac{1}{m(N_0 + 1)} \sum_{i=1}^{N_D} \text{tr} \left((\Sigma_v)^{-1} \Sigma \right) \\ \left. \frac{\partial L_{p3}}{\partial \sigma^2} \right|_{\bar{\Sigma}_0, \bar{\sigma}^2} &= -\frac{1}{m(N_0 + 1)} \sum_{i=1}^{N_D} \text{tr} \left((\Sigma_v)^{-1} \Sigma \right) \end{aligned} \quad (\text{A.18})$$

Subsequently, the last formula in Eq. (A.15) can be rewritten as:

$$\frac{m(N_0 + 1)N_D}{2\bar{\sigma}^2} - \frac{m(N_0 + 1)N_D}{2\bar{\sigma}^4} J_D - \frac{1}{2\bar{\sigma}^2} N_\theta = 0 \quad (\text{A.19})$$

Then the MPV of parameter σ^2 can be solved as:

$$\bar{\sigma}^2 = \frac{m(N_0 + 1)}{m(N_0 + 1) - N_\theta} J_D \quad (\text{A.20})$$

which gives $\bar{\sigma}^2 = J_D$ for $m(N_0 + 1)$ sufficiently larger than N_θ .

2.6 References

- [1] M.I. Friswell, Model updating in structural dynamics: a survey, *Journal of Sound and Vibration*. 167 (1993) 347–375.
- [2] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, *Journal of Engineering Mechanics*. 124 (1998) 455–461.
- [3] K.V. Yuen, *Bayesian Methods for Structural Dynamics and Civil Engineering*, 2010.
- [4] J.L. Beck, Bayesian system identification based on probability logic, *Structural Control and Health Monitoring*. 17(7) (2010) 825–847.
- [5] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, *Mechanical Systems and Signal Processing*. 56 (2015) 123–149.
- [6] J.L. Beck, K.V. Yuen, Model Selection Using Response Measurements: Bayesian Probabilistic Approach, *Journal of Engineering Mechanics*. 130. No. 2 (2004) 192–203.
- [7] M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, *Journal of Vibration and Control*. 14 (2008) 7–34.
- [8] K.V. Yuen, Recent developments of Bayesian model class selection and applications in civil engineering, *Structural Safety*. 32 (2010) 338–346.

- [9] K.-V. Yuen, L.S. Katafygiotis, Bayesian modal updating using complete input and incomplete response noisy measurements, *Journal of Engineering Mechanics*. 128 (2002) 340–350.
- [10] W.-J. Yan, L.S. Katafygiotis, A novel Bayesian approach for structural model updating utilizing statistical modal information from multiple setups, *Structural Safety*. 52 (2015) 260–271.
- [11] C. Papadimitriou, J.L. Beck, L.S. Katafygiotis, Updating robust reliability using structural test data, *Probabilistic Engineering Mechanics*. 16 (2001).
- [12] J.L. Beck, A.A. Taflanidis, Prior and Posterior Robust Stochastic Predictions for Dynamical Systems Using Probability Logic, *International Journal for Uncertainty Quantification*. 3 (2013) 271–288.
- [13] S.K. Au, F.L. Zhang, Fundamental two-stage formulation for Bayesian system identification, Part I: General theory, *Mechanical Systems and Signal Processing*. 66–67 (2016) 31–42.
- [14] I. Behmanesh, S. Yousefianmoghadam, A. Nozari, B. Moaveni, A. Stavridis, Uncertainty quantification and propagation in dynamic models using ambient vibration measurements, application to a 10-story building, *Mechanical Systems and Signal Processing*. 107 (2018) 502–514.
- [15] S. Alampalli, Effects of testing, analysis, damage, and environment on modal parameters, *Mechanical Systems and Signal Processing*. 14 (2000) 63–74.
- [16] J.F. Clinton, S.C. Bradford, T.H. Heaton, J. Favela, The observed wander of the natural frequencies in a structure, *Bulletin of the Seismological Society of America*. 96 (2006) 237–257.
- [17] H.H. Khodaparast, J.E. Mottershead, M.I. Friswell, Perturbation methods for the estimation of parameter variability in stochastic model updating, *Mechanical Systems and Signal Processing*. 22 (2008) 1751–1773.
- [18] P. Moser, B. Moaveni, Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge, *Mechanical Systems and Signal Processing*. 25 (2011) 2336–2357.
- [19] L.D. Avendaño-Valencia, E.N. Chatzi, Multivariate GP-VAR models for robust structural identification under operational variability, *Probabilistic Engineering Mechanics*. 60 (2020) 103035.

- [20] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, *Mechanical Systems and Signal Processing*. 64–65 (2015) 360–376.
- [21] S. Wu, P. Angelikopoulos, C. Papadimitriou, R. Moser, P. Koumoutsakos, A hierarchical Bayesian framework for force field selection in molecular dynamics simulations, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 374 (2016) 20150032.
- [22] X.-W. Liu, D.-G. Lu, P.C.J. Hoogenboom, Hierarchical Bayesian fatigue data analysis, *International Journal of Fatigue*. 100 (2017) 418–428.
- [23] S. Wu, P. Angelikopoulos, J.L. Beck, P. Koumoutsakos, Hierarchical Stochastic Model in Bayesian Inference for Engineering Applications: Theoretical Implications and Efficient Approximation, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*. 5 (2019).
- [24] S. Wu, P. Angelikopoulos, G. Tauriello, C. Papadimitriou, P. Koumoutsakos, Fusing heterogeneous data for the calibration of molecular dynamics force fields using hierarchical Bayesian models, *The Journal of Chemical Physics*. 145 (2016) 244112.
- [25] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, *Journal of Sound and Vibration*. 374 (2016) 92–110.
- [26] M. Song, B. Moaveni, C. Papadimitriou, A. Stavridis, Accounting for amplitude of excitation in model updating through a hierarchical Bayesian approach: Application to a two-story reinforced concrete building, *Mechanical Systems and Signal Processing*. 123 (2019) 68–83.
- [27] C.M. Carlo, Markov chain monte carlo and gibbs sampling, *Lecture Notes for EEB*. 581 (2004).
- [28] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Modeling error estimation and response prediction of a 10-story building model through a hierarchical bayesian model updating framework, *Frontiers in Built Environment*. 5 (2019).
- [29] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Accounting for Modeling Errors and Inherent Structural Variability through a Hierarchical Bayesian Model Updating Approach: An Overview, *Sensors*. 20 (2020) 3874.
- [30] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Probabilistic hierarchical Bayesian framework for time-domain model updating and robust predictions, *Mechanical Systems and Signal Processing*. 123 (2019) 648–673.

- [31] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Data-driven uncertainty quantification and propagation in structural dynamics through a hierarchical Bayesian framework, *Probabilistic Engineering Mechanics*. 60 (2020) 103047.
- [32] O. Sedehi, L.S. Katafygiotis, C. Papadimitriou, Hierarchical Bayesian operational modal analysis: Theory and computations, *Mechanical Systems and Signal Processing*. 140 (2020) 106663.
- [33] S. Au, *Operational Modal Analysis*, Springer, 2017.
- [34] Y.-C. Zhu, S.-K. Au, Bayesian operational modal analysis with asynchronous data, Part II: Posterior uncertainty, *Mechanical Systems and Signal Processing*. 98 (2018) 920–935.
- [35] Y.C. Zhu, S.K. Au, Bayesian data driven model for uncertain modal properties identified from operational modal analysis, *Mechanical Systems and Signal Processing*. 136 (2020) 106511.
- [36] H. Jensen, C. Papadimitriou, Bayesian Finite Element Model Updating, in: *Sub-Structure Coupling for Dynamic Analysis*, Springer, 2019: pp. 179–227.
- [37] E. Ntotsios, C. Papadimitriou, Multi-objective optimization algorithms for finite element model updating, in: *23rd International Conference on Noise and Vibration Engineering 2008, ISMA 2008*, 2008: pp. 1895–1909.
- [38] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern classification*, John Wiley & Sons, 2012.
- [39] P. a Bromiley, *Products and Convolutions of Gaussian Probability Density Functions*, *Tina Memo*. (2003) No. 2003-003.
- [40] B.P. Carlin, S. Chib, Bayesian model choice via Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society: Series B (Methodological)*. 57 (1995) 473–484.
- [41] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, *Journal of Engineering Mechanics*. 133 (2007) 816–832.
- [42] S. Wu, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian annealed sequential importance sampling: an unbiased version of transitional Markov chain Monte Carlo, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*. 4 (2018).
- [43] M. Coughlin, N. Christensen, J. Gair, S. Kandhasamy, E. Thrane, Method for estimation of gravitational-wave transient model parameters in frequency-time maps, *Classical and Quantum Gravity*. 31 (2014).

- [44] H.A. Jensen, Structural optimization of linear dynamical systems under stochastic excitation: A moving reliability database approach, *Computer Methods in Applied Mechanics and Engineering*. 194 (2005) 1757–1778.
- [45] H. Jensen, C. Papadimitriou, *Sub-structure Coupling for Dynamic Analysis: Application to Complex Simulation-Based Problems Involving Uncertainty* (Vol. 89). Springer, 2019.
- [46] S.K. Au, J.L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, *Probabilistic Engineering Mechanics*. 16 (2001) 263–277.
- [47] S.K. Au, J. Ching, J.L. Beck, Application of subset simulation methods to reliability benchmark problems, *Structural Safety*. 29 (2007) 183–193.
- [48] I. Papaioannou, W. Betz, K. Zwirgmaier, D. Straub, MCMC algorithms for Subset Simulation, *Probabilistic Engineering Mechanics*. 41 (2015) 89–103.
- [49] D. Patsialis, A.P. Kyprioti, A.A. Taflanidis, Bayesian calibration of hysteretic reduced order structural models for earthquake engineering applications, *Engineering Structures*. 224 (2020) 111204.
- [50] S.K. Au, F.L. Zhang, Y.C. Ni, Bayesian operational modal analysis: Theory, computation, practice, *Computers and Structures*. 126 (2013) 3–14.

Chapter 3. Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework

Original Paper:

X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework, *Computer Methods in Applied Mechanics and Engineering*. Submitted. (2021). <https://doi.org/10.5281/zenodo.5520607>.

ABSTRACT

A new time-domain probabilistic technique based on hierarchical Bayesian modeling (HBM) framework is proposed for calibration and uncertainty quantification of hysteretic type nonlinearities of dynamical systems. Specifically, probabilistic hyper models are introduced respectively for material hysteretic model parameters as well as prediction error variance parameters, aiming to consider both the uncertainty of the model parameters as well as the prediction error uncertainty due to unmodelled dynamics. A new asymptotic approximation is developed to simplify the process of nonlinear model updating and substantially reduce the computational burden of the HBM framework. Two numerical examples are conducted to verify the accuracy and performance of the proposed method considering Bouc-Wen (BW) hysteretic type nonlinearities. Model error is manifested as uncertainty due to variability in the measured data from multiple datasets. Results from a five-story numerical structure indicate that the model error is the main source of error that can affect the uncertainty in the model parameters due to the variability in the experimental data. It is also demonstrated that the parameter uncertainty due to the variability arising from model error depends on the sensor locations. It is shown that the proposed approach is robust for not only quantifying uncertainties of structural parameters and prediction error parameters, but also predicting the system quantities of interests (QoI) with reasonable accuracy and providing reliable

uncertainty bounds, as opposed to the conventional Bayesian approach which often severely underestimates the uncertainty bounds.

3.1 Introduction

Finite element (FE) models are extensively employed for representing structural systems and predicting their responses to dynamic loads [1–4]. Discrepancies between the predicted responses from FE models and the measured responses from the physical structures are often inevitable. To achieve a more authentic model, model updating has received considerable attentions in recent decades using deterministic (e.g. [5–8]) and probabilistic approaches (e.g. [9–12]). Updating linear models has been widely applied and is shown to achieve a great progress in the field of structural dynamics [6,9,13]. However, most physical structures are inherently characterized by nonlinear behaviours with higher uncertainties when subjected to large loads due to material and/or geometric nonlinearities. Linear systems often neglect such nonlinearities and therefore cause a considerably large modeling error between the real structures and the updated models. Characterizations of such nonlinearities may provide more information for accurate and efficient representations of real structures. To this end, updating nonlinear models is essential for accurate response and reliability predictions [14–16] and assessment of structures subjected to large loads such as earthquakes [17–19].

The core idea of model updating techniques is to find the most probable values of the structural model parameters by minimizing the difference between the FE-predicted and the measured responses [20]. The non-probabilistic approaches, or refereed as deterministic methods, can be applied to address such problems. Several studies have already demonstrated a good performance of the deterministic strategies [21,22]. However, a common shortcoming exists in those methods as well. Although the most plausible values of the model parameters can be estimated, the effect of the parameter uncertainties is often neglected. Such uncertainties can be arisen from the model error, the measurement noise or the changing environmental and ambient conditions. Quantifying the uncertainties is necessary for understanding the statistical characteristics of model parameters as well as propagating those uncertainties to predict the system quantities of interests (QoI). A remedy for the deterministic approaches is to apply the probabilistic means in model updating process. Due to their rigorous probabilistic framework, Bayesian inference methods have been widely used for quantifying and propagating the uncertainties in model updating [23–27]. Several contributions based on Bayesian strategies have already been proposed for nonlinear model

updating. Muto and Beck developed a Bayesian updating method for estimating the hysteretic material model parameters using stochastic simulation [28]. Ebrahimian et al. presented a framework for damage identification of dynamical structures with material nonlinearities using batch Bayesian estimations [29]. Song et al. proposed a Bayesian model updating methodology for dynamical systems with geometric nonlinearities based on the nonlinear normal modes extracted from broadband vibration data [30]. Ceravolo et al. employed a Bayesian uncertainty quantification framework for the identification of hysteretic parameters with consideration of the model discrepancy in seismic structural health monitoring [31]. More investigations for nonlinear model updating based on the Bayesian techniques can be found in the literature [32–39].

Bayesian inference provides a powerful probabilistic tool for updating nonlinear models and handling the uncertainties of nonlinear model parameters. However, the conventional Bayesian inference framework cannot properly account for an underlying variability in model parameters and uncertainties arising from multiple data sets under different excitations, operational, environmental and experimental conditions. The variability in the model parameters can originate from the presence of model and experimental error [40]. The uncertainty of the model parameters due to these variabilities is irreducible, in contrast to the identification uncertainty which is usually inversely proportional to the amount of data considered in a data set. To properly account for this irreducible variability, uncertainties can be embedded in the model parameters by assigning a parameterized hyper prior distribution. The hyper parameters in this distribution are assumed to be unknown quantities and to be estimated from the available multiple data sets.

A hierarchical Bayesian modeling (HBM) framework has recently been introduced in various engineering fields [41–47]. In the field of structural dynamics, it was initially proposed by Behmanesh et al. for structural identification based upon a mainly full simulation HBM approach [42]. For improving the efficiency of the HBM framework, Sedehi et al. and Jia et al. respectively developed asymptotic approximation-included HBM approaches for model updating based on time-domain [48,49] and frequency-domain linear models [50]. Patsialis et al. applied the HBM framework for reduced order structural models in earthquake engineering [51]. Such improvements have been successfully applied in updating linear models.

In this paper, a new time-domain hierarchical Bayesian modeling framework is developed for the identification of nonlinear models, aiming to quantifying the uncertainties of the nonlinear model parameters and the prediction error parameters, and further propagating the

overall uncertainties to the system output QoI. The contributions of this work are in the following aspects. A methodology is proposed for nonlinear model updating using response time history data which can characterize the nonlinear behaviors in the real structures. More importantly, the proposed methodology can capture the uncertainties due to model error by embedding uncertainties in the model and prediction error parameters. This is achieved by assigning a parameterized distribution in these parameters with hyper parameters to be estimated from the multiple datasets. Moreover, the presented methodology adopts a novel asymptotic approximation approach which can significantly improve the computational efficiency of the HBM framework.

The paper is organized as follows. Section 3.2 presents the detailed mathematical formulation of the proposed HBM framework and the approximations used. Section 3.3 applies the HBM framework to identify parameters of a nonlinear system with nonlinearities modelled by the Bouc-Wen hysteresis law. Section 3.4 provides two numerical examples to demonstrate the effectiveness of the proposed approach to account for uncertainties due to model error. Section 3.5 reports the conclusions of this study.

3.2 Hierarchical Bayesian Nonlinear Model Updating

3.2.1 Proposed hierarchical models

Suppose that N_D data sets of measured vibration time histories $\mathbf{D} = \{D_i, i = 1, \dots, N_D\}$ subjected to N_D known input loadings are available from a nonlinear structure. Let $D_i = \{\hat{\mathbf{Y}}_i(j) \in R^{N_0}, j = 1, 2, \dots, N_i\}$ be the i -th experimental data set consisting of a sequence of response data measured at N_0 degrees of freedom (DOF), where the notation j corresponds to a time instant $t_j = j\Delta t_i$ and N_i is the number of the sampled data using the sampling rate Δt_i . Let also $\mathbf{X}_i = \{\mathbf{X}_i(j), j = 1, 2, \dots, N_i\}$ be the i -th input loading which corresponds to the i -th data set D_i . Consider a parameterized class of nonlinear model $M(\boldsymbol{\theta})$ that is used to characterize the nonlinear behavior (e.g. material nonlinearities) of dynamical systems, where $\boldsymbol{\theta} \in R^{N_0}$ is the set of material/structural model parameters to be estimated using the measured response time histories, and N_0 is the total number of the unknown parameters in the set $\boldsymbol{\theta}$. Let also $\mathbf{g}_i(\boldsymbol{\theta}) = \{\mathbf{g}_i(j; \boldsymbol{\theta}) \in R^{N_s}, j = 1, 2, \dots, N_i\}$ be the response time histories under the same input loading \mathbf{X}_i and the same sampling rate Δt_i .

predicted from the model $M(\boldsymbol{\theta})$, where N_s denotes the number of DOFs. Subsequently, the discrepancy $\boldsymbol{\varepsilon}_i$ between the i -th experimental data set and the i -th predicted response time histories can be defined based on the prediction error equation:

$$\boldsymbol{\varepsilon}_i = \hat{\mathbf{Y}}_i(j) - L\mathbf{g}_i(j; \boldsymbol{\theta}) \quad (3.1)$$

where L is a selection matrix, usually containing elements of zeroes and ones, that associates the DOF of the model with the measured DOF.

In the analysis that follows, the prediction error is modelled by a random vector with Gaussian probability distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}$. A probabilistic model $p(\boldsymbol{\Sigma})$ is used herein to describe the uncertainty of the prediction error term. Realization of the prediction error parameters is free to vary across the different data sets, with the realization $\boldsymbol{\Sigma}_i$ that corresponds to the data set D_i considered to be an independent sample of the distribution $p(\boldsymbol{\Sigma})$. For the i -th data set D_i , $\boldsymbol{\Sigma}_i$ is assumed as a diagonal matrix

with l -th diagonal element $(\sigma_i a_{i,l})^2$, where $a_{i,l} = \sqrt{\frac{1}{N} \sum_{j=1}^N \mathbf{g}_{i,l}^2(j; \boldsymbol{\theta})}$ denotes the intensity of the

model predictions in the l -th DOF corresponding to the i -th data set. Thus, the probabilistic model $p(\boldsymbol{\Sigma})$ can be equally represented as the probabilistic model $p(\sigma^2)$, and it is modelled by the inverse gamma (IG) distribution given by:

$$p(\sigma^2) = \text{IG}(\sigma^2 | \lambda_1, \lambda_2) = \frac{(\lambda_2)^{\lambda_1}}{\Gamma(\lambda_1)} (\sigma^2)^{-\lambda_1-1} \exp\left(-\frac{\lambda_2}{\sigma^2}\right) \quad (3.2)$$

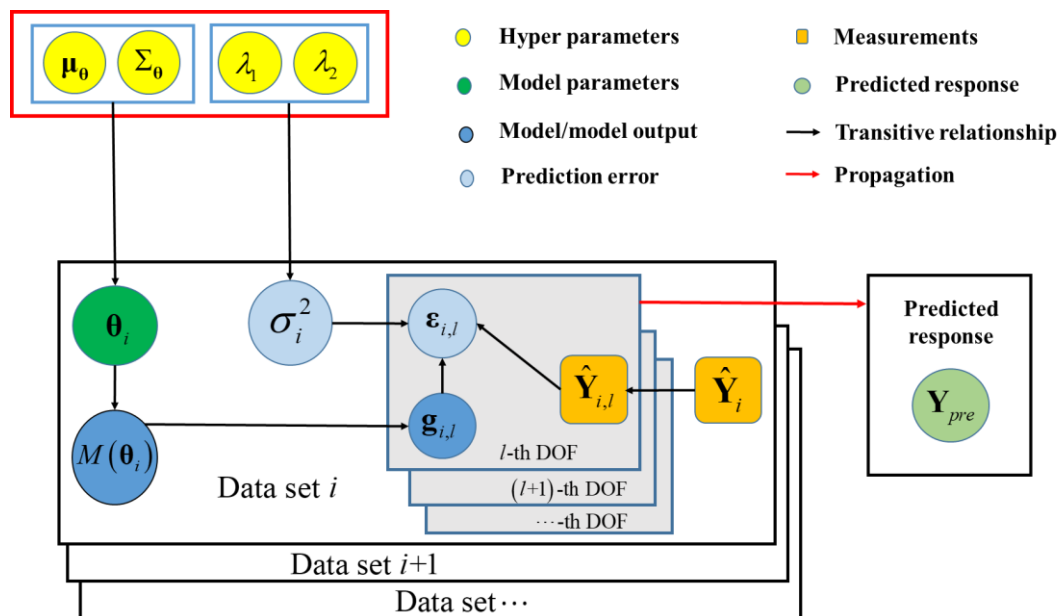
where the parameters λ_1 and λ_2 are the shape hyper-parameter and scale hyper-parameter, respectively.

Similarly, the uncertainty of model parameters $\boldsymbol{\theta}$ is probabilistically modeled using a Gaussian distribution [42,48]:

$$p(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) = N(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) \quad (3.3)$$

with unknown mean $\boldsymbol{\mu}_\theta \in R^{N_\theta}$ and covariance matrix $\boldsymbol{\Sigma}_\theta \in R^{N_\theta \times N_\theta}$. The parameter set $\{\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta\}$ is considered to be an uncertain hyper parameter set to be estimated using the available multiple data sets. Realization of $\boldsymbol{\theta}$ from the Gaussian distribution $N(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ can vary across the different data set, where the realization $\boldsymbol{\theta}_i$ is considered to be an independent sample of the distribution $N(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ that corresponds to the data set D_i .

This constitutes a hierarchy model that has two classes of parameters. The first class of model parameters comprises the $N_D(N_\theta + 1)$ experiment-specific parameters $\boldsymbol{\theta}_i$ and σ_i^2 , $i = 1, \dots, N_D$, while the second class model parameters comprises at most $N_\theta + 2 + N_\theta(N_\theta + 1)$ hyper-parameters involved in $\boldsymbol{\mu}_\theta$, $\boldsymbol{\Sigma}_\theta$, λ_1 and λ_2 . The number of parameters in the first set increases linearly with the number of datasets making the parameter estimation problem challenging when the number of data sets increases. The full set Θ of all parameters to be identified is $\Theta = \{\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \{\sigma_i^2\}_{i=1}^{N_D}, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \lambda_1, \lambda_2\}$.



3.2.2 Formulation for the proposed HBM framework

3.2.2.1 Posterior distribution of full set of parameters

The full parameters include the model parameters $\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}$, the prediction error parameters $\{\sigma_i^2\}_{i=1}^{N_D}$ and the hyper-parameters $\{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2\}$. In the case of considering N_D independent data sets, the joint prior distribution of the full parameters is expressed as:

$$\begin{aligned} p(\boldsymbol{\Theta}) &= p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \{\sigma_i^2\}_{i=1}^{N_D}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2) \\ &= p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2) \prod_{i=1}^{N_D} p(\boldsymbol{\theta}_i | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) p(\sigma_i^2 | \lambda_1, \lambda_2) \end{aligned} \quad (3.4)$$

where $p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2)$ denotes the prior distribution of hyper-parameters, and $p(\boldsymbol{\theta}_i | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and $p(\sigma_i^2 | \lambda_1, \lambda_2)$ are introduced in Eqs. (3.3) and (3.2), respectively. In developing Eq. (3.4), the hierarchy structure in Fig. 3-1 is assumed. Specifically, the experiment-specific parameter set $\boldsymbol{\theta}_i$ and σ_i^2 are independent, the distribution of $\boldsymbol{\theta}_i$ is independent of the values of the prediction error hyper parameters λ_1 and λ_2 , while the distribution for σ_i^2 is independent of the structural model hyper parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$. According to Bayes' theorem, the posterior distribution of full parameters is proportional to the prior distribution and the likelihood function:

$$p(\boldsymbol{\Theta} | \mathbf{D}) \propto p(\mathbf{D} | \boldsymbol{\Theta}) p(\boldsymbol{\Theta}) \quad (3.5)$$

Due to the independence of individual data set, and the fact that the i -th data set D_i depends only on the i -th model parameter $\boldsymbol{\theta}_i$ and its prediction error parameter σ_i^2 , the likelihood function can be simplified as:

$$p(\mathbf{D} | \boldsymbol{\Theta}) = p(\mathbf{D} | \{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \{\sigma_i^2\}_{i=1}^{N_D}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2) = \prod_{i=1}^{N_D} p(D_i | \boldsymbol{\theta}_i, \sigma_i^2) \quad (3.6)$$

Herein $p(D_i | \boldsymbol{\theta}_i, \sigma_i^2)$ is the likelihood function for a specific data set i , which can be readily obtained based on the prediction error equation in Eq. (3.1):

$$p(D_i | \boldsymbol{\theta}_i, \sigma_i^2) \propto (\sigma_i^2)^{-\frac{N_0 N_i}{2}} \exp \left\{ -\frac{N_0 N_i}{2 \sigma_i^2} J(\boldsymbol{\theta}_i) \right\} \quad (3.7)$$

where $J(\boldsymbol{\theta}_i)$ is stated as:

$$J(\boldsymbol{\theta}_i) = \frac{1}{N_0 N_i} \sum_{l=1}^{N_0} \frac{1}{a_{i,l}^2} \sum_{j=1}^{N_i} \left(\hat{\mathbf{Y}}_i(j) - \mathbf{g}_i(j; \boldsymbol{\theta}_i) \right)^2 \quad (3.8)$$

Note that $J(\boldsymbol{\theta}_i)$ tends to a finite value representing the average discrepancy between the measurements and the model predictions. Subsequently, using the expressions in Eqs. (3.2) and (3.3), and substituting the simplified likelihood function from Eq. (3.6) into Eq. (3.5) yields:

$$p(\boldsymbol{\Theta} | \mathbf{D}) = p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \{\sigma_i^2\}_{i=1}^{N_D}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2 | \mathbf{D})$$

$$\propto p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2) \prod_{i=1}^{N_D} p(D_i | \boldsymbol{\theta}_i, \sigma_i^2) N(\boldsymbol{\theta}_i | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) IG(\sigma_i^2 | \lambda_1, \lambda_2) \quad (3.9)$$

3.2.2.2 Marginal posterior distribution of hyper-parameters

For obtaining the posterior distribution of hyper parameters, the computational procedure proposed by the previous study [48] will be followed in this section, namely the joint distribution in Eq. (3.9) firstly will be marginalized over σ_i^2 's and then over $\boldsymbol{\theta}_i$'s.

Marginalizing the joint distribution of full parameters over σ_i^2 's, one can obtain:

$$p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2 | \mathbf{D})$$

$$\propto p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2) \prod_{i=1}^{N_D} \left[N(\boldsymbol{\theta}_i | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \int_{\sigma_i^2} p(D_i | \boldsymbol{\theta}_i, \sigma_i^2) IG(\sigma_i^2 | \lambda_1, \lambda_2) d\sigma_i^2 \right] \quad (3.10)$$

The integral in Eq. (3.10) is evaluated analytically in Appendix A, resulting in:

$$\int_{\sigma_i^2} p(D_i | \boldsymbol{\theta}_i, \sigma_i^2) IG(\sigma_i^2 | \lambda_1, \lambda_2) d\sigma_i^2 \propto \frac{(\lambda_2)^{\lambda_1}}{\Gamma(\lambda_1)} \Gamma(f(\lambda_1)) \left[\frac{N_0 N_i}{2} J(\boldsymbol{\theta}_i) + \lambda_2 \right]^{-f(\lambda_1)} \quad (3.11)$$

where $\Gamma(\cdot)$ is the Gamma function and $f(\lambda_1) = (N_0 N_i + 2\lambda_1) / 2$. Hence, Eq. (3.10) can be rewritten as:

$$p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_D}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2 | \mathbf{D}) \propto p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \lambda_1, \lambda_2) \left[\frac{(\lambda_2)^{\lambda_1}}{\Gamma(\lambda_1)} \Gamma(f(\lambda_1)) \right]^{N_D}$$

$$\prod_{i=1}^{N_D} \left[N(\boldsymbol{\theta}_i | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \exp(-L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)) \right] \quad (3.12)$$

where $L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)$ is defined in the form:

$$L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2) = f(\lambda_1) \ln \left[\frac{N_0 N_i}{2} J(\boldsymbol{\theta}_i) + \lambda_2 \right] \quad (3.13)$$

To simplify the analysis and derive analytical expressions for the posterior PDF of the hyper parameters, a key asymptotic approximation, valid for large number of data, is next introduced. The function $\exp(-L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2))$ can be approximated by using Taylor expansion when a large number of data points are available [41]:

$$\exp(-L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)) \approx \exp\left(-L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2) - \frac{1}{2}(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)^T \mathbf{H}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)\right) \quad (3.14)$$

Herein $\hat{\boldsymbol{\theta}}_i$ is the most probable values (MPV) computed by minimizing the objective function $L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)$, while $L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)$ and $\mathbf{H}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)$ are respectively the function value and the hessian matrix of the function $L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)$ with respect to the parameters $\boldsymbol{\theta}_i$ evaluated at the MPV $\hat{\boldsymbol{\theta}}_i$. The calculation of the MPV $\hat{\boldsymbol{\theta}}_i$ and the hessian matrix $\mathbf{H}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)$ are derived in Appendix B and are shown to be given by

$$\hat{\boldsymbol{\theta}}_i = \arg \min_{\boldsymbol{\theta}_i} (J(\boldsymbol{\theta}_i)) \quad (3.15)$$

$$\mathbf{H}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2) = \frac{N_0 N_i f(\lambda_1)}{N_0 N_i J(\hat{\boldsymbol{\theta}}_i) + 2\lambda_2} \mathbf{H}(\hat{\boldsymbol{\theta}}_i) \quad (3.16)$$

where $\mathbf{H}(\hat{\boldsymbol{\theta}}_i) = \nabla \nabla^T J(\boldsymbol{\theta}_i) \big|_{\boldsymbol{\theta}_i = \hat{\boldsymbol{\theta}}_i}$ is the hessian matrix of the measure of fit function $J(\boldsymbol{\theta}_i)$ evaluated at $\hat{\boldsymbol{\theta}}_i$. It is noted that the MPV value can be readily solved by minimizing the function $J(\boldsymbol{\theta}_i)$ which is independent of the unknown hyper parameters λ_1 and λ_2 . Moreover, an explicit expression is found for the hessian matrix $\mathbf{H}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)$ in terms of the hessian of the function $J(\boldsymbol{\theta}_i)$ and the hyper-parameters λ_1 and λ_2 . Based on the calculations of Eq. (3.15) and Eq. (3.16), Eq. (3.14) can be rewritten as:

$$\exp(-L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)) \propto \left[\frac{N_0 N_i}{2} J(\hat{\boldsymbol{\theta}}_i) + \lambda_2 \right]^{-f(\lambda_1)} \sqrt{|\boldsymbol{\Sigma}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)|} N(\boldsymbol{\theta}_i | \hat{\boldsymbol{\theta}}_i, \boldsymbol{\Sigma}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)) \quad (3.17)$$

where the covariance matrix

$$\boldsymbol{\Sigma}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2) = \frac{1}{N_0 N_i} \frac{N_0 N_i J(\hat{\boldsymbol{\theta}}_i) + 2\lambda_2}{f(\lambda_1)} \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}_i) = \frac{2}{N_0 N_i} \frac{J(\hat{\boldsymbol{\theta}}_i) + \frac{2}{N_0 N_i} \lambda_2}{1 + \frac{2}{N_0 N_i} \lambda_1} \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}_i) \quad (3.18)$$

is the inverse of the hessian matrix $\mathbf{H}_L(\hat{\boldsymbol{\theta}}_i, \lambda_1, \lambda_2)$. One can view expressions (3.15) and (3.18) as quantifying the most probable value and the identification uncertainty corresponding to Bayesian parameter inference of $\boldsymbol{\theta}_i$ using the dataset D_i under uniform prior distribution for $\boldsymbol{\theta}_i$.

It is noted that for large number of data points $N_0 N_i$, Eq. (3.18) can be approximately written as:

$$\Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2) = \tilde{\Sigma}_L(\hat{\theta}_i) \approx \frac{2}{N_0 N_i} J(\hat{\theta}_i) \mathbf{H}^{-1}(\hat{\theta}_i) \quad (3.19)$$

demonstrating that the covariance matrix $\Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2) = \tilde{\Sigma}_L(\hat{\theta}_i)$ is independent of the hyper parameters λ_1 and λ_2 . More importantly, Eq. (3.19) shows the uncertainty in the estimates of θ_i is inversely proportional to the square root of the number of data points, reflecting the fact that the identification uncertainty decreases as the number of data points in a dataset increases. Substituting the expression from Eq. (3.17) into Eq. (3.12) yields:

$$p(\{\theta_i\}_{i=1}^{N_D}, \mu_0, \Sigma_0, \lambda_1, \lambda_2 | \mathbf{D}) \propto p(\mu_0, \Sigma_0, \lambda_1, \lambda_2) T(\hat{\theta}, \lambda_1, \lambda_2) \prod_{i=1}^{N_D} \left[N(\theta_i | \mu_0, \Sigma_0) N(\theta_i | \hat{\theta}_i, \Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2)) \right] \quad (3.20)$$

Herein the function $T(\hat{\theta}, \lambda_1, \lambda_2)$ is defined as:

$$T(\hat{\theta}, \lambda_1, \lambda_2) = \left[\frac{(\lambda_2)^{\lambda_1}}{\Gamma(\lambda_1)} \Gamma(f(\lambda_1)) \right]^{N_D} \prod_{i=1}^{N_D} \left[\frac{N_0 N_i}{2} J(\hat{\theta}_i) + \lambda_2 \right]^{-f(\lambda_1)} \sqrt{|\Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2)|} \quad (3.21)$$

Subsequently, by marginalizing the distribution in Eq. (3.20) over θ_i 's space for $i = 1, \dots, N_D$, and noting that

$$\int_{\theta_i} N(\theta_i | \mu_0, \Sigma_0) N(\theta_i | \hat{\theta}_i, \Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2)) d\theta_i = N(\mu_0 | \hat{\theta}_i, \Sigma_0 + \Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2)) \quad (3.22)$$

one readily obtains the posterior distribution of the hyper parameters as follows:

$$p(\mu_0, \Sigma_0, \lambda_1, \lambda_2 | \mathbf{D}) \propto p(\mu_0, \Sigma_0, \lambda_1, \lambda_2) T(\hat{\theta}, \lambda_1, \lambda_2) \prod_{i=1}^{N_D} N(\mu_0 | \hat{\theta}_i, \Sigma_0 + \Sigma_L(\hat{\theta}_i, \lambda_1, \lambda_2)) \quad (3.23)$$

As seen from Eq. (3.23), the hyper parameters corresponding to both model parameters and prediction error hyper parameters can be computed together through the proposed framework. Any Markov Chain Monte Carlo (MCMC) algorithm such as the transitional MCMC (TMCMC) [52,53] or the nested sampling algorithm [54] can be used to draw the samples from the posterior distribution of the hyper parameters. Note that sampling from the posterior distribution $p(\mu_0, \Sigma_0, \lambda_1, \lambda_2 | \mathbf{D})$ is not a time consuming operation since it does no longer require computationally expensive model runs. The model runs are required only to estimate and store the values of $\hat{\theta}_i$ and $\mathbf{H}(\hat{\theta}_i)$ before the sampling approach is initiated.

3.2.2.3 Posterior distribution of hyper parameters for large number of data

For large number of data ($N_0 N_i$ is large) within each data set, the values of $f(\lambda_1) = N_0 N_i / 2$ and $\Gamma(f(\lambda_1)) = \Gamma(N_0 N_i / 2)$ are independent of the parameter λ_1 . Using Eq. (3.19) and assuming that prior to data the hyper parameter sets $\{\mu_0, \Sigma_0\}$ and $\{\lambda_1, \lambda_2\}$ are independent, i.e. $p(\mu_0, \Sigma_0, \lambda_1, \lambda_2) = p(\lambda_1, \lambda_2) p(\mu_0, \Sigma_0)$, the posterior distribution in Eq. (3.23) takes the simplified form

$$p(\mu_0, \Sigma_0, \lambda_1, \lambda_2 | \mathbf{D}) \approx p(\lambda_1, \lambda_2 | \mathbf{D}) p(\mu_0, \Sigma_0 | \mathbf{D}) \quad (3.24)$$

where

$$p(\mu_0, \Sigma_0 | \mathbf{D}) \propto \prod_{i=1}^{N_D} N(\mu_0 | \hat{\theta}_i, \Sigma_0 + \tilde{\Sigma}_L(\hat{\theta}_i)) p(\mu_0, \Sigma_0) \quad (3.25)$$

and

$$p(\lambda_1, \lambda_2 | \mathbf{D}) \propto T(\hat{\theta}, \lambda_1, \lambda_2) \propto \left[\frac{(\lambda_2)^{\lambda_1}}{\Gamma(\lambda_1)} \right]^{N_D} \prod_{i=1}^{N_D} \left[\frac{N_0 N_i}{2} J(\hat{\theta}_i) + \lambda_2 \right]^{-N_0 N_i / 2} p(\lambda_1, \lambda_2) \quad (3.26)$$

According to Eq. (3.24) the hyper parameter sets $\{\mu_0, \Sigma_0\}$ and $\{\lambda_1, \lambda_2\}$ remain independent given the data sets. This independence is useful to analytically derive the most probable values and the uncertainties of the hyper parameters $\{\mu_0, \Sigma_0\}$ and $\{\lambda_1, \lambda_2\}$ by separately considering the distribution $p(\mu_0, \Sigma_0 | \mathbf{D})$ and $p(\lambda_1, \lambda_2 | \mathbf{D})$.

It has been shown in [55] that the form in Eq. (3.25) of the posterior distribution of the hyper parameters $\{\mu_0, \Sigma_0\}$ yields the most probable values of the hyper parameters to be

$$\hat{\mu}_0 = \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\theta}_i \quad (3.27)$$

and using the assumption that all $\tilde{\Sigma}_L(\hat{\theta}_i)$ are approximately equal, the hyper-parameter covariance matrix to be

$$\hat{\Sigma}_0 = \frac{1}{N_D} \sum_{i=1}^{N_D} (\hat{\mu}_0 - \hat{\theta}_i)(\hat{\mu}_0 - \hat{\theta}_i)^T - \Sigma_0 \quad (3.28)$$

where $\Sigma_0 = \frac{1}{N_D} \sum_{i=1}^{N_D} \tilde{\Sigma}_L(\hat{\theta}_i)$ is taken as the average of the identification uncertainty of each

dataset. The uncertainty in the hyper parameters, approximated by the inverse of the hessian of $-\ln p(\mu_0, \Sigma_0 | \mathbf{D})$ evaluated at the most probable values $\{\hat{\mu}_0, \hat{\Sigma}_0\}$, can be derived to be

$$\mathbf{\Sigma}(\hat{\boldsymbol{\mu}}_{\theta}, \hat{\boldsymbol{\Sigma}}_{\theta}) = \frac{1}{N_D} \mathbf{H}_h^{-1}(\hat{\boldsymbol{\mu}}_{\theta}, \hat{\boldsymbol{\Sigma}}_{\theta}) \quad (3.29)$$

where $\mathbf{H}_h(\boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta})$ is a 2x2 block matrix with the individual blocks given by [55]

$$\begin{aligned} \mathbf{H}_{\mu\mu}(\boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta}) &= \frac{1}{N_D} \sum_{i=1}^{N_D} \left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} \\ \mathbf{H}_{\mu\Sigma}(\boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta}) &= -\frac{1}{N_D} \sum_{i=1}^{N_D} \left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} \otimes \left[\left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} (\boldsymbol{\mu}_{\theta} - \hat{\boldsymbol{\theta}}_i) \right] \\ \mathbf{H}_{\Sigma\Sigma}(\boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta}) &= -\frac{1}{2N_D} \sum_{i=1}^{N_D} \left[\left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} \otimes \left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} \right] \\ &\quad + \frac{1}{N_D} \sum_{i=1}^{N_D} \left[\left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} \otimes \left[\left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} (\boldsymbol{\mu}_{\theta} - \hat{\boldsymbol{\theta}}_i)(\boldsymbol{\mu}_{\theta} - \hat{\boldsymbol{\theta}}_i)^T \left[\boldsymbol{\Sigma}_{\theta} + \tilde{\boldsymbol{\Sigma}}_L(\hat{\boldsymbol{\theta}}_i) \right]^{-1} \right] \right] \end{aligned} \quad (3.30)$$

where \otimes denotes the Kronecker tensor product of two matrices. Note that the matrices in the left-hand-side of Eq. (3.30) tend to finite values as the number of datasets N_D increases.

Using Eq. (3.26), the most probable values $\hat{\lambda}_1$ and $\hat{\lambda}_2$ can be obtained by minimizing the $-\ln p(\lambda_1, \lambda_2 | \mathbf{D})$ with respect to the parameters λ_1 and λ_2 , while an asymptotic estimate of the uncertainty in the estimates of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ can readily be derived by calculating the inverse of the hessian of $-\ln p(\lambda_1, \lambda_2 | \mathbf{D})$, evaluated at the most probable values $\hat{\lambda}_1$ and $\hat{\lambda}_2$, in the form:

$$\mathbf{\Sigma}_{\lambda}(\hat{\lambda}_1, \hat{\lambda}_2) = \frac{1}{N_D} \mathbf{H}_{\lambda}^{-1}(\hat{\lambda}_1, \hat{\lambda}_2) \quad (3.31)$$

where the elements of the 2x2 hessian matrix are

$$\begin{aligned} \mathbf{H}_{\lambda,11}(\lambda_1, \lambda_2) &= \frac{d^2 \ln \Gamma(\lambda_1)}{d\lambda_1^2} = \psi'(\lambda_1) = \psi(1, \lambda_1) \\ \mathbf{H}_{\lambda,12}(\lambda_1, \lambda_2) &= -\frac{1}{\lambda_2} \\ \mathbf{H}_{\lambda,22}(\lambda_1, \lambda_2) &= \frac{\lambda_1}{\lambda_2^2} - \frac{2}{N_0 N_i} \frac{1}{N_D} \sum_{i=1}^{N_D} \frac{1}{\left[J_i(\hat{\boldsymbol{\theta}}_i) + \frac{2}{N_0 N_i} \lambda_2 \right]^2} \approx \frac{\lambda_1}{\lambda_2^2} \end{aligned} \quad (3.32)$$

where $\psi(\cdot)$ is the Digamma function and $\psi(k, \lambda_1)$ is the k -th derivative of the digamma function at λ_1 , which is readily evaluated by a Matlab-based function ‘psi’. It is noted that Eq. (3.32) does not explicitly depend on the estimates $\hat{\boldsymbol{\theta}}_i$ for large number of data per dataset (large $N_0 N_i$).

The aforementioned formulations (3.29) and (3.31) for the covariance matrices of the estimates of all hyper parameters to be inversely proportional to the number N_D of the datasets suggest that the uncertainties in the hyper parameter estimates reduces as the number of datasets increases. Also, the forms (3.29) and (3.31) for the covariance matrices suggest that the posterior uncertainty in the hyper parameters tends to be approximated by a Gaussian distribution as the number N_D increases.

3.2.2.4 Probability distribution function (PDF) of model and prediction error parameters

The PDF of model parameters $\boldsymbol{\theta}$ given the data \mathbf{D} can be simplified to

$$p(\boldsymbol{\theta} | \mathbf{D}) = \int \int \int \int p(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) p(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \lambda_1, \lambda_2 | \mathbf{D}) d\boldsymbol{\mu}_\theta d\boldsymbol{\Sigma}_\theta d\lambda_1 d\lambda_2 \quad (3.33)$$

where we used the fact that $p(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \lambda_1, \lambda_2) = p(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$, i.e. the conditional distribution of the model parameter $\boldsymbol{\theta}$ given the values of the hyper parameters $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\Sigma}_\theta$ is independent of hyper-parameters λ_1, λ_2 . Sampling estimates can be employed to compute the integral in the form:

$$p(\boldsymbol{\theta} | \mathbf{D}) \approx \frac{1}{M} \sum_{k=1}^M N(\boldsymbol{\theta} | \boldsymbol{\mu}_\theta^{(k)}, \boldsymbol{\Sigma}_\theta^{(k)}) \quad (3.34)$$

by drawing the samples $\boldsymbol{\mu}_\theta^{(k)}, \boldsymbol{\Sigma}_\theta^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)}$ from the distribution $p(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \lambda_1, \lambda_2 | \mathbf{D})$, where M is the number of samples.

Similarly, the PDF of prediction error parameter σ^2 can be computed by following the same procedure:

$$\begin{aligned} p(\sigma^2 | \mathbf{D}) &= \int \int \int \int p(\sigma^2 | \lambda_1, \lambda_2) p(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \lambda_1, \lambda_2 | \mathbf{D}) d\boldsymbol{\mu}_\theta d\boldsymbol{\Sigma}_\theta d\lambda_1 d\lambda_2 \\ &\approx \frac{1}{M} \sum_{k=1}^M IG(\sigma^2 | \lambda_1^{(k)}, \lambda_2^{(k)}) \end{aligned} \quad (3.35)$$

Thus, the samples of structural parameters and prediction error parameter can be obtained according to Eqs. (3.34) and (3.35), respectively using any MCMC algorithm.

3.2.2.5 Predictions of output QoI

After the structural model is calibrated, the uncertainty of the structural parameters and the uncertainty from the prediction error parameter can be used to predict the uncertainty in output QoI using Monte Carlo simulations. Specifically, let $\boldsymbol{\theta}^{(q)}$ denote the q -th sample of

model parameter $\boldsymbol{\theta}$ generated from its distribution $p(\boldsymbol{\theta} | D)$ in Eq. (3.34), $\sigma^{2(q)}$ is the q -th sample of prediction error variance parameter which can be generated from its distribution $p(\sigma^2 | D)$ in Eq. (3.35), and $\boldsymbol{\varepsilon}_l^{(q)}$ is the prediction error term sampled from a Gaussian distribution with zero mean and covariance matrix $\boldsymbol{\Sigma} = (\sigma^{2(q)} a_{pre,l}^2) \mathbf{I}$, where $a_{pre,l}$ is the intensity of the predicted response time histories in the l -th DOF. The uncertainty in a response time history $\mathbf{Y}_{pre}(j)$ can be obtained by analyzing the samples generated according to the following expression:

$$\mathbf{Y}_{pre,l}^{(q)}(j) = \mathbf{g}_l(j; \boldsymbol{\theta}^{(q)}) + \boldsymbol{\varepsilon}_l^{(q)} \quad (3.36)$$

where l defines the Bayesian Hierarchical Learning Framework for Multi-Level Modeling using Multi-level Data in Structural Dynamics-th DOF of the system and j is the data point. Eq. (3.35) is applied to both the observed and unobserved QoI. Observations of a specific type of response time histories (e.g. accelerations) at a limited number of DOF provides estimates of the prediction error parameter σ^2 which can then be used in (3.36) to make predictions of response time histories (e.g. accelerations) at unobserved DOF of the model. For the unobserved QoI of different type (e.g. velocities, strain, displacements) an estimate of the variance σ^2 involved in the definition of the zero mean Gaussian error term $\boldsymbol{\varepsilon}_l$ is not available from the observations and has to be subjectively postulated. Here, the responses to unobserved QoI of different type than that used for parameter estimation are computed based on only the structural parameters uncertainty. This is reasonable as the estimates of the prediction error are not available for such unobserved QoI.

3.2.3 Computational procedure of the proposed algorithm

The procedure for the parameter identification along with the predictions of QoI is summarized in Algorithm 1. The computational process of the proposed HBM framework includes two steps for the identification of the hyper parameters. In the first step the MPV and the hessian matrix are obtained according to each data set. In this step it requires the model runs and thus it is the most computationally expensive step. While in the second step, it does not need the model runs, and only the MPV and the Hessian matrix evaluated from the step are used for sampling the posterior distribution of the hyper parameters.

It is reminded that an alternative full sampling (FS)-based HBM approach [43,51] can be also utilized to obtain the samples of the hyper parameters. This is achieved by drawing the

samples of the model and prediction error parameters from the likelihood function for each data set in the first step, and then use all the available samples to compute the posterior distribution of the hyper parameters in the second step. Therefore, the computational cost of the FS approach is more expensive than the proposed asymptotic approximation approach. The comparison of the computational effort between the FS and the proposed methods is conducted in the second application of Section 3.4, where both methods are used for estimating the posterior distributions of the hyper parameters.

Algorithm 1:

Proposed HBM framework for parameter identification and response prediction

1. Identify hyper parameters

First Step: Find MPV and hessian of function $L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)$ for each data set

1.1) Minimize $J(\boldsymbol{\theta}_i)$ to compute MPV $\hat{\boldsymbol{\theta}}_i$ using Eq. (3.15)

1.2) Evaluate Hessian matrix $\mathbf{H}(\hat{\boldsymbol{\theta}}_i)$ at MPV $\hat{\boldsymbol{\theta}}_i$

Second Step: Compute posterior distribution $p(\boldsymbol{\mu}_\theta, \Sigma_\theta, \lambda_1, \lambda_2 | \mathbf{D})$ using Eq. (3.23)

2. Identify model parameters and prediction error parameter

2.1) Compute PDF $p(\boldsymbol{\theta} | \mathbf{D})$ of model parameters using Eq. (3.34)

2.2) Compute PDF $p(\sigma^2 | \mathbf{D})$ of prediction error parameter using Eq. (3.35)

3. Predictions of QoI

3.1) Draw samples $\boldsymbol{\theta}^{(q)} (q = 1, \dots, N_q)$ from $p(\boldsymbol{\theta} | \mathbf{D})$, and $\sigma^{2(q)}$ from $p(\sigma^2 | \mathbf{D})$

3.2) Draw samples from $N(\mathbf{0}, \sigma^{2(q)} a_{pre,i}^2 \mathbf{I})$ for $\boldsymbol{\varepsilon}_i$

3.3) Calculate the predictions using Eq. (3.36)

3.4) Estimate the statistical properties using Monte Carlo simulations

3.3 Application to Nonlinear Systems using Bouc-Wen hysteresis

The Bouc-Wen (BW) model is widely used in dynamical structures to represent the hysteretic behavior of nonlinear systems [22,31]. It was initially proposed by Bouc [56], subsequently modified by Wen [57] and thereafter extended by other researchers in the literature [58–60]. Details of the formulation in civil infrastructure can be found in references [59,61]. Herein,

for demonstration purposes, it will be applied to a multistory building represented by a shear model with BW hysteretic-type inter-story nonlinearities.

Specifically, the differential equation of motion is written in the form:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{F}_R(\mathbf{u}(t), \mathbf{z}(t)) = \mathbf{P}(t) \quad (3.37)$$

where \mathbf{M} and \mathbf{C} are the mass and the viscous damping matrices, $\mathbf{u}(t)$ is the relative displacement response, $\mathbf{F}_R(\mathbf{u}(t), \mathbf{z}(t))$ is the nonlinear restoring force vector at time t , and $\mathbf{P}(t)$ is the input force vector. According to the shear building model, the i -th component of the vector $\mathbf{F}_R(\mathbf{u}(t), \mathbf{z}(t))$ is given by [28,59]:

$$F_R^i = F_r^i - F_r^{i+1} \quad (3.38)$$

$i = 1, \dots, N_0$, with $F_r^{i+1} = 0$ for $i = N_0$. According to the BW model the quantity F_r^i in Eq. (3.38) is the nonlinear inter-story restoring force given as [59]:

$$F_r^i = \alpha_i k_i u_i(t) + (1 - \alpha_i) k_i z_i(t) \quad (3.39)$$

where $\alpha_i k_i u_i(t)$ corresponds to the elastic component whereas $(1 - \alpha_i) k_i z_i(t)$ represents to the hysteretic component, k_i denotes the stiffness, α_i defines the share of linear part while $1 - \alpha_i$ define the share of nonlinear hysteretic part, and $z_i(t)$ is the virtual hysteretic displacement which comprises the hysteretic component of the system. Without including the pinching effect and degradation functions (stiffness or strength degradation), the formulation of the implemented hysteretic displacement $z_i(t)$ can be simplified as:

$$\dot{z}_i(t) = A_i \dot{u}_i(t) - \beta_i |\dot{u}_i(t)| |z_i(t)|^{n-1} - \gamma_i \dot{u}_i(t) |z_i(t)|^{n_i} \quad (3.40)$$

where the parameter A_i determines the tangent stiffness, and the parameters β_i , γ_i and n_i affect the shape and smoothness of the hysteretic model, respectively.

For solving the differential equation of the BW model, the state vector in state space form is given as:

$$\mathbf{y}(t) = \begin{Bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \\ \mathbf{y}_3(t) \end{Bmatrix} = \begin{Bmatrix} \mathbf{u}(t) \\ \dot{\mathbf{u}}(t) \\ \mathbf{z}(t) \end{Bmatrix} \quad (3.41)$$

The derivative of the state vector $\mathbf{y}(t)$ is readily obtained as:

$$\dot{\mathbf{y}}(t) = \begin{Bmatrix} \dot{\mathbf{y}}_1(t) \\ \dot{\mathbf{y}}_2(t) \\ \dot{\mathbf{y}}_3(t) \end{Bmatrix} = \begin{Bmatrix} \mathbf{y}_2(t) \\ \mathbf{M}^{-1}\mathbf{P}(t) - \mathbf{M}^{-1}(\mathbf{C}\mathbf{y}_2(t) + \mathbf{F}_R(\mathbf{y}_1(t), \mathbf{y}_3(t))) \\ \mathbf{A}\mathbf{y}_2(t) - \boldsymbol{\beta}|\mathbf{y}_2(t)| \otimes \mathbf{y}_3(t) \otimes |\mathbf{y}_3(t)|^{n-1} - \boldsymbol{\gamma}\mathbf{y}_2(t) \otimes |\mathbf{y}_3(t)|^n \end{Bmatrix} \quad (3.42)$$

The ordinary differential equations in (3.42) can then be solved numerically using the fourth-fifth order Runge-Kutta method [62]. The implemented BW model at each structural element can be fully parameterized by 6 parameters, namely the linear stiffness parameter $\mathbf{k} = \{k_i\}_{i=1}^n$ and the nonlinear parameters $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^n$, $\mathbf{A} = \{A_i\}_{i=1}^n$, $\boldsymbol{\beta} = \{\beta_i\}_{i=1}^n$, $\boldsymbol{\gamma} = \{\gamma_i\}_{i=1}^n$, $\mathbf{n} = \{n_i\}_{i=1}^n$. Selected parameters can be incorporated into the model parameters set $\boldsymbol{\theta}$ and subsequently identified based on the proposed HBM framework.

In order to find the most probable values $\hat{\boldsymbol{\theta}}_i$ needed in the HBM formulation in Eq. (3.23), one needs to solve an optimization problem of minimizing $J(\boldsymbol{\theta}_i)$ in Eq. (3.8) with respect to the model parameters $\boldsymbol{\theta}_i$. Gradient based optimization techniques are used which require the gradient of $J(\boldsymbol{\theta}_i)$. These gradients require the knowledge of the derivatives $\mathbf{y}_\theta(t)$ of the response vector $\mathbf{y}(t)$ in (3.42) with respect to a model parameter θ in the set $\boldsymbol{\theta}_i$. Using the form in Eq.(3.42), analytical expressions can readily be developed for these derivatives. Specifically, the expressions for the derivatives $\mathbf{y}_\theta(t)$ constitute a set of differential equations that is also solved using the fourth-fifth order Runge-Kutta method. The formulation for the analytical derivatives is not included in this work, however, similar study can be found in [63].

3.4 Illustrative Examples

3.4.1 Case study 1: calibration of a SDOF nonlinear system using a linear model

This case study aims to demonstrate that uncertainties from processing multiple data sets arise due to model error and to investigate the effectiveness of the HBM framework for quantifying uncertainties in the presence of model error. To introduce model error the physical system is assumed to be nonlinear and the model of the system to be linear. Specifically, a single degree of freedom (SDOF) nonlinear structure modeled by Bouc-Wen hysteretic type nonlinearity is employed as the physical system, as shown in Fig. 3-2(a). The mass m and the viscous damping ratio ζ are assigned the nominal values of 1 kg and 2%, respectively. The initial stiffness k which corresponds to the linear component is set to the

nominal value of 40 N/m. The values of the parameters associated with the nonlinear component are set to $\alpha=0.5, \beta=2, \gamma=2, n=1$ and $A=1$. The excitations are Gaussian sequences with mean zero and standard deviation of 5. Recorded measured data consisting of acceleration time histories that are simulated from the nonlinear model. Specifically, 100 data sets consisting of acceleration time histories are simulated from the nominal nonlinear model using different realizations of the white noise excitation. For each data set, the sampling rate is taken as 0.01s corresponding to a sampling frequency 100Hz for a total of 10 seconds. The linear model, shown in Fig. 3-2(b), used to represent the nonlinear system is parameterized with 2 model parameters $\theta=(\theta_1, \theta_2)^T$ associated with the stiffness and damping ratio. The parameters θ_1 and θ_2 multiply the nominal stiffness and damping ratio, respectively.

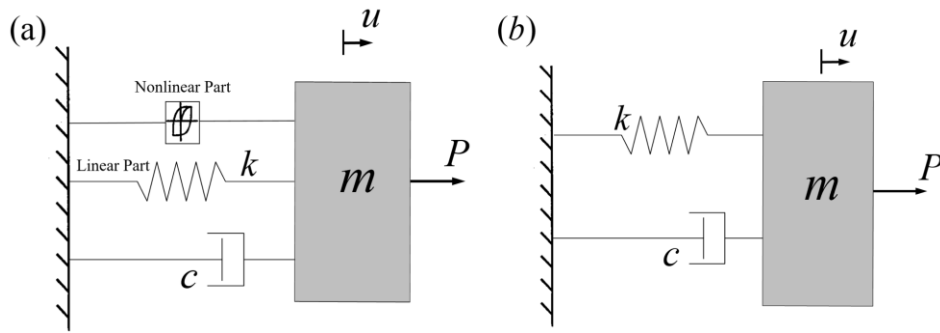


Fig. 3-2 (a) SDOF system with BW hysteresis (b) SDOF linear model of the system in (a)

The proposed HBM framework is next applied to identify the hyper parameters of the model and to predict the responses of the linear model. The MPVs of model parameters θ_i for each data set is computed according to Eq. (3.15), and the hessian matrix evaluated at θ_i is then calculated. The posterior distribution of the hyper-parameters $\{\mu_\theta, \Sigma_\theta, \lambda_1, \lambda_2\}$ can then be estimated using Eq. (3.23). Nested sampling algorithm [54] is employed here to draw the samples from the posterior distribution. The number of the initial samples and the tolerance value in the sampler are set to 500 and 0.01, respectively. The results of the posterior distribution of the hyper-parameters corresponding to $N_D=100$ is shown in Fig. 3-3, where the diagonal figures show the marginal PDF of each hyper parameter, and the lower diagonal sub-figures show the contour plots for each pair of the hyper parameters. The mean of the hyper parameters are also reported in Table 3-1. It can be seen from Fig. 3-3 that a clear peak appears for all the hyper parameters. The first two hyper parameters (hyper mean) have mean values of 0.9455 and 2.2607, respectively. These values deviate from their nominal values of

$\mu_0 = (1,1)^T$ especially for the second hyper parameter due to the existence of substantial model error. Regarding the uncertainty of the model parameters, the variability of damping ratio with a coefficient of variation (cov) of 13.7% is much larger than the variability of the stiffness parameter with a cov of 1.47%. This is reasonable since the model response is more sensitive to the stiffness parameter than the damping ratio parameter. Furthermore, the marginal distribution of the hyper parameters in Fig. 3-3 provide the uncertainty in the estimates of the hyper parameters $\{\mu_0, \Sigma_0\}$ with cov, computed from the samples, equal to (0.0015,0.0138) for μ_0 and (0.0722,0.0708) for the diagonal elements of Σ_0 . Such uncertainties are large for the elements of the hyper covariance Σ_0 and are expected to affect propagation of uncertainties into output QoI. As mentioned in the theoretical formulation, the uncertainty in such estimates is inversely proportional to the square root of the number of datasets and thus is expected to be higher for smaller number of datasets used. The last two parameters λ_1 and λ_2 aim to capture the uncertainty of the prediction error parameter. It is obvious that the two parameters exhibit a strong correlation as shown in Fig. 3-3 by the subplot corresponding to the projection of samples and contour to the two parameter space (λ_1, λ_2) . With those parameters one can calculate the probability of the prediction error based on Eq. (3.35) and the mean of the prediction error is also reported in Table 3-1.

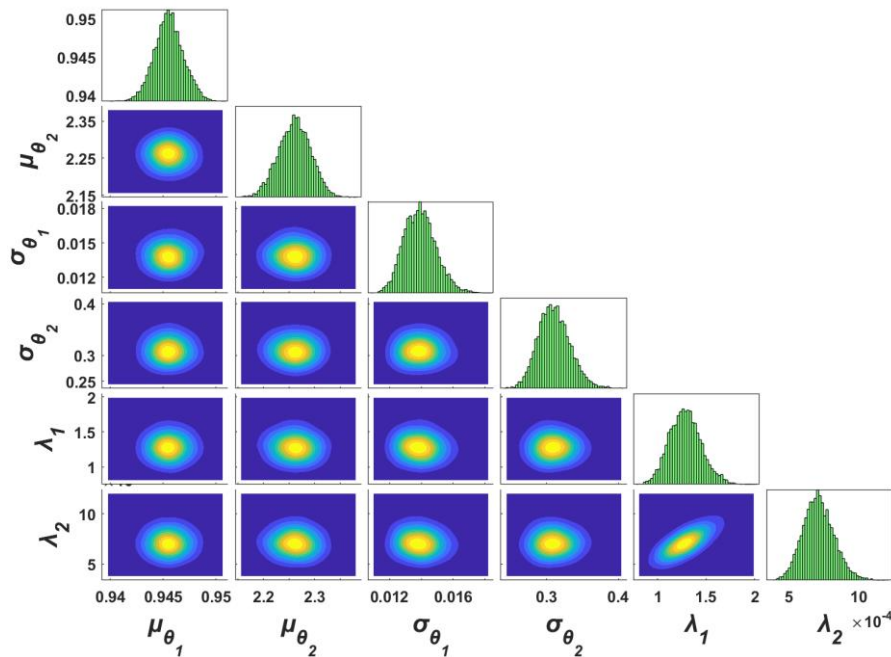


Fig. 3-3 Posterior distribution of the hyper-parameters

Table 3-1 Estimates of mean and standard deviation of the model parameters and prediction error parameters

	Model parameters						Prediction error parameters		
	$\hat{\mu}_{\theta_1}$	$\hat{\sigma}_{\theta_1}$	$\frac{\hat{\sigma}_{\theta_1}}{\hat{\mu}_{\theta_1}}$	$\hat{\mu}_{\theta_2}$	$\hat{\sigma}_{\theta_2}$	$\frac{\hat{\sigma}_{\theta_2}}{\hat{\mu}_{\theta_2}}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\sigma}$
HBM	0.9455	0.0139	1.47%	2.2607	0.3105	13.7%	1.2863	0.0007	0.0260
CBM	0.9390	0.0001	0.01%	2.4410	0.0019	0.078%	-	-	0.0690

For the purpose of the comparisons between the proposed HBM approach and the conventional Bayesian method (CBM), 100 data sets used in HBM are incorporated in a single data set for investigating the performance of the CBM. Table 3-1 includes results of the mean estimates and the identification uncertainty of the model parameters obtained from the CBM. As seen, the CBM provides estimates of the mean of the model parameters that are quite close to the proposed HBM approach. However, a substantial difference is revealed regarding the extent of variability of the model parameters as quantified by the standard deviations in the two approaches. Specifically, the proposed HBM framework offers larger uncertainty estimates as opposed to the CBM which yields extremely small evaluations of the model parameter standard deviations. This feature of the proposed HBM approach is beneficial for predictions of realistic uncertainty bounds of the unobserved quantities of interest. Moreover, the proposed approach can predict the hyper parameters of the prediction error, elucidating explicitly a probability distribution of the prediction error parameter. It should be noted that the mean of the prediction error computed by CBM is larger than that by HBM. This is because the prediction error in CBM include both the expected variability to the model parameters due to multiple datasets and uncertainty for each data set due to the model error. In contrast, the proposed HBM split the two uncertainties, embedding part of the uncertainty in the model parameters, and therefore present a more reasonable framework for model predictions and structure assessments.

The identification uncertainty of model parameter in each data set can be estimated according to Eq.(3.18), where the parameters λ_1 and λ_2 are assumed to be their mean evaluations from Table 1. Fig. 3-4 shows the plots of the identification uncertainty obtained based on individual data set and the ensemble uncertainty observed over multiple data sets for the model and prediction error parameters as a function of the number of data sets N_D

ranging from 1 to 100. Apparently, the identification uncertainty for each dataset is fairly small which is attributed to the fact that a large amount of data points is used here in measurements. Such uncertainties can be negligible for further predictions. However, the ensemble uncertainty arising from the variability due to model error, is large and irreducible, tending to a constant value after approximately 10 datasets. More importantly, this uncertainty covers almost all the mean values of the individual model parameters, representing that the model parameter values have significant variabilities due to the presence of model error.

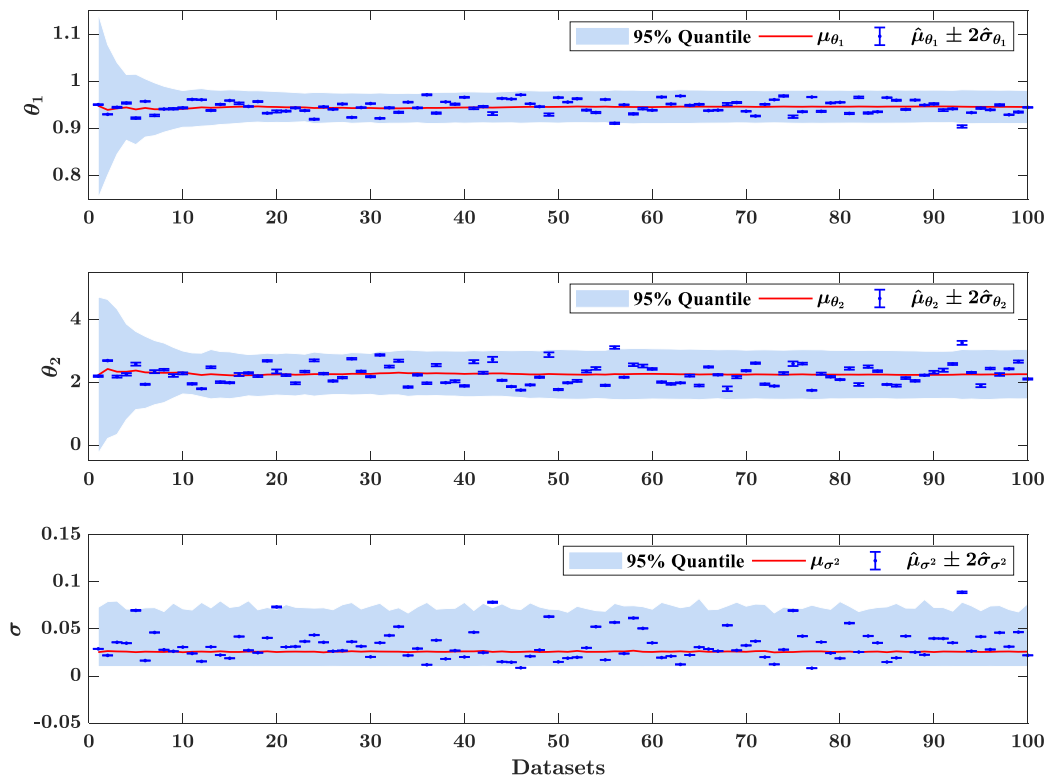


Fig. 3-4 Estimates of identification uncertainty and ensemble uncertainty

The above mentioned uncertainties are propagated to predict output quantities of interest. The predictions are performed for observed and unobserved QoI in order to highlight the importance of the different uncertainties. For comparison purpose, the CBM is also employed to predict the QoI. Fig. 3-5 shows the predictions of acceleration (observed QoI) where the uncertainties from both the model parameters and the prediction error parameter are considered. Despite the conceptual differences in HBM and CBM in quantifying and handling uncertainties, it is seen that by accounting for the overall uncertainties both methods can provide reliable uncertainty bounds (UB) and the actual response (as shown in the red

line of Fig. 3-5(a)) which generates from a new input excitation falls inside of the UB. However, a significant difference between the HBM and CBM is shown in Fig. 3-6 for the predictions of displacement (unobserved QoI). Due to the fact that the prediction error is available only for the observed QoI, only the uncertainty of the model parameters is propagated to the predictions of the unobserved QoI. It is evident that owing to an underestimation of the identification uncertainty of model parameters, the CBM provides a thin UB if only the uncertainties of model parameters are considered, and the actual response falls outside of the thin UB. In contrast, the proposed HBM approach provides reasonable UB by accounting for only the uncertainties of the model parameters as most parts of the actual displacement falls within the produced UB, as depicted in Fig. 3-6.

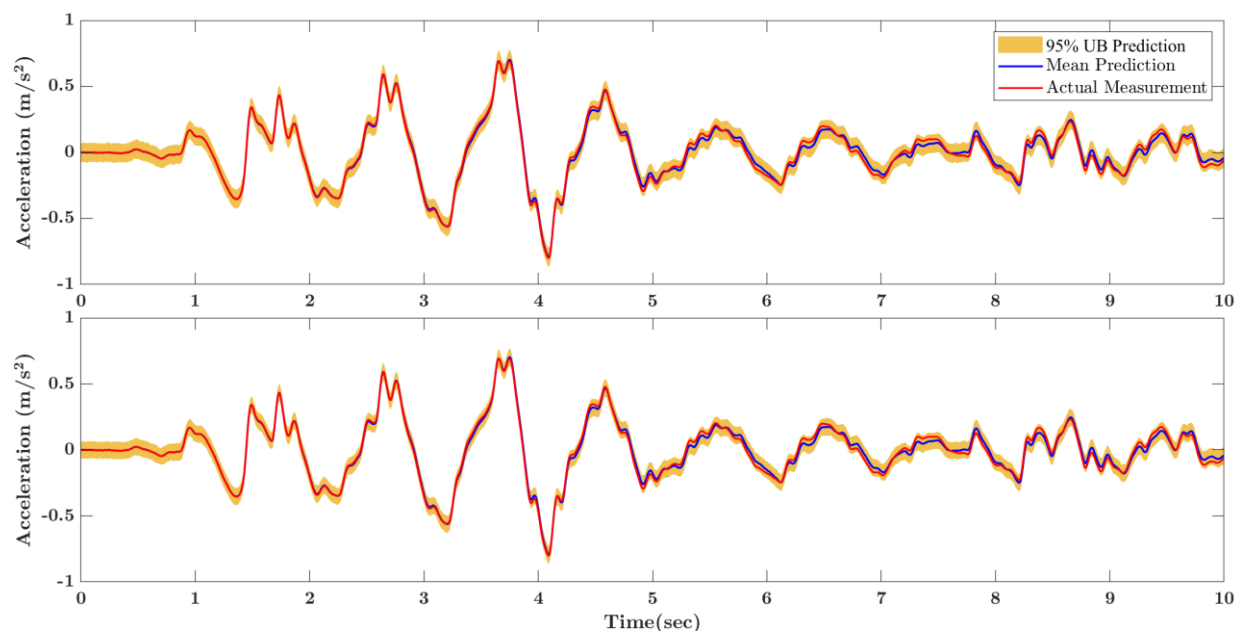


Fig. 3-5 Predictions of acceleration using CBM (top) and HBM (bottom)

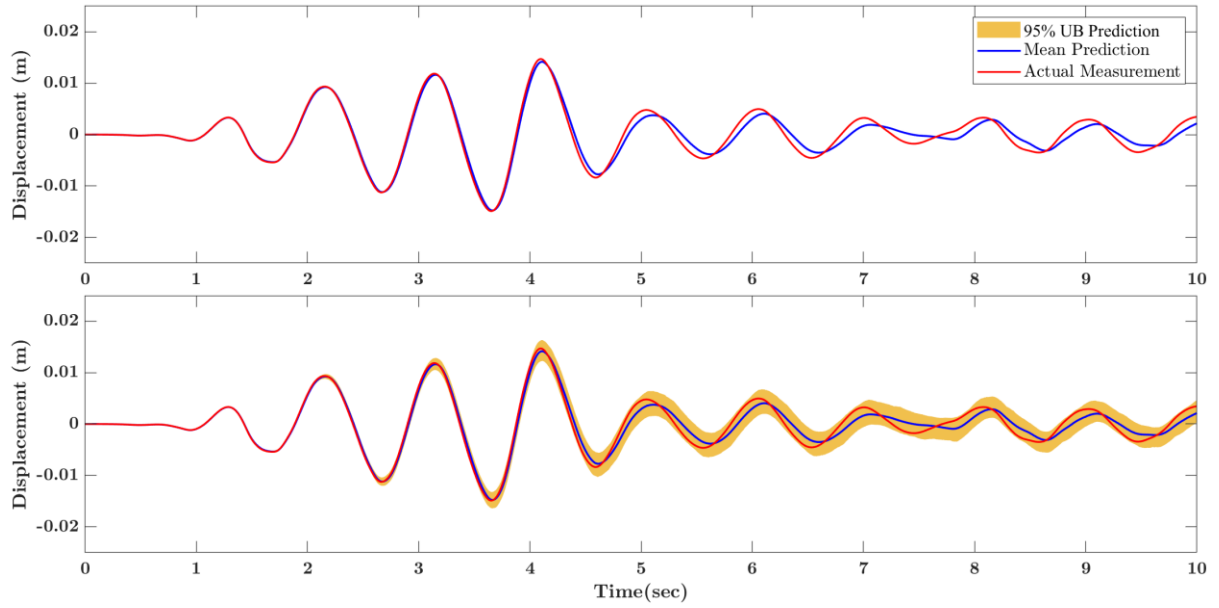


Fig. 3-6 Predictions of displacement using CBM (top) and HBM (bottom)

3.4.2 Case study 2: calibration of a 5 DOF nonlinear system with BW hysteretic model

Fig. 3-7(a) shows a 5-DOF shear building model of a structure with mass $m=1$ kg, stiffness $k_0=500$ N/m, and hysteretic nonlinearities for the inter-story stiffness of each floor. The nominal values of the parameters of the nonlinear BW hysteretic model for each story are assumed to be $\alpha_0=0.1$, $\beta_0=1$, $\gamma_0=1$, $A=2$ and $n=1$, respectively. The physical system is assumed to be the model shown in Fig. 3-7(a) with the values of the model parameters corresponding to mass m , stiffness k and the nonlinear parameters α , β and γ in all five floors of the structure independently perturbed by 5% from their nominal values. The perturbed value of each parameter is different for each floor of the building. The other two parameters A and n are set to be their nominal values. The system is excited at the base with $N_D=25$ different earthquake excitations taken from the engineering strong motion (ESM) database [64]. One of the earthquake excitations is shown in Fig. 3-2(b). 25 data sets of acceleration time histories are generated and used as measurements for identifying the model of the system. Most of the analysis and results that follow are based on time history measurement obtained from a single acceleration sensor located at the first story.

The model of the system is parameterized using the four unknown parameters $\theta=(\theta_1, \theta_2, \theta_3, \theta_4)$ representing normalized k, α, β and γ , denoted with red in Fig. 3-7(a). The parameters $\theta=(\theta_1, \theta_2, \theta_3, \theta_4)$ are assumed to be the same for all stories and are normalized by

the nominal values of the corresponding model properties such as $k = \theta_1 k_0$, $\alpha = \theta_2 \alpha_0$, $\beta = \theta_3 \beta_0$ and $\gamma = \theta_4 \gamma_0$. Therefore, the identification will involve estimating the stiffness parameter θ_1 and evaluating the material nonlinear parameters $(\theta_2, \theta_3, \theta_4)$. The other quantities are assumed to be deterministic with their values set to their nominal values.

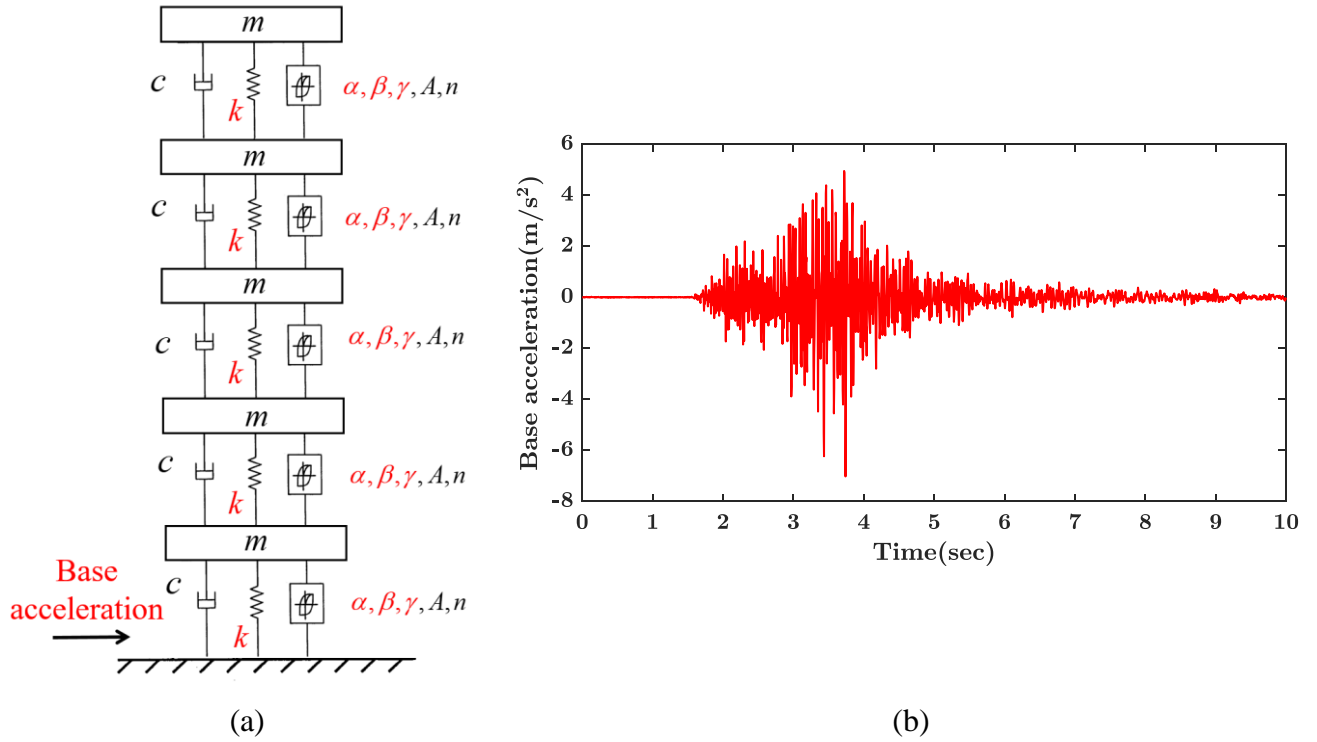


Fig. 3-7 (a) 5-DOF shear model of a building system (b) Base earthquake excitation

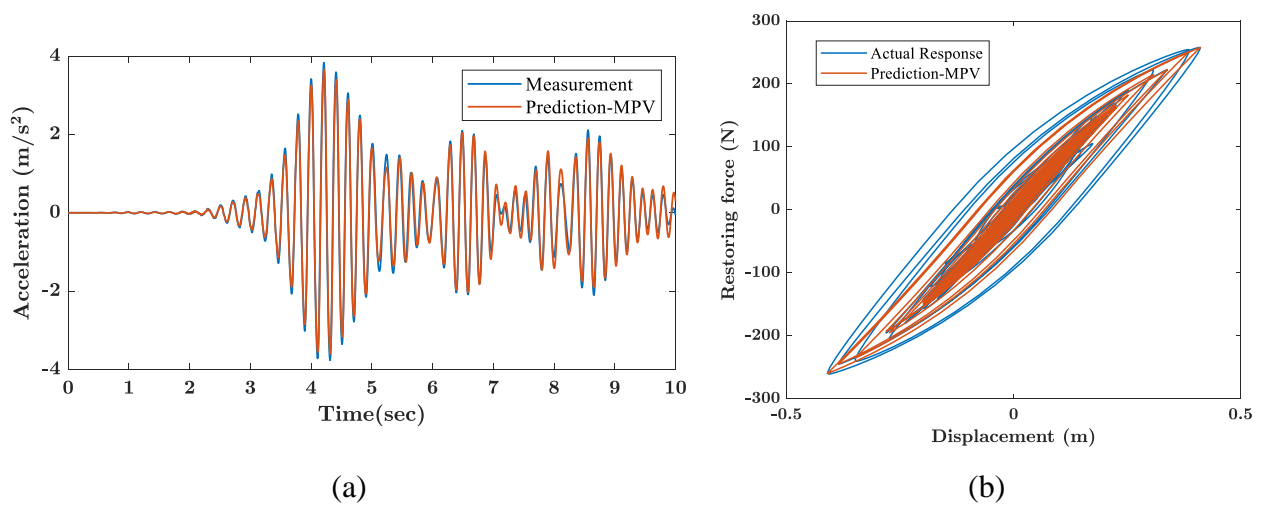


Fig. 3-8 (a) Comparisons between the measured acceleration response and prediction and (b) the actual hysteresis loop and prediction from optima model of the first floor.

The proposed hierarchical Bayesian nonlinear model updating approach is applied to identify the stiffness and the parameters of the hysteretic nonlinearity. The MPV of the model parameters are estimated from each data set according to the first step of the proposed algorithm. To identify whether the updated model can match the real structure, the measured accelerations subjected to the base excitation in Fig. 3-7(b) and the corresponding hysteresis loops of the first floor are plotted along with the responses from its optimal calibrated model, as shown in Fig. 3-8. As seen, there is a mismatch between the measurements and the predictions due to the presence of model error. Such discrepancy will be captured by the prediction error term.

Subsequently, the posterior distributions of the hyper-parameters corresponding to 25 data sets are computed based on Eq. (3.23), and the results by using a single acceleration sensor located at the first story are shown in Fig. 3-9. The deviation of the estimated hyper mean values μ_0 from unity indicates their departure from the nominal values of the real structure. It is indicated that the hyper means of the model parameters (linear and nonlinear) are not fairly close to the nominal values of the real structure due to model error, since the updated model cannot match the simulated measurements from the real structure perfectly. It is also noted that the values of hyper standard deviations are considerable for all the model parameters, corresponding to coefficient of variations $\hat{\sigma}_{\theta_i} / \hat{\mu}_{\theta_i}$ ranging from 1.21% to 13.65%, as shown in Table 3-2. These values indicate a relatively large variability in the model parameters from data set to data set. For the hyper parameters of prediction error, again a strong relationship between λ_1 and λ_2 is observed since the posterior distribution of prediction error depends on both of them.

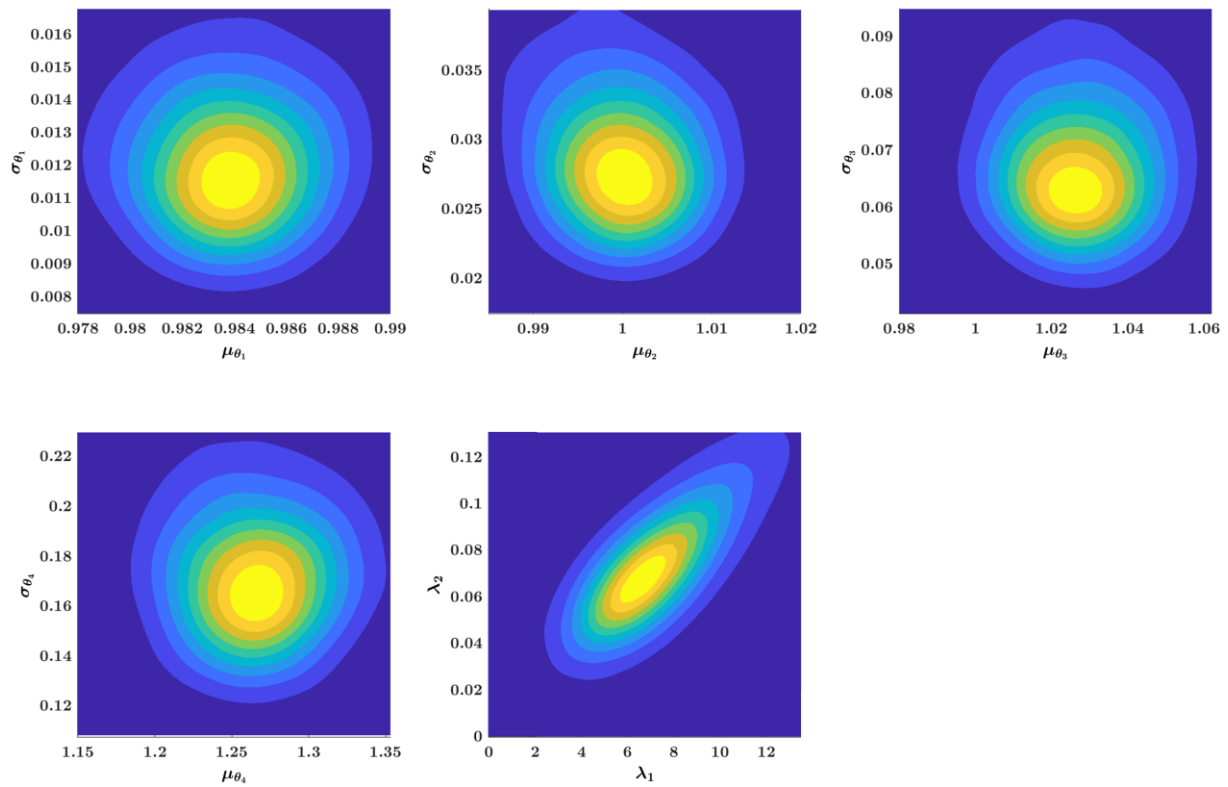


Fig. 3-9 Posterior distributions of hyper parameters computed based on Eq. (3.23)

Table 3-2 Estimates of means of hyper parameters

$\hat{\mu}_{\theta_1}$	0.9838	$\hat{\mu}_{\theta_2}$	1.0004	$\hat{\mu}_{\theta_3}$	1.0268	$\hat{\mu}_{\theta_4}$	1.2661	$\hat{\lambda}_1$	7.2965
$\hat{\sigma}_{\theta_1}$	0.0121	$\hat{\sigma}_{\theta_2}$	0.0286	$\hat{\sigma}_{\theta_3}$	0.0675	$\hat{\sigma}_{\theta_4}$	0.1728	$\hat{\lambda}_2$	0.0750
$\frac{\hat{\sigma}_{\theta_1}}{\hat{\mu}_{\theta_1}}$	1.23%	$\frac{\hat{\sigma}_{\theta_2}}{\hat{\mu}_{\theta_2}}$	2.86%	$\frac{\hat{\sigma}_{\theta_3}}{\hat{\mu}_{\theta_3}}$	6.57%	$\frac{\hat{\sigma}_{\theta_4}}{\hat{\mu}_{\theta_4}}$	13.65%	-	-

The posterior distributions of the hyper parameters are also computed according to the analytical solutions derived in Eqs. (3.27)-(3.32), as shown in Fig. 3-10. Results from the analytical solutions are in good agreement with the ones from Eq. (3.23). It is also noted that a slightly difference for the uncertainty bounds of the hyper parameters can be found between Fig. 3-9 and Fig. 3-10. This is due to the fact that the posterior distribution of the hyper parameters computed by the analytical solutions are under the conditions where the posterior distributions are approximated as a Gaussian distribution. The uncertainties of the hyper parameters are neglected in the analytical solution while the uncertainties of sampling based on Eq. (3.23) are considered for the hyper parameters. However, based on the theoretical results developed in this work, such differences are expected to be reduced for large number of data sets.

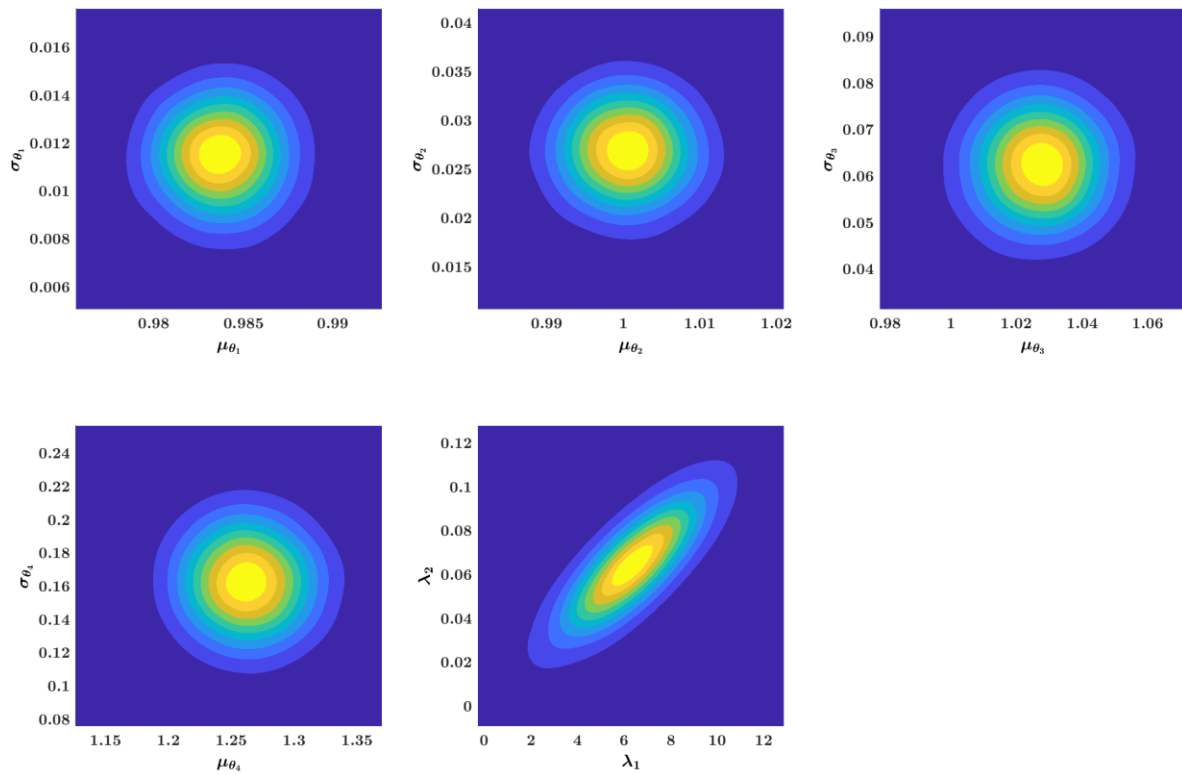


Fig. 3-10 Posterior distributions of hyper parameters computed based on the analytical solutions

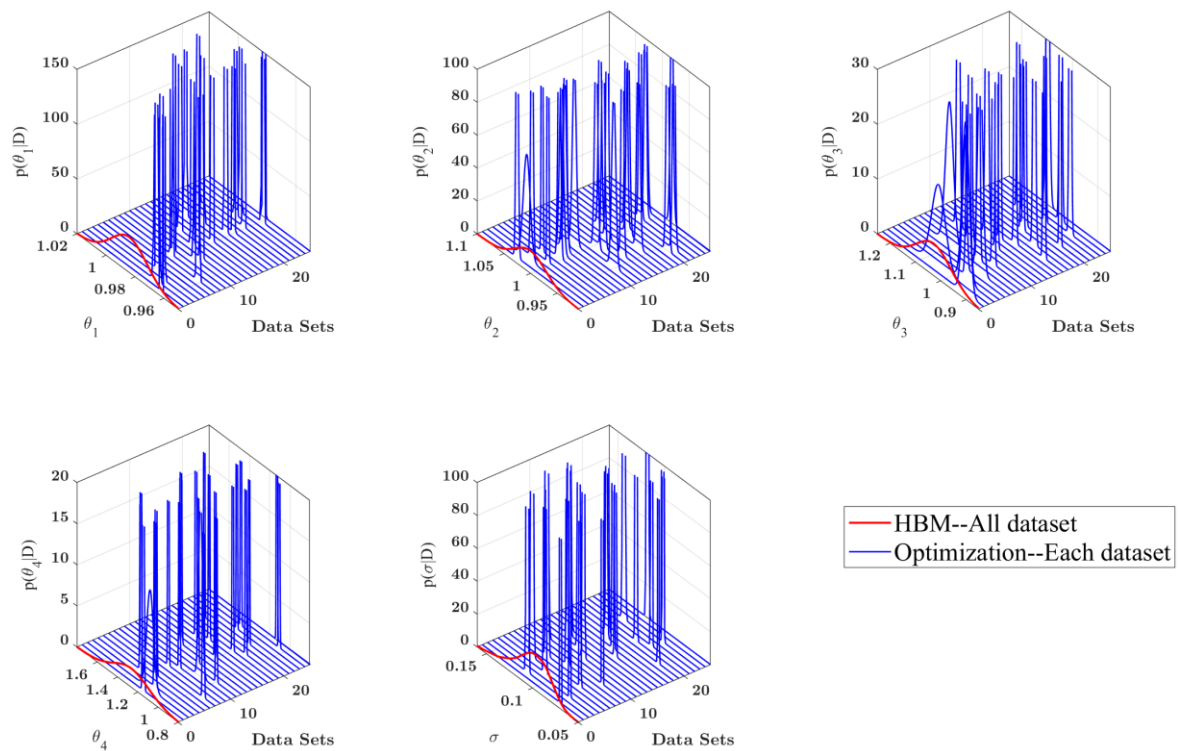


Fig. 3-11 Estimates of identification uncertainty in each dataset and uncertainty due to variability overall datasets

Moreover, the identification uncertainty for each data set calculated by the optimization procedure is compared to the uncertainty due to variability computed by the whole HBM framework. Results for the marginal PDF of the model parameters for each dataset and the marginal PDF of the model parameters estimated from HBM are presented in Fig. 3-11. It is shown that the identification uncertainty of the model and prediction error parameters in each data set (with blue) is much smaller than the ensemble uncertainty over all data sets (with red), and therefore the identification uncertainty can be negligible with sufficient number of data points. It is also seen that the mean of the parameters in each data set varies over different data sets due to the presence of model error. However, most of the mean values fall within the ensemble uncertainty bounds derived from the proposed HBM framework.

For the purpose of comparing the proposed method with the full sampling (FS) method [43,51] as well as the conventional Bayesian method (CBM), results of the mean values of the hyper parameters alongside the prediction error parameters are depicted in Fig. 3-12. It is clear that the proposed approach can capture the ensemble uncertainty of the model parameters and provide clear estimates of the model error while the conventional Bayesian approach fails to reach the same values. Although the CBM offers a relatively good accuracy in terms of the means of the model parameters, it severely underestimates the uncertainty of the model parameters, incorporating such uncertainties into the prediction error term together with the identification uncertainty and the model error uncertainty. As a result, CBM overestimates the value of the prediction error parameter. The proposed method can accurately estimate almost the same values as the FS approach but at a significantly reduced computational effort. The computational effort for the proposed and the FS method, carried out in a computer with a 32-core processor, are reported in Table 3-3. It is clear that the computational effort of the proposed method is around 10 times faster than that with the FS approach in the first step, and around 5 times faster in the second step. Therefore, the proposed method can not only guarantee the computational accuracy but also improve the computational efficiency to a great extent.

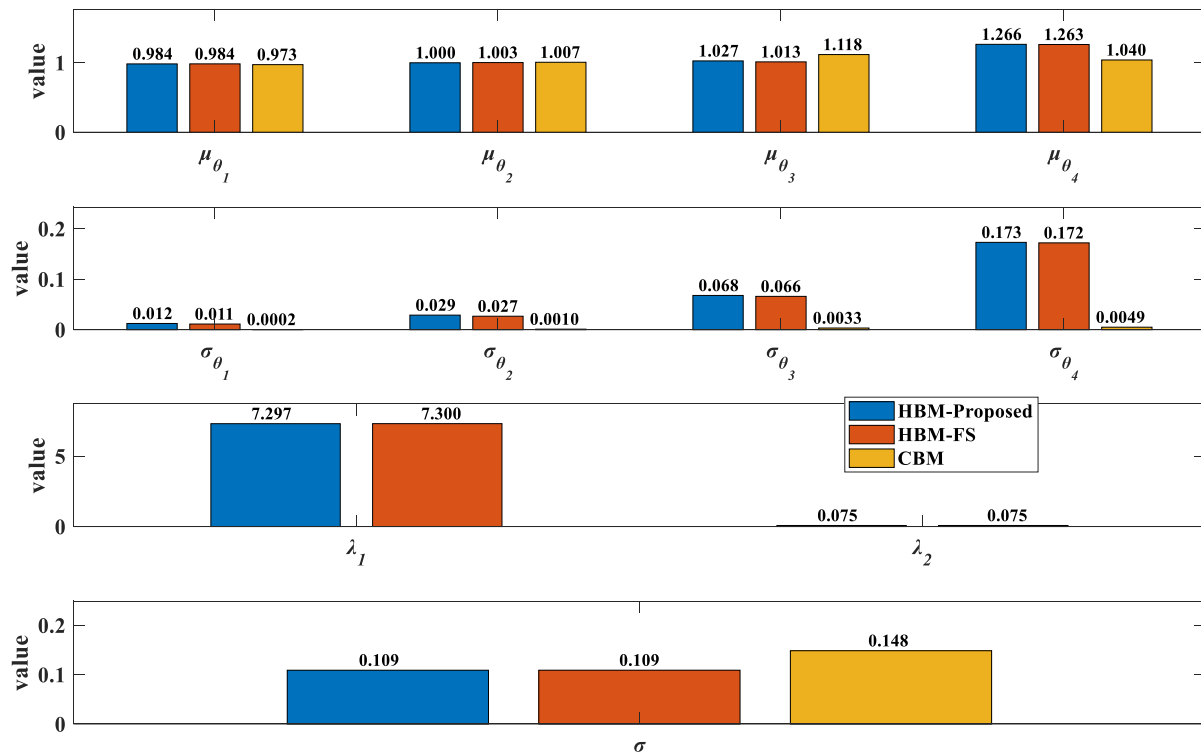


Fig. 3-12 Mean estimates of the hyper parameters and prediction error parameters by using the proposed method, FS and CBM

Table 3-3 Computational effort of the proposed method and FS method

	Step 1	Step 2	Total
Proposed	52s	23s	75s
FS	652s	115s	766s

The estimated uncertainty of the model parameters and the uncertainty from prediction error parameters are next used to predict output QoI. Results are presented for two cases. In the first case, relevant to observed QoI such as accelerations in this example, the uncertainty in both the structural model and prediction error parameters is propagated. In the second case, relevant to unobserved quantities of interest such as displacement, only the uncertainty in the structural model parameters is propagated to the responses since the prediction error term and its uncertainty is not known. Fig. 3-13 depicts the predicted accelerations along with the measured accelerations of the first floor using CBM and HBM. As expected, both methods can offer a good accuracy and most of the actual response is contained within the uncertainty bounds. However, it should be noted that the sources of uncertainties from both CBM and HBM are conceptually different, with significant part of the uncertainty in the HBM method

to be embedded in the structural model parameter, while most of the uncertainty in the CBM to be quantified in the model prediction term. Fig. 3-14 shows uncertainty propagation results of considering only the structural model parameter uncertainty for the predictions. The displacements of the first floor are predicted as the unobserved quantities using CBM and HBM. Fig. 3-15 shows the results for the unobserved displacement of the nonlinear part z_i of the response using CBM and the proposed HBM. It is observed that the proposed method delivers a reasonably accurate uncertainty bound which contains most of the measurements while the conventional Bayesian method provides an extremely thin uncertainty bound where the measurement falls outside.

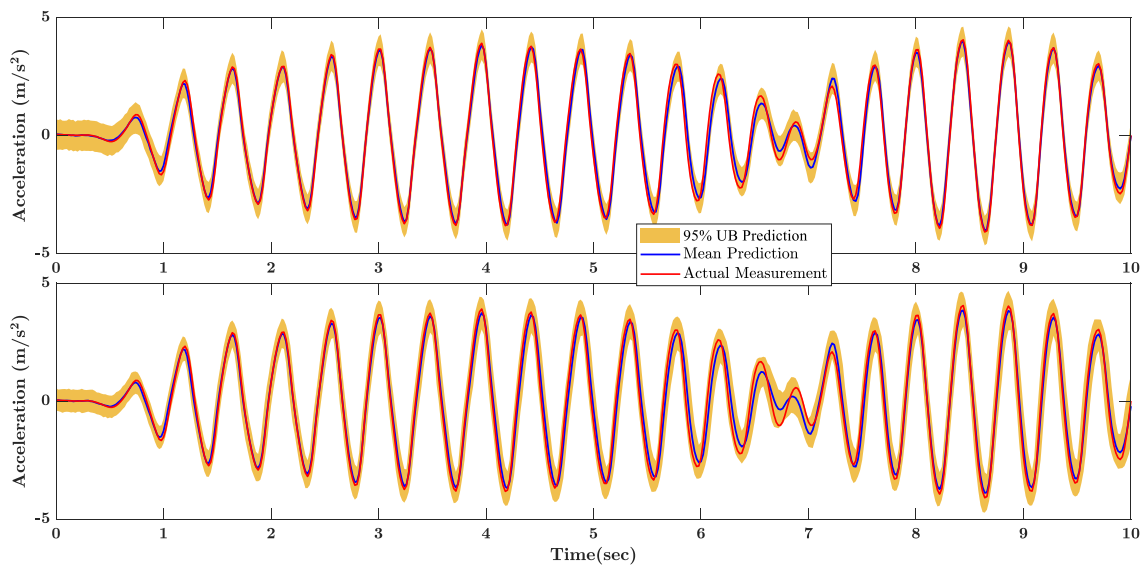


Fig. 3-13 Predictions of accelerations of the first floor using CBM (top) and HBM (bottom)

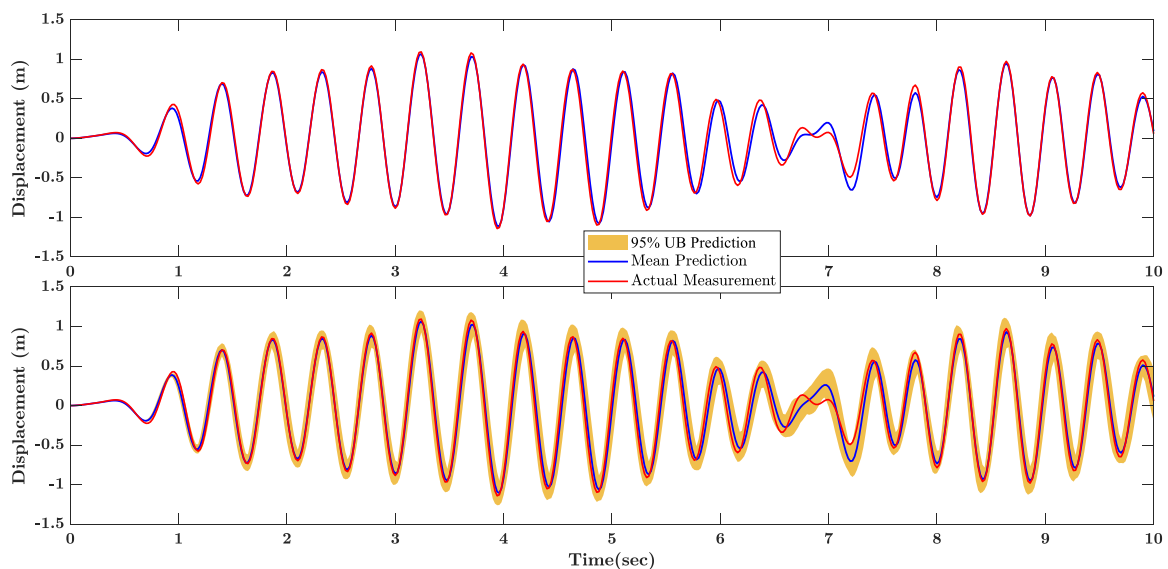


Fig. 3-14 Predictions of displacement of the first floor CBM (top) and HBM (bottom)

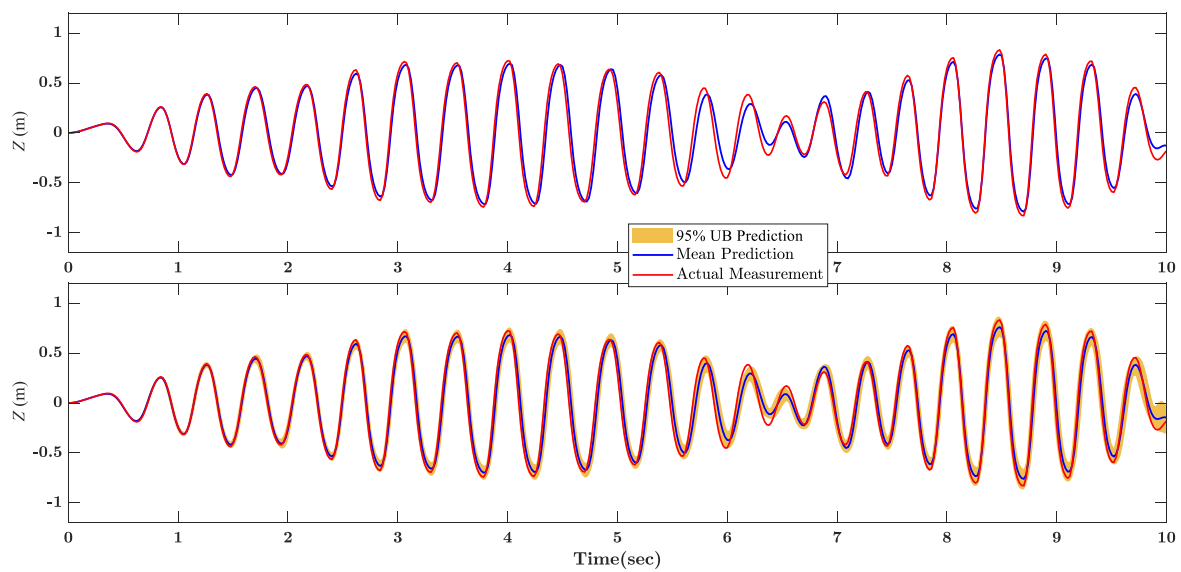


Fig. 3-15 Predictions of nonlinear displacements of the first floor CBM (top) and HBM (bottom)

Finally, the effect of sensor locations and number of sensors on the estimation of the parameters and uncertainties is investigated. Results for the most probable values of the hyper parameters are reported in Table 3-4 for one acceleration sensor placed either at first or fifth floor, 2 acceleration sensors placed at first and fifth floors, as well as 5 acceleration sensors placed at all the floors. It can be seen that different sensor configurations affect the estimates of the hyper parameters. These differences in the estimates from different sensor configurations are expected to affect uncertainties in structural and prediction error model parameters. Results for the posterior PDF of the model and prediction error parameters, computed according to Eqs. (3.34) and (3.35), are shown in Fig. 3-16 for the different sensor configurations considered in Table 3-4. These results suggest that the mean estimate of the parameters as well as the spread of uncertainty in the parameters depend on the number and location of sensors. Measurements from different sensor locations may cause different prediction errors and thus affect the parameter uncertainty due to the presence of model error. Placing a single sensor at the first floor reduces the prediction error. Increasing the number of sensors and thus the number of measurements, although is expected to reduce the identification accuracy by reducing the uncertainties in the estimates of the model parameters for each data set, it does not significantly affect the spread of uncertainty due to the variability from multiple datasets.

Table 3-4 Statistical information of model parameters and prediction error parameter
corresponding to Fig. 3-16

Sensor(s)	μ_{θ_1}	σ_{θ_1}	μ_{θ_2}	σ_{θ_2}	μ_{θ_3}	σ_{θ_3}	μ_{θ_4}	σ_{θ_4}	σ
1 st floor	0.9838	0.0121	1.0004	0.0284	1.0261	0.0653	1.2605	0.1728	0.1040
5 th floor	1.0159	0.0173	0.9201	0.0376	0.8594	0.0991	1.8660	0.2486	0.1615
1 st and 5 th floors	1.0078	0.0139	0.9728	0.0320	0.9253	0.0720	1.8204	0.2399	0.1582
All the floors	1.0085	0.0102	0.9880	0.0197	0.9822	0.0926	1.8730	0.2278	0.1686

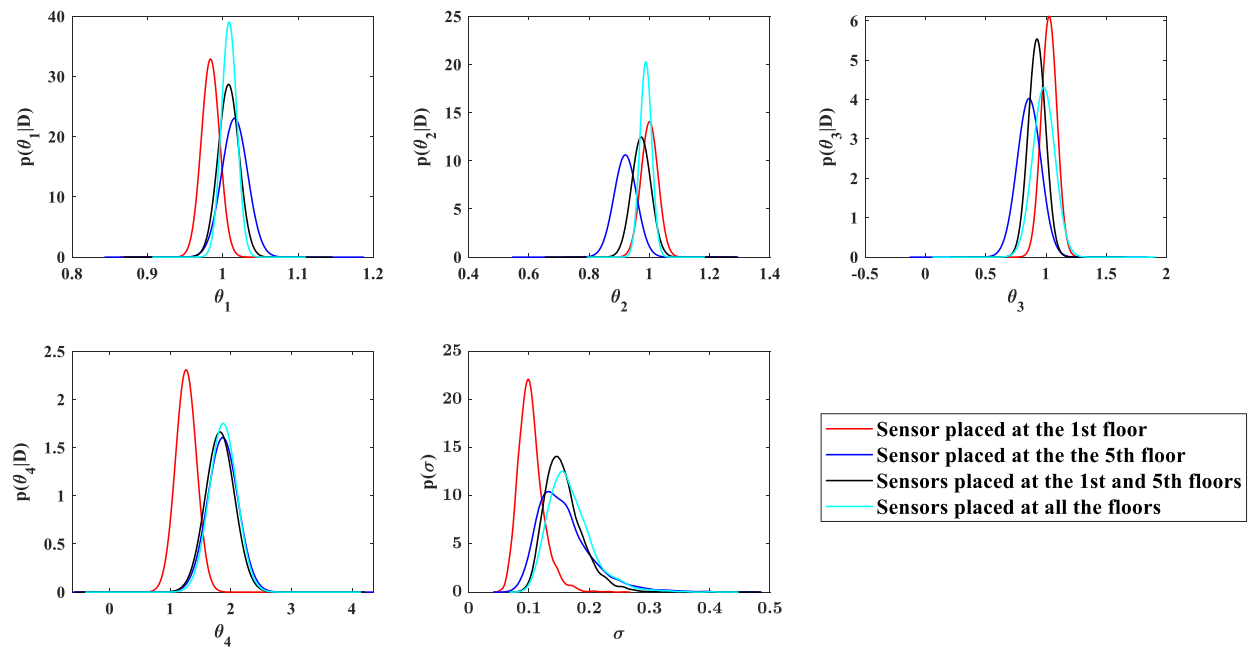


Fig. 3-16 Probabilistic distributions of the model and prediction error parameters

3.5 Conclusion

A nonlinear model updating strategy based on HBM framework is proposed and evaluated when applied to two numerical models of nonlinear systems characterized by Bouc-Wen hysteresis type nonlinearities. According to the HBM framework, uncertainties due to the variability arising from multiple datasets are embedded in the model parameters by assigning a Gaussian distribution to the model parameters with hyper parameters defined as the mean and the covariance matrix of the Gaussian distribution. The unmodelled dynamics are included in Gaussian prediction error by assigning a hyper distribution to the prediction error parameter. Asymptotic approximations are introduced to obtain insightful analytical expression for the posterior distribution of the hyper parameters. The accuracy of the asymptotic approximations is guaranteed due to the large number of data involved in the time history measurements for each dataset. Sampling from this posterior distribution of the hyper

parameters is computationally very efficient since there are no expensive model runs involved at this stage. The resulting analytical expressions provide valuable insight into the mean and covariance of the hyper parameters and their dependence of the most probable values and identification uncertainty of the model parameters estimated for each dataset, as well as the variability in the model parameter estimates arising from multiple datasets. In particular, the analytical expressions were used to show that the structural model hyper parameters are independent from the model prediction error hyper parameters for large number of data within each data set.

Two numerical examples are used to demonstrate the effectiveness and applicability of the proposed HBM framework. It is clearly demonstrated that the irreducible uncertainty arising from the variability of the multiple datasets is due to model error, while the identification uncertainty for each dataset can be reduced as the number of data in each dataset increases. The HBM framework is capable of accounting for the irreducible uncertainty, while conventional Bayesian inference fails to account for such uncertainty, resulting to underestimation of uncertainty bounds for the structural model and prediction error parameters. In particular, embedding the uncertainties due to variability from multiple datasets to the structural model parameters has the effect of obtaining reasonable uncertainty bounds for unobserved response QoI as opposed to significantly underestimated uncertainty bounds obtained from the classical Bayesian inference framework. Moreover, although the number and location of sensors seems to slightly affect the values of the hyper parameters and thus affect the model parameter estimates and uncertainties, increasing the number of sensors does not substantially affect the irreducible uncertainty arising from the variability due to multiple datasets. Finally, the asymptotic approximation is shown to provide one order of magnitude reduction of the computational effort compared to existing full sampling approach for the considered case study.

Appendix A. Calculation for the integral in Eq. (3.10)

Substituting the likelihood function $p(D_i | \theta_i, \sigma_i^2)$ from Eq. (3.7) and the inverse gamma distribution from Eq. (3.6) to the integral in Eq. (3.10) yields:

$$\int_{\sigma_i^2} p(D_i | \theta_i, \sigma_i^2) IG(\sigma_i^2 | \lambda_1, \lambda_2) d\sigma_i^2 \propto \frac{(\lambda_2)^{\lambda_1}}{\Gamma(\lambda_1)} \int_{\sigma_i^2} \sigma_i^{-N_0 N_i - 2\lambda_1 - 2} \exp \left\{ -\frac{\frac{N_0 N_i}{2} J(\theta_i) + \lambda_2}{\sigma_i^2} \right\} d\sigma_i^2 \quad (A.1)$$

The integral in Eq. (A.1) can be obtained analytically as:

$$\int_{\sigma_i^2} \sigma_i^{-N_0 N_i - 2\lambda_1 - 2} \exp \left\{ -\frac{\frac{N_0 N_i}{2} J(\boldsymbol{\theta}_i) + \lambda_2}{\sigma_i^2} \right\} d\sigma_i^2 = \Gamma(f(\lambda_1)) \left[\frac{N_0 N_i}{2} J(\boldsymbol{\theta}_i) + \lambda_2 \right]^{-f(\lambda_1)} \quad (\text{A.2})$$

where $\Gamma(\bullet)$ is the Gamma function and $f(\lambda_1)$ is defined as:

$$f(\lambda_1) = \frac{N_0 N_i + 2\lambda_1}{2} \quad (\text{A.3})$$

Substituting (A.2) into (A.1) one derives Eq. (3.11).

Appendix B. Calculation for the MPV and hessian matrix in Eq. (3.15) and Eq. (3.16)

The first order derivative of function $L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)$ in Eq. (3.13) with respect to parameters $\boldsymbol{\theta}_i$ can be computed as:

$$\nabla_{\boldsymbol{\theta}_i} L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2) = \frac{N_0 N_i f(\lambda_1)}{N_0 N_i J(\boldsymbol{\theta}_i) + 2\lambda_2} \nabla_{\boldsymbol{\theta}_i} J(\boldsymbol{\theta}_i) \quad (\text{B.1})$$

Setting the derivatives equal to zero one can obtain the MPV of parameters $\boldsymbol{\theta}_i$ as:

$$\hat{\boldsymbol{\theta}}_i = \arg \min_{\boldsymbol{\theta}_i} (J(\boldsymbol{\theta}_i)) \quad (\text{B.2})$$

The hessian of the function $L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2)$ can be also calculated as follows:

$$\nabla_{\boldsymbol{\theta}_i} \nabla_{\boldsymbol{\theta}_i}^T L(\boldsymbol{\theta}_i, \lambda_1, \lambda_2) = N_0 N_i f(\lambda_1) \frac{\mathbf{H}(\boldsymbol{\theta}_i) (N_0 N_i J(\boldsymbol{\theta}_i) + 2\lambda_2) - N_0 N_i \nabla_{\boldsymbol{\theta}_i} J(\boldsymbol{\theta}_i) \nabla_{\boldsymbol{\theta}_i}^T J(\boldsymbol{\theta}_i)}{(N_0 N_i J(\boldsymbol{\theta}_i) + 2\lambda_2)^2} \quad (\text{B.3})$$

where $\mathbf{H}(\boldsymbol{\theta}_i) = \nabla_{\boldsymbol{\theta}_i} \nabla_{\boldsymbol{\theta}_i}^T J(\boldsymbol{\theta}_i)$ is the hessian of the function $J(\boldsymbol{\theta}_i)$. Substituting $\hat{\boldsymbol{\theta}}_i$ calculated from Eq. (B.2) into Eq. (B.3) and noting that $\nabla_{\boldsymbol{\theta}_i} J(\hat{\boldsymbol{\theta}}_i) = \mathbf{0}$ one can get the hessian matrix evaluated at the MPV $\hat{\boldsymbol{\theta}}_i$ of the parameters in the form given in Eq. (3.16).

3.6 References

- [1] T.J.R. Hughes, The finite element method: linear static and dynamic finite element analysis, Courier Corporation, 2012.
- [2] O.C. Zienkiewicz, R.L. Taylor, Finite Element Method for Solid and Structural Mechanics, Elsevier, 2005.
- [3] V. Papadopoulos, M. Papadrakakis, G. Deodatis, Analysis of mean and mean square response of general linear stochastic finite element systems, Comput. Methods Appl. Mech. Eng. 195 (2006) 5454–5471.

- [4] J. Guggenberger, H. Grundmann, Stochastic response of large FEM models with hysteretic behaviour in beam elements, *Comput. Methods Appl. Mech. Eng.* 194 (2005) 1739–1756.
- [5] J.E. Mottershead, M.I. Friswell, Model updating in structural dynamics: A survey, *J. Sound Vib.* 167 (1993) 347–375.
- [6] M. Friswell, J.E. Mottershead, *Finite element model updating in structural dynamics*, Springer Science & Business Media, 2013.
- [7] N. Distefano, A. Rath, System identification in nonlinear structural seismic dynamics, *Comput. Methods Appl. Mech. Eng.* 5 (1975) 353–372.
- [8] N. Distefano, A. Rath, Sequential identification of hysteretic and viscous models in structural seismic dynamics, *Comput. Methods Appl. Mech. Eng.* 6 (1975) 219–232.
- [9] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, *J. Eng. Mech.* 124 (1998) 455–461.
- [10] L.S. Katafygiotis, J.L. Beck, Updating models and their uncertainties. II: Model identifiability, *J. Eng. Mech.* 124 (1998) 463–467.
- [11] E. Simoen, C. Papadimitriou, G. Lombaert, On prediction error correlation in Bayesian model updating, *J. Sound Vib.* 332 (2013) 4136–4152.
- [12] H. Jensen, C. Papadimitriou, Bayesian Finite Element Model Updating, in: *Sub-Structure Coupling Dyn. Anal.*, Springer, 2019: pp. 179–227.
- [13] M. Song, S. Yousefianmoghadam, M.E. Mohammadi, B. Moaveni, A. Stavridis, R.L. Wood, An application of finite element model updating for damage assessment of a two-story reinforced concrete building and comparison with lidar, *Struct. Heal. Monit.* 17 (2018) 1129–1150.
- [14] M. Papadrakakis, V. Papadopoulos, N.D. Lagaros, Structural reliability analysis of elastic-plastic structures using neural networks and Monte Carlo simulation, *Comput. Methods Appl. Mech. Eng.* 136 (1996) 145–163.
- [15] H.A. Jensen, C. Vergara, C. Papadimitriou, E. Millas, The use of updated robust reliability measures in stochastic dynamical systems, *Comput. Methods Appl. Mech. Eng.* 267 (2013).
- [16] S. Bansal, S.H. Cheung, A new stochastic simulation algorithm for updating robust reliability of linear structural dynamic systems subjected to future Gaussian excitations, *Comput. Methods Appl. Mech. Eng.* 326 (2017) 481–504.

- [17] E. Asgarieh, B. Moaveni, A. Stavridis, Nonlinear finite element model updating of an infilled frame based on identified time-varying modal parameters during an earthquake, *J. Sound Vib.* 333 (2014) 6057–6073.
- [18] Z.C. Wang, Y. Xin, W.X. Ren, Nonlinear structural joint model updating based on instantaneous characteristics of dynamic responses, *Mech. Syst. Signal Process.* 76–77 (2016) 476–496.
- [19] J. Ma, G. Chen, L. Ji, L. Qian, S. Dong, A general methodology to establish the contact force model for complex contacting surfaces, *Mech. Syst. Signal Process.* 140 (2020) 106678.
- [20] W.-X. Ren, H.-B. Chen, Finite element model updating in structural dynamics by using the response surface method, *Eng. Struct.* 32 (2010) 2455–2465.
- [21] M. Kurt, M. Eriten, D.M. McFarland, L.A. Bergman, A.F. Vakakis, Methodology for model updating of mechanical components with local nonlinearities, *J. Sound Vib.* 357 (2015) 331–348.
- [22] E. Asgarieh, B. Moaveni, A.R. Barbosa, E. Chatzi, Nonlinear model calibration of a shear wall building using time and frequency data features, *Mech. Syst. Signal Process.* 85 (2017) 236–251.
- [23] J.L. Beck, S.-K. Au, Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation, *J. Eng. Mech.* 128 (2002) 380–391.
- [24] K.V. Yuen, *Bayesian Methods for Structural Dynamics and Civil Engineering*, 2010.
- [25] K.V. Yuen, Recent developments of Bayesian model class selection and applications in civil engineering, *Struct. Saf.* 32 (2010) 338–346.
- [26] W. Betz, I. Papaioannou, J.L. Beck, D. Straub, Bayesian inference with Subset Simulation: Strategies and improvements, *Comput. Methods Appl. Mech. Eng.* 331 (2018) 72–93.
- [27] I. Babuška, F. Nobile, R. Tempone, A systematic approach to model validation based on Bayesian updates and prediction related rejection criteria, *Comput. Methods Appl. Mech. Eng.* 197 (2008) 2517–2539.
- [28] M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, *J. Vib. Control.* 14 (2008) 7–34.
- [29] H. Ebrahimian, R. Astroza, J.P. Conte, R.A. de Callafon, Nonlinear finite element model updating for damage identification of civil structures using batch Bayesian estimation, *Mech. Syst. Signal Process.* 84 (2017) 194–222.

- [30] M. Song, L. Renson, J.P. Noël, B. Moaveni, G. Kerschen, Bayesian model updating of nonlinear systems using nonlinear normal modes, *Struct. Control Heal. Monit.* 25 (2018) 1–20.
- [31] R. Ceravolo, A. Faraci, G. Miraglia, Bayesian calibration of hysteretic parameters with consideration of the model discrepancy for use in seismic structural health monitoring, *Appl. Sci.* 10 (2020).
- [32] K.-V. Yuen, J.L. Beck, Updating Properties of Nonlinear Dynamical Systems with Uncertain Input, *J. Eng. Mech.* 129 (2003) 9–20.
- [33] J. Ching, J.L. Beck, K.A. Porter, Bayesian state and parameter estimation of uncertain dynamical systems, *Probabilistic Eng. Mech.* 21 (2006) 81–96.
- [34] E.N. Chatzi, A.W. Smyth, S.F. Masri, Experimental application of on-line parametric identification for nonlinear hysteretic systems with model uncertainty, *Struct. Saf.* 32 (2010) 326–337.
- [35] E.N. Chatzi, Andrew W. Smyth, The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing, *Struct. Control Heal. Monit.* (2011) 99–123.
- [36] M. Wu, A.W. Smyth, Application of the unscented Kalman filter for real-time nonlinear structural system identification, *Struct. Control Heal. Monit.* 14(7) (2007) 971–990.
- [37] P. Liu, S.K. Au, Bayesian parameter identification of hysteretic behavior of composite walls, *Probabilistic Eng. Mech.* 34 (2013) 101–109.
- [38] R. Sandhu, C. Pettit, M. Khalil, D. Poirel, A. Sarkar, Bayesian model selection using automatic relevance determination for nonlinear dynamical systems, *Comput. Methods Appl. Mech. Eng.* 320 (2017) 237–260.
- [39] M. Ebrahimzadeh Hassanabadi, A. Heidarpour, S. Eftekhari Azam, M. Arashpour, Recursive principal component analysis for model order reduction with application in nonlinear Bayesian filtering, *Comput. Methods Appl. Mech. Eng.* 371 (2020) 113334.
- [40] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Accounting for Modeling Errors and Inherent Structural Variability through a Hierarchical Bayesian Model Updating Approach: An Overview, *Sensors*. 20 (2020) 3874.
- [41] G.C. Ballesteros, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian hierarchical models for uncertainty quantification in structural dynamics, in: *Vulnerability, Uncertainty, Risk Quantif. Mitigation, Manag.*, 2014: pp. 1615–1624.

- [42] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, *Mech. Syst. Signal Process.* 64–65 (2015) 360–376.
- [43] S. Wu, P. Angelikopoulos, G. Tauriello, C. Papadimitriou, P. Koumoutsakos, Fusing heterogeneous data for the calibration of molecular dynamics force fields using hierarchical Bayesian models, *J. Chem. Phys.* 145 (2016) 244112.
- [44] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, *J. Sound Vib.* 374 (2016) 92–110.
- [45] X.-W. Liu, D.-G. Lu, Survival analysis of fatigue data: Application of generalized linear models and hierarchical Bayesian model, *Int. J. Fatigue.* 117 (2018) 39–46.
- [46] S. Wu, P. Angelikopoulos, J.L. Beck, P. Koumoutsakos, Hierarchical Stochastic Model in Bayesian Inference for Engineering Applications: Theoretical Implications and Efficient Approximation, *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B Mech. Eng.* 5 (2019).
- [47] M. Song, B. Moaveni, C. Papadimitriou, A. Stavridis, Accounting for amplitude of excitation in model updating through a hierarchical Bayesian approach: Application to a two-story reinforced concrete building, *Mech. Syst. Signal Process.* 123 (2019) 68–83.
- [48] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Probabilistic hierarchical Bayesian framework for time-domain model updating and robust predictions, *Mech. Syst. Signal Process.* 123 (2019) 648–673.
- [49] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Data-driven uncertainty quantification and propagation in structural dynamics through a hierarchical Bayesian framework, *Probabilistic Eng. Mech.* 60 (2020) 103047.
- [50] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Properties, (2021). <https://zenodo.org/record/5078051#.YOV0LrUzbD4>.
- [51] D. Patsialis, A.P. Kyprioti, A.A. Taflanidis, Bayesian calibration of hysteretic reduced order structural models for earthquake engineering applications, *Eng. Struct.* 224 (2020) 111204.

- [52] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, *J. Eng. Mech.* 133 (2007) 816–832.
- [53] S. Wu, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian annealed sequential importance sampling: an unbiased version of transitional Markov chain Monte Carlo, *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B Mech. Eng.* 4 (2018).
- [54] M. Coughlin, N. Christensen, J. Gair, S. Kandhasamy, E. Thrane, Method for estimation of gravitational-wave transient model parameters in frequency-time maps, *Class. Quantum Gravity.* 31 (2014).
- [55] O. Sedehi, L.S. Katafygiotis, C. Papadimitriou, Hierarchical Bayesian operational modal analysis: Theory and computations, *Mech. Syst. Signal Process.* 140 (2020) 106663.
- [56] R. Bouc, A mathematical model for hysteresis, *Acta Acust. United with Acust.* 24 (1971) 16–25.
- [57] Y.-K. Wen, Method for random vibration of hysteretic systems, *J. Eng. Mech. Div.* 102 (1976) 249–263.
- [58] F. Ikhoulane, V. Mañosa, J. Rodellar, Dynamic properties of the hysteretic Bouc-Wen model, *Syst. Control Lett.* 56 (2007) 197–205.
- [59] F. Ikhoulane, J. Rodellar, *Systems with Hysteresis: Analysis, Identification and Control using the Bouc-Wen Model*, 2007.
- [60] P. Sengupta, B. Li, Modified Bouc-Wen model for hysteresis behavior of RC beam-column joints with limited transverse reinforcement, *Eng. Struct.* 46 (2013) 392–406.
- [61] M. Ismail, F. Ikhoulane, J. Rodellar, The hysteresis Bouc-Wen model, a survey, *Arch. Comput. Methods Eng.* 16 (2009) 161–188.
- [62] J.R. Cash, A.H. Karp, A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides, *ACM Trans. Math. Softw.* 16 (1990) 201–222.
- [63] Z. Cao, Q. Fei, D. Jiang, D. Zhang, H. Jin, R. Zhu, Dynamic sensitivity-based finite element model updating for nonlinear structures using time-domain responses, *Int. J. Mech. Sci.* 184 (2020).
- [64] Luzi L., Lanzano G., Felicetta C., D’Amico M. C., Russo E., Sgobba S., Pacor, F., & ORFEUS Working Group 5 (2020). Engineering Strong Motion Database (ESM) (Version 2.0). Istituto Nazionale di Geofisica e Vulcanologia (INGV). <https://doi.org/10.13127/ESM.2>

Chapter 4. Hierarchical Bayesian Learning Framework for Multi-Level Modeling using Multi-level Data

Original Paper:

X. Jia, C. Papadimitriou, Hierarchical Bayesian Learning Framework for Multi-Level Modeling using Multi-level Data, Mechanical Systems and Signal Processing. Submitted. (2021). <https://doi.org/10.5281/zenodo.5702385>

ABSTRACT

A hierarchical Bayesian learning framework is proposed to account for multi-level modeling in structural dynamics. In multi-level modeling the system is considered as a hierarchy of lower-level models, starting at the lowest material level, progressing to the component level, then the subsystem level, before ending up to the system level. Bayesian modeling and uncertainty quantification techniques based on measurements that rely on data collected only at the system level cover a quite limited number of component/subsystem operating conditions that are far from representing the full spectrum of system operating conditions. In addition, the large number of models and parameters involved from the lower to higher modeling levels of the system, constitutes this approach inappropriate for simultaneously and reliably quantifying the uncertainties at the different modeling levels. In this work, comprehensive hierarchical Bayesian learning tools are proposed to account for uncertainties through the multi-level modeling process. The uncertainty is embedded within the structural model parameters by introducing a probability model for these parameters that depend on hyperparameters. An important issue that has to be accounted for is that parameters of models at lower levels are shared at the subsystem and system levels. This necessitates a parameter inference process that takes into account data from different modeling levels. Accurate and insightful asymptotic approximations are developed, substantially reducing the computational effort required in the parameter uncertainty quantification procedure. The uncertainties inferred based on datasets available from the different levels of model hierarchy are

propagated through the different levels of the system to predict uncertainties and confidence levels of output quantities of interest. A simple dynamical system consisting of components and subsystems is employed to demonstrate the effectiveness of the proposed method.

4.1 Introduction

Computational finite element (FE) models are extensively employed to represent physical structures and perform analyses in structural dynamics [1,2]. FE model updating has been an essential step to assess and improve the accuracy of the computational models [3,4]. In the process of modeling a structure, uncertainties arise from variabilities in experimental tests, environmental and operational conditions [5], manufacturing variabilities [6], material properties variabilities, as well as model error [5,7]. However, physical structures consist of multiple components, similar or dissimilar, that are assembled to subsystems that comprise the overall system [8]. Such structures often exhibit complexities resulting in the uncertainty due to variability arising from the assembling process, manufacturing process as well as nonlinearities manifested at various modeling levels during operation under harsh environments [9,10].

Several contributions have already been made for capturing such uncertainties and improving the accuracy of the computational models in multi-level structures. Statistical model calibration technique for a hierarchy system was proposed to improve the predictive capability of the computational models from the lowest level to the highest level [11]. Model parameters at a lower level was calibrated using a conventional Bayesian modeling approach from test data at the same level. Data from higher levels of model hierarchy had no effect on learning parameters involved at lower modeling level. An alternative Bayesian learning procedure was also used in [12], where data at the lowest level were used to construct prior probability distribution of the model parameters and then data from the next level of model hierarchy were used to update the model parameters. A hierarchical model updating strategy using uncorrelated modes was also developed for complex assembled structure to ensure a high accuracy of the updated FE model [13]. Uncertainties inferred at different levels were propagated to uncertainties to quantify the confidence in predictions at different levels [11] and at the system level [14]. Design of validation experiments for multi-level models was also proposed in [8]. The aforementioned developments in hierarchical modeling based on Bayesian calibration methods do not use experimental data from different levels of model hierarchy to simultaneously inform the parameters of models at all levels of hierarchy and

infer the variability of the parameter estimates even at the lowest level of model hierarchy. This is proposed to be addressed in this work using a hierarchical Bayesian modeling framework.

The hierarchical Bayesian modeling (HBM) framework was recently proposed in various engineering fields to capture the uncertain variability [15–18]. Full sampling-based HBM framework is commonly used to calibrate the model parameters, quantify the uncertainties and predict their uncertainties to the system quantities of interest (QoI) using data obtained from the system level [19–23]. The up-to-date full-sampling HBM approach is also applied to a two-level masonry structural [24] as well as a red blood cell multi-level model to inform the parameters of models at all levels of model hierarchy [25]. Undoubtedly, the above-mentioned approaches show a good performance on quantifying the uncertainties or improving the models for a multi-level structure. However, there is still room for improvement about the process of calibrating the models in a multi-level structure. For example, model error is neglected in a multi-level structure leading to the missing information of the discrepancy between the responses of physical structures and the outputs of computational models [11]. Additionally, the sampling procedure utilized during the process of updating the models makes the HBM approach inefficient and often provides no insight into the estimated uncertainties. For improving the efficiency and obtaining valuable insights of HBM, approaches based on asymptotic approximations have been developed to handle the uncertainty quantification and propagation in structural dynamics problems [26–29]. The existing HBM framework is conducted using information collected from the system level. However, the framework has great potential to be implemented to a multi-level modeling of the system using multi-level data for the purpose of model validation, uncertainty quantification and propagation.

To this end, this paper extends the recently proposed HBM framework [29] to a hierarchy structure, aiming to take into account the diverse sources of the uncertainty through the hierarchy of models of the systems and propagate it through the computation models for improving the confidence in the response predictions. It is noted that, given the hierarchy of models for a structure, the information from a lower level of the structure is also inferred from the data at the higher levels. This guides us to build reliable models at all system levels that account in the modeling and parameter estimation process for the uncertain variability arising from component manufacturing and assembling process. This is achieved by assigning a probabilistic distribution for the model parameters, where the hyper parameters are inferred from datasets available at lower and higher levels of modeling. The estimation

process of the hyper parameters is conducted based on the available datasets, and the response QoI are predicted according to the uncertainties obtained from the parameter estimation process. Asymptotic approximations are introduced to the framework, providing insightful expressions and efficiently improving the computational cost of the proposed framework.

This work is structured as follows. Section 4.2 describes the proposed hierarchical Bayesian modeling framework for multi-level models, including the formulations for quantifying the uncertainties and propagating the uncertainties through the computational models manifested in the system hierarchy. Section 4.3 demonstrates the effectiveness and accuracy of the proposed method using a carefully designed numerical case study. Section 4.4 reports the conclusions of this study.

4.2 Hierarchical Bayesian Modeling Framework for Multi-level Models

The HBM framework developed in [29] is used as a tool to integrate the multiple experimental datasets and the hierarchy of models within a dynamical structure. The major difference between this work and the work in [29] is that, instead of using the available information from a full-scale system or a single component, the proposed strategy aims to take into account the diverse sources of the uncertainty through the multi-level structure to estimate model parameters and their uncertainties as well as propagate uncertainties through the computational models for the response predictions. The details for the multi-level models along with their uncertainties are introduced below.

4.2.1 Multi-level models and data

For the purpose of demonstrating the methodology and without loss of generality, consider a system hierarchy where the higher level includes two lower levels. Specifically, it is assumed that the system is assembled from two subsystems (subsystem SS1 and subsystem SS2) with the subsystem SS1 assembled from three components C1, C2 and C3, as shown in Fig. 4-1(a). Experimental data are available at the subsystem SS1 level and the component C1 and C2 levels. There are no measurements available at the system level, the subsystem SS2 level and component C3 level. Parameterized models are introduced to represent the components C1, C2 and C3, with parameters to be identified using the available experimental data from components C1, C2 and subsystem SS1. Specifically, parameterized physics-based models M_1 , M_2 and M_3 are considered for each structural component C1, C2 and C3 with the

structural model parameters denoted by $\theta^{(1)}$, $\theta^{(2)}$ and $\theta^{(3)}$, respectively. No experimental data sets are available for the subsystem SS2 so that the models associated with this subsystem and its possible components that is composed of are not updated using data. The parameterized model for subsystem SS1 level and the corresponding component levels are depicted in Fig. 4-1(b). Details for the available testing procedures, experimental data and the process for estimating the parameters of the parameterized models are given in the next subsections.

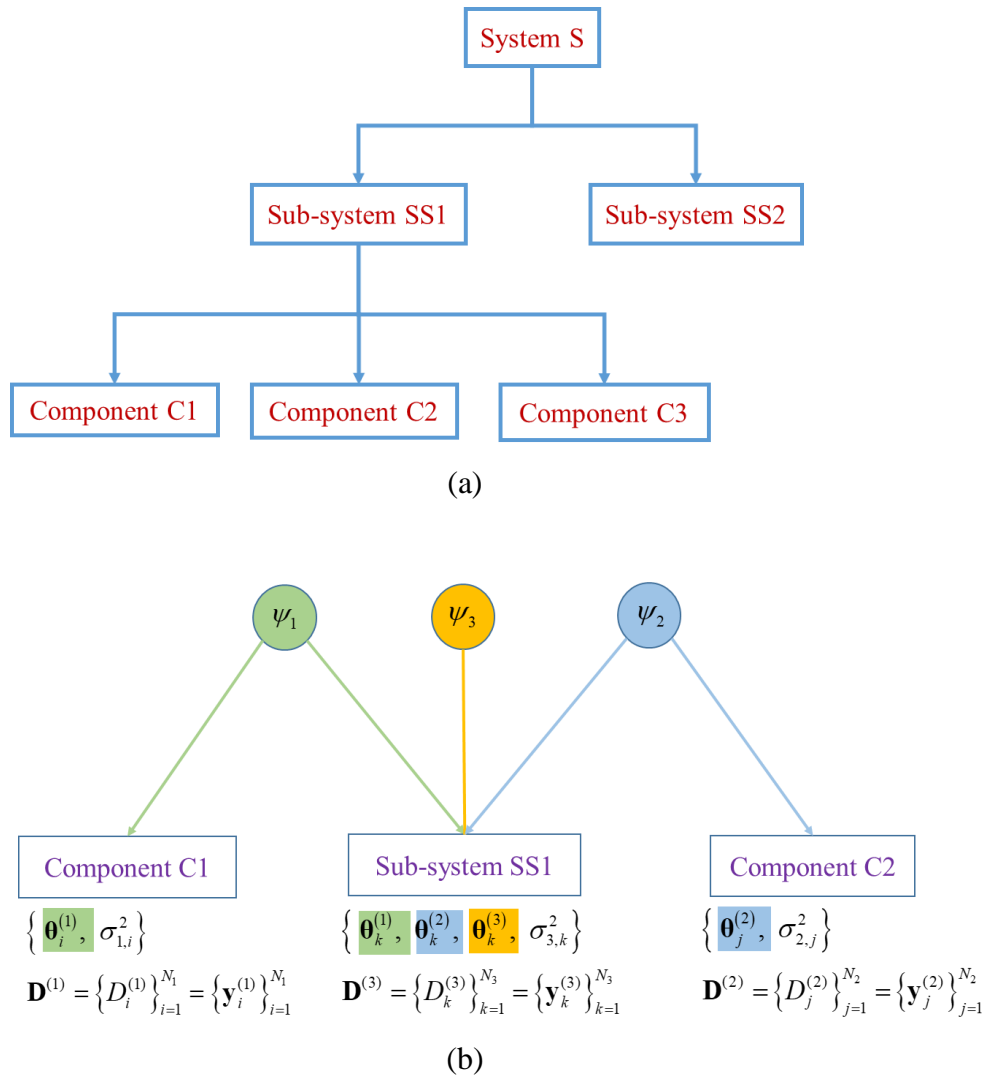


Fig. 4-1 Graphical representation of (a) decomposition of a physical structure into components, subsystems and system and (b) parameterized models and datasets available

4.2.1.1 Test data at component level

Multiple datasets are considered to be available for each one of the two components C1 and C2. For component C1, the available datasets are denoted by $D_i^{(1)}$, where the superscript (1) represents the first component, and the subscript i corresponds to the i -th data set, $i = 1, 2, \dots, N_1$, while N_1 is the number of data sets available for component C1. The datasets are obtained by testing the component C1 as an isolated component before it is assembled into the subsystem and eventually to the system. The model of the test configuration of the isolated component C1 includes the parameter set $\theta^{(1)}$. Two types of testing can be performed. In the first type, denoted as testing Type A, it is considered that a single structural component is available and multiple tests are carried out for this component. Due to model error, measurement error and variability in the experimental testing, the parameter estimates are expected to vary from dataset to dataset [29,30]. The measured data collected in this way are used to quantify the uncertainty in the parameter set $\theta^{(1)}$ due to variability in the datasets. In the second type of testing, denoted by testing Type B, the component is considered to be a member of a population of components manufactured/assembled to be identical. However, differences from component to component in the population exist due to manufacturing/assembled variabilities. Each dataset $D_i^{(1)}$ is referred to a test carried out at member i of the population of available structural components C1. Tests over a subset of N_1 components in the population are used to infer the values and the uncertainties of the model parameter set $\theta^{(1)}$ due to the variabilities in all components in the population (tested and non-tested). The component C1 in this case is considered to be one of the members in the population. To account for the uncertainty due to variability in the datasets for a single component (testing Type A) or due to the manufacturing/assembled variability of selected component from the population (testing Type B), uncertainties are embedded in the model parameters by assigning a parameterized probability distribution $p(\theta^{(1)}|\psi_1)$ with the hyperparameter set ψ_1 to be inferred from the available multiple datasets.

A similar testing approach is followed for component C2 to infer the corresponding hyperparameters ψ_2 of the uncertainty, quantified by the prior PDF $p(\theta^{(2)}|\psi_2)$, assigned to the model parameter set $\theta^{(2)}$, where the available datasets are denoted by $D_j^{(2)}$, $j = 1, 2, \dots, N_2$.

Similarly to components C1 and C2, component C3 can be a single structural component, or it can be a member of a population of identically manufactured/assembled components.

Uncertainties are embedded in the model parameters $\theta^{(3)}$ by assigning a prior probability distribution $p(\theta^{(3)}|\psi_3)$ for the parameter set $\theta^{(3)}$. No datasets are available at the level of the component C3 to identify the parameter set $\theta^{(3)}$ and the hyperparameter set ψ_3 . The parameter set $\theta^{(3)}$ and the hyperparameter set ψ_3 of the model of this component will be inferred using the multiple experimental datasets available at the subsystem SS1 level as described in the next subsection.

4.2.1.2 Test data at sub-system level

It is assumed that datasets $D_k^{(3)}$, $k=1,2,\dots,N_3$, are available for subsystem SS1, while no datasets are available for subsystem SS2. The datasets are obtained by testing the subsystem SS1 as an isolated subsystem before it is assembled with subsystem SS2 to form the system. The model of the test configuration of the isolated subsystem SS1 includes the parameter set $\theta^{(1)}$, $\theta^{(2)}$ and $\theta^{(3)}$. These datasets at the subsystem SS1 level provide information for inferring the parameters $\theta^{(1)}$, $\theta^{(2)}$ and $\theta^{(3)}$ of the models of all three components associated with this subsystem. Noting that the parameter set $\theta^{(1)}$ of component C1 is shared at the component C1 and the subsystem SS1 levels, the data from the component C1 and subsystem SS1 levels are used to infer the values and its uncertainties of this parameter set. Similar is the case for the model parameter set $\theta^{(2)}$ shared at component level C2 and subsystem SS1 level with measured data available from both levels. For the parameter set $\theta^{(3)}$ of the model M_3 introduced at the component level C3, there are no measurements available at the component level. The measurements at the subsystem level are used to infer the parameter set $\theta^{(3)}$, jointly with the model parameter sets $\theta^{(1)}$ and $\theta^{(2)}$ of the other two components. Finally, there are no measurements available to infer the parameters of models introduced to represent the subsystem SS2 and its possible components that is comprised of. In this case the models at subsystem SS2 level can be considered known or prior information can be used to subjectively assign uncertainties in the parameters of the models involved.

Similar to the component level, two types of testing can be performed at the subsystem level. In the Type A testing, a single subsystem SS1 is manufactured from the components C1, C2 and using a single component C3. Multiple tests are carried out for this specific subsystem, providing the datasets $D_k^{(3)}$, $k=1,2,\dots,N_3$. Due to model and measurement errors

and variabilities in the experimental testing and/or environmental/operational conditions, the parameter estimates are expected to vary from dataset to dataset [29,30]. In the Type B testing, the subsystem SS1 is considered as a member of a population of subsystems manufactured to be identical by an assembling process that involves the assemblance of components C1, C2 and C3, where C3 is a member of a population of identically manufactured components. Due to manufacturing/assembled variabilities there are variations in the properties of the components C3 used from the population of components, resulting in variations of the subsystems SS1 that are assembled from the three components C1, C2 and C3. Each dataset $D_k^{(3)}$ is referred to a single test carried out at member k of the population of available subsystems SS1. Tests over a subset of N_3 subsystems in the population are used to infer the values and the uncertainties of the model parameter sets $\theta^{(1)}$, $\theta^{(2)}$ and $\theta^{(3)}$. The uncertainties in the parameter set $\theta^{(3)}$ are due to the variabilities in all members in the population of components C3 (tested and non-tested). The subsystem SS1 in this case is considered to be one of the members in the population of subsystems. To account for the uncertainty due to variability in the datasets for a single subsystem SS1 (testing Type A) or due to the manufacturing/assembled variability of selected subsystem SS1 from the population (testing Type B), uncertainties are embedded in the parameters of the model involved by assigning a parameterized prior probability distribution

$$p(\theta|\psi) = p(\theta^{(1)}, \theta^{(2)}, \theta^{(3)}|\psi_1, \psi_2, \psi_3) = p(\theta^{(1)}|\psi_1)p(\theta^{(2)}|\psi_2)p(\theta^{(3)}|\psi_3) \quad (4.1)$$

with the hyperparameter set $\psi=[\psi_1, \psi_2, \psi_3]$, where independence is assumed for the model parameters of the three different components. The hyperparameters ψ are then inferred from the available multiple datasets at the subsystem SS1 level and the two component C1 and C2 levels.

4.2.1.3 System level

For the system level, it is assumed that there are no experimental data available. The uncertainties of the model parameters $\theta=[\theta^{(1)}, \theta^{(2)}, \theta^{(3)}]$ estimated based on the experimental data $\mathbf{D}=\{\{D_i^{(1)}\}_{i=1}^{N_1}, \{D_j^{(2)}\}_{j=1}^{N_2}, \{D_k^{(3)}\}_{k=1}^{N_3}\}$, are considered to predict the output quantities of interest (QoI) at the system level.

Note that each individual dataset can consist of either response time histories for linear and nonlinear models of components and/or subsystems, or modal properties identified from

the response time histories for linear models of components or subsystems. For the ease of the further use, both types of datasets are denoted by the notation \mathbf{y} . As an example, the i -th dataset in the component C1 is represented as $D_i^{(1)} = \mathbf{y}_i^{(1)}$. Similar notations $\mathbf{y}_j^{(2)}$ and $\mathbf{y}_k^{(3)}$ are used for the datasets of component C2 and subsystem SS1, respectively.

4.2.2 Bayesian learning of model parameters and hyperparameters

The objective in this work is to fuse all available information in the multi-level modeling approach in order to quantify different sources of uncertainties and to propagate such uncertainties to predict output QoI.

4.2.2.1 Prior hypothesis for model parameters

To take into account in the modeling and parameter estimation process the uncertainties of model parameters due to experimental, environmental, operational and manufacturing variabilities, these uncertainties are embedded within the model by introducing a hierarchy in the model parameters. This is accomplished by postulating the probabilistic model in Eq. (4.1) for the parameters that depend on hyperparameters. Herein, a Gaussian prior PDF model is assigned for the model parameters, given as:

$$p(\boldsymbol{\theta}|\boldsymbol{\psi}) = N(\boldsymbol{\theta} | \boldsymbol{\mu}_{\boldsymbol{\psi}}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}}) = N(\boldsymbol{\theta}^{(1)} | \boldsymbol{\mu}_{\boldsymbol{\psi}_1}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}_1}) N(\boldsymbol{\theta}^{(2)} | \boldsymbol{\mu}_{\boldsymbol{\psi}_2}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}_2}) N(\boldsymbol{\theta}^{(3)} | \boldsymbol{\mu}_{\boldsymbol{\psi}_3}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}_3}) \quad (4.2)$$

where $\boldsymbol{\psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \boldsymbol{\psi}_3]$ denotes the hyperparameters of the model parameters, with $\boldsymbol{\psi}_i = [\boldsymbol{\mu}_{\boldsymbol{\psi}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}_i}]$, $i=1,2,3$, involving the hyper mean and hyper covariance matrix of the normal distribution assigned to the prior distribution of the parameter set $\boldsymbol{\theta}^{(i)}$. Realization of $\boldsymbol{\theta}$ varies across the datasets, where $\boldsymbol{\theta}_i$ corresponds to the i -th dataset. Specifically, the hyper parameter $\boldsymbol{\psi}_1$ corresponds to the model parameter $\boldsymbol{\theta}^{(1)}$, denoted with green in Fig. 4-1(b). The parameter $\boldsymbol{\theta}^{(1)}$ and hyperparameter $\boldsymbol{\psi}_1$ are shared at the component C1 and subsystem SS1 levels so they are inferred from the datasets $D_i^{(1)}$, $i=1,2,\dots,N_1$ and $D_k^{(3)}$, $k=1,2,\dots,N_3$ available at the component C1 and the subsystem SS1 levels. Similar definition is described by hyperparameter $\boldsymbol{\psi}_2$, denoted with blue in Fig. 4-1(b). The parameter $\boldsymbol{\theta}^{(2)}$ and hyperparameter $\boldsymbol{\psi}_2$ are shared at the component C2 and subsystem SS1 levels so they are inferred from the datasets $D_j^{(2)}$, $j=1,2,\dots,N_2$ and $D_k^{(3)}$, $k=1,2,\dots,N_3$ available at the component C2 and the subsystem SS1 levels. The hyper parameter $\boldsymbol{\psi}_3$, denoted with orange

in Fig. 4-1(b), is assigned for model parameter $\theta^{(3)}$. They are inferred only from the datasets $D_k^{(3)}$, $k=1,2,\dots,N_3$ available at the subsystem SS1 level. Note that each element ψ_i ($i=1,2,3$) consists of the hyper mean μ_{ψ_i} and the hyper covariance matrix Σ_{ψ_i} to be identified using the multiple data sets collected from either the component levels or subsystem level or both of them. Formulations for estimating the parameters and predicting the responses are presented below.

4.2.2.2 Parameter estimation

Assume that the model outputs from either the components or subsystem levels are linked to their computational models corresponding to a particular value of the model parameters. The discrepancy between the model outputs and the measured data can be modeled based on the prediction error equation, given as follows:

$$\begin{aligned} \text{Component C1: } \mathbf{y}_i^{(1)} &= \mathbf{g}_1(\theta_i^{(1)}) + \boldsymbol{\varepsilon}_i^{(1)}, & i=1,2,\dots,N_1 \\ \text{Component C2: } \mathbf{y}_j^{(2)} &= \mathbf{g}_2(\theta_j^{(2)}) + \boldsymbol{\varepsilon}_j^{(2)}, & j=1,2,\dots,N_2 \\ \text{Sub-system SS1: } \mathbf{y}_k^{(3)} &= \mathbf{g}_3(\theta_k) + \boldsymbol{\varepsilon}_k^{(3)}, & k=1,2,\dots,N_3 \end{aligned} \quad (4.3)$$

where $\mathbf{g}_1(\cdot)$, $\mathbf{g}_2(\cdot)$ and $\mathbf{g}_3(\cdot)$ are the responses from components C1, C2 and subsystem SS1. Note that the model parameters θ_k contain the overall parameters $[\theta_k^{(1)}, \theta_k^{(2)}, \theta_k^{(3)}]$ which is used for the predicted output through the subsystem level. Notations $\boldsymbol{\varepsilon}_i^{(1)}$, $\boldsymbol{\varepsilon}_j^{(2)}$ and $\boldsymbol{\varepsilon}_k^{(3)}$ are respectively the prediction errors for the components C1, C2 and subsystem SS1, and they are modeled as a zero-mean Gaussian variables with different covariance matrices Σ_1 , Σ_2 and Σ_3 . It is assumed herein that $\Sigma_1 = \sigma_1^2 \mathbf{I}$, $\Sigma_2 = \sigma_2^2 \mathbf{I}$ and $\Sigma_3 = \sigma_3^2 \mathbf{I}$, where σ_1^2 , σ_2^2 and σ_3^2 are defined as the prediction error variance to be estimated using the multiple datasets.

Realization of $\theta_i^{(1)}$, $i=1,2,\dots,N_1$, from the Gaussian distribution $N(\theta^{(1)} | \mu_{\psi_1}, \Sigma_{\psi_1})$ can vary across the different data set $D_i^{(1)}$, $i=1,2,\dots,N_1$, available for component C1, where the realization $\theta_i^{(1)}$ is considered to be an independent sample of the distribution $N(\theta^{(1)} | \mu_{\psi_1}, \Sigma_{\psi_1})$ that corresponds to the data set $D_i^{(1)}$. Similar is the case for realization of $\theta_j^{(2)}$, $j=1,2,\dots,N_2$ and θ_k , $k=1,2,\dots,N_3$ from the distributions $N(\theta^{(2)} | \mu_{\psi_2}, \Sigma_{\psi_2})$ and $N(\theta | \mu_{\psi_3}, \Sigma_{\psi_3})$ for component C1 and subsystem SS1, respectively, that can vary across the different data sets $D_j^{(2)}$, $j=1,2,\dots,N_2$ and $D_k^{(3)}$, $k=1,2,\dots,N_3$.

This constitutes a hierarchy of model parameters that has two classes of parameters. The first class of model parameters comprises the experiment-specific model and prediction error parameters $\Theta = \{\{\theta_i^{(1)}\}_{i=1}^{N_1}, \{\theta_j^{(2)}\}_{j=1}^{N_2}, \{\theta_k\}_{k=1}^{N_3}\}$ and $\sigma^2 = \{\sigma_1^2, \sigma_2^2, \sigma_3^2\}$, while the second class model parameters comprises the hyper-parameters involved in $\Psi = [\psi_1, \psi_2, \psi_3]$. The number of parameters in the first set increases linearly with the number of datasets making the parameter estimation problem challenging when the number of data sets increases.

To integrate the available data sets \mathbf{D} and the model information, Bayesian inference is performed to identify the datasets-specific model parameters Θ , the hyper parameters Ψ , as well as the prediction error parameters σ^2 . Based on Bayes' theorem, the posterior distribution of the overall parameters given the available datasets, $p(\Theta, \Psi, \sigma^2 | \mathbf{D})$, is proportional to the product of the likelihood function $p(\mathbf{D} | \Theta, \Psi, \sigma^2)$ and the joint prior distribution $p(\Theta, \Psi, \sigma^2)$, given by:

$$p(\Theta, \Psi, \sigma^2 | \mathbf{D}) \propto p(\mathbf{D} | \Theta, \Psi, \sigma^2) p(\Theta, \Psi, \sigma^2) \quad (4.4)$$

Due to the independence of the prediction error variance σ^2 and the parameters $\{\Theta, \Psi\}$, the prior distribution $p(\Theta, \Psi, \sigma^2)$ can be written as:

$$p(\Theta, \Psi, \sigma^2) = p(\Theta | \Psi) p(\Psi) p(\sigma^2) \quad (4.5)$$

The first factor $p(\Theta | \Psi)$ in the right-hand-side (RHS) of Eq. (4.5) depends on the prior assumption in Eq. (4.2). It is reminded that the realization of the model parameter θ is free to vary from data set to data set, with the realization θ_i corresponding to the data set D_i considered to be an independent sample of its prior distribution $p(\theta | \Psi)$. Therefore, due to the fact that the model parameters from dataset to dataset in the components or subsystem levels are independent, the joint prior distribution takes the form:

$$\begin{aligned} p(\Theta, \Psi, \sigma^2) &\propto p(\Psi) p(\sigma^2) \prod_{i=1}^{N_1} p(\theta_i^{(1)} | \psi_1) \prod_{j=1}^{N_2} p(\theta_j^{(2)} | \psi_2) \prod_{k=1}^{N_3} p(\theta_k | \psi) \\ &\propto p(\Psi) p(\sigma^2) \prod_{i=1}^{N_1} N(\theta_i^{(1)} | \mu_{\psi_1}, \Sigma_{\psi_1}) \prod_{j=1}^{N_2} N(\theta_j^{(2)} | \mu_{\psi_2}, \Sigma_{\psi_2}) \prod_{k=1}^{N_3} N(\theta_k | \mu_{\psi}, \Sigma_{\psi}) \end{aligned} \quad (4.6)$$

where the hyper mean μ_{ψ} and the hyper covariance matrix Σ_{ψ} in the last factor of Eq. (4.6) contain respectively the elements $[\mu_{\psi_1}, \mu_{\psi_2}, \mu_{\psi_3}]$ and the elements $[\Sigma_{\psi_1}, \Sigma_{\psi_2}, \Sigma_{\psi_3}]$.

From the structure of the prediction error equations in Eq. (4.3), the likelihood function $p(\mathbf{D} | \Theta, \Psi, \sigma^2)$ depends on the values of the parameter set Θ and it is independent of the

hyper parameters Ψ . Thus $p(\mathbf{D}|\Theta, \Psi, \sigma^2)$ is determined by the overall data sets \mathbf{D} , the model parameters Θ as well as the prediction error variance σ^2 . Also, owing to the independence of different datasets and the independence of model and prediction error parameters over different datasets, the likelihood function $p(\mathbf{D}|\Theta, \Psi, \sigma^2)$ can be expressed as the product of the individual likelihood function for each dataset as follows

$$p(\mathbf{D}|\Theta, \Psi, \sigma^2) = \prod_{i=1}^{N_1} p(D_i^{(1)} | \theta_i^{(1)}, \sigma_1^2) \prod_{j=1}^{N_2} p(D_j^{(2)} | \theta_j^{(2)}, \sigma_2^2) \prod_{k=1}^{N_3} p(D_k^{(3)} | \theta_k, \sigma_3^2) \quad (4.7)$$

where the three factors in the RHS of Eq. (4.7) correspond to the likelihood function for component C1, C2 and subsystem SS1. The likelihood function corresponding to each dataset in the components or subsystem can be analytically constructed based on the prediction error equation (4.3), given as:

$$\begin{aligned} \text{Component C1: } p(D_i^{(1)} | \theta_i^{(1)}, \sigma_1^2) &\propto |\Sigma_1|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[\mathbf{y}_i^{(1)} - g_1(\theta_i^{(1)})]^T \Sigma_1^{-1} [\mathbf{y}_i^{(1)} - g_1(\theta_i^{(1)})]\right\} \\ \text{Component C2: } p(D_j^{(2)} | \theta_j^{(2)}, \sigma_2^2) &\propto |\Sigma_2|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[\mathbf{y}_j^{(2)} - g_2(\theta_j^{(2)})]^T \Sigma_2^{-1} [\mathbf{y}_j^{(2)} - g_2(\theta_j^{(2)})]\right\} \\ \text{Sub-system SS1: } p(D_k^{(3)} | \theta_k, \sigma_3^2) &\propto |\Sigma_3|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[\mathbf{y}_k^{(3)} - g_3(\theta_k)]^T \Sigma_3^{-1} [\mathbf{y}_k^{(3)} - g_3(\theta_k)]\right\} \end{aligned} \quad (4.8)$$

Substituting (4.8) and (4.6) into (4.4) one derives the posterior distribution of all parameters which can be populated with samples using available sampling algorithms such as Gibbs sampling [15] or transitional Markov Chain Monte Carlo (TMCMC) [23,31,32]. However, the large number of parameters that may be involved in structural model parameter set Θ results in computationally very expensive sampling procedures.

4.2.2.3 Posterior distribution of hyperparameters using Asymptotic Approximations

Two asymptotic approximations for the likelihood functions, developed in [29] and asymptotically valid for large number of data within a dataset, are employed herein to simplify the analysis and derive analytical expressions for the marginal posterior distribution of the hyperparameters Ψ . These approximations leverage the use of Taylor expansion, and approximate each likelihood function involved in Eq. (4.7) to a simplified form of a Gaussian distribution by expanding the negative logarithm of the likelihood function for each dataset with respect to either the model and prediction error variance parameters (approximation 1, A-1 for short) or only the model parameters (approximation 2, A-2 for short) about their maximum likelihood estimate (MLE). Details are presented in [29]. By utilizing

approximation A-1, Eq. (4.7) can be rewritten as the product of several Gaussian distributions as follows:

$$\begin{aligned}
 p(\mathbf{D} | \boldsymbol{\Theta}, \boldsymbol{\Psi}, \boldsymbol{\sigma}^2) = & \prod_{i=1}^{N_1} N(\boldsymbol{\theta}_i^{(1)} | \hat{\boldsymbol{\theta}}_i^{(1)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)}) N(\sigma_1^2 | \hat{\sigma}_{1,i}^2, \hat{\Sigma}_{\sigma_{1,i}^2}^2) \\
 & \prod_{j=1}^{N_2} N(\boldsymbol{\theta}_j^{(2)} | \hat{\boldsymbol{\theta}}_j^{(2)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)}) N(\sigma_2^2 | \hat{\sigma}_{2,j}^2, \hat{\Sigma}_{\sigma_{2,j}^2}^2) \\
 & \prod_{k=1}^{N_3} N(\boldsymbol{\theta}_k | \hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k}) N(\sigma_3^2 | \hat{\sigma}_{3,k}^2, \hat{\Sigma}_{\sigma_{3,k}^2}^2)
 \end{aligned} \quad (4.9)$$

where the MLEs $\hat{\boldsymbol{\theta}}_i^{(1)}$, $\hat{\boldsymbol{\theta}}_j^{(2)}$, $\hat{\boldsymbol{\theta}}_k$, $\hat{\sigma}_{1,i}^2$, $\hat{\sigma}_{2,j}^2$, and $\hat{\sigma}_{3,k}^2$ are computed by minimizing the negative logarithm of the likelihood functions $-\ln p(D_i^{(1)} | \boldsymbol{\theta}_i^{(1)}, \sigma_{1,i}^2)$, $-\ln p(D_j^{(2)} | \boldsymbol{\theta}_j^{(2)}, \sigma_{2,j}^2)$ and $-\ln p(D_k^{(3)} | \boldsymbol{\theta}_k, \sigma_{3,k}^2)$ with respect to the model parameters $\boldsymbol{\theta}_i^{(1)}$, $\boldsymbol{\theta}_j^{(2)}$, $\boldsymbol{\theta}_k$ and the prediction error variance parameters σ_1^2 , σ_2^2 and σ_3^2 , and the covariance matrices $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)}$, $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)}$, $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k}$, $\hat{\Sigma}_{\sigma_{1,i}^2}^2$, $\hat{\Sigma}_{\sigma_{2,j}^2}^2$, and $\hat{\Sigma}_{\sigma_{3,k}^2}^2$, quantifying the parameter identification or estimation uncertainties, are the inverse of the hessian matrices of the negative logarithm of the likelihood functions evaluated at their corresponding MLEs. The parameter uncertainties $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)}$, $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)}$ and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k}$ are identified based on each dataset, and they are referred to as the identification uncertainty.

Likewise, by utilizing approximation A-2 in [29], Eq. (4.7) takes the form:

$$\begin{aligned}
 p(\mathbf{D} | \boldsymbol{\Theta}, \boldsymbol{\Psi}, \boldsymbol{\sigma}^2) = & T(\sigma_1^2) \prod_{i=1}^{N_1} N(\boldsymbol{\theta}_i^{(1)} | \hat{\boldsymbol{\theta}}_i^{(1)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)}(\sigma_1^2)) \\
 & T(\sigma_2^2) \prod_{j=1}^{N_2} N(\boldsymbol{\theta}_j^{(2)} | \hat{\boldsymbol{\theta}}_j^{(2)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)}(\sigma_2^2)) \\
 & T(\sigma_3^2) \prod_{k=1}^{N_3} N(\boldsymbol{\theta}_k | \hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k}(\sigma_3^2))
 \end{aligned} \quad (4.10)$$

For data consisted of modal properties at N_d degrees of freedoms (DOF) with N_m observed modes, the function $T(\sigma_r^2)$ is defined in [29], where $r=1,2,3$. For time histories data measured at N_d degrees of freedoms (DOF) with N_t data points at each DOF, the function $T(\sigma_r^2)$, is defined as:

$$T(\sigma_r^2) = (\sigma_r^2)^{-\frac{[N_t N_d - N_{\boldsymbol{\theta}^{(r)}}] N_r}{2}} \exp\left(-\frac{N_t N_d N_r}{2 \sigma_r^2} J_r\right) \quad (4.11)$$

where $N_{\theta^{(r)}}$ is the number of model parameters corresponding to the r -th model. For example, $N_{\theta^{(1)}}$ denotes the number of model parameters corresponding to component C1. The notation J_r is the average of the measure of fit over all datasets in the r -th model, defined as $J_r = \frac{1}{N_r} \sum_{w=1}^{N_r} J_{r,w}$, where $J_{r,w}$ is the measure of fit between the measurements $\mathbf{y}_w^{(r)}$ and the model predictions $\mathbf{g}_w^{(r)}$ for the w -th dataset in the r -th model, given as:

$$J_{r,w} = \frac{1}{N_t N_d} [\mathbf{y}_w^{(r)} - \mathbf{g}_r(\boldsymbol{\theta}_{r,w})]^T [\mathbf{y}_w^{(r)} - \mathbf{g}_r(\boldsymbol{\theta}_{r,w})] \quad (4.12)$$

It is noted that, for the model of component levels, the model parameters $\boldsymbol{\theta}_{r,w}$ equals to $\boldsymbol{\theta}_w^{(r)}$, while for the subsystem SS1 level, it corresponds to the overall parameters $\boldsymbol{\theta}_w = [\boldsymbol{\theta}_w^{(1)}, \boldsymbol{\theta}_w^{(2)}, \boldsymbol{\theta}_w^{(3)}]$ due to the fact the model of subsystem level consists of not only the model parameters $\boldsymbol{\theta}_w^{(3)}$, but also the parameters $\boldsymbol{\theta}_w^{(1)}$ and $\boldsymbol{\theta}_w^{(2)}$. The covariance matrix $\hat{\Sigma}_{\boldsymbol{\theta}_w}^{(r)}(\sigma_r^2)$ in Eq. (4.10) for modal properties data is defined in [29], while for time histories data, it is defined as:

$$\hat{\Sigma}_{\boldsymbol{\theta}_w}^{(r)}(\sigma_r^2) = \frac{2\sigma_r^2}{N_t N_d} \mathbf{H}_J^{-1}(\hat{\boldsymbol{\theta}}_w^{(r)}) \quad (4.13)$$

where $\mathbf{H}_J(\hat{\boldsymbol{\theta}}_w^{(r)})$ is the hessian matrix of measure of fit J_r evaluated at the corresponding MLE $\hat{\boldsymbol{\theta}}_w^{(r)}$.

Once the joint prior distribution and the likelihood function are obtained, the posterior distribution can be readily computed. For approximation A-1, substituting the expressions of the joint prior distribution in Eq. (4.6) and the likelihood function in Eq. (4.9) into Eq. (4.4) yields:

$$p(\boldsymbol{\Theta}, \boldsymbol{\Psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) \propto p(\boldsymbol{\Psi}) p(\boldsymbol{\sigma}^2) \prod_{i=1}^{N_1} N(\boldsymbol{\theta}_i^{(1)} | \boldsymbol{\mu}_{\boldsymbol{\Psi}_1}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}_1}) N(\boldsymbol{\theta}_i^{(1)} | \hat{\boldsymbol{\theta}}_i^{(1)}, \hat{\Sigma}_{\boldsymbol{\theta}_i}^{(1)}) N(\sigma_1^2 | \hat{\sigma}_{1,i}^2, \hat{\Sigma}_{\sigma_{1,i}^2}^2) \\ \prod_{j=1}^{N_2} N(\boldsymbol{\theta}_j^{(2)} | \boldsymbol{\mu}_{\boldsymbol{\Psi}_2}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}_2}) N(\boldsymbol{\theta}_j^{(2)} | \hat{\boldsymbol{\theta}}_j^{(2)}, \hat{\Sigma}_{\boldsymbol{\theta}_j}^{(2)}) N(\sigma_2^2 | \hat{\sigma}_{2,j}^2, \hat{\Sigma}_{\sigma_{2,j}^2}^2) \\ \prod_{k=1}^{N_3} N(\boldsymbol{\theta}_k | \boldsymbol{\mu}_{\boldsymbol{\Psi}}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}}) N(\boldsymbol{\theta}_k | \hat{\boldsymbol{\theta}}_k, \hat{\Sigma}_{\boldsymbol{\theta}_k}) N(\sigma_3^2 | \hat{\sigma}_{3,k}^2, \hat{\Sigma}_{\sigma_{3,k}^2}^2) \quad (4.14)$$

Note that the model parameter $\boldsymbol{\theta}$ in Eq. (4.14) consists of the model parameters for each dataset, and in effect it can be calibrated alongside the hyper parameter $\boldsymbol{\Psi}$ and prediction error variance parameter $\boldsymbol{\sigma}^2$ using any Markov Chain Monte Carlo (MCMC) algorithm. However, the large number of parameters involved in Eq. (4.14) poses difficulties for the

calculation process of sampling methods. Also, for the purpose of getting analytical expressions of the posterior distributions of hyper parameters, a marginalization procedure is preferred. To this end, the marginal distribution of the hyper and prediction error variance parameters is first calculated. It is achieved by integrating the posterior distribution in Eq. (4.14) over the model parameter space Θ , resulting in

$$\begin{aligned}
 p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) &\propto p(\boldsymbol{\psi})p(\boldsymbol{\sigma}^2) \prod_{i=1}^{N_1} \int_{\boldsymbol{\theta}_i^{(1)}} N(\boldsymbol{\theta}_i^{(1)} | \boldsymbol{\mu}_{\boldsymbol{\psi}_1}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}_1}) N(\boldsymbol{\theta}_i^{(1)} | \hat{\boldsymbol{\theta}}_i^{(1)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)}) d\boldsymbol{\theta}_i^{(1)} N(\sigma_1^2 | \hat{\sigma}_{1,i}^2, \hat{\Sigma}_{\sigma_{1,i}^2}^2) \\
 &\quad \prod_{j=1}^{N_2} \int_{\boldsymbol{\theta}_j^{(2)}} N(\boldsymbol{\theta}_j^{(2)} | \boldsymbol{\mu}_{\boldsymbol{\psi}_2}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}_2}) N(\boldsymbol{\theta}_j^{(2)} | \hat{\boldsymbol{\theta}}_j^{(2)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)}) d\boldsymbol{\theta}_j^{(2)} N(\sigma_2^2 | \hat{\sigma}_{2,j}^2, \hat{\Sigma}_{\sigma_{2,j}^2}^2) \quad (4.15) \\
 &\quad \prod_{k=1}^{N_3} \int_{\boldsymbol{\theta}_k} N(\boldsymbol{\theta}_k | \boldsymbol{\mu}_{\boldsymbol{\psi}}, \boldsymbol{\Sigma}_{\boldsymbol{\psi}}) N(\boldsymbol{\theta}_k | \hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k}) d\boldsymbol{\theta}_k N(\sigma_3^2 | \hat{\sigma}_{3,k}^2, \hat{\Sigma}_{\sigma_{3,k}^2}^2)
 \end{aligned}$$

A large number of multidimensional integrals over the model parameters for each dataset is involved. Using the fact that the integrals of the product of two Gaussian distributions in Eq. (4.15) result in a Gaussian distribution [33], the marginal distribution $p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D})$ can be readily derived in the analytical form:

$$\begin{aligned}
 p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) &\propto p(\boldsymbol{\psi})p(\boldsymbol{\sigma}^2) \prod_{i=1}^{N_1} N(\boldsymbol{\mu}_{\boldsymbol{\psi}_1} | \hat{\boldsymbol{\theta}}_i^{(1)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)} + \boldsymbol{\Sigma}_{\boldsymbol{\psi}_1}) N(\sigma_1^2 | \hat{\sigma}_{1,i}^2, \hat{\Sigma}_{\sigma_{1,i}^2}^2) \\
 &\quad \prod_{j=1}^{N_2} N(\boldsymbol{\mu}_{\boldsymbol{\psi}_2} | \hat{\boldsymbol{\theta}}_j^{(2)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)} + \boldsymbol{\Sigma}_{\boldsymbol{\psi}_2}) N(\sigma_2^2 | \hat{\sigma}_{2,j}^2, \hat{\Sigma}_{\sigma_{2,j}^2}^2) \quad (4.16) \\
 &\quad \prod_{k=1}^{N_3} N(\boldsymbol{\mu}_{\boldsymbol{\psi}} | \hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k} + \boldsymbol{\Sigma}_{\boldsymbol{\psi}}) N(\sigma_3^2 | \hat{\sigma}_{3,k}^2, \hat{\Sigma}_{\sigma_{3,k}^2}^2)
 \end{aligned}$$

It is indicated that, except from the presence of the identification uncertainty, the uncertainties due to variability and the model structure uncertainty are also included in Eq. (4.16), denoted as $\boldsymbol{\Sigma}_{\boldsymbol{\psi}_1}$ and σ_1^2 for component C1, $\boldsymbol{\Sigma}_{\boldsymbol{\psi}_2}$ and σ_2^2 for component C2, as well as $\boldsymbol{\Sigma}_{\boldsymbol{\psi}}$ and σ_3^2 for the subsystem SS1.

Similarly, following a similar procedure and substituting the expressions of the joint prior distribution in Eq. (4.6) and the likelihood function in Eq. (4.10) into Eq. (4.4), one can finally derive the marginal distribution $p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D})$ based on the second approximation A-2 as follows

$$\begin{aligned}
p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) &\propto p(\boldsymbol{\psi}) p(\boldsymbol{\sigma}^2) T(\sigma_1^2) \prod_{i=1}^{N_1} N(\boldsymbol{\mu}_{\boldsymbol{\psi}_1} | \hat{\boldsymbol{\theta}}_i^{(1)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_i}^{(1)}(\sigma_1^2) + \boldsymbol{\Sigma}_{\boldsymbol{\psi}_1}) \\
&\quad T(\sigma_2^2) \prod_{j=1}^{N_2} N(\boldsymbol{\mu}_{\boldsymbol{\psi}_2} | \hat{\boldsymbol{\theta}}_j^{(2)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_j}^{(2)}(\sigma_2^2) + \boldsymbol{\Sigma}_{\boldsymbol{\psi}_2}) \\
&\quad T(\sigma_3^2) \prod_{k=1}^{N_3} N(\boldsymbol{\mu}_{\boldsymbol{\psi}} | \hat{\boldsymbol{\theta}}_k, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_k}(\sigma_3^2) + \boldsymbol{\Sigma}_{\boldsymbol{\psi}})
\end{aligned} \tag{4.17}$$

According to Eq. (4.16) or (4.17), the samples of the hyper and prediction error variance parameters can be generated through any Markov Chain Monte Carlo (MCMC) algorithm such as the transitional MCMC (TMCMC) [31,32] or the nested sampling algorithm [34]. It should be pointed out that the number of parameters involved in the marginal posterior distributions (4.16) or (4.17) is substantially less than the number of parameters involved in the original posterior distribution (4.6), which makes the sampling approach for the distributions (4.16) or (4.17) computationally very efficient.

4.2.2.4 Most probable values (MPV) of hyperparameters

The MPVs of the hyper parameters and prediction error variance parameters can be estimated by minimizing the negative logarithm of the joint distribution $p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D})$ in Eqs. (4.16) and (4.17). The analytical solutions are derived by selecting the identification uncertainty to be the same for each dataset and equal to the average of the identification uncertainty over all datasets. Details are given in Appendix A. For approximation A-1, the results are given as:

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{\psi}_1} = \frac{1}{N_1 + N_3} \left(\sum_{i=1}^{N_1} \hat{\boldsymbol{\theta}}_i^{(1)} + \sum_{k=1}^{N_3} \hat{\boldsymbol{\theta}}_k^{(1)} \right); \quad \hat{\boldsymbol{\mu}}_{\boldsymbol{\psi}_2} = \frac{1}{N_2 + N_3} \left(\sum_{j=1}^{N_2} \hat{\boldsymbol{\theta}}_j^{(2)} + \sum_{k=1}^{N_3} \hat{\boldsymbol{\theta}}_k^{(2)} \right); \quad \hat{\boldsymbol{\mu}}_{\boldsymbol{\psi}_3} = \frac{1}{N_3} \sum_{k=1}^{N_3} \hat{\boldsymbol{\theta}}_k^{(3)} \tag{4.18}$$

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\psi}_1} = \boldsymbol{\Sigma}_v^{(1)} - \boldsymbol{\Sigma}^{(1)}; \quad \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\psi}_2} = \boldsymbol{\Sigma}_v^{(2)} - \boldsymbol{\Sigma}^{(2)}; \quad \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\psi}_3} = \boldsymbol{\Sigma}_v^{(3)} - \boldsymbol{\Sigma}^{(3)} \tag{4.19}$$

$$\hat{\sigma}_1^2 = \sum_{i=1}^{N_1} \frac{1}{\hat{\sigma}_{1,i}^2} / \sum_{i=1}^{N_1} \frac{1}{\hat{\sigma}_{1,i}^4}; \quad \hat{\sigma}_2^2 = \sum_{j=1}^{N_2} \frac{1}{\hat{\sigma}_{2,j}^2} / \sum_{j=1}^{N_2} \frac{1}{\hat{\sigma}_{2,j}^4}; \quad \hat{\sigma}_3^2 = \sum_{k=1}^{N_3} \frac{1}{\hat{\sigma}_{3,k}^2} / \sum_{k=1}^{N_3} \frac{1}{\hat{\sigma}_{3,k}^4} \tag{4.20}$$

where

$$\begin{aligned}\Sigma_v^{(1)} &= \frac{1}{N_1 + N_3} \left(\sum_{i=1}^{N_1} (\hat{\mu}_{\psi_1} - \hat{\theta}_i^{(1)}) (\hat{\mu}_{\psi_1} - \hat{\theta}_i^{(1)})^T + \sum_{k=1}^{N_3} (\hat{\mu}_{\psi_1} - \hat{\theta}_k^{(1)}) (\hat{\mu}_{\psi_1} - \hat{\theta}_k^{(1)})^T \right); \\ \Sigma^{(1)} &= \frac{1}{N_1 + N_3} \left(\sum_{i=1}^{N_1} \hat{\Sigma}_{\theta_i}^{(1)} + \sum_{k=1}^{N_3} \hat{\Sigma}_{\theta_k}^{(1)} \right)\end{aligned}\quad (4.21)$$

$$\begin{aligned}\Sigma_v^{(2)} &= \frac{1}{N_2 + N_3} \left(\sum_{j=1}^{N_2} (\hat{\mu}_{\psi_2} - \hat{\theta}_j^{(2)}) (\hat{\mu}_{\psi_2} - \hat{\theta}_j^{(2)})^T + \sum_{k=1}^{N_3} (\hat{\mu}_{\psi_2} - \hat{\theta}_k^{(2)}) (\hat{\mu}_{\psi_2} - \hat{\theta}_k^{(2)})^T \right); \\ \Sigma^{(2)} &= \frac{1}{N_2 + N_3} \left(\sum_{j=1}^{N_2} \hat{\Sigma}_{\theta_j}^{(2)} + \sum_{k=1}^{N_3} \hat{\Sigma}_{\theta_k}^{(2)} \right)\end{aligned}\quad (4.22)$$

$$\Sigma_v^{(3)} = \frac{1}{N_3} \sum_{k=1}^{N_3} (\hat{\mu}_{\psi_3} - \hat{\theta}_k^{(3)}) (\hat{\mu}_{\psi_3} - \hat{\theta}_k^{(3)})^T; \quad \Sigma^{(3)} = \frac{1}{N_3} \sum_{k=1}^{N_3} \hat{\Sigma}_{\theta_k}^{(3)} \quad (4.23)$$

For approximation A-2, the MPVs of all the parameters are given as follows:

$$\hat{\mu}_{\psi_1} = \frac{1}{N_1 + N_3} \left(\sum_{i=1}^{N_1} \hat{\theta}_i^{(1)} + \sum_{k=1}^{N_3} \hat{\theta}_k^{(1)} \right); \quad \hat{\mu}_{\psi_2} = \frac{1}{N_2 + N_3} \left(\sum_{j=1}^{N_2} \hat{\theta}_j^{(2)} + \sum_{k=1}^{N_3} \hat{\theta}_k^{(2)} \right); \quad \hat{\mu}_{\psi_3} = \frac{1}{N_3} \sum_{k=1}^{N_3} \hat{\theta}_k^{(3)} \quad (4.24)$$

$$\hat{\Sigma}_{\psi_1} = \Sigma_v^{(1)} - \Sigma^{(1)}(\hat{\sigma}_1^2); \quad \hat{\Sigma}_{\psi_2} = \Sigma_v^{(2)} - \Sigma^{(2)}(\hat{\sigma}_2^2); \quad \hat{\Sigma}_{\psi_3} = \Sigma_v^{(3)} - \Sigma^{(3)}(\hat{\sigma}_3^2) \quad (4.25)$$

$$\begin{aligned}\hat{\sigma}_1^2 &= \frac{N_m(N_d + 1)}{N_m(N_d + 1) - N_\theta} J_1 = \frac{N_t N_d}{N_t N_d - N_\theta} J_1 \\ \hat{\sigma}_2^2 &= \frac{N_m(N_d + 1)}{N_m(N_d + 1) - N_\theta} J_3 = \frac{N_t N_d}{N_t N_d - N_\theta} J_3 \\ \hat{\sigma}_3^2 &= \frac{N_m(N_d + 1)}{N_m(N_d + 1) - N_\theta} J_3 = \frac{N_t N_d}{N_t N_d - N_\theta} J_3\end{aligned}\quad (4.26)$$

where

$$\begin{aligned}\Sigma^{(1)}(\hat{\sigma}_1^2) &= \frac{1}{N_1 + N_3} \left(\sum_{i=1}^{N_1} \hat{\Sigma}_{\theta_i}^{(1)}(\hat{\sigma}_1^2) + \sum_{k=1}^{N_3} \hat{\Sigma}_{\theta_k}^{(1)}(\hat{\sigma}_3^2) \right) \\ \Sigma^{(2)}(\hat{\sigma}_2^2) &= \frac{1}{N_2 + N_3} \left(\sum_{j=1}^{N_2} \hat{\Sigma}_{\theta_j}^{(2)}(\hat{\sigma}_2^2) + \sum_{k=1}^{N_3} \hat{\Sigma}_{\theta_k}^{(2)}(\hat{\sigma}_3^2) \right) \\ \Sigma^{(3)}(\hat{\sigma}_3^2) &= \frac{1}{N_3} \sum_{k=1}^{N_3} \hat{\Sigma}_{\theta_k}^{(3)}(\hat{\sigma}_3^2)\end{aligned}\quad (4.27)$$

and $J_r = \frac{1}{N_r} \sum_{w=1}^{N_r} J_{r,w}$, $r = 1, 2, 3$, is the average of measure of fit given in Eq. (4.12) for time

history measurements.

The analytical but approximate expressions derived for the MPV of the hyperparameters provide valuable insight into the values expected for the hyper means and hyper covariance matrices. For example, the parameter set $\theta^{(1)}$ is shared at the component C1 and the

subsystem SS1 levels. Thus datasets from component C1 and subsystem SS1 provide information for estimating the hyper parameter set $\boldsymbol{\Psi}_1 = [\boldsymbol{\mu}_{\Psi_1}, \boldsymbol{\Sigma}_{\Psi_1}]$. Specifically, the first of Eq.(4.18) or (4.24) indicates that the hyper mean $\hat{\boldsymbol{\mu}}_{\Psi_1}$ of the parameter set $\boldsymbol{\theta}^{(1)}$ corresponding to the component C1 is the average value of the most probable estimates $\hat{\boldsymbol{\theta}}_i^{(1)}$, $i = 1, \dots, N_1$ and $\hat{\boldsymbol{\theta}}_k^{(1)}$, $k = 1, \dots, N_3$, computed for the datasets available at the component C1 and the subsystem SS1. The most probable value $\hat{\boldsymbol{\Sigma}}_{\Psi_1}$ of the hyper covariance is built from the difference between two terms as indicated in the first equation of Eq. (4.19) or Eq. (4.25). The first term $\boldsymbol{\Sigma}_v^{(1)}$ given in the first of Eq. (4.21) is the covariance due to the variability of the most probable estimates $\hat{\boldsymbol{\theta}}_i^{(1)}$, $i = 1, \dots, N_1$ and $\hat{\boldsymbol{\theta}}_k^{(1)}$, $k = 1, \dots, N_3$ of the parameters set computed using the datasets from component C1 and subsystem SS1, while the second term $\boldsymbol{\Sigma}^{(1)}$ in the second of Eq. (4.21) or Eq. (4.27) is the average of the identification uncertainty $\hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)}$, $i = 1, \dots, N_1$, and $\hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)}$, $k = 1, \dots, N_3$, computed for each dataset available at the component C1 and subsystem SS1. Similar is the interpretation for the MPV of hyper mean $\hat{\boldsymbol{\mu}}_{\Psi_2}$ and hyper covariance $\hat{\boldsymbol{\Sigma}}_{\Psi_2}$ of the second component C2, computed only from the datasets available for component C2 and subsystem SS1. For the MPV of the hyper parameters $\hat{\boldsymbol{\mu}}_{\Psi_3}$ and $\hat{\boldsymbol{\Sigma}}_{\Psi_3}$ similar expressions are obtained which involve only the datasets at subsystem SS1 level. It is reminded that the identification uncertainty reduces as the number of data in a data set increases [29]. This is evident also in (4.13), where the parameter identification uncertainty for each data set is inversely proportional to the number $N_i N_d$ of data points in each dataset. Therefore, given a sufficient number of data points, the identification uncertainties become negligible and thus the MPVs of the hyper covariance equal to the uncertainty due to variability arising from the multiple datasets. This is consistent with a frequentist point of view where the hyper covariance is defined as a variation of the parameters computed from one dataset to another dataset.

4.2.2.5 Posterior distribution of model parameters

Once the samples of the hyper parameters and the prediction error variance parameters are available, the posterior predictive distribution of new model parameters can be calculated through the following procedure:

$$p(\boldsymbol{\theta}_{new} | \mathbf{D}) = \int \int_{\boldsymbol{\Psi} \boldsymbol{\sigma}^2} p(\boldsymbol{\theta}_{new} | \boldsymbol{\Psi}) p(\boldsymbol{\Psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) d\boldsymbol{\Psi} d\boldsymbol{\sigma}^2 \quad (4.28)$$

where $\boldsymbol{\theta}_{new}$ defines the new model parameters whose samples will be used to predict the output QoI. The multi-dimensional integral in Eq. (4.28) can be estimated based on the samples obtained from Eq. (4.16) or Eq. (4.17), given as:

$$\begin{aligned} p(\boldsymbol{\theta}_{new} | \mathbf{D}) &\approx \frac{1}{N} \sum_{l=1}^N p(\boldsymbol{\theta}_{new} | \boldsymbol{\Psi}^{(l)}) = \frac{1}{N} \sum_{l=1}^N N(\boldsymbol{\theta}_{new} | \boldsymbol{\mu}_{\boldsymbol{\Psi}}^{(l)}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}}^{(l)}) \\ &= \frac{1}{N} \sum_{l=1}^N N(\boldsymbol{\theta}_{new}^{(1)} | \boldsymbol{\mu}_{\boldsymbol{\Psi}_1}^{(l)}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}_1}^{(l)}) N(\boldsymbol{\theta}_{new}^{(2)} | \boldsymbol{\mu}_{\boldsymbol{\Psi}_2}^{(l)}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}_2}^{(l)}) N(\boldsymbol{\theta}_{new}^{(3)} | \boldsymbol{\mu}_{\boldsymbol{\Psi}_3}^{(l)}, \boldsymbol{\Sigma}_{\boldsymbol{\Psi}_3}^{(l)}) \end{aligned} \quad (4.29)$$

where each sample $\boldsymbol{\Psi}^{(l)}$ consists of the samples of the hyper means $\boldsymbol{\mu}_{\boldsymbol{\Psi}_1}^{(l)}$, $\boldsymbol{\mu}_{\boldsymbol{\Psi}_2}^{(l)}$ and $\boldsymbol{\mu}_{\boldsymbol{\Psi}_3}^{(l)}$, and also the samples of the hyper covariance matrices $\boldsymbol{\Sigma}_{\boldsymbol{\Psi}_1}^{(l)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\Psi}_2}^{(l)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\Psi}_3}^{(l)}$. N samples in total are used to estimate the probability density function $p(\boldsymbol{\theta}_{new} | \mathbf{D})$. Note also that the uncertainty in the model parameter set $\boldsymbol{\theta}$ can be described by sampling the distribution in Eq. (4.29).

4.2.2.6 Response predictions

Fig. 4-2 shows the framework for propagating the uncertainties to the response predictions. It is noted that both the uncertainties from the model parameters and the prediction error parameters are considered for the observed output QoI, while for the unobserved output QoI only the uncertainty from the model parameters can be considered because of the absence of information for the prediction error term. The probability distribution of the unobserved QoI \mathbf{y}^{Unob} can be estimated by evaluating the conditional distribution for a new output of the model given the observations, as follows:

$$p(\mathbf{y}^{Unob} | \mathbf{D}) = \int_{\boldsymbol{\theta}_{new}} p(\mathbf{y}^{Unob} | \boldsymbol{\theta}_{new}) p(\boldsymbol{\theta}_{new} | \mathbf{D}) d\boldsymbol{\theta}_{new} \approx \frac{1}{N_u} \sum_{i=1}^{N_u} p(\mathbf{y}^{Unob} | \boldsymbol{\theta}_{new}^{(i)}) \quad (4.30)$$

where $\boldsymbol{\theta}_{new}^{(i)}$ is the i -th sample obtained from the probability distribution $p(\boldsymbol{\theta}_{new} | \mathbf{D})$ in Eq. (4.29), N_u is the number of samples. Similarly, for the observed QoI, by considering the overall uncertainties, the probability distribution $p(\mathbf{y}^{ob} | \mathbf{D})$ of the observed QoI \mathbf{y}^{ob} can be readily estimated in the form:

$$p(\mathbf{y}^{ob} | \mathbf{D}) \approx \frac{1}{N_o} \sum_{i=1}^{N_o} p(\mathbf{y}^{ob} | \boldsymbol{\theta}_{new}^{(i)}, \boldsymbol{\sigma}^{2,(i)}) \quad (4.31)$$

where the number of the samples N_o can be either the same or different with the number of samples N_u .

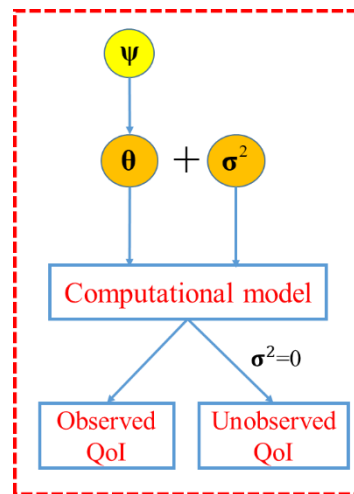


Fig. 4-2 Uncertainty propagation for response predictions

Some points for the prediction in a multi-level modeling approach using multi-level datasets are outlined as follows:

- For the predictions at component levels C1 and C2, only the uncertainty of the model parameters $\theta^{(1)}$ or $\theta^{(2)}$ is considered for the unobserved QoI in component C1 or component C2. However, due to the fact that both parameters are shared in the subsystem level SS1, they are also inferred from the datasets available in the subsystem level.
- For the predictions at subsystem level, all three sets of model parameters are involved for predicting the observed and unobserved QoI, with the parameters informed from the datasets available at the components C1, C2 and subsystem SS1 levels.
- For the system level, since there is no dataset available, the prediction error variance parameter may not be calibrated, and all the QoI in this level will only take into account the uncertainty from the model parameters. The uncertainties in the parameters are informed from the dataset available at component C1, C2 and subsystem SS1 levels.

4.2.2.7 Algorithms

Algorithm 1 shows the steps for estimating the hyper and prediction error parameters using the proposed hierarchical Bayesian modeling framework, and Algorithm 2 summarizes the

process of propagating uncertainties for predicting response QoI. In Algorithm 1, it is reminded that a full sampling (FS) procedure [23,31,32] could be used as an alternative to obtain the samples of the hyper and prediction error parameters. This can be accomplished by drawing the samples of structural model and prediction error parameters from the likelihood function for each dataset available at components C1, C2 and subsystem SS1, and the available samples are subsequently utilized for computing the posterior distributions of the hyper parameters and prediction error parameters.

Algorithm 1: Estimations of hyper and prediction error parameters

Step 1: Components and sub-system model runs are required

For A-1:

- (1.1) Minimize the negative logarithm of the likelihood functions $p(D_i^{(1)} | \theta_i^{(1)}, \sigma_1^2)$, $p(D_j^{(2)} | \theta_j^{(2)}, \sigma_2^2)$ and $p(D_k^{(3)} | \theta_k, \sigma_3^2)$ given in Eq. (4.8) with respect to model and prediction error parameters to compute the MLEs $\hat{\theta}_i^{(1)}$, $\hat{\theta}_j^{(2)}$, $\hat{\theta}_k$, $\hat{\sigma}_{1,i}^2$, $\hat{\sigma}_{2,j}^2$ and $\hat{\sigma}_{3,k}^2$ for $i=1,2,\dots,N_1$, $j=1,2,\dots,N_2$, $k=1,2,\dots,N_3$
- (1.2) Evaluate the hessian matrix of the negative logarithm of the likelihood functions given in Eq. (4.8) and calculate the covariance matrices $\hat{\Sigma}_{\theta_i}^{(1)}$, $\hat{\Sigma}_{\theta_j}^{(2)}$ and $\hat{\Sigma}_{\theta_k}$ as the inverse of the hessian matrices evaluated at the obtained MLEs for $i=1,2,\dots,N_1$, $j=1,2,\dots,N_2$, $k=1,2,\dots,N_3$

For A-2:

- (1.1) Minimize the negative logarithm of the likelihood functions given in Eq. (4.8) to with respect to model parameters to compute the MLEs $\hat{\theta}_i^{(1)}$, $\hat{\theta}_j^{(2)}$, $\hat{\theta}_k$, where $i=1,2,\dots,N_1$, $j=1,2,\dots,N_2$, $k=1,2,\dots,N_3$
- (1.2) Evaluate the hessian matrix $\mathbf{H}_J(\hat{\theta}_w^{(r)})$ of the measure of fit $J_r = \frac{1}{N_r} \sum_{w=1}^{N_r} J_{r,w}$ at the MLEs $\hat{\theta}_i^{(1)}$, $\hat{\theta}_j^{(2)}$, $\hat{\theta}_k$, with $J_{r,w}$ given in (4.12), and obtain $\hat{\Sigma}_{\theta_w}^{(r)}(\sigma_r^2)$ from (4.13) for $r=1,2,3$

Step 2: No model runs are required

For A-1:

- (1.3) Represent the posterior distribution of hyper and prediction error parameters $p(\Psi, \sigma^2 | \mathbf{D})$ defined in Eq. (4.16) using samples $\Psi^{(l)}$ and $\sigma^{2,(l)}$ generated from a sampling approach like TMCMC or nested sampling

For A-2:

- (1.3) Represent the posterior distribution of hyper and prediction error parameters $p(\Psi, \sigma^2 | \mathbf{D})$ defined in Eq. (4.17) using samples $\Psi^{(l)}$ and $\sigma^{2,(l)}$ generated from a sampling approach like TMCMC or nested sampling
-

Algorithm 2: Uncertainty propagation for response predictions

(2.1) Draw samples $\boldsymbol{\theta}_{new}^{(q)}$, $l = 1, \dots, N$, from $p(\boldsymbol{\theta}_{new}^{(q)} | \mathbf{D})$ in Eq. (4.29)

For unobserved QoI:

(2.2) Draw samples from the posterior distribution of unobserved QoI using (4.30), and use the samples to estimate the statistical properties, such as the mean, standard deviation and quantiles

For observed QoI:

(2.3) Draw samples from the posterior distribution of unobserved QoI using (4.31), and use the samples to estimate the statistical properties, such as the mean, standard deviation and quantiles

4.3 Numerical Example

This section illustrates the proposed method using a numerical case study. A six degree of freedom (DOF) spring-mass chain model of a mechanical system is presented in Section 4.3.1, where the model with its decompositions in subsystems and components and the available observations are described in detail. The parameter estimation process is conducted based on the available datasets in the different levels of model hierarchy, and the response QoI are predicted according to the uncertainties obtained from the parameter estimation process. Results are discussed in detail in Section 4.3.2.

4.3.1 Problem description

As shown in Fig. 4-3, the six DOFs model is represented as the system in this study. The system is decomposed into one subsystem and three components. Component C1, denoted with green, consists of the first two links of the model, component C2, denoted with blue, contains the third and the fourth links, and component C3, denoted by orange, contains the fifth link. The sub-system SS1 in Fig. 4-3 consists of the components C1, C2 and C3. The system S is assembled from the subsystem SS1 and a second subsystem that consist of the sixth link with known mechanical properties. Parameterized computational models for the system, subsystem and components are introduced to make simulation-based predictions of structural responses. The information for the computational model and the parameterization is given in Table 4-1. As shown, the model is fully parameterized using three stiffness-related model parameters $\boldsymbol{\theta} = [\theta^{(1)}, \theta^{(2)}, \theta^{(3)}]^T$. Specifically, the first parameter $\theta^{(1)}$ is linked to the stiffness of the first two springs of the component C1. The parameter $\theta^{(1)}$ is shared between

component C1 and subsystem SS1. The second parameter $\theta^{(2)}$ is linked to the stiffness of the third and the fourth spring, shared between component C2 and subsystem SS1. The last parameter $\theta^{(3)}$ corresponds only to the stiffness of the fifth spring, shared between the component C3 and subsystem SS1. Each parameter is multiplied by the corresponding nominal stiffness values reported in Table 4-1, representing the model stiffness which is used for the process of updating the models.

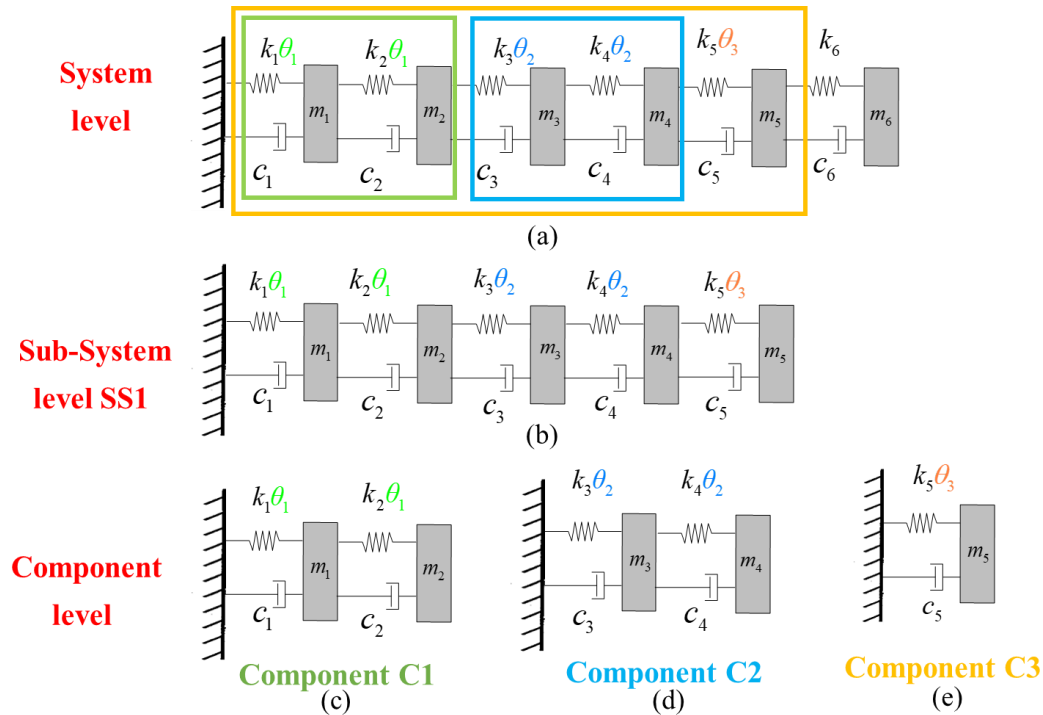


Fig. 4-3 (a) Six-DOF model of a mechanical system, (b) test configuration of subsystem SS1, (c,d) test configurations of the components C1 and C2, (e) component C3

Table 4-1 Information for computational models

Link i	Mass m_i (Kg)	Stiffness k_i (N/m)	Damping ratio ζ_i	Existed Level	Parameters for stiffness
1	1	1800	0.02	C1/SS1/S	$\theta^{(1)}$
2	1	1800	0.02	C1/SS1/S	$\theta^{(1)}$
3	1	1500	0.02	C2/SS1/S	$\theta^{(2)}$
4	1	1500	0.02	C2/SS1/S	$\theta^{(2)}$
5	1	1800	0.02	SS1/S	$\theta^{(3)}$
6	1	1800	0.02	S	-

Multiple datasets are recorded for the components C1 and C2 and the subsystem SS1. The configurations of the components C1 and C2 and the subsystem SS1 under which the experiments are performed are shown in Fig. 4-3(b-d). Different types of tests are carried out for components C1, C2 and subsystem SS1, reporting different sources of the uncertainty due to variability. In lieu of experimental data, datasets are simulated as follows.

For component C1 a type A testing is performed based on the testing configuration in Fig. 4-3c. Responses are simulated under the condition where model error is present. For this, the first mass is perturbed by 5% from the nominal mass value reported in Table 4-1 and the resulting perturbed model of component C1 is used to produce the responses that will be used as datasets. $N_1=10$ datasets of acceleration time histories are simulated subjected to 10 independent base excitations. The excitation herein is assumed as a Gaussian sequence with mean zero and standard deviation equal to one. Sensors for recording the data are assumed to locate at both the first and second link, and therefore 2 sets of acceleration time histories exist in a dataset. The sampling rate for each set is taken as 0.01s corresponding to a sampling frequency 100Hz for total of 5.2 seconds. The test-to-test variability in component C1 is attributed to the presence of model error. Different loading conditions subjected to the same model result in different responses, and will further lead to different identification values of the model parameters from dataset to dataset.

For component C2 a Type A testing is also performed. Data sets consisted of modal properties are generated under the assumption of zero model error. $N_2=20$ datasets in total are simulated based on a Gaussian distribution for $\theta^{(2)}$ with mean 1 and standard deviation 0.05. Specifically, to generate a dataset D_j , $j=1, \dots, N_2$, a realization $\theta_j^{(2)}$ is drawn from the distribution $N(\theta^{(2)}|1, 0.05^2)$ and the modal properties can then be produced through the computational model of component C2. Due to the fact that the realizations of $\theta^{(2)}$ from dataset to dataset are statistically independent, different datasets of the modal properties are thus statistically independent. For component C2, the test-to-test variability can be due to the alterations of environmental conditions and material properties. For example, temperature or humidity changes make the material properties different, and it can further affect the modal features. Such changes are represented as the variations of the model parameters, following a Gaussian distribution with a significant variance.

For subsystem SS1 level a type B testing is performed. The subsystem is assembled from the fixed component C1, the fixed component C2, and a component C3 that is a member of a

population of identically manufactured components (links). The subsystems generated through this process form a population of identically assembled subsystems, although variability exists due to manufacturing errors caused by the manufacturing variabilities arising from the component C3. Since the subsystem SS1 contains the component C1, it is apparent that the model error also exists in any one of the assembled subsystems. $N_3=20$ subsystems are assembled through this process and then a single test is performed for representative members in the population of subsystems SS1. Acceleration datasets are collected from the test carried out for each subsystem. Datasets $D_k, k=1, \dots, N_3, N_3=20$, are generated based on samples of the model parameters $[\theta^{(2)}, \theta^{(3)}]^T$ generated from a Gaussian distribution with mean $[1 \ 1]^T$ and a diagonal covariance matrix with the same diagonal elements 0.05^2 . For each acceleration dataset, 3 sensors are placed at the first, third and fifth links, and thus 3 sets of the time histories data are included in a dataset. Each set consists of $N_t=520$ data points in the increment of 0.01s, the same with the one in component C1.

The process to assemble the system is to connect the subsystem SS1 (with the component C3 chosen from the population of identically manufactured components C2) with a fixed link (the 6th link in Figure 3) representing the second subsystem to form the final system S. The uncertainty at the system level arises due to the variability from the manufacturing process of components C3 used to assemble the subsystem SS1, due to the variability from the different environmental conditions under which the component C2 was tested, and due to the model error present in component C1. For system level, no dataset is available.

Results and discussions for identifying and propagating uncertainties are presented in the next subsection. Although the components, subsystems and systems are simple spring mass chain models, the model is representative of more involved models of industrial components that are assembled to form the subsystems and eventually the systems. The differences are in the complexity of the models for the components, the number of parameters per component and the higher computational cost that would be required to perform computations. A simple model was purposely chosen to represent the features for the proposed method and provide valuable insight into uncertainty quantification and propagation process.

4.3.2 Results and discussions

4.3.2.1 Parameter estimation

The foregoing asymptotic approximation A-1 for the likelihood function in Eq. (4.9) is utilized to compute the MLE and the identification uncertainty in each dataset. It is noted that

the first two parameters $\theta^{(1)}$ and $\theta^{(2)}$ are included in both the components and subsystem levels. The MLEs of these two parameters $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$ along with their identification uncertainties $\hat{\sigma}_{\theta^{(1)}}$ and $\hat{\sigma}_{\theta^{(2)}}$ can be calculated according to each dataset from either the components or subsystem levels. Results are shown in Fig. 4-4. The blue dots and error bars indicate the MLEs and the identification uncertainties from the components level, while the red ones display the results from the subsystem SS1 level. It is seen that the MLEs of the first two parameters for a specific dataset computed from the components level differ from those from the subsystem SS1 level. Also, the MLEs of the model parameters from either the components or subsystem levels significantly fluctuate across the datasets. For the first parameter $\theta^{(1)}$, the identified values from the different datasets vary considerably due to the presence of model error. Similar variations are observed for the second and third parameter $\theta^{(2)}$ and $\theta^{(3)}$ which are due to the changes of environmental conditions and manufacturing process. It is also seen that the identification uncertainties keep rather constant and extremely small values for each parameter. It is reminded that for the first and third parameters, such small uncertainties are produced by the sufficient number of data points available for each dataset due to the fact that experimental data consist of response time histories. For the second parameter, the small uncertainties are attributed to the assumptions of zero model error in component C2. It should be observed that such small identification uncertainties cannot capture the variations of the MLEs, failing to represent the uncertainty of model parameters due to variability.

For describing the uncertainty due to variability, the distributions of the hyper parameters can be computed based on the approximation A-1 in Eq. (4.15) and the approximation A-2 in Eq. (4.17). For exploring the effect of the datasets at different levels of hierarchy on the posterior distribution of the hyper parameters, results are presented for three cases where the datasets are considered from components only, subsystem only and both of them. Results based on the approximation A-1 are shown in Fig. 4-5. Observing the hyper mean values, it is noted that the hyper mean values inferred using datasets from C1 only and the subsystem SS1 only are shifted from the nominal values, while the hyper mean value inferred using datasets from C2 is fairly close to the nominal values. This is due to the effect of the model error present in C1 and SS1, so that no model parameter values provide predictions that exactly match the measurements. However for the case of component C2, the perfect model assumed can yield the exact nominal values when given a sufficient large number of datasets.

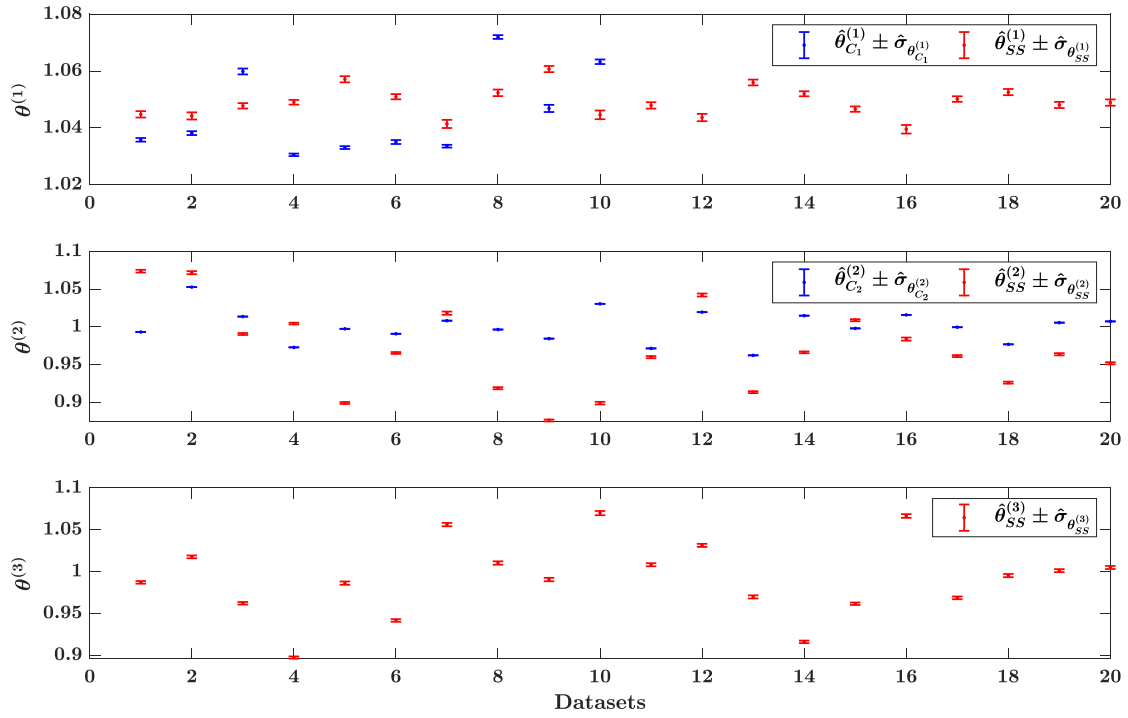


Fig. 4-4 The MLEs and the identification uncertainties of model parameters computed based on the individual dataset from components and sub-system

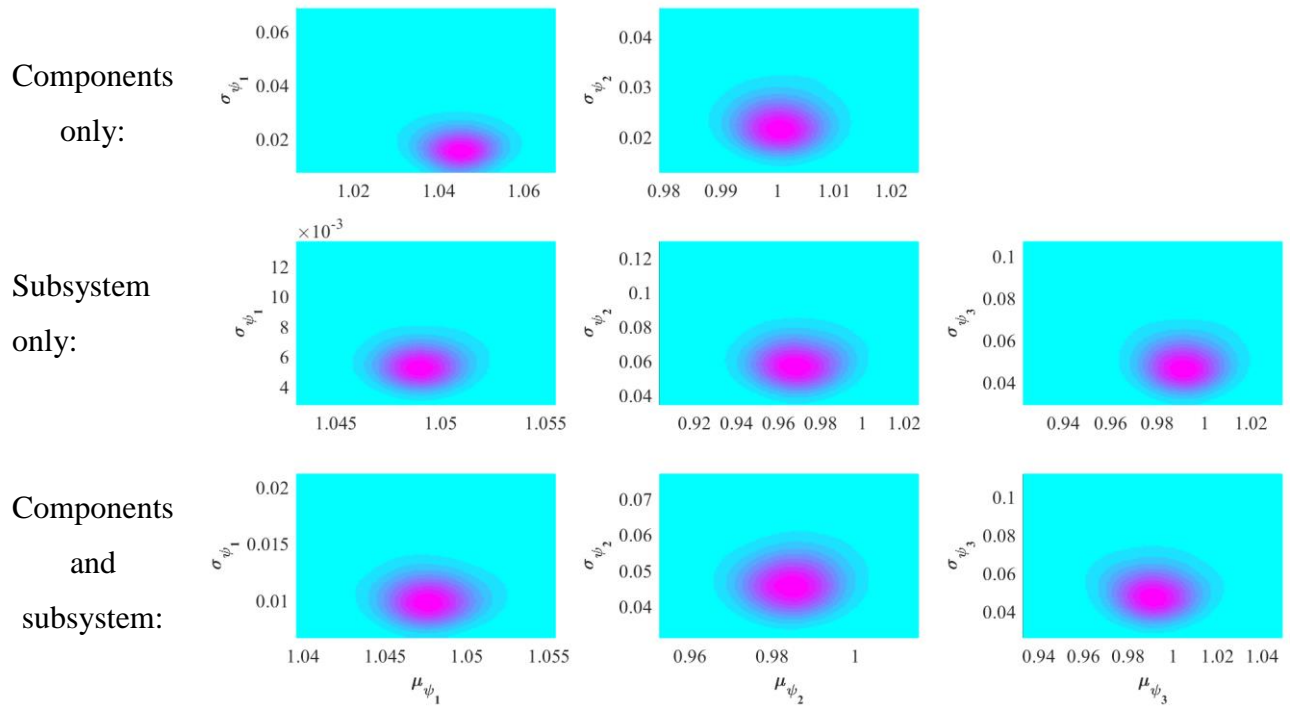


Fig. 4-5 Posterior distributions of the hyper parameters considering the datasets from components only (top), sub-system only (middle) and both the components and sub-system (bottom)

It is also seen that the MPV of the hyper mean μ_{ψ_1} from both components and subsystem is around 1.047 which corresponds approximately to an average value of the MPV 1.041 and 1.049 computed respectively from only the component C1 and only the subsystem SS1. For the hyper standard deviations (square root of the diagonal entry in hyper covariance), it is seen that a relatively larger value is computed (around 1.5%) for the first hyper parameter σ_{ψ_1} in C1 while a smaller one (0.5%) is obtained in SS1, and the hyper standard deviations from both the components and subsystem SS1 is around 1% by considering the overall information. Similar remarks can be concluded for other hyper parameters, which indicates that the overall hyper parameter values consider the information from both the components and subsystem levels. This scenario is consistent with the analytical expression for the hyper means in Eq. (4.18) and the hyper covariance in Eq. (4.19). Finally, it is also noted that the identification uncertainties shown in Fig. 4-4 are small compared to the parameter variability shown by the hyper standard deviations in Fig. 4-5.

The posterior distribution of model error for C1, C2 and SS1 can be computed together with the posterior distribution of the hyper parameters according to the approximation A-1 in Eq. (4.15). Results for the posterior distribution of model error are depicted in Fig. 4-6. As seen, due to the fact that the model error is present in the first mass, both the component C1 and subsystem SS1 exhibit a large model error. It is also seen that the model error in C1 is larger than that of SS1. This can be attributed to the fact that at the subsystem level the number of parameters involved are more than the number of the parameters involved at the component level, making the three-parameter model more flexible to provide a better fit to the datasets than the one-parameter model involved at the component level, thus reducing the model error that is quantified by the prediction error parameters. The model error for C2 is fairly close to zero implying excellent matches between the updated model and the measurements. This is due to the absence of the model error assigned in component C2.

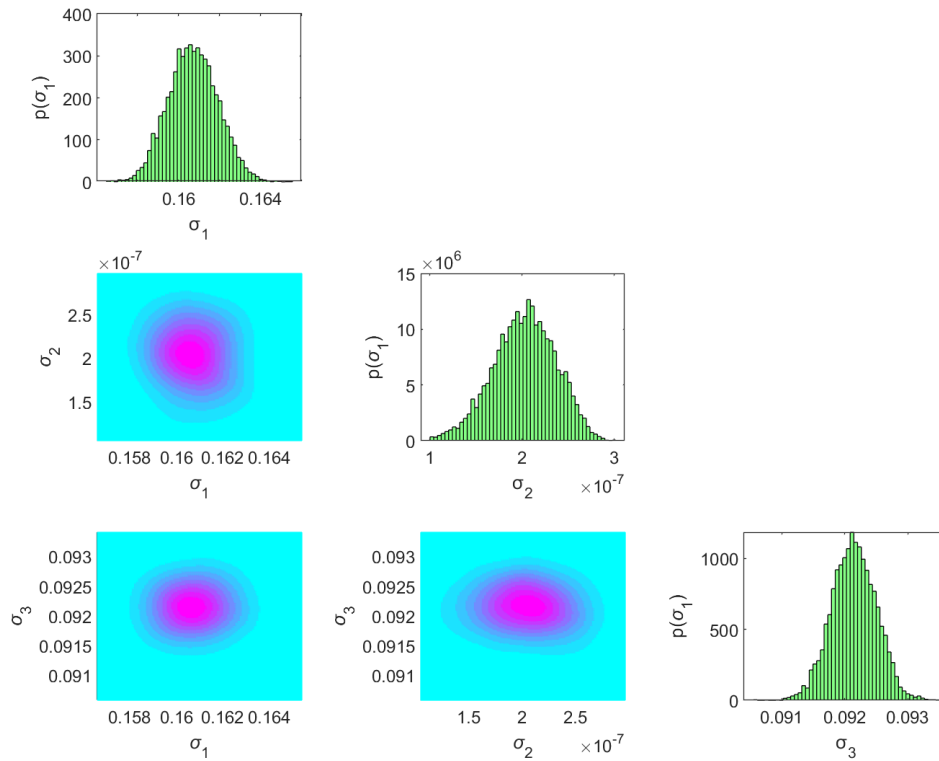


Fig. 4-6 Posterior distributions of model error for C1, C2 and SS1

For the purpose of comparing the accuracy of the proposed approximation A-1 and A-2, the results of the hyper parameters alongside the model error from A-1 and A-2 as well as results for these hyperparameters obtained from the analytical solutions derived in Subsection 4.2.2.4 are compared to the full simulation (FS) approach [25] based on TMCMC algorithm [31,32] and reported in Table 4-2. The FS procedure provides an accurate solution for a large number of samples. It is seen that both the results from A-1 and A-2 are in good agreements with the FS method. This is due to the fact that the asymptotic approximation holds true when the identification uncertainty is small compared to the parameter variability. The likelihood of the model parameters can be asymptotically approximated as a Gaussian distribution when given a large number of data points in a dataset (for C1 and SS1) or given a perfect computational model (for C2). The lightly difference between A-2 and the FS shown in Table 4-2 may arise from the number of insufficient samples used for FS. Most importantly, both the results from A-1 and A-2 exhibit acceptable accuracy. The A-2 exhibits higher accuracy than A-1 due to the fact that the asymptotic approximation for the prediction error parameters considered in A-1 is less accurate. Compared to the analytical results with A-1 and A-2, it is found that the analytical results are also slightly different from the results obtained by the A-1 and A-2 approximations. This is due to the assumptions of the identification uncertainty that

was used to simplify derivations and obtain the analytical expressions. For gaining more insightful expressions, the analytical solutions was derived by selecting the identification uncertainty to be the same for each dataset and equal to the average of the identification uncertainty over all datasets. However, the identification uncertainty may not be exactly the same in each dataset. These assumptions have an effect on the accuracy of the analytical expressions. The analytical solutions do provide a better understanding of the estimates of the hyper parameters and the model error parameters. However, from the results in Table 4-2 they also provide very good estimates of the hyperparameter and model error parameter values.

The computational gain of using the approximations A-1 and A-2 compared with FS is also investigated and reported in Table 4-3. Herein the sample procedure employs the parallel version of the TMCMC algorithm [32] so that the parallel tools can be used during the computation. For guaranteeing the accuracy of FS, the number of the samples is chosen to be 10000. All the calculations are made in a 32-core computer. It is clear that the computational effort of A-1 and A-2 are much faster than FS, although the parallel tools are used for the FS method. Specifically, the running time for A-1 and A-2 are relatively close and less than 1.5 minutes. However, it takes around 38.3 minutes to get the same results by using FS. As a consequence, the proposed approximations A-1 and A-2 not only guarantee good results, but more importantly they can significantly improve the computational efficiency, providing faster computational tools for quantifying the uncertainties in a multi-level modeling approach.

Table 4-2 Estimates of the means of hyper parameters and model error

Parameters	$\hat{\mu}_{\psi_1}$	$\hat{\mu}_{\psi_2}$	$\hat{\mu}_{\psi_3}$	$\hat{\sigma}_{\psi_1}$	$\hat{\sigma}_{\psi_2}$	$\hat{\sigma}_{\psi_3}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$
A-1	1.0476	0.9846	0.9930	0.0101	0.0471	0.0490	0.1610	1.9×10^{-7}	0.0923
A-2	1.0475	0.9828	0.9933	0.0100	0.0460	0.0477	0.1714	2.3×10^{-7}	0.0997
Analytical A-1	1.0476	0.9851	0.9922	0.0099	0.0468	0.0491	0.1621	1.9×10^{-7}	0.0917
Analytical A-2	1.0476	0.9851	0.9922	0.0098	0.0446	0.0453	0.1737	2×10^{-7}	0.0998
FS	1.0476	0.9861	0.9922	0.0104	0.0460	0.0482	0.1703	2.3×10^{-7}	0.0986

Table 4-3 Computational effort of A-1, A-2 and FS methods

Computational cost (s)	
A-1	80
A-2	89
FS	2301

Given the hyperparameter uncertainties in Fig. 4-5, the posterior distribution of each model parameter is obtained under the same cases with Fig. 4-5 where the datasets are considered from components only, subsystem only and both of them. Results are shown in Fig. 4-7. It is seen that the means of the parameter $\theta^{(1)}$ and $\theta^{(2)}$ computed by considering the overall information is a kind of average of the values obtained using datasets from components only and subsystem only. It is also seen that the uncertainty of parameter $\theta^{(1)}$ using datasets available at component C1 only is larger than that using both the datasets available at component C1 and subsystem SS1, while the opposite case occurs for the uncertainty of the parameter $\theta^{(2)}$. This is attributed to the effect that datasets from the subsystem level has on the estimates of these parameters. As shown in Fig. 4-5, a small standard deviation is computed for parameter $\theta^{(1)}$ using the subsystem level datasets, while a larger one is obtained for parameter $\theta^{(2)}$ using datasets from the subsystem level. For the third parameter $\theta^{(3)}$, since there is no dataset available for component C3, the results are expected to be exactly the same using the datasets available at subsystem only and at both components and subsystem.

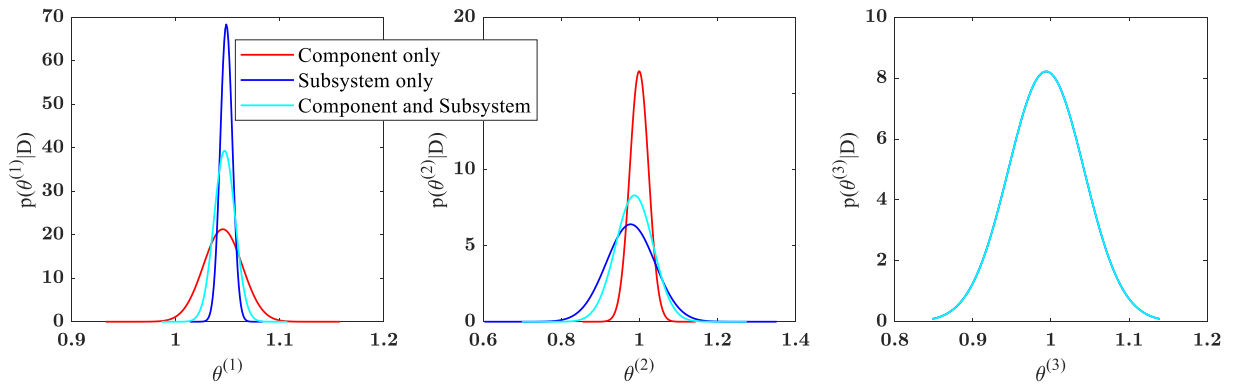


Fig. 4-7 Posterior distribution of model parameters considering the datasets from components only, sub-systemonly and both the components and sub-system

The marginal and joint posterior distributions of the model parameters, computed according to Eq. (4.29), are shown in Fig. 4-8. Such posterior distributions are calculated based on all the datasets from all levels of hierarchy, and thus the model parameters involve all the information from the components and subsystem levels. It is apparent that large uncertainties are obtained for the model parameters which are irreducible as the number of datasets at all levels of hierarchy increases. In contrast, the identification uncertainties

obtained from each dataset in Fig. 4-4 yield unreasonable thin uncertainty bounds for the model parameters. Most importantly, the irreducible uncertainties can be propagated to the un-observed QoI, which may provide a realistic uncertainty bound for the response predictions. Results and discussions for predicting the responses through the system level are conducted in the next subsection.

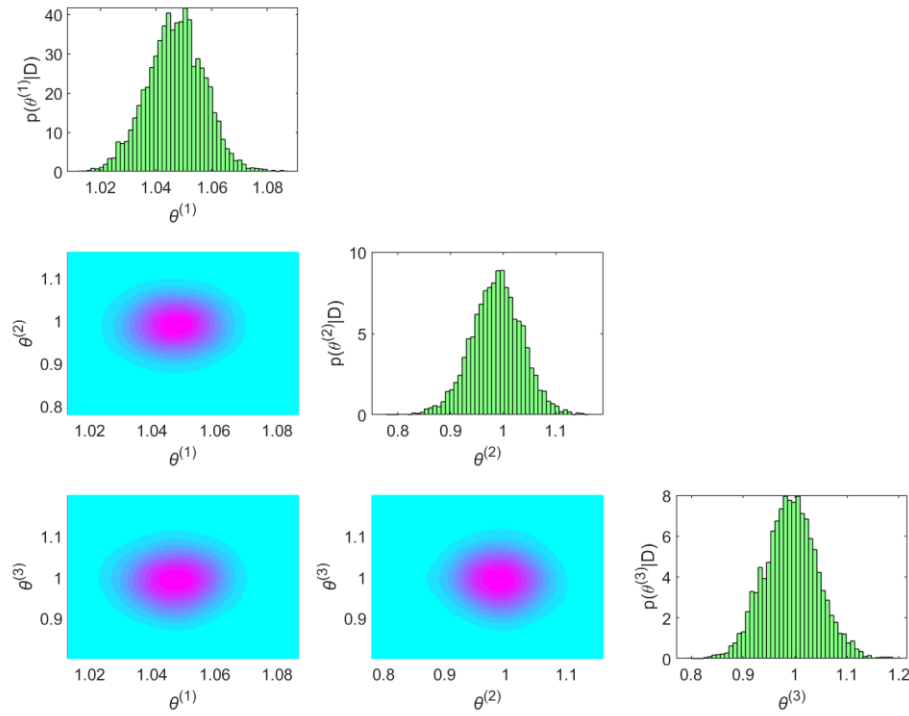


Fig. 4-8 Posterior distributions of model parameters

4.3.2.2 Response predictions through the computational model at system level

The parameter uncertainties are next propagated through the hierarchy of computational models for computing the response predictions at the system level. The acceleration input with 5.2s and 100Hz sampling frequency shown in Fig. 4-9 is applied to the base of the system model in Fig. 4-3. For exploring the effect of datasets from the different levels of hierarchy on the uncertainty in the response predictions, results for the response predictions considering datasets from components/sub-system only are compared with the ones considering the datasets from both the components and sub-system. Fig. 4-10 shows the predicted results of displacement of the second and fourth DOFs by propagating the uncertainties through the system model. The light red shaded area shows the 95% uncertainty bounds (UB) of displacements considering the uncertainty of model parameter $\theta^{(1)}$ inferred from the datasets of C1 only, and the uncertainties of $\theta^{(2)}$ and $\theta^{(3)}$ inferred from the overall

datasets of components and subsystem, while the light blue shaded area displays the 95% UB of displacements considering the uncertainties of model parameters θ inferred from all datasets available from the components and subsystem. The cyan and blue solid lines show the mean predictions corresponding to the light red shaded area and light blue shaded area, respectively, and the red line shows the actual measurements generated from the exact system model subjected to the base input in Fig. 4-9. It is observed that a slightly difference can be found for the mean predictions between the case where the model parameter $\theta^{(1)}$ is inferred from the dataset at component C1 level only and the case where the model parameter $\theta^{(1)}$ is inferred from the datasets available from both component C1 and subsystem SS1 levels. This is due to the fact that the hyper mean estimated from component C1 level is slightly different with the one identified from both component C1 and subsystem SS1 levels, as seen in the first column of Fig. 4-5. Regarding the uncertainty bounds of the displacements, it is observed that the 95% UB of displacement of the 2nd DOF calculated for the case where $\theta^{(1)}$ is inferred from component C1 level datasets is somewhat larger than that computed for the case where $\theta^{(1)}$ is inferred from both component C1 and subsystem SS1 level datasets. This is reasonable as the hyper standard deviation of the first model parameter computed from component C1 only is larger than the one estimated from both component C1 and subsystem SS1 (see Fig. 4-5). This scenario reveals that the uncertainty in response predictions depends on the datasets used at the different levels of modeling hierarchy to infer the parameters of the models involved. Herein, the uncertainties of the model parameter $\theta^{(1)}$, estimated based on a comprehensive effect considering both the information from component C1 and subsystem SS1, provides different uncertainty bounds than the ones provided by the uncertainty in the model parameter $\theta^{(1)}$ inferred by neglecting the data at the subsystem level. The same finding is revealed for the displacement of the fourth DOF.

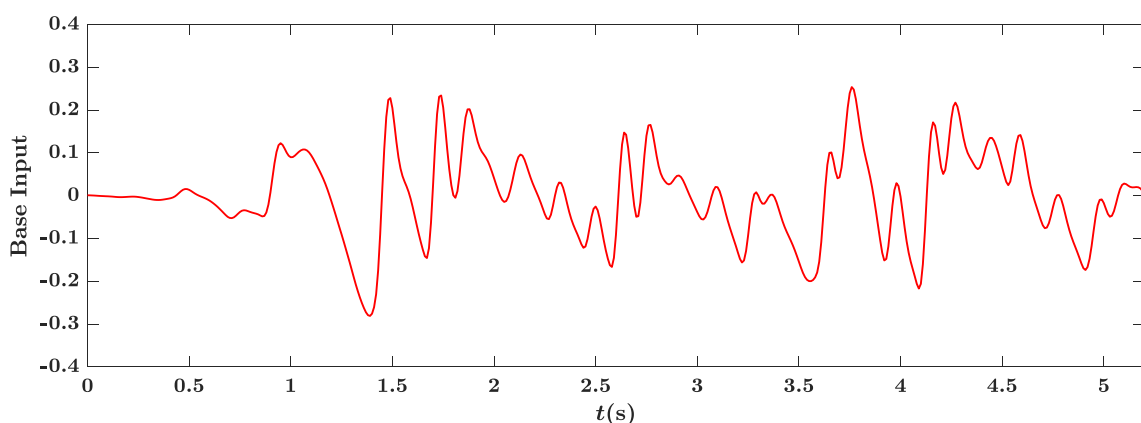


Fig. 4-9 Base input

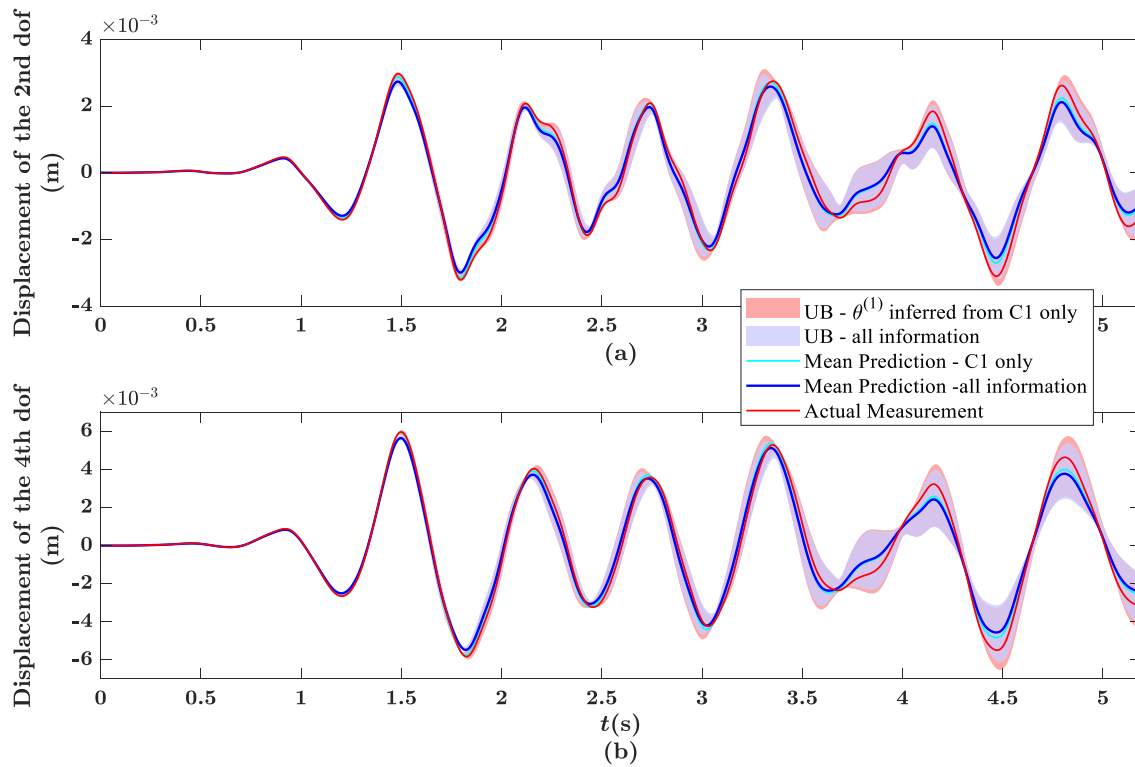


Fig. 4-10 Predictions of displacement of (a) the second and (b) the fourth DOFs by considering the uncertainty of parameter $\theta^{(1)}$ inferred from C1 only and also inferred from all information

Results for the predictions of displacement of the 2nd and 4th DOFs are also conducted by considering the uncertainty of parameter $\theta^{(2)}$ inferred from component C2 level datasets only, as shown in Fig. 4-11. It is found that the mean predictions of displacements considering the uncertainty of parameter $\theta^{(2)}$ inferred from component C2 datasets only are different from the ones inferred from the datasets available from all levels. Again, this is attributed to the difference of the identified values of the model parameters between component C2 level datasets only and both component C2 and subsystem SS1 level datasets (see Fig. 4-5). For the UB of displacements, different scenario is found compared to the cases in Fig. 4-10. It is obvious that the 95% UB of displacements considering only the uncertainties of model parameters from component C2 are much smaller than the uncertainty bounds computed based on the uncertainties from the overall information, revealing that the information of the model parameters is also taken from the subsystem SS1. This finding is crucial for a multi-level modeling approach, especially when the uncertainties of the model parameters identified from the components level is insignificant. Besides, due to the large variability

obtained from component C2 and subsystem SS1 levels, the actual displacement generated based on a new dataset through system model falls within the 95% UB of predicted displacement.

Cases are also investigated for the UB of the modal properties. Fig. 4-12 and Fig. 4-13 show the predictions of modal frequencies and mode shapes under the same conditions with Fig. 4-11. For the ease of comparison, the modal frequencies are normalized by the nominal values of the modal frequencies. It is seen that the uncertainty bounds of the predictions, obtained considering only the model parameter uncertainties inferred from component C2 datasets, contain the real measurement. This is due to the fact that the parameter variability in component C2 is significant compared with the zero model error assumed in component C2. When the uncertainties of model parameters estimated from both component C2 and subsystem SS1 are considered, larger uncertainty bounds of modal frequencies are observed in Fig. 4-12(b), indicating that datasets from different levels may significantly affect prediction uncertainties. The results also show the measured modal properties corresponding to a new dataset. It is clear that the new measurement is inside of the 95% UB of the predictions. For the predictions of mode shapes, similar results are obtained although the uncertainty seems to be smaller, especially for the lowest four modes, than the uncertainty observed for modal frequencies. This could be due to the insensitivity of the lowest four modeshapes to the model parameters.

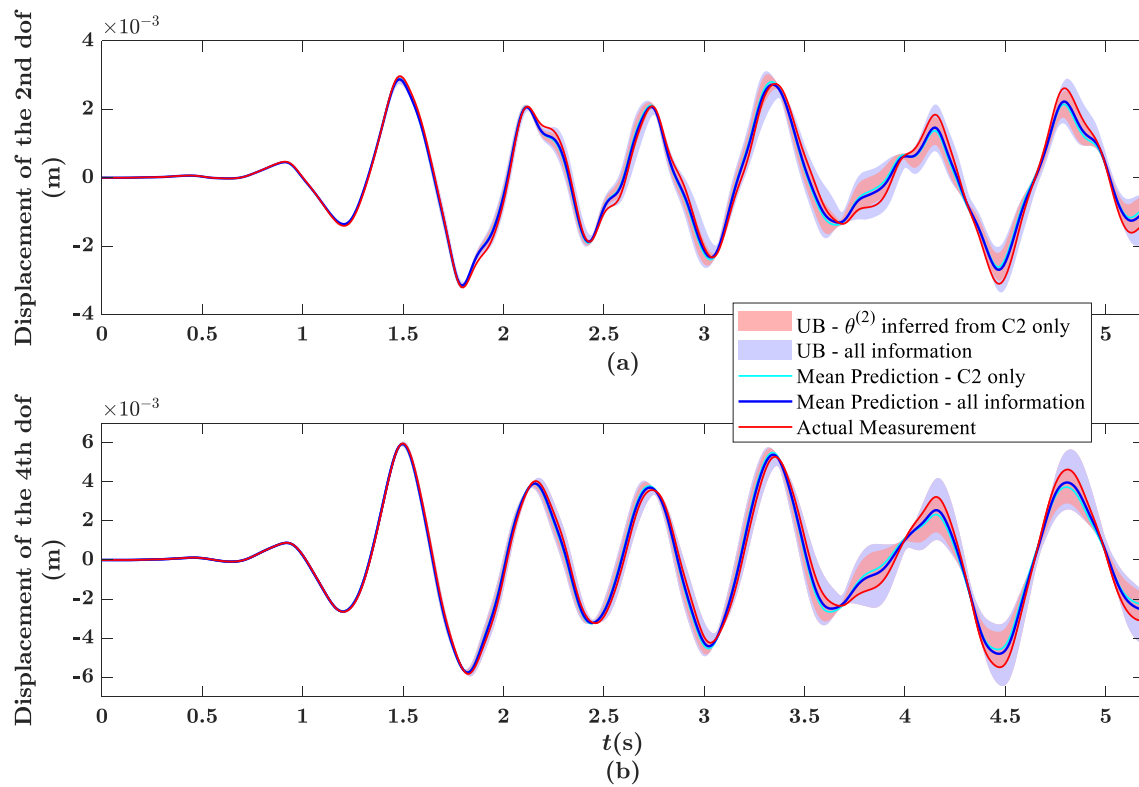


Fig. 4-11 Predictions of displacement of (a) the second and (b) the fourth DOFs by considering the uncertainty of parameter $\theta^{(2)}$ inferred from C2 only and inferred from all information

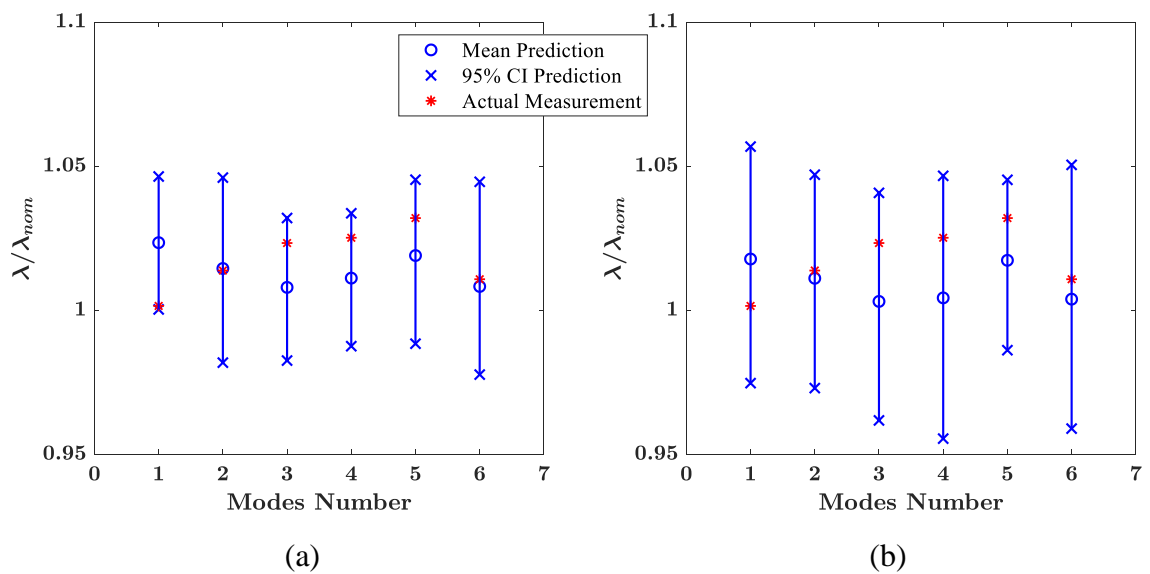


Fig. 4-12 Prediction of the ratio of modal frequencies to their nominal values obtained by propagating the uncertainty of the model parameter $\theta^{(2)}$ (a) inferred from C2 level datasets only (b) inferred from all available datasets

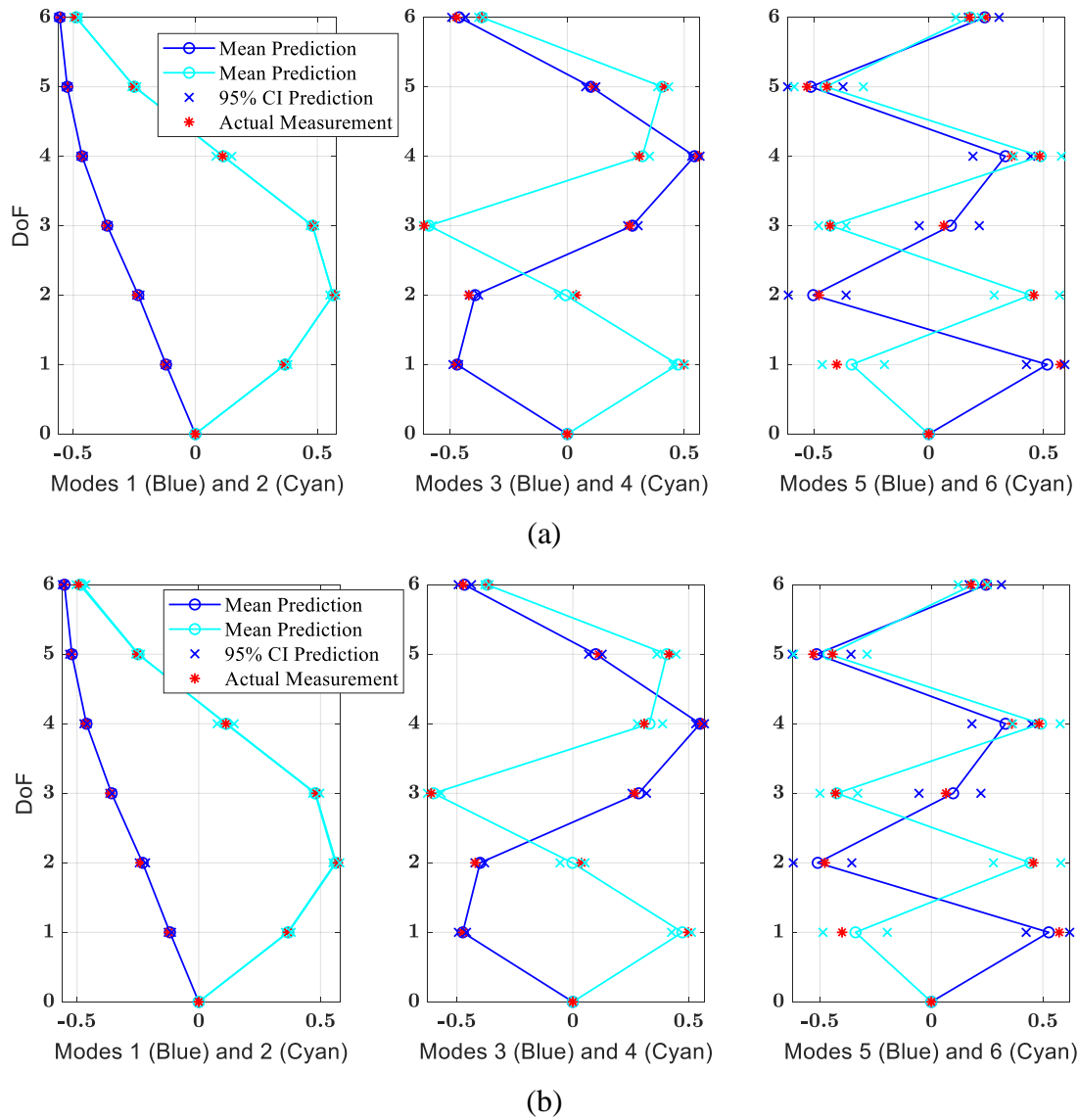


Fig. 4-13 Prediction of the mode shapes by propagating the uncertainty of the model parameter $\theta^{(2)}$ (a) inferred from C2 level datasets only, and (b) inferred from all available datasets

Fig. 4-14 depicts the predicted acceleration of the fifth DOF by propagating different types of uncertainties through the system model. The light red shaded area shows the 95% UB of predictions considering the parameter uncertainties inferred from subsystem SS1 level datasets only, while the blue one shows the 95% UB considering the parameter uncertainties inferred from all available datasets. The model error is also considered for the predicted 5th DOF acceleration, as shown in the green shaded area in Fig. 4-14(b). It is reminded that in Fig. 4-5, the hyper standard deviation σ_{ψ_1} from subsystem SS1 datasets only is smaller than

that from both subsystem SS1 and component C1 datasets, while the opposite case is observed for hyper standard deviation σ_{ψ_2} . Interestingly, the 95% UB of acceleration considering the uncertainties of model parameters from subsystem SS1 only is wider than that from both subsystem SS1 and components. This can be due to the fact that the hyper standard deviation σ_{ψ_2} is larger than the hyper standard deviation σ_{ψ_1} , and the former one contributes more on the uncertainty bounds of the predictions. When the model error is considered for the prediction affect the UB from 0 to 1.5 sec while the UB for higher that 1.5 sec remain unaffected, indicating that in this case that a large part of the uncertainty is captured by the model parameters.

The parameter uncertainties computed from the overall information are also used to predict the UB of displacement and acceleration of the 6th DOF which is at the system level and not at the levels of the tested components or subsystem. Actual measurements are also generated from the perturbed model. Results shown in Fig. 4-15 depict the 95% UB of the predictions along with the mean predictions and measurements. It is expected that the measurements are far away from the mean predictions due to the presence of model error assumed in the first mass. However, wide uncertainty bounds are obtained for both the displacement and acceleration. The measurements over almost the whole duration fall within the 95% UB of the predictions, which demonstrates the validity and effectiveness of the proposed approach.

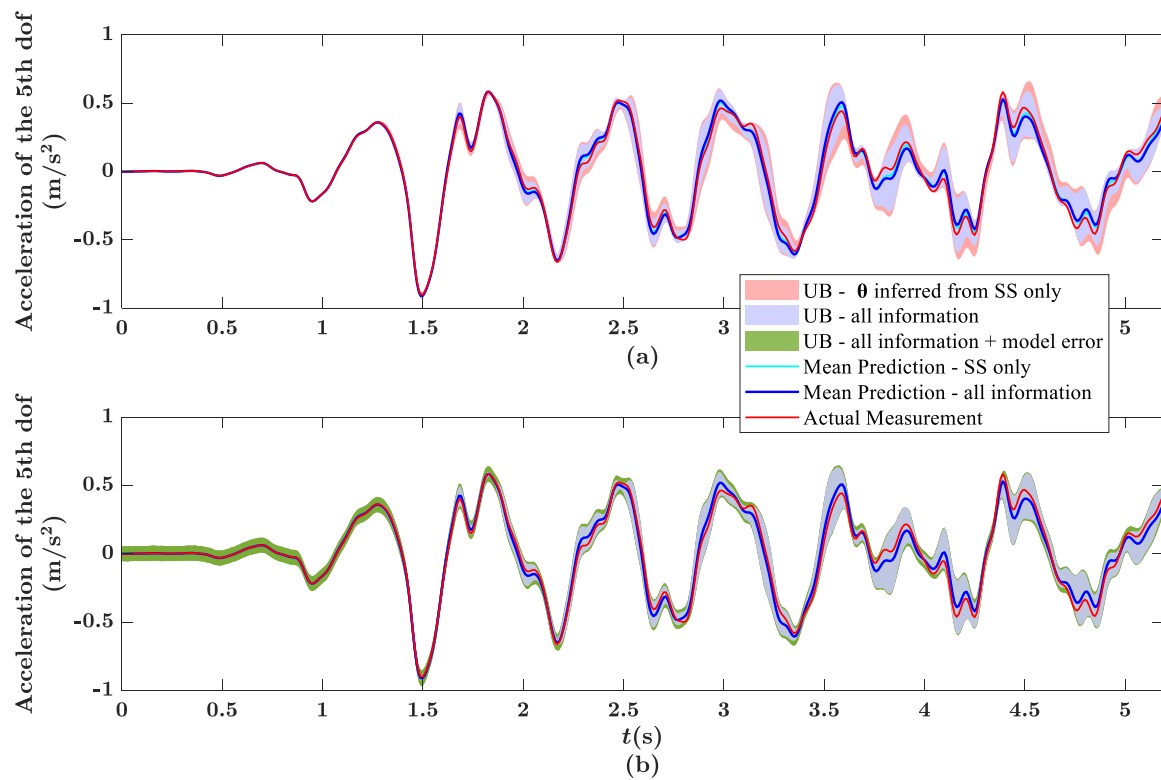


Fig. 4-14 Predictions of acceleration of the fifth DOF by propagating the uncertainties of (a) model parameters inferred from SS only and inferred from all information and (b) model parameters and prediction error

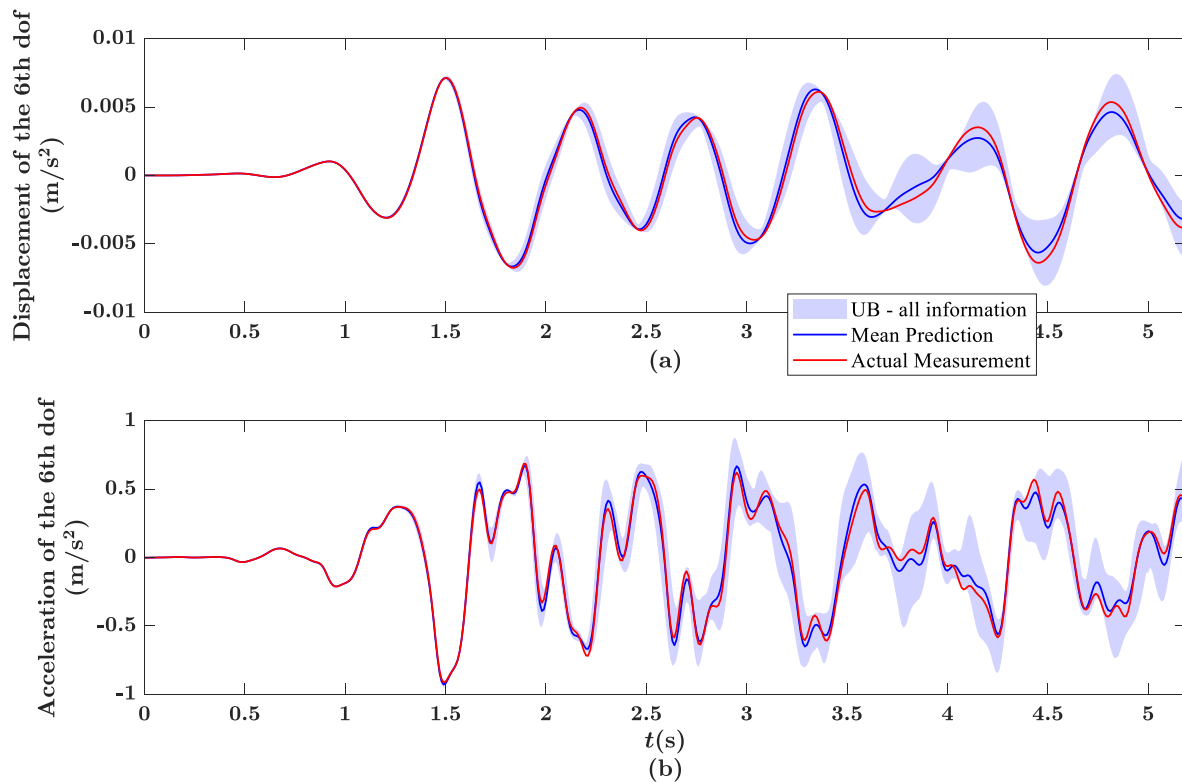


Fig. 4-15 Displacement and acceleration of the sixth DOF by propagating the parameter uncertainties through the system model

4.4 Concluding Remarks

A systematic hierarchical Bayesian learning framework is developed to account for model hierarchy in structural dynamics utilizing multiple datasets obtained from different levels of the hierarchy. The framework can incorporate the uncertainty in the model parameters manifested in the different levels of model hierarchy by assigning hyper distributions for model parameters and inferring the hyperparameters using datasets collected from testing at the different levels of model hierarchy. The proposed approach captures the uncertainty in the model parameters due to variabilities in experimental data, environmental conditions, material properties, manufacturing process, assembling process as well as different mechanisms activated under different loading conditions. Asymptotic approximations developed in [29] are integrated within the framework, considerably simplifying the computational process of estimating the uncertainties in the hyper parameters and model prediction error parameters. The accuracy of the approximations is guaranteed for sufficiently large number of data within a dataset. Analytical expressions are derived for the most probable values of the hyper parameters and model error parameters, providing useful insight on the effect of datasets from the different levels of model hierarchy on the estimates of the model hyper parameters. Ultimately, the uncertainties are propagated for predicting confidence levels for output QoI by considering the overall information provided from the datasets available at the different levels of model hierarchy.

The effectiveness of the proposed framework is demonstrated by applying it to a simple system that includes two lower levels of hierarchy, with datasets available at the component level for two out of the three components considered and also at the subsystem level. Selective results are presented to demonstrate various issues related to the dependence of the uncertainties in the parameters and the predictions obtained by propagating the uncertainties through the model hierarchy on the availability of datasets at different levels of model hierarchy, the effect of the level of model error and level of manufacturing variability. The accuracy of the asymptotic approximations and the analytical expressions for estimating the MPV of the hyperparameters were also investigated by comparing results with the ones obtained from the full sampling approach. Good accuracy is observed which makes the asymptotic approximations suitable to address the computational burden that is involved in uncertainty estimation and propagation problems. Specifically, for the example application presented, the computational effort of the proposed approach is significantly less, by one to two order of magnitude, than the computation effort required from full sampling approaches.

Predictions of the (un)observed QoI are conducted by considering different cases of uncertainty modeling in the model hierarchy.

The proposed method properly accounts for the uncertainties by considering the information contained in all datasets available from the different levels of hierarchy, providing a realistic tool for calibrating the parameters manifested at different levels of model hierarchy, and propagating the data-informed uncertainties through the model hierarchy to compute output QoI at the system level, thus improving confidence in response predictions. In this work, the formulation is presented using a specific structure of model hierarchy for the system assembled from two subsystems and one of the subsystems is assembled by three components. Datasets are available at component level for the two out of the three components and also at the subsystem level for one subsystem. General structures of model and data hierarchies were not treated. Moreover, the 6-DOF spring-mass chain model used to demonstrate the framework was selected due to its simplicity, with the procedure be representative of more complex models. The application of the framework in more complex computational models is only expected to substantially increase the computational effort due to the presence of more parameters per component level and the time-consuming operations for performing model simulations at various levels of model hierarchy. It is the subject of future work to generalize the proposed approach to account for complex computational models and data dependencies.

Appendix A. Calculations of the MPVs of hyper parameters and prediction error variance parameters for approximations A-1 and A-2

The MPVs of hyper parameters and prediction error variance parameters can be obtained by minimizing the negative logarithm of the joint distributions $p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D})$ in Eqs. (4.16) for approximation A-1 and (4.17) for approximation A-2, given by (within a constant):

$$\begin{aligned}
L_{A-1} &= -\ln p(\boldsymbol{\psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) \\
&= \frac{1}{2} \sum_{i=1}^{N_1} \ln(|\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)}|) + \frac{1}{2} \sum_{i=1}^{N_1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)}) - \ln \prod_{i=1}^{N_1} N(\sigma_1^2 | \hat{\sigma}_{1,i}^2, \hat{\Sigma}_{\sigma_{1,i}^2}) \\
&+ \frac{1}{2} \sum_{j=1}^{N_2} \ln(|\boldsymbol{\Sigma}_{\psi_2} + \hat{\boldsymbol{\Sigma}}_{\theta_j}^{(2)}|) + \frac{1}{2} \sum_{j=1}^{N_2} (\boldsymbol{\mu}_{\psi_2} - \hat{\boldsymbol{\theta}}_j^{(2)})^T (\boldsymbol{\Sigma}_{\psi_2} + \hat{\boldsymbol{\Sigma}}_{\theta_j}^{(2)})^{-1} (\boldsymbol{\mu}_{\psi_2} - \hat{\boldsymbol{\theta}}_j^{(2)}) - \ln \prod_{j=1}^{N_2} N(\sigma_2^2 | \hat{\sigma}_{2,j}^2, \hat{\Sigma}_{\sigma_{2,j}^2}) \quad (\text{A.1}) \\
&+ \frac{1}{2} \sum_{k=1}^{N_3} \ln(|\boldsymbol{\Sigma}_{\psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k}|) + \frac{1}{2} \sum_{k=1}^{N_3} (\boldsymbol{\mu}_{\psi} - \hat{\boldsymbol{\theta}}_k)^T (\boldsymbol{\Sigma}_{\psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k})^{-1} (\boldsymbol{\mu}_{\psi} - \hat{\boldsymbol{\theta}}_k) - \ln \prod_{k=1}^{N_3} N(\sigma_3^2 | \hat{\sigma}_{3,k}^2, \hat{\Sigma}_{\sigma_{3,k}^2})
\end{aligned}$$

$$\begin{aligned}
L_{A-2} &= -\ln p(\boldsymbol{\Psi}, \boldsymbol{\sigma}^2 | \mathbf{D}) \\
&= -\frac{1}{2} \sum_{i=1}^{N_1} \ln \left(\left| \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)}(\sigma_1^2) \right| \right) + \frac{1}{2} \sum_{i=1}^{N_1} \ln \left(\left| \boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)}(\sigma_1^2) \right| \right) + \frac{1}{2} \sum_{i=1}^{N_1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)}(\sigma_1^2))^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)}) \\
&\quad - \frac{1}{2} \sum_{j=1}^{N_2} \ln \left(\left| \hat{\boldsymbol{\Sigma}}_{\theta_j}^{(2)}(\sigma_2^2) \right| \right) + \frac{1}{2} \sum_{j=1}^{N_2} \ln \left(\left| \boldsymbol{\Sigma}_{\Psi_2} + \hat{\boldsymbol{\Sigma}}_{\theta_j}^{(2)}(\sigma_2^2) \right| \right) + \frac{1}{2} \sum_{j=1}^{N_2} (\boldsymbol{\mu}_{\Psi_2} - \hat{\boldsymbol{\theta}}_j^{(2)})^T (\boldsymbol{\Sigma}_{\Psi_2} + \hat{\boldsymbol{\Sigma}}_{\theta_j}^{(2)}(\sigma_2^2))^{-1} (\boldsymbol{\mu}_{\Psi_2} - \hat{\boldsymbol{\theta}}_j^{(2)}) \quad (\text{A.2}) \\
&\quad - \frac{1}{2} \sum_{k=1}^{N_3} \ln \left(\left| \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(3)}(\sigma_3^2) \right| \right) + \frac{1}{2} \sum_{k=1}^{N_3} \ln \left(\left| \boldsymbol{\Sigma}_{\Psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(3)}(\sigma_3^2) \right| \right) + \frac{1}{2} \sum_{k=1}^{N_3} (\boldsymbol{\mu}_{\Psi} - \hat{\boldsymbol{\theta}}_k)^T (\boldsymbol{\Sigma}_{\Psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(3)}(\sigma_3^2))^{-1} (\boldsymbol{\mu}_{\Psi} - \hat{\boldsymbol{\theta}}_k) \\
&\quad - \ln T(\sigma_1^2) - \ln T(\sigma_2^2) - \ln T(\sigma_3^2)
\end{aligned}$$

For approximation A-1, the first derivatives of the objective function L_{A-1} with respect to the parameters $\boldsymbol{\mu}_{\Psi_1}$ and $\boldsymbol{\Sigma}_{\Psi_1}$ depend on the terms of $\boldsymbol{\mu}_{\Psi_1}$, $\boldsymbol{\mu}_{\Psi}$ and $\boldsymbol{\Sigma}_{\Psi_1}$, $\boldsymbol{\Sigma}_{\Psi}$, respectively, given by:

$$\begin{aligned}
\frac{\partial L_{A-1}}{\partial \boldsymbol{\mu}_{\Psi_1}} &= \frac{1}{2} \frac{\partial \sum_{i=1}^{N_1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})}{\partial \boldsymbol{\mu}_{\Psi_1}} \\
&\quad + \frac{1}{2} \frac{\partial \sum_{k=1}^{N_3} (\boldsymbol{\mu}_{\Psi} - \hat{\boldsymbol{\theta}}_k)^T (\boldsymbol{\Sigma}_{\Psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(3)})^{-1} (\boldsymbol{\mu}_{\Psi} - \hat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\mu}_{\Psi_1}} \quad (\text{A.3})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial L_{A-1}}{\partial \boldsymbol{\Sigma}_{\Psi_1}} &= \frac{1}{2} \frac{\partial \sum_{i=1}^{N_1} \ln \left(\left| \boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)} \right| \right) + \sum_{i=1}^{N_1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})}{\partial \boldsymbol{\Sigma}_{\Psi_1}} \\
&\quad + \frac{1}{2} \frac{\partial \sum_{k=1}^{N_3} \ln \left(\left| \boldsymbol{\Sigma}_{\Psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(3)} \right| \right) + \sum_{k=1}^{N_3} (\boldsymbol{\mu}_{\Psi} - \hat{\boldsymbol{\theta}}_k)^T (\boldsymbol{\Sigma}_{\Psi} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(3)})^{-1} (\boldsymbol{\mu}_{\Psi} - \hat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\Sigma}_{\Psi_1}} \quad (\text{A.4})
\end{aligned}$$

It has been shown in [28,29] that the first terms in Eqs. (A.3) and (A.4) simplify to

$$\frac{1}{2} \frac{\partial \sum_{i=1}^{N_1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})}{\partial \boldsymbol{\mu}_{\Psi_1}} = \sum_{i=1}^{N_1} (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)}) \quad (\text{A.5})$$

$$\begin{aligned}
&\frac{1}{2} \frac{\partial \sum_{i=1}^{N_1} \ln \left(\left| \boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)} \right| \right) + \sum_{i=1}^{N_1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})}{\partial \boldsymbol{\Sigma}_{\Psi_1}} \quad (\text{A.6}) \\
&= \frac{1}{2} \sum_{i=1}^{N_1} (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} - \frac{1}{2} \sum_{i=1}^{N_1} (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)}) (\boldsymbol{\mu}_{\Psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\Psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1}
\end{aligned}$$

while the second term in Eq. (A.3) can be readily solved by using the chain rules [35], resulting in the form:

$$\begin{aligned} \frac{1}{2} \frac{\partial \sum_{k=1}^{N_3} (\boldsymbol{\mu}_\psi - \hat{\boldsymbol{\theta}}_k)^T (\boldsymbol{\Sigma}_\psi + \hat{\boldsymbol{\Sigma}}_{\theta_k})^{-1} (\boldsymbol{\mu}_\psi - \hat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\mu}_{\psi_1}} &= (\mathbf{I} \mathbf{0} \mathbf{0}) \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_\psi + \hat{\boldsymbol{\Sigma}}_{\theta_k})^{-1} (\boldsymbol{\mu}_\psi - \hat{\boldsymbol{\theta}}_k) \\ &= \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_k^{(1)}) \end{aligned} \quad (\text{A.7})$$

where the size of the identity matrix \mathbf{I} and zero matrix $\mathbf{0}$ depend on the size of vector $\boldsymbol{\mu}_\psi$ and the size of entries matrix in $\boldsymbol{\mu}_\psi$, and $\hat{\boldsymbol{\theta}}_k^{(1)}$ and $\hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)}$ are the entries of vector $\hat{\boldsymbol{\theta}}_k$ and matrix $\hat{\boldsymbol{\Sigma}}_{\theta_k}$.

The second term of Eq. (A.4) can be also solved using the chain rules, given as:

$$\begin{aligned} \frac{1}{2} \frac{\partial \sum_{k=1}^{N_3} \ln(\|\boldsymbol{\Sigma}_\psi + \hat{\boldsymbol{\Sigma}}_{\theta_k}\|) + \sum_{k=1}^{N_3} (\boldsymbol{\mu}_\psi - \hat{\boldsymbol{\theta}}_k)^T (\boldsymbol{\Sigma}_\psi + \hat{\boldsymbol{\Sigma}}_{\theta_k})^{-1} (\boldsymbol{\mu}_\psi - \hat{\boldsymbol{\theta}}_k)}{\partial \boldsymbol{\Sigma}_{\psi_1}} \\ = \frac{1}{2} \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} - \frac{1}{2} \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_k^{(1)}) (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_k^{(1)})^T (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} \end{aligned} \quad (\text{A.8})$$

Therefore, Eqs. (A.3) and (A.4) can be derived in the form

$$\frac{\partial L_{A-1}}{\partial \boldsymbol{\mu}_{\psi_1}} = \sum_{i=1}^{N_1} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)}) + \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_k^{(1)}) \quad (\text{A.9})$$

$$\begin{aligned} \frac{\partial L_{A-1}}{\partial \boldsymbol{\Sigma}_{\psi_1}} &= \frac{1}{2} \sum_{i=1}^{N_1} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} - \frac{1}{2} \sum_{i=1}^{N_1} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)}) (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_i^{(1)})^T (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)})^{-1} \\ &\quad + \frac{1}{2} \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} - \frac{1}{2} \sum_{k=1}^{N_3} (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_k^{(1)}) (\boldsymbol{\mu}_{\psi_1} - \hat{\boldsymbol{\theta}}_k^{(1)})^T (\boldsymbol{\Sigma}_{\psi_1} + \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)})^{-1} \end{aligned} \quad (\text{A.10})$$

For getting more insightful expressions, the identification uncertainties $\hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)}$ and $\hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)}$ are assumed to be equal to average uncertainty for all datasets from component C1 and subsystem SS1, given by:

$$\boldsymbol{\Sigma}^{(1)} = \frac{1}{N_1 + N_3} \left(\sum_{i=1}^{N_1} \hat{\boldsymbol{\Sigma}}_{\theta_i}^{(1)} + \sum_{k=1}^{N_3} \hat{\boldsymbol{\Sigma}}_{\theta_k}^{(1)} \right) \quad (\text{A.11})$$

Letting $\frac{\partial L_{A-1}}{\partial \boldsymbol{\mu}_{\psi_1}}$ and $\frac{\partial L_{A-1}}{\partial \boldsymbol{\Sigma}_{\psi_1}}$ be equal to zero, one can get the MPVs of $\boldsymbol{\mu}_{\psi_1}$ and $\boldsymbol{\Sigma}_{\psi_1}$ in the form

given by the first of Eq. (4.18) and first of Eq. (4.19).

Following the same procedure for the calculations of the MPVs of $\boldsymbol{\mu}_{\psi_1}$ and $\boldsymbol{\Sigma}_{\psi_1}$, one can also get the MPVs of other hyper parameters to be given by the Eqs. (4.18) and (4.19).

For the first derivatives of the objective function L_{A-1} with respect to the prediction error parameters, the product of multiple normal distribution of scalar variable σ_r^2 , $r = 1, 2, 3$, can

be calculated firstly. As an example, the product $\prod_{i=1}^{N_1} N(\sigma_1^2 | \hat{\sigma}_{1,i}^2, \hat{\Sigma}_{\sigma_1^2})$ can be given by a normal distribution with mean given by [36]:

$$\hat{\sigma}_1^2 = \sum_{i=1}^{N_1} \frac{1}{\hat{\sigma}_{1,i}^2} / \sum_{i=1}^{N_1} \frac{1}{\hat{\sigma}_{1,i}^4} \quad (\text{A.12})$$

Similar results can be found for other two prediction error parameters, given as:

$$\hat{\sigma}_2^2 = \sum_{j=1}^{N_2} \frac{1}{\hat{\sigma}_{2,j}^2} / \sum_{j=1}^{N_2} \frac{1}{\hat{\sigma}_{2,j}^4}; \quad \hat{\sigma}_3^2 = \sum_{k=1}^{N_3} \frac{1}{\hat{\sigma}_{3,k}^2} / \sum_{k=1}^{N_3} \frac{1}{\hat{\sigma}_{3,k}^4} \quad (\text{A.13})$$

The formulas (4.24) to (4.27) for the MPVs of all parameters for the approximation A-2 are obtained in a similar manner by utilizing the chain rule and existing solutions of the MPVs of hyper and prediction error parameters for a full-scale system [29].

4.5 References

- [1] O.C. Zienkiewicz, R.L. Taylor, Finite Element Method for Solid and Structural Mechanics, Elsevier, 2005.
- [2] T.J.R. Hughes, The finite element method: linear static and dynamic finite element analysis, Courier Corporation, 2012.
- [3] J.E. Mottershead, M.I. Friswell, Model updating in structural dynamics: A survey, Journal of Sound and Vibration. 167 (1993) 347–375. <https://doi.org/10.1006/jsvi.1993.1340>.
- [4] M. Friswell, J.E. Mottershead, Finite element model updating in structural dynamics, Springer Science & Business Media, 2013.
- [5] M. Song, I. Behmanesh, B. Moaveni, C. Papadimitriou, Accounting for Modeling Errors and Inherent Structural Variability through a Hierarchical Bayesian Model Updating Approach: An Overview, Sensors. 20 (2020) 3874. <https://doi.org/10.3390/s20143874>.
- [6] G.C. Ballesteros, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian hierarchical models for uncertainty quantification in structural dynamics, in: Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management, 2014: pp. 1615–1624. <https://doi.org/10.1061/9780784413609.162>.

- [7] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, *Mechanical Systems and Signal Processing*. 56 (2015) 123–149. <https://doi.org/10.1016/j.ymssp.2014.11.001>.
- [8] B.P. Smarslok, D. Villanueva, G. Bartram, Design of multi-level validation experiments for multi-physics systems, 19th AIAA Non-Deterministic Approaches Conference, 2017. (2017) 1–10. <https://doi.org/10.2514/6.2017-1774>.
- [9] B.C. Jung, H. Yoon, H. Oh, G. Lee, M. Yoo, B.D. Youn, Y.C. Huh, Hierarchical model calibration for designing piezoelectric energy harvester in the presence of variability in material properties and geometry, *Structural and Multidisciplinary Optimization*. 53 (2016) 161–173. <https://doi.org/10.1007/s00158-015-1310-4>.
- [10] A.B. Nellippallil, J.K. Allen, P. Mohan, F. Mistree, Robust concept exploration of materials, products and associated manufacturing processes, *Proceedings of the ASME Design Engineering Technical Conference*. 2B-2018 (2018) 1–17. <https://doi.org/10.1115/DETC201885913>.
- [11] B.D. Youn, B.C. Jung, Z. Xi, S.B. Kim, W.R. Lee, A hierarchical framework for statistical model calibration in engineering product development, *Computer Methods in Applied Mechanics and Engineering*. 200 (2011) 1421–1431. <https://doi.org/10.1016/j.cma.2010.12.012>.
- [12] A. Urbina, S. Mahadevan, T.L. Paez, a Bayes Network Approach To Uncertainty Quantification in Hierarchically Developed Computational Models, *International Journal for Uncertainty Quantification*. 2 (2012) 173–193. <https://doi.org/10.1615/int.j.uncertaintyquantification.v2.i2.70>.
- [13] C. Fei, H. Liu, R. Patricia Liem, Y. Choy, L. Han, Hierarchical model updating strategy of complex assembled structures with uncorrelated dynamic modes, *Chinese Journal of Aeronautics*. (2021). <https://doi.org/10.1016/j.cja.2021.03.023>.
- [14] X. Jiang, S. Mahadevan, A. Urbina, Bayesian nonlinear structural equation modeling for hierarchical validation of dynamical systems, *Mechanical Systems and Signal Processing*. 24 (2010) 957–975. <https://doi.org/10.1016/j.ymssp.2009.10.002>.
- [15] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, *Mechanical Systems and Signal Processing*. 64 (2015) 360–376. <https://doi.org/10.1016/j.ymssp.2015.03.026>.
- [16] S. Wu, P. Angelikopoulos, G. Tauriello, C. Papadimitriou, P. Koumoutsakos, Fusing heterogeneous data for the calibration of molecular dynamics force fields using

- hierarchical Bayesian models, *The Journal of Chemical Physics*. 145 (2016) 244112. <https://doi.org/10.1063/1.4967956>.
- [17] S. Wu, P. Angelikopoulos, C. Papadimitriou, R. Moser, P. Koumoutsakos, A hierarchical Bayesian framework for force field selection in molecular dynamics simulations, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 374 (2016) 20150032. <https://doi.org/10.1098/rsta.2015.0032>.
- [18] X.-W. Liu, D.-G. Lu, P.C.J. Hoogenboom, Hierarchical Bayesian fatigue data analysis, *International Journal of Fatigue*. 100 (2017) 418–428. <https://doi.org/10.1016/j.ijfatigue.2017.03.043>.
- [19] J.B. Nagel, B. Sudret, A unified framework for multilevel uncertainty quantification in Bayesian inverse problems, *Probabilistic Engineering Mechanics*. 43 (2016) 68–84. <https://doi.org/10.1016/j.probengmech.2015.09.007>.
- [20] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, *Journal of Sound and Vibration*. 374 (2016) 92–110. <https://doi.org/10.1016/j.jsv.2016.03.022>.
- [21] S. Wu, P. Angelikopoulos, J.L. Beck, P. Koumoutsakos, Hierarchical Stochastic Model in Bayesian Inference for Engineering Applications: Theoretical Implications and Efficient Approximation, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*. 5 (2019). <https://doi.org/10.1115/1.4040571>.
- [22] M. Song, B. Moaveni, C. Papadimitriou, A. Stavridis, Accounting for amplitude of excitation in model updating through a hierarchical Bayesian approach: Application to a two-story reinforced concrete building, *Mechanical Systems and Signal Processing*. 123 (2019) 68–83. <https://doi.org/10.1016/j.ymssp.2018.12.049>.
- [23] D. Patsialis, A.P. Kyprioti, A.A. Taflanidis, Bayesian calibration of hysteretic reduced order structural models for earthquake engineering applications, *Engineering Structures*. 224 (2020) 111204. <https://doi.org/10.1016/j.engstruct.2020.111204>.
- [24] J.B. Nagel, N. Mojsilovic, B. Sudret, Bayesian assessment of the compressive strength of structural masonry, *12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP 2015*. (2015). <https://doi.org/10.14288/1.0076072>.
- [25] A. Economides, G. Arampatzis, D. Alexeev, S. Litvinov, L. Kulakova, C. Papadimitriou, P. Koumoutsakos, Hierarchical Bayesian Uncertainty Quantification

- for a Red Blood Cell Model, *Physical Review Applied*. 10 (2020) 1. <https://doi.org/10.1103/PhysRevApplied.15.034062>.
- [26] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Probabilistic hierarchical Bayesian framework for time-domain model updating and robust predictions, *Mechanical Systems and Signal Processing*. 123 (2019) 648–673. <https://doi.org/10.1016/j.ymssp.2018.09.041>.
- [27] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Data-driven uncertainty quantification and propagation in structural dynamics through a hierarchical Bayesian framework, *Probabilistic Engineering Mechanics*. 60 (2020) 103047. <https://doi.org/10.1016/j.probengmech.2020.103047>.
- [28] O. Sedehi, L.S. Katafygiotis, C. Papadimitriou, Hierarchical Bayesian operational modal analysis: Theory and computations, *Mechanical Systems and Signal Processing*. 140 (2020) 106663. <https://doi.org/10.1016/j.ymssp.2020.106663>.
- [29] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Hierarchical Bayesian Modeling Framework for Model Updating and Robust Predictions in Structural Dynamics using Modal Properties, *Mechanical Systems and Signal Processing*. Submitted. (2021). <https://zenodo.org/record/5078051#.YOV0LrUzbD4>.
- [30] X. Jia, O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, B. Moaveni, Nonlinear Model Updating through a Hierarchical Bayesian Modeling Framework, *Computer Methods in Applied Mechanics and Engineering*. Submitted. (2021). <https://doi.org/10.5281/zenodo.5520607>.
- [31] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, *Journal of Engineering Mechanics*. 133 (2007) 816–832. [https://doi.org/10.1061/\(ASCE\)0733-9399\(2007\)133:7\(816\)](https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816)).
- [32] S. Wu, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian annealed sequential importance sampling: an unbiased version of transitional Markov chain Monte Carlo, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*. 4 (2018). <https://doi.org/10.1115/1.4037450>.
- [33] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern classification*, John Wiley & Sons, 2012. <https://doi.org/10.1007/BF01237942>.
- [34] M. Coughlin, N. Christensen, J. Gair, S. Kandhasamy, E. Thrane, Method for estimation of gravitational-wave transient model parameters in frequency-time maps,

- Classical and Quantum Gravity. 31 (2014). <https://doi.org/10.1088/0264-9381/31/16/165012>.
- [35] M.S. Pedersen, B. Baxter, B. Templeton, C. Rishøj, D.L. Theobald, E. Hoegh-rasmussen, G. Casteel, J. Bin Gao, K. Dedecius, K. Strim, L. Christiansen, L.K. Hansen, L. Wilkinson, L. He, M. Bar, O. Winther, P. Sakov, S. Hattinger, K.B. Petersen, C. Rishøj, The Matrix Cookbook, Matrix. M (2008) 1–71. <https://doi.org/10.1111/j.1365-294X.2006.03161.x>.
- [36] P. A Bromiley, Products and Convolutions of Gaussian Probability Density Functions Density Functions, Tina Memo. (2003) No. 2003-003.

Chapter 5. Data Features-based Bayesian Learning for Time-domain Model Updating and Robust Predictions in Structural Dynamics

ABSTRACT

Bayesian inference has been demonstrated as a rigorous tool for updating models and predicting responses in structural dynamics. Most often, the likelihood function within the Bayesian framework is formulated based on a point-to-point probabilistic description of the discrepancy between the measurements and model predictions. This description results in an underestimation of uncertainties due to the inherent reduction of the parameter uncertainty as the number of data points increases. In this paper, the problem of estimating the uncertainty of parameters is re-visited using time-domain responses. Specifically, spatially and temporally uncorrelated/correlated prediction models are developed to re-formulate the likelihood function based on data features between the measurements and model predictions. Relation functions between the proposed probabilistic models and the likelihood-free approximate Bayesian computation (ABC) strategy are investigated, analytically demonstrating that the proposed data features models can offer reasonable and consistent uncertainties for the model parameters. A shear building model is employed to validate the effectiveness of the proposed approach. Results indicate that the proposed models provide reasonable parameter uncertainties as well as realistic uncertainty bounds of output quantities of interest (QoI) which is independent of the sampling rate used for a long time span response, in contrast to the classical Bayesian formulation which often severely underestimates the parameter uncertainties and provides thin and unrealistic uncertainty bounds of response predictions.

5.1 Introduction

Bayesian model updating has gained more interest because of its effectiveness in structural dynamical inverse problems [1–8]. In Bayesian model updating, the prior probability density function (PDF) of model parameters is updated to the posterior PDF by accounting for the information obtained from the measurements. Using probability models for the prediction errors, often formulated as the discrepancy between model predictions and the measurements, the likelihood function is developed [9]. Asymptotic approach [10] and sampling techniques [11] have been developed to solve the parameter inference problem. In particular, sampling methods include versions of Markov Chain Monte Carlo (MCMC), e.g. adaptive MCMC [12], sequential Monte Carlo Sampler [13] as well as Transitional MCMC [14,15]. For likelihood-free parameter inference, the approximate Bayesian computation (ABC) has been developed [16–18]. Among the algorithms proposed to solve the ABC, the subset simulation [17,19] is shown to be computational effective alternative.

Bayesian model updating in structural dynamics using response time histories measurements such as accelerations, displacements or strains is often formulated by introducing point-to-point probabilistic descriptions of the discrepancy between the measurements and model predictions [1,20,21]. Spatially and temporally uncorrelated prediction error models used to quantify these discrepancies, result in very peaked posterior probability distributions for the model parameters due to the large number of data points available from high sampling rates. Spatially and temporally correlated prediction error models are more reasonable for quantifying uncertainties [22,23]. However, the uncertainty depends on the correlation structure assumed which is often unknown and needs to be selected from a family of user-introduced correlation structures that might not be representative for the application. In general, the uncertainty quantified by the posterior probability distribution depends highly on the prediction error models and the correlation structure introduced between time instances as well as between measurements at different locations.

This paper revisits the problem of Bayesian learning given response time history measurements. It is expected that for sufficiently small sampling rate, the information contained in the response time histories is independent of the sampling rate used to represent the time histories. Conventional techniques fail to quantify such independence and also give unrealistically small uncertainties due to the large number of data points used to represent the time histories. To properly quantify such uncertainties, new formulations for likelihood-

informed Bayesian inference based on probability models introduced for the features between the measured data and model predictions are developed. Specifically, a probability model is assigned to the square of the discrepancy of the response time history between the measurement and the model prediction. Different probability models are investigated, such as truncated normal model and exponential distribution model. It is demonstrated that reasonable uncertainties are obtained for the model parameters that are independent of the sampling rate used to represent the response time histories. A relation between likelihood-informed and likelihood-free Bayesian computations is also established, demonstrating that both formulations yield reasonable and consistent uncertainties for the model parameters. A correlation prediction error model is also established, which can ensure the robustness of the correlation structural system.

The rest of this paper is organized as follows. Section 5.2 reviews the conventional Bayesian parameter estimation techniques. In Section 5.3, new likelihood-informed formulations for Bayesian model updating are proposed and compared with ABC formulation, a correlated prediction error model is also developed in this section. The effectiveness of the proposed method is demonstrated using a shear building model in Section 5.4. Section 5.5 reports the conclusions of this study.

5.2 Classical Formulation for Bayesian Model Updating

Consider a parameterized class of structure models $\mathbf{g}(\boldsymbol{\theta}; M)$, where M is the model, $\boldsymbol{\theta}$ is the set of model parameters which can be estimated using measurements D . Let $D = \{\hat{y}_j(k\Delta t) \in R^{N_0}, j = 1, 2, \dots, N_0; k = 1, 2, \dots, N\}$ be the measured response time histories data from the structure, where N_0 is the number of observed degrees of freedom (DOF) of the models, N is the number of the sampled data using a sampling rate Δt , j and k denote the j -th DOF and time index at time $k\Delta t$, respectively.

In Bayesian parameter estimation, the probability of unknown parameter sets $\boldsymbol{\theta}$ in the model class M can be first estimated from the prior probability density functions (PDF), and then updated based on the following Bayesian formula when measurements D are available:

$$p(\boldsymbol{\theta} | D, M) = c p(D | \boldsymbol{\theta}, M) p(\boldsymbol{\theta} | M) \quad (5.1)$$

where $p(\boldsymbol{\theta} | D, M)$ is the posterior PDF of the model parameters given the measurements D and the model class M , $p(\boldsymbol{\theta} | M)$ is the prior PDF, c is the constant that ensures the posterior PDF integrates to one and $p(D | \boldsymbol{\theta}, M)$ is the likelihood function of observing the data from

the model class. The prior PDF $p(\boldsymbol{\theta}|M)$ is a reflection of the probability of model parameters $\boldsymbol{\theta}$ before any measurements are available. Theoretically, it could be any types of distributions, while in practice it can be chosen in light of engineering experience. The likelihood function plays a key role in the process of parameter estimations. To construct the likelihood function, one needs to consider the relationship between the model predictions and measured data. Conventional methods for constructing the likelihood using direct response time history measurements are based on the prediction error equations formulated at time instant $t = k\Delta t$, given as follows:

$$\hat{y}(k) = L\mathbf{g}(k; \boldsymbol{\theta}, M) + \varepsilon(k; \boldsymbol{\theta}) \quad (5.2)$$

where $L \in R^{N_0 \times N_d}$ is the selective matrix consisting of zeroes and ones that maps the N_d model DOFs to the measured DOFs. A zero-mean Gaussian model is often assumed for the prediction errors $\varepsilon(k; \boldsymbol{\theta})$, $k = 1, 2, \dots, N$, with covariance matrix $\boldsymbol{\Sigma} \in R^{N_0 \times N_0}$ [1,20,24]. Given such assumption, the likelihood function takes the form:

$$p(D|\boldsymbol{\theta}, M) \propto |\boldsymbol{\Sigma}|^{-\frac{N}{2}} \exp\left\{-\frac{N_0 N}{2} J(\boldsymbol{\theta}; \boldsymbol{\Sigma}, D)\right\} \quad (5.3)$$

where

$$J(\boldsymbol{\theta}; \boldsymbol{\Sigma}, D) = \frac{1}{NN_0} \sum_{k=1}^N [\hat{y}(k) - L\mathbf{g}(k; \boldsymbol{\theta}, M)]^T \boldsymbol{\Sigma}^{-1} [\hat{y}(k) - L\mathbf{g}(k; \boldsymbol{\theta}, M)] \quad (5.4)$$

defines the average measure of fit between the model predictions and measurements. Specially, the prediction error model can consider both the uncorrelated and correlated cases. When the case of the uncorrelated prediction error model is assumed, the covariance matrix $\boldsymbol{\Sigma}$ takes the diagonal form $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$, where σ is the standard deviation of the prediction error to be determined by the Bayesian calibration, \mathbf{I} is the unit matrix. If a prediction error correlation model is considered, the correlation is taken into account by selecting a non-diagonal matrix. More details can be found in literatures [22,23]. Once the prior PDF and likelihood function are determined, the posterior PDF can be computed based on any types of Markov chain Monte Carlo (MCMC) algorithm, such as sequential MCMC [13], Transitional MCMC [14,15] and nested sampling [25], etc. Discussions on the sampling approaches are not involved in this paper. However, a tutorial on the sampling methods for solving the Bayesian parameter estimation problem can be found in literature [26].

It is notable that although the classical Bayesian formulation is widely used to update the models and estimate the model parameters, the outcome of classical formulation has been criticized for underestimating the uncertainties of model parameters. In particular for a long

time history response, information contained in the response is independent of the sampling rate used to represent the time histories. Conventional formulation using a Gaussian prediction error model may not consider such independence, leading to an underestimation with the sampling rate increasing. A recent likelihood-informed formulation developed for the case where the modal frequencies and mode shape components are available as the measured data [27]. This paper pushes forward the development of novel formulations for the likelihood function within the Bayesian framework using time histories measurements. Probabilistic prediction models based on data features between the model predictions and measurements are introduced into the formulation in order to provide reasonable uncertainty for the parameters that are independent of the sampling rate used for the time histories response. Details for the proposed prediction models are given below.

5.3 Data Features Models for Bayesian Model Updating and Robust Predictions

5.3.1 Uncorrelated data features models

5.3.1.1 Two likelihood formulations for uncorrelated case

In the proposed Bayesian model updating, new formulations for the likelihood functions are presented by introducing probabilistic models for the features between the measured data and the model predictions. For uncorrelated case, data feature e_j for the j -th DOF of the structure is assumed as the average of the square of the discrepancy between the measurements $\hat{y}_j(k)$ and the model predictions $g_j(k; \boldsymbol{\theta} | M)$, $k = 1, 2, \dots, N$, satisfying the following equation:

$$e_j = \frac{1}{N} \sum_{k=1}^N [\hat{y}_j(k) - g_j(k; \boldsymbol{\theta}, M)]^2 \quad (5.5)$$

It is noted that the notation e_j keeps positive value and thus the zero-mean Normal distribution is not well-suited for the realizations of data features. For this, truncated normal (TN) and exponential (EXP) distributions are introduced to further build the likelihood function within the Bayesian framework.

Given the TN distribution for data feature e_j , the PDF of each variable e_j can be written as [28]:

$$p(e_j) = \frac{\sqrt{2}}{\sqrt{\pi}\sigma} \exp\left(-\frac{e_j^2}{2\sigma^2}\right) \quad (5.6)$$

where σ is the standard deviation of the TN model to be estimated using the Bayesian inference. Given the assumption that data features between the DOFs are independent, the likelihood function can then be derived as:

$$p(\mathbf{e}|\mathbf{\theta}, M) = \prod_{j=1}^{N_0} p(e_j|\mathbf{\theta}, M) \propto \sigma^{-N_0} \exp\left\{-\frac{N_0}{2\sigma^2} J_{TN}(\mathbf{\theta}; \sigma^2; M)\right\} \quad (5.7)$$

where

$$J_{TN}(\mathbf{\theta}; M) = \frac{1}{NN_0} \sum_{j=1}^{N_0} \sum_{k=1}^N [\hat{y}_j(k) - g_j(k; \mathbf{\theta}, M)]^2 \quad (5.8)$$

defines the average measure of fit between the measurements and model predictions, stabilizing to a constant value as the number of data increases. It is noticed that the negative logarithmic of the likelihood function $L_{TN}(\mathbf{\theta}, \sigma^2)$ is more convenient to integrate into the sampling algorithm, defined as:

$$L_{TN}(\mathbf{\theta}, \sigma^2) = -\ln p(\mathbf{e}|\mathbf{\theta}, M) \propto N_0 \ln \sigma + \frac{N_0}{2\sigma^2} J_{TN}(\mathbf{\theta}; \sigma^2, M) \quad (5.9)$$

Given the EXP distribution for data feature e_j , the PDF of each variable e_j is given by:

$$p(e_j) = \begin{cases} \lambda \exp(-\lambda e_j) & e_j \geq 0 \\ 0 & e_j < 0 \end{cases} \quad (5.10)$$

where the parameter λ is reparameterized by $\lambda = \frac{1}{2\sigma^2}$, which can make the exponent term equal to that of the truncated normal distribution. Similarly, the negative logarithmic likelihood function $L_{EXP}(\mathbf{\theta}, \sigma^2)$ is calculated as:

$$L_{EXP}(\mathbf{\theta}, \sigma^2) = -\ln p(\mathbf{e}|\mathbf{\theta}, M) \propto 2N_0 \ln \sigma + \frac{N_0}{2\sigma^2} J_{EXP}(\mathbf{\theta}; M) \quad (5.11)$$

where $J_{EXP}(\mathbf{\theta}; M) = J_{TN}(\mathbf{\theta}; M)$.

When the prior PDF and the proposed likelihood function are available, the posterior PDF of the model parameters $\mathbf{\theta}$ can be solved according to Eq. (5.1). TMCMC algorithm is applied to estimate the parameters alongside their uncertainties as well as selecting the models which will be discussed in Section 5.3.3.

5.3.1.2 Asymptotic approximations of the posterior PDF using TN and EXP models

Based on the Bayesian central limit theorem, when a large number of measured data is available, the posterior PDF can be asymptotically approximated by the Gaussian distribution centered at the most probable value (MPV) of the model parameters $\hat{\boldsymbol{\theta}}$ and characterized by the covariance matrix $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$ which is the inverse of hessian matrix $\mathbf{H}(\boldsymbol{\theta})$ evaluated at the MPV $\hat{\boldsymbol{\theta}}$ [10,20]:

$$p(\boldsymbol{\theta} | \mathbf{e}, M) \approx N(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})) \quad (5.12)$$

where the MPV $\hat{\boldsymbol{\theta}}$ can be obtained by minimizing the negative logarithmic posterior PDF. Equivalently, when the prior PDF is assigned as a uniform distribution, the MPV can be computed by minimizing the negative logarithmic likelihood function. For the proposed TN and EXP models, the MPVs of the model parameters $\hat{\boldsymbol{\theta}}$ can be calculated through the following optimization problem:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \{J_{TN}(\boldsymbol{\theta}; M)\} = \arg \min_{\boldsymbol{\theta}} \{J_{EXP}(\boldsymbol{\theta}; M)\} \quad (5.13)$$

The covariance matrix $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$ can be also readily computed by solving the second derivatives of the negative logarithmic likelihood function with respect to the model parameters, which leads to:

$$\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) = \frac{2\sigma^2}{N_0} \mathbf{H}_J^{-1}(\hat{\boldsymbol{\theta}}) \quad (5.14)$$

where $\mathbf{H}_J(\hat{\boldsymbol{\theta}})$ defines the second derivatives of $J_{TN}(\boldsymbol{\theta}; M)$ with respect to $\boldsymbol{\theta}$ for TN model and the second derivatives of $J_{EXP}(\boldsymbol{\theta}; M)$ with respect to $\boldsymbol{\theta}$ for EXP model, also stabilizing to a constant value as the number of data increases.

Additionally, the most probable value (MPV) of the prediction error variance parameter $\hat{\sigma}^2$ can be also obtained by minimizing the function $L_{TN}(\boldsymbol{\theta}, \sigma^2)$ with respect to σ^2 for TN model, given as:

$$\hat{\sigma}^2 = J_{TN}(\hat{\boldsymbol{\theta}}; M) = \frac{1}{NN_0} \sum_{j=1}^{N_0} \sum_{k=1}^N [\hat{y}_j(k) - g_j(k; \hat{\boldsymbol{\theta}}, M)]^2 \quad (5.15)$$

and $L_{EXP}(\boldsymbol{\theta}, \sigma^2)$ with respect to σ^2 for EXP model, given as:

$$\hat{\sigma}^2 = \frac{1}{2} J_{EXP}(\hat{\boldsymbol{\theta}}; M) = \frac{1}{2NN_0} \sum_{j=1}^{N_0} \sum_{k=1}^N [\hat{y}_j(k) - g_j(k; \hat{\boldsymbol{\theta}}, M)]^2 \quad (5.16)$$

It is expected that the MPV of the variance parameter computed from TN model is larger than that from EXP model when given larger number of data points NN_0 . Also, such values will affect the uncertainty of model parameters as indicated from Eq. (5.14), showing that larger uncertainties can be obtained from TN model than that from EXP model. It is also noted that within a time history response, the uncertainty of the model parameters is independent of the number of data points N . In other words, the parameter uncertainties is independent of the sampling rate used for the response. Such features are important especially for the case where a long time span history response is available and a higher sampling rate will not get more information for the representation of the response, and thus the parameter uncertainties will not be changed as the sampling rate increases/decreases. Indeed, the parameter uncertainties decrease as the number of measured DOFs (sensor numbers) increases. However, increasing the number of sensors may not provide more information and thus keeping the same number of sensors is adequate and effective in some cases.

5.3.1.3 Relation between the proposed TN, EXP models and ABC

This sub-section provides the relation between the proposed models and the approximate Bayesian computation (ABC) strategy. In ABC algorithm, a summary statistics S and a tolerance level δ are introduced [18,29]:

$$\rho(S(X), S(D)) \leq \delta \quad (5.17)$$

where $X \in D$ denotes a simulated dataset from $p(\cdot | \theta, M)$, and ρ is a measure of the closeness between the model predictions and the measured data. Relation between the proposed models and ABC approach is investigated under the condition where the measure ρ is chosen to be the least square measure of the distance between the measurements and the model prediction from a parameterized class of structures models. Specifically for the model with predictions $\mathbf{g}(\theta; M)$, it is given as:

$$\rho = \frac{1}{N_0 N} \sum_{j=1}^{N_0} \sum_{k=1}^N [\hat{y}_j(k) - g_j(k; \theta, M)]^2 \quad (5.18)$$

Equivalently, the measure ρ can be written as:

$$\rho = J_{TN}(\theta; M) = 2J_{EXP}(\theta; M) \leq \delta \quad (5.19)$$

It is shown that the measure ρ fluctuates as the samples of model parameters changes, and thus the choice of the tolerance level δ would affect the quality and efficiency of the ABC algorithm. An excessive choice may result in an inaccurate estimate of the model parameters

or redundant iterations of the algorithm. Therefore, a suitable choice of the tolerance level δ is preferred in ABC algorithm.

It is noted that from Eqs. (5.15) and (5.16), when the best estimates of $J_{TN}(\hat{\theta}; M)$ and $J_{EXP}(\hat{\theta}; M)$ are achieved, the model parameters are chosen as their MPVs and meanwhile the best discrepancy between the model predictions and measurements yield to the MPV of the prediction error variance $\hat{\sigma}^2$. Such features provide benefits for the choice of the tolerance level δ , given as:

$$\delta = \begin{cases} J_{TN}(\hat{\theta}; M) = \hat{\sigma}^2 & \text{For TN model} \\ J_{EXP}(\hat{\theta}; M) = 2\hat{\sigma}^2 & \text{For EXP model} \end{cases} \quad (5.20)$$

It turns out that a tolerance level δ can be selected as the best estimate of prediction error variance $\hat{\sigma}^2$ computed from TN model or chosen as twice the value of $\hat{\sigma}^2$ obtained from EXP model. The proposed relation provides an effective way about how to choose a suitable tolerance value for ABC method. It also points out that, once a tolerance level δ is assigned based on such relations, TN and EXP models will offer similar uncertainties for the model parameters with ABC algorithm. It is also reminded that obtaining an accurate $\hat{\sigma}^2$ is under the condition where N_0 is sufficient. Limited number of sensors may slightly result in different results of $\hat{\sigma}^2$ and thus will affect the results of parameters uncertainties provided by ABC algorithm.

5.3.2 Correlated data features model

This section proposes a data features model for correlated case. Consider the presence of correlation between the adjacent data features e_j and e_{j+1} , the multi-dimensional truncated normal (MTN) distribution with zero-mean is assumed for data features $\mathbf{e} = (e_1, e_2, \dots, e_j, \dots, e_{N_0})$. The likelihood function $p(\mathbf{e} | \theta, M)$ can be then expressed as:

$$p(\mathbf{e} | \theta, M) = \frac{\exp\left(-\frac{1}{2} \mathbf{e}^T \mathbf{C}^{-1} \mathbf{e}\right)}{\int_0^{+\infty} \dots \int_0^{+\infty} \exp\left(-\frac{1}{2} \mathbf{e}^T \mathbf{C}^{-1} \mathbf{e}\right) d\mathbf{e}} \quad (5.21)$$

where the covariance matrix \mathbf{C} of total data features is given in the form:

$$\mathbf{C} = \left\{ C_{ij} = \sigma_i \sigma_j R(x_i - x_j) \mid i, j = 1, 2, \dots, N_0 \right\} \quad (5.22)$$

where $x_i - x_j$ is the spatial distance between the DOFs i and j . The function R is introduced as:

$$R(\bullet) = \exp\left(-\frac{|\bullet|}{\eta}\right) \quad (5.23)$$

where η is the measure of spatial correlation length. Given an extensive spatial distance, the value of C_{ij} equals to zero, while when it is close, $C_{ij} = \sigma_i \sigma_j$.

For simplifying the likelihood of Eq. (5.21), the multi-dimensional integration is solved as follows. Let $f = \mathbf{e}^T \mathbf{A} \mathbf{e}$, where $\mathbf{A} = \mathbf{C}^{-1}$, and introduce the function $\mathbf{e} = \mathbf{U} \mathbf{q}$. Thus, a matrix \mathbf{U} can be found that satisfies the form:

$$f = \mathbf{q}^T \mathbf{\Omega} \mathbf{q} \quad (5.24)$$

where $\mathbf{\Omega}$ is a diagonal matrix with the element ω_j ($j = 1, 2, \dots, N_0$). Subsequently, the integration in Eq. (5.21) can be transformed as:

$$\int_0^{+\infty} \dots \int_0^{+\infty} \exp\left(-\frac{1}{2} \sum_{j=1}^{N_0} \omega_j q_j^2\right) d\mathbf{U} \mathbf{q} = |\mathbf{U}| \prod_{j=1}^{N_0} \int_0^{+\infty} \exp\left(-\frac{1}{2} \omega_j q_j^2\right) dq_j = |\mathbf{U}| \prod_{j=1}^{N_0} \sqrt{\frac{\pi}{\omega_j}} \quad (5.25)$$

Therefore, Eq. (5.21) can be expressed as:

$$p(\mathbf{e} | \boldsymbol{\theta}, M) = |\mathbf{U}| \prod_{j=1}^{N_0} \sqrt{\frac{\pi}{\omega_j}} \exp\left(-\frac{1}{2} \mathbf{e}^T \mathbf{C}^{-1} \mathbf{e}\right) \quad (5.26)$$

After obtaining the likelihood and the prior PDF, the posterior PDF can be updated using TMCMC algorithm.

5.3.3 Model class selection

To further investigate the preference between the proposed models, model selection is discussed in this section. Bayesian inference technique allows to assess the plausibility of several model classes using the measured data. Let M_i define the i -th model class, $i = 1, 2, \dots, N_M$, where N_M represents the number of model classes. The posterior PDF of model class M_i is expressed using the Bayes' theorem in the form [30,31]:

$$p(M_i | D) = \frac{p(D | M_i) p(M_i)}{\sum_{i=1}^{N_M} p(D | M_i) p(M_i)} \quad (5.27)$$

where $p(M_i)$ is the prior PDF of model class M_i , $p(D | M_i)$ is the likelihood function of the model class M_i and also the evidence for the model class M_i provided by the

measurements D corresponding to the constant c in Eq. (5.1). When a uniform prior is assigned for the model class M_i , the posterior PDF is proportional to the evidence $p(D|M_i)$ for the model class M_i which can be estimated from the TMCMC algorithm. Therefore once the model parameters are identified based on the proposed models, the model classes can be ranked according to their evidence values.

5.3.4 Response predictions of output QoI

The calibrated model based on the proposed data features is then used to predict the output QoI. Let \mathbf{Q}_{unob} and \mathbf{Q}_{ob} be the unobserved and observed QoI, respectively. For the unobserved QoI \mathbf{Q}_{unob} , the uncertainties from the model parameters can be propagated through the calibrated model, while for the observed QoI \mathbf{Q}_{ob} , both the uncertainties from model parameters and prediction error can be considered. The PDF of the quantity \mathbf{Q}_{unob} given the data features can be computed based on the well-known total probability theory:

$$\begin{aligned} p(\mathbf{Q}_{\text{unob}} | \mathbf{e}) &= \int_{\boldsymbol{\theta}} p(\mathbf{Q}_{\text{unob}} | \boldsymbol{\theta}) p(\boldsymbol{\theta} | \mathbf{e}) d\boldsymbol{\theta} \\ &\approx \frac{1}{N_s} \sum_{i=1}^{N_s} p(\mathbf{Q}_{\text{unob}} | \boldsymbol{\theta}^{(i)}) \end{aligned} \quad (5.28)$$

where $\boldsymbol{\theta}^{(i)}$ is the i -th sample drawn from the posterior distribution $p(\boldsymbol{\theta} | \mathbf{e})$, N_s is the number of the samples. Similarly for observed QoI, the PDF of the quantity \mathbf{Q}_{ob} given the data features can be estimated in the form:

$$p(\mathbf{Q}_{\text{ob}} | \mathbf{e}) \approx \frac{1}{N_s} \sum_{i=1}^{N_s} p(\mathbf{Q}_{\text{ob}} | \boldsymbol{\theta}^{(i)}, \boldsymbol{\sigma}^{2,(i)}) \quad (5.29)$$

where $\boldsymbol{\sigma}^{2,(i)}$ can be obtained along with the samples of model parameters $\boldsymbol{\theta}^{(i)}$ through TMCMC algorithm. After the PDFs of \mathbf{Q}_{unob} and \mathbf{Q}_{ob} are available, the mean, standard deviation and quantiles, quantifying the response uncertainties can be estimated.

5.4 Illustrative Example

5.4.1 Description of a 10-story building model and measurements

A 10-story shear model of a building system is utilized to demonstrate the effectiveness of the proposed approach. The structure is shown in Fig. 5-1(a), where the system is fixed at the base. Fig. 5-1(b) depicts the parameterized model using 2 model parameters. Details for the

building model are given as follows. The equation of motion with base excitation $\ddot{\mathbf{y}}_g(t)$ of the structure takes the form:

$$\mathbf{M}\ddot{\mathbf{v}}(t) + \mathbf{C}\dot{\mathbf{v}}(t) + \mathbf{K}(\boldsymbol{\theta})\mathbf{v}(t) = -\mathbf{M}\mathbf{1}\ddot{\mathbf{y}}_g(t) \quad (5.30)$$

where $\mathbf{1} = [1, 1, \dots, 1]^T$ is a 10×1 vector. The system is created based on the following assumptions:

- The mass matrix \mathbf{M} is a diagonal matrix having the same element equaling to 1kg.
- The stiffness in each floor is assumed to be the same equaling to 1800N/m. The model is parameterized using two parameters $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$ associated with the stiffness, representing the ratio of the stiffness to its nominal values. Specifically, the parameter θ_1 is linked to the stiffness of the first three floors, while the second one corresponds to the stiffness from the fourth to sixth floors, as shown in Fig. 5-1(b). Given the nominal model $\boldsymbol{\theta} = (1, 1)^T$, the stiffness matrix \mathbf{K} is given as:

$$\mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 & \dots & 0 \\ -k_2 & k_2 + k_3 & -k_3 & 0 & 0 \\ 0 & -k_3 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & k_9 + k_{10} & -k_{10} \\ 0 & 0 & 0 & -k_{10} & k_{10} \end{bmatrix} \quad (5.31)$$

- Rayleigh damping is assumed for the damping matrix, written as

$$\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K} \quad (5.32)$$

where the coefficient α and β are taken to be 0.2265 and 6.7515×10^{-4} , respectively, corresponding to given damping ratios $\zeta_1 = \zeta_5 = 0.02$ for the first and fifth modes of the system.

- Given the above system properties, the natural frequencies of the first three modes are estimated to be 1.0Hz, 3.0Hz and 4.9Hz.

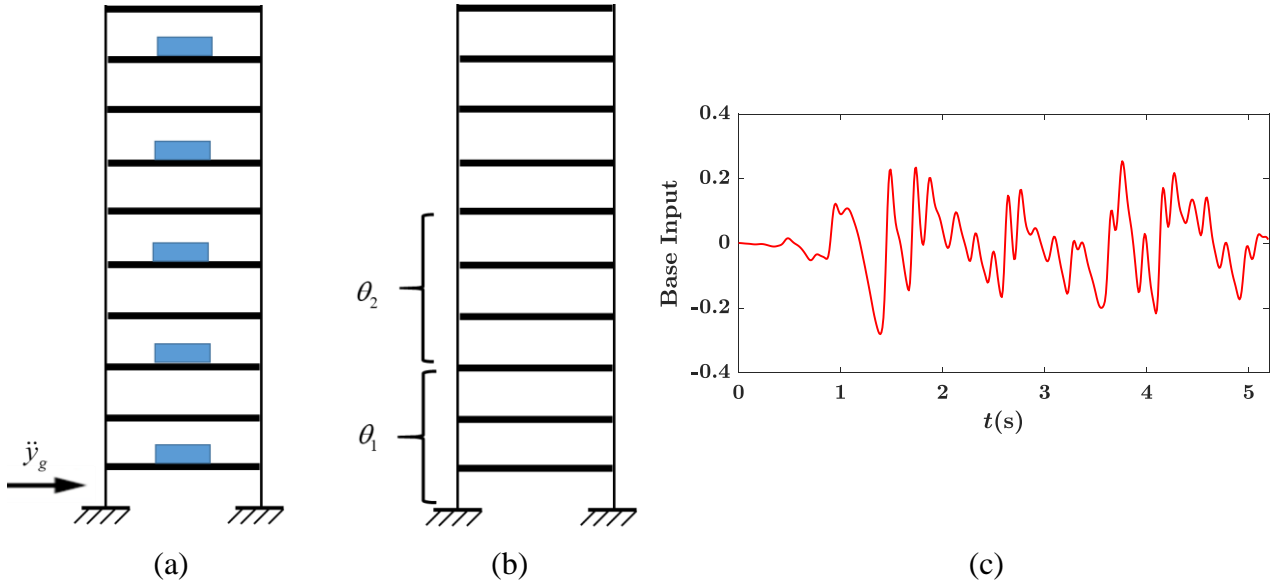


Fig. 5-1 (a) 10-story physical model (b) parameterized model (c) base excitation

For solving the equation of motion in Eq. (5.30), the transformation $\mathbf{v}(t) = \Phi \xi(t)$ is used for the expression with respect to the modal coordinates $\xi(t)$:

$$\ddot{\xi}(t) + \mathbf{C}^* \dot{\xi}(t) + \mathbf{\Omega} \xi(t) = -\Phi^T \mathbf{M} \ddot{y}_g(t) \quad (5.33)$$

where Φ is the mode shapes, \mathbf{C}^* and $\mathbf{\Omega}$ are two diagonal matrices with elements $2\zeta\omega_i$ and ω_i^2 , respectively, where ω_i defines the i -th modal frequency. The state-space form is next constructed:

$$\dot{\mathbf{x}}(t) = \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_c \mathbf{p}(t) \quad (5.34)$$

where \mathbf{x} is state vector, $\dot{\mathbf{x}}$ is first derivative of the state, \mathbf{A}_c is system state matrix and \mathbf{B}_c is the input to state matrix given as:

$$\mathbf{x}(t) = \begin{bmatrix} \xi(t) \\ \dot{\xi}(t) \end{bmatrix}, \quad \mathbf{A}_c = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{\Omega} & -\mathbf{C}^* \end{bmatrix}, \quad \mathbf{B}_c = \begin{bmatrix} \mathbf{0} \\ -\Phi^T \mathbf{M} \mathbf{I} \end{bmatrix}, \quad \mathbf{p}(t) = \ddot{y}_g(t) \quad (5.35)$$

The observation equation can also be written in the form:

$$\mathbf{d}(t) = \mathbf{G}_c \mathbf{x}(t) + \mathbf{J}_c \mathbf{p}(t) \quad (5.36)$$

where \mathbf{G}_c and \mathbf{J}_c are the observation that define the output quantities of the system. As an example, matrices \mathbf{G}_c and \mathbf{J}_c for acceleration measurements are defined as:

$$\mathbf{G}_c = \begin{bmatrix} -\mathbf{S}_a \Phi \mathbf{\Omega} & -\mathbf{S}_a \Phi \mathbf{C}^* \end{bmatrix}, \quad \mathbf{J}_c = \begin{bmatrix} \mathbf{S}_a (\mathbf{I} - \Phi \Phi^T \mathbf{M} \mathbf{I}) \end{bmatrix} \quad (5.37)$$

where \mathbf{S}_a is a selection matrix. For obtaining other output quantities, one can modify the matrices \mathbf{G}_c and \mathbf{J}_c .

An earthquake excitation $\ddot{y}_g(t)$ shown in Fig. 5-1(c) is applied to the physical structure to simulate the measurements, where the excitation corresponds to a sampling rate 100Hz for total 5.2 seconds. In order to add some modeling error between the physical and nominal structures, the nominal stiffness values of all the stories are polluted with 5% Gaussian white noise. Five accelerometers, denotes as blue cubic in Fig. 5-1(a), are placed at the first, third, fifth, seventh and ninth floors to collect the acceleration measurements, and therefore five sets of data consisting of 520 data points in each dataset are recorded.

5.4.2 Model updating and response predictions using proposed data features models

5.4.2.1 Parameter estimation

The acceleration measurements are used to estimate the associated stiffness parameters. For comparison purpose, the conventional Bayesian formulation using a Normal (NORM) distribution model is also applied to update the model parameters. TMCMC algorithm is employed to generate samples from the posterior distribution of the model and prediction error parameters. Specifically, the number of samples in the algorithm is chosen as 10000 and the tolerance value is selected as 0.5 for all the approaches. Details for the choice of such values are referred to [15]. The prior distributions of all the parameters are assumed to be uniform with bounds in the domain $[0.5, 1.5]$ for model parameters θ_1 and θ_2 and $[0.001, 0.5]$ for prediction error parameters σ .

Fig. 5-2 shows the posterior distributions along with the contour plots of the model and prediction error parameters using NORM, TN and EXP models. Table 5-1 also reports the mean, 5%, and 95% quantiles of samples from the posterior distributions using these models. It is seen that the mean of the model parameters computed by the proposed models are in good agreements with the ones calculated from the conventional formulation. This is expected from Eq. (5.13) where the most probable values of model parameters are found in a sense of optimization of the average measure of fit between the measurements and model predictions. However, a significant difference of the parameter uncertainties is observed between the proposed models and conventional formulation. It is obvious that narrow uncertainties of the model parameters are offered by the conventional formulation NORM, while wider uncertainties are obtained using the proposed models TN and EXP. This is due to the fact that the parameter uncertainties in NORM model depend on the number of data

points as opposite to the proposed models in which the uncertainties are independent of the number of data points within a dataset. For the results of prediction error parameters, it is found that the mean values from TN and EXP models show a difference from the one of NORM model. This is due to the limited number of sensors used in this case. However, the uncertainties computed from proposed models do contain the values calculated from NORM model.

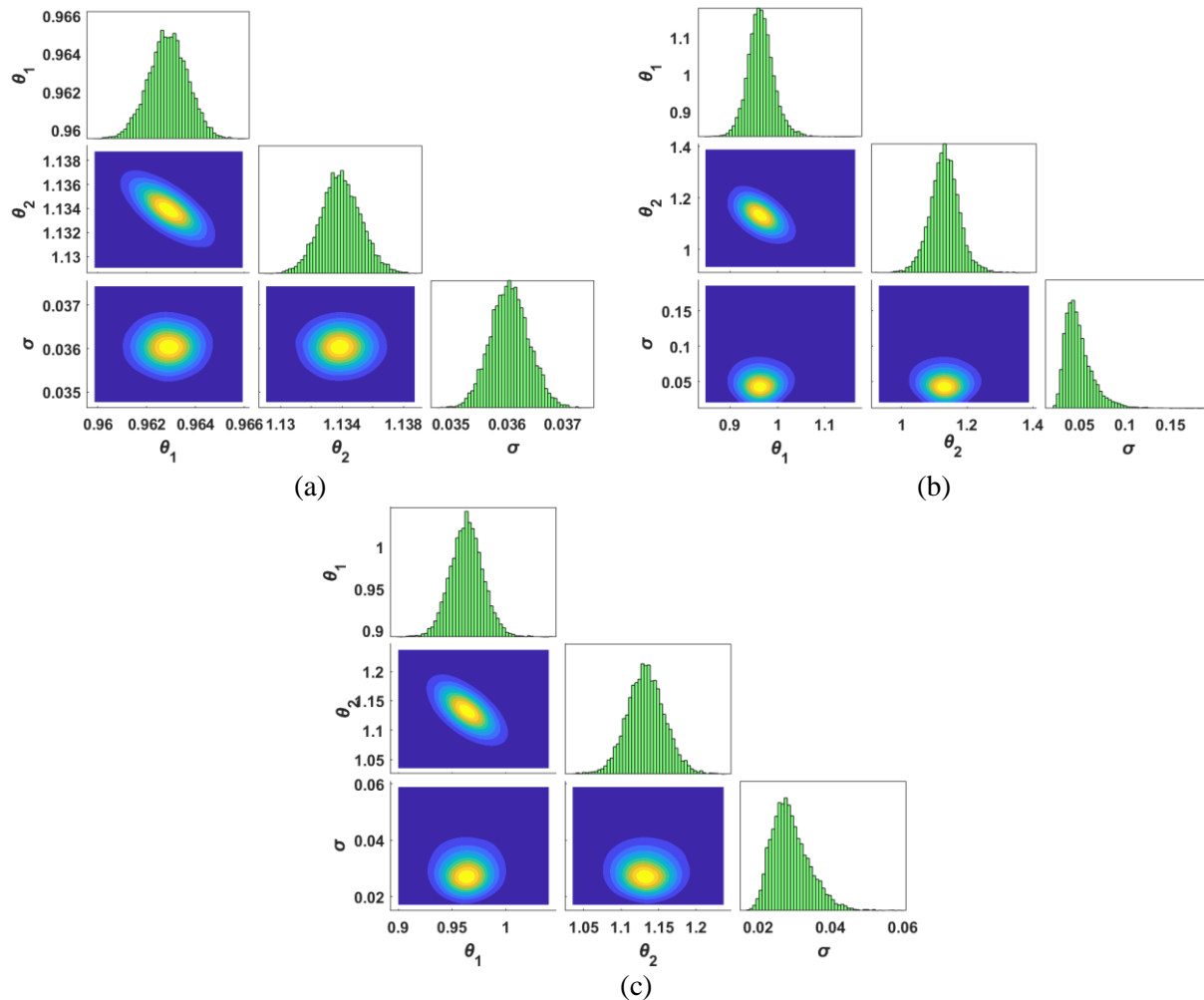


Fig. 5-2 Posterior distributions of the model and prediction error parameters using (a) NORM (b) TN (c) EXP models

Table 5-1 Mean and 5%, 95% quantiles of model and prediction error parameters using
NORM, TN and EXP models

	NORM			TN			EXP		
	θ_1	θ_2	σ	θ_1	θ_2	σ	θ_1	θ_2	σ
Mean	0.9629	1.1339	0.0361	0.9652	1.1321	0.0493	0.9638	1.1331	0.0291
5% Quantile	0.9615	1.1318	0.0355	0.9218	1.0588	0.0306	0.9391	1.0924	0.0220
95% Quantile	0.9642	1.1361	0.0366	1.0123	1.2049	0.0797	0.9891	1.1750	0.0389

To further investigate the differences between the TN, EXP and NORM models, different sampling rates are used to represent the same measured acceleration data. For this, sampling rates are chose from 20Hz to 1000Hz, representing the same information of the measurements. Such measurements are used to estimate the parameters using those models. Fig. 5-3 shows the parameter estimations results alongside their 90% confidence interval (CI). It can be found that the mean values of model parameters are independent of the sampling rates and all of the models provide good agreements on the estimates of the means of model parameters. However, the uncertainty bounds are substantially different for the NORM model and the proposed methods. Specifically, the conventional Bayesian method gives very small uncertainties that decrease as the number of sampling points increase. It is contrary to intuition since there is not extra information contained in the time history with higher sampling rate. Although the results obtained from the sampling rate 20Hz can have a relatively larger uncertainty, it is still a small value which is unrealistic in the engineering problems. For the results from TN and EXP models, it becomes evident that the different sampling rates chosen do not affect the information contained in the data. The proposed method based on the data features provides much higher uncertainties that are independent on the number of data points used. This is consistent with intuition since the information contained in the acceleration time history is almost independent of the sampling rate used in the measurements. It is noted that the parameter uncertainties using TN model are larger than the uncertainties computed from EXP model. This is expected from Eq. (5.14) since the parameter uncertainties depend also on the prediction error term. As indicated from Eqs. (5.15) and (5.16), the values of prediction error obtained from EXP are smaller than those from TN model. Results in Fig. 5-2(b) and Fig. 5-2(c) also report the same conclusion. A discussion of the model selection of EXP and TN models will be given later. It is also noticed that the uncertainty of prediction error is neglected herein since propagating uncertainties considers more on the prediction error itself rather than its uncertainty. In effect, uncertainty

of prediction error will also be decreased as the sampling rate increases using NORM model, while it will keep the constant by using TN and EXP models.

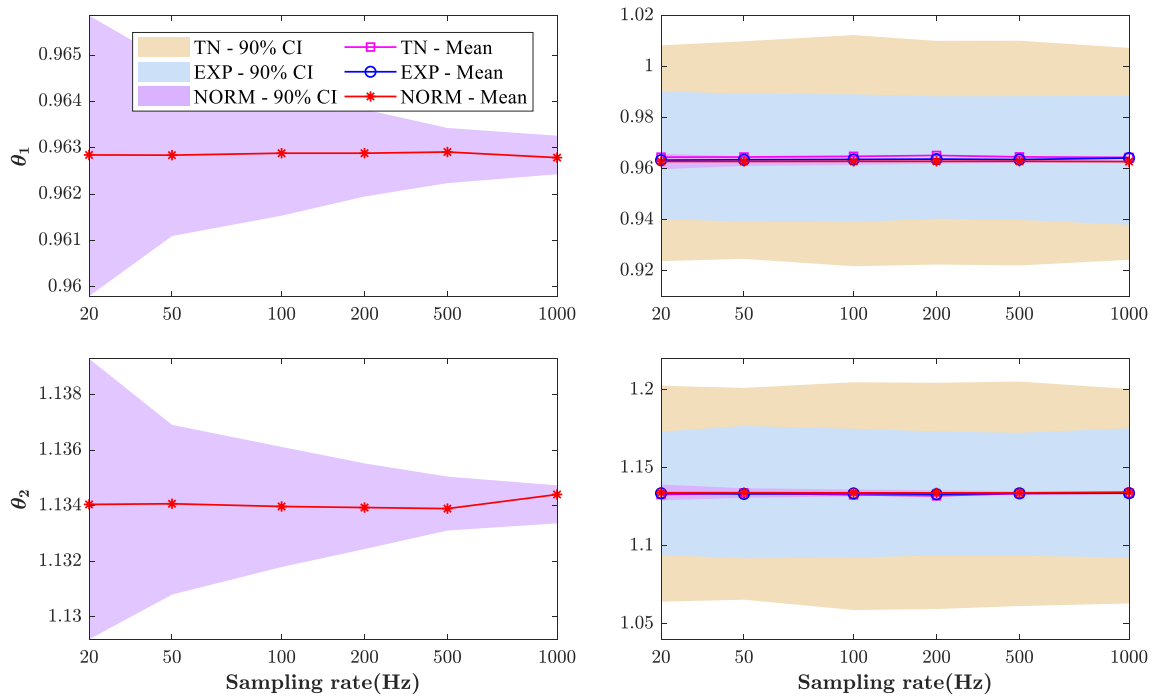


Fig. 5-3 Parameter estimation using different sampling rates of the measurements

Next, parameter estimation results along with their uncertainties are compared for the proposed TN model and ABC method. For ABC method, the measure is chosen to be the least square of the distance between the measurements and the model prediction, and the tolerance value δ is calculated through Eq. (5.20) based on the mean values of prediction error using TN model. Again the sampling rates are selected from 20 Hz to 1000Hz of the same measurements to conduct the parameter estimation results, as shown in Fig. 5-4. It is expected that the means of the model parameters computed from ABC are almost the same with the ones from TN model. Also, the parameter uncertainties predicted by the proposed likelihood-informed method is similar to the uncertainty estimated by the ABC method. Both methods (TN and ABC) provide uncertainty bounds that are independent on the sampling rate. It is seen that the parameter uncertainties from ABC are larger than that from TN. This could be due to the selection of the best estimate $\hat{\sigma}$. Attributed to the limited number of sensors, the best estimate $\hat{\sigma}$ differs with its mean values. This is also reported from Fig. 5-2(b). The discrepancy of the parameter uncertainties between ABC and TN models is expected to be decreased given the situation where the best estimate of prediction error is closer to its mean

value. For comparison purpose, the PDFs of model parameters using TN, EXP, ABC and NORM models are plotted, where the results correspond to the use of sampling rate 100Hz of the measurements. It is clear that peaked distribution of model parameters are provided by NORM, while the proposed TN and EXP models offer a larger uncertainty bounds similar to the results from ABC algorithm.

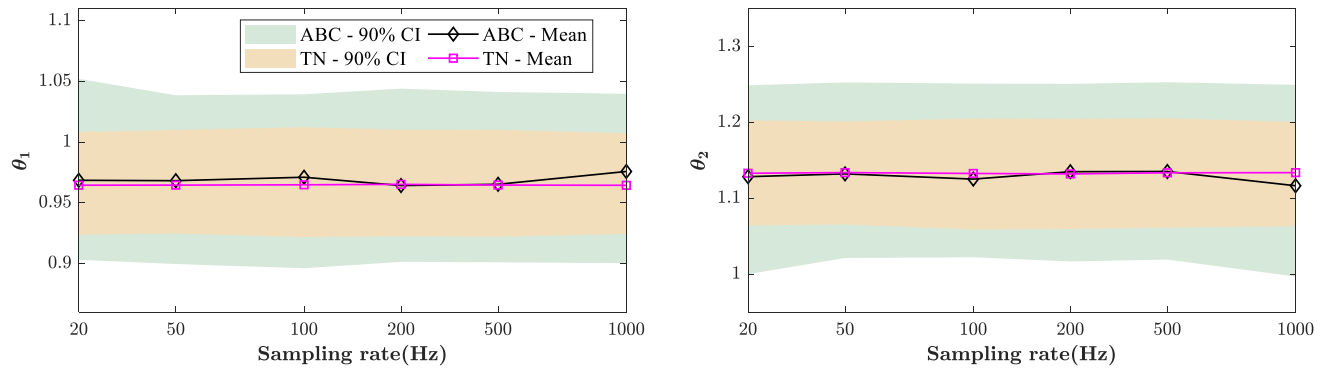


Fig. 5-4 Parameter estimation using TN and ABC models

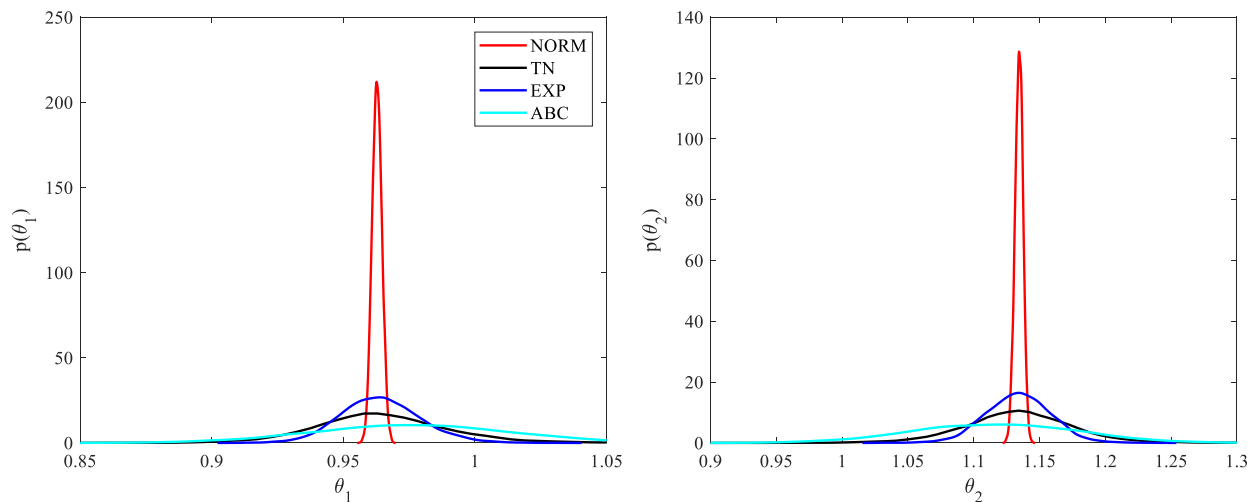


Fig. 5-5 PDFs of model parameters corresponding to the sampling rate 100Hz of measurements

The aforementioned results are based on the case of the absence of correlation. However, spatial correlation of the prediction error between one DOF to another DOF plays an important role for updating models. The proposed correlated data features model MTN is then applied to estimating the model and prediction error parameters. The spatial distance between the adjacent DOFs j and $j+1$ is equal to one meter, and the correlation length η is

defined as $\eta=1$. The standard deviation $\sigma_j = \sigma_{j+1}$ in the covariance matrix is assumed to be the same σ . The data features correlation model can be then constructed based on Eq. (5.21). For comparison purpose, the conventional Bayesian formulation assuming Normal distribution for the correlated prediction error is also studied, named as CONORM herein. The correlation properties (spatial distance and correlation length) in CONORM are assumed to be the same with those in MTN model.

TMCMC algorithm is also applied to sample the posterior distributions of the model and prediction error parameters. Sampling rates are taken from 20 Hz to 1000Hz of the same time history as well. Results of the parameter estimation are shown in Fig. 5-6. Similar to the case without correlation, the parameter uncertainties decrease as the sampling rates increase when using CONORM model, while when the proposed MTN is used, the parameter uncertainties keep stable with the sampling rates increasing, demonstrating that the proposed model provides result that is independent of the sampling rate. Also, the proposed MTN model can ensure larger uncertainty bounds for the model parameters, while the conventional one provides extremely small parameter uncertainties even lower sampling rate is used for the measurements. Interestingly, when the correlation is considered for the prediction error, the mean values obtained from MTN show a slightly difference with the ones from CONORM model, however, the mean values keep constant with the large number of data points contained in the time history.

For comparison between the correlated and uncorrelated cases, the posterior distributions of the model and prediction error parameters corresponding to the sampling rate 100Hz of the measurements are plotted in Fig. 5-7. Table 5-2 also reports the mean and 5%, 95% quantiles of all the parameters using the CONORM and MTN models. It is apparent that the model error reduces when the correlation is considered, especially for the proposed TN model. This also offers an explanation regarding the difference of the mean values between CONORM and MTN in Fig. 5-6. The model parameters calculated from MTN provides a better match between the model predictions and measurements than those provided by CONORM. Also, by considering the correlation of the prediction error in TN model, the uncertainties of the parameters are slightly smaller than those from the uncorrelated case. Rather, the parameter uncertainties produced by MTN are much wider than the ones yielded by CONORM.

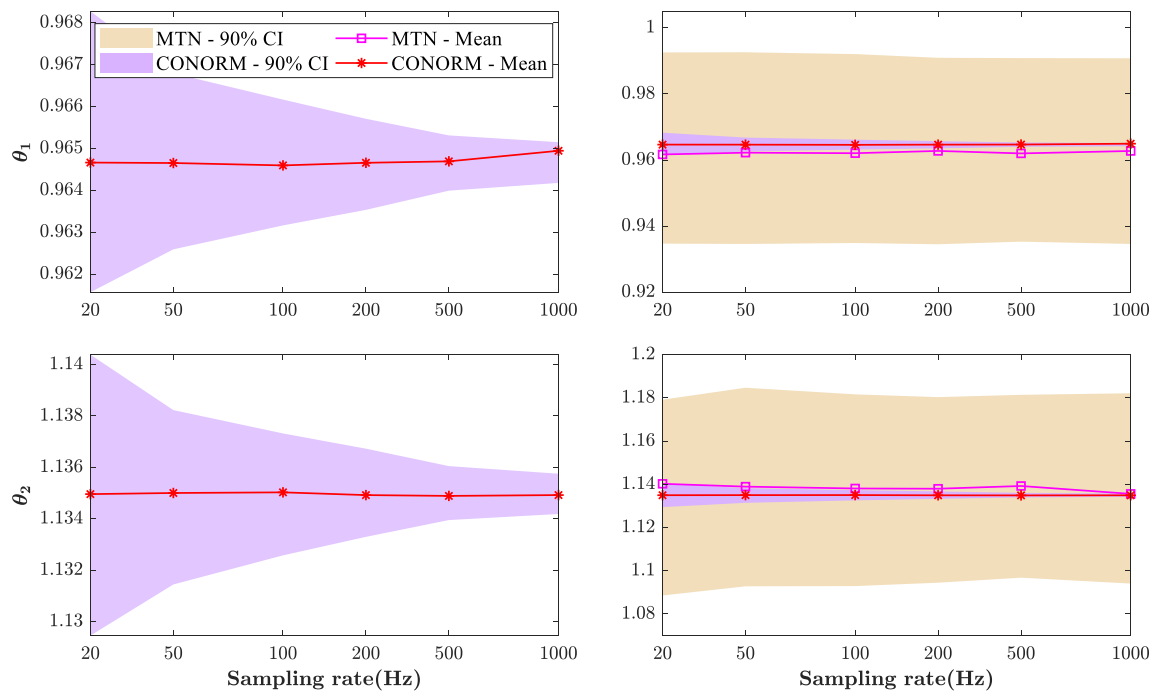


Fig. 5-6 Parameter estimation using CONORM and MTN

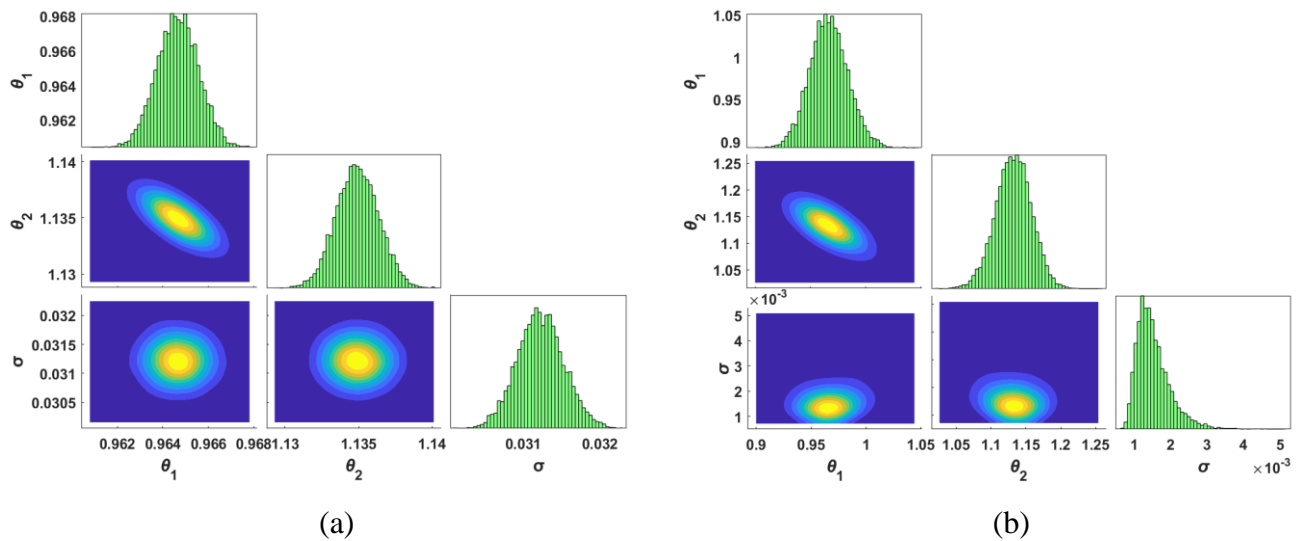


Fig. 5-7 Posterior distributions of model and prediction error parameters using (a) CONORM and (b) MTN models

Table 5-2 Mean and 5%, 95% quantiles of model and prediction error parameters using
CONORM and MTN models

	CONORM			MTN		
	θ_1	θ_2	σ	θ_1	θ_2	σ
Mean	0.9647	1.1350	0.0312	0.9628	1.1380	0.0016
5% Quantile	0.9631	1.1326	0.0307	0.9349	1.0929	0.0009
95% Quantile	0.9662	1.1373	0.0317	0.9921	1.1816	0.0025

5.4.2.2 Model selection

As discussed earlier, three data features models are proposed in this paper. Two data features models TN and EXP are presented for uncorrelated case, while the last one, MTN model, is developed for considering the correlation case of prediction error. In order to select a better model class, a model selection procedure is performed based on the theory in Section 5.3.3. TMCMC allows the procedure to obtain the log evidence directly and further to compare the model classes. Due to the fact that the random samples are chosen from TMCMC, it will lead to a slightly difference of every runs. Therefore, 10 times running of TMCMC is applied to all the methods for obtaining the log evidence values. For exploring if the sampling rates will have an influence on each model, aforementioned sampling rates are used for the measurements. Results of the log evidence values with different sampling rates are reported Table 5-3. It is evident that the correlation model MTN provides a better one than the uncorrelated model TN, indicating the importance of the correlation of prediction error. However, comparisons between MTN and EXP reveal that although MTN takes into account the correlation case, the proposed EXP model shows a better performance given the data in the structural model. Last but not least, it is obvious that the log evidence is independent on the sampling rate. In this case, the sampling rate will not affect the most probable models.

Table 5-3 Logarithm evidence values with different sampling rates

Sampling rate	20Hz	50Hz	100Hz	200Hz	500Hz	1000Hz
TN	15.579	15.736	15.494	15.692	15.609	15.902
MTN	42.418	41.462	42.025	41.485	40.450	43.140
EXP	51.293	51.509	51.408	51.257	51.200	51.580

5.4.2.3 *Response predictions*

The uncertainties computed from the model updating results are then considered for response predictions. Due to the fact the proposed EXP model provides the best one, samples of the parameters estimated from EXP model are then used for predicting the observed QoI and unobserved QoI. Specifically, samples of the posterior distribution in Fig. 5-2(c) are used to make predictions. For comparison purpose, samples in Fig. 5-2(a) obtained from NORM are also used to predict the QoI. Acceleration of the first story is considered as the observed QoI while displacement of the second story is selected as the unobserved QoI. Results for the observed acceleration are conducted considering the samples from both the model and prediction error parameters, as shown in Fig. 5-8. The 95% uncertainty bounds of predicted acceleration along with the mean predictions are depicted. The measured acceleration of the first story is also shown in the figure. It is observed that there is a discrepancy between the measurements and the mean prediction due to the presence of model error. It can be also seen that both NORM and EXP models can get larger uncertainty bounds for the observed QoI, where most parts of the measured acceleration falls within the predicted 95% uncertainty bounds. This is expected since the uncertainties of model parameters and prediction error are taken into account for the observed QoI. However, it should be reminded that the uncertainties from both NORM and EXP are formally different, with both the uncertainty from model parameters and prediction error in EXP model, while most of the uncertainty in NORM to be quantified in the prediction error term. Fig. 5-9 depicts the results for the predictions of displacement at the second floor. For this case, only the uncertainties of the model parameters are propagated since there is no information for the displacement of the measurements. However, a new displacement is also simulated from the physical structure which is only for comparison purpose. It is obvious that the NORM model obtains an extremely thin and unrealistic uncertainty bound for the predicted displacement, and indeed the measurements fall outside of the 95% uncertainty bounds. However, the result of uncertainty bound obtained from the proposed EXP model is relatively larger, and most of the response contains of the 95% uncertainty bounds. A slightly discrepancy could be improved by using a sufficient confidence interval of the predictions.

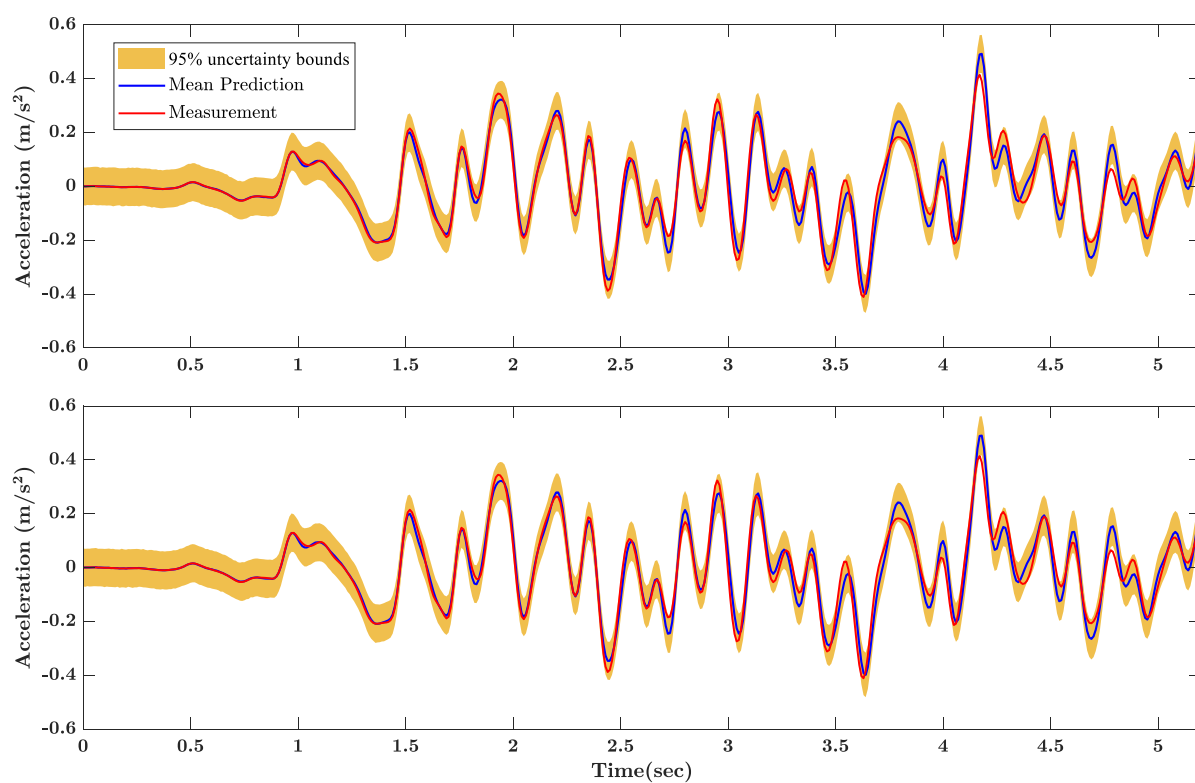


Fig. 5-8 Predictions of acceleration of the first story based on the results from NORM model (top) and from EXP model (bottom)

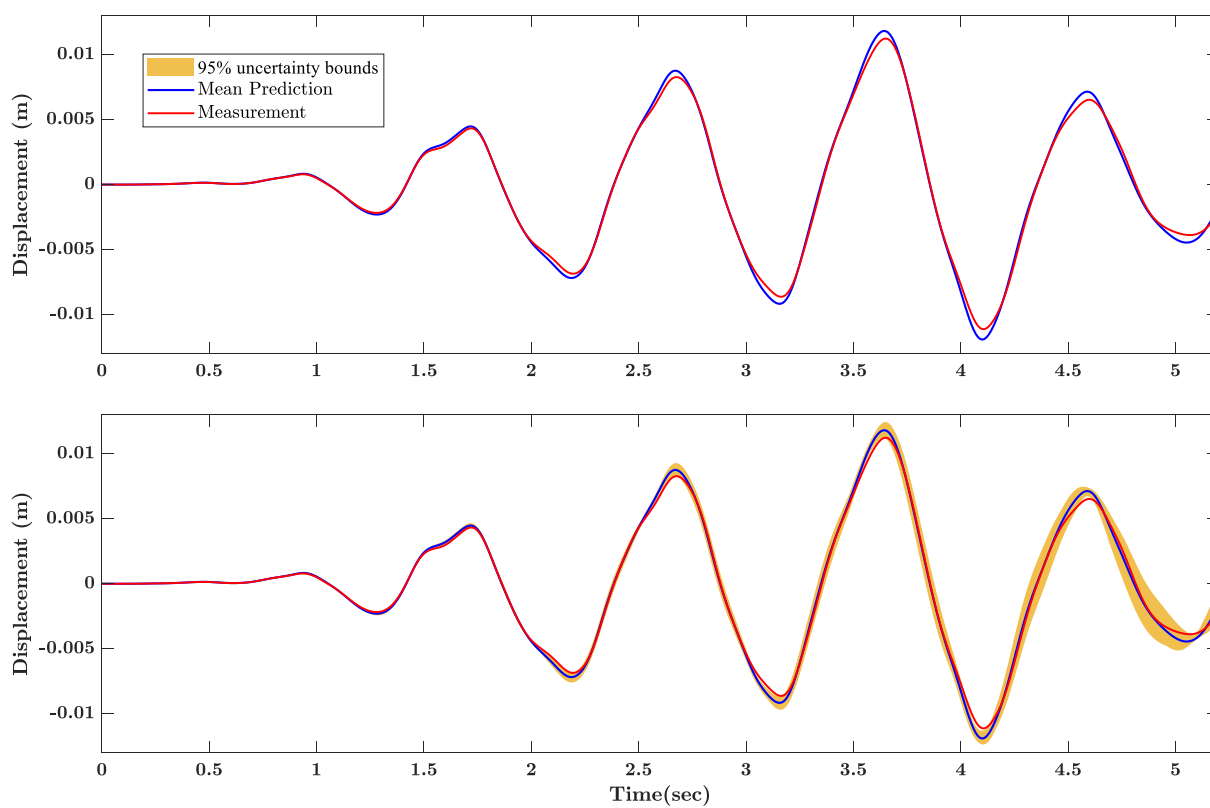


Fig. 5-9 Predictions of displacement of the second story based on the results from NORM model (top) and from EXP model (bottom)

5.5 Conclusion

In this paper, new formulations based on the data features models of prediction error has been proposed for likelihood-informed Bayesian inference. Three models are developed to take into account the uncorrelated and correlated cases of the prediction error model. Asymptotic approximations are presented to get some insights on the proposed models. Model selection procedure is also performed for selecting a better model over the proposed data features models. A 10-story building system is used to demonstrate the effectiveness of the proposed formulations. Comparisons between the conventional formulations and the proposed models are performed through the whole study. Main conclusions are drawn as follows.

- The proposed data-features likelihood-based Bayesian methodology correctly accounts for the uncertainty in the model parameters, making such uncertainty independent of sampling rate of the measured response time histories. In contrast, the uncertainty in the model parameters obtained from conventional Bayesian inference formulation depends on the sampling rate of the response time histories, despite the fact that the information contained in the response time history data is independent of the sampling rate.
- The proposed likelihood-informed Bayesian formulation provides results that are consistent with the ones obtained from likelihood-free ABC formulations.
- The proposed method can consider the correlation case, and it can ensure that the uncertainty is independent on the different sampling rates, which thus can provide an effective tool for the practical applications.
- The proposed method applied herein to linear structural systems but could be extended to non-linear structural systems given response time history measurements.

5.6 References

- [1] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, *J. Eng. Mech.* 124 (1998) 455–461.
- [2] M.W. Vanik, J.L. Beck, S.-K. Au, Bayesian probabilistic approach to structural health monitoring, *J. Eng. Mech.* 126 (2000) 738–745.

- [3] M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation, *JVC/Journal Vib. Control*. 14 (2008) 7–34.
- [4] E. Ntotsios, C. Papadimitriou, P. Panetsos, G. Karaiskos, K. Perros, P.C. Perdikaris, Bridge health monitoring system based on vibration measurements, *Bull. Earthq. Eng.* 7 (2009).
- [5] K.V. Yuen, *Bayesian Methods for Structural Dynamics and Civil Engineering*, 2010.
- [6] W.J. Yan, L.S. Katafygiotis, A novel Bayesian approach for structural model updating utilizing statistical modal information from multiple setups, *Struct. Saf.* 52 (2015) 260–271.
- [7] W.J. Yan, D. Chronopoulos, C. Papadimitriou, S. Cantero-Chinchilla, G.S. Zhu, Bayesian inference for damage identification based on analytical probabilistic model of scattering coefficient estimators and ultrafast wave scattering simulation scheme, *J. Sound Vib.* 468 (2020) 115083.
- [8] M. Song, L. Renson, J.P. Noël, B. Moaveni, G. Kerschen, Bayesian model updating of nonlinear systems using nonlinear normal modes, *Struct. Control Heal. Monit.* 25 (2018) 1–20.
- [9] J.L. Beck, Bayesian system identification based on probability logic, *Struct. Control Heal. Monit.* 17 (2010) 825–847.
- [10] C. Papadimitriou, J.L. Beck, L.S. Katafygiotis, Asymptotic expansions for reliability and moments of uncertain systems, *J. Eng. Mech.* 123 (1997) 1219–1229.
- [11] J.L. Beck, S.-K. Au, Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation, *J. Eng. Mech.* 128 (2002) 380–391.
- [12] S.K. Au, J.L. Beck, A new adaptive importance sampling scheme for reliability calculations, *Struct. Saf.* 21 (1999) 135–158.
- [13] J.S. Liu, R. Chen, Sequential monte carlo methods for dynamic systems, *J. Am. Stat. Assoc.* 93 (1998) 1032–1044.
- [14] J. Ching, Y.-C. Chen, Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, *J. Eng. Mech.* 133 (2007) 816–832.
- [15] S. Wu, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian annealed sequential importance sampling: an unbiased version of transitional Markov chain Monte Carlo, *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B Mech. Eng.* 4 (2018).

- [16] R.D. Wilkinson, Approximate Bayesian computation (ABC) gives exact results under the assumption of model error, *Stat. Appl. Genet. Mol. Biol.* 12 (2013) 129–141.
- [17] M. Chiachio, J.L. Beck, J. Chiachio, G. Rus, Approximate Bayesian computation by subset simulation, *SIAM J. Sci. Comput.* 36 (2014) A1339–A1358.
- [18] Z. Feng, Y. Lin, W. Wang, X. Hua, Z. Chen, Probabilistic Updating of Structural Models for Damage Assessment Using Approximate Bayesian Computation, *Sensors*. 20 (2020) 3197.
- [19] M.K. Vakilzadeh, Y. Huang, J.L. Beck, T. Abrahamsson, Approximate Bayesian Computation by Subset Simulation using hierarchical state-space models, *Mech. Syst. Signal Process.* 84 (2017) 2–20.
- [20] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, *Mech. Syst. Signal Process.* 56 (2015) 123–149.
- [21] I. Behmanesh, S. Yousefianmoghadam, A. Nozari, B. Moaveni, A. Stavridis, Uncertainty quantification and propagation in dynamic models using ambient vibration measurements, application to a 10-story building, *Mech. Syst. Signal Process.* 107 (2018) 502–514.
- [22] C. Papadimitriou, G. Lombaert, The effect of prediction error correlation on optimal sensor placement in structural dynamics, *Mech. Syst. Signal Process.* 28 (2012) 105–127.
- [23] E. Simoen, C. Papadimitriou, G. Lombaert, On prediction error correlation in Bayesian model updating, *J. Sound Vib.* 332 (2013) 4136–4152. <https://doi.org/10.1016/j.jsv.2013.03.019>.
- [24] P.E. Hadjidoukas, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Π4U: A high performance computing framework for Bayesian uncertainty quantification of complex models, *J. Comput. Phys.* 284 (2015).
- [25] M. Coughlin, N. Christensen, J. Gair, S. Kandhasamy, E. Thrane, Method for estimation of gravitational-wave transient model parameters in frequency-time maps, *Class. Quantum Gravity*. 31 (2014).
- [26] A. Lye, A. Cicirello, E. Patelli, Sampling methods for solving Bayesian model updating problems: A tutorial, *Mech. Syst. Signal Process.* 159 (2021) 107760.
- [27] C. Argyris, C. Papadimitriou, P. Panetsos, P. Tsopela, Bayesian model-updating using features of modal data: Application to the Metsovo bridge, *J. Sens. Actuator Networks*. 9 (2020).

- [28] S. Wilhelm, B.G. Manjunath, tmvtnorm: A Package for the Truncated Multivariate Normal Distribution Generation of random numbers computation of marginal densities, *R J.* 2 (2010) 25–29.
- [29] A. Ben Abdesslem, N. Dervilis, D. Wagg, K. Worden, Model selection and parameter estimation of dynamical systems using a novel variant of approximate Bayesian computation, *Mech. Syst. Signal Process.* 122 (2019) 364–386.
- [30] J.L. Beck, K.-V. Yuen, Model Selection Using Response Measurements: Bayesian Probabilistic Approach, *J. Eng. Mech.* 130 (2004) 192–203.
- [31] K.V. Yuen, Recent developments of Bayesian model class selection and applications in civil engineering, *Struct. Saf.* 32 (2010) 338–346.

Chapter 6. Statistics-based Bayesian Modeling Framework for Uncertainty Quantification and Propagation

Original Paper:

M. Ping, X. Jia, C. Papadimitriou, X. Han, C. Jiang, Statistics-based Bayesian modeling framework for uncertainty quantification and propagation, Mechanical Systems and Signal Processing. Submitted. (2021). <https://doi.org/10.5281/zenodo.5545922>.

ABSTRACT

A new Bayesian modeling framework is proposed to account for the uncertainty in the model parameters arising from model and measurements errors, as well as experimental, operational, environmental and manufacturing variabilities. Uncertainty is embedded in the model parameters using a single level hierarchy where the uncertainties are quantified by Normal distributions with the mean and the covariance treated as hyperparameters. Unlike existing hierarchical Bayesian modelling frameworks, the likelihood function for each observed quantity is built based on the Kullback–Leibler divergence used to quantify the discrepancy between the probability density functions (PDFs) of the model predictions and measurements. The likelihood function is constructed assuming that this discrepancy for each measured quantity follows a truncated normal distribution. For Gaussian PDFs of measurements and response predictions, the posterior PDF of the model parameters depends on the lower two moments of the respective PDFs. This representation of the posterior is also used for non-Gaussian PDFs of measurements and model predictions to approximate the uncertainty in the model parameters. The proposed framework can tackle the situation where only PDFs or statistical characteristics are available for measurements. The propagation of uncertainties is accomplished through sampling. Two applications demonstrate the use and effectiveness of the proposed framework. In the first one, structural model parameter inference is considered using simulated statistics for the modal frequencies and mode shapes. In the second one,

uncertainties in the parameters of the probabilistic $S-N$ curves used in fatigue are quantified based on experimental data.

6.1 Introduction

The general Bayesian statistical framework proposed by Beck and Katafygiotis [1] provides a rigorous mathematical means to address the model updating problem under uncertainty. Based on this framework, there have been a lot of works for various applications, such as parameter estimation [2–7], model selection [8,9], damage identification [10–12], and robust uncertainty propagation [13,14], among which the parameter estimation application serving as the foundation of other applications has kept the overwhelming attention. The Bayesian parameter inference is accomplished by embedding a parameterized probabilistic model to describe the discrepancy between model predictions and measurements. Then the formulation for the posterior distribution of the structural and prediction error model parameters is provided as a product of the likelihood function and the prior distribution of the model parameters. To estimate the posterior distribution of the model parameters, the likelihood function is usually built based on a relation function between model predictions and measurements by defining a probabilistic structure of the prediction error.

For industrial applications, the parameter estimation results from different measurements show distinct variations. The variation usually arises from load uncertainty, model error, measurement noise, and changing environmental/operational conditions [15–17]. Variations in the parameters of a model introduced to simulate a population of identically manufactured structures are also obtained due to manufacturing variabilities [18,19]. Therefore, it is important to describe these variations. The hierarchical Bayesian modeling framework (HBM) [20–26] has been proposed to quantify the uncertainty in the model parameters and prediction errors due to the aforementioned variabilities. The core of HBM is using a parameterized prior distribution of model parameters by introducing an extra layer involving hyper parameters to describe the variation of model parameters.

In this paper, a new probabilistic model is proposed on the basis of Bayesian framework, and the main difference is the principle to build the likelihood function, which is based on the relationship of statistics between model predictions and measurements for each model output. Compared with HBM, it does not require collecting datasets, and measurements can be the PDFs or statistics of measured quantities, so its application is more universal. It can also handle cases for which only the statistics like mean and variance are available from the

measurements. In such cases it is unreasonable to build datasets in terms of samples generated from them and then apply existing conventional or hierarchical Bayesian modeling frameworks.

This paper is organized as follow. In Section 6.2, the new proposed Bayesian modeling framework is described in detail, including construction of the proposed probabilistic model, uncertainty quantification of parameters, and uncertainty propagation to quantities of interest (QoI). The application to structural dynamics based on measured modal properties is presented in Section 6.3. In Section 6.4, a three-DOF spring mass chain system is taken as a simulated example to illustrate the effectiveness of the proposed framework. The application to the parameter inference and uncertainty quantification of probabilistic S - N curves used in fatigue damage accumulation is given in Section 6.5. Conclusions are presented in Section 6.6.

6.2 Proposed Bayesian Modeling Framework

6.2.1 Probabilistic model

Fig. 6-1 shows the structure of the proposed probabilistic modeling framework. Assume a parameterized model of a structural system and let $q_k(\boldsymbol{\theta})$, $k=1, \dots, n_q$ be the model predictions for n_q output quantities, where $\boldsymbol{\theta}$ is model parameter vector to be identified by measurements available for these output quantities. To account for model error and environmental/operational variabilities in the model predictions, an additive error term e_k is considered so that the predictions from the model are taken as

$$\tilde{q}_k(\boldsymbol{\theta}) = q_k(\boldsymbol{\theta}) + w_k e_k \quad (6.1)$$

where w_k is a weighting factor that scales the error terms e_k .

Uncertainties are embedded in the model parameter set $\boldsymbol{\theta}$ by assigning to the set $\boldsymbol{\theta}$ a Gaussian distribution with mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$. To account for the unmodelled dynamics, the error term e_k is assumed to follow a Gaussian distribution $N(e_k | 0, \sigma_e^2)$ with zero mean and variance σ_e^2 , where the general notation $N(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is introduced to denote a multivariable normal distribution evaluated at \mathbf{x} with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. It should be noted that the weights w_k introduced in Eq. (6.1) provides the flexibility to scale differently the error term e_k for each quantity of interest

$\tilde{q}_k(\boldsymbol{\theta})$. This is needed when the intensities of the measured or response quantities involved in Eq. (6.1) have considerable differences for each k . In this case it is best to choose w_k values to correspond to a measure of the intensity of the respective measured or response quantity, such as the mean of the measured or response quantity.

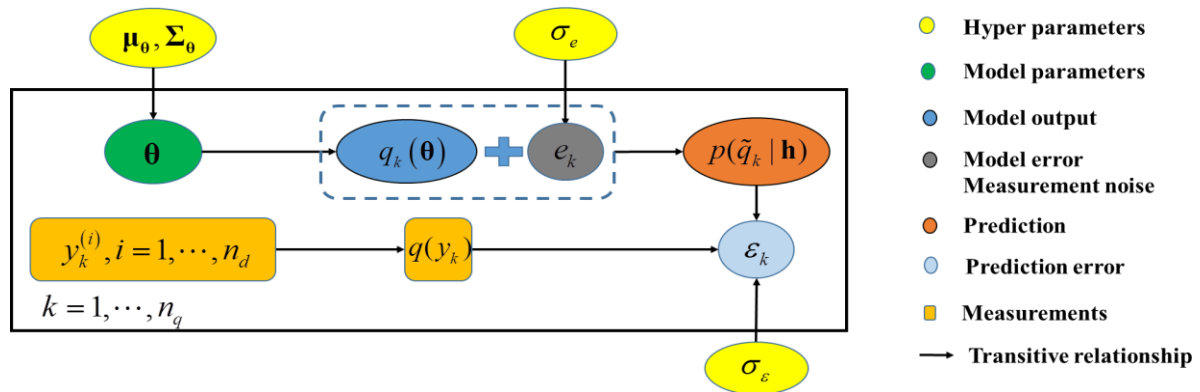


Fig. 6-1 Proposed probabilistic modeling framework

To distinguish from the model parameters $\boldsymbol{\theta}$ and the error term e_k , the parameters $\mathbf{h} = \{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e\}$ are called hyper parameters. Given the hyper parameters, one can obtain the conditional PDF $p(\tilde{q}_k | \mathbf{h})$ of k -th model output quantity \tilde{q}_k given the values of the hyper parameters \mathbf{h} by propagating the uncertainties in the model parameters $\boldsymbol{\theta}$ and the error term e_k .

For a conditional PDF that can be approximated by a Gaussian distribution, one has $p(\tilde{q}_k | \mathbf{h}) = \mathcal{N}(\tilde{q}_k | \mu_{\tilde{q}_k | \mathbf{h}}, \sigma_{\tilde{q}_k | \mathbf{h}}^2)$, where $\mu_{\tilde{q}_k | \mathbf{h}}$ is the mean and $\sigma_{\tilde{q}_k | \mathbf{h}}^2$ is the variance of \tilde{q}_k given \mathbf{h} . In particular, a Gaussian distribution arises for the case of a linear model $\mathbf{q}(\boldsymbol{\theta}) = \mathbf{A}\boldsymbol{\theta} + \mathbf{b}$, where $\mathbf{q} = [q_1, \dots, q_{n_q}]^T$, $\mathbf{A} = [\mathbf{a}_1^T, \dots, \mathbf{a}_{n_q}^T]^T$, $\mathbf{b} = [b_1, \dots, b_{n_q}]^T$ and \mathbf{a}_k , $k = 1, \dots, n_q$, is the k -th row (vector) of \mathbf{A} with the same dimension as $\boldsymbol{\theta}$. Given $\boldsymbol{\mu}_0$, $\boldsymbol{\Sigma}_0$ and σ_e , one readily derives that the mean $\mu_{\tilde{q}_k | \mathbf{h}}$ and the variance $\sigma_{\tilde{q}_k | \mathbf{h}}^2$ of the quantity \tilde{q}_k in terms of the hyperparameters \mathbf{h} as follows

$$\mu_{\tilde{q}_k | \mathbf{h}} = \mathbf{a}_k \boldsymbol{\mu}_0 + b_k \quad (6.2)$$

$$\sigma_{\tilde{q}_k | \mathbf{h}}^2 = \mathbf{a}_k \boldsymbol{\Sigma}_0 \mathbf{a}_k^T + w_k^2 \sigma_e^2 \quad (6.3)$$

The PDF of k -th model output $P_{\tilde{q}_k | \mathbf{h}}$ can be directly obtained in the form

$$p(\tilde{q}_k | \mathbf{h}) = N(\tilde{q}_k | \mu_{\tilde{q}_k | \mathbf{h}}, \sigma_{\tilde{q}_k | \mathbf{h}}^2) = N(\tilde{q}_k | \mathbf{a}_k \boldsymbol{\mu}_0 + b_k, \mathbf{a}_k \boldsymbol{\Sigma}_0 \mathbf{a}_k^T + w_k^2 \sigma_e^2) \quad (6.4)$$

For nonlinear models, the PDF $p(\tilde{q}_k | \mathbf{h}) = p(\tilde{q}_k | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e)$ of k -th model output can be derived by the total probability theorem as

$$\begin{aligned} p(\tilde{q}_k | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e) &= \int_{\boldsymbol{\theta}} p(\tilde{q}_k | \boldsymbol{\theta}, \sigma_e) N(\boldsymbol{\theta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) d\boldsymbol{\theta} \\ &\approx \frac{1}{n_s} \sum_j^{n_s} N(\tilde{q}_k | q_k(\boldsymbol{\theta}^{(j)}), w_k^{2(j)} \sigma_e^2) \end{aligned} \quad (6.5)$$

where $\boldsymbol{\theta}^{(j)}$, $j = 1, \dots, n_s$, are sampled from $N(\boldsymbol{\theta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. The structure of the model error in Eq. (6.1) was used to replace conditional PDF $p(\tilde{q}_k | \boldsymbol{\theta}, \sigma_e)$ in Eq. (6.5) by a normal PDF $N(\tilde{q}_k | q_k(\boldsymbol{\theta}), w_k^2 \sigma_e^2)$ with mean $q_k(\boldsymbol{\theta})$ and variance $w_k^2 \sigma_e^2$.

Let $\pi(y_k)$ denotes the PDF of k -th measured quantity y_k , so that the PDFs for all the measurements are $\{\pi(y_k), k = 1, \dots, n_d\}$. For the situation where the measurements are the statistics of measured quantities, the $\pi(y_k)$ can be directly obtained by moments-based PDF simulation methods, like maximum entropy method (MEM) [27,28]. For the situation where only measured data $D_k = \{y_k^{(i)}, i = 1, \dots, n_d\}$ are available for the k -th measured quantity y_k , where n_d denotes the number of measured data of k -th measured quantity, the kernel density estimation (KDE) [29,30] can be used to simulate $\pi(y_k)$ for sufficient large number of measurements n_d , as follows

$$\pi(y_k) = \frac{1}{n_d \times h} \sum_i^{n_d} K\left(\frac{y_k - y_k^{(i)}}{h}\right) \quad (6.6)$$

Also, if n_d is not big enough, $\pi(y_k)$ can be assumed to be a Gaussian distribution with mean and variance calculated by the data set D_k .

Given $\pi(y_k)$ and $p(\tilde{q}_k | \mathbf{h})$, the discrepancy between them is then quantified by the Kullback–Leibler divergence (KL-div) [31]. However, considering the asymmetry of KL-div, a symmetric measure of the discrepancy can be used, defined as

$$\begin{aligned} \varepsilon_k(y_k, \tilde{q}_k | \mathbf{h}) &= D_{KL}[\pi, p] + D_{KL}[p, \pi] \\ &= \frac{1}{2} \int_x \pi(x) \log\left(\frac{\pi(x)}{p(x | \mathbf{h})}\right) dx + \frac{1}{2} \int_x p(x | \mathbf{h}) \log\left(\frac{p(x | \mathbf{h})}{\pi(x)}\right) dx \end{aligned} \quad (6.7)$$

For a Gaussian distribution $p(\tilde{q}_k | \mathbf{h}) = \mathcal{N}(\tilde{q}_k | \mu_{\tilde{q}_k | \mathbf{h}}, \sigma_{\tilde{q}_k | \mathbf{h}}^2)$ of \tilde{q}_k given \mathbf{h} and for k -th measurement that is Gaussian distributed, i.e. $\pi(y_k) = \mathcal{N}(y_k | \mu_{y_k}, \sigma_{y_k}^2)$ for all k , the KL-div simplifies to the analytical form

$$\begin{aligned} \varepsilon_k(y_k, \tilde{q}_k | \mathbf{h}) &= \frac{1}{4} \left\{ \log \left[\frac{\sigma_{\tilde{q}_k | \mathbf{h}}^2}{\sigma_{y_k}^2} \right] + \frac{\sigma_{y_k}^2}{\sigma_{\tilde{q}_k | \mathbf{h}}^2} + \frac{[\mu_{y_k} - \mu_{\tilde{q}_k | \mathbf{h}}]^2}{\sigma_{\tilde{q}_k | \mathbf{h}}^2} - 1 \right\} \\ &\quad + \frac{1}{4} \left\{ \log \left[\frac{\sigma_{y_k}^2}{\sigma_{\tilde{q}_k | \mathbf{h}}^2} \right] + \frac{\sigma_{\tilde{q}_k | \mathbf{h}}^2}{\sigma_{y_k}^2} + \frac{[\mu_{y_k} - \mu_{\tilde{q}_k | \mathbf{h}}]^2}{\sigma_{y_k}^2} - 1 \right\} \\ &= \frac{1}{4} \left[\frac{\sigma_{\tilde{q}_k | \mathbf{h}}^2}{\sigma_{y_k}^2} - 1 \right] + \frac{1}{4} \left[\frac{\sigma_{y_k}^2}{\sigma_{\tilde{q}_k | \mathbf{h}}^2} - 1 \right] + \frac{1}{4} \frac{[\mu_{y_k} - \mu_{\tilde{q}_k | \mathbf{h}}]^2}{\sigma_{\tilde{q}_k | \mathbf{h}}^2} \left[1 + \frac{\sigma_{\tilde{q}_k | \mathbf{h}}^2}{\sigma_{y_k}^2} \right] \end{aligned} \quad (6.8)$$

Note that the first two terms give a measure of the error between the variance of the experimental value and the variance predicted from the model. These two terms become zero when the variance $\sigma_{y_k}^2$ of the experimental value equals the variance $\sigma_{\tilde{q}_k | \mathbf{h}}^2$ of the model prediction. Also, the last term gives the error between the mean of the experimental value and the mean of the model predictions. When the mean of the experimental value is equal to the mean of the model predictions, then the third term disappears. The discrepancy as defined by KL-div is a weighted sum of the discrepancies between the variances of the two PDFs and the means of the two PDFs. However, the KL-div measure is a rational method to assign the weights which otherwise one would have to select arbitrarily.

For non-Gaussian PDFs arising from nonlinear models, Eq. (6.8) for the KL-div can also be used as an approximate measure of the discrepancy between the two PDFs in terms of the first two moments of the PDFs. Alternatively, for nonlinear models, the integral can be approximated by Monte Carlo (MC) sample estimates

$$\varepsilon_k(y_k, \tilde{q}_k | \mathbf{h}) \approx \frac{1}{2n} \sum_i^n \log \left(\frac{\pi(y_k^{(i)})}{p(y_k^{(i)} | \mathbf{h})} \right) + \frac{1}{2n} \sum_i^n \log \left(\frac{p(\tilde{q}_k^{(i)})}{\pi(\tilde{q}_k^{(i)})} \right) \quad (6.9)$$

where $y_k^{(i)}$ and $\tilde{q}_k^{(i)}$, $i=1, \dots, n$, are the samples distributed as $\pi(y_k)$ and $p(\tilde{q}_k | \mathbf{h})$, respectively. Estimating the KL-div from Eq. (6.9) requires a large number of samples and can be a computationally very tedious procedure. Simplified approximations, such as Eq. (6.8), based on the first two moments of the non-Gaussian PDFs are preferred.

6.2.2 Estimation of hyper parameters uncertainty

The first task of uncertainty quantification is to identify the hyper parameters in the proposed probabilistic model. This is accomplished by introducing a probabilistic model to represent the variables $\varepsilon = \{\varepsilon_k, k=1, \dots, n_q\}$ quantifying the discrepancy between the PDF of the model predictions and the measurements. Specifically, the variables in the set $\varepsilon = \{\varepsilon_k, k=1, \dots, n_q\}$ are assumed to follow a truncated normal distribution $\text{TN}(\varepsilon_k | 0, \sigma_\varepsilon^2), \varepsilon_k \geq 0$ because of the non-negative values of KL-div, with the parameter σ_ε be another hyper parameter to be inferred from the data. Thus, in the proposed probabilistic model, the hyper parameters can be categorized into two types: one is to describe the uncertainty of the prediction model, which comprises $\mathbf{h} = \{\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \sigma_e\}$ and the other is σ_ε .

According to the relationship between measurements and model predictions described by Eq. (6.7), the Bayes theorem is applied to infer the posterior distribution of hyper parameters as

$$p(\mathbf{h}, \sigma_\varepsilon | \varepsilon) \propto p(\varepsilon | \mathbf{h}, \sigma_\varepsilon) p(\mathbf{h}) p(\sigma_\varepsilon) \quad (6.10)$$

where $p(\mathbf{h}, \sigma_\varepsilon | \varepsilon)$ is joint posterior distribution; $p(\mathbf{h})$ and $p(\sigma_\varepsilon)$ are the prior distributions assuming that \mathbf{h} and σ_ε are independent; and $p(\varepsilon | \mathbf{h}, \sigma_\varepsilon)$ is the likelihood function. Assuming that ε_k in $\varepsilon = \{\varepsilon_k, k=1, \dots, n_q\}$ are independent, the likelihood function takes the form

$$p(\varepsilon | \mathbf{h}, \sigma_\varepsilon) = \prod_{k=1}^{n_q} p(\varepsilon_k | \mathbf{h}, \sigma_\varepsilon) \quad (6.11)$$

Using the assumed truncated normal distribution for ε_k , one has

$$p(\varepsilon_k | \mathbf{h}, \sigma_\varepsilon) = \text{TN}(\varepsilon_k | 0, \sigma_\varepsilon^2) = \frac{2}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left(-\frac{\varepsilon_k^2(y_k, \tilde{q}_k | \mathbf{h})}{2\sigma_\varepsilon^2}\right) \quad (6.12)$$

with all $\varepsilon_k \geq 0$. Substituting Eq. (6.12) and (6.11) into Eq. (6.10), the posterior distribution takes the form

$$p(\mathbf{h}, \sigma_\varepsilon | \varepsilon) \propto \frac{1}{\sigma_\varepsilon^{n_q}} \exp\left(-\frac{n_q}{2\sigma_\varepsilon^2} J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})\right) p(\mathbf{h}) p(\sigma_\varepsilon) \quad (6.13)$$

where the notations $\{y_k\}$ and $\{\tilde{q}_k\}$ are defined as the sets $\{y_k\} = \{y_1, \dots, y_{n_q}\}$ and $\{\tilde{q}_k\} = \{\tilde{q}_1, \dots, \tilde{q}_{n_q}\}$, and

$$J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\}) = \frac{1}{n_q} \sum_{k=1}^{n_q} \varepsilon_k^2(y_k, \tilde{q}_k | \mathbf{h}) \quad (6.14)$$

is the mean square discrepancy function formed from the individual discrepancies for each measurement property. It should be noted that $J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})$ stabilizes to a finite value as the number of output measured quantities increases. Given the prior distribution, samples $\{\mathbf{h}^{(i)}, \sigma_\varepsilon^{(i)}; i=1, \dots, n\} = \{\boldsymbol{\mu}_0^{(i)}, \boldsymbol{\Sigma}_0^{(i)}, \sigma_e^{(i)}, \sigma_\varepsilon^{(i)}; i=1, \dots, n\}$ of hyper parameters distributed proportional to the joint posterior distribution can be generated by using any sampling algorithm. Herein, the nested sampling algorithm [32] is used to generate samples.

For a large number of output quantities n_q , with the prior distribution $p(\mathbf{h})$ selected to be uniform, Eq. (6.13) can be approximated as

$$p(\mathbf{h}, \sigma_\varepsilon | \varepsilon) \propto N(\mathbf{h} | \hat{\mathbf{h}}, \sigma_\varepsilon^2 \hat{\boldsymbol{\Sigma}}_{\mathbf{h}}) p(\sigma_\varepsilon) \quad (6.15)$$

where $\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} [J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})]$, $\hat{\boldsymbol{\Sigma}}_{\mathbf{h}} = \frac{1}{n_q} \mathbf{H}^{-1}(\hat{\mathbf{h}})$, $\mathbf{H}(\hat{\mathbf{h}}) = \frac{\partial^2 J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})}{\partial \mathbf{h}^T \partial \mathbf{h}} \Big|_{\mathbf{h}=\hat{\mathbf{h}}}$. The

derivation is given in Appendix A. Noting also that $\mathbf{H}(\hat{\mathbf{h}})$ stabilizes to a finite value as the number of output measured quantities increases, the uncertainty in the estimates of the hyperparameters \mathbf{h} quantified by $\hat{\boldsymbol{\Sigma}}_{\mathbf{h}}$ is inversely proportional to the number of measured variables n_q , which implies that the uncertainty decreases as the number of measured variables increases. For uniform prior PDF $p(\sigma_\varepsilon)$, it can be readily shown using Eq. (A.1) that the most probable value of the parameter σ_ε is given by $\hat{\sigma}_\varepsilon = \sqrt{J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})}$ which is an overall measure of discrepancy between the model predictions and the measurements.

Note that the estimated σ_ε can be used to evaluate the accuracy of the prediction model quantitatively. As ε_k , $k=1, \dots, n_q$, are assumed to follow a truncated Gaussian distribution with zero mean, σ_ε quantifies the average distance of all KL-div values from zero. So it can be treated as an index of prediction accuracy between the measured and model predicted PDFs. The smaller the value of σ_ε , the better the accuracy.

6.2.3 Uncertainty quantification of structural model parameters and error term

The posterior distribution of model parameters can be derived using the total probability theorem

$$p(\boldsymbol{\theta} | \varepsilon) = \int \int \int \int N(\boldsymbol{\theta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e, \sigma_\varepsilon | \varepsilon) d\boldsymbol{\Sigma}_0 d\boldsymbol{\mu}_0 d\sigma_e d\sigma_\varepsilon \quad (6.16)$$

where use was made of the fact that conditional PDF $p(\boldsymbol{\theta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e, \sigma_\varepsilon, \varepsilon) = N(\boldsymbol{\theta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ depends only on the values of $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$. Based on the samples of $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$, the integral can be approximated as

$$p(\boldsymbol{\theta} | \varepsilon) \approx \frac{1}{n} \sum_i^n N(\boldsymbol{\theta} | \boldsymbol{\mu}_0^{(i)}, \boldsymbol{\Sigma}_0^{(i)}) \quad (6.17)$$

Similarly, the posterior distribution of the error term is

$$p(e_k | \varepsilon) = \int \int \int \int N(e_k | 0, \sigma_e^2) p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e, \sigma_\varepsilon | \varepsilon) d\boldsymbol{\Sigma}_0 d\boldsymbol{\mu}_0 d\sigma_e d\sigma_\varepsilon \quad (6.18)$$

where use was made of the fact that $p(e_k | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \sigma_e, \sigma_\varepsilon, \varepsilon) = N(e_k | 0, \sigma_e^2)$ depends only on the values of σ_e^2 . The integral can be approximated by the following sample estimate

$$p(e_k | \varepsilon) \approx \frac{1}{n} \sum_i^n N(e_k | 0, [\sigma_e^{(i)}]^2) \quad (6.19)$$

6.2.4 Uncertainty propagation to output QoI

Let $z_j(\tilde{\mathbf{q}}(\boldsymbol{\theta}))$, $j=1, \dots, n_t$ be an output QoI that depends on the quantities $\tilde{\mathbf{q}}(\boldsymbol{\theta}) = \{\tilde{q}_1(\boldsymbol{\theta}), \dots, \tilde{q}_{n_q}(\boldsymbol{\theta})\}$. For example, $\tilde{\mathbf{q}}(\boldsymbol{\theta})$ can be related to the modal frequencies and mode shape components and $z_j(\tilde{\mathbf{q}}(\boldsymbol{\theta}))$, $j=1, \dots, n_t$ can be the response time histories (displacement, acceleration, strain, stresses) that are computed from the modal properties using modal analysis. Using samples $\boldsymbol{\theta}^{(l)}$ and $e_k^{(l)}$ generated from the PDFs defined in Eqs. (6.17) and (6.19), respectively, then the samples of $z_j(\tilde{\mathbf{q}}(\boldsymbol{\theta}))$ can be obtained as $z_j(\mathbf{q}(\boldsymbol{\theta}^{(l)}) + \mathbf{w}^{(l)} * \mathbf{e}^{(l)})$ from which the $\alpha\%$ to $1-\alpha\%$ quantiles of the response QoI can be estimated, where $\mathbf{w}^{(l)} * \mathbf{e}^{(l)}$ expresses element-wise product. Also the statistics of the response QoI, such as mean and higher moments, can be predicted by the following Monte Carlo estimates

$$\begin{aligned} E[z_j^m(\tilde{\mathbf{q}}(\boldsymbol{\theta}))] &= E[z_j^m(\mathbf{q}(\boldsymbol{\theta}) + \mathbf{w} * \mathbf{e})] \\ &= \int \int z_j^m(\mathbf{q}(\boldsymbol{\theta}) + \mathbf{w} * \mathbf{e}) p(\boldsymbol{\theta} | \varepsilon) p(\mathbf{e} | \varepsilon) d\boldsymbol{\theta} d\mathbf{e} \\ &\approx \frac{1}{n} \sum_{l=1}^n z_j^m(\mathbf{q}(\boldsymbol{\theta}^{(l)}) + \mathbf{w}^{(l)} * \mathbf{e}^{(l)}) \end{aligned} \quad (6.20)$$

The mean is given by $\mu_{z_j} = E[z_j(\mathbf{q}(\boldsymbol{\theta}) + \mathbf{w} * \mathbf{e})]$ and the variance by $\sigma_{z_j}^2 = E[z_j^2(\mathbf{q}(\boldsymbol{\theta}) + \mathbf{w} * \mathbf{e})] - \mu_{z_j}^2$. Note that for $z_j(\mathbf{q}(\boldsymbol{\theta}) + \mathbf{w} * \mathbf{e}) = q_j(\boldsymbol{\theta}) + w_j e_j$, $j = 1, \dots, n_q$, one obtains the predictions of the observed QoI.

6.3 Application to Structural Dynamics using Measured Modal Properties

For a linear system with n degrees of freedom (DOF), the PDF of measured modal properties are expressed as $\{\pi(\hat{f}_r), \pi(\hat{\phi}_{r,j}), r = 1, 2, \dots, R; j = 1, 2, \dots, n_0\}$, where \hat{f}_r represents the r -th modal frequency, $\hat{\phi}_r = [\hat{\phi}_{r,1}, \hat{\phi}_{r,2}, \dots, \hat{\phi}_{r,n_0}]$ is the r -th normalized mode shape vector at n_0 measured locations, and R is the number of contributed modes. Treating each modal property to be a model output, as there are R modal frequencies and $R \times n_0$ mode shapes, the number of model outputs is $R \times (n_0 + 1)$. Consider a model parameterized by $\boldsymbol{\theta}$, the model outputs corresponding to the measurements are $\{f_r(\boldsymbol{\theta}), \phi_r(\boldsymbol{\theta}), r = 1, 2, \dots, R\}$, and the predictions from the model that take into account model errors are defined as

$$\begin{aligned}\tilde{f}_r(\boldsymbol{\theta}) &= f_r(\boldsymbol{\theta}) + w_r e_r \\ \tilde{\phi}_{r,j}(\boldsymbol{\theta}) &= \phi_{r,j}(\boldsymbol{\theta}) + w_{r,j} e_{r,j}\end{aligned}\tag{6.21}$$

where e_r and $e_{r,j}$ are assigned to follow an identical Gaussian distribution $N(0, \sigma_e^2)$. To take into account the different intensities of the modal frequencies and the mode shape components, the weight factors w_r and $w_{r,j}$ are respectively selected to be the mean of \hat{f}_r and $\hat{\phi}_{r,j}$.

Based on the measurements and probabilistic model described above, the proposed probabilistic modeling framework can be implemented. The posterior distribution of hyperparameters is given by Eq. (6.13) where the discrepancy function in Eq. (6.14) becomes

$$J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\}) = \frac{1}{R(n_0 + 1)} \left[\sum_{r=1}^R \varepsilon_r^2(\hat{f}_r, \tilde{f}_r | \mathbf{h}) + \sum_{r=1}^R \sum_{j=1}^{n_0} \varepsilon_{r,j}^2(\hat{\phi}_{r,j}, \tilde{\phi}_{r,j} | \mathbf{h}) \right]\tag{6.22}$$

and $n_q = R(n_0 + 1)$. Approximating the PDFs $\pi(\hat{f}_r)$ and $\pi(\hat{\phi}_{r,j})$ by Gaussian distributions, the formula of KL-div can be simplified using Eq. (6.8), so that the expressions $\varepsilon_r(\{\hat{f}_r\}, \{\tilde{f}_r\} | \mathbf{h})$ and $\varepsilon_{r,j}(\{\hat{\phi}_{r,j}\}, \{\tilde{\phi}_{r,j}\} | \mathbf{h})$ depend on the mean and variances of the model predictions conditional on the values of the hyperparameters \mathbf{h} , given by

$$\begin{aligned}
\mu_{\tilde{f}_r|\mathbf{h}} &= \int_{\mathbf{h}} [f_r(\boldsymbol{\theta}) + w_r e_r] p(\boldsymbol{\theta}|\mathbf{h}) p(e_r|\mathbf{h}) d\mathbf{h} = \int_{\mathbf{h}} f_r(\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathbf{h}) d\mathbf{h} = \mu_{f_r|\mathbf{h}} \\
\sigma_{\tilde{f}_r|\mathbf{h}}^2 &= \int_{\mathbf{h}} [f_r(\boldsymbol{\theta}) + w_r e_r - \mu_{\tilde{f}_r|\mathbf{h}}]^2 p(\boldsymbol{\theta}|\mathbf{h}) p(e_r|\mathbf{h}) d\mathbf{h} \\
&= \int_{\mathbf{h}} [f_r(\boldsymbol{\theta}) - \mu_{f_r|\mathbf{h}}]^2 p(\boldsymbol{\theta}|\mathbf{h}) d\mathbf{h} + w_r^2 \sigma_e^2 \\
\mu_{\tilde{\varphi}_{r,j}|\mathbf{h}} &= \int_{\mathbf{h}} [\varphi_{r,j}(\boldsymbol{\theta}) + w_{r,j} e_{r,j}] p(\boldsymbol{\theta}|\mathbf{h}) p(e_{r,j}|\mathbf{h}) d\mathbf{h} = \int_{\mathbf{h}} \varphi_{r,j}(\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathbf{h}) d\mathbf{h} = \mu_{\varphi_{r,j}|\mathbf{h}} \\
\sigma_{\tilde{\varphi}_{r,j}|\mathbf{h}}^2 &= \int_{\mathbf{h}} [\varphi_{r,j}(\boldsymbol{\theta}) + w_{r,j} e_{r,j} - \mu_{\tilde{\varphi}_{r,j}|\mathbf{h}}]^2 p(\boldsymbol{\theta}|\mathbf{h}) p(e_{r,j}|\mathbf{h}) d\mathbf{h} \\
&= \int_{\mathbf{h}} [\varphi_{r,j}(\boldsymbol{\theta}) - \mu_{\varphi_{r,j}|\mathbf{h}}]^2 p(\boldsymbol{\theta}|\mathbf{h}) d\mathbf{h} + w_{r,j}^2 \sigma_e^2
\end{aligned} \tag{6.23}$$

To calculate the multi-dimensional integrals efficiently, the univariate dimension reduction method (UDRM) [33] is introduced to approximate them. The UDRM involves an additive decomposition of a multivariate response function into multiple univariate functions, so the multi-dimensional integral required by response moments are approximated by a series of one-dimensional integral of these univariate functions. According to UDRM, $f_r(\boldsymbol{\theta})$ and $\varphi_{r,j}(\boldsymbol{\theta})$ can be approximated as

$$\begin{aligned}
f_r(\boldsymbol{\theta}) &\approx -(n-1)f_r(\boldsymbol{\mu}_{\boldsymbol{\theta}}) + \sum_{i=1}^n f_r(\theta_i, \boldsymbol{\mu}_{\boldsymbol{\theta}_{-i}}) \\
\varphi_{r,j}(\boldsymbol{\theta}) &\approx -(n-1)\varphi_{r,j}(\boldsymbol{\mu}_{\boldsymbol{\theta}}) + \sum_{i=1}^n \varphi_{r,j}(\theta_i, \boldsymbol{\mu}_{\boldsymbol{\theta}_{-i}})
\end{aligned} \tag{6.24}$$

where $\boldsymbol{\mu}_{\boldsymbol{\theta}_{-i}}$ represents the mean vector of $\boldsymbol{\theta} \in \mathbb{R}^n$ that excludes the component μ_{θ_i} , and $f_r(\theta_i, \boldsymbol{\mu}_{\boldsymbol{\theta}_{-i}})$ is defined as $f_r(\theta_i, \boldsymbol{\mu}_{\boldsymbol{\theta}_{-i}}) = f_r(\mu_{\theta_1}, \mu_{\theta_2}, \dots, \mu_{\theta_{i-1}}, \theta_i, \mu_{\theta_{i+1}}, \dots, \mu_{\theta_n})$. Substituting Eq. (6.24) into Eq. (6.23), the mean and variance can be approximated into a series of one-dimensional integral of univariate functions in Eq. (6.24) [33], which can be easily solved by numerical integral methods, hereby a lot of computation can be saved.

6.4 Simulated Example

A population of 3-DOF linear systems manufactured to be identical is taken as a simulated example. Due to manufacturing variabilities, the properties of the system, such as stiffness vary for each member in the population. The modal properties of the members of the population are chosen as measured quantities to study the effectiveness of the proposed probabilistic modeling framework.

6.4.1 Model description

Consider a three DOF model, shown in Fig. 6-2, introduced to represent each member in the population. The nominal values of the mass for the model DOFs are set to be $m_1 = 6\text{kg}$, $m_2 = 5\text{kg}$ and $m_3 = 5\text{kg}$, while the nominal values of the stiffness of each link are set to $k_{10} = 22\text{kN/m}$, $k_{20} = 21\text{kN/m}$ and $k_{30} = 20\text{kN/m}$. With these assumptions, the modal properties of the nominal model are listed in Table 6-1.

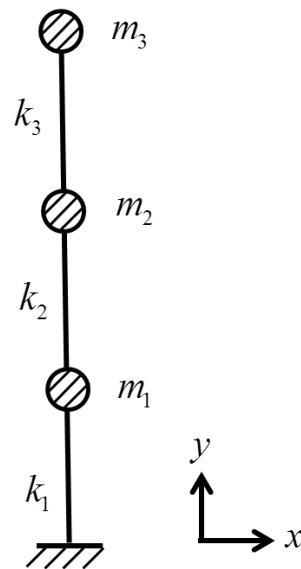


Fig. 6-2 3-DOF spring mass chain system

Table 6-1 Modal properties

	Frequency (Hz)	Mode shape at DOF 1	Mode shape at DOF 2	Mode shape at DOF 3
Mode 1	4.59	0.324	0.587	0.741
Mode 2	12.2	0.740	0.281	-0.611
Mode 3	17.9	-0.498	0.787	-0.362

6.4.2 Uncertainty quantification

The properties of each member in the population are simulated as follows. To consider the variation of model parameters from member to member due to manufacturing variability, the stiffness of the link i of each member is generated from a Gaussian distribution

$k_i(x) \sim N\left(x | k_{i0}, (k_{i0} * 0.03)^2\right)$, $i = 1, 2, 3$, corresponding to approximately 3% variation of the

stiffness parameters about their nominal values. The modal properties are then simulated from Eq. (6.21) using 10^3 MCs samples. The error terms defined in Eq. (6.21) are assumed to follow the Normal distribution with zero mean and standard deviation equals to 0.05, corresponding to a 5% model error. The mean and variance of modal properties are then computed and listed in Table 6-2 serving as known statistics of measurements. Then the uncertainty quantification for the stiffness parameters can be conducted according to the methodology presented in Section 6.3.

Table 6-2 Mean and variance of simulated modal properties

	Frequency		Mode shape at DOF 1		Mode shape at DOF 2		Mode shape at DOF 3	
	mean	variance	mean	variance	mean	variance	mean	variance
Mode 1	4.59	0.015	0.324	3.1×10^{-4}	0.587	8.8×10^{-4}	0.741	1.4×10^{-3}
Mode 2	12.2	0.106	0.740	1.5×10^{-3}	0.281	5.4×10^{-4}	-0.611	9.8×10^{-4}
Mode 3	17.9	0.233	-0.498	8.3×10^{-4}	0.787	1.6×10^{-3}	-0.362	5.2×10^{-4}

The 3-DOF model shown in Fig. 6-2 is used to represent each member in the group. To take into account the variation in the model properties, the stiffness of each link is parameterized by $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \theta_3]^T$ such as $k_i(\theta_i) = k_{i0} * \theta_i$, where $\boldsymbol{\theta}$ are model parameters to be identified. Assuming the prior distributions of all the hyper parameters to be uniform distributions with their upper and lower boundaries listed in Table 6-3, the nested sampling algorithm [32] is implemented to generate samples of hyper parameters. The results are shown in Fig. 6-3, while the mean and standard deviation of the posterior distributions of hyper parameters are summarized in Table 6-4. As expected, the estimated mean values of all the hyper parameters $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\Sigma}_\theta$ (defined to be diagonal), as well as the prediction error parameter σ_ϵ , show good agreement with the nominal values used to simulate the measurements. The standard deviations of the hyperparameters are small, which means the uncertainty of identified results is small. Also, the values of the samples for σ_ϵ are very small, which means the prediction results have a good accuracy compared with measurements. The closeness of the results to the values used to simulate the measurements

also demonstrates that the Gaussian approximation of modal properties assumed in Section 6.3 is reasonable.

Table 6-3 Upper and lower boundaries of hyper parameters

	μ_{θ_1}	σ_{θ_1}	μ_{θ_2}	σ_{θ_2}	μ_{θ_3}	σ_{θ_3}	σ_e	σ_ε
Upper boundary	0.8	0	0.8	0	0.8	0	0	0
Lower boundary	1.2	0.1	1.2	0.1	1.2	0.1	0.1	1

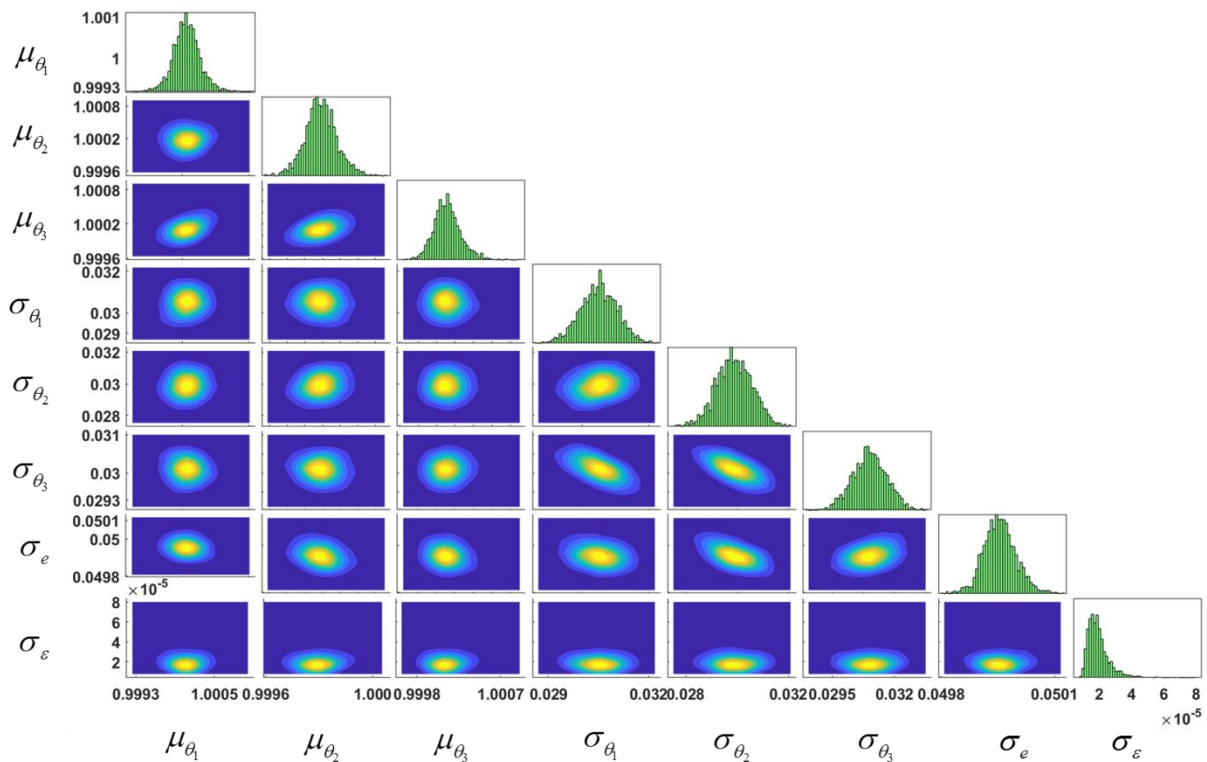


Fig. 6-3 Samples of joint posterior distribution of hyper parameters

Table 6-4 Estimates of mean and standard deviation of hyper parameters

	μ_{θ_1}	μ_{θ_2}	μ_{θ_3}	σ_{θ_1}	σ_{θ_2}	σ_{θ_3}	σ_e
Nominal value	1	1	1	0.03	0.03	0.03	0.05
Mean	1.0001	1.0002	1.0001	0.0305	0.0299	0.0301	0.0500
Standard deviation	2.1×10^{-4}	2.2×10^{-4}	2.1×10^{-4}	6×10^{-4}	7×10^{-4}	3×10^{-4}	5×10^{-5}

Based on Eq. (6.16) and (6.18), the posterior distributions of $\boldsymbol{\theta}$ and $e_r, e_{r,j}$ are computed using sampling and shown in Fig. 6-4. As $e_r, e_{r,j}$ are identically distributed, there is only one figure for their posterior distribution. The estimated statistics of $\boldsymbol{\theta}$ and e are summarized in Table 6-5. The values are compared with the nominal values assigned to simulate the measurements. It can be seen that the mean and standard deviation of posterior distributions of model parameters and error terms are very close to the nominal values. The samples of ε are close to zeros, which indicates that the discrepancy between the PDFs of measurements and predictions is small enough.

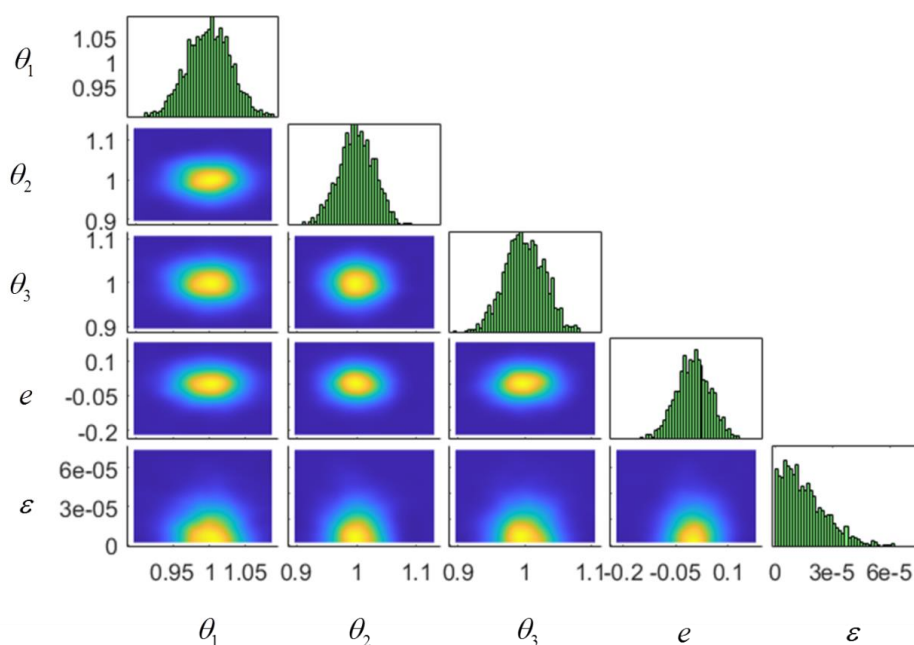


Fig. 6-4 Samples of posterior distribution of $\boldsymbol{\theta}$, $e = \{e_r, e_{r,j};, r = 1, 2, \dots, R; , j = 1, 2, \dots, n_0\}$ and ε

Table 6-5 Estimates of mean and standard deviation of model parameters and error terms

	θ_1	θ_2	θ_3	e
Mean	1.0010	1.0010	1.0000	-0.0012
Standard deviation	0.0307	0.0293	0.0301	0.0494

6.4.3 Uncertainty propagation

Using the samples of θ , e_r and $e_{r,j}$, the mean and variance of modal properties are predicted and listed in Table 6-6. These values should be compared with the mean and variance of the measurements in Table 6-2. The predicted results are of good accuracy, and the maximum relative error is less than 4%. The predicted PDFs of modal properties are also computed and compared with Gaussian PDFs of the measurements in Fig. 6-5. A very good agreement is also observed, validating the effectiveness of the proposed methodology in identifying the model parameters. It should be noted that such uncertainty bounds are expected to be thin for classical Bayesian framework based on multiple datasets used for the modal properties [34]. The level of uncertainty is expected in classical Bayesian approaches to decrease as the number of data increases.

Table 6-6 Predicted mean and variance of the modal properties

	Frequency		Mode shape at DOF 1		Mode shape at DOF 2		Mode shape at DOF 3	
	mean	variance	mean	variance	mean	variance	mean	variance
Mode 1	4.59	0.015	0.324	3.1×10^{-4}	0.587	8.8×10^{-4}	0.743	1.4×10^{-3}
Mode 2	12.2	0.107	0.741	1.5×10^{-3}	0.280	5.7×10^{-4}	-0.611	9.7×10^{-4}
Mode 3	17.9	0.225	-0.498	8.4×10^{-4}	0.787	1.6×10^{-3}	-0.363	5.0×10^{-4}

Furthermore, based on the identified modal properties, the response time history of displacement or acceleration or velocity can also be predicted. For this, a zero mean discrete Gaussian white noise base excitation with standard deviation 1, shown in Fig. 6-6, is applied and the mean and variance of time history response of displacement of the third DOF is estimated taking into account the uncertainties in the model parameters and error terms assumed to simulate the measurements in Section 6.4.2. The modal analysis is used to perform the corresponding predictions based on the predicted modal properties with error terms taken into consideration. To make a comparison, the mean and 90% credible interval boundaries obtained respectively from measurements and predictions are shown in Fig. 6-7. As we can see, the model prediction results match very well the measurements. For the situation that the error terms of modal properties are not considered in the predictions the predicted uncertainty intervals are smaller than that of measurement uncertainty intervals, as

shown in Fig. 6-8, signifying that the error terms are necessary to be included in the propagation analysis for accurate model predictions.

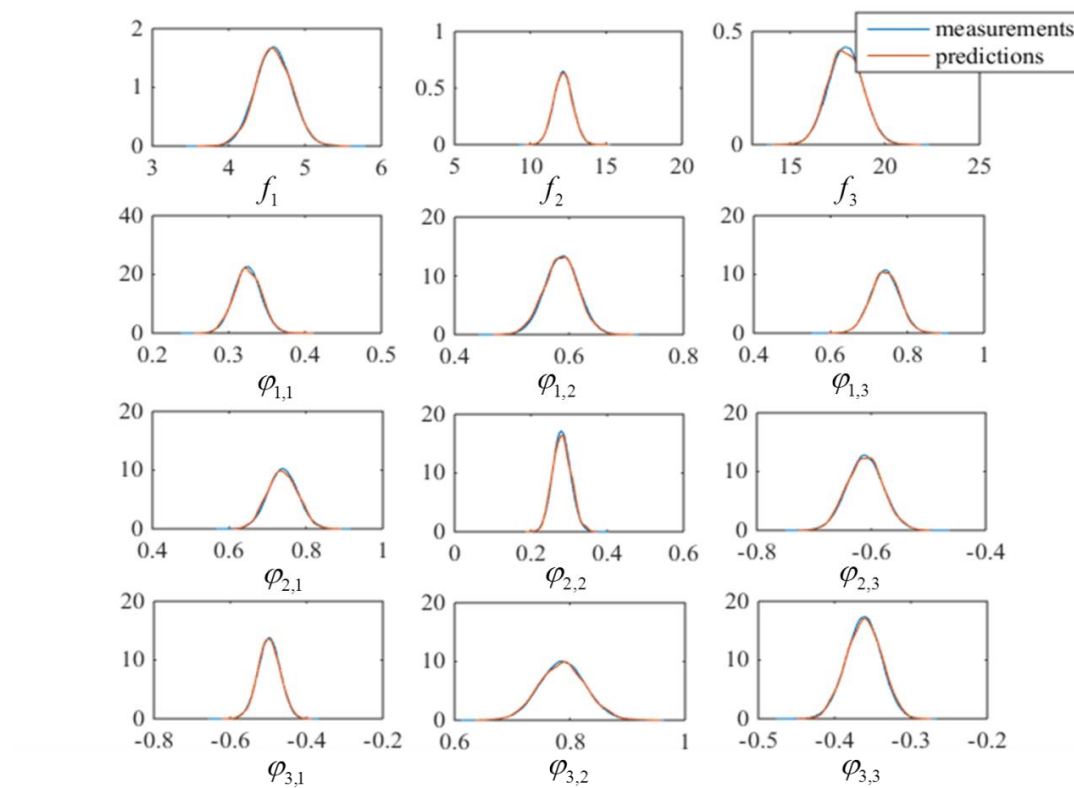


Fig. 6-5 Comparison between measured and predicted PDFs for the modal properties

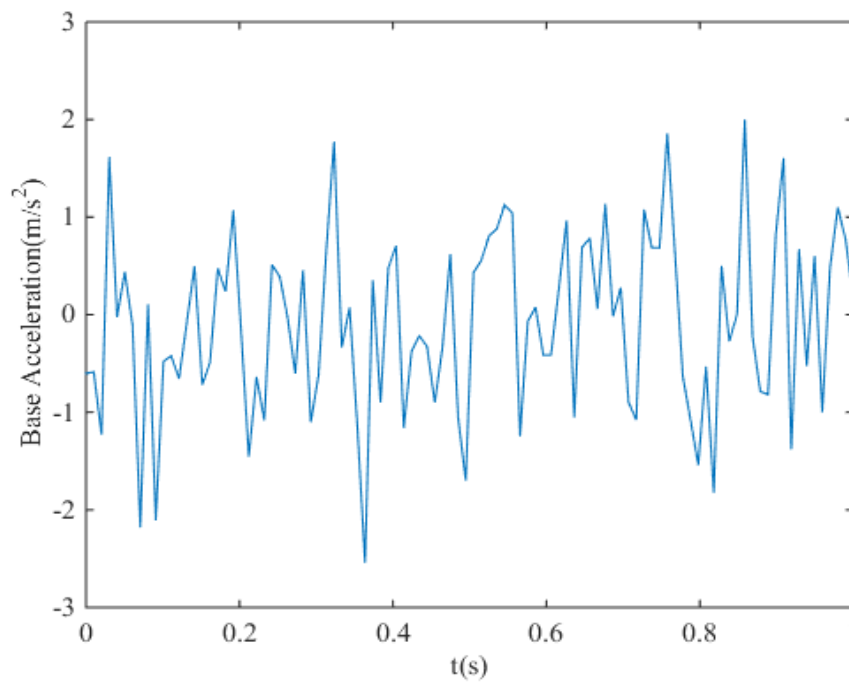


Fig. 6-6 Base excitation

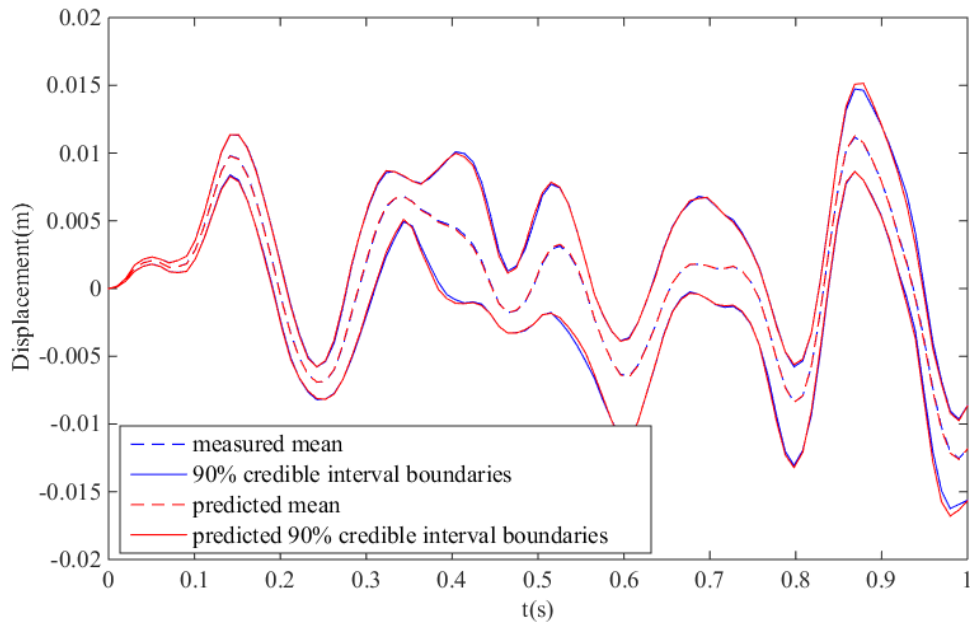


Fig. 6-7 Comparison between measured and predicted results for the displacement time history at DOF 3

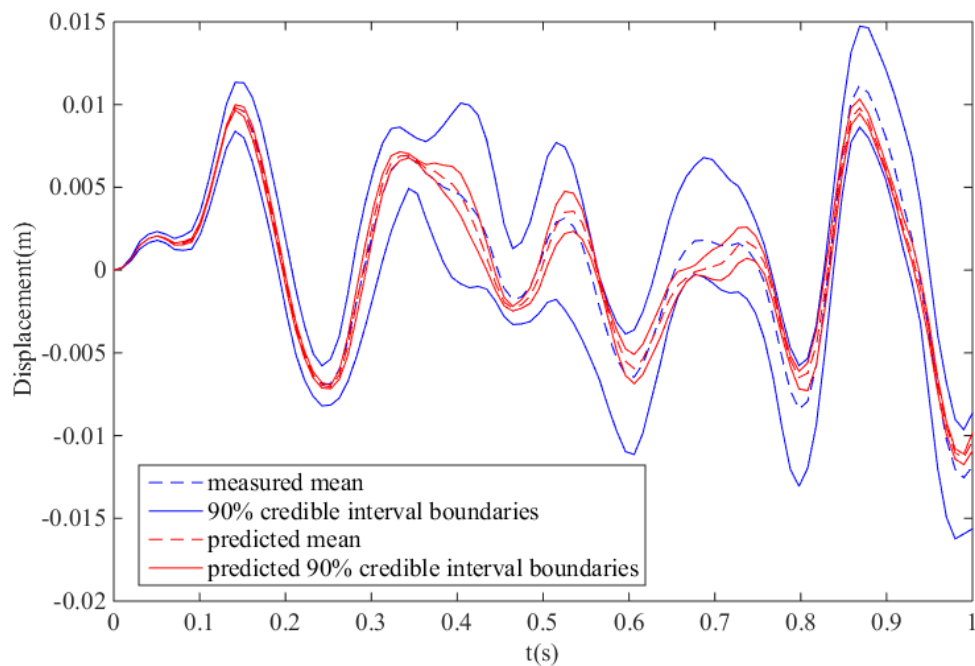


Fig. 6-8 Comparison between measured and predicted results for the displacement time history at DOF 3 without consideration of error terms in prediction

6.5 Estimation of S-N Curve Model Parameters using Fatigue Data

In this section, the experimental data from fatigue tests are used to infer the uncertainties in the model parameters of the S - N curves. As the S - N data is usually modelled by a linear model, the equations derived for the linear model in Section 6.2.1 are directly applicable.

6.5.1 Model description

Materials fatigue performance is commonly characterized in the form of an S - N curve, which is usually simulated by the Basquin's relation [35]

$$N_k = AS_k^{-B} \quad (6.25)$$

where N_k expresses the fatigue life at the k -th stress level, S_k expresses corresponding stress level, while $A > 0$ and $B > 0$ are material parameters to be estimated using experimental data. For the laboratory fatigue tests, the fatigue life dispersion always exists due to many factors, such as uncertainties of mechanical properties of structure and material, changing environmental factors in laboratory, and random error in observations, etc. Taking the variability into consideration, Eq. (6.25) can be expressed as

$$N_k = AS_k^{-B}\eta_k \quad (6.26)$$

where η_k quantifies the randomness of fatigue life in stress level S_k . For the sake of simplicity, Eq. (6.26) is usually rewritten in log scale as [36]

$$\tilde{q}_k = \alpha + \beta x_k + e_k \quad (6.27)$$

where $\tilde{q}_k = \log_{10}(N_k)$, $x_k = \log_{10} S_k - \frac{1}{n} \sum_{j=1}^n \log_{10} S_j$, $\alpha = \log_{10}(A) - B \frac{1}{n} \sum_{j=1}^n \log_{10} S_j$, $\beta = B$,

and $e_k = \ln \eta_k$ is assumed to follow a zero mean normal distribution with dispersion σ .

Eq. (6.27) is usually adopted as a probabilistic model to estimate probabilistic S - N curves. To consider the variability of the model parameter set $\theta = \{\alpha, \beta\}$ for different specimens, the parameters α and β are respectively assumed to follow Gaussian distributions as $N(\alpha | \mu_\alpha, \sigma_\alpha^2)$ and $N(\beta | \mu_\beta, \sigma_\beta^2)$, where μ_α , σ_α , μ_β and σ_β are the hyperparameters. To avoid unidentifiability issues with respect to the model parameters σ_α and σ due to the presence of the additive terms $\alpha + e_k$ in Eq. (6.27), the term e_k and its uncertainty is absorbed in α by replacing Eq. (6.27) with the model

$$\tilde{q}_k = \alpha + \beta x_k \quad (6.28)$$

Based on this model, both the variability of different specimens in one stress level and across stress levels can be comprehensively considered in σ_α and σ_β .

6.5.2 Uncertainty quantification using fatigue tests

The data listed in Table 6-7, taken from [37], are used to infer the model parameters. Fatigue tests were conducted with standard plate specimens of aluminum alloy 2524-T3 under four stress levels with about 15 observations each. The data from each stress level is assumed to follow a Gaussian distribution, and the mean and standard deviation are solved as measured statistics. Based on the probabilistic model described above, the proposed probabilistic modeling framework can be implemented. Given the prior distributions of all the hyper parameters as uniform distribution with their upper and lower bounds listed in Table 6-8, the nested sampling algorithm is implemented to generate samples of the hyper parameters $\{\mu_\alpha, \sigma_\alpha, \mu_\beta, \sigma_\beta\}$ as shown in Fig. 6-9. Moreover, the most probable values of $\{\mu_\alpha, \sigma_\alpha, \mu_\beta, \sigma_\beta\}$ computed according to Eq. (A.3), as well as the standard deviation of the estimates of the hyper-parameters computed using the samples, are listed in Table 6-9. Based on Eq. (6.17), the posterior distribution of α and β are shown in Fig. 6-10. Using the samples in Fig. 6-9, the uncertainties in the estimates of the hyperparameters are of the order of 0.83% and 12% for the hyper-mean parameters $\{\mu_\alpha, \mu_\beta\}$ and of the order of 46% and 70% for the hyper-standard deviation parameters $\{\sigma_\alpha, \sigma_\beta\}$. Also, from Table 6-9 the coefficient of variation of the parameters α and β based on the mean values of the hyper parameters are $\hat{\sigma}_\alpha / \hat{\mu}_\alpha = 1.7\%$ and $\hat{\sigma}_\beta / \hat{\mu}_\beta = 21\%$, respectively. It can be seen that there is considerable uncertainty in the values of the parameters α and β that can affect fatigue predictions.

Table 6-7 Fatigue life test data of aluminum alloy 2524-T3

S_i (MPa)	$\log_{10} N_i$
200	5.603, 5.544, 5.528, 5.630, 5.594, 5.540, 5.581, 5.548, 5.426, 5.567, 5.554, 5.627, 5.630, 5.596, 5.626
300	5.028, 5.074, 5.016, 4.894, 4.993, 5.071, 5.024, 5.035, 4.954, 5.039, 5.098, 5.057, 5.092, 5.082, 5.005
350	4.784, 4.842, 4.776, 4.813, 4.813, 4.860, 4.798, 4.776, 4.758, 4.770, 4.755, 4.837, 4.736, 4.842, 4.796
400	4.477, 4.400, 4.426, 4.462, 4.592, 4.411, 4.447, 4.402, 4.665, 4.475, 4.458, 4.551, 4.525, 4.641

Table 6-8 Upper and lower bounds of hyper parameters

	μ_α	σ_α	μ_β	σ_β	σ_ε
Upper bound	-30	-30	0	0	0
Lower bound	30	30	10	10	10

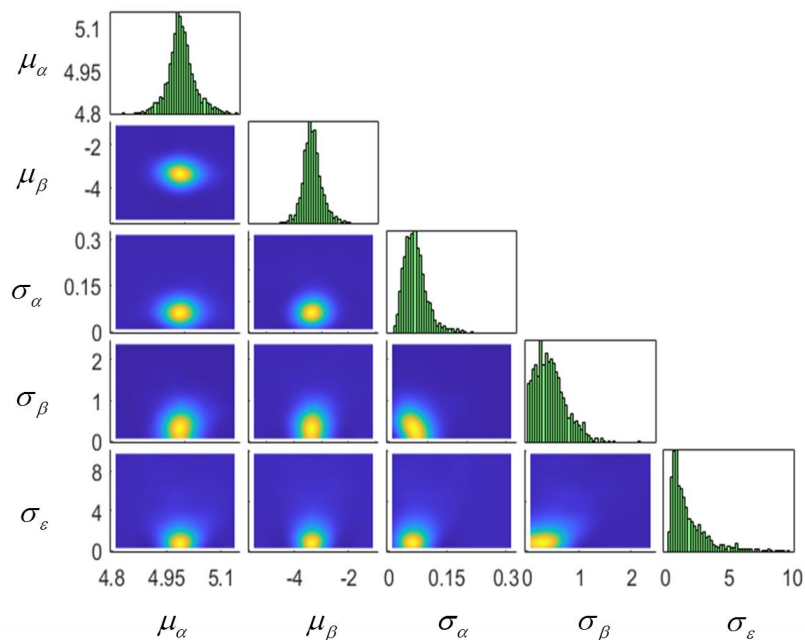
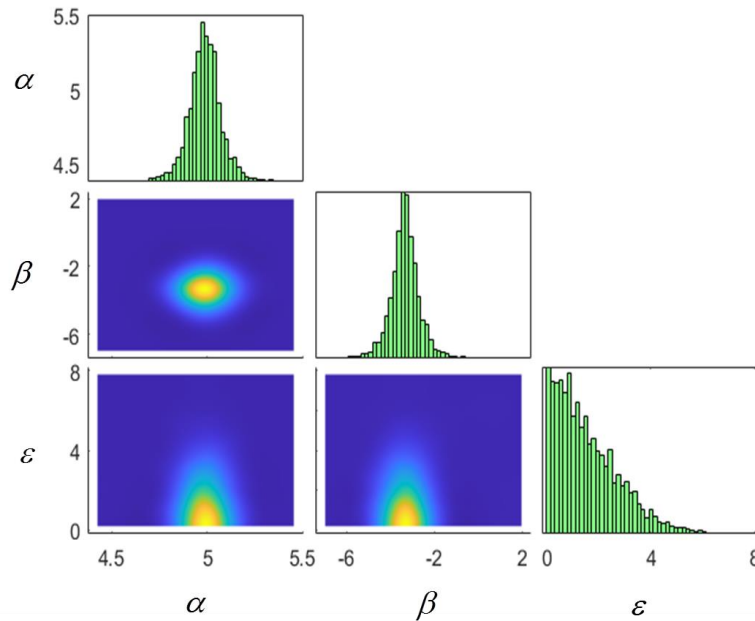


Fig. 6-9 Samples of joint posterior distribution of hyper parameters

Table 6-9 Most probable values of hyper parameters

	μ_α	μ_β	σ_α	σ_β	σ_ε
MPV	4.984	-3.435	0.511	0.372	0.387
Standard deviation	0.0127	0.1533	0.0166	0.1958	0.1367

Fig. 6-10 Samples of posterior distribution of α , β and ε

6.5.3 Uncertainty propagation

Using the samples of α and β generated by Eq. (6.17), the samples of q_k for various stress levels are predicted and the 90% credible interval is estimated. Results are shown in Fig. 6-11 and compared with measured data available for the four stress levels. Predictions of the 90% credible intervals take into account the uncertainty in the hyper-parameters. Results are also presented for the 90% credible intervals estimated by ignoring the uncertainties in the hyper-parameters. This is achieved by drawing samples from the distribution $N(\boldsymbol{\theta} | \hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0)$ and propagating these samples for predicting the fatigue life for different stress levels in Fig. 6-11. It can be seen that the 90% uncertainty intervals considering the uncertainties in the hyper parameters are narrow enough and contain the fatigue data available at the four stress levels. Moreover, these uncertainty intervals are comparable to uncertainty intervals obtained by methods based on HBM [38,39]. The uncertainty intervals computed using the MPV of the

hyper parameters, ignoring the uncertainties in the hyper parameters, are narrower and contain well a large percentage of fatigue data. It is evident, however, that propagation based on the MPV of the hyper parameters fail to fully contain all the data. The discrepancy between the two credible intervals is expected to decrease as one includes fatigue data from more than four stress levels.

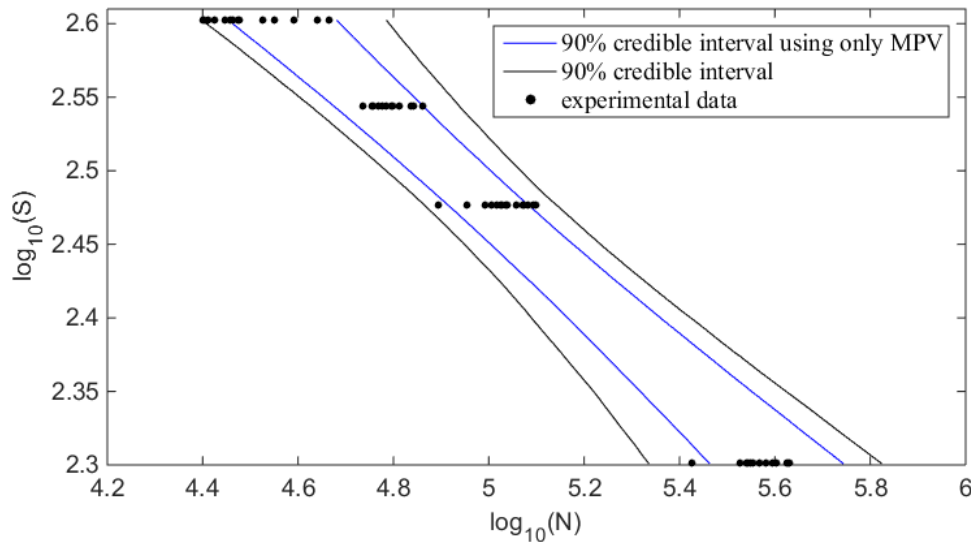


Fig. 6-11 Predicted 90% credible intervals and comparison with measured fatigue data

Finally, from the results in Fig. 6-9, it is observed that the values of σ_ε are not close to zero, which means there are some discrepancies between predicted PDFs and the measured PDFs. This is also depicted in Fig. 6-11, as well as in Fig. 6-12 comparing, for each stress level, the prediction of the Gaussian PDF of q_k based on the model to the Gaussian PDF based on the measurements. For the model predictions, the PDF corresponding to the most probable values of the hyperparameters is presented along with the PDF taking into account the uncertainties in the hyper parameters. It can be seen that the most probable values of two PDFs in each figure are consistent, while the uncertainty predicted using only MPV is narrower than that considering uncertainties of hyper parameters. The reason for the discrepancies between measured and model predicted PDFs is that the variation of the available fatigue data from the four stress levels deviates from the linear model, so that the predictions from the linear model cannot exactly account for the mean and variance of experimental data for all four stress levels simultaneously.

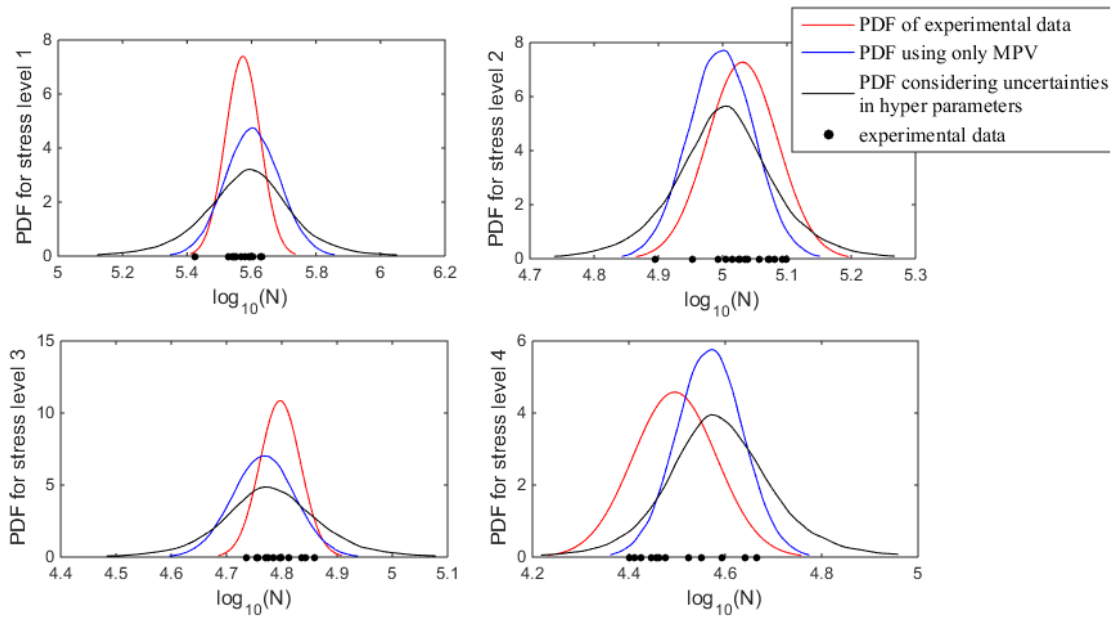


Fig. 6-12 Comparison of PDF of experimental data and model predicted PDFs for the four stress levels

6.6 Conclusion

The new Bayesian inference method proposed in this work addresses the issue of underestimation of the uncertainty in the model parameters due to model error, mentioned in [24], arising from multiple measurements available for a structure or measurements available for members of a population of identically manufactured structures [18,19]. The proposed method offers an alternative to HBM methods recently proposed in the literature [21,25] to correctly address the uncertainty in the model parameters. Based on the proposed framework, uncertainties are embedded in the model parameters by assigning a Normal distribution with mean and covariance constituting the hyperparameters to be estimated using Bayesian inference. The posterior distribution of hyper parameters of the model parameters is directly computed by Bayes theorem applied on KL-div measures between the model predicted PDF and the PDF of the experimental data. In particular, the proposed framework is applicable when the statistics of the measured quantities are available. Computationally efficient and insightful analytical expressions for the posterior distribution of the hyperparameters were developed for the case for which the PDFs are approximated by Normal distributions. In particular, Normal distributions for the predictions arise when the output QoI depend linearly on the model parameters. In this case the posterior PDF of the model parameters depends on the lower two moments of the respective PDFs. This representation of the posterior is also used for non-Gaussian PDFs to approximate the uncertainty in the model parameters. For

nonlinear relations between the output QoI and the model parameters, the univariate dimension reduction method (UDRM) is used to efficiently estimate the involved multi-dimensional integrals for the lower two moments of model predicted PDF.

An application to structural dynamics based on measured modal properties from a group of identically manufactured 3-DOF systems is presented based on simulated data to illustrate the proposed framework. The effectiveness of the methodology is demonstrated by noting that the estimates of hyperparameters, model parameters, and uncertainties recover the values used to simulate the data. Also, the method is applied to the quantification of uncertainties of the parameters of $S-N$ curves based on the experimental data from fatigue tests, demonstrating that the proposed framework can also obtain competitive results to alternative methods based on HBM.

Appendix A

Introduce the function $L(\mathbf{h}, \sigma_\varepsilon)$ defined as the negative of the logarithm of the posterior distribution in Eq. (6.13)

$$L(\mathbf{h}, \sigma_\varepsilon) = -\ln p(\mathbf{h}, \sigma_\varepsilon | \varepsilon) = n_q \ln \sigma_\varepsilon + \frac{n_q}{2\sigma_\varepsilon^2} J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\}) \quad (\text{A.1})$$

where the prior PDF $p(\mathbf{h})$ is assumed to be uniform. Using Taylor series expansion with respect to variables \mathbf{h} about the most probable value $\hat{\mathbf{h}}$ of $p(\mathbf{h}, \sigma_\varepsilon | \varepsilon)$, given by

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} [J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})] \quad (\text{A.2})$$

and keeping the first two non-zero terms in the expansion, the posterior PDF $p(\mathbf{h}, \sigma_\varepsilon | \varepsilon) = \exp(-L(\mathbf{h}, \sigma_\varepsilon))$ can be approximated as (valid for larger number of output quantities)

$$p(\mathbf{h}, \sigma_\varepsilon | \varepsilon) \propto p(\sigma_\varepsilon) N(\mathbf{h} | \hat{\mathbf{h}}, \sigma_\varepsilon^2 \hat{\Sigma}_{\mathbf{h}}) \quad (\text{A.3})$$

where the covariance matrix $\sigma_\varepsilon^2 \hat{\Sigma}_{\mathbf{h}}$ equals to the inverse of Hessian matrix $\mathbf{H}(\hat{\mathbf{h}}, \sigma_\varepsilon)$ of the function $L(\mathbf{h}, \sigma_\varepsilon)$ evaluated at $\hat{\mathbf{h}}$ and given by $\frac{n_q}{\sigma_\varepsilon^2} \mathbf{H}(\hat{\mathbf{h}})$, where

$$\mathbf{H}(\hat{\mathbf{h}}) = \frac{\partial^2 J(\mathbf{h}; \{y_k\}, \{\tilde{q}_k\})}{\partial \mathbf{h}^T \partial \mathbf{h}} \bigg|_{\mathbf{h}=\hat{\mathbf{h}}} \quad (\text{A.4})$$

So the covariance matrix $\hat{\Sigma}_{\mathbf{h}}$ is given by $\hat{\Sigma}_{\mathbf{h}} = \frac{1}{n_q} \mathbf{H}^{-1}(\hat{\mathbf{h}})$. The MPV can be directly computed using an optimization tool, while the hessian matrix can be computed numerically or analytically [40,41].

6.7 References

- [1] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, *Journal of Engineering Mechanics*. 124 (1998) 455–461.
- [2] L.S. Katafygiotis, C. Papadimitriou, H.-F. Lam, A probabilistic approach to structural model updating, *Soil Dynamics and Earthquake Engineering*. 17, (1998) 495–507.
- [3] M.W. Vanik, J.L. Beck, S.-K. Au, Bayesian probabilistic approach to structural health monitoring, *Journal of Engineering Mechanics*. 126 (2000) 738–745.
- [4] J.L. Beck, S.-K. Au, Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation, *Journal of Engineering Mechanics*. 128 (2002) 380–391.
- [5] E.A. Johnson, H.F. Lam, L.S. Katafygiotis, J.L. Beck, Phase I IASC-ASCE Structural Health Monitoring Benchmark Problem Using Simulated Data, *Journal of Engineering Mechanics*. 130 (2004) 3–15.
- [6] J.L. Beck, Bayesian system identification based on probability logic, *Structural Control and Health Monitoring*. 17 (2010) 825–847.
- [7] D. Straub, I. Papaioannou, Bayesian Updating with Structural Reliability Methods, *Journal of Engineering Mechanics*. 141 (2015) 04014134.
- [8] J.L. Beck, K.V. Yuen, Model Selection Using Response Measurements: Bayesian Probabilistic Approach, *Journal of Engineering Mechanics*. 130. No. 2 (2004) 192–203.
- [9] K.V. Yuen, *Bayesian Methods for Structural Dynamics and Civil Engineering*, 2010.
- [10] H. Sohn, K.H. Law, A Bayesian probabilistic approach for structure damage detection, *Earthquake Engineering and Structural Dynamics*. 26 (1997) 1259–1281.
- [11] K.V. Yuen, J.L. Beck, S.K. Au, Structural damage detection and assessment by adaptive Markov chain Monte Carlo simulation, *Structural Control and Health Monitoring*. 11 (2004) 327–347.
- [12] K.V. Yuen, J.L. Beck, L.S. Katafyglotis, Unified probabilistic approach for model updating and damage detection, *Journal of Applied Mechanics, Transactions ASME*. 73 (2006) 555–564.

- [13] C. Papadimitriou, J.L. Beck, L.S. Katafygiotis, Updating robust reliability using structural test data, *Probabilistic Engineering Mechanics*. 16 (2001) 103–113.
- [14] J.L. Beck, A.A. Taflanidis, Prior and Posterior Robust Stochastic Predictions for Dynamical Systems Using Probability Logic, *International Journal for Uncertainty Quantification*. 3 (2013) 271–288.
- [15] P. Cornwell, C.R. Farrar, S.W. Doebling, H. Sohn, Environmental variability of modal properties, *Experimental Techniques*. 23 (1999) 45–48.
- [16] S. Alampalli, Effects of testing, analysis, damage, and environment on modal parameters, *Mechanical Systems and Signal Processing*. 14 (2000) 63–74.
- [17] P. Moser, B. Moaveni, Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge, *Mechanical Systems and Signal Processing*. 25 (2011) 2336–2357.
- [18] G.C. Ballesteros, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian hierarchical models for uncertainty quantification in structural dynamics, in: *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, 2014: pp. 1615–1624.
- [19] J.B. Nagel, B. Sudret, A unified framework for multilevel uncertainty quantification in Bayesian inverse problems, *Probabilistic Engineering Mechanics*. 43 (2016) 68–84.
- [20] J.N. Rouder, L.U. Jun, An introduction to Bayesian hierarchical models with an application in the theory of signal detection, *Psychonomic Bulletin and Review*. 12 (2005) 573–604.
- [21] I. Behmanesh, B. Moaveni, G. Lombaert, C. Papadimitriou, Hierarchical Bayesian model updating for structural identification, *Mechanical Systems and Signal Processing*. 64–65 (2015) 360–376.
- [22] I. Behmanesh, B. Moaveni, Accounting for environmental variability, modeling errors, and parameter estimation uncertainties in structural identification, *Journal of Sound and Vibration*. 374 (2016) 92–110.
- [23] S. Wu, P. Angelikopoulos, J.L. Beck, P. Koumoutsakos, Hierarchical Stochastic Model in Bayesian Inference for Engineering Applications: Theoretical Implications and Efficient Approximation, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*. 5 (2019).
- [24] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Probabilistic hierarchical Bayesian framework for time-domain model updating and robust predictions, *Mechanical Systems and Signal Processing*. 123 (2019) 648–673.

- [25] O. Sedehi, C. Papadimitriou, L.S. Katafygiotis, Data-driven uncertainty quantification and propagation in structural dynamics through a hierarchical Bayesian framework, *Probabilistic Engineering Mechanics*. 60 (2020).
- [26] D. Patsialis, A.P. Kyprioti, A.A. Taflanidis, Bayesian calibration of hysteretic reduced order structural models for earthquake engineering applications, *Engineering Structures*. 224 (2020) 111204.
- [27] E.T. Jaynes, Information theory and statistical mechanics, *Physical Review*. 106 (1957) 620.
- [28] K. Sobczyk, J. Trębicki, Approximate probability distributions for stochastic systems: Maximum entropy method, *Computer Methods in Applied Mechanics and Engineering*. 168 (1999) 91–111.
- [29] D.W. Scott, *Multivariate density estimation: theory, practice, and visualization*, John Wiley & Sons, 2015.
- [30] C. Yang, R. Duraiswami, N.A. Gumerov, L. Davis, Improved fast gauss transform and efficient kernel density estimation, *Proceedings of the IEEE International Conference on Computer Vision*. 2003-Janua (2003) 664–671.
- [31] Kullback S, Leibler R A. On information and sufficiency. *The annals of mathematical statistics*, 1951, 22(1): 79-86.
- [32] M. Coughlin, N. Christensen, J. Gair, S. Kandhasamy, E. Thrane, Method for estimation of gravitational-wave transient model parameters in frequency-time maps, *Classical and Quantum Gravity*. 31 (2014).
- [33] S. Rahman, H. Xu, A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics, *Probabilistic Engineering Mechanics*. 19 (2004) 393–408.
- [34] O. Sedehi, L.S. Katafygiotis, C. Papadimitriou, Hierarchical Bayesian operational modal analysis: Theory and computations, *Mechanical Systems and Signal Processing*. 140 (2020).
- [35] O.H. Basquin, The exponential law of endurance tests, in: *Proc Am Soc Test Mater*, 1910: pp. 625–630.
- [36] M. Guida, F. Penta, A Bayesian analysis of fatigue data, *Structural Safety*. 32 (2010) 64–76.
- [37] L. Xie, J. Liu, N. Wu, W. Qian, Backwards statistical inference method for P-S-N curve fitting with small-sample experiment data, *International Journal of Fatigue*. 63 (2014) 62–67.

- [38] X.W. Liu, D.G. Lu, P.C.J. Hoogenboom, Hierarchical Bayesian fatigue data analysis, *International Journal of Fatigue*. 100 (2017) 418–428.
- [39] J. Chen, S. Liu, W. Zhang, Y. Liu, Uncertainty quantification of fatigue S-N curves with sparse data using hierarchical Bayesian data augmentation, *International Journal of Fatigue*. 134 (2020) 105511.
- [40] H. Jensen, C. Papadimitriou, Bayesian Finite Element Model Updating, in: *Sub-Structure Coupling for Dynamic Analysis*, Springer, 2019: pp. 179–227.
- [41] E. Ntotsios, C. Papadimitriou, Multi-objective optimization algorithms for finite element model updating, in: *23rd International Conference on Noise and Vibration Engineering 2008, ISMA 2008*, 2008: pp. 1895–1909.

Chapter 7. Conclusions and Future Work

7.1 Concluding Remarks

This thesis has presented novel methodologies for performing uncertainty quantification and propagation in structural dynamical simulations, with Bayesian inference as the platform for integrating physics-based models of dynamical structures, and measurement information collected during system operation. The presented methodologies provided a versatile and efficient Bayesian probabilistic framework to properly account for and quantify the parameter uncertainties, realistically predict the output responses as well as efficiently assess system reliability and robustness to uncertainty. Key conclusions and innovations of this thesis are outlined as follows.

A hierarchical Bayesian modeling (HBM) framework is developed to account for the uncertain variabilities due to model error, based on modal properties for linear systems and time histories responses in nonlinear systems. This framework introduces probabilistic models for the model parameters as well as the prediction error parameter that depend on their hyper parameters to capture the uncertainties due to variabilities from test to test. Asymptotic approximations have been developed within the HBM framework, providing insights for the interpretations of different sources of uncertainties. Specifically, according to the framework, uncertainty due to model error arising from experimental, operational, environmental and manufacturing variabilities is embedded in the structural model parameters by assigning a parameterized prior probability distribution in the model parameters with hyperparameters to be inferred from multiple datasets available. The unmodelled dynamics is quantified by a Gaussian prediction error term with covariance to be either estimated from the available datasets or further described by parameterized probability distributions to account for the variability of the prediction error covariance matrix over the available datasets. Analytical expressions for the hyper parameters reveal that the uncertainties due to variabilities that arise from the multiple data sets are irreducible and dominant as compared to the identification uncertainties estimated within a data set.

Conventional Bayesian technique fails to take into account such uncertainties, resulting in an underestimation of the parameter uncertainties, while the proposed HBM framework provides more reasonable account for the parameter uncertainties. Besides the insightful expressions, the asymptotic approximations considerably improve the computational efficiency of the HBM framework compared to the framework using a full sampling procedure. Within this framework, uncertainties of model parameters and prediction error parameter can be propagated to predict more reliable uncertainty bounds for observed and unobserved output quantities of interests. Reliability analysis perspective has been also incorporated into the framework to account for the data-driven uncertainty in the model parameters in the prediction of the probability of failure of dynamical systems. Simulated and experimental studies with identified modal properties are utilized for calibrating models, predicting responses and updating system reliability. Results from these studies demonstrate that the proposed HBM framework provides better understanding of parameter uncertainties and offers realistic uncertainty bounds of response predictions as compared to the conventional Bayesian method which severely underestimates the parameter uncertainties and lumps all the uncertainties into the prediction error function.

An innovation of this thesis lies in the capability of capturing diverse sources of uncertainty in the model parameters within a multi-level modeling approach for a structure consisting of components, subsystem and systems levels. Test and monitoring data from each level in the modeling hierarchy provide essential information required to build confidence in the models used, guiding the process of selecting, calibrating and improving the mathematical models. Data-driven modeling is important for narrowing the uncertainties through the modeling and simulation process of the system, leading to improved and accurate predictions of the system performance and safety under various operational and environmental conditions. To build reliable models from multi-level data one should also take into account that models set up at component, sub-system and system levels may share common parameters and thus these common parameters should be inferred from data obtained from more than one modeling level. The aforementioned HBM framework has the great potential to be extended to the multi-level modeling approach of a structure. In such a model hierarchy, hyper parameters are defined as the sharing information between lower levels and higher levels of a system. The estimation process of the hyper parameters is conducted based on the available datasets in all levels of model hierarchy, and the response QoI are predicted according to the uncertainties obtained from the parameter estimation process. A simple dynamical system example is used as a verification of the concept and a

benchmark of the proposed method in the context of real engineering applications. Results indicated that the proposed framework can consider the overall information of a multi-level structure to account for all datasets available at different levels of hierarchy, thus building more reliable models and obtaining more accurate response predictions.

This thesis also revisits the underestimation problem of parameter uncertainties within the conventional Bayesian framework, proposing three data features-based models for likelihood informed Bayesian model updating. Data features herein are introduced as the average of the square of the discrepancy between the measurements and model predictions. Two data features-based models are developed for the uncorrelated prediction error case, while one data features-based model is presented for spatial correlated case. Analytical expressions are carried out for the model parameters with the uncorrelated data features models. Based on such analytical solutions, relations between the proposed models and the likelihood free ABC method are investigated and obtained. Deriving such relations are beneficial especially for the choices of the tolerance values used in ABC algorithm. A dynamical system is employed for demonstration purposes. Model updating, response predictions and model selections are conducted using the proposed models. Conventional Bayesian formulations using a Gaussian prediction error model are also performed and compared with the proposed models. Results showed that the proposed data features-based model inference using an exponential distribution provides the best one over the three proposed distributions. It is also indicated that all the proposed models provide consistent confidence intervals for the model parameters which are independent of the sampling rate that is used to accurately represent the same time histories response data. In contrast, the conventional Bayesian formulation fails to properly account for different sampling rates, resulting to unrealistically small uncertainties as the sampling rate increases or the number of data points representing the time history increases.

For the case where the multiple data sets are not available, but only given by PDF or statistics of the measurements, this thesis proposed a statistics-based Bayesian framework for uncertainty quantification and propagation in engineering simulations. As in the HBM framework, uncertainties are embedded in the structural model by assigning a parameterized prior probability distribution of the model parameters, with the hyperparameters to be estimated from the multiple statistics available for measured QoI. In this framework, the discrepancy of the PDFs or statistics between the measurements and model predictions are quantified using a measure of KL-divergence. The posterior PDF of the model parameters depends on the lower two moments of the respective PDFs. This representation of the posterior is also used for non-Gaussian PDFs of measurements and model predictions to approximate the

uncertainty in the model parameters. The proposed framework is advantageous to engineering problems especially for S-N curve fatigue analysis where difficulties arise from the selection of multiple data sets. The proposed framework can also be viewed as an alternative to the HBM framework to tackle the situation where only PDFs or statistical characteristics are available for measurements. Comparisons are also made between the proposed method and conventional Bayesian approach. It has been shown that the proposed method is more flexible to quantify the parameter uncertainties and obtain better predictions than that from the conventional Bayesian approach.

7.2 Future Work

According to the current development of Bayesian methodologies for uncertainty quantification and propagation in structural dynamical simulations in this thesis, several potential future work could be followed:

1. Hierarchical modeling framework for multi-level models with model nonlinearities.

Although the proposed Bayesian methodologies in this thesis are conceptually distinct, mathematically elegant, theoretically rigorous, it has not been tested in the real changing dynamical problem. Specifically, for more complicated multi-level models, the effectiveness for model nonlinearities needs to be demonstrated using the hierarchical modeling framework integrating real data from components, subsystems and systems. The computational issues for such a complicated model could be addressed using model reduction techniques or surrogate models.

2. Hierarchical Bayesian neural network for model calibration and verification in structural dynamics.

The methodologies presented in this thesis are based on physics-based models and measured data during system operation. A pure data-driven approach is also preferred as one of the future directions incorporating into the HBM framework. For this, Bayesian neural network could be potentially applied to link the model input and output. Analytical expressions could be derived and further developed for general cases in structural dynamics. This might be an important step forward on the development of HBM framework.

3. Hierarchical Bayesian learning for damage identification.

This thesis proposed a novel hierarchical Bayesian modeling framework for model updating and response predictions in structural dynamics. Damage identification is of great importance in the field of dynamical structures, especially in the area of structural health monitoring. This thesis lacks sufficient discussions and applications on this topic. However, it has the great potential to be applied to this area to consider the local damage affected by the uncertainty due to variability arisen from environmental conditions and different loading excitations. Sparse Bayesian learning techniques can be incorporated within the HBM framework to effectively account for the sparseness of damage and address the ill-conditioning associated with the large number of model parameters required to be introduced in order to cover all possible damage scenarios.