ABSTRACT

TADINADA, SASHI KANTH. A Bayesian Framework for Probabilistic Seismic Fragility Assessment of Structures (Under the direction of Dr. Abhinav Gupta.)

A statistical framework is proposed to address the problem of evaluating and updating estimates of seismic fragility for structural systems in nuclear power plants and other critical infrastructure facilities. The main objective of the framework is to enable (i) integration of fragility estimates from existing studies with the new fragility data from experiments or finite element simulations; (ii) planning and allocation of resources for efficient collection of fragility data. Seismic fragility analyses of large-scale engineering structures is performed by conducting large-scale simulations involving multiple nonlinear time-history analyses using experimentally-validated finite element models. Accurate estimation of fragility often demands a formidable number of nonlinear time-history simulations which can be computationally prohibitive. A computational methodology based on Bayesian inference techniques is proposed to reduce the computational effort and efficient allocation of computing resources. Application of the proposed methodology to large-scale piping systems with localized nonlinearities is simplified further by proposing an equivalent linearization technique. A fragility model is developed for these systems where linearization of the localized nonlinearity enables us to evaluate a robust prior estimate of fragility using only elastic time-history analyses. The model is incorporated into the Bayesian framework to evaluate posterior estimates by updating the prior using a much reduced number of nonlinear time-history analyses. Modeling of uncertainties in the interdependencies between input parameters (material and model) is an important step in setting up probabilistic simulations. A sampling algorithm is suggested in order to facilitate simulation of closely-spaced random variables subjected to inequality constraints.

A Bayesian Framework for Probabilistic Seismic Fragility Assessment of Structures

by
Sashi Kanth Tadinada

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Civil Engineering

Raleigh, North Carolina

2012

APPROVED BY:

Dr. Vernon Matzen	Dr. Mervyn Kowalsky
Dr. Shamim Rahman	Dr. Abhinav Gupta Chair of Advisory Committee

DEDICATION

To *Amma* and *Nanni* …

BIOGRAPHY

Sashi Kanth Tadinada was born on August 26th, 1983 in Kakinada, India. His family hails from the town of Rajahmundry in India. He joined the undergraduate program in civil engineering at Indian Institute of Technology, Roorkee in July 2000 and received the Bachelors of Technology (B.Tech) degree in May 2004. Following graduation, he worked for two years at Larsen & Toubro Ltd., - Engineering Construction and Contracts Division (L&T-ECCD) as a senior engineer in their Transportation Infrastructure (TI) Business Unit in New Delhi, India. In August 2006, he came to the U.S. to join the graduate school pursuing both MS (2008) and PhD (2012) degrees from North Carolina State University.

Sashi's research interests include structural dynamics, Bayesian statistics, probabilistic methods, risk assessment, optimization techniques and scientific computing.

ACKNOWLEDGEMENTS

The research was supported by the Center for Nuclear Power Plant Structures, Equipment and Piping at North Carolina State University. Resources for the center came from dues of the member organizations, National Science Foundation, Department of Civil, Construction and Environmental Engineering and College of Engineering at NCSU. I am greatly indebted to the center for proving me with all the facilities and support throughout my graduate school. I would also like to thank all the staff members of the Civil Engineering department for their help and support extended over the years.

I would like to express deepest and sincere gratitude for my advisor, Dr. Abhinav Gupta for giving me this opportunity to pursue my graduate studies and teaching me everything about scientific research & writing. His constant guidance and inspiration throughout my research as well as in the preparation of this manuscript is simply invaluable. I would like to thank Dr. Mervyn Kowalsky, Dr. Vernon Matzen and Dr. Shamim Rahman for graciously agreeing to serve on my doctoral committee and for their valuable suggestions. Special thanks to my colleagues and friends – Sami, Ju and Yonghee for sharing the results from their work and for their constant input and feedback without which this collaborative research could not have been possible.

None of this effort would have been remotely possible without the unconditional love, support and encouragement from my dear parents and my kid brother. I appreciate the sacrifices they made for me to have this education. A special shout-out to all the amazing

friends I made during my stay at NCSU. Last but not the least, I want to thank the workshop "*Medhajananam*" and its creator Krishna Sharma for lending a unique and beautiful perspective to my world-view.

TABLE OF CONTENTS

LIST OF FIGURES

PART II : A BAYESIAN FRAMEWORK FOR ESTIMATING AND UPDATING SEISMIC FRAGILITY OF STRUCTURES

PART III : STRUCTURAL FRAGILITY OF LARGE-SCALE PIPING SYSTEMS USING EQUIVALENT ELASTIC TIME-HISTORY SIMULATIONS

PART IV: SIMULATION OF CONSTRAINED, CLOSELY-SPACED RANDOM VARIABLES IN ENGINEERING RISK ANALYSES

LIST OF TABLES

PART I

INTRODUCTION

1. BACKGROUND

1.1 Probabilistic Seismic Risk Analysis (PSRA)

The fundamental goal in the seismic design of any Nuclear Power Plant (NPP) is the minimization of radiological risk (release of radioactive emissions into the environment) and ensuring the safety during and after an earthquake. It is critical that engineers be able to estimate the associated risks due to seismic events to a high degree of confidence. In the nuclear industry, this is achieved using a detailed probabilistic framework that carefully accounts for uncertainties in the expected intensity of ground motions, the predicted response of the structures and ultimately, the chain of events that could lead to the catastrophic failure. The regulatory guide (USNRC, 1983) is a comprehensive document establishing a detailed framework for Probabilistic Risk Analysis (PRA) of Nuclear Power Plants. It identifies the following analyses as key elements of seismic risk assessment for any nuclear power plant structure:

- *Seismic Hazard Analysis at the site* – The seismic hazard $H(\lambda)$ at a site describes the frequency i.e. the probability per year that an earthquake with a Peak Ground Acceleration (PGA) value equal to λ is likely to occur. The seismic hazard $H(\lambda)$ is usually determined from a probabilistic seismic hazard model that incorporates the seismological and geological characteristics of the region (Kennedy et al, 1980).

- *The Responses of Plant Systems and Structures* – In order to obtain the probability of damage given an earthquake of PGA level equal to λ, the statistical properties of the seismic response of each structural system in the power plant need to be evaluated.
- *Component Fragilities* – The fragility of a structural component $P_f(\lambda)$ is defined as the probability of exceedence that the component's response will exceed a specified limit-state when subjected to seismic excitation of PGA equal to λ. Specifying the limit-state or defining what constitutes performance criteria is a key step in evaluating the component fragilities. The performance criteria are determined from typical failure modes and sequence of events that lead to the overall catastrophic failure.
- *Plant Systems and Accident Sequences* – The constituent component fragilities are used to evaluate the overall system-level fragility $P_{f,SYS}(\lambda)$ based on initiating events, failure sequence and event-tree models which carefully consider all possibilities leading to radiological release. The overall risk P_F of the structure can be determined as the convolution of seismic hazard $H(\lambda)$ and the system fragility curve $P_{f,SYS}(\lambda)$ as

$$P_F = \int_0^\infty H(\lambda) \frac{dP_{f,SYS}(\lambda)}{d\lambda} d\lambda \qquad (1)$$

- *Consequence Analyses* – Lastly, a consequence analysis is carried out to estimate the total impact of a seismically induced reactor accident in terms of property damage, economic loss and fatalities. The final output of the consequence analysis is a family

of risk curves which plots the frequency of exceedence per reactor-year vs. the number of fatalities (USNRC, 1983).

1.2 Seismic design criteria

Consistent with the PRA framework of USNRC (1983), ASCE (2005) specified the seismic design criteria for nuclear power plant structures based on a Performance-Goal based (Risk Informed) approach. ASCE (2005) enumerates five different Seismic Design Categories (SDCs). The SDCs range from 1 for conventional buildings to 5 for critical facilities like nuclear reactors. Additionally, four design limit states were also defined - Limit State A ("large permanent distortion for near-collapse conditions") to Limit State D ("essentially elastic behavior"). Each SDC has a target performance goal P_F for any structure being designed in that particular SDC. For example, $P_F = 1 \times 10^{-5}/yr$ for SDC – 5D implies that for any structural design that belongs to Seismic Design Category 5 and failure Limit State D, the overall seismic risk evaluated using Equation (1) cannot exceed the performance goal of $P_F = 1 \times 10^{-5}/yr$. In addition to the overall performance goal P_F, ASCE (2005) also prescribes that the structure must have sufficient levels of conservatism in form of minimum seismic margin factors $F_{p\%}$ such that the following probability goals are reasonably achieved:

1. Less than About a 1% Probability of Unacceptable Performance for the Design Basis Earthquake (DBE) Ground Motion ; and

2. Less than About a 10% Probability of Unacceptable Performance for a Ground Motion Equal to 150% of the Design Basis Earthquake (DBE) Ground Motion

The seismic margin factor, $F_{p\%}$ is defined as a ratio of the PGA corresponding to the *p%* probability of failure on the fragility curve to the PGA of the Design Basis Earthquake (DBE) as per Equation (2):

$$F_{p\%} = \frac{P_f^{-1}(p\%)}{DBE} \tag{2}$$

where $P_f(\lambda)$ is the fragility curve of the structure with respect to a specified limit state. From the above probability goals, all Structures, Systems and Components (SSC) designed according to ASCE (2005) criteria must demonstrate seismic margin factors - $F_{1\%}$ = 1.0 and $F_{10\%}$ = 1.5 corresponding to the "Unacceptable Performance" defined as the "onset of significant inelastic deformation" limit state.

USNRC (2010) and USNRC (1993) prescribe an additional plant-level seismic margin by adopting risk-informed acceptance criteria for all New Standard Plant designs. According to USNRC (1993), it is required that all new plant designs must demonstrate that the seismic margin corresponding to 1% probability of failure on "Core Damage" fragility curve must be at least 1.67 i.e. $F_{1\%}$ = 1.67. At this point, a distinction must be made between the seismic margins prescribed by ASCE (2005) - $F_{1\%}$ = 1.0 and USNRC (1993) - $F_{1\%}$ = 1.67 which are different from each other owing to the fact that they correspond to different design

limit states and different fragility curves altogether. The ASCE (2005) seismic margin factors $F_{1\%} = 1.0$; $F_{10\%} = 1.5$ pertain to the component-level fragilities for a limit-state characterized as "onset of significant inelastic deformation" limit state whereas the USNRC (1993) seismic margin factor should be evaluated from a "Core Damage" plant-level fragility curve. Kennedy (2007) argues that ASCE (2005) values of component-level seismic margins could be substantially too low as a design criteria for achieving the more stringent seismic margin criteria laid down by USNRC (1993). For this reason, USNRC (2001) adopted a component-level design seismic margins for the "Core Damage" limit state as $F_{1\%} = 1.67$ to be consistent with the overriding evaluation criteria of USNRC (1993). Using the overall target performance goal P_F specifications, Kennedy (2007) presented a detailed discussion on the risk-informed approach for establishing the design ground motion or the Site Specific Response Spectrum (SSRS) corresponding to the Safe Shutdown Earthquake (SSE) for future nuclear plant designs defined on the basis of ASCE (2005) Standard.

In summary, the acceptability of seismic design for a NPP structure depends on whether it satisfies the criteria in terms of Target Performance Goal P_F corresponding to its Seismic Design Category (SDC) and whether the design has sufficient "Core Damage" seismic margin factor $F_{1\%}$. From Equations (1) and (2), it can be inferred that both the quantities P_F and $F_{1\%}$ are evaluated from the overall system-level fragility curve, $P_{f,SYS}(\lambda)$ which in turn depends on the fragilities of individual components. A careful examination of the PSRA framework and the seismic design criteria indicates that the evaluation of

structural fragilities is one of most important and influential steps in the entire framework. Moreover, the minimum seismic margins such as $F_{1\%}$ = 1.67 laid down by USNRC (1993) necessitates that the fragility of NPP structures be evaluated with adequate accuracy especially in the low-probability (tail) region of the fragility curve.

In this research, we propose development of a comprehensive framework to accurately evaluate the system-level seismic fragility of any structural system in a NPP. The framework aims to integrate existing approaches and results from probabilistic finite-element simulations using innovative statistical inference techniques to better facilitate establishing the USNRC (1993) criteria of $F_{1\%}$ = 1.67 for NPP structures. The next section discusses some of the existing approaches in fragility analysis of structures and elaborates on the specific objectives of the proposed research.

2. STRUCTURAL FRAGILITY

2.1 Definition

Seismic structural fragility $P_f(\lambda)$ is defined as the probability of a structure to attain or exceed a specified damage state $G(.)$ at a given measure of ground motion intensity λ e.g. Peak Ground Acceleration (PGA) or Spectral Acceleration (S_a) etc. Mathematically, it is often expressed as

$$P_f(\lambda) = P[G(.) < 0 \mid \lambda] \tag{3}$$

where the damage state $G(\overline{\Theta})$ is function of the vector $\overline{\Theta} = [\theta_1, \theta_2 ... \theta_k]$ where $\theta_1, \theta_2 ... \theta_k$ are the random variables representing various material, modeling and loading uncertainties for the structure. In its most simplistic form, the damage state $G(.)$ is often described by a function of at least two variables as:

$$G(C, D) = C - D \tag{4}$$

where C represents the capacity of the structure corresponding to the specified damage condition and D denotes response demand on the structure at a given ground motion intensity λ. In this case, the fragility is

$$P_f(\lambda) = P[C - D < 0 \mid \lambda] \tag{5}$$

The capacity C (or "Strength") of a structure may be defined as a limit seismic load that the structure can endure before the specified damage occurs. It is typically evaluated from material properties and other strength parameters like compression strength, yield stress, design code equations etc. The demand, D (or "Load") at any given ground motion intensity λ is a function of modeling and analysis variables like finite element models, ground motion histories, damping, soil-structure interaction etc. In a seismic fragility analysis, Equation (5) has to be evaluated by accounting for uncertainty in all the significant capacity and demand variables for the structure.

2.2 Log-normal model based on design factors of safety

The fragility $P_f(\lambda)$ is commonly interpreted as being a cumulative distribution function for the structural acceleration capacity. Current state of practice in the nuclear industry (EPRI, 1994) considers that the acceleration capacity, *A* of a structure can be modeled as follows:

$$A = a_m \cdot \varepsilon_U \cdot \varepsilon_R \tag{6}$$

where a_m is the median acceleration capacity of the structure and ε_R denotes the inherent randomness of the capacity about the median and ε_R denotes the uncertainty in our knowledge of the median capacity itself. The ε_R and ε_U are modeled as log-normally distributed random variables with a median of unity and log-standard deviations of β_R and β_U respectively. Kennedy et al. (1980) have derived the expression for a fragility curve corresponding to a confidence level α as follows:

$$P_f(\lambda) = \Phi\left(\frac{\ln\left(\lambda / a_m\right) + \beta_U \Phi^{-1}(\alpha)}{\beta_R}\right) \tag{7}$$

The equation (7) is used to construct confidence bands of fragility curves. In this case, the median fragility curve is given by

$$P_f(\lambda) = \Phi\left(\frac{\ln\left(\lambda / a_m\right)}{\sqrt{\beta_R{}^2 + \beta_U{}^2}}\right) \tag{8}$$

Furthermore, Kennedy et al. (1980) simplified the fragility estimation by expressing the acceleration capacity A as a product of safety margin factors and the PGA of the design earthquake, λ_{design}:

$$A = \left(\prod F_i\right) \cdot \lambda_{design} \tag{9}$$

The various F_i represent various safety margin factors accounting for inherent conservatism in strength and uncertainty in various design variables like capacity, ground motion, damping, soil-structure interaction, modal combination, response spectrum, modeling etc. Each F_i is characterized as a lognormal variable similar to that of Equation (6) with a median value of $\hat{F}_i$ and log-standard deviations β_R^i and β_U^i. Thus, the parameters for the fragility curve of Equation (5) i.e. a_m, β_R and β_U can be finally given by

$$\begin{aligned} a_m &= \left(\prod \hat{F}_i\right) \cdot \lambda_{design} \\ \beta_R &= \sqrt{\sum \beta_R^{i\,2}} \; ; \; \beta_U = \sqrt{\sum \beta_U^{i\,2}} \end{aligned} \tag{10}$$

This model provides a quick heuristic estimate of seismic fragility of a structure that is straightforward to compute. Determining the values of $\hat{F}_i$, β_R^i and β_U^i in this model often requires expert judgment and professional experience data (Kennedy et al, 1980) which may not be available for all kinds of structures. Also, this approach often yields conservative

estimates thereby leading to wide confidence bands associated with the median fragility curves.

In order to minimize the width of the confidence bands in the tail region, we need more fragility data. In such cases, many researchers have used experimental test results, empirical damage data, incremental dynamic analyses and numerical simulations to develop sophisticated structural fragility models for a various kinds of structures. Shinozuka et al (2000) utilized bridge damage data from 1995 Kobe earthquake to estimate the parameters of the lognormal fragility model using Maximum Likelihood Estimation (MLE) methods. Zentner (2010) used a similar approach to obtain the parameters for the fragility of a reactor coolant system from numerical simulations. Serder and Polat (2006) used incremental dynamic analyses to develop capacity models for the mid-rise R/C frame buildings and suggested regression relationships between fragility parameters and the number of storeys in the building. Carausu and Vulpe (1996) proposed a fragility estimation method for seismically isolated NPP structures by deriving bi-linear regression relationships using data on input Peak Ground Velocity (PGV) values and maximum displacements of the isolation layer. Ramamoorthy et al (2006), Do-Eun et al (2008) and Huang et al (2010) etc. suggested developing complex probabilistic models for predicting the strengths ("Capacity") and responses ("Demand") for a variety of reinforced concrete structural systems using experimental data and finite element model simulations. The above methods for direct statistical estimation of the fragility parameters can be quite effective for evaluating

component level fragilities where experimental and damage data is abundantly available. But, in the SPRA framework where the system-level fragilities are of prime interest, the lack of any meaningful quantity of experience data and high costs involved in conducting dynamic experiments for collecting fragility data on large-scale structural systems pose a limitation in directly applying these existing approaches. For large systems, one way to estimate system-level fragilities is to use nonlinear time history analyses of an experimentally-validated full-scale Finite Element model of the entire system. A Monte-Carlo simulation is typically implemented to account for various kinds of uncertainties in material, modeling and ground motion input variables.

2.3 Fragility using Monte-Carlo simulations

Availability of computational resources can enable us to evaluate the fragility functions within a simulation-based framework until additional experience or empirical data becomes available. Monte-Carlo based fragility estimation appears quite straightforward as it involves performing a large number of multiple dynamic analyses for the structure adequately accounting for various uncertainties in the material, modeling and loading parameters. The fragility curve can be generated by evaluating the probability of failure at different levels of ground motion intensities - $\lambda_1, \lambda_2, \ldots \lambda_k$ using the expression in Equation (3):

$$P_f\left(\lambda_j\right) = P\left[G(.) < 0 \mid \lambda_j\right] \cong \frac{\sum 1\left[G(.) < 0 \mid \lambda_j\right]}{n_j};\ n_j >> 0;\ j = 1...k \qquad (11)$$

where 1(.) is the indicator function and n_j is number of simulations conducted at ground motion intensity λ_j. The *k* levels of ground motion intensities $\lambda_1, \lambda_2, ... \lambda_k$ must be carefully chosen so as to be able to predict the complete nature of the fragility function satisfactorily. If n_j is sufficiently large, the fragility $P_f(\lambda)$ can be simply obtained by computing the ratio of the number of times a particular damage state *G*(.) has been attained by the structure to the total number of simulations, n_j.

The accuracy of the Monte-Carlo simulation approach depends entirely on two variables: (i) the vector of PGA values, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ... \lambda_k]^T$; (ii) the vector $\mathbf{N} = [n_1, n_2, ... n_k]^T$ where n_j denotes the number of time-history analyses performed at each PGA λ_j for *j = 1, ... ,k*. Therefore in real-life structural fragility estimation, a trade-off must be made between the required accuracy of the fragility estimates and the total number of analyses $N_{sim} = \sum_{j=1}^{k} n_j$ that can be performed efficiently. For simple structural systems with a small number of Degrees of Freedoms (DOFs), performing a large number of multiple time history analyses does not impose any serious constraint. However, for a large non-linear structural system with many DOFs, each time-history analysis may demand significant computational resources.

To reduce the total number of time history analyses required, it is sometimes practical to use conventional EPRI (1994) approach or other existing studies to obtain a preliminary

estimate of fragility curves for the structure and then use limited time-history simulation data to update the fragility curves. The Bayesian Inference techniques are a popular and powerful tool to effectively integrate existing information or "prior knowledge" on the random parameters in the fragility model with "newly collected information" from the time-history simulations in order to obtain the updated or "posterior" fragility curves. The next section introduces the Bayesian updating approach and discusses how it can be used to update the random parameters in fragility models.

3. BAYESIAN INFERENCE

Suppose that the structural fragility model $P_f(\lambda)$ depends on a vector of random variables $\boldsymbol{\Theta} = [\theta_1, \theta_2 ....]^T$ such that $P_f(\lambda) = f(\lambda; \boldsymbol{\Theta})$. Also assume that our current knowledge of $\boldsymbol{\Theta}$ can be represented by a joint density function - $f'_{\boldsymbol{\Theta}}(\theta_1, \theta_2 ...)$. This $f'_{\boldsymbol{\Theta}}(\theta_1, \theta_2 ...)$ is often referred to as "prior" information and is generally obtained from professional knowledge of experts, past studies etc. Suppose a vector of *m* observations - $\mathbf{y} = [y_1, y_2 ... y_m]^T$ of structural fragility data have been collected. The Bayes Theorem can then be used to update our knowledge of the random parameters using the observed data **y** to obtain the "posterior" information $f''_{\boldsymbol{\Theta}}(\theta_1, \theta_2 ... | \mathbf{y})$ as

$$f''_{\boldsymbol{\Theta}}(\theta_1,\theta_2...|\mathbf{y}) = \frac{P(\theta_1,\theta_2...|\mathbf{y})\, f'_{\boldsymbol{\Theta}}(\theta_1,\theta_2...)}{P(\mathbf{y})}$$
$$P(\mathbf{y}) = E_{\boldsymbol{\Theta}}\left[P(\theta_1,\theta_2...|\mathbf{y})\right] = \int P(\theta_1,\theta_2...|\mathbf{y})\, f'_{\boldsymbol{\Theta}}(\theta_1,\theta_2...)d\boldsymbol{\Theta} \quad (12)$$

In Equation (12), $P(\theta_1,\theta_2...|\mathbf{y})$ is often referred to as the "Likelihood function" of the random parameters. The integrations involved in the evaluation of $P(\mathbf{y})$ can be quite formidable to be evaluated by even numerical techniques especially for higher-order $\boldsymbol{\Theta}$. So, the posterior statistics $f''_{\boldsymbol{\Theta}}(\theta_1,\theta_2...|\mathbf{y})$ are typically computed by numerically generating a large number of posterior samples employing a special class of computational algorithms called Markov Chain Monte Carlo (MCMC) methods (Winkler, 2003; Congdon, 2006).

Bayesian methods provide the flexibility to incorporate different types of data (experimental, empirical, numerical simulations, etc.) into a single fragility model and help reduce the variability that might otherwise exist in predicted values due to limited data. A detailed updating procedure is illustrated in which we start with the initial fragility estimates based on the conventional lognormal approach of Section 2.2 as "prior" information and then use the results from the time-history analyses of an experimentally-validated FE model to update the fragility curves based on Equation (12). Another advantage of the Bayesian inference techniques is that they can be instrumental in optimizing the number of time history analyses required as well as the region of PGA values in which to conduct the time history analyses. Instead of performing an arbitrary number of time history analyses at various levels of PGAs, one can start with a relatively small number of time history observations and infer from the

posterior statistics as to how many more observations might be necessary to achieve specified convergence criterion. While studies like Hamada et al (2008) discuss methods based on Genetic Algorithms (GA) to plan for reliability data collection, no explicit planning approach exists for conducting time history analyses and design of Monte-Carlo experiments in structural fragility applications. A comprehensive methodology for a Bayesian statistics based planning of the Monte-Carlo simulations is also presented for reducing the total computational effort involved.

4. RESEARCH OBJECTIVES

The primary objective of this research is to develop a comprehensive framework for evaluating seismic fragilities for large-scale structural systems as well as updating them based on the availability of additional fragility data. The key aspects desired in the proposed framework are:

- The framework should be able to predict the structural fragility estimates to desired limits of accuracy with a high degree of confidence.
- The framework must aim to reduce the total number of nonlinear time history analyses required for a specified degree of accuracy in the fragility curve.
- The framework should be robust enough to account for uncertainties in complex structures where interdependence among various physical variables (structural and material) must be accounted for.

5. PROPOSED RESEARCH

The specific tasks needed to accomplish the objectives of this research can be characterized in the following sub-sections:

5.1 Evaluate the limitations of existing approaches for evaluating structural fragility

- Study the current framework of the Seismic Probabilistic Risk Assessment (SPRA) in the nuclear industry (Kennedy, 2007) Demonstrate the conventional approach of structural fragility calculation with the lognormal model based on design factors of safety (EPRI, 1994) using an example of a Box-Shaped reinforced concrete shear wall. The purpose of this exercise is to determine if the existing approach yields structural fragility estimates with sufficient degree of confidence required to establish the USNRC (1993) criteria.
- Stipulate maximum acceptable value for width of the 95% confidence bands as an indicator of the level of the accuracy desired.
- Identify the limitations of the conventional approach:
    - Relies on expert judgment and professional experience independent of the specific design of a structural system
    - The resultant fragility curves tend to be very sensitive to the industry-recommended design factors and are quite conservative with a high degree of variability, thereby possessing large 95% confidence interval widths.

- No procedure exists to update fragility or reduce confidence bands using additionally available data

5.2 Development of a Bayesian framework for updating structural fragility

System level fragilities are often calculated by a Monte-Carlo simulation in which the experimentally validated finite element model of one or more components is incorporated into a system-level FE model. The non-linear large-scale system level model is then used to conduct multiple time history analyses within a Monte-Carlo framework to evaluate system-level fragilities. The excessive computational time needed in conducting a single time history analysis renders this framework impractical if many different earthquake input motions are considered in the fragility evaluation. A Bayesian framework can be used quite effectively to determine the number of time history simulations needed in order to achieve a desired degree of accuracy. Specific tasks in such a framework can be enumerated as:

- Determine the parameters associated with the fragility curve assuming the conventional lognormal fragility model from the EPRI (1994) approach or any other existing method. The prior information on the random parameters of the model - median capacity and the log-standard deviation must be specified based on this preliminary fragility analysis.
- Conduct a relatively small number of time-history analyses of the experimentally-validated FE model.

- From the responses of each time history analysis at a given PGA, determine if the structure has exceeded the specified limit state. Consider that the structure has "failed" if the maximum responses fr**o**m the analysis exceed the specified limit state.
- Using the entire failure data collected above, determine the "likelihood" function of the random parameters of the lognormal model.
- Evaluate the "posterior" statistics of the random parameters - median capacity and the log-standard deviation, from the "prior" fragility and the "likelihood" functions employing the Bayes Theorem using Markov Chain Monte Carlo methods.
- Use the posterior statistics to compute the median fragility curve and the widths of the associated 95% confidence bands. If the 95% confidence bands do not satisfy the specified convergence criteria, more simulations may be needed.
- In order to minimize the total number of time history analyses required for convergence, an optimal plan for the Monte-Carlo simulation may be effective compared to conducting an arbitrary number of time history analyses at various levels of PGA. It is proposed to develop such a plan based on Bayesian Inference and Markov Chain Monte Carlo methods.

5.3 Simulation of constrained, closely-spaced random variables in engineering risk analyses

Any Monte-Carlo simulation involving a large-scale structure needs to account for uncertainties in several input variables simultaneously. Under certain circumstances, the value of an input variable could be dependent on or constrained by the values of other input

variables. In such cases, care must be taken to accommodate these constraints while generating random sample sets of input variables for time history analyses. Consequently, we need a sampling technique to generate sets of random variables so that they may be: (a) characterized by a continuous Probability Density Function (pdf); (b) subject to a non-linear constraint. The various steps for achieving this objective are outlined below:

- Illustrate the limitations associated with ignoring the constraints between the input parameters of a structure in a Monte-Carlo simulation. For example, closely-spaced natural frequencies of a structure with the constraint that a simulated set of natural frequencies must always be an ordered set. Sampling of closely-spaced frequencies as independent random variables may violate this constraint.
- Evaluate if simply rejecting the disordered or incorrect sets is an acceptable solution to the above problem. This can be verified by comparing the shapes of the simulated density functions with the expected distributions. If the overlap between the pdfs is significant, rejecting the improper sets may lead to considerable bias.
- Establish if using the explicit correlation coefficients can resolve the problem of disordered/incorrect sampling.
- Evaluate the possibility of deriving an analytical solution from basic probability principles.

- If an analytical solution is not possible, suggest a numerical procedure to obtain a joint probability density function surface that can be used to the sample closely-spaced constrained random-variables.

5.4 A Bayesian model for large-scale piping systems based on equivalent linear analyses

Sometimes, the fragility of certain structural systems which exhibit predominantly elastic deformations is often dictated by localized non-linearity present in connections and joints within the system. Typical examples include connections in piping systems of a nuclear power plant or a hospital where the failures of T-Joints and/or bracing members often lead to the damage of the piping system. To evaluate the seismic fragility in such situations, the conventional EPRI (1994) approach to obtain the "prior" fragility curve cannot be implemented directly and tends to suffer from lack of sufficient information. In this section, we propose a method to obtain improved estimates of "prior" fragility by conducting multiple linear time-history analyses of an equivalent linear system. The use of linear time-history analyses to estimate the prior information has two distinct advantages:

(i) For any system with predominantly linear behavior, equivalent linearization of localized nonlinearities can lead to system fragilities that are quite close to the fragility that may be evaluated from multiple non-linear analyses. Therefore, this technique can be used for obtaining a good estimate of the prior fragility curve which can then be updated in a Bayesian framework.

(ii) The time-history response of a linear system at any PGA can be obtained by scaling the responses of the system calculated for a unit PGA case. The total computational effort can therefore be minimized as only a set of elastic responses at a single PGA value are needed for computing the complete fragility curve.

Specific tasks needed to implement the proposed method are:

- Study the applicability of existing equivalent linearization techniques to large scale piping systems. Identify the chief limitation(s) of extending the approach to such systems.
- Introduce the concept of equivalent linear limit state: for a fragility curve of a nonlinear system given by $P_f^{NL}(\lambda) = P[G(.) > 0]$, determine an equivalent linear limit state, $G^{lin}(.)$ corresponding to the actual limit state $G(.)$ such that fragility evaluated from using the equivalent linearized system - $P_f^{lin}(\lambda) = P[G^{lin}(.) > 0]$ is close to $P_f^{NL}(\lambda)$.
- Perform an exhaustive set of numerical simulations using representative systems to propose general guidelines on evaluating $G^{lin}(.)$ as a function of structural properties and the desired limit state $G(.)$
- Incorporate the calculated prior $P_f^{lin}(\lambda)$ based on $G^{lin}(.)$ into the generalized Bayesian framework prposed in Section 5.2.

6. ORGANIZATION

This dissertation consists primarily of three manuscripts which will be submitted for possible publications in peer reviewed journals. The first manuscript, Part II of the dissertation, presents the Bayesian statistical framework for seismic fragility evaluation of structures. The various features of the framework are illustrated using a detailed fragility analysis of a Reinforced Concrete (RC) box-shaped shear wall. The second manuscript in this dissertation, Part III, discusses the evaluation of seismic fragility of nonlinear connections in large piping systems using equivalent linearization techniques. The third manuscript in this dissertation, Part IV, addresses the problem of sampling closely-spaced random variables subjected to inequality constraints for engineering risk analysis applications. Part V of this dissertation gives a summary and conclusions of this research and the anticipated future research on these topics.

REFERENCES

ASCE/SEI 43-05, "Seismic Design Criteria for Structures, Systems, and Components in Nuclear Facilities," American Society of Civil Engineers and Structural Engineering Institute, 2005.

Carausu, A. and Vulpe, A. "Fragility estimation for seismically isolated nuclear structures by high confidence low probability of failure values and bi-linear regression," Nuclear Engineering and Design, 160 p. 287-297, 1996.

Congdon, P., "Bayesian Statistical Modelling," Wiley Series in Probability and Statistics, 2nd Edition, 2006.

Do-Eun C., Gardoni, P., Rosowsky, D., Haukaas, T., "Probabilistic capacity models and seismic fragility estimates for RC columns subject to corrosion", Reliability Engineering and System Safety, v 93, n 3, p 383-93, 2008

Electric Power Research Institute (EPRI), "Methodology for Developing Seismic Fragilities", EPRI TR-103959 (Principal Investigators: Reed, J.W. and Kennedy, R.P.), Electric Power Research Institute, Palo Alto, California, 1994.

Hamada, M.S., Wilson, A., Reese, C. S., Martz, H., "Bayesian Reliability," Springer Series in Statistics, 2008.

Huang, Q. Gardoni, P., Hurlebaus, S., "Probabilistic seismic demand models and fragility estimates for reinforced concrete highway bridges with one single-column bent," Journal of Engineering Mechanics, v 136, n 11, p 1340-1353, 2010.

Kennedy, R.P., "Performance-Goal Based (Risk Informed) Approach for Establishing the SSE Site Specific Response Spectrum for Future Nuclear Power Plants," *Jaeger Lecture*, SMIRT-19, Toronto, Canada, 2007.

Kennedy, R.P., Cornell, C.A., Campbell, R.D., Kaplan, S., Perla, H.F., "Probabilistic seismic safety of an existing nuclear power plant," Nuclear Engineering and Design, 59, 315–338, 1980.

Petelet M., Iooss B., Asserin O., Loredo A., "Latin hypercube sampling with inequality constraints," *Advances in Statistical Analysis*, submitted. Available at: http://hal.archives-ouvertes.fr/docs/00/52/00/21/PDF/AStA10_iooss_rev2.pdf, 2010.

Ramamoorthy, S.K., Gardoni, P., Bracci, J.M., "Probabilistic demand models and fragility curves for reinforced concrete frames," v 132, n 10, p 1563-1572, 2006

Serdar K.M., Polat, Z., "Fragility analysis of mid-rise R/C frame buildings," Engineering Structures, v 28, n 9, p 1335-1345, 2006.

Shinozuka, M., Feng, M.Q., Lee, J., Naganuma, "Statistical analysis of fragility curves," Journal of Engineering Mechanics, v 126, n 12, p 1224-1231, 2000.

Tadinada S.K., "Consideration of Uncertainties in Seismic Analysis of Non-classically damped coupled systems," Master's Thesis, North Carolina State University, Available at: http://repository.lib.ncsu.edu/ir/bitstream/1840.16/2134/1/etd.pdf , 2009.

USNRC, "PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants," NUREG/CR-2300, 1983.

USNRC, "Policy, Technical, and Licensing Issues Pertaining to Evolutionary and Advanced Light-Water Reactor (ALWR) Designs, (Item II Q)," SECY-93-087, 1993.

USNRC, "Technical Basis for Revision of Regulatory Guidance on Design Ground Motions: Harzard- and Risk- Consistent Ground Motion Spectra Guidelines," NUREG/CR-6728, 2001.

USNRC, "Interim Staff Guidance on Implementation of a Probabilistic Risk Assessment-Based Seismic Margin Analysis for New Reactors," DC/COL-ISG-020, 2010.

Winkler, R.P., "An Introduction to Bayesian Inference and Decision," Probabilistic Publishing, 2nd edition, 2003.

Zentner, I., "Numerical computation of fragility curves for NPP equipment," Nuclear Engineering and Design 240, 1614–1621, 2010.

PART II

A BAYESIAN FRAMEWORK FOR ESTIMATING AND UPDATING SEISMIC FRAGILITY OF STRUCTURES

Sashi Kanth Tadinada and Abhinav Gupta

To be submitted to Computers and Structures

ABSTRACT

The fundamental goal in the seismic design of any Nuclear Power Plant (NPP) is the minimization of radiological risk (release of radioactive emissions into the environment) and ensuring the operational safety immediately following an earthquake. It is therefore, critical that engineers be able to estimate the associated risks due to seismic events to a high degree of confidence. In the nuclear industry, this is done using a detailed probabilistic framework that carefully accounts for all possible uncertainties in the expected intensity of ground motions, the predicted response of the structures and ultimately, the chain of events that could lead to the catastrophic failure. For real-life complex systems with non-standard designs, one realistic way of estimating seismic risk of the system is to use a robust experimentally-validated Finite-Element (FE) simulation model and perform multiple dynamic analyses by considering uncertainties in material, modeling and loading variables. But this Probabilistic approach to seismic risk estimation has one important drawback - it requires a large number of computationally intensive simulations to be performed for accurate estimates of risk. Typically, this demands for a trade-off between the desired accuracy of the risk estimates and the total number of simulations that can be performed in a reasonable time.

In this work, we present a comprehensive statistical framework based on Bayesian inference techniques to minimize the total computational effort required in conducting Finite Element

(FE) simulations for seismic risk assessment of large-scale engineering structures. The salient features of the framework are:

a. Using Bayesian Inference to flexibly combine preliminary risk data from a variety of information sources like experiments, existing studies or simplified approaches with data from FE simulations thus reducing the overall dependence on simulation data alone;
b. Embedding Bayesian methods within a stochastic optimization environment to help forecast, plan and allocate adequate computing resources such that the desired accuracy is achieved using minimum possible number of simulations.

We illustrate the application of this framework using the example of a Box-Shaped Reinforced Concrete shear wall in a nuclear reactor. The framework is shown to have wide applicability and it generalizes well for a diverse range of engineering risk analysis problems.

1. BACKGROUND

Evaluation of seismic fragilities for structural systems is a key step in the Seismic Probabilistic Risk Analysis (SPRA) of a Nuclear Power Plant (NPP). In the SPRA framework, a fragility curve describes the relationship between a ground motion intensity parameter like Peak Ground Acceleration (PGA) and the corresponding probability of failure for specified performance criteria.

The widely used approach to evaluate structural fragilities in the nuclear industry is the lognormal model based on design factors of safety (Kennedy et al, 1980; EPRI, 1994) which assumes a lognormal distribution for the acceleration capacity of the structure. USNRC (1993) has adopted a risk-informed acceptance criteria for new plant designs which requires that all designs must demonstrate a safety factor of 1.67 over the 1% probability of failure ($F_{1\%}$) on the core damage fragility curve i.e. the probability of core damage in any NPP should be less than 1% in the event of an earthquake whose PGA is 1.67 times the design basis level (Kennedy, 2007). Such a criterion necessitates that fragility of nuclear power plant structures be determined with adequate accuracy especially in low probability regions (tail) of the fragility curve. However, the conventional EPRI (1994) approach relies heavily on data from past studies and expert judgment which may not always be available for newer structural systems being designed for future NPPs. Moreover, the logarithmic standard deviations of safety factors in the model as recommended by the industry professionals are

generally conservative leading to a wide band of confidence intervals for the median fragility curve and consequently, greater variability in the fragility estimates.

In order to minimize the width of the confidence bands in the tail region, we need to acquire additional fragility data. In such cases, many researchers use experimental test results, empirical damage data, incremental dynamic analyses and numerical simulations to develop sophisticated structural fragility models. Shinozuka et al (2000) utilized bridge damage data from 1995 Kobe earthquake to estimate the parameters of the lognormal fragility model using Maximum Likelihood Estimation (MLE) methods. Zentner (2010) used likelihood estimation methods to obtain the parameters for the fragility of a reactor coolant system from numerical simulation data. Kircil and Polat (2006) used incremental dynamic analyses to develop capacity models for the mid-rise Reinforced Concrete frame buildings and suggested regression relationships between fragility parameters and the number of storeys in the building. Carausu and Vulpe (1996) proposed a fragility estimation method for seismically isolated NPP structures by deriving bi-linear regression relationship using data on input Peak Ground Velocity (PGV) values and maximum displacements of the isolation layer. Mehrdad and Der Kiureghian (2001), Ramamoorthy et al (2006), Choe et al (2008) and Huang et al (2010) etc. suggested developing complex probabilistic models for predicting the strengths ("Capacity") and responses ("Demand") for a variety of reinforced concrete structural systems using experimental data and finite element simulations. These methods for direct statistical estimation of the fragility parameters can be quite effective for evaluating

component level fragilities where experimental and field data is abundant. However in the SPRA framework, it is the system-level fragilities of large structural systems that are of prime interest. Experimental testing for fragility data on large scale systems in most cases is uneconomical and impractical thereby limiting the amount of data that can be collected. Moreover, the fragility models at the component-level cannot always be extended in a straight-forward manner to the system level (Shinozuka et al, 2000). The system-fragility is typically determined from the constituent component fragilities using event and fault tree models which consider the failure of a component as an independent event. However, seismic behavior of large-scale structural systems depends significantly on the dynamic interactions between the sub-systems. The dynamic interaction effects cannot be considered in a fault- and event-tree based system reliability approach.

For large structural systems, a generally used approach for computing structural fragility is to use probabilistic simulations involving multiple time history analyses of an experimentally validated Finite Element (FE) model of the structure (Komura et al, 1989; Shinozuka and Kishimoto, 1989; Garcia and Soong, 2003). In this approach, the structural properties of the individual components and their failure mechanisms are studied in detail through careful testing. The validated models are then integrated into the large-scale finite element model of the complete system. This approach allows for consideration of uncertainties in the physical variables, modeling nonlinearities, inter-dependencies between various structural parameters and stochastic models for earthquake excitations. Multiple time

history analyses are conducted at various levels of PGA to evaluate the fragility values. An important drawback of this approach is that it requires a large number of dynamic analyses to accurately predict fragility in the low-probability region which can be highly inefficient and computationally prohibitive. Thus, a trade-off exists between the required accuracy of the fragility estimates and the total number of analyses that can be completed in a reasonable manner.

To reduce the total number of time history analyses required, it is sometimes practical to use conventional or existing approaches to derive preliminary estimates of fragility curves for the structure and then use simulation data from the time history analyses to update the fragility curves as per desired accuracy. In this paper, we propose a framework based on Bayesian Inference techniques to update preliminary estimates of the fragility curves and reduce the width of the confidence bands associated with them. The purpose is to improve the confidence in establishing the USNRC specified criteria of $F_{1\%} = 1.67$ mentioned earlier. Bayesian Inference is a powerful statistical technique to combine existing or "prior" knowledge with newly available data for improving the process of generating additional data.

A number of researchers have applied Bayesian techniques to address structural fragility problems. Gardoni et al (2002) presented a Bayesian framework to develop multi-variate probabilistic capacity and demand models and applied it to Reinforced concrete bridge components. Straub and Der Kiureghian (2008) employed Bayesian methods to update the fragility models of electrical sub-station equipment to account for the statistical

dependence between empirical observations of seismic performance. Koursourelakis (2010) suggested a Bayesian framework to develop fragility surfaces simultaneously considering four different ground motion intensity parameters instead of just the PGA and applied it to geotechnical problems. Many other researchers have used Bayesian approaches to update and develop the probabilistic capacity and demand models using experimental data or time history analyses for a variety of structural systems such as Reinforced Concrete (RC) structural walls (Mehrdad and Der Kiureghian, 2001), RC bridge components (Gardoni et al, 2002; Zhong et al, 2009), RC frames (Ramamoorthy et al, 2008), woodframe structures (Pei et al, 2009), power plant equipment (Yuichi et al, 2005), multi-storey buildings (Bai et al, 2011) and so on. A common theme in all these studies is that the availability of additional data is used to update the median fragility curve. In this paper, we illustrate that while a median fragility curve can converge with the availability of only limited additional data, the corresponding confidence bands associated with aleotary uncertainties can be significantly large. This means that accurate evaluation of a design parameter such as $F_{1\%}$ with respect to an acceptable criterion may require much more data.

Bayesian methods provide the flexibility to incorporate different types of data (experimental, empirical, numerical etc.) into a single fragility model and to draw reliable inferences even with limited data. We present a procedure in which we start with the initial fragility estimates based on the conventional lognormal approach (EPRI, 1994) as "prior" information. Next, the results of a relatively small number of time history analyses conducted

at various PGA levels are used to evaluate the updated or "posterior" fragility curves. It is illustrated that the Bayesian framework, even with a relatively small number of time-history analyses data, is quite effective in achieving the convergence criteria. Another advantage of the Bayesian Inference technique is that it can be effective in reducing the number of time history analyses required as well as identifying the region of PGA values in which to conduct the time history analyses. Instead of performing an arbitrary number of time history analyses at various levels of PGAs, one can start with a relatively small number of time history analyses and infer from the posterior statistics as to how many more observations might be necessary to satisfy a specified convergence criterion. The advantages of the proposed framework are illustrated by evaluating the seismic fragility of Reinforced Concrete (RC) Box-shaped shear wall.

2. THE LOG-NORMAL FRAGILITY MODEL

Current state-of-practice in the nuclear industry (Reed and Kennedy, 1994; EPRI, 1994) considers that the acceleration capacity *A* of a structure can be modeled as a random variable:

$$A = a_m \cdot \varepsilon_U \cdot \varepsilon_R \qquad (1)$$

where a_m is the median acceleration capacity of the structure and ε_R is the random variable denoting the inherent randomness of the capacity about the median and the random variable ε_R denotes the uncertainty in our knowledge of the median capacity itself. The ε_R and ε_U

are modeled using the log-normal distribution with a median of unity and log-standard deviations of β_R and β_U respectively. Using the relationship in Equation (1), the expression for the fragility curve corresponding to a confidence level α is given by Kennedy et al. (1980) as follows:

$$P_f(\lambda) = \Phi\left(\frac{\ln\left(\lambda / a_m\right) + \beta_U \Phi^{-1}(\alpha)}{\beta_R}\right) \tag{2}$$

The median fragility curve can be written as

$$P_f(\lambda) = \Phi\left(\frac{\ln\left(\lambda / a_m\right)}{\sqrt{\beta_R{}^2 + \beta_U{}^2}}\right) \tag{3}$$

In order to determine the parameters of the model, Kennedy et al. (1980) expressed the acceleration capacity, *A* as a product of safety margin factors and the design level PGA - λ_{design}:

$$A = \left(\prod F_i\right) \cdot \lambda_{design} \tag{4}$$

The various F_i 's represent different safety margin factors accounting for conservatism and uncertainty in the design variables like capacity, ground motion, damping, soil-structure interaction, modal combination, response spectrum, modeling etc. Each F_i is

itself characterized as a log-normal variable with a median value of $\hat{F}_i$ and its corresponding logarithmic-standard deviations β_R^i and β_U^i. Thus, the parameters for the fragility curve of Equation (2) i.e. a_m, β_R and β_U can be evaluated as:

$$a_m = \left(\prod \hat{F}_i\right) \cdot \lambda_{design}$$
$$\beta_R = \sqrt{\sum {\beta_R^i}^2} \; ; \beta_U = \sqrt{\sum {\beta_U^i}^2} \qquad (5)$$

2.1 Example –Reinforced Concrete box-type shear wall

The structural fragility of a Reinforced Concrete Box-shaped shear wall shown in Figure 1 is calculated using the EPRI (1994) approach. It is desired to evaluate the seismic fragility of this structure subjected to unidirectional earthquake excitations. For brevity, the detailed calculations involved in evaluation of the various safety margin factors and their log-standard deviations - $\hat{F}_i$, β_R^i and β_U^i are not presented. The parameters of the resultant lognormal fragility curve were determined for the shear wall are determined as: $a_m = 1.658g$, $\beta_R = 0.288$ and $\beta_U = 0.265$.

Figure 2 gives the median fragility curve $P_f(\lambda)$ calculated using Equation (3) and the corresponding 90% confidence bands based on Equation (2). It is evident from Figure 2 that the EPRI (1994) approach results in large confidence intervals associated with the median fragility curves. This can be also be seen by examining the width of the 90% confidence interval for the median capacity a_m which evaluates to [0.99*g*, 2.78*g*]. This means that the

value of a_m can lie anywhere between 0.99g and 2.78g with a probability of 90%. One reason for such a conservatism in the fragility estimates is that the values of the safety factors, $\hat{F}_i$, β_R^i and β_U^i are based on capacity equations from design codes, expert judgment and existing professional experience pertaining to typical design practices in the industry.

One direct way to overcome the wide confidence bands of the fragility curves from the EPRI (1994) approach is to estimate fragility from the results of nonlinear seismic analysis of the structure using an experimentally validated Finite Element (FE) model by considering uncertainties in material, modeling and loading variables. But this approach to risk estimation has one important drawback - it requires a large number of simulations to be performed for accurate estimates of risk. This can be a limitation in the case of large FE models which often take significant computational time even for a single analysis. Therefore, a trade-off is often required between the desired accuracy of the risk estimates and the total number of analyses that can be performed in a reasonable time.

2.2 Objective

The primary objective of this study is to develop a comprehensive framework for evaluating system-level structural fragilities as well as updating them based on the availability of additional fragility data. The key aspects desired in the proposed framework are:

- The framework should be able to predict the structural fragility estimates to desired limits of accuracy with a high degree of confidence

- The framework should be capable of considering the effects of various kinds of uncertainties on the seismic response of the structure
- The framework must aim to reduce the total number of nonlinear time history analyses required for a specified degree of accuracy in the fragility curve.

The proposed framework implements Bayesian inference techniques in order to evaluate the seismic fragilities. To minimize the total computational effort, the framework updates the initial fragility curves obtained from the conventional (EPRI, 1994) approach (called "prior fragility") using the data from time history simulations to develop fragility curves of large-scale systems with an intent to achieve relatively much narrower confidence bands. The next section introduces the Bayesian updating approach and discusses how one can implement it to update the fragility curves when new information becomes available.

3. BAYESIAN FRAMEWORK FOR UPDATING STRUCTURAL FRAGILITY

3.1 Bayesian inference

In general, any structural fragility model can be represented as a function of the earthquake intensity λ, i.e. $P_f(\lambda) = f(\lambda;\boldsymbol{\Theta})$ where $\boldsymbol{\Theta} = [\theta_1, \theta_2 ....]^T$ is a vector of random variables that define the shape of the fragility curve. Let us consider that our current knowledge of $\boldsymbol{\Theta}$ be denoted by a joint density function $f'_{\boldsymbol{\Theta}}(\theta_1, \theta_2 ...)$. This joint density function $f'_{\boldsymbol{\Theta}}(\theta_1, \theta_2 ...)$ is often referred to as "prior" information which is obtained using existing data and

professional expertise. Any new data either from experimental or numerical studies can be represented by a vector of observations or data - $\mathbf{y} = [y_1, y_2 ... y_m]^T$. The Bayes Theorem can be used to update our knowledge of the random parameters using the new data $\mathbf{y}$ to obtain the "posterior" information $f''_{\mathbf{\Theta}}(\theta_1, \theta_2 ... | \mathbf{y})$ as

$$f''_{\mathbf{\Theta}}(\theta_1, \theta_2 ... | \mathbf{y}) = \frac{P(\theta_1, \theta_2 ... | \mathbf{y}) f'_{\mathbf{\Theta}}(\theta_1, \theta_2 ...)}{P(\mathbf{y})}$$
$$P(\mathbf{y}) = E_{\mathbf{\Theta}}[P(\theta_1, \theta_2 ... | \mathbf{y})] = \int P(\theta_1, \theta_2 ... | \mathbf{y}) f'_{\mathbf{\Theta}}(\theta_1, \theta_2 ...) d\mathbf{\Theta} \quad (6)$$

In Equation (6), $P(\theta_1, \theta_2 ... | \mathbf{y})$ is often referred to as the "Likelihood function".

3.2 Methodology for updating structural fragility

In this paper, we consider the two-parameter lognormal fragility model (Kennedy et al, 1980) presented in Section 2 given as follows:

$$P_f(\lambda; \tilde{a}, \beta_r) = \Phi\left[\frac{\ln \lambda - \ln \tilde{a}}{\beta_r}\right] \quad (7)$$

where λ denotes the peak ground acceleration (PGA) level and $\mathbf{\Theta} = [\theta_1, \theta_2]^T = [\tilde{a}, \beta_r]^T$ represent the median and logarithmic standard deviation of the acceleration capacity of the structure.

3.2.1 Prior information

To begin with, we assume a distribution for the joint density function $f_{\tilde{A},B_r}(\tilde{a},\beta_r)$ which represents the "current state of knowledge" about the parameters - $\tilde{a}$, β_r . The $f'_{\tilde{A},B_r}(\tilde{a},\beta_r)$ is called the "prior" distribution of the random parameters i.e. $f'_{\boldsymbol{\Theta}}(\theta_1,\theta_2) = f'_{\tilde{A},B_r}(\tilde{a},\beta_r)$ and is usually estimated from existing studies of similar structures, simplified analyses methods or professional experience etc. For the RC box-shaped shear wall example, the prior distributions are assumed from the fragility estimates of the ERPI (1994) approach as evaluated in Section 2.1. The EPRI (1994) method does not directly specify any variability in parameter β_r . So, a uniform distribution within ± 15% of its mean value assumed in this study. The prior distributions for the shear wall fragility are summarized in Equation (8):

$$\begin{aligned} \tilde{a} &\sim \log N\left(median, a_m = 1.658g\ ;\ \beta_u = 0.265\right) \\ \beta_r &\sim Uniform\left(0.2448\ ,\ 0.3312\right) \end{aligned} \tag{8}$$

3.2.2 Likelihood function

A Bayesian framework seeks to update the "state-of-knowledge" of the parameters - $\tilde{a}$, β_r i.e. we obtain a more informed density function as we gather data from experiments or numerical simulations. Typically, the fragility data $\mathbf{y}$ when obtained from experimental tests or numerical simulations can be described as a set of three vectors, as in

$$\mathbf{y} \equiv \left[\begin{Bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{Bmatrix} \quad \begin{Bmatrix} n_1 \\ n_2 \\ \vdots \\ n_k \end{Bmatrix} \quad \begin{Bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{Bmatrix} \right] \tag{9}$$

The data set $\mathbf{y}$ summarizes the number of times r_i that the damage has been observed out of the total number of structural specimens n_i analyzed or tested at a given PGA level λ_i. If the experiments or numerical simulations are conducted at *k* different levels of PGA, the likelihood function $L(\tilde{a}, \beta_r \mid \mathbf{y})$ is written as:

$$\begin{gathered} \mathbf{y} \equiv [y_1, y_2 ... y_k]^T ; y_i = [\text{PGA} = \lambda_i, n_i \text{ analyses}, r_i \text{ failures}] \quad i = 1,2...k \\ L(\tilde{a}, \beta_r \mid \mathbf{y}) \propto \prod_{i=1}^{k} P_f{}'(\lambda_i; \tilde{a}, \beta_r)^{r_i} \cdot \left(1 - P_f{}'(\lambda_i; \tilde{a}, \beta_r)\right)^{n_i - r_i} \end{gathered} \tag{10}$$

For evaluating the RC box-shaped shear wall fragility, 30 acceleration time-histories consistent with the site design response spectra were selected to model the uncertainty in the horizontal ground motion. Additionally, uncertainties in material and modeling properties of the experimentally-validated FE model were considered by configuring the inputs of each time-history analyses according to the assumed distributions. All the 30 acceleration histories were scaled to *k* = 5 different PGA levels $[\lambda_1, \lambda_2, ..., \lambda_5]^T = [0.5g, 1.0g, 1.2g, 1.5g, 2.0g]^T$ to conduct a total of 150 time-history analyses for obtaining the fragility data. In each time-history analysis, the maximum base shear was recorded and compared with the ASCE 43-05 design base shear at the corresponding PGA level. The structure is assumed to have been

"damaged" or "failed" if the maximum base shear from a time-history analysis at a given PGA exceeds the corresponding design base shear. The number of failures recorded out of a total of 30 analyses at each of the 5 respective PGA levels is summarized in a vector form as $[r_1, r_2, ..., r_5]^T = [0, 17, 26, 29, 30]^T$. With this information, we can now write the likelihood function based on Equation (10).

3.2.3 Posterior distribution

The likelihood function evaluated using Equation (10) and the prior density function $f'_{\tilde{A},B_r}(\tilde{a}, \beta_r)$ can be used to determine the posterior distribution $f''_{\tilde{A},B_r}(\tilde{a}, \beta_r)$ of the random variables - $\tilde{a}, \beta_r$ as

$$f''_{\tilde{A},B_r}(\tilde{a}, \beta_r \mid \mathbf{y}) = \frac{L(\tilde{a}, \beta_r \mid \mathbf{y}) \times f'_{\tilde{A},B_r}(\tilde{a}, \beta_r)}{\iint \left(L(\tilde{a}, \beta_r \mid \mathbf{y}) \times f'_{\tilde{A},B_r}(\tilde{a}, \beta_r) \right) d\tilde{a}\, d\beta_r} \qquad (11)$$

Using the results from 150 time-history analyses presented in Section 3.2.2, we can evaluate the posterior densities of $\tilde{a}$ & β_r using Equation (11). The comparison between prior and posterior density functions of $\tilde{a}$ & β_r are given in Figures 3a. and 3b.

3.2.4 Updated fragility curve

The updated median fragility curve $P_f''(\lambda \mid \mathbf{y})$ can be obtained by integrating out the random parameters using the posterior density function:

$$P_f^{''}(\lambda \mid \mathbf{y}) = \iint \Phi\left[\frac{\ln \lambda - \ln \tilde{a}}{\beta_r}\right] \cdot f_{\tilde{A},B_r}^{''}(\tilde{a}, \beta_r)\, d\tilde{a} \cdot d\beta_r \qquad (12)$$

It must be noted that the integrations involved in Equations (11) – (12) are too arduous to be trivially evaluated by numerical integration techniques. Therefore, the likelihood function $L(\tilde{a}, \beta_r \mid \mathbf{y})$ and the posterior density $f_{\tilde{A},B_r}^{''}(\tilde{a}, \beta_r \mid \mathbf{y})$ are typically evaluated using a special class of computational algorithms called Markov Chain Monte Carlo (MCMC) methods (Winkler, 2003; Congdon, 2006) which eliminate the need for explicit evaluation of intricate integrals. Figure 4 plots the prior and the updated median posterior fragility curves. It can be observed that the posterior fragility curve is very close to the point-estimate of fragility, $P_f(\lambda_i) = \frac{r_i}{n_i}$ as evaluated from the results of the time-history analyses.

3.2.5 Confidence bands

One of the important objectives of this work is to improve the confidence associated with the median fragility curves so as better facilitate the establishment of the USNRC criteria of $F_{1\%}$ = 1.67. Figure 5 depicts the change in 95% confidence bands between the EPRI (1994) prior and updated posterior fragility estimates. As we can see, the Bayesian updating methodology can be quite instrumental in significant reduction of the 95% confidence band region upon incorporating the data from the time-history simulations. The 95% confidence interval width

of median capacity $\tilde{a}$ (i.e. the PGA corresponding to 50% probability of failure) shrinks to $0.158g$ from the initial EPRI estimate of $1.794g$ following the Bayesian updating process.

Therefore, the Bayesian fragility model presented provides us with a consistent way to incorporate newly available fragility data to update the existing seismic fragility estimates as long as the data can be organized into the format given in Equation (9). Not only that, Bayesian updating can be instrumental in keeping a track of confidence bands which indicate the current level of accuracy in the fragility estimate and thereby, help us decide if further data may be needed to achieve the desired accuracy.

4. PLANNING OF PROBABILISTIC SIMULATIONS

In this section, we explore two important and frequent questions in the planning the Probabilistic simulations for structural fragility applications – (a) How many time history analyses are required for estimating the structural fragility within sufficient limits of accuracy? (b) How can we minimize the overall computational cost incurred in conducting the simulations? In the Bayesian updating example of Section 3, the input variables for the design of probabilistic simulations $\mathbf{P} \equiv \left[\overline{\lambda}, \overline{N}\right]$ have been arbitrarily chosen. 150 non-linear time-history analyses have been conducted at k=5 PGA levels with n=30 analyses at each PGA. Clearly, performing more time-history analyses i.e. $n > 30$ at each PGA will result in even narrower confidence bands for posterior fragility than those observed in Figure 5.

Currently, no guidelines exist on how to choose the various PGA values or the number of analyses to be performed at a given PGA value in performing the probabilistic simulations. Since even a single time-history analysis of the experimentally validated finite element model can take a formidable amount of computational time, availability of adequate computational resources to perform hundreds of hours of simulations can be a limiting factor. Advanced planning and allocation can help use the computational resources more efficiently. Therefore, it is of interest for engineers to obtain a reasonable estimate of total number of analyses that might be required in order to estimate seismic fragility of the structure to sufficient accuracy. Before we proceed, let us formally present the key input variables for designing any probabilistic simulation:

4.1 Input variables for time-history simulations

Two input variables are typically considered in any time-history simulation:

(i) $\{\bar{\lambda}\} = [\lambda_1 \ \lambda_2 \ ... \ \lambda_k]^T$, a vector of k earthquake intensities or PGA levels at which the time-history analyses have to be conducted;

(ii) $\{\bar{N}\} = [n_1 \ n_2 \ ... \ n_k]^T$, a vector denoting the number of analyses n_i to be conducted at the i^{th} PGA level λ_i.

Consequently, any design plan **P** for the time-history simulations can be characterized in terms of the input variables as $\mathbf{P} \equiv [\bar{\lambda}, \bar{N}]$. So, we would like to determine the optimum

design plan $\mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]$ such that the total number of time-history simulations $N_{sim} = \sum_{j=1}^{k} n_j$ is as low as possible. In order to do so, we need to first establish a convergence criterion which is a measure of desired accuracy of fragility estimates.

4.2 Convergence criteria

In the lognormal model of structural fragility, the random parameter $\tilde{a}$ representing the median capacity of the structure is more influential to the posterior fragility curve than the log-standard deviation parameter β_r. Therefore, it can be assumed that the resultant fragility curve will have converged if we are reasonably confident about the mean value of the median acceleration capacity $\tilde{a}$ i.e. if the length of the 95% confidence interval for $\tilde{a}$ is sufficiently small. The mean-centered 95% confidence interval of $\tilde{a}$, $L_{0.95}(\tilde{a})$ can be calculated by

$$L_{0.95}(\tilde{a}) = F_{\tilde{A}}^{-1}(0.975) - F_{\tilde{A}}^{-1}(0.025) \tag{13}$$

where $F_{\tilde{A}}(\tilde{a})$ denotes the cumulative distribution function of $\tilde{a}$. Using the prior information on $\tilde{a}$ from Equation (8) yields an initial 95% confidence interval of $L_{0.95}(\tilde{a} \mid prior)$ = 1.794*g*. In the illustrative example of Section 3, the 95% confidence interval after incorporating the data from 30 analyses at each of *k*=5 PGA levels is 0.158*g* which is a considerable reduction in comparison to the prior confidence interval length of 1.794*g*. Therefore, a reasonable criterion to indicate convergence of posterior fragility curves is to be able to estimate the value of the median capacity $\tilde{a}$ within a specified limit of accuracy, say

$\pm 0.1g$ with a confidence of 95%. Such a criteria is equivalent to stipulating that the maximum length of the 95% confidence interval for $\tilde{a}$ shall be less than or equal to $L_{\max} = 0.2g$.

4.2.1 Problem statement

We now formally state the problem of planning the probabilistic simulations. The goal of planning the simulations is to minimize the overall computational effort required in determining the structural seismic fragility curve to a desired limit of accuracy. This can be expressed as:

$$\begin{aligned} &\text{Find } \mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right] \text{such that} \\ &\text{minimize} \sum_{j=1}^{k} n_i \\ &\text{subject to } L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]\right) \le L_{\max} \end{aligned} \tag{14}$$

In Equation (14), $L_{\max}$ denotes the maximum allowed 95% confidence interval for $\tilde{a}$ as a measure of desired accuracy; $\{\bar{\lambda}\}$ denotes the vector of various PGA values at which the simulations are to be conducted; $L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]\right)$ denotes the posterior 95% confidence interval value of $\tilde{a}$ after incorporating the fragility data obtained once the plan $\mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]$ is executed. In other words, we are trying to obtain an estimate of amount of data/number of analyses that might be necessary to satisfy the convergence criteria $L_{\max}$. The following sections discuss the estimation of the required amount of fragility data.

4.3 Expected number of analyses required

4.3.1 Sampling distributions

The theory of sampling distributions is commonly used to estimate the number of samples or data points required in order to predict a given statistic within a specified margin of accuracy. Consider a random variable *X* with a mean of μ_X and a known standard deviation σ_X. For a *n*-sized population sample of *X*, the sample mean $\overline{X}$ follows a distribution that given by

$$\overline{X} \sim N\left(\mu_X, {\sigma_X}^2 \Big/ n\right) \tag{15}$$

The quantity $\sigma_X \Big/ \sqrt{n}$ represents the "standard error" and is equal to the standard deviation of the sample mean $\overline{X}$. Clearly, we can see that $\sigma_X \Big/ \sqrt{n} \to 0$ when $n \to \infty$ implying that the larger the sample size, smaller is the variance of the sample mean and greater is our accuracy in estimating the mean μ_X. Suppose that we are interested in determining the mean a_m of the median capacity $\tilde{a}$ within an accuracy of ± 0.1*g* with a confidence of 95%. From the prior information on the median capacity in Equation (8), one can determine that the standard deviation of $\tilde{a}$ is $\sigma_{\tilde{a}} = 0.4631$ using the properties of the Lognormal distribution. Similar to the relationship given in Equation (15), the sample mean $\overline{a}_m$ for a sample size of *s* follows the distribution: $\overline{a}_m \sim N\left(a_m, {\sigma_{\tilde{a}}}^2 \Big/ s\right)$. If a_m must be estimated within ± 0.1*g* with a

probability of $(1-\alpha)$= 0.95, we have that its 95% confidence interval should be less than or equal to 0.2g. Using this criterion, we can estimate the approximate number of samples $s_{\min}$ that must be collected:

$$L_{0.95}\left(\overline{a_m}\right) \le 0.2$$
$$\left(\Phi^{-1}\left(1-\alpha/2\right)-\Phi^{-1}\left(\alpha/2\right)\right)\frac{\sigma_{\tilde{a}}}{\sqrt{s_{\min}}} \le 0.2 \Rightarrow s_{\min} \ge 83 \quad (16)$$

Based on the prior distribution on $\tilde{a}$, we need 83 independent samples of the median capacity in order to estimate a_m within ± 0.1g with 95% confidence. A key limitation of this approach is that it yields the number of median capacity samples that are required to be collected. Using the theory of sampling distributions needs observations of median acceleration capacity of the structure. However, it must be remembered that we cannot observe directly any samples of the median capacity from the time history simulation data. This is because the median capacity is not a directly observable quantity but rather a statistically inferred parameter from the simulation data. The results of a time history simulation are typically compiled in the format of Equation (9). The $s_{\min}$ calculated in Equation (16) cannot be directly related to some design plan $\mathbf{P} \equiv \left[\{\overline{\lambda}\}, \{\overline{N}\}\right]$ that can guide us in conducting the time-history simulations. Since the posterior statistics of $\tilde{a}$ on which the convergence criteria is based on, are evaluated via the Bayesian updating process, any experimental design approach to determine the design plan $\mathbf{P} \equiv \left[\{\overline{\lambda}\}, \{\overline{N}\}\right]$ should be again based on Bayesian inference.

4.3.2 Expected Minimum Number of Analyses Using Bayesian Inference

Suppose we ought to determine the optimum configuration of the number of time-history analyses $\{\overline{N}\}$ are needed at each of the k PGA levels - $\lambda_1, \lambda_2, ...\lambda_k$ to ensure that a convergence criteria $L_{0.95}(\tilde{a} \mid \mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}]) \leq L_{\max}$ is met. Let the current state of knowledge on the random parameters $\tilde{a}, \beta_r$ of the structural fragility model be represented by the probability density $f_{\tilde{A},B_R}(\tilde{a}, \beta_r)$. In practice, it is reasonable to preset the vector of the PGA values $\{\overline{\lambda}\}$ based on the current or prior knowledge beforehand and assume it to be a known component in the design plan $\mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}]$. Now, the optimization problem of Equation (14) can be rewritten as

$$\begin{aligned} &\text{Find } \{\overline{N}\} = [n_1, \ldots, n_k]^T \text{ such that} \\ &\text{minimize} \sum_{j=1}^{k} n_j \\ &\text{subject to } E\left[L_{0.95}(\tilde{a} \mid \mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}])\right] \leq L_{\max} \end{aligned} \tag{17}$$

In Equation (17), $E\left[L_{0.95}(\tilde{a} \mid \mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}])\right]$ represents the expected or average length of the 95% confidence interval from the updated posterior density of $\tilde{a}$ using the fragility data from executing the plan $\mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}]$. Remember that determining the exact posterior statistics of $\tilde{a}$ require additional information $\{\overline{R}\} = [r_1, \ldots, r_k]^T$ which is a vector of the number of analyses r_j in which the structure is predicted as damaged out of a total number of analyses n_j conducted at a PGA level of λ_j. During the planning stage, we do not know the results of

the simulations $\{\overline{R}\}$ yet. In such a case, $\{\overline{R}\}$ can be modeled as a vector of random variables where each r_j in $\{\overline{R}\}$ is a binomial random variable invoking the prior density $f_{\tilde{A},B_R}(\tilde{a},\beta_r)$. The binomial distribution follows directly from the likelihood function in Equation (10) of Section 3. Equation (18) describes how $\{\overline{R}\}$ can be characterized as a vector of random variables based on current or prior state of knowledge:

$$\begin{aligned} &\{\overline{R}\} = [r_1,\ldots,r_k]^T \\ &r_j \sim Bin\left(n_j, P_f\left(\lambda_j;\breve{a},\beta_r\right)\right)\ \forall j = 1...k \\ &(\breve{a},\beta_r) \sim f'_{\tilde{A},B_R}(\tilde{a},\beta_r) \end{aligned} \tag{18}$$

In Equation (18), $P_f\left(\lambda_j;\breve{a},\beta_r\right)$ is the lognormal fragility function of Equation (7) whereas $f'_{\tilde{A},B_R}(\tilde{a},\beta_r)$ represents the prior or current state of knowledge on the random parameters - $(\breve{a},\beta_r)$. Now, for any given plan $\mathbf{P} \equiv \left[\{\overline{\lambda}\},\{\overline{N}\}\right]$, we can determine the expected posterior 95% confidence interval $E\left[L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\overline{\lambda}\},\{\overline{N}\}\right]\right)\right]$ by averaging out the random vector $\{\overline{R}\}$ as shown in Equation (19).

$$\begin{aligned} \mathrm{E}\left[L_{0.95}(\breve{a}) \mid \mathbf{P} \equiv \{\overline{N}\},\{\overline{\lambda}\}\right] &= \mathrm{E}_{\{\overline{R}\}}\left[L_{0.95}\left(\breve{a} \mid \{\overline{N}\},\{\overline{\lambda}\},\{\overline{R}\}\right)\right] \\ &= \int_{\{\overline{R}\}} L_{0.95}\left(\breve{a} \mid \{\overline{N}\},\{\overline{\lambda}\},\{\overline{R}\}\right)\cdot P\left[\{\overline{R}\} = \{\overline{r}\}\right]\cdot d\{\overline{r}\} \end{aligned} \tag{19}$$

Notice that the posterior 95% confidence interval of $\tilde{a}$ - $L_{0.95}\left(\breve{a} \mid \{\overline{N}\},\{\overline{\lambda}\},\{\overline{R}\}\right)$ can be precisely determined using the Bayesian updating approach of Section 3. However, the

integral in the Equation (19) cannot be explicitly evaluated as the Bayesian updating methodology is quite intricate and requires a formidable effort to determine the posterior densities in analytical form. In such a case, we can still determine $E\left[L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]\right)\right]$ by randomly sampling the simulation results $\{\bar{R}\}$ for the given plan $\mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]$ using the prior information and thereby using the MCMC techniques to infer the posterior convergence criteria in each case (Hamada et al, 2004). The steps to determine the expected convergence criteria $E\left[L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]\right)\right]$ given $\mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]$ using the current information - $f'_{\tilde{A},\mathrm{B}_R}(\tilde{a}, \beta_r)$ are:

Step 1: Draw a sample set for $(\tilde{a}, \beta_r)$ using the prior density $f'_{\tilde{A},\mathrm{B}_R}(\tilde{a}, \beta_r)$

Step 2: Using the sample set $(\tilde{a}, \beta_r)$ from Step (1), generate an instance of fragility data $\{\bar{R}\}$ for the given plan $\mathbf{P} \equiv \left[\{\bar{\lambda}\}, \{\bar{N}\}\right]$ following Equation (19). This can be done using the steps (a) through (c) for each $j = 1,2, \ldots k$ where k is the length of the vector of PGA values.

(a) Evaluate p_j using the lognormal fragility function: $p_j = P_f\left(\lambda_j; \breve{a}, \beta_r\right)$

(b) Generate an instance of r_j by drawing a sample from $Bin\left(n_j, P_f\left(\lambda_j; \breve{a}, \beta_r\right)\right)$

(c) Assemble the random vector $\{\bar{R}\} = [r_1, \ldots, r_k]^T$

Step 3: Using the data $\mathbf{y} \equiv \{\overline{N}\}, \{\overline{\lambda}\}, \{\overline{R}\}$, generate the posterior statistics $f''_{\tilde{A},B_R}(\tilde{a}, \beta_r)$ using the Bayesian updating methodology of Section 3. The posterior density $f''_{\tilde{A}}(\tilde{a})$ can be used to evaluate the convergence criteria - $L_{0.95}(\breve{a} \mid \{\overline{N}\}, \{\overline{\lambda}\}, \{\overline{R}\})$ from Equation (13).

Step 4: Repeat *m* times the steps 1 through Step 3 to evaluate *m* values of the convergence criteria. In this paper, the value of *m* is taken as 500.

Step 5: Determine the average of the *m* values of $L_{0.95}(\breve{a}) \mid \{\overline{N}\}, \{\overline{\lambda}\}, \{\overline{R}\}$ to evaluate the $E\left[L_{0.95}(\tilde{a} \mid \mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}])\right]$ for a given plan $\mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}]$.

Determining the expected convergence criteria $E\left[L_{0.95}(\tilde{a} \mid \mathbf{P} \equiv [\{\overline{\lambda}\}, \{\overline{N}\}])\right]$ for a given plan **P** enables us to solve the optimization problem presented in Equation (17) to determine the optimal $\{\overline{N}\}$ that is expected to minimize the overall computational effort in conducting the simulations while achieving the desired level of accuracy.

4.4 Determining the optimal simulation plan

Since $\{\overline{N}\} = [n_1, \ldots, n_k]^T$ is a *k*-dimensional vector, one simple way to obtain an approximate solution is to plan the same number of analyses at each of the *k* PGA levels – $\lambda_1, \lambda_2, \ldots \lambda_k$ i.e. $n_1 = n_2 = \ldots n_k = q$. Such an assumption transforms the optimization problem in a *k*-D variable space into a 1-D problem which can be solved graphically or using simple techniques like mid-point search etc. To illustrate this planning solution for the RC box-

shaped shear wall example, we use the EPRI (1994) prior densities of the random parameters $\tilde{a}, \beta_r$ as given in Equation (8). The vector of PGA values $\{\bar{\lambda}\}$ is selected as $\{\bar{\lambda}\} = [0.5g \ 1.0g \ 1.2g \ 1.5g \ 2.0g]^T$ such that k = 5. If q analyses were to be conducted at each of the 5 PGA levels i.e. $\{\bar{N}\} = [q \ q \ q \ q \ q]^T$, the expected value of convergence criteria $E\left[L_{0.95}(\tilde{a} \mid \{\bar{\lambda}\}, q)\right]$ can be calculated for different values of q using the steps (1)-(5) of Section 4.3.2 for the shear wall example. Figure 7 plots the expected convergence criteria $E\left[L_{0.95}(\tilde{a} \mid \{\bar{\lambda}\}, q)\right]$ as a function of number of analyses, q planned to be conducted at each PGA level. We can see that $E\left[L_{0.95}(\tilde{a} \mid \{\bar{\lambda}\}, q)\right]$ monotonically decreases as the number of simulations conducted increase. This is expected as more analyses imply a larger amount of collected simulation data collected which leads to much improved confidence in the evaluated fragility curve as reflected by a smaller 95% confidence interval of the random parameter $\tilde{a}$. If $L_{\max} = 0.2g$ *i.e.* if we desire to predict the value of $\tilde{a}$ within ±0.1*g*, we can see from Figure (6) that the corresponding $q \approx 57$. This means that we must plan to conduct at least q = 57 analyses at each of the k = 5 PGA levels i.e. a total of 285 analyses in order to meet the convergence criteria of $E\left[L_{0.95}(\tilde{a} \mid \{\bar{\lambda}\}, q)\right] \leq 0.2g$.

While the simplified assumption of conducting the same number of analyses at all the k PGA levels i.e. $n_1 = n_2 = \ldots n_k = q$ gives us a quick estimate of the total number of simulations required, it is not the most optimal solution possible to the optimization problem of Equation (17). We present here the following qualitative argument as to why this approach

may not be the most optimal solution: suppose that, following a similar analysis as presented above, we conclude that about 100 analyses might have to be conducted at each PGA level. At sufficiently high levels of PGA, we can expect high probabilities of failure *i.e.* close to one. Let us say that we completed 20 analyses at a high PGA value and that each of the 20 simulations predicted damage to the structure indicating 100% probability of failure thus far. It can be reasoned based on this intermediate result that performing the remaining 80 analyses as planned may not significantly contribute any new information in comparison to the computational costs incurred in completing the 80 analyses. A similar argument can be presented for extremely low PGA values where the probabilities of failure are extremely low. Since our convergence criteria is based on the median acceleration capacity $\tilde{a}$, performing more analyses relatively around the intermediate levels of PGA or around the median region (corresponding to 50% probability of failure) as compared to the tail regions of the fragility curve can maximize the "information gained" and hence could be a better solution. Therefore, performing different number of analyses at different levels of PGA or treating $\{\overline{N}\}$ as a vector of k independent variables instead of assuming $n_1 = n_2 = \ldots n_k = q$ for the optimization problem of Equation (17), one can expect to arrive at a solution that results in a greater economy of the total computing effort, $\sum_{j=1}^{k} n_j$.

4.4.1 Optimal simulation plan using heuristic techniques

In Equation (17), the minimization objective and the constraint can be combined into a single objective function by introducing a penalty function as follows:

$$\begin{aligned} &\text{Find } \{\overline{N}\} = [n_1, \ldots, n_k]^T \text{ such that} \\ &\text{minimize} \left(\sum_{j=1}^{k} n_j \right) \cdot 1(L^* \le L_{\max}) + (\mu + \varepsilon(L^* - L_{\max})) \cdot 1(L^* > L_{\max}) \\ &\text{where } L^* = E\left[L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\overline{\lambda}\}, \{\overline{N}\}\right]\right)\right] \end{aligned} \tag{20}$$

where μ and ε are penalty constants and 1(.) is the indicator function. In Equation (20), the first term of the objective function represents the total computational cost which is the total number of analyses in case the candidate solution $\{\overline{N}\} = [n_1, \ldots, n_k]^T$ is expected to satisfy the convergence criteria. Else, the penalty function is invoked for the set of plans $\mathbf{P} \equiv \left[\{\overline{\lambda}\}, \{\overline{N}\}\right]$ whose expected convergence criteria doesn't meet the target convergence criteria, $L_{\max}$. Each calculation of this objective function involves subsequent evaluation of $E\left[L_{0.95}\left(\tilde{a} \mid \mathbf{P} \equiv \left[\{\overline{\lambda}\}, \{\overline{N}\}\right]\right)\right]$ using Steps (1)-(5) of Section 4.3.2. As a result, determining the full description of the objective function in a *k*-dimensional space is computationally challenging. In such cases, a reasonably good solution to the above optimization problem may be obtained by using population-based heuristic or meta-heuristic techniques such as genetic algorithms or swarm intelligence algorithms etc.

In order to obtain an optimal $\{\overline{N}\}$ for planning of probabilistic simulations for seismic fragility assessment of the RC box-shaped shear wall, we use a technique called the Particle Swarm Optimization (PSO) to solve the minimization problem described in Equation (20). PSO is a meta-heuristic computational method that proposes a swarm of candidate solutions (called "particles") and the solution is arrived at by iteratively moving the population (or "swarm") of candidate solutions over the search space based on simple rules. It is relatively easy to implement and can be conveniently used to solve optimization problems in a *k*-D space even when *k* is moderately large. Of many variants of PSO, the "canonical particle swarm optimizer" algorithm was used in this study. The various parameters for configuring the PSO were taken from the study by Montes de Oca et al (2006). The values of the penalty constant used in the paper are: $\mu = 500$ and $\varepsilon = 0.1$.

To implement the planning of simulations for the shear wall example or any application, two computer programs are required:

(a) Program to execute the Particle Swarm Optimization (PSO) to find the optimum plan

(b) Program that evaluates the objective function in Equation (20) given a specific instance of a plan $\mathbf{P} \equiv \left[\{\overline{\lambda}\}, \{\overline{N}\}\right]$ using Steps (1)-(5) of Section 4.3.2. This program will be called for execution by (a) as it keeps generating various candidate solutions.

Executing (a) would enable us to obtain the optimal solution for $\{\overline{N}\}=[n_1,\ldots,n_k]^T$ as the design plan according to the current state of knowledge of random parameters $(\tilde{a},\beta_r)$, the selected vector of PGA values $\{\overline{\lambda}\}$ and the target convergence criteria, $L_{\max}$.

4.5 Incremental procedure to conduct probabilistic simulations until convergence

In this section, we propose and illustrate a methodical approach to conduct the time-history simulations till required accuracy is achieved in the fragility curve. Section 4.4 discusses how one can determine an optimal plan $\mathbf{P}\equiv\left[\{\overline{\lambda}\},\{\overline{N}\}\right]$ to conduct the simulations can be determined based on the current knowledge of the random parameters of the fragility model. Consequently, the quality of the optimal solution for $\{\overline{N}\}=[n_1,\ldots,n_k]^T$ is itself predominantly influenced by the quality of the prior information. In such a case, it is often inefficient to conduct all the simulations according to the optimum plan based on a single prior. Instead, it is more effective to conduct the simulations incrementally in stages wherein only a fraction of simulations based on the optimal plan are conducted and the fragility data obtained is used to update the current knowledge of the parameters of the fragility model. The optimal plan is again revised by solving the optimization problem of Section 4.4 using the prior from the updated knowledge on the random parameters - $(\tilde{a},\beta_r)$. This revised optimal plan helps us conduct the next stage of simulations and this process is repeated gradually until the desired accuracy is achieved in the fragility estimates. A flow chart for the incremental approach is depicted in Figure 7.

4.5.1 Illustrative example for planning of probabilistic simulations

To illustrate the implementation of probabilistic simulations using the above approach, we again consider the example of the RC box-shaped shear wall example described in Section 2.1. We start with the EPRI (1994) fragility estimates as the prior knowledge of the fragility model parameters. Each stage is implemented based on the procedure detailed in Figure 7. The vector of PGA values is $\{\bar{\lambda}\} = [0.5g \;\; 1.0g \;\; 1.2g \;\; 1.5g \;\; 2.0g]^T$ and is assumed to remain the same for all the stages. The fragility curve was assumed to have converged if the 95% posterior confidence interval of the median capacity parameter $\tilde{a}$ is less than or equal to 0.2*g*. The various details of the stage-wise execution of the probabilistic simulations and the required number of analyses estimated at the end of each stage are presented in Tables 1(a) and 1(b).

Table 1(a) gives the details of the optimal plan, actual number of simulations conducted and the subsequent fragility data for each stage. We can infer from Table 1(b) that based on the EPRI (1994) prior fragility information, the optimal plan predicts that we might need up to a total of 157 simulations for convergence. After Stage 1 which comprised of 25 simulations, the optimal plan incorporating the results of the Stage 1 predicts that we need only a total of 75 simulations for convergence indicating a significant drop in the expected total number of simulations required. The simulations can be carried out in incremental stages until the seismic fragility curves are evaluated to the desired accuracy level. We can see that the convergence was reached after 80 time-history analyses in order to predict the

value of median capacity $\tilde{a}$ within $\pm 0.1g$. Figure 8 compares the median fragility curves evaluated at the end of each phase.

The Bayesian framework presented in this paper is an effective approach to update the prior fragility curves using additional fragility data from Experiments or Probabilistic simulations with multiple time history analyses based on experimentally validated FE model. The framework present is proven to produce reliable estimates with quick convergence. Furthermore, the ability of the framework to produce fragility estimates regardless of the amount of experimental or simulation data available can be very helpful in subsequent planning of resource allocation for collecting the data. This feature of the framework can have far-reaching implications in saving substantial costs in computation, resources as well as planning for experiments and data collection while updating the fragility estimates of old and existing structural systems in NPPs to reevaluate their seismic risk.

5. CONCLUSION

The paper addresses the problem of evaluating seismic fragility of large-scale engineering structures to a desired level of accuracy while minimizing the overall computational effort to achieve the same. For real-life structural and nonstructural systems, modeling the various uncertainties in material, model and earthquake loading variables enables us to perform multiple nonlinear time history analyses from which the seismic risk or fragility at a

particular intensity level of earthquake can be determined statistically. We proposed a comprehensive framework based on Bayesian inference methods to facilitate the same. The proposed Bayesian framework has three important features:

1. The framework provides a consistent way to incorporate different kinds of seismic risk/fragility data and dynamically update the structural fragility of engineering structures as and when data becomes available.
2. The framework enables quick convergence by reducing the confidence intervals even with limited amount of time-history data. Not only that, the framework can be used to continually monitor the current accuracy of fragility estimates and also forecast how much more data might be required in order to achieve desired accuracy in seismic fragility curves.
3. The framework combines optimization techniques with the ability of the Bayesian fragility model to forecast the required amount of data for desired accuracy to form an efficient planning tool that is capable of minimizing the total computational cost in collecting the data itself.

The above features of the framework are illustrated with an example of a box-shaped reinforced concrete shear wall.

ACKNOWLEDGEMENTS

The authors on behalf of the Center for Nuclear Power Plant Structures, Equipment and Piping at North Carolina State University sincerely wish to thank Kajima Corporation of Japan and NCSU College of Engineering for their continued support. Resources for the Center come from sponsoring industrial organizations, National Science Foundation, the Department of Civil, Construction and Environmental Engineering and College of Engineering in the University.

REFERENCES

ASCE/SEI 43-05, "Seismic Design Criteria for Structures, Systems, and Components in Nuclear Facilities," American Society of Civil Engineers and Structural Engineering Institute, 2005.

Carausu, A. and Vulpe, A. “Fragility estimation for seismically isolated nuclear structures by high confidence low probability of failure values and bi-linear regression,” Nuclear Engineering and Design, 160 p. 287-297, 1996.

Congdon, P., “Bayesian Statistical Modelling,” Wiley Series in Probability and Statistics, 2nd Edition, 2006.

Do-Eun C., Gardoni, P., Rosowsky, D., Haukaas, T., “Probabilistic capacity models and seismic fragility estimates for RC columns subject to corrosion”, Reliability Engineering and System Safety, v 93, n 3, p 383-93, 2008

Electric Power Research Institute (EPRI), “Methodology for Developing Seismic Fragilities”, EPRI TR-103959 (Principal Investigators: Reed, J.W. and Kennedy, R.P.), Electric Power Research Institute, Palo Alto, California, 1994.

Hamada, M.S., Wilson, A., Reese, C. S., Martz, H., “Bayesian Reliability,” Springer Series in Statistics, 2008.

Huang, Q. Gardoni, P., Hurlebaus, S., “Probabilistic seismic demand models and fragility estimates for reinforced concrete highway bridges with one single-column bent,” Journal of Engineering Mechanics, v 136, n 11, p 1340-1353, 2010.

Kennedy, R.P., “Performance-Goal Based (Risk Informed) Approach for Establishing the SSE Site Specific Response Spectrum for Future Nuclear Power Plants,” *Jaeger Lecture*, SMIRT-19, Toronto, Canada, 2007.

Kennedy, R.P., Cornell, C.A., Campbell, R.D., Kaplan, S., Perla, H.F., “Probabilistic seismic safety of an existing nuclear power plant,” Nuclear Engineering and Design, 59, 315–338, 1980.

Montes de Oca, M. A., Stutzle, T., Birattari, M., Dorigo, M., "A comparison of particle swarm optimization algorithms based on run-length distributions", Proc. 5th Int. Workshop, Ant Colony Optimization Swarm Intell. (ANTS '06), vol. LNCS 4150, pp.1, 2006

Petelet M., Iooss B., Asserin O., Loredo A., "Latin hypercube sampling with inequality constraints," *Advances in Statistical Analysis*, submitted. Available at: http://hal.archives-ouvertes.fr/docs/00/52/00/21/PDF/AStA10_iooss_rev2.pdf, 2010.

Ramamoorthy, S.K., Gardoni, P., Bracci, J.M., "Probabilistic demand models and fragility curves for reinforced concrete frames," v 132, n 10, p 1563-1572, 2006

Serdar K.M., Polat, Z., "Fragility analysis of mid-rise R/C frame buildings," Engineering Structures, v 28, n 9, p 1335-1345, 2006.

Shinozuka, M., Feng, M.Q., Lee, J., Naganuma, "Statistical analysis of fragility curves," Journal of Engineering Mechanics, v 126, n 12, p 1224-1231, 2000.

Tadinada S.K., "Consideration of Uncertainties in Seismic Analysis of Non-classically damped coupled systems," Master's Thesis, North Carolina State University, Available at: http://repository.lib.ncsu.edu/ir/bitstream/1840.16/2134/1/etd.pdf , 2009.

USNRC, "PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants," NUREG/CR-2300, 1983.

USNRC, "Policy, Technical, and Licensing Issues Pertaining to Evolutionary and Advanced Light-Water Reactor (ALWR) Designs, (Item II Q)," SECY-93-087, 1993.

USNRC, "Technical Basis for Revision of Regulatory Guidance on Design Ground Motions: Harzard- and Risk- Consistent Ground Motion Spectra Guidelines," NUREG/CR-6728, 2001.

Winkler, R.P., "An Introduction to Bayesian Inference and Decision," Probabilistic Publishing, 2nd edition, 2003.

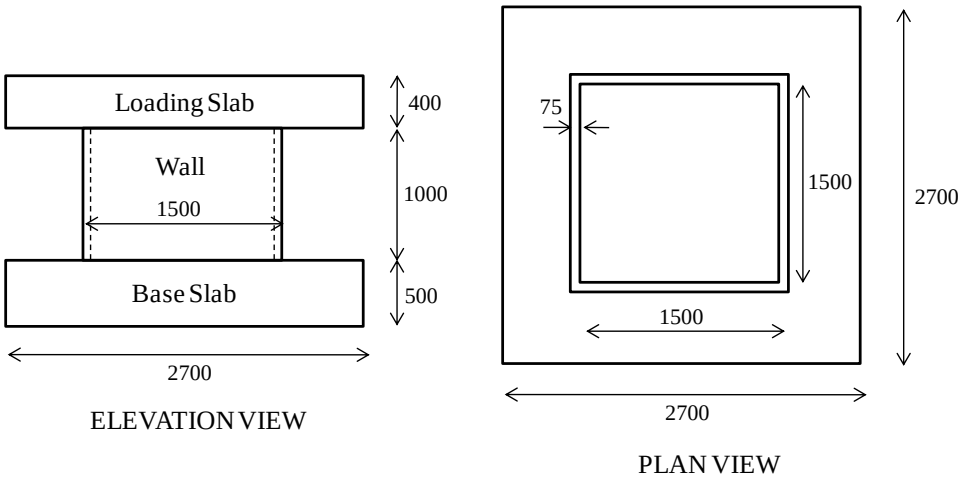
Zentner, I., "Numerical computation of fragility curves for NPP equipment," Nuclear Engineering and Design 240, 1614–1621, 2010.

Table 1(a): Results of the Stage-wise Implementation of Probabilistic Simulations for Shear Wall Fragility

Stage 1	PGA	Optimal Plan $\{\overline{N}\}$	Simulations Conducted $\{\overline{N}\}$	No. of Failures, $\{\overline{R}\}$	$L_{0.95}(\tilde{a})$ at the end of stage
	0.5	5	5	0	0.409 g
	1	5	5	1	
	1.2	10	5	4	
	1.5	51	5	4	
	2	86	5	5	
Stage 2	PGA	Optimal Plan $\{\overline{N}\}$	Simulations Conducted $\{\overline{N}\}$	No. of Failures, $\{\overline{R}\}$	$L_{0.95}(\tilde{a})$ at the end of stage
	0.5	7	7	0	0.260 g
	1	17	10	5	
	1.2	30	15	13	
	1.5	14	10	9	
	2	7	7	7	
Stage 3	PGA	Optimal Plan $\{\overline{N}\}$	Simulations Conducted $\{\overline{N}\}$	No. of Failures, $\{\overline{R}\}$	$L_{0.95}(\tilde{a})$ at the end of stage
	0.5	7	7	0	0.206 g
	1	17	17	8	
	1.2	30	30	26	
	1.5	14	14	13	
	2	7	7	7	
Stage 4	PGA	Optimal Plan $\{\overline{N}\}$	Simulations Conducted $\{\overline{N}\}$	No. of Failures, $\{\overline{R}\}$	$L_{0.95}(\tilde{a})$ at the end of stage
	0.5	7	7	0	0.197 g
	1	22	22	12	
	1.2	30	30	26	
	1.5	14	14	13	
	2	7	7	7	

Table 1(b): Summary of the Stage-wise Implementation of Probabilistic Simulations

Stage	Total Number of Time History Analyses Completed	Expected Total Number of Simulations Required	$L_{0.95}(\tilde{a})$ at the end of stage	Remarks
0	0	157	1.794 g	EPRI (1994) as prior
1	25	75	0.409 g	
2	50	75	0.260 g	
3	75	75	0.206 g	
4	**80**	**80**	0.197 g	Converged



(All dimensions are in mm)

Figure 1: Box-shaped Reinforced Concrete Shear Wall

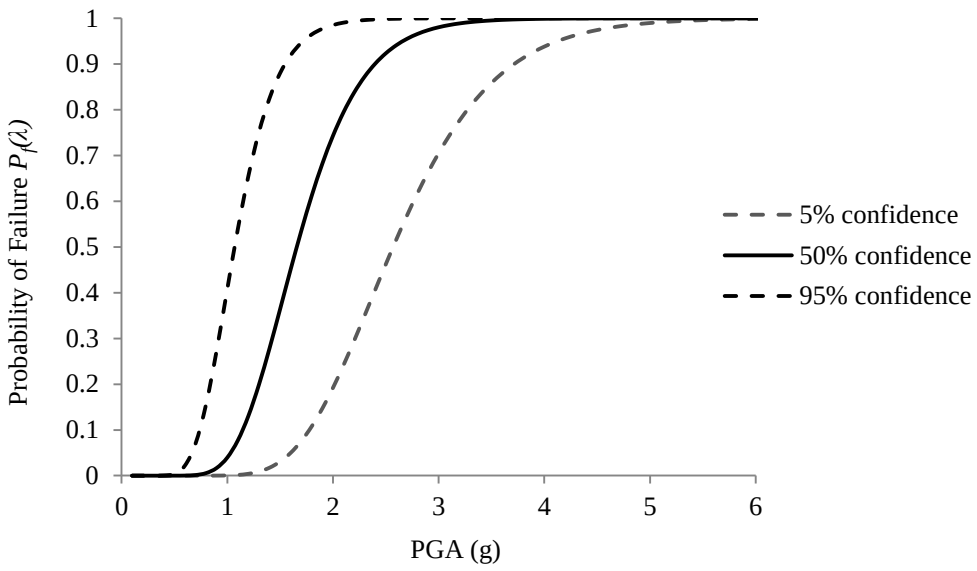


Figure 2: Fragility Curve of the Box-type Shear Wall Structure

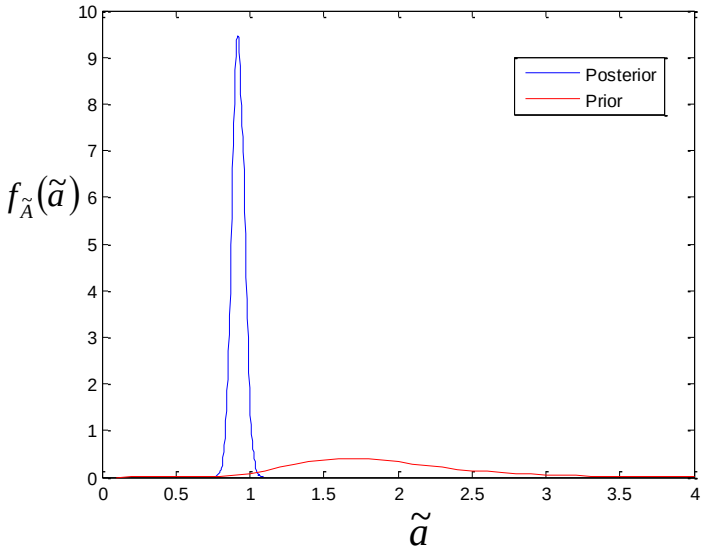


Figure 3(a): Prior and Posterior Densities of $\tilde{a}$

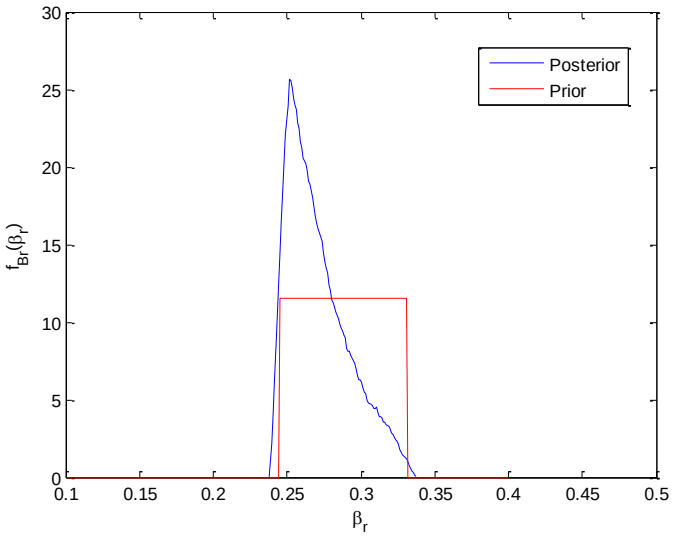


Figure 3(b): Prior and Posterior Densities of β_r

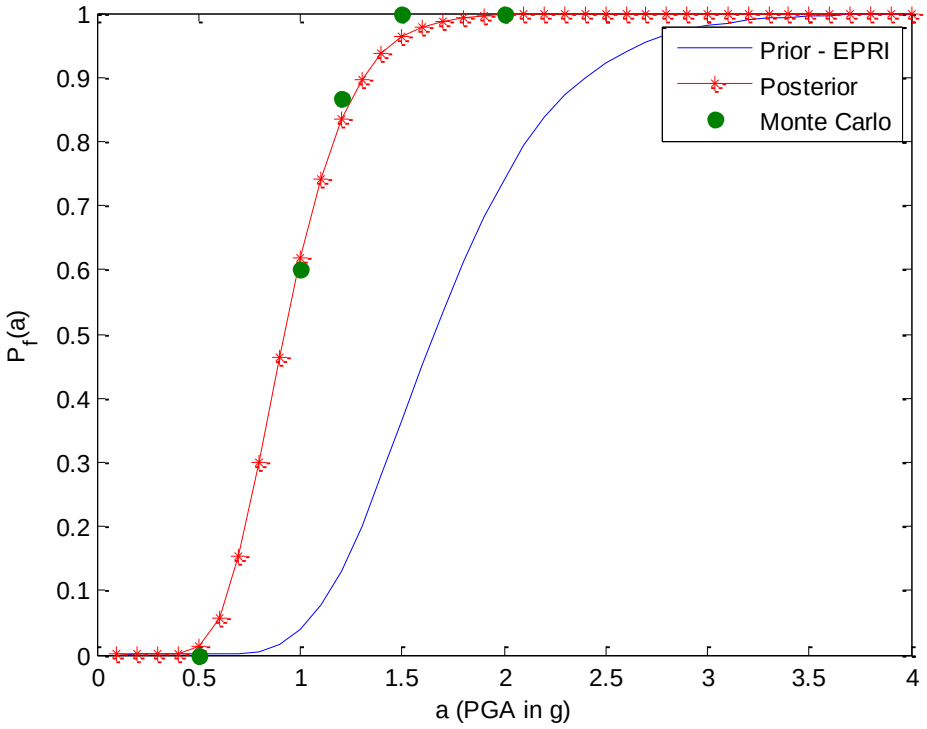


Figure 4: Comparison of Prior and Updated Median Fragility Curves

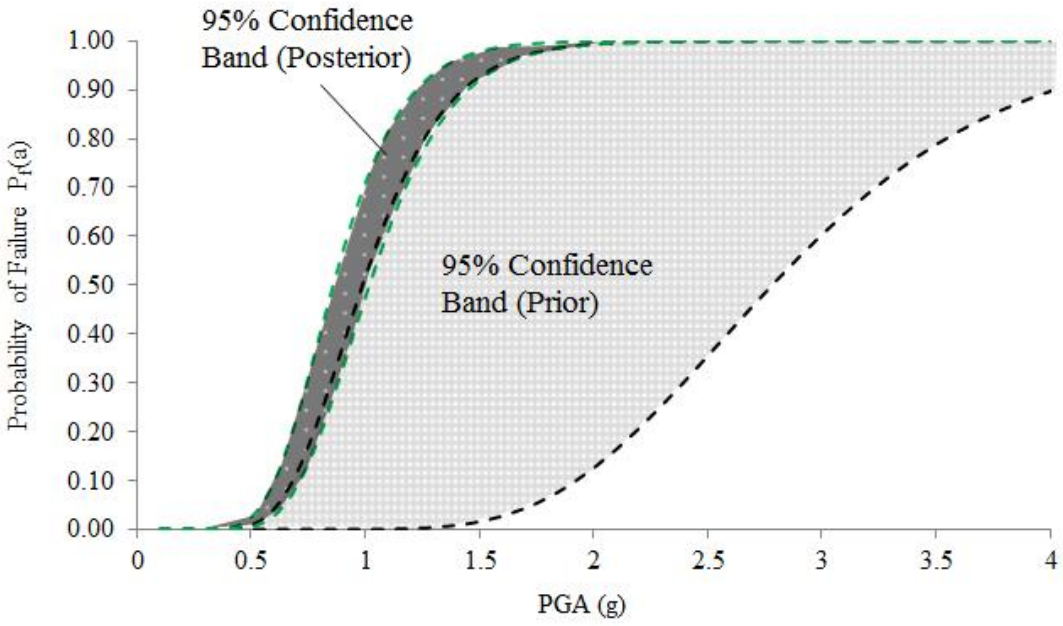


Figure 5: 95% Confidence Bands on Prior and Posterior Fragility Curves

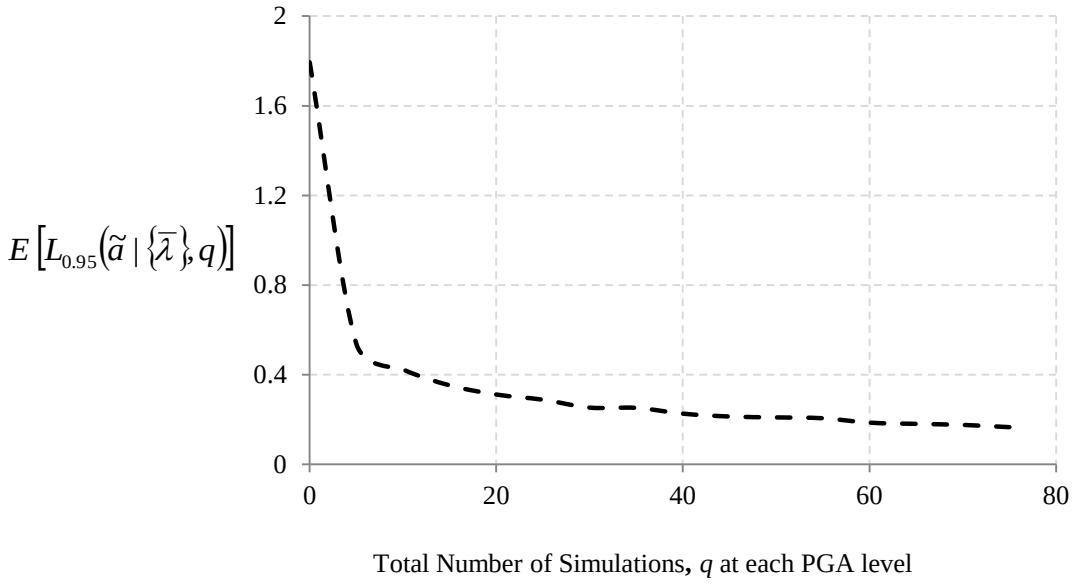


Figure 6: Convergence Criteria as a Function of q based on EPRI (1994) Prior information

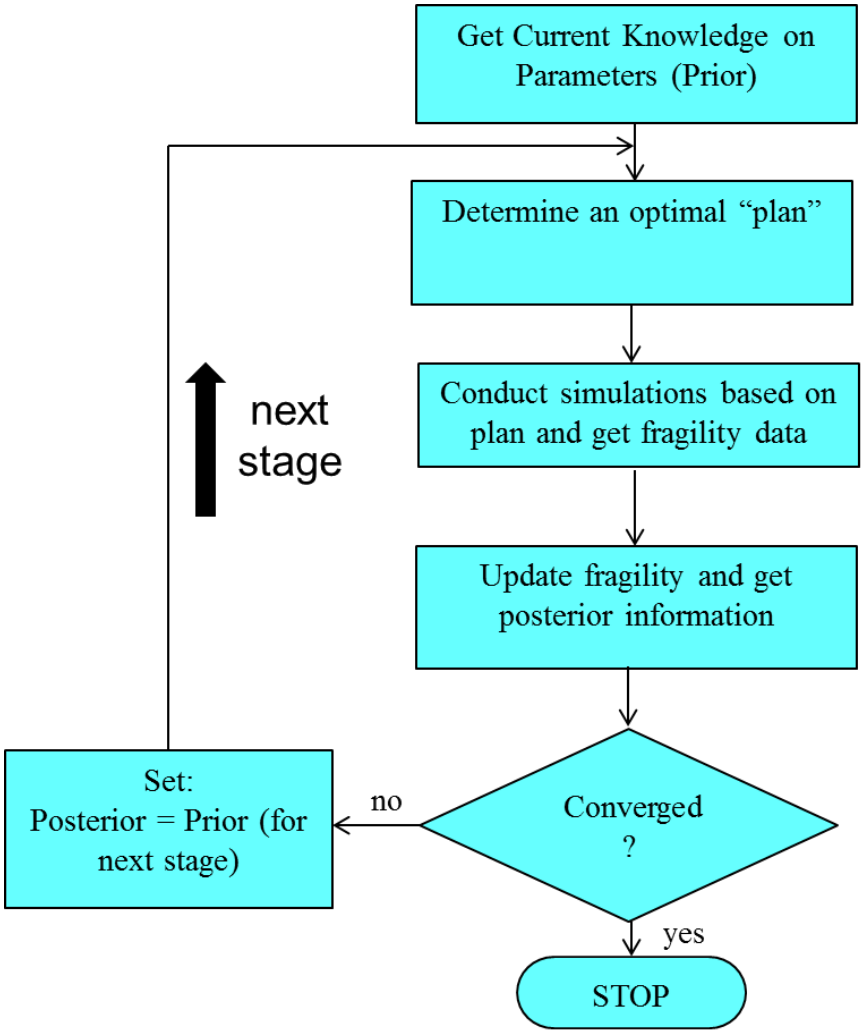


Figure 7: Flowchart for the Incremental Procedure of Planning and Conducting Probabilistic Simulations

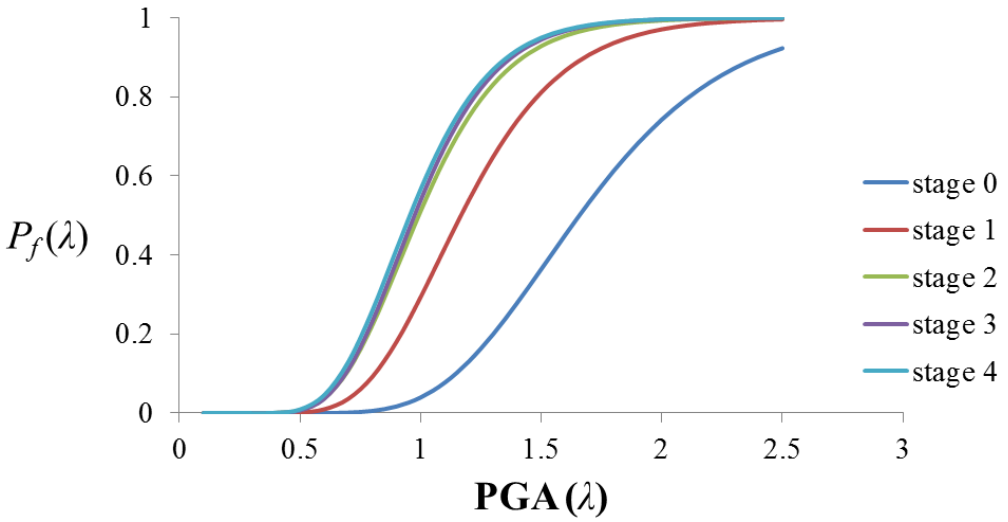


Figure 8: Median Fragility Curves from Posterior Statistics at the End of Each Stage

PART III

STRUCTURAL FRAGILITY OF LARGE-SCALE PIPING SYSTEMS USING EQUIVALENT ELASTIC TIME-HISTORY SIMULATIONS

Sashi Kanth Tadinada and Abhinav Gupta

To be submitted to International Journal of Nuclear Science and Engineering

ABSTRACT

The main focus of the paper is to evaluate seismic fragility of nonlinear connections in large piping systems using elastic analyses. Many studies assessing the damage from the 1971 San Fernando and 1994 North Ridge earthquake report that one of the leading causes of damage in critical facilities like hospitals is because of flooding due to failure of sprinkler piping systems, collapse of nonstructural systems like ventilation ducts, partition walls etc. This is an important problem because many studies analyzing the damage from 1994 North Ridge and 1971 San Fernando earthquakes conclude that 80% of the total economic loss is due to failure of nonstructural systems. Typical piping systems are characterized as: (a) essentially elastic stress-strain behavior; (b) damage due to excessive inelastic deformation at specialized inelastic locations such as T-joint connections, elbows, valves etc. ; (c) low ductility (<3.0) corresponding to failure. Since the structure is largely elastic, we evaluate the effectiveness of using linearization techniques to describe the localized nonlinearities with equivalent elastic properties. The chief motivation to do so is that one of the main limitations in the seismic fragility analyses of structures using probabilistic simulations is the enormous computational costs associated with real life-structures described by intensive nonlinear Finite Element (FE) models.

Most of the current Equivalent Linearization (EQL) techniques are not directly extendable to risk analyses applications as they seek to minimize the average error between the overall

responses of the nonlinear and linear systems. Since it is the maximum responses that indicate the damage in seismic risk analysis, any EQL technique must seek to minimize the error between the maximum responses and not the average response. To address this problem, we introduce a concept called "Equivalent Elastic Limit State". By studying a large number of representative nonstructural systems which are essentially linear but characterized by localized fragile nonlinearities, we propose a model to compute seismic fragility curves using only elastic time-history simulations. It is shown that not only is the seismic fragility predicted using elastic analyses is shown to be quite close to the actual fragility, fragility estimated using this approach can be used as a very reliable prior estimate for Bayesian fragility models so that the final fragility curve can be obtained using a minimum number of nonlinear time-history analyses. The efficacy of the model is demonstrated for seismic fragility analysis of a full-scale floor piping model of San Diego University Hospital with fragile 2" T-joint connections.

1. INTRODUCTION

Seismic reliability of nonstructural systems in critical facilities like hospitals and nuclear power plants is a well-recognized safety and economic consideration. The installation and construction cost of non-structural components like fire protection or water-sprinkler piping, Heating, Ventilating and Air Conditioning (HVAC) ducts etc. can exceed over 80% of the total cost in same facility. Many studies assessing the damage from the 1974 San Fernando and the 1994 North Ridge earthquakes (Ayres and Phillips, 1998; Kircher, 2003; Reitherman, 2009) report that nearly 85% of all economic impact is from the damage due to nonstructural systems alone. In addition, the failure of nonstructural systems has led to flooding and subsequent water damage to equipment thereby, escalating the costs for clean-up, replacement and loss of function. Kennedy (1999) argues that in nuclear power plants, it is the piping systems that control the overall HCLPF capacity (a measure of seismic radiological risk) of the plant because of their wide dispersion and large variability in details leading to multiple number of potential failure locations. Consequently, many engineers have recognized the need to address the problem in design stages such that the nonstructural components remain operational or functional after an earthquake. In recent years, probabilistic seismic fragility studies for these types of structural configurations have become necessary to mitigate risk and achieve reliable designs.

In the framework of Probabilistic Seismic Risk Assessment (PSRA), the reliability of a nonstructural system is expressed in terms of "seismic fragility". Seismic fragility of a

component or a system is described by a curve that gives the conditional probability of failure for a given intensity of earthquake design parameter. Seismic behavior of a nonstructural system depends not only on its own dynamic properties but also significantly on its interaction with the primary structure or the building it is supported on. Influence of mass interaction, non-classical damping, higher mode effects as well as the possibility of tuning of natural frequencies with those of the primary system play a significant role in the resultant response of a nonstructural system subjected to seismic loading (Gupta, 1992). Estimating seismic risk of a large scale piping system in a nuclear power plant or a hospital is therefore unique in each case.

Based on numerous studies of piping behavior in earthquakes, major concerns for damage to piping systems are factors like seismic anchor motion, brittle joints and poor supports (IAEA, 1993). For instance, the Olive View hospital in Los Angeles had to be shut down for several days after 1994 North Ridge earthquake because of water leakage owing to the rupture of elbow joints in Fire-sprinkler piping (FEMA, 1994). In general, it is widely acknowledged that the failure locations that govern the seismic fragility of piping systems are the connection subassemblies that include mechanical couplings, welded joints, threaded screws, elbows, tees as well as nozzles to pump, vessel and valves etc. These critical locations often exhibit complex inelastic mechanical behavior that is quite difficult to model or predict using simple finite element models. In recent years, many experimental test programs were developed in order to understand and characterize the complex behavior of

these localized connections subjected to seismic loading. For example, DeVita (1988) describes a seismic fragility testing facility designed and built to test different kinds of representative piping components as a part of an component fragility test programs commissioned by USNRC and EPRI; Hirose et al (2008) performed seismic fragility tests on motor-operated pressure valves as a part of JNES initiative to obtain seismic capacity data and ensure operability of valves at target acceleration levels. Filiatrault et al (2010) on behalf of MCEER conducted a full-scale experimental seismic fragility test program to evaluate the seismic response and failure mechanisms of different types of steel-studded gypsum partition walls and fire-sprinkler piping systems. The main objective of these tests is to obtain seismic capacities of each component as well as help us understand its force-deformation responses under seismic loading. An accurate FE model for a connection like a threaded piping joint involves dense meshing and use of special contact elements which makes the problem computationally inefficient for risk assessment purposes. Data from fragility based testing can be instrumental in formulating simpler but experimentally-validated nonlinear finite element models that describe the seismic response of the connections with sufficient accuracy. Ju et al (2011) used experimental results from monotonic and cyclic loading tests on 2" black iron threaded T-joint connections performed at SUNY-Buffalo (Dow, 2010) to characterize the (a) maximum allowable value of rotational deformation at the T-joint section to prevent “first leak” damage state; (b) force-displacement and moment-rotation relationships at T-joint section. A non-linear component FE model of the T-joint connection was formulated and validated against the experimental results. In a similar way, it is

customary to integrate the individual experimentally validated non-linear FE models of all the critical components in order to develop a full-scale system level piping FE model.

One common observation in most of the experimental fragility studies is that large-scale piping systems have three distinct properties: (1) the system is predominantly elastic i.e. a substantial proportion of the total system mass participating in the dominant modes of vibration has deformations below the yield limits; (2) the risk of damage is typically due to excessive deformations at the localized nonlinear joint connections; (3) the damage limit states like "First Leak" etc. tend to occur at low ductility levels of ductility ($\mu \approx 2.0$ - 3.0). This means that piping systems can be modeled as essentially elastic structures with fragile connections that exhibit localized nonlinear behavior. The seismic fragility of such a full-scale piping system subject to damage at the inelastic connections or joint locations can be now evaluated by conducting probabilistic simulations. The experimentally-validated Finite Element (FE) model is used to perform nonlinear multiple time-history analyses at different levels of Peak Ground Acceleration (PGA) upon propagating uncertainties in the material, modeling and loading variables of the FE model. The fragility is determined as the ratio of number of analyses in which failure has been observed to the total number of simulations performed. But the Monte-Carlo simulation approach to fragility estimation has one important drawback: it requires a large number of time-history analyses to be performed for accurate estimates of probabilities of failure. This can be a disadvantage in the case of large piping systems whose intricate geometry often leads to high-order, wide-banded mass and

stiffness matrices. In such a case, the FE model can demand significant computational resources even for a single nonlinear time-history analysis. This means that a trade-off is required between the desired accuracy of the risk estimates and the total number of analyses that can be performed in a reasonable time.

Rather than relying solely on the results of the nonlinear time-history simulations, it is sometimes practical to use data from existing studies or simplified approaches to obtain a preliminary estimate of fragility curves for the structure and then use the actual nonlinear time-history simulation data to update these preliminary estimates. The Bayesian Inference techniques are a popular and powerful tool to effectively integrate existing "prior knowledge" of a fragility model with "newly collected data" from the time-history simulations in order to obtain the updated or "posterior" fragility curves. While conventional techniques for fragility like the "factor of safety" method (EPRI, 1994) work as a very good prior estimate for primary structural systems like buildings, bridges, shear walls etc., there is currently no forthright way to obtain reliable preliminary fragility estimates for nonstructural systems like piping, equipment, partition walls etc. Kennedy (1980) computed nominal values for the fragility of piping and equipment systems based on a review of representative seismic design specifications and design-analyses. However, owing to large variability in details as well as considerable influence of the supporting structure on the seismic behavior of nonstructural components like large-scale piping, modeling and generalizing these systems to obtain prior estimates may not be very accurate.

Since the nonlinearity in the piping is lumped among relatively small number of inelastic elements describing the fragile connection, we presume that suitable equivalent linearization of the nonlinear joint can effectively reduce the system to a completely elastic structure. The advantage in using an elastic system for fragility analysis is that it suffices to evaluate the structural responses at a unit PGA such that the responses and fragilities at other PGAs can be obtained by linear scaling thereby significantly reducing the total number of time-history analyses for obtaining the parameters of the fragility curve. As the damage state for fragility of piping connections is specified at low ductility values, we are typically concerned with only marginal levels of nonlinearity until failure. Therefore, equivalent linearization seems a justifiable technique to obtain good initial estimates for the seismic fragility of a full-scale piping system. However, most of the current equivalent linearization techniques work well for Single Degree of Freedom (SDOF) systems or for structures whose response is dominated by a single mode whereas in large MDOF piping systems, different modes tend to contribute to the damage at different locations. Not only that, existing Equivalent Linearization (EQL) techniques are not directly extendible to seismic risk applications as they seek to minimize the average error between the overall responses of the nonlinear and linear systems. Since it is the maximum responses that indicate the damage in a seismic risk analysis, any EQL technique must seek to minimize the error between the maximum responses and not the average response. To address this issue, a concept called equivalent elastic limit state is introduced which proposes to disregard the localized nonlinearity but instead modify the failure capacity of the component such that the eventual

damage probabilities are close to that of the actual nonlinear system. By studying a large number of representative nonstructural systems which are essentially linear but characterized by localized fragile nonlinearities, a comprehensive model is developed to predict the value of the modified failure capacity or "equivalent elastic limit state" of the linearized member.

In this paper, we develop a comprehensive methodology to evaluate the seismic fragility of nonlinear connections in a large-scale piping system based on equivalent elastic time-history simulations. The various steps in the proposed approach are:

(1) Formulate an experimentally-validated FE model of the full-scale piping system

(2) Propagate various kinds of uncertainties in material, modeling and loading variables

(3) Create an equivalent elastic FE model by replacing all the nonlinear elements at the connections with elastic members whose stiffness is equal to the initial stiffness of the nonlinear element.

(4) Calculate the equivalent elastic limit state for each nonlinear element in the structure corresponding to its actual limit deformation for damage

(5) Perform linear time-history analyses using the equivalent elastic FE model from (3) at PGA =1g.

(6) Evaluate fragility estimates at each specified PGA levels using the elastic responses from (5) with respect to the equivalent elastic limit state calculated in (4).

(7) By employing Bayesian methods, update the prior fragility curve obtained from (6) with actual nonlinear time-history simulation data using FE model of (1).

Before studying the technique of equivalent linearization to model nonlinear piping connections, let us briefly review the Bayesian framework for updating the seismic fragilities of a structure and the importance of having as good prior information as possible.

2. BAYESIAN UPDATING OF STRUCTURAL FRAGILITY

Seismic structural fragility $P_f(\lambda)$ is defined as the conditional probability of damage to a structure according to specified failure criteria $G(.)$ at a given intensity λ of earthquake ground motion parameter. The ground motion parameter λ is typically measured in terms of Peak Ground Acceleration (PGA) defined at the support level or at the base (ground level) of the structure although seismic limit load values like Peak Ground Velocity (PGV) or Spectral Acceleration (SA) etc. have been proposed as well.

$$P_f(\lambda) = P[G(.) < 0 \mid PGA = \lambda] \tag{1}$$

The current state of practice in Probabilistic Seismic Risk Assessment (PSRA) framework uses the two-parameter lognormal fragility model (Kennedy et al 1980; Reed et al, 1994) where the real acceleration capacity, A of a structure is modeled as a product of two lognormal random variables as:

$$A = \tilde{a} \cdot \varepsilon_R$$
$$\tilde{a} = a_m \cdot \varepsilon_U \tag{2}$$

where $\tilde{a}$ is the true median acceleration capacity of the structure and ε_R is the random variable denoting the inherent randomness of the capacity about the median. Since the true median capacity $\tilde{a}$ is not exactly known, it is in turn modeled as a product of an estimate of the median capacity a_m and the random variable ε_U denoting the uncertainty in our knowledge of a_m. Both ε_R and ε_U are modeled using the log-normal distribution with a median of unity and log-standard deviations of β_R and β_U respectively. The expression for the fragility curve corresponding to a confidence level α is given by:

$$P_f(\lambda) = \Phi\left(\frac{\ln\left(\lambda / a_m\right) + \beta_U \Phi^{-1}(\alpha)}{\beta_R} \right) \tag{3}$$

The median fragility curve can be written as

$$P_f(\lambda) = \Phi\left(\frac{\ln\left(\lambda / a_m\right)}{\sqrt{\beta_R{}^2 + \beta_U{}^2}} \right) \tag{4}$$

In more general terms, the fragility is expressed as a function of ground motion parameter λ and a vector of random variables - $\boldsymbol{\Theta} = [\theta_1, \theta_2, \ldots]^T$. The shape of the fragility

curve $P_f(\lambda;\boldsymbol{\Theta})$ is determined by our current knowledge of $\boldsymbol{\Theta}$. For example, the random parameter vector $\boldsymbol{\Theta}$ for the two-parameter lognormal model of Equation (2) becomes $\boldsymbol{\Theta}=[\theta_1,\theta_2]^T=[\tilde{a},\beta_r]^T$. Let us consider that our current knowledge of $\boldsymbol{\Theta}$ be denoted by a joint density function $f'_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})$. This joint density function $f'_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})$ is often referred to as "prior" information which is obtained using existing data and professional expertise. If the data either from experimental or numerical studies can be represented by a vector of observations or data - $\mathbf{y}=[y_1,y_2...y_m]^T$. The Bayes theorem can be used to update our knowledge of the random parameters by incorporating the new data $\mathbf{y}$ to obtain the "posterior" information denoted by its updated density function $f''_{\boldsymbol{\Theta}}(\boldsymbol{\Theta}\,|\,\mathbf{y})$:

$$f''_{\boldsymbol{\Theta}}(\boldsymbol{\Theta}\,|\,\mathbf{y})=\frac{L(\mathbf{y}\,|\,\boldsymbol{\Theta})\,f'_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})}{\int L(\mathbf{y}\,|\,\boldsymbol{\Theta})\,f'_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})d\boldsymbol{\Theta}} \tag{5}$$

In Equation (5), $L(\mathbf{y}\,|\,\boldsymbol{\Theta})$ is referred to as the "Likelihood function". The likelihood function denotes the probability that data $\mathbf{y}$ can be observed given that the prior information is $f'_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})$. The fragility data $\mathbf{y}$ when obtained from experimental tests or numerical simulations can be described as a set of three vectors, as in

$$\mathbf{y} \equiv \begin{bmatrix} \begin{Bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{Bmatrix} & \begin{Bmatrix} n_1 \\ n_2 \\ \vdots \\ n_k \end{Bmatrix} & \begin{Bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{Bmatrix} \end{bmatrix} \tag{6}$$

The i^{th} row of the data set $\mathbf{y}$ which is $[\lambda_i, n_i, r_i]$ summarizes the number of times r_i that the damage has been observed out of the total number of structural specimens n_i analyzed or tested at a given PGA = λ_i. In order to obtain the updated fragility given some prior information on the model $f'_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})$ and simulation or experimental fragility data **y**, the first step is to formulate the likelihood function $L(\mathbf{y} \mid \boldsymbol{\Theta})$. The outcome of each experiment at a PGA = λ_i is realized as a binomial random variable with the probability that the structure is likely to be damaged being equal to $P_f{}'(\lambda_i; \boldsymbol{\Theta})$ and the probability that the structure is not expected to be damaged being $\left(1 - P_f{}'(\lambda_i; \boldsymbol{\Theta})\right)$. If the experiments or numerical simulations are conducted at *k* different levels of PGA, the likelihood function $L(\boldsymbol{\Theta} \mid \mathbf{y})$ is written as:

$$L(\boldsymbol{\Theta} \mid \mathbf{y}) = \prod_{i=1}^{k} \binom{n_i}{r_i} P_f{}'(\lambda_i; \boldsymbol{\Theta})^{r_i} \cdot \left(1 - P_f{}'(\lambda_i; \boldsymbol{\Theta})\right)^{n_i - r_i} \tag{7}$$

Substituting Equation (7) into Equation (5) gives us the posterior or updated information of the random parameters - $f''_{\boldsymbol{\Theta}}(\boldsymbol{\Theta} \mid \mathbf{y})$. An updated or posterior density $f''_{\boldsymbol{\Theta}}(\boldsymbol{\Theta} \mid \mathbf{y})$ implies that the fragility estimates have been revised. The updated median fragility curve can be obtained from $f''_{\boldsymbol{\Theta}}(\boldsymbol{\Theta} \mid \mathbf{y})$ by using Equation (7).

$$P_f^{''}(\lambda \mid \mathbf{y}) = \int P_f\ (\lambda;\boldsymbol{\Theta}) f_{\Theta}^{''}(\boldsymbol{\Theta} \mid \mathbf{y})\, d\boldsymbol{\Theta} \tag{8}$$

The rest of the paper concentrates on determining the prior density function of the random parameters $f_{\boldsymbol{\Theta}}^{'}(\boldsymbol{\Theta})$ using Equivalent linearization techniques. Having a good prior fragility estimates within the Bayesian framework are advantageous in three significant ways:

(1) Rather than starting with no information, a prior fragility curve gives us an idea of the range of PGA values where the fragility experiments or nonlinear simulations can be conducted in an optimized fashion in order to maximize the accuracy and reduce total number of simulations required.

(2) Bayesian updating can be very instrumental in making predictions even with limited amounts of nonlinear data and make it possible to gradually monitor the convergence and accuracy level of the fragility curves thus making the planning and execution of nonlinear simulations very efficient.

(3) Engineers are often interested in a particular portion of the fragility curve instead of the entire curve. For example, risk-informed seismic design of components needs involves values of median capacity or the High Confidence Low Probability of Failure (HCLPF) capacity from the lower tail portion of the fragility curve etc. Having an initial estimate of fragility can better point us towards the specific region of fragility curve and perform simulations accordingly.

3. EQUIVALENT LINEARIZATION

Let us first evaluate if Equivalent Linearization (EQL) techniques can be directly applied to obtain a reasonable prior fragility curve for a full-scale piping system with fragile connections with localized nonlinearities. The Stochastic equivalent linearization of non-linear structures subjected to dynamic excitation has been a well-researched topic for over 50 years in the field of random vibrations. The basic idea of equivalent linearization is as follows: Suppose we have a nonlinear Single Degree of Freedom (SDOF) oscillator of mass *m* whose restoring action is described by a function of its displacement and velocity - $F_s(u,\dot{u})$. The equation of motion for the response $u(t)$ may be expressed as

$$m\ddot{u} + F_s(u,\dot{u}) = f(t) \tag{9}$$

where $f(t)$ is a forcing function. In EQL, we seek to determine the coefficients k_{eq} and c_{eq} for a linear SDOF system whose response $v(t)$ is described as:

$$m\ddot{v} + c_{eq}\dot{v} + k_{eq}v = f(t) \tag{10}$$

The k_{eq} and c_{eq} are so determined such that the average error $\varepsilon(u,v)$ between the responses of the nonlinear and linear systems is minimized i.e.

$$\begin{aligned} &\text{Find } k_{eq}, c_{eq} : \\ &\min\ \varepsilon(u,v) = E_t\left[(u(t) - v(t))^2\right] \end{aligned} \tag{11}$$

Equation (11) describes the statistical minimization of the mean squared error between the responses of nonlinear and equivalent linear system but in a more general sense, one can use any appropriate description for the error $\varepsilon(u,v)$ while setting up the problem. The EQL techniques were first formulated independently by Booton (1954), Kazakov (1956) for the problems in control theory and Caughey (1963) for vibratory systems using the criteria similar to Equation (11). Zhang et al (1991) used the mean-squared difference between the potential energies stored in linear and nonlinear springs as the criterion for minimizing the error $\varepsilon(u,v)$ to determine the equivalent stiffness. By 1998, it is estimated that nearly 400 papers had been published on the subject of statistical linearization of nonlinear dynamical systems (Crandall, 2006). Recent advances in the EQL techniques now enable us in principle to linearize arbitrary large FE models. The review paper by Proppe et al (2003) provides an excellent commentary on the relevance and applicability of a variety of existing theoretical and empirical EQL techniques for FE models of structures subjected to stochastic excitations. Proppe et al (2003) acknowledges that most nonlinear models available in commercial FE-codes in practice seem unsuitable for computing the linearization coefficients. Furthermore, almost all existing EQL techniques seek to minimize the "average error" between the responses of actual and the linearized system. But in the seismic fragility analysis, it is the maximum values of response that indicate damage and are of prime interest for evaluating the probabilities of failure. Therefore, an ideal linearization technique for seismic risk analysis problems must attempt to minimize the error between the maximum responses not the error between overall responses. This means most existing EQL techniques do not

generalize very well for obtaining the prior fragility curves of nonlinear structures. Another important technique for linearization of nonlinear structures with hysteretic stiffness elements is based on introducing an "Equivalent Viscous Damper" by equating the energy dissipated per vibration cycle in the nonlinear system to the energy dissipated by the added viscous damper in the linear system (Jacobsen, 1930; Hudson, 1965; Shibata and Sozen, 1976; Dwairi and Kowalsky, 2004). The equivalent viscous damper approach has been widely adapted and is proven to work very well for SDOF systems or structures whose response is dominated by a single mode. But in piping systems which manifest significant higher mode effects, introducing a viscous damper to substitute for localized nonlinearity into an MDOF system is not a straight-forward process. Consequently, any equivalent linearization scheme for probabilistic seismic analysis applications must be able to satisfy two important constraints:

(a) Must aim to minimize the error between the maximum responses of the actual and the linearized systems.

(b) Must be easily extendable to large MDOF structural systems.

Keeping these constraints in mind, we now introduce a concept called "Equivalent elastic limit state" that would enable us to use linear analyses for obtaining reliable initial estimates of fragility.

4. CONCEPT OF EQUIVALENT ELASTIC LIMIT STATE

Consider a nonlinear oscillator and its corresponding equivalent linear approximation as shown in Figure 1. Let the seismic fragility $P_f^{NL}(\lambda)$ of the system in Figure 1(a) be defined as

$$P_f^{NL}(\lambda;U^*) = P\left[u_{\max}(\lambda) > U^* \mid PGA = \lambda\right] \tag{12}$$

where $u_{\max}(\lambda)$ is the random variable representing the peak deformation of the structure when subjected to an earthquake of PGA = λ and U^* denotes the damage limit value or the maximum allowable deformation for the structure to remain undamaged. Consider an approximation to the nonlinear system such that the coefficients k_{eq} and c_{eq} of the equivalent linear system in Figure 1(b) are given as:

$$\begin{aligned} k_{eq} &= k_0 \\ c_{eq} &= c \end{aligned} \tag{13}$$

Equation (13) suggests that the linear approximation of Figure 1(b) for the system in Figure 1(a) is comprised of same damping characteristics and a spring constant equal to the initial stiffness k_0 as the nonlinear system. The force-displacement relationships of both the structures are shown schematically in Figure 2.

Obviously, using simplified linear approximations of Equation (13) without any rigorous analysis will most likely result in inaccurate fragility estimates if the failure

probabilities were to be evaluated with respect to the original damage state U^* that was specified for the nonlinear spring. If the fragility using the responses from the equivalent linear system were to be close to the actual fragility from Equation (12), it necessitates that we might have to evaluate the fragility at a different damage limit state for the linear system. This leads us to defining a fictitious quantity called an "equivalent elastic limit state" corresponding to the original nonlinear limit state U^*. The "equivalent elastic limit state" is the value V^* where the damage limit state for the linear system such that the resultant fragility evaluated from the linear system responses is close to the actual fragility of nonlinear system. Using a damage limit state value of V^*, the fragility of the equivalent linear system may be written as

$$P_f^{lin}\left(\lambda;V^*\right)=P\left[v_{\max}\left(\lambda\right)>V^* \mid PGA=\lambda\right] \tag{14}$$

where $v_{\max}\left(\lambda\right)$ is the random variable denoting the peak deformation in the equivalent linear system when subjected to an earthquake of PGA = λ. Directly using a linear system as per the simplified approximations precludes us of any effort to determine k_{eq} and c_{eq} along the lines of existing EQL techniques. Instead a new unknown quantity of "equivalent elastic limit state" V^* must be determined in order to arrive at satisfactory fragility estimates of the nonlinear system using elastic analyses results of a linear system. Following the definition of "equivalent elastic limit state", the problem is now restated as that of estimating V^*:

$$\begin{aligned} &\textit{Find } V^* \text{ such that :} \\ &P_f^{lin}(\lambda; V^*) \cong P_f^{NL}(\lambda; U^*) \\ &P\left[u_{\max}(\lambda) > U^* \mid \lambda\right] \cong P\left[v_{\max}(\lambda) > V^* \mid \lambda\right] \end{aligned} \tag{15}$$

Equation (15) can be further expressed as an optimization problem which minimizes the total square error between the fragility curves from nonlinear and linear systems in Equations (12) and (14) respectively:

$$\begin{aligned} &\text{find } v = V^* \text{ such that :} \\ &\underset{v \in \Re^+}{\text{minimize}} \int_{\lambda} \left(P_f^{NL}(\lambda; U^*) - P_f^{lin}(\lambda; v)\right)^2 d\lambda \end{aligned} \tag{16}$$

Solving this minimization problem needs *a priori* knowledge of the fragility $P_f^{NL}(\lambda; U^*)$ of the nonlinear system which is not available. However, it might be feasible to obtain a rough estimate of the equivalent elastic limit state V^* corresponding to the actual damage limit state U^* using rudimentary analytical concepts. We now discuss possible approaches for estimating the equivalent elastic limit state and later, the accuracy and applicability of these approaches is examined using simple illustrative examples.

5. ESTIMATING THE EQUIVALENT ELASTIC LIMIT STATE

One can infer from Equation (15) that the value of the equivalent elastic limit state *V** depends on the pre-supposed capacity of the nonlinear spring, U^* but also on how the random variables $u_{\max}(\lambda)$ and $v_{\max}(\lambda)$ are interrelated. Based on the linearization rules

assumed in Equation (13), the relationship between $u_{\max}(\lambda)$ and $v_{\max}(\lambda)$ is simply the relationship between the actual peak inelastic deformation in the structure and the likely resultant peak elastic deformation if the nonlinear stress-strain behavior were to be neglected. Determining the maximum displacement of a nonlinear system from the maximum displacement of the linear system is of particular interest for applications in earthquake-resistant structural design of structures. Engineers develop inelastic spectra for obtaining the maximum displacements of a nonlinear system from the elastic response spectrum to calculate the force reduction factors to be applied in seismic design. We now investigate a simple technique of determining inelastic spectra and discuss how it can be modified to estimate the value of equivalent elastic limit state.

5.1 Equal strain energy approximation

The seminal work in the area of inelastic spectra determination was by Veletsos and Newmark (1960) who studied earthquake response of simple elasto-plastic SDOF systems and compared the peak responses between the inelastic and elastic systems. They proposed what became popular as "Equal Displacement" and "Equal Strain Energy" rules for relating the peak deformation of an inelastic system, $u_{\max}(\lambda)$ with that of peak displacement, $v_{\max}(\lambda)$ of the corresponding elastic system. According to the "Equal Displacement" rule for structures in the low-frequency and medium-frequency range, an elastic and inelastic system shall have approximately the same maximum displacement. The "Equal Strain Energy" rule valid for structures with moderately-high natural frequency, recommends that the principle of

conservation of energies can be used by which the monotonic load-deformation diagram of the elastic system up to the maximum deformation is same as that of elastic-perfectly plastic system subjected to the same excitation. These observations laid the foundation for the widely-adapted Veletsos-Newmark-Hall (VNH) procedure to construct design inelastic spectra from the elastic spectra which forms the basis for all seismic design today.

Piping connections are typically rigid components with low ductility failure capacity. This property of the piping systems qualifies the criteria for the applicability of "Equal Strain Energy" rule which has been shown with some consistency for high-frequency systems as an acceptable measure to obtain inelastic displacement from elastic response values for design purposes. Let us see how the "Equal Strain Energy" criterion can be used for estimating the "equivalent elastic limit state" value. Figure 3 depicts a schematic representation of the Equal Strain Energy rule.

In the "Equal Strain Energy" rule, the approximate nonlinear displacement is obtained by equating the strain energy i.e. the areas under the force-deformation curve as shown in Figure 3. Since we know the value of the actual damage limit state U^*, the same idea can be used to determine a value of V^* by equating the areas as $A_1 = A_2$.

$$A_1 = A_2$$
$$A_2 = \frac{1}{2} k_0 \left(V^*\right)^2 \qquad (17)$$
$$V^* = \sqrt{\frac{2A_1}{k_0}}$$

5.2 Illustrative examples

We now examine the accuracy and robustness of the "Equal Strain Energy" estimator for the equivalent elastic limit state using some illustrative examples. Initially, we consider a small-scale representative structure modeled as a SDOF primary - SDOF secondary type coupled system. The heavy primary system has a mass m_p and is modeled as an elastic structure with constant spring stiffness K_p. The lighter secondary system of mass m_s is connected with a nonlinear spring whose force-displacement relationship is modeled as a bilinear curve with a secondary slope that is $\alpha = 0.1$times the initial slope k_s. The bilinear spring has a damage limit value U^* which is twice the value of yield displacement $\Delta_y = 0.06m$. A damping ratio of $\xi = 0.05$ is assumed for all the modes. The properties of the system are summarized in Figure 4. The properties are chosen to replicate similar dynamic properties as a large piping system with fragile localized nonlinearities: (a) the ratio of masses of secondary to primary systems is quite small (= 0.001); (b) there is a fragile nonlinear connection between the heavy primary and the light secondary systems ensuring that the failure is limited to a small portion of the overall structure; and (c) the failure limit U^* is defined at low levels of ductility $\mu = 2$.

It is desired to calculate the seismic fragility curve for this structure subjected to variations in input ground motion. For this example, a total of 75 real earthquake records normalized to unit PGA are selected for performing the time-history simulations of both nonlinear and linear systems. The first step is to predict the fragility using linear analyses. Multiple elastic analyses of the linearized system are performed at unit PGA by simply replacing the nonlinear spring within the structure with a linear spring of same initial stiffness. N = 75 nonlinear time-history analyses were performed at PGA = 1g. For a bilinear force-displacement relationship as shown in Figure 4, the equivalent elastic limit state based on "Equal Strain Energy" approximation can be derived based on Equation (17) to be:

$$V^* = \Delta_y \sqrt{1+\left(\mu^*-1\right)\left(2+\alpha\left(\mu^*-1\right)\right)} \text{ where } \mu^* = \frac{U^*}{\Delta_y} \tag{18}$$

where μ^* is the ductility value corresponding to the damage limit state displacement U^*. From Equation (18), the value of equivalent elastic limit state for this example is $V^* = 0.1056m$ using $U^* = 0.12m$ or $\mu^* = 2$. From V^*, the fragility curve may be evaluated from Equation (19).

$$\begin{aligned} P_f^{lin}\left(\lambda; V^*\right) &= P\left[v_{\max}\left(\lambda\right) > V^* \mid PGA = \lambda\right] \\ &\approx \frac{\sum_{i=1}^{N} 1\left(\lambda \cdot v_{\max}^i\left(1g\right) > V^*\right)}{N} \end{aligned} \tag{19}$$

where $v_{\max}^i\left(1g\right)$ is the maximum relative displacement of the secondary system from the i^{th} elastic time-history analysis conducted at PGA = 1.0g and 1(.) is the indicator function. Note

that for a full description of the fragility curve, we need $N = 75$ responses at only a single PGA value as the responses at other PGAs may be obtained by scaling. To compare the accuracy of the estimate, we now determine the actual fragility curve using data from multiple nonlinear time history analyses. To obtain a full description of fragility curve, 12 distinct levels of PGA values were selected between 0.25*g* and 3.0*g* at an interval of 0.25*g*. *N* = 75 nonlinear time-history analyses were performed at each PGA level. In each analysis, the maximum response $u_{\max}^{i}(\lambda)$ was recorded to be compared with the damage limit value of U^* to determine failure. At a given PGA level, the actual fragility can now be estimated using the following equation:

$$P_f^{NL}(\lambda;U^*) = P\left[u_{\max}(\lambda) > U^* \mid PGA = \lambda\right] \approx \frac{\sum_{i=1}^{N} 1\left(u_{\max}^{i}(\lambda) > U^*\right)}{N} \tag{20}$$

In Equation (20), $u_{\max}^{i}(\lambda)$ denotes the peak relative displacement of the secondary system in the i^{th} nonlinear time-history analysis conducted at a PGA level of λ. 900 nonlinear time-history analyses (75 earthquakes & 12 PGA levels) were performed to obtain the full description of the fragility curve.

The two fragility curves are plotted in Figure 5 from which one can observe that the fragility curve from 75 elastic analyses is not very close to the actual fragility curve estimated from nonlinear analyses. This implies that the estimate of equivalent elastic limit state as predicted by the "Equal Strain Energy" rule may not the most optimum value. As the

true fragility curve of the structure for this example is available *a priori*, we can solve the optimization problem stated in Equation (16) to determine the how close was the estimate from "Equal Strain Energy" rule to the most optimum value of the equivalent elastic limit state V_{opt}^{*} .

Figure 6 plots the root square error between the fragility curve from the elastic analyses and the actual fragility curve of the structure in this example as a function of the equivalent elastic limit state value V^{*}. For the specific example of Figure 4, it can be seen that $V_{opt}^{*} = 0.250m$ in comparison to the "Equal Strain Energy" criteria value of $0.1056m$. Before we can generalize about overall error in prediction of V^{*} made by the "Equal Strain Energy" criteria, it will be useful to study a wide range of representative systems.

5.3 Effect of tuning

For this exploratory study to assess the robustness of the "Equal Strain Energy" criteria, we considered a variety of SDOF primary – SDOF secondary systems. Coupled systems such as the one considered in Figure (5) have been studied in Gupta and Jaw (1986) and Gupta (1992) who identified that the response of the secondary systems is significantly affected due to factors like (a) tuning between the natural frequencies of primary and secondary systems $\left(\omega_p \approx \omega_s\right)$; (b) non-classical damping; (c) higher mode effects etc. Consequently for the 2-DOF coupled system being considered, it is reasonable to expect that tuning between the primary and secondary systems can affect the value of equivalent elastic limit state as well.

In order to study the effects of tuning, 70 different systems with various combinations of natural frequencies (ω_p, ω_s) were selected to denote varying degrees of tuning between the primary and secondary systems. The fragile nonlinear spring element connecting the primary and secondary system is assumed to be bilinear with the slope ratio $\alpha = 0.1$ similar to the example in Figure 4. The same set of 75 real earthquake records was used to represent the variation in ground motion characteristics. The damage ductility level μ^* for all the cases is considered to be 2.0. For each of the 70 systems, the optimum value of equivalent elastic limit state V_{opt}^* was determined using data from a total of 900×70 = 63000 nonlinear analyses and 75×70=5250 elastic analyses in a similar manner as followed in Section 5.2. The results of this exploratory study are presented in Figure 7. In the figure, the X-axis represents the tuning parameter $\eta = \dfrac{\omega_p - \omega_s}{\omega_p + \omega_s}$ a measure of degree of tuning between primary and secondary systems while the Y-axis plots the non-dimensional parameter $\dfrac{V_{opt}^*}{\Delta_y}$ signifying the ductility level for the optimum equivalent elastic limit state calculated from each fragility analysis. The figure also plots the "Equal Strain Energy" estimate of V^* along with the observed mean trend between $\dfrac{V_{opt}^*}{\Delta_y}$ and η. It can be noted that the "Equal Strain Energy" approximation predicts the V_{opt}^* reasonably well for the detuned systems $(|\eta| > 0.25)$ but is

very inaccurate for tuned systems. The values of $\frac{V_{opt}^*}{\Delta_y}$ are relatively high in the tuned regions i.e. as the absolute value of η approaches 0. For the example structure detailed in Figure 4, the value of tuning parameter is $\eta = 0.05$ (indicating a nearly-tuned system). Hence, the fragility curve from elastic analyses is substantially different from the true fragility curve when the equivalent elastic limit state was calculated from the "equal strain energy" criteria. For $\eta = 0$ or a perfectly tuned system, the equivalent elastic limit state displacement is nearly 7.6 times the yield displacement indicating that tuning with the primary system amplifies the maximum response of an elastic secondary system due to resonance. The high values $\frac{V_{opt}^*}{\Delta_y}$ in the tuned regions is because an inelastic system detunes itself away from the primary system after yield changes activating the secondary stiffness while the tuned elastic system continues to resonate with the primary system thus resulting in a much greater displacements for the elastic system compared to the inelastic system for the same earthquake loading. Since the elastic displacements of tuned coupled systems at all PGAs are greater than the corresponding inelastic displacements, the value of equivalent elastic limit state V_{opt}^* must be greater than the original damage limit state U^* so that Equation (15) is satisfied. For completeness of the equivalent elastic displacement model, we need to sufficiently account for the case where the primary and secondary systems might be tuned or nearly-tuned.

5.4 Using Equivalent Viscous Damping (EVD)

In order to address the large errors pertaining to the "Equal Strain Energy" estimator for equivalent elastic limit state values especially in case of tuning between primary and secondary systems, we now present a formulation for another estimator of V^* that is based on an equivalent linearization technique that explicitly considers vibration of an oscillator at its resonating frequency. In the class of linearization techniques called "Equivalent Viscous Damping (EVD)" methods, the nonlinear spring is replaced by an elastic system with equivalent stiffness and an additional equivalent viscous damper. The additional viscous damping is introduced to account for the energy dissipation in the nonlinear systems due to the hysteretic loops after yielding. The equivalent viscous damping concept was first introduced by Jacobsen (1930). In this work, Jacobsen approximated the steady state solution of a nonlinear SDOF system by equating the energy dissipated by one cycle of sinusoidal response at resonant frequency of a linear system with equivalent viscous damping.

According to Jacobsen (1930), a nonlinear spring with a hysteresis loop shown shaded in Figure 8(a) will have the same peak displacement U^* as an elastic system with the same initial stiffness and an additional equivalent viscous damper whose damping ratio ξ_{eq} can be calculated using Equation (21) can be added to the initial damper of damping ratio ξ.

$$\xi_{eq} = \frac{2}{\pi}\left(\frac{A_1}{A_0}\right) \tag{21}$$

In Equation (21), A_1 is the shaded area in Figure 8(a) representing the total energy dissipated by the nonlinear spring in one cycle of sinusoidal motion at resonance and A_0 is area of the rectangular block enclosing the hysteresis loop. The expression in Equation (21) is derived by equating the area of the shaded regions in Figures 8(a) and 8(b). If the same amount of energy must be dissipated in one cycle, a linear system with neither nonlinearity nor additional viscous damping must vibrate with increased amplitude to account for the energy dissipation in the hysteresis loop of the nonlinear system. These Jacobsen criteria of Equal Strain Energy dissipation in one cycle of sinusoidal motion yields us the value of the required amplitude of motion for the system in Figure 8(c). Note that the system in Figure 8(c) is same as the equivalent system of Figure 1(b) with properties of Equation (13) that was considered to introduce the concept of equivalent elastic limit state. The resultant amplitude V^* corresponding to a given peak displacement U^* of the nonlinear system can be used as an estimator for the equivalent elastic limit state. The energy dissipated ΔE by a viscous damper of damping ratio ξ in a SDOF system of spring stiffness *k* in one cycle of sinusoidal motion at resonant frequency with amplitude *X* is given by:

$$\Delta E = \pi \xi k X^2 \qquad (22)$$

Based on Equation (22), if the area of the shaded regions in Figures 4(b) and 4(c) must be equal, the expression for V^* can be derived as:

$$\pi\left(\xi+\xi_{eq}\right)kU^{*2}=\pi\,\xi\,kV^{*2}=A_1$$

$$V^{*}=U^{*}\cdot\left(1+\frac{\xi_{eq}}{\xi}\right)^{1/2} \tag{23}$$

where ξ_{eq} is the equivalent viscous damping is evaluated from Equation (18) and ξ is the initial damping ratio of the nonlinear system. For the bilinear type of inelasticity shown in Figure 4, Equation (23) reduces to the following expression:

$$\xi_{eq}=\frac{2}{\pi}\left\{\frac{\left(\mu^{*}-1\right)\left(1-\alpha\right)}{\mu^{*}\left(1+\alpha\left(\mu^{*}-1\right)\right)}\right\}$$
$$V^{*}=\mu^{*}\Delta_y\cdot\left(1+\frac{\xi_{eq}}{\xi}\right)^{1/2} \tag{24}$$

In order to evaluate the performance of this estimator for *V**, we can plot the value of *V** from Equation (24) over the simulation data in Figure 7. Substituting the values $\mu^{*}=2\,;\,\alpha=0.1$ as used for the simulations, Equation (24) yields $\frac{V^{*}}{\Delta_y}=4.983$ as an estimate for equivalent elastic limit state ductility based on Jacobsen's EVD approach. This value is plotted in Figure 9 along the same data used in Figure 7. As we can see, the EVD estimator value for V^{*} of 4.98 matches very well with the average $\frac{V^{*}_{opt}}{\Delta_y}$ in the nearly tuned region i.e. for very small values of $\eta=0.025-0.05$. However, for the isolated case of perfect tuning between primary and secondary systems $(\eta=0)$, the EVD approach still produces

substantial errors. Since the observations of Figure 9 are based on the specific case in which the failure in the structure is defined at a damage limit state ductility $\mu^* = 2.0$. Let us examine if the above observations are consistent when we consider a different ductility level for damage. Figure 10 presents a similar plot as Figure 9 comparing the accuracy of both "Equal Strain Energy" and EVD estimators with the simulation results with failure defined at $\mu^* = 3.0$ using the data from the same dynamic analyses of the 70 systems considered in Section 5.3. It can be inferred from the figure that the results from the simulations at damage ductility of 3.0 are identical to the observations made for damage ductility of 2.0 which are: (a) The "Equal Strain Energy" estimator produces good predictions for the equivalent elastic limit state values for detuned systems when $|\eta| > 0.3$; (b) The "equivalent viscous damping" criteria is accurate for nearly-tuned structures when $|\eta| = 0.025 - 0.05$.

5.5 Composite model for predicting the equivalent elastic limit state

Figures 9-10 suggest that we now have two criteria for predicting the values of equivalent elastic limit state V^*, one of which is accurate for tuned systems and the other for the detuned systems. Using the criteria that work well for boundary values of tuning, we now propose a simple interpolation model that makes close predictions for V^* at intermediate levels of tuning between primary and secondary systems. In other words, the equivalent elastic limit state V^* can be varied according to the tuning parameter η using a function of the form:

$$
\begin{aligned}
&V^* = V^*_{ESE} + \left(V^*_{EVD} - V^*_{ESE}\right)\cdot W(\eta) \\
&V^*_{ESE} = V^* \text{ value from "Equal Strain Energy" criteria} \\
&V^*_{EVD} = V^* \text{ value from "Equivalent Viscous Damping" criteria}
\end{aligned}
\tag{25}
$$

where $W(\eta)$ is an interpolating function that varies with the tuning parameter η. In accordance with the results of the simulation study, it is desirable that any candidate function $W(\eta)$ should approach zero for large values of $|\eta|$ and should approach unity when $|\eta|$ is slightly away from zero. Looking at the average trend of $\frac{V^*_{opt}}{\Delta_y}$ in Figures 9-10, we can see that it resembles the shape of a heavy-tailed probability distribution. In this paper, we use the kernel of a Cauchy probability density function can be used to describe $W(\eta)$:

$$
W(\eta) = \left(\frac{1}{r\left(1 + \left(s\eta / r\right)^2\right)} \right) \tag{26}
$$

In Equation (26), *r* and *s* are empirical constants given in Equation (27) determined using least squares fit from the results of 140 fragility analyses comprising of 70 different systems and 2 damage ductility limit states.

$$
\begin{aligned}
r &= 0.54 \\
s &= 15
\end{aligned}
\tag{27}
$$

The values of *r* and *s* are found to be the same for both the damage ductility levels. Using this information, the composite model for predicting the equivalent elastic limit state can be written as a function of degree of tuning η:

$$V^* = V^*_{ESE} + \left(V^*_{EVD} - V^*_{ESE}\right) \cdot \frac{1}{0.54\left(1 + \left(15\eta / 0.54\right)^2\right)} \tag{28}$$

One can use the value of V^* from Equation (28) to evaluate the fragility $P_f^{lin}(\lambda)$ using Equation (19). Figures 11 and 12 indicate that the composite model for evaluating the equivalent elastic limit state values from Equation (28) is in good agreement with the results of the simulation study using the representative primary-secondary type coupled systems. The model performs well even for the perfectly-tuned case of $\eta = 0$. Employing the composite model for equivalent linear limit state, Figure 13 plots the revised predicted fragility curve predicted from linear analyses for the example structure given in Figure 4. It can be noted that using Equation (28) results in a much better prior estimate than using only "equal strain energy" criteria for this structure.

Another notable conclusion from Figures 11 and 12 is that the model is very effective in predicting V^* when $\eta < 0$ than if $\eta > 0$. This is partially due to the choice of $W(\eta)$ as a symmetric function around the origin even though the errors are unsymmetrical. From the definition of the tuning parameter, $\eta < 0$ implies that the natural frequency of the secondary system is greater than the natural frequency of the primary system implying the secondary

system is relatively stiffer than the primary system. Most real-life piping connections like tee-joints, valves etc. are exceedingly rigid compared to the overall piping system. It is unrealistic for a typical piping system to be more rigid than a connection within it. Since the main focus of this paper is to evaluate fragility for failure at rigid localized connections, most practical applications pertain to detuned structures (with a value of η close to -1) while the model is comprehensive and addresses all degrees of tuning.

5.6 Accuracy of resultant prior fragility curves

In this section, we evaluate the accuracy of the fragility curve $P_f^{lin}\left(\lambda;V^*\right)$ predicted using the equivalent elastic limit state concept. Using the standard lognormal approximation for the fragility curve, the parameters median capacity a_m and the log-standard deviation β_R of $P_f^{lin}\left(\lambda;V^*\right)$ is likely to differ from the respective parameters corresponding to the actual fragility $P_f^{Nl}\left(\lambda;U^*\right)$. One main reason for the deviations may be due to the error between V^* from the predicted model and the observed V_{opt}^* from simulation results as can be noted from Figures 11-12. The predictive model of Equation (28) for equivalent elastic limit states is based on two popular approaches of predicting nonlinear response using elastic systems: (a) inelastic spectra; (b) equivalent viscous damping. Keeping in mind that we are primarily interested in only an initial estimate, the model presented in the paper uses basic seminal concepts that are simple to implement and were so chosen in order to avoid any additional computational effort. Of course, many phenomenal improvements have been suggested over

the years in both these areas. The predictive model can be significantly improved by implementing some of the state-of-the art techniques. Another reason for the error can be ascribed to the concept of equivalent elastic limit state itself. The definition of the equivalent elastic limit state implies the intrinsic assumption that for a given earthquake input, if the nonlinear system peak response $u_{\max}(\lambda)$ exceeds U^*, then the corresponding peak response of the elastic system $v_{\max}(\lambda)$ will exceed V^* i.e. $P[v_{\max}(\lambda) > V^* \mid u_{\max}(\lambda) > U^*] = 1.0$. While this is a reasonable assumption, this may not be true at all times. This intrinsic error can be noticed in Figure 6 where the minimum possible error corresponding to V^*_{opt} is equal to zero.

In order to effectively integrate the prior fragility into Bayesian framework, it is necessary to assess the reliability of the predictions pertaining to the equivalent elastic limit state model being used to ascertain the values of fragility of the structure. One way to estimate the error in predictions is by computing the quantity ε_U defined in Equation (2) as the ratio between the true median capacity $\tilde{a}$ (calculated from nonlinear simulations) and the prior median capacity a_m (estimated from elastic analyses). The deviation of ε_U from unity gives us the degree of confidence in the estimated median capacity a_m. We can estimate the statistics of ε_U using the results from the 140 fragility analyses (70 different representative systems and 2 ductility levels) considered earlier in order to validate the model for equivalent elastic limit state model of Equation (28). In each case, ε_U is estimated as

$$\varepsilon_U = \frac{\tilde{a}\left(= \text{median from } P_f^{NL}\left(\lambda; U^*\right)\right)}{a_m\left(= \text{median from } P_f^{lin}\left(\lambda; V^*\right)\right)} \tag{29}$$

Figure 14 shows the cumulative distribution function of ε_U thus obtained. From a set of 140 values, it was found out that the variations in ε_U have a log-standard deviation of $\beta_U = 0.24$. This β_U value represents the uncertainty in the median value a_m when the fragility is estimated using the equivalent elastic limit state model. So, it is recommended that a value of $\beta_U = 0.24$ may be used in Equation (3) for the complete description of fragility curve. In this approach, a constant value of β_U is calculated over the values of ε_U corresponding to all degrees of tuning. From Figures 11-12, one can see that the error is different at different values of tuning. It is possible to formulate a precise model to evaluate β_U as a function of the tuning parameter η. Such a model would require a more detailed validation study with additional example cases.

6. APPLICATION TO HOSPITAL PIPING SYSTEM FRAGILITY

To illustrate the concept of equivalent elastic limit state to a real-life example, we consider the FE model of a hospital piping system shown in Figure 15. Also shown in the figure is a threaded T-joint connection which is a part of the 2" branch of the water-sprinkler system. It is intended to evaluate the seismic fragility for failure at the T-joint connection. A comprehensive fragility analysis of this structure was performed by Ju et al (2011) using

nonlinear time-history simulations. In this section, we repeat the same fragility analysis but within a Bayesian framework using the prior fragility based on the equivalent elastic limit state concept discussed in the previous section.

Detailed experiments conducted on similar type of 2" black iron T-joint components indicate that the "First Leak" failure occurs due to excessive rotations at the connection. The experiments revealed that the "First Leak" in the threaded connection occurs when the rotation at the T-joint section exceeds 0.022 radians (Dow, 2010). Furthermore, Ju et al (2011) used the monotonic and cyclic loading test data to create a FE model for the threaded T-joint connection using nonlinear rotational springs. The force-deformation response of the rotational spring at the T-joint section is shown in Figure 16.

The first step in determining the prior estimate of seismic fragility of the connection is to linearize the structure by replacing the nonlinear spring of Figure 16 in the FE model of the structure with an elastic spring of the same initial stiffness. Elastic analyses using 75 real earthquake records are conducted at the PGA of 1.0g. The next step is determining the equivalent elastic limit state corresponding to the specified damage limit state of $U^* = 0.022$ radians. Since, the mass ratio of the T-joint to the overall piping is very small, it is reasonable to consider the idealization of the piping to a coupled system such as the one considered in Figure 5. Also, the fundamental frequency of the piping system is 1.82 Hz whereas the natural frequency of the threaded T-joint connection assembly is 60 Hz thus making the connection very rigid compared to the main piping system. The tuning parameter η is

computed as $\frac{1.82-60}{1.82+60}=-0.94$. For very low negative values of η, Equation (28) indicates that the $V^* \approx V_{ESE}^*$. From Figure (14), we can see that the Moment-Rotation response up to the "First Leak" failure is essentially a bilinear curve. In this case, the failure damage level ductility for "First Leak" is $\mu^* = U^*/\Delta_y = 3.49$. We now use Equation (18) to evaluate the equivalent elastic limit state value, $V_{ESE}^* = 0.159\ rad$.

The prior fragility is now readily computed employing Equation (19) upon substituting the value of the equivalent elastic limit state as 0.159 *rad*. A lognormal fit for the prior data is also plotted in Figure 17 as the lognormal fragility model is commonly used in practice. Using the prior information in conjunction with the lognormal fragility model of Equation (3), we define the following prior distributions for the random parameters $\tilde{a}$ and β_R representing median capacity and log-standard deviation representing randomness of the capacity about the median value respectively.

$$\begin{aligned} &\tilde{a} \sim \log N\left(median = a_m, \beta_U\right) \\ &a_m = 2.009g \\ &\beta_U = 0.24 \\ &\beta_R \sim N\left(0.4097, COV = 0.2\right) \end{aligned} \tag{30}$$

Nonlinear time-history analyses are performed at 6 different PGA levels – 1, 1.2, 1.5, 2.0, 2.5 and 3.0*g* and the responses are used to revise the probability density functions of random

parameters $\tilde{a}$ and β_R following the Bayesian updating process detailed in Section 2 leading us to the updated or posterior median curves.

Figure 17 shows the final posterior median fragility curve after incorporating results of 360 nonlinear analyses at 6 different PGA levels. Also, shown in the figure is the fragility curve obtained by updating the prior with data from 90 nonlinear simulations. First of all, it can be seen that the prior fragility curve is a good initial estimate in comparison to the final posterior curve considering that the total effort required is merely 75 elastic analyses of the piping system. Secondly, the updated fragility after 90 analyses is quite close to the final posterior curve. By enabling us to compute fragility even with limited amount of simulation data, Bayesian updating is very instrumental in gradual monitoring of the convergence of the posterior fragility curve. This feature can help us assess and conduct of the adequate number of nonlinear time-history analyses until convergence in a measured manner.

7. CONCLUSION

In this paper, we present a technique to obtain preliminary estimates for seismic fragility of nonstructural systems like piping, fire-protection assemblies, air-ventilation ducts etc. using elastic time-history analyses. Typical piping systems are essentially elastic structures but are seismically vulnerable at connections like Tee, elbow joints, valves etc. These localized connections often exhibit inelastic behavior and are prone to damage at low ductility (<3.0)

levels. Hence, piping and other nonstructural systems are usually modeled as large inelastic systems with localized nonlinearities characterized by low ductility limit states for damage. Since the overall structure behaves predominantly elastic and that nonlinearity is restricted to small regions, equivalent linearization of the nonlinear components in the structure seems like an attractive possibility to estimate the fragility of the system using only elastic time-history simulations. In addition to being less computationally intensive compared to nonlinear dynamic analyses, using elastic analyses reduces the total effort since it is sufficient to perform the simulations at a single PGA level. The fragilities at other PGA levels can be readily evaluated by simply scaling the responses. The main goal of this exercise is arrive at a reasonably quick but accurate seismic fragility estimates which can be used as reliable prior information for a Bayesian updating framework. The Bayesian updating framework seeks to update the prior fragility using only limited results from actual nonlinear simulations or real-life experimental fragility data thus minimizing the total effort required in estimating the seismic fragility to a desired degree of accuracy.

Extending the existing equivalent linearization techniques to large MDOF systems is computationally intensive and is many a time not a straight-forward procedure. Moreover, most equivalent linearization techniques aim to minimize the error in mean response between the systems whereas in seismic fragility analyses, it is the peak responses that determine the risk. To address this issue, a concept called equivalent elastic limit state is introduced which proposes to disregard the localized nonlinearity instead modify the failure capacity of the

component such that the damage probabilities evaluate to be close to that of the actual nonlinear system. By studying a large number of representative nonstructural systems which are essentially linear but characterized by localized fragile nonlinearities, a comprehensive model is developed to predict the value of the modified failure capacity or "equivalent elastic limit state" of the linearized member.

Using results from nearly 70000 nonlinear and linear time-history analyses, the suggested model was validated with about 140 different fragility analyses involving 70 representative structural systems. It was shown that it can produce reliable initial or prior fragility estimates of seismic fragility for large piping or nonstructural systems. Also, the degree of accuracy in the model's predictions is also analyzed. The application of the model is illustrated by considering the real-life example of a hospital main piping system which comprises of a 2" water-sprinkler branch with a fragile threaded tee-joint. The prior fragilities obtained using exclusively elastic time-history simulations were found to be quite close to the fragility from actual nonlinear simulations of the piping system. The process of evaluating the updated posterior fragility curve of the piping system using Bayesian inference is presented. Lastly, it is demonstrated the methodology based on prior estimates using equivalent elastic analyses in conjunction with Bayesian updating by incorporating nonlinear simulation data is an efficient approach to minimize the total computational effort of evaluating seismic fragility of critical nonlinear connections in large piping systems.

ACKNOWLEDGEMENTS

The authors on behalf of the Center for Nuclear Power Plant Structures, Equipment and Piping at North Carolina State University sincerely wish to thank National Science Foundation and NCSU College of Engineering for their continued support. Resources for the Center come from sponsoring industrial organizations, National Science Foundation, the Civil Engineering Department, and College of Engineering in the University.

REFERENCES

Ayres, J. M. and Phillips, R. J., "Water Damage in Hospital Resulting from the Northridge Earthquake," ASHRAE Transactions Part 1B, SF-98-14-2, 1286-1296, 1998.

Booton R.C., "Nonlinear control systems with random inputs," IEEE Transactions on Circuit Theory CT-1: 9–18, 1954.

Caughey T.K., "Equivalent linearization techniques," Journal of the Acoustical Society of America, 35:1706–1711, 1963.

Crandall S.H., "A half-century of stochastic equivalent linearization," Structrual Control and Health Monitoring, 13:27-40, 2006.

DeVita, V., "Seismic fragility testing of representative nuclear piping systems," American Society of Mechanical Engineers, Pressure Vessels and Piping Division (Publication) PVP, v 144, p 237-242, 1988.

Dow, J., "Testing and Analysis of Iron and Plastic T-joint in Sprinkler Systems," NEESR-GC: "Simulation of the Seismic Performance of Nonstructural Systems", Available at http://nees.org/site/resources/pdfs/REU2009_Dow_Paper.pdf, 2010.

Dwairi H. and Kowalsky M., "Investigation of Jacobsen's Equivalent Viscous Damping Approach as applied to Displacement-based Seismic Design," Paper No. 228, Proc. 13th World Conference on Earthquake Engineering, Canada, 2004.

Electric Power Research Institute (EPRI), "Methodology for Developing Seismic Fragilities", EPRI TR-103959 (Principal Investigators: Reed, J.W. and Kennedy, R.P.), Electric Power Research Institute, Palo Alto, California, 1994.

Federal Emergency Management Agency (FEMA), "Reducing the Risks of Nonstructural Earthquake Damage: A Practical Guide," FEMA-74, Third Edition, 1994.

Filiatrault A., Mosqueda G., Retamales R., Davies R., Tian Y., and Fuchs J., "Experimental Seismic Fragility of Steel Studded Gypsum Partition Walls and Fire Sprinkler Piping Subsystems," ASCE Conference Proceedings, Structures Congress, pp. 2633-2644, 2010.

Gupta A.K., Response Spectrum Method in Seismic Analysis and Design of Structures, CRC Press, Inc., Boca Raton, FL, 1992.

Gupta A.K. and Jaw J.K., "Seismic Response of non-classically damped systems," Nuclear Engineering and Design, Vol 01, 153-159, 1986.

Hirose, J., Suzuki, K., Ueki, T., Yoshika, N., "Seismic fragility tests of motor operated valve assemblies with safety-related function," American Society of Mechanical Engineers, Pressure Vessels and Piping Division (Publication) PVP, v 8, p 17-23, 2008.

Hudson D.E., "Equivalent Viscous Friction for Hysteretic Systems with Earthquake-Like Excitations," 3rd World Conference on Earthquake Engineering, New Zealand, ppII-185/II-201, 1965.

International Atomic Energy Agency (IAEA), "Probabilistic safety assessment for seismic events," IAEA-TECDOC-724, 1993.

Jacobsen L.S., "Steady Forced Vibrations as Influenced by Damping." ASME Transactions, 52(1): 169-181, 1930.

Ju, B., Tadinada, S.K. and Gupta, A., "Fragility Analysis of Threaded T-joint connections in Hospital Piping Systems," ASME Pressure Vessels and Piping (PVP) Conference, Baltimore, July 2011.

Kazakov I.E., "An approximate method for the statistical investigation of nonlinear systems," *Trudi Voenna-Vozdushnoi Inzheneroi Akademii imeni Professora N.E. Zhukovskogo* (in Russian) 394:1–52, 1954.

Kennedy R.P., "Risk based seismic design criteria," Nuclear Engineering and Design, Volume 192, Issue 2, Pages 117-135, 1999.

Kennedy R.P., Cornell C.A., Campbell R.D., Kaplan, S., Perla H.F., "Probabilistic seismic safety of an existing nuclear power plant," Nuclear Engineering and Design, 59, 315–338, 1980.

Kircher, C. A., "It Makes Dollars and Sense to Improve Nonstructural System Performance," Proceedings of Seminar on Seismic Design, Performance, and Retrofit of Nonstructural Components in Critical Facilities, ATC-29-2, Newport Beach, California, 2003.

Proppe C., Pradlwarter H.J., Schueller G.I., "Equivalent linearization and Monte Carlo simulation in stochastic dynamics," Probabilistic Engineering Mechanics, 18:1–15, 2003.

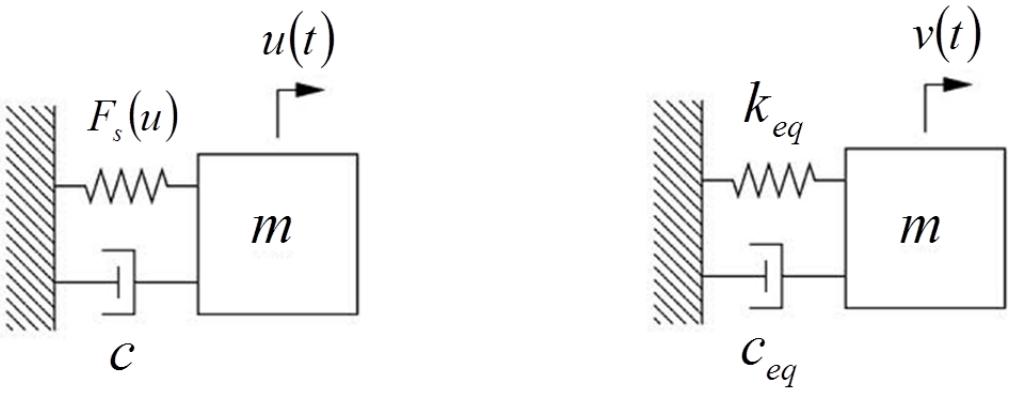
Reitherman, R., "Nonstructural Earthquake Damage," Construction of University for Research in Earthquake Engineering (CUREE), Available at:

http://www.curee.org/image_gallery/calendar/essays/2010-CUREE_excerpt.pdf, 2009.

Shibata A. and Sozen M., "Substitute Structure Method for Seismic Design in R/C." Journal of the Structural Division, ASCE, 102(ST1): 1-18, 1976.

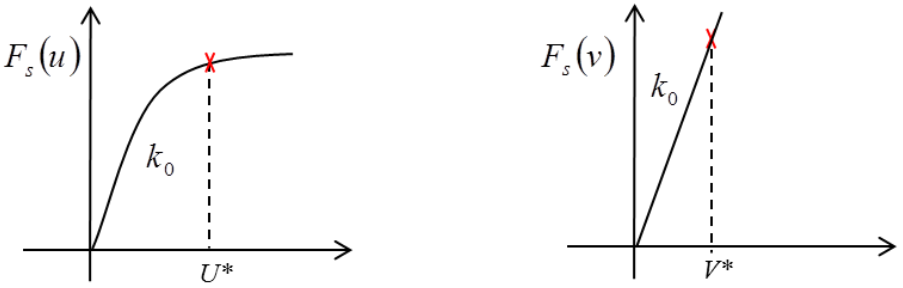
Veletsos, A.S. and Newmark, N.M., "Effects of Inelastic Behavior on the Response of Simple Systems to Earthquake Ground Motions," Proc. 2nd World Conference on Earthquake Engineering, Japan, Vol. II, pp. 895-912, 1960.

Zhang X.T., Elishakoff I, Zhang R.C., "A new stochastic linearization technique based on minimum mean square deviation of potential energies,". Stochastic Structural Dynamics: New Theoretical Developments, Lin YK, Elishakoff I (eds). Springer: Berlin, 327–338, 1991.



(a) *Nonlinear system* (b) *Linear system*

Figure 1: Representative Nonlinear and Equivalent Linear Systems



(a) *Nonlinear system* (b) *Linear system*

Figure 2: Force-Displacement Relationships for both Nonlinear and Linear Systems

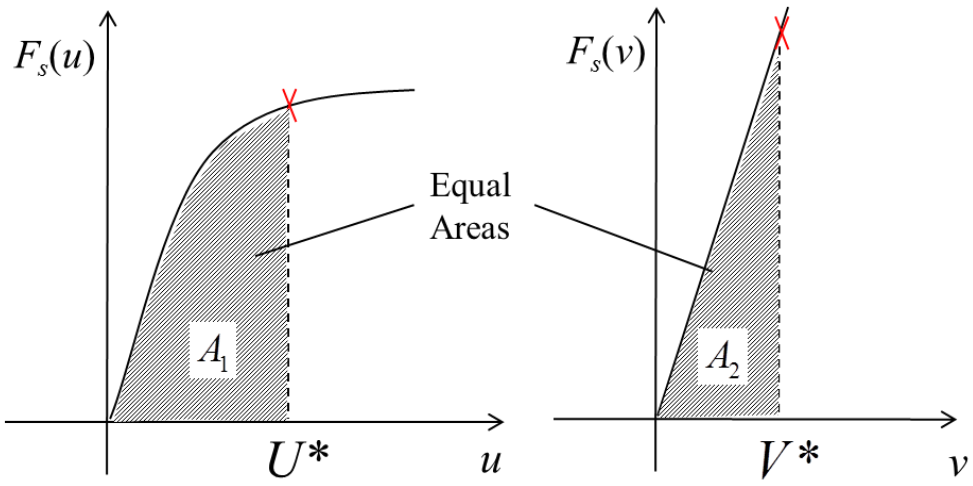


Figure 3: “Equal Areas” approximation for determining the equivalent elastic limit state value

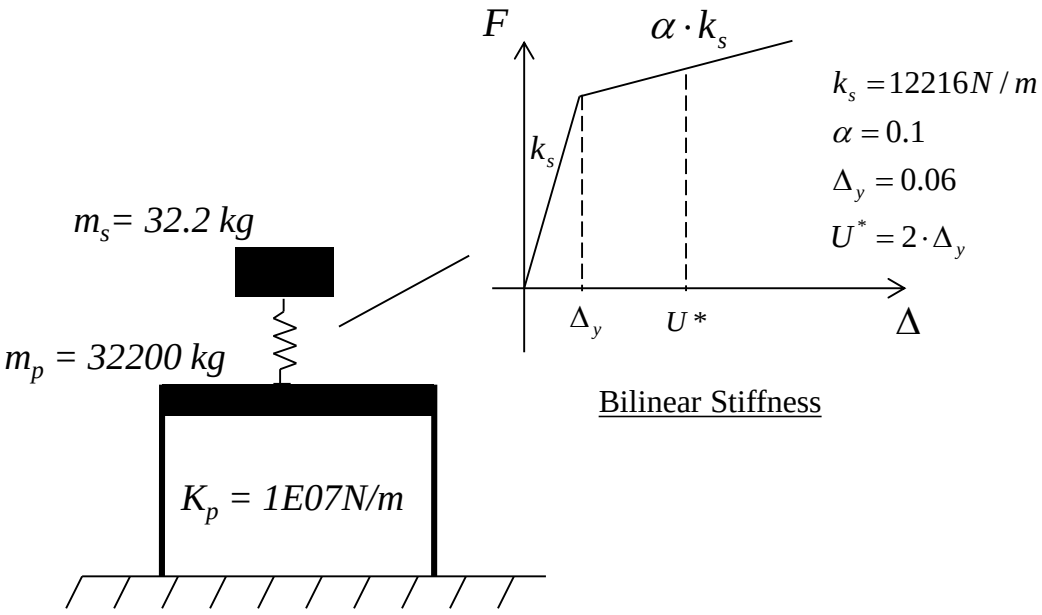


Figure 4: Representative Primary-Secondary Coupled System for the Illustrative Example

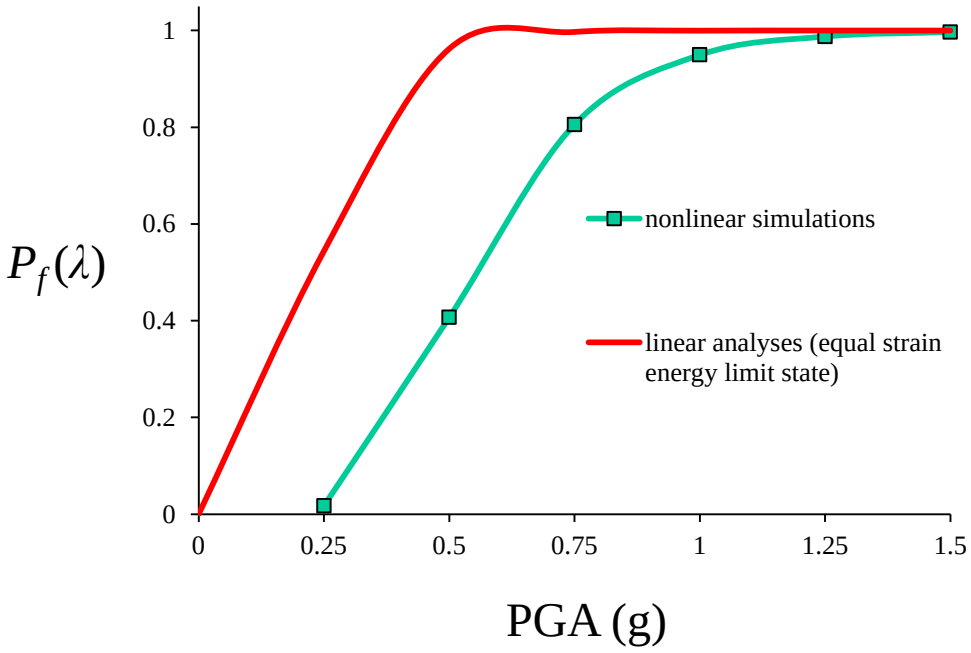


Figure 5: Seismic Fragility Curves from Nonlinear and Linear Time-History Simulations for the Example Structure

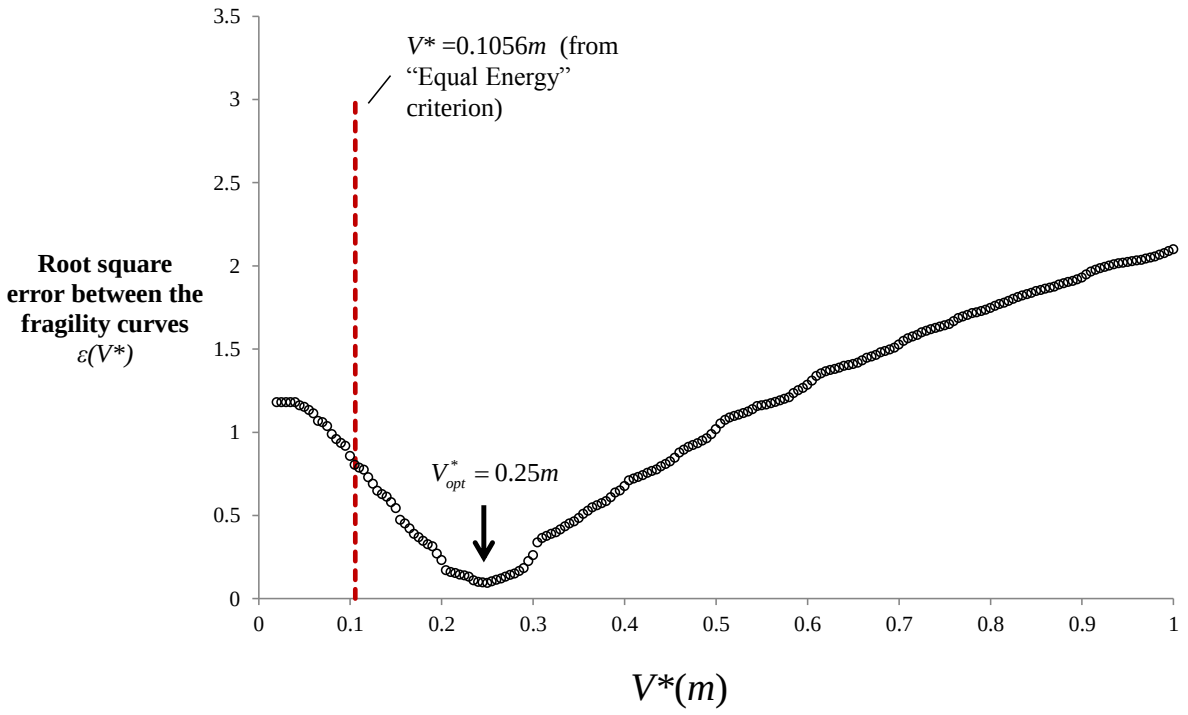


Figure 6: Error between fragility from elastic analyses and the actual fragility for the example structure

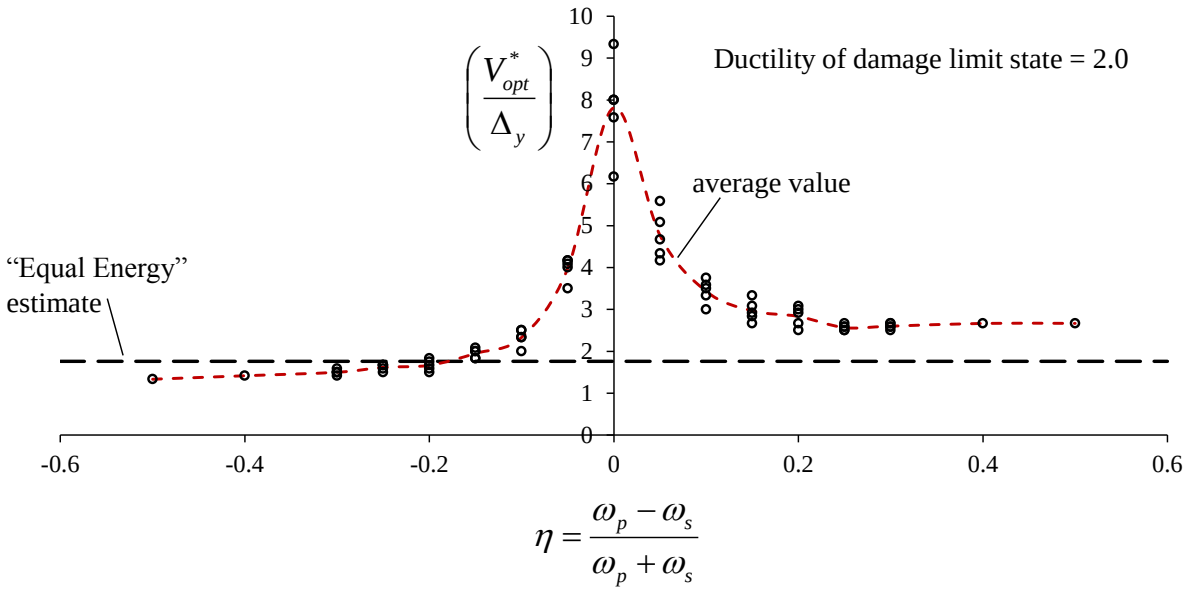


Figure 7: Effect of tuning between primary and secondary systems on the equivalent elastic limit state

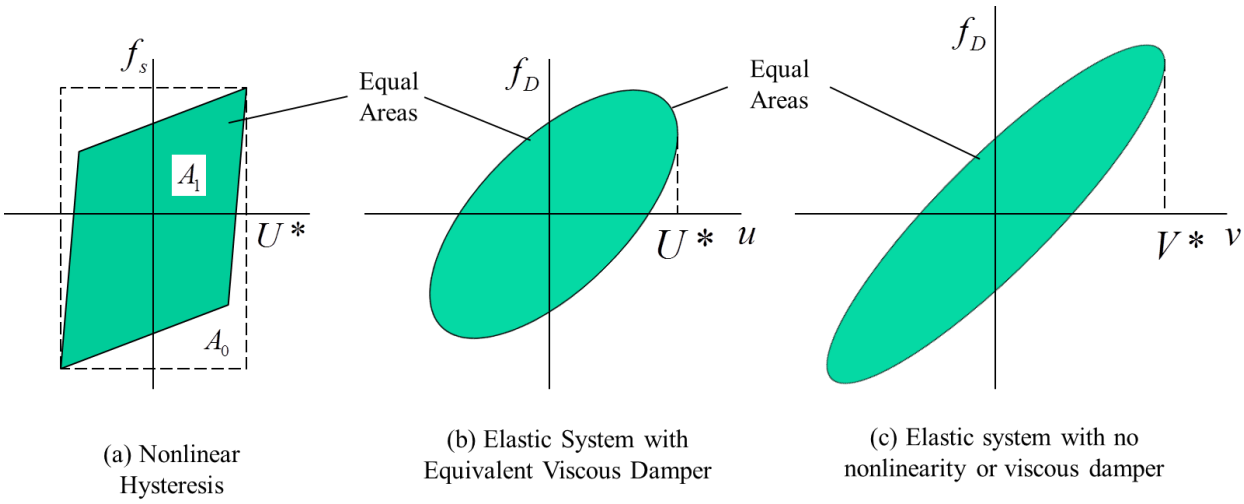


Figure 8: Estimation of equivalent elastic limit state value using the equivalent viscous damping concept

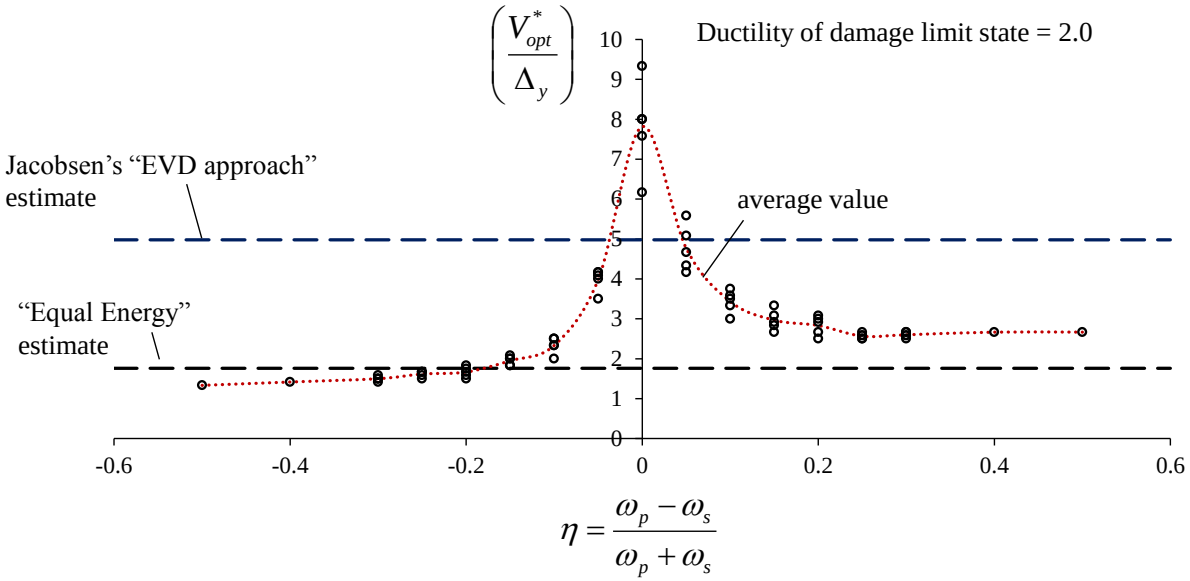


Figure 9: Comparison of equivalent elastic limit states from simulations and estimators (Failure Ductility = 2)

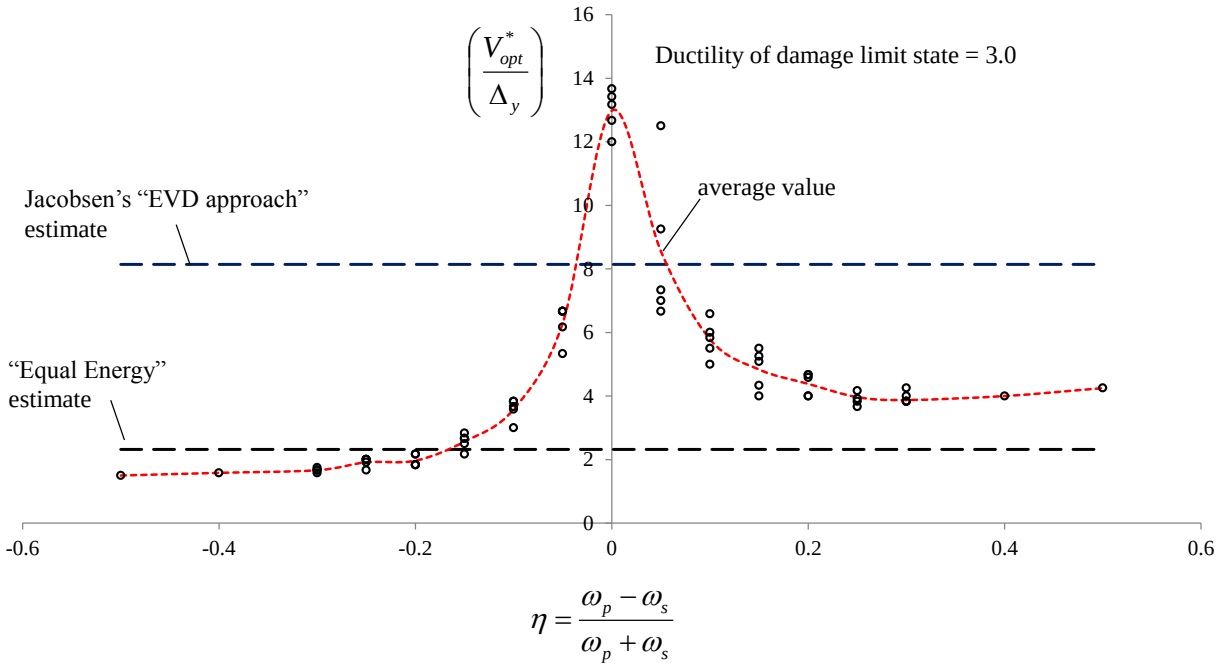


Figure 10: Comparison of equivalent elastic limit states from simulations and estimators (Failure Ductility = 3)

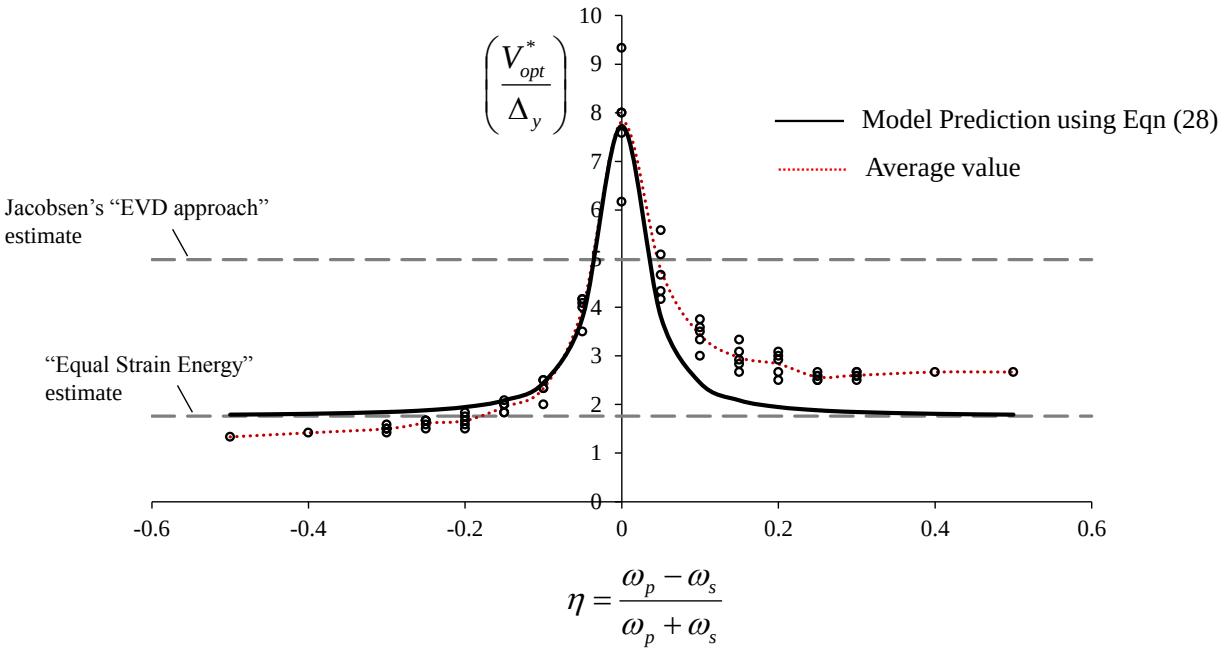


Figure 11: Validation of the model for damage ductility level = 2.0

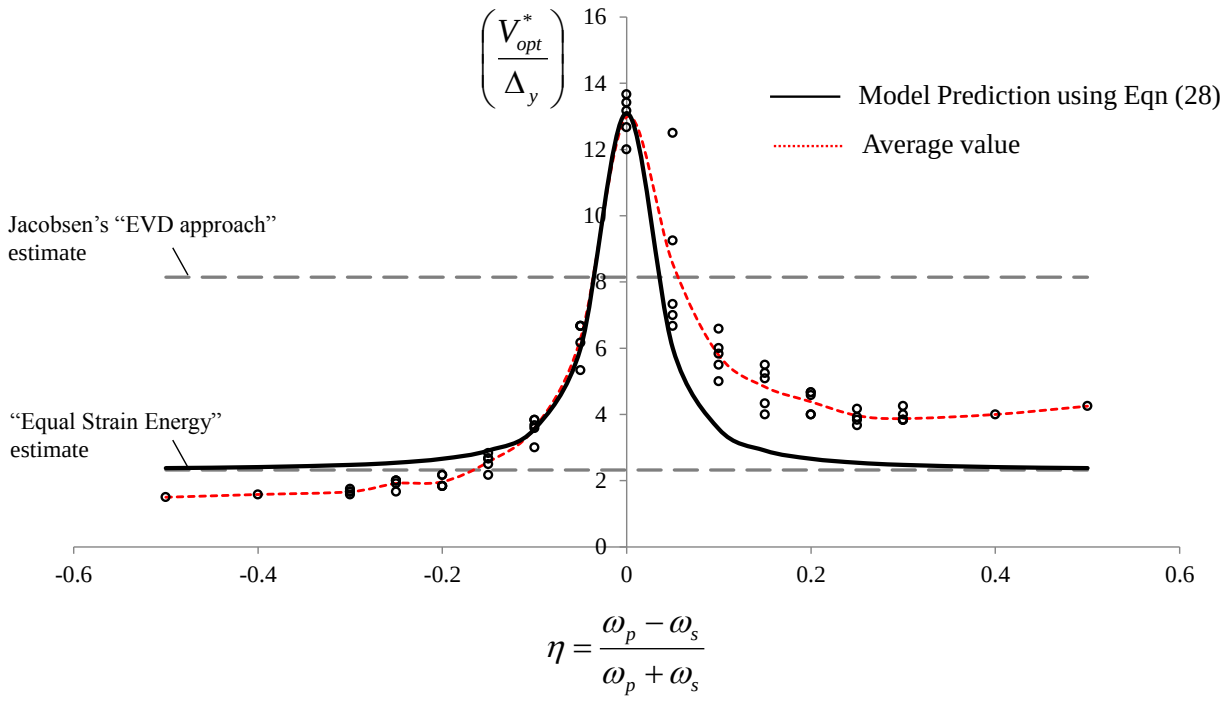


Figure 12: Validation of the model for damage ductility level = 3.0

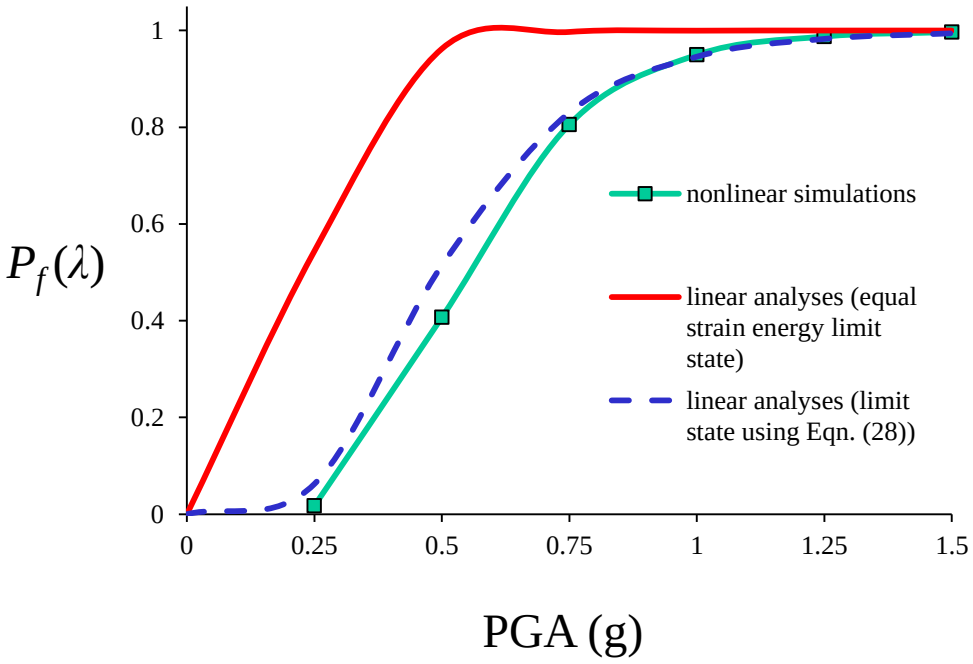


Figure 13: Seismic Fragility Curves from Nonlinear and Equivalent Linear Analyses for the Example Structure

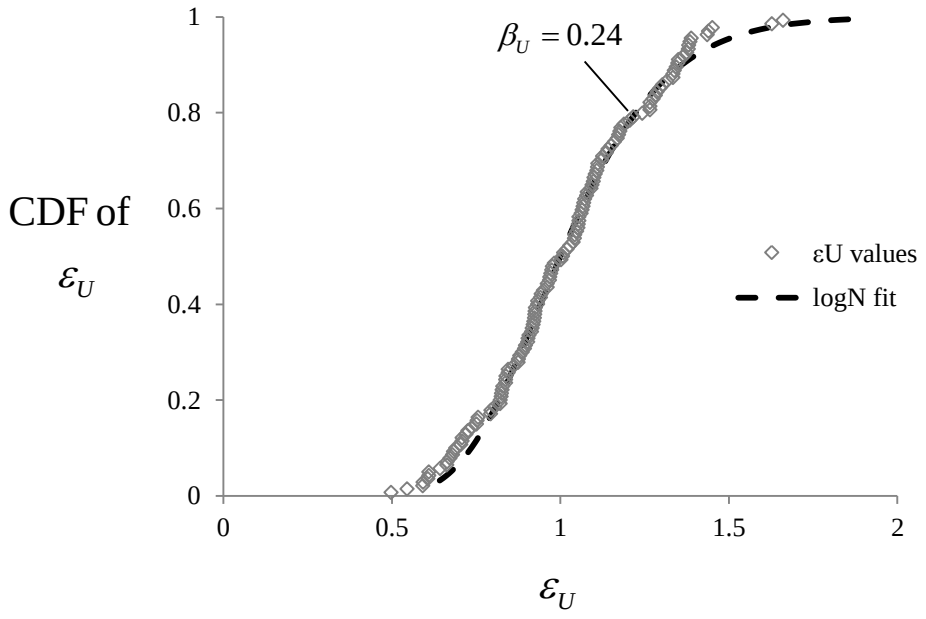


Figure 14: Variations in ε_U

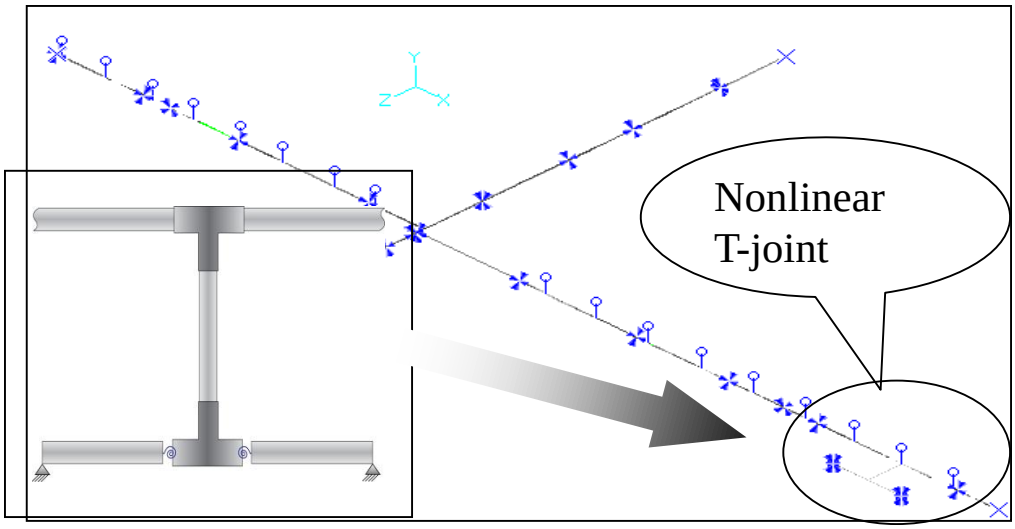


Figure 15: Hospital Main Piping System with a 2-inch Branch Containing the Threaded T-joint

Courtesy: Ju et al (2011)

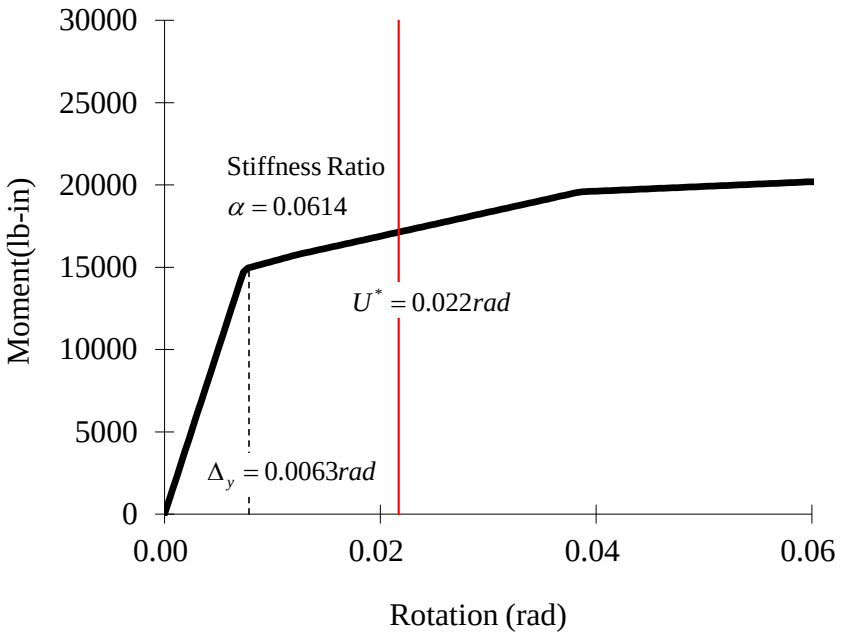


Figure 16: Moment – Rotation Response at the Threaded T-joint Section

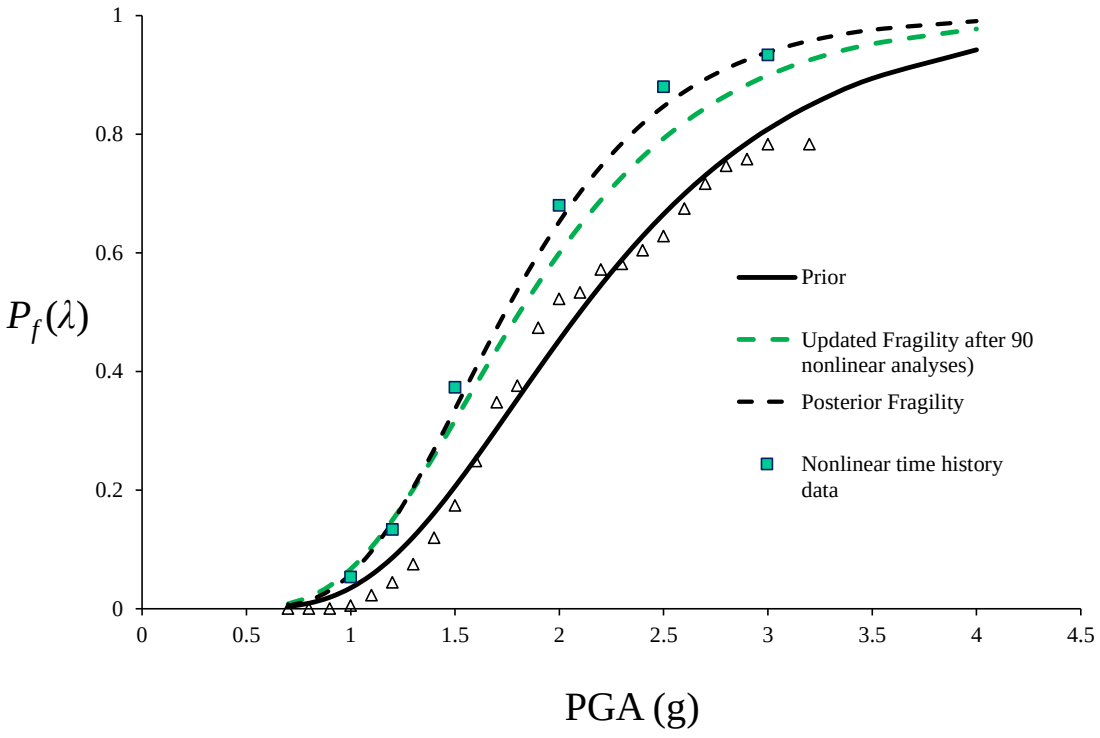


Figure 17: Prior and Updated Median Fragility Curves for the Hospital Piping System

PART IV

SIMULATION OF CONSTRAINED, CLOSELY-SPACED RANDOMA VARIABLES IN ENGINEERING RISK ANALYSES

Sashi Kanth Tadinada and Abhinav Gupta

To be submitted to International Journal of Probabilistic Engineering Mechanics

ABSTRACT

The paper addresses the problem of sampling closely-spaced random variables subject to inequality constraints. Optimal sampling design is one of the key issues in engineering risk analysis. Often, the random parameters and physical variables are either inter-related or ordered and this mutual dependence imposes constraints on the random variables representing the input parameters. Ignoring the interdependencies and sampling these related parameters as independent random variables can lead to inconsistency in sampling with respect to originally assumed distributions for the parameters if the incorrectly simulated sample sets are simply rejected as is usually the case in most engineering risk assessment studies. While solutions have been proposed for sampling of uniformly distributed variables with simple linear inequality constraints, many analyses require characterizing the random variables with probability distributions other than the Uniform distribution. In this paper, we suggest a numerical iterative procedure to derive a joint probability density function surface that can be used to sample closely-spaced random variables characterized by any continuous probability density functions and subject to any kind of non-linear constraints as long as certain conditions required for a solution to exist are satisfied. The sampling procedure is illustrated for three cases where the random variables are respectfully described by Uniform, Lognormal and Gaussian distributions subject to a linear constraint; and one case of uniformly distributed random variables with a non-linear constraint region in shape of a circle. The procedure is shown to work satisfactorily in all cases.

1. BACKGROUND

Availability of increased computational power over the past several years has prompted engineers to conduct large-scale simulations within a Monte Carlo. For example, risk assessment in structural engineering is often based on large-scale finite element models of complex real-life structures and involves consideration of uncertainty/variability in numerous parameters/physical variables. As the size of simulation models and number of random parameters increases, the overall efficiency of a Monte Carlo decreases especially if a single realization itself consumes significant time and computational resources. Another important aspect in such a study involving many different parameters is to ensure that the variability in each parameter is suitably accounted for without a bias in the probabilistic models used to characterize them. Therefore, a trade-off exists between the number of analyses required and the extent to which the randomness in the various parameters is accounted for. This trade-off in the field of probabilistic risk assessment is often addressed by selection of appropriate "Sampling" techniques (Melchers, 1990; Fang et al, 2006). Sampling involves generation of multiple instances of every parameter in accordance with the probabilistic model used to characterize a particular parameter/variable. The premise for selecting a particular "sampling" scheme is typically focused on evaluating as few instances as possible in order to limit the number of large-scale analyses and yet maintain the accuracy if the evaluated risk/fragilities.

Often, the random variables representing various parameters/physical variables in the model tend to be correlated. However, the degree of correlation is either not known or difficult to evaluate. Furthermore, certain variables must be simulated as ordered sets. Consequently, constraints are introduced to maintain certain relationships between the random variables during the sampling process. Some examples of such constraints are

- *Generating a set of natural frequencies of a structure:* The natural frequencies of any structure are ordered and therefore, sampling the frequencies as independent random variables can lead to disordered sets especially if the frequencies are closely-spaced.

- *Simulating the stress-strain curves of a bi-linear material:* Suppose random samples of the yield and ultimate strains in a material modeled with a bi-linear stress curve (with primary slope, k_1 and a secondary slope, k_2) are being generated. Once again, sampling the yield and ultimate strains as independent random variables could lead to a secondary stiffness being higher than the primary stiffness i.e. $k_2 > k_1$ which is unrealistic for most structural materials. Thus, a constraint that $k_1 > k_2$ must be considered.

Simulation of ordered sets of random variables can become formidable particularly if the individual probability density functions (pdfs) overlap significantly.

A key motivation for this study is authors' research on simulation of natural frequencies of a structure. The associated constraint that the simulated set of natural

frequencies be ordered can be addressed directly if a complete description of a joint probability density function of natural frequencies of a structure can be evaluated. Many researchers propose approaches to evaluate the probability density functions of the natural frequencies of a structure by taking into account the uncertainties in the structural mass and stiffness matrices. Some of these approaches include perturbation methods (Soong and Bogdanoff, 1963; Collins and Thomson, 1969; Hart 1973; Song et al, 1995), possibilistic approaches (Gao, 2006; Moderes et al, 2006; Sim et al, 2007), maximum entropy approaches (Sobczyk and Trebicki, 1990; Soize, 2001; Adhikari and Friswell, 2007; Mignolet and Soize, 2008) and random matrix theory (Adhikari, 2006).

Often, there are parameters other than elements of mass and stiffness matrices that can influence a significant variation in the natural frequencies. For example, USNRC (1978) suggests considering a variation of ±15% in the natural frequencies due to factors like soil-structure interaction (Hamilton and Hadjian, 1976; Clinton et al 2006). Therefore, the existing methods can be used accurately to determine the exact density functions if the uncertainty in the structural model is dominated by variations in mass and stiffness matrices but cannot be directly applied in cases when other factors contribute significantly. Moreover, determining a complete joint density function by considering uncertainties in all physical and geometrical parameters of a real-life structure is computationally intensive, problem-specific and impractical in most real-life applications.

At first, it appears that one can use correlation coefficients as a measure of statistical dependence to satisfy the constraints. However, Tadinada (2009) illustrates that although the problem of constraints is minimized by employing an explicit correlation coefficient between the random variables, it does not ensure correct sampling unless the random variables are perfectly correlated.

Therefore, any sampling technique must consider the applicable constraints explicitly. Let Ω_1 and Ω_2 represent two random variables with closely-spaced means $(\mathrm{E}[\Omega_1] < \mathrm{E}[\Omega_2])$ such that their pdfs overlap significantly. Random samples for Ω_1 and Ω_2 are to be generated such that, in any sampled set (ω_1, ω_2), ω_2 is always greater than ω_1. This constraint may be alternatively represented as $P(\omega_1 > \omega_2) = 0$. One obvious way to enforce this constraint is to sample Ω_1 and Ω_2 as independent random variables and ignore all the disordered sample sets i.e. *ignore* $\{(\omega_1, \omega_2): \omega_1 > \omega_2\}$. However, such a direct consideration of the above constraints gives incorrect results in the sense that the individual density functions of Ω_1 and Ω_2 will no longer remain the same as initially assumed. To illustrate this, let us consider $\mathrm{E}[\Omega_1] = 10$ and $\mathrm{E}[\Omega_2] = 11$ to be the means of random variables Ω_1 and Ω_2 representing the two closely-spaced frequencies ω_1 and ω_2. The pdfs of Ω_1 and Ω_2 are simulated by assuming uniform distribution within ±15% of $\mathrm{E}[\Omega_1]$ and $\mathrm{E}[\Omega_2]$ respectively. Those samples (ω_1, ω_2) in which $\omega_1 > \omega_2$ are rejected. As seen in this Figure 1, the simulated distributions are not uniform.

A solution to the problem with direct rejection of incorrect samples was addressed by both Tadinada (2009) as well as Petelet et al (2010). Tadinada (2009) derived an analytical formulation for a joint probability density function $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ between Ω_1 and Ω_2 based on the assumption that the surface of the joint pdf is centrally-symmetric and subsequently used the joint pdf to sample the random variables representing natural frequencies of a structure. Petelet et al (2010) implemented a novel Latin Hypercube Sampling (LHS) algorithm which is based on finding and executing permutations within the sampled set such that the monotonic constraint is satisfied and applied it to sample various random instances of material models for Steel. However, it must be noted that the methods enunciated in both the above referred works are applicable only when the closely-spaced random variables Ω_1 and Ω_2 are uniformly distributed and for a linear monotonic inequality constraint $P(\omega_1 > \omega_2) = 0$

However, a uniform distribution may not always be suitable to model the parameters. In this paper, we present a simple numerical procedure to arrive at a description of the joint probability density function $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ where Ω_1 and Ω_2 can described by any generalized continuous distribution not just a uniform distribution. Also, we illustrate that the same algorithm can be used to sample random variables for any generalized constraint boundary and not just a monotonic linear inequality of the type $\Omega_2 > \Omega_1$.

2. SAMPLING RANDOM VARIABLES WITH AN ORDER CONSTRAINT

Given that Ω_1 and Ω_2 are two real valued continuous random variables defined over (a_1, b_1) and (a_2, b_2) respectively where $-\infty < a_1, a_2, b_1, b_2 < \infty$. If the pdf of Ω_1 is denoted by $f_{\Omega_1}(\omega_1)$ and the pdf of Ω_2 by $f_{\Omega_2}(\omega_2)$, we address the problem of sampling a set of two constrained variables (Ω_1, Ω_2) such that Ω_2 is always greater than Ω_1. No other dependence or correlations between Ω_1 and Ω_2 is assumed to be known. As discussed in the preceding section, implementation of such a constraint requires sampling from a suitably defined joint density function $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2)$ between Ω_1 and Ω_2. The desired properties of the joint density function $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2)$ must be:

- Any random sampled set (ω_1, ω_2) is always ordered i.e. $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2) = 0 \; \forall \omega_1 > \omega_2$.
- The initially assumed distributions for Ω_1 and Ω_2 are maintained.

The problem of determining such a function $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2)$ can be stated more formally as: Find joint density function $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2)$ corresponding to the random variables Ω_1 and Ω_2 such that

$$f_{\Omega_1,\Omega_2}(\omega_1, \omega_2) \geq 0 \; \forall \omega_1, \omega_2 \in \Re \tag{1}$$

$$f_{\Omega_1,\Omega_2}(\omega_1, \omega_2) = 0 \; \textit{if} \; \omega_1 > \omega_2 \tag{2}$$

$$\int_{-\infty}^{\infty} f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)d\omega_1 = f_{\Omega_2}(\omega_2)\,\forall\,\omega_2 \in \Re \tag{3}$$

$$\int_{-\infty}^{\infty} f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)d\omega_2 = f_{\Omega_1}(\omega_1)\,\forall\,\omega_1 \in \Re \tag{4}$$

2.1 Necessary conditions for a solution to exist

It must be noted that a solution for the problem of determining $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ subject to the constraint that $P(\Omega_1 > \Omega_2) = 0$ may not exist for any arbitrary random variables Ω_1 and Ω_2. This section discusses some of the necessary conditions that must be satisfied by the domains (a_1,b_1) and (a_2,b_2) of Ω_1 and Ω_2, so that a solution $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ can exist. More specifically, it is claimed that a solution for $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ exists if and only if: (i) $a_2 > a_1$; and (ii) $b_2 > b_1$. This can be proved as follows:

Suppose $a_2 \le a_1$. We have,

$$\begin{aligned} &F_{\Omega_1}(a_1) = 0;\ F_{\Omega_2}(a_2) = 0 \\ &i.e.\ \ F_{\Omega_2}(a_1) > 0\ if\ \ a_1 > a_2 \end{aligned} \tag{5}$$

It can now be argued that there exists some $a \in (a_2, a_1]$ such that

$$\int_{a_1}^{b_1} f_{\Omega_1,\Omega_2}(\omega_1, a)d\omega_1 = 0 \;\; if \;\; a \le a_1 \;\&\; \omega_1 \in [a_1, b_1]$$
$$\Rightarrow \int_{a_1}^{b_1} f_{\Omega_1,\Omega_2}(\omega_1, a)d\omega_1 \neq f_{\Omega_2}(a) > 0 \tag{6}$$

This is a contradiction since the LHS of Equation (6) evaluates to zero thus violating the condition in Equation (2). Therefore, the condition $a_2 > a_1$ must always be met.

A similar argument can be presented to prove the second condition. Suppose $b_2 \le b_1$. We have,

$$F_{\Omega_1}(b_1) = 1; F_{\Omega_2}(b_2) = 1$$
$$i.e. F_{\Omega_1}(b_2) < 1 \; if \;\; b_1 > b_2 \tag{7}$$

Again, it can be argued that there exists some $b \in (b_2, b_1]$ such that

$$\int_{a_2}^{b_2} f_{\Omega_1,\Omega_2}(b, \omega_2)d\omega_2 = 0 \;\; if \;\; b \ge b_2 \;\&\; \omega_2 \in [a_2, b_2]$$
$$\Rightarrow \int_{a_2}^{b_2} f_{\Omega_1,\Omega_2}(b, \omega_2)d\omega_2 \neq f_{\Omega_1}(b) > 0 \tag{8}$$

Equation (8) again presents a contradiction similar to Equation (6). Hence, a solution for $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2)$ cannot exist if $b_2 < b_1$. Therefore, it can be concluded that the necessary conditions for a solution $f_{\Omega_1,\Omega_2}(\omega_1, \omega_2)$ are: (i) $a_2 > a_1$; and (ii) $b_2 > b_1$.

2.2 Sample space and coordinate transform

Consider the relations defined as follows:

$$\begin{aligned} u &= F_{\Omega_1}(\omega_1) \Rightarrow \omega_1 = F_{\Omega_1}{}^{-1}(u) \\ v &= F_{\Omega_2}(\omega_2) \Rightarrow \omega_2 = F_{\Omega_2}{}^{-1}(v) \end{aligned} \tag{9}$$

where $F_{\Omega_1}(\omega_1)$ and $F_{\Omega_2}(\omega_2)$ are cumulative distribution functions (cdf) of Ω_1 and Ω_2. If we assume that the cdfs are invertible functions (since all continuous distributions have invertible cdfs), then the transformations defined in Equation (9) between (Ω_1, *U*) and (Ω_2, *V*) exhibit one-on-one correspondence. It can be seen that the corresponding random variables U and V are individually U(0,1) i.e. the marginal densities of U and V are given by $f_U(u) = 1 = f_V(v) \; \forall \; u, v \in [0,1]$. Hence, it will suffice to determine the joint density between U and V - $f_{U,V}(u,v)$ and since the domain of U-V space in $\Re^2$ is always $[0,1]^2$, it is more convenient to determine $f_{U,V}(u,v)$ than evaluating $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ directly. Such a joint distribution formulated in terms of corresponding uniformly distributed variables defined using cdfs is referred to as a "Copula". Having determined $f_{U,V}(u,v)$, one can use the transformations in Equation (9) to evaluate back $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$.

From Equation (2), it is evident that no sampled point (ω_1,ω_2) can lie below the line $\omega_2 = \omega_1$. In the transformed co-ordinates, this line is represented by a curve defined by the relation $v = g(u)$ given by

$$g(u) = F_{\Omega_2}\left(F_{\Omega_1}{}^{-1}(u)\right) \forall u \in [0,1] \tag{10}$$

Now, the problem can be restated in terms of U and V as follows:

$$f_{U,V}(u,v) \geq 0 \quad \forall u,v \in [0,1] \tag{11}$$

$$f_{U,V}(u,v) = 0 \ \ if \ \ v < F_{\Omega_2}\left(F_{\Omega_2}{}^{-1}(u)\right) \tag{12}$$

$$\int_0^1 f_{U,V}(u,v)du = f_V(v) = 1 \quad \forall u,v \in [0,1] \tag{13}$$

$$\int_0^1 f_{U,V}(u,v)dv = f_U(u) = 1 \quad \forall u,v \in [0,1] \tag{14}$$

Figure 2 shows a schematic representation of the domain for the joint density function $f_{U,V}(u,v)$. Curve *DE* represents the constraint boundary $v = g(u)$ which separates the constrained region **R** over which $f_{U,V}(u,v)$ = 0. Consequently, the effective sample space for (u,v) is denoted by *ODEBC*. Let us examine the case where U and V are simulated as independent variables, i.e. there is no constrained region **R**. In this case, the joint density function would simply be $f_0(u,v) = f_U(u) \cdot f_V(v) = 1 \ \forall \ u,v \in [0,1]$. The volume enclosed by $f_0(u,v)$ over the entire domain is 1.0. If a constraint exists, then the volume under $f_0(u,v)$ is no longer unity. Therefore, the definition of $f_{U,V}(u,v)$ requires that additional surface is needed to be added over the unconstrained space i.e. *ODEBC* in Figure 2 to make the volume

under $f_{U,V}(u,v)$ equal to unity. If this additional surface or "layer" is denoted by $f_1(u,v)$, we can write

$$f_{U,V}(u,v) = f_0(u,v) + f_1(u,v) \tag{15}$$

It can be said that the surface $f_1(u,v)$ 'makes up' for the volume lost under $f_0(u,v)$ due to the constraint.

Suppose that Ω_1 and Ω_2 are two closely-spaced uniform random variables with respective means μ_1 and μ_2 can be described as follows:

$$\begin{aligned} &\Omega_1 \sim Uniform\left((1-\theta)\mu_1, (1+\theta)\mu_1\right) \\ &\Omega_2 \sim Uniform\left((1-\theta)\mu_2, (1+\theta)\mu_2\right) \\ &where\ \theta \in (0,1);\ 0 < \mu_1 < \mu_2 \end{aligned} \tag{16}$$

The individual density functions $f_{\Omega_1}(\omega_1)$ and $f_{\Omega_2}(\omega_2)$ for this case as shown in Figure 3 are

$$\begin{aligned} f_{\Omega_1}(\omega_1) &= \frac{1}{\left((1+\theta)\mu_1 - (1-\theta)\mu_1\right)} = \frac{1}{B_1} = \frac{1}{2\theta\mu_1} \\ f_{\Omega_2}(\omega_2) &= \frac{1}{\left((1+\theta)\mu_2 - (1-\theta)\mu_2\right)} = \frac{1}{B_2} = \frac{1}{2\theta\mu_2} \end{aligned} \tag{17}$$

In Figure 3, $\overline{\omega}$ denotes the overlap between the two density function and can be expressed as $\overline{\omega} = (1+\theta)\mu_1 - (1-\theta)\mu_2$. Considering the nontrivial case where $\overline{\omega} > 0$ implying that there is a definite overlap between the density functions, we now write the Cumulative Distribution Functions (CDFs) of Ω_1 and Ω_2.

$$F_{\Omega_1}(\omega_1)=\begin{cases}0 \text{ if } \omega_1<(1-\theta)\mu_1\\ \dfrac{(\omega_1-(1-\theta)\mu_1)}{B_1} \text{ if } \omega_1\in[(1-\theta)\mu_1,(1+\theta)\mu_1]\\ 1 \text{ if } \omega_1>(1+\theta)\mu_1\end{cases}$$

$$F_{\Omega_2}(\omega_2)=\begin{cases}0 \text{ if } \omega_2<(1-\theta)\mu_2\\ \dfrac{(\omega_2-(1-\theta)\mu_2)}{B_2} \text{ if } \omega_2\in[(1-\theta)\mu_2,(1+\theta)\mu_2]\\ 1 \text{ if } \omega_2<(1-\theta)\mu_2\end{cases} \tag{18}$$

Substituting Equation (18) into Equation (10) gives us the following expression for the constraint boundary $v=g(u)$:

$$g(u)=\begin{cases}0 \quad \text{for } u\in[0,\rho)\\ \dfrac{(u-\alpha)}{(1-\alpha)}\cdot\dfrac{\overline{\omega}}{B_2} \text{ for } u\in(\rho,1]\end{cases} \tag{19}$$

$$\text{where } \rho=\frac{(1-\theta)\times(\mu_2-\mu_1)}{B_1}$$

Equation (19) indicates that the constraint boundary $g(.)$ or the curve *DE* of Figure 2 is a straight line when Ω_1 and Ω_2 are characterized by uniform distributions. For the instance where $g(.)$ is linear, Tadinada and Gupta (2011) derived the analytical expressions for the joint density $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ by assuming that the surface $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ is centrally-symmetric. Similar to the Equation (15), Tadinada and Gupta (2011) considered the joint

density $f_{U,V}(u,v)$ to be a summation of N additional density surfaces over $f_0(u,v)$ which make up the volume lost due to the constraint $P(\omega_1 > \omega_2) = 0$:

$$f_{U,V}(u,v) = f_0(u,v) + \sum_{k=1}^{N} f_k(u,v)$$

$$N = 1 + \text{int}\left[\frac{\dfrac{1+\theta}{1-\theta} - \left(\dfrac{\mu_2}{\mu_1}\right)}{\left(\dfrac{\mu_2}{\mu_1}\right) - 1} \right] \tag{20}$$

The k^{th} layer $f_k(u,v)$ in the expansion for $f_{U,V}(u,v)$ can be evaluated by the expressions presented in Equation (21).

$$\begin{aligned}
&f_k(u,v) = (B_1 \cdot B_2)\, p_k(x = B_1.u\ ,\ y = B_2.v) \\
&p_k(x,y) = C_k \cdot \int_{y+\alpha}^{\bar{\omega}-(k-1)\alpha} p_{k-1}(x,y)\, dx \ \times \cdots \\
&\qquad\qquad \cdots \int_{x+\alpha}^{\bar{\omega}-(k-1)\alpha} p_{k-1}(x,y)\, dy \quad for\ 0 < x, y < \bar{\omega} - (k-1)\alpha \\
&\text{where } C_k = \frac{1}{\displaystyle\int_0^{\bar{\omega}-(k-1)\alpha} \int_{x+\alpha}^{\bar{\omega}-(k-1)\alpha} p_{k-1}(x,y)\, dx\, dy} \quad \text{and} \quad \alpha = B_1 - \bar{\omega} \\
&p_k(x,y) = p_k(\bar{\omega} - (k-1)\alpha - x, \bar{\omega} - (k-1)\alpha - y) \\
&p_k(x,y) = p_k(B_1 - x, B_2 - y)
\end{aligned} \tag{21}$$

However, if $f_{\Omega_1}(\omega_1)$ or $f_{\Omega_2}(\omega_2)$ are characterized by other probability distributions like Gaussian or Lognormal etc., the constraint boundary $g(.)$ can be a highly non-linear function where obtaining analytical expressions is not so straightforward. In such cases, one

might have to solve for $f_{U,V}(u,v)$ using numerical techniques. This paper proposes and studies an iterative procedure to numerically evaluate a solution surface for $f_{U,V}(u,v)$ such that the conditions of Equations (11) - (14) are satisfied. Having obtained a description of $f_{U,V}(u,v)$, we may use the joint density function to sample Ω_1 and Ω_2 using the relationships in Equation (9).

2.3 Numerical procedure to solve for $f_{U,V}(u,v)$

The domain of both the transformed variables *U* and *V* is [0,1]. Let the domain space in $\Re^2$ for the surface $f_{U,V}(u,v)$ be discretized into a grid as shown in Figure 4. The grid comprises of N_1 divisions along the *U*–axis and N_2 divisions along the *V*-axis. The spacing between the grid lines should be sufficiently close enough for achieving desired accuracy. The spacings between the gridlines along the U-axis are denoted by $\Delta u_i, i = 1,2,\ldots N_1$ and those along the V-axis are denoted by Δv_j, $j = 1,2,\ldots N_2$. The constraint boundary demarcating the constrained region **R** is also shown in Figure 4.

The i^{th} gridline along the *U*-axis is located at $U = u_i$ where $u_i = \sum_{k=1}^{i} \Delta u_k, p = 1,\ldots N_1$ and similarly, the j^{th} gridline along the *V*-axis is located at $V = v_j$ where $v_j = \sum_{k=1}^{j} \Delta v_k, j = 1,\ldots N_2$. Trivially, it can be seen that $u_{N_1} = v_{N_2} = 1$. Also, the entire domain

space (of a total area of 1.0) is divided into $N_1 \times N_2$ grid areas. Each such rectangular division of area can be considered as a "*grid-object*". Each *grid-object* can be indexed by its relative position in the entire grid. Each *grid-object* has the following properties:

- Centroid of the grid object, $C_{ij} \equiv \left(u_i^c, v_j^c\right) \equiv \left(\frac{u_{i-1} + u_i}{2}, \frac{v_{j-1} + v_j}{2} \right)$
- Area, $A_{ij} = \Delta u_i \cdot \Delta v_j$
- An indicator δ_{ij} indicating if the centroid of *grid-object* is in the constrained region **R** or not.

$$\begin{aligned} \delta_{ij} &= 0, C_{ij} \in \mathbf{R} \\ &= 1, C_{ij} \notin \mathbf{R} \end{aligned}$$

Figure 4 displays the location and various properties that can be associated with a $(i,j)^{\text{th}}$ *grid-object*. Having defined the grid and determined the properties of all the associated grid-objects, we now proceed to evaluate the value of the function $f_{U,V}(u,v)$ at centroids of all the $N_1 \times N_2$ grid-objects in the grid layout i.e. at all C_{ij}. It shall be assumed that the value of $f_{U,V}(u,v)$ is constant over the area of each grid-object. This can be written as

$$f_{U,V}(u,v) \cong f_{U,V}\left(C_{ij}\right) \ \forall \ (u,v) \in (i,j)^{\text{th}} \text{ grid} - \text{object} \qquad (22)$$

where C_{ij} is the centroid of $(i,j)^{\text{th}}$ grid-object. In the rest of the paper, $f_{U,V}(C_{ij})$ shall be referred to as f_{ij} for simplicity of notation.

2.3.1 Governing Equations

Using the grid-based approximation for $f_{U,V}(u,v)$ defined in Equation (22), the problem statement as given in Equations (11)-(14) can be expressed as:

$$f_{ij} \geq 0 \quad \forall i = 1 \ldots N_1 \, , \; j = 1 \ldots N_2 \tag{23}$$

$$\delta_{ij} = \begin{cases} 0 \;\; \textit{if} \;\; v_j^c < F_{\Omega_2}\left(F_{\Omega_2}{}^{-1}\left(u_j^c\right)\right) \\ 1 \;\; \textit{otherwise} \end{cases} \quad \forall \; i = 1 \ldots N_1 \, , \; j = 1 \ldots N_2$$

$$f_{ij} = 0 \;\; \textit{whenever} \; \delta_{ij} = 0 \tag{24}$$

$$\sum_{i=1}^{N_1} f_{ij} \cdot A_{ij} = \Delta v_j = \sum_{i=1}^{N_1} A_{ij} \quad \forall j = 1 \ldots N_2 \tag{25}$$

$$\sum_{j=1}^{N_2} f_{ij} \cdot A_{ij} = \Delta u_i = \sum_{j=1}^{N_2} A_{ij} \quad \forall i = 1 \ldots N_1 \tag{26}$$

Equation (25) yields N_2 conditions for each $j = 1 \ldots N_2$ and Equation (26) yields N_1 conditions for each $i = 1 \ldots N_1$ that must be satisfied by any candidate solution for $f_{U,V}(u,v)$.

2.4 The algorithm

1) Divide the entire U-V space in $\Re^2$ into a grid with n_u and n_v gridlines along U and V axes respectively. The spacing of the gridlines should be close enough for sufficient accuracy.

2) Demarcate the constrained region **R** as shown in Figure 4.

3) For each of the $N_1 \times N_2$ grid-objects, determine the centroid C_{ij}, Area A_{ij}, and the constraint-indicator δ_{ij}.

4) Set the initial solution $f_{ij}^{(0)}$ as zero.

$$f_{ij}^{(0)} = 0 \ \forall i = 1 \ldots n_u ; j = 1 \ldots n_v \tag{27}$$

Steps 5 and 6 detail the procedure to evaluate $f_{ij}^{(m)}$ from $f_{ij}^{(m-1)}$ where m = 1,2, … till convergence.

5) Do (i) to (iii) below for each $j = 1,2,\ldots N_2$

i. Calculate the volume $V_{1,j}^{(m-1)}$ enclosed by the $f_{ij}^{(m-1)}$ over the region bounded between (j-1)$^{\text{th}}$ and j^{th} horizontal gridlines i.e. the region between lines $V = v_{j-1}$ and $V = v_j$. Notice that this is the LHS of the Equation (25).

$$V_{1,j}^{(m-1)} = \sum_{i=1}^{N_1} f_{ij}^{(m-1)} A_{ij} \tag{28}$$

ii. In order to check if Equation (25) is satisfied, we now compute the RHS of the Equation (25) which signifies the "target volume" that must be enclosed within horizontal region bounded by $(j-1)^{\text{th}}$ and j^{th} horizontal gridlines. The difference between the RHS and LHS gives the excess volume $\Delta V_{1,j}^{(m-1)}$ needed:

$$\begin{aligned} \Delta V_{1,j}^{(m-1)} &= \sum_{i=1}^{N_1} A_{ij} - V_{1,j}^{(m-1)} \\ &= \sum_{i=1}^{N_1} A_{ij} \left(1 - f_{ij}^{(m-1)}\right) \end{aligned} \tag{29}$$

iii. Distribute $\Delta V_{1,j}^{(m-1)}$ equally over the total entire unconstrained area A_j over the horizontal region bounded by $(j-1)^{\text{th}}$ and j^{th} horizontal gridlines to obtain the new surface $f_{ij}^{(m_1)}$ after correcting the errors in the conditions of Equation (25):

$$\begin{aligned} A_j &= \sum_{i=1}^{N_1} A_{ij} \delta_{ij} \\ f_{ij}^{(m_1)} &= f_{ij}^{(m-1)} + \delta_{ij} \frac{\Delta V_{1,j}^{(m-1)}}{A_j} \\ f_{ij}^{(m_1)} &= f_{ij}^{(m-1)} + \delta_{ij} \cdot \frac{\sum_{i=1}^{N_1} A_{ij} \left(1 - f_{ij}^{(m-1)}\right)}{\sum_{i=1}^{N_1} A_{ij} \delta_{ij}} \end{aligned} \tag{30}$$

6) Do (i) to (iii) below for each $i = 1,2,\ldots N_1$

i. Calculate the volume $V_{2,i}^{(m-1)}$ enclosed by the surface $f_{U,V}^{(m-1)}(u,v)$ over the region bounded by $(i-1)^{\text{th}}$ and i^{th} vertical gridlines i.e. the region between lines $U = u_{i-1}$ and $U = u_i$. Notice that this is the LHS of the Equation (26).

$$V_{2,i}^{(m-1)} = \sum_{j=1}^{N_2} f_{ij}^{(m-1)} A_{ij} \tag{31}$$

ii. In order to check if Equation (26) is satisfied, we now compute the RHS of the Equation (26) which signifies the "target volume" that must be enclosed within vertical region bounded by $(i-1)^{\text{th}}$ and i^{th} vertical gridlines. The difference between the RHS and LHS gives the excess volume $\Delta V_{2,i}^{(m-1)}$ needed:

$$\begin{aligned} \Delta V_{2,i}^{(m-1)} &= \sum_{j=1}^{N_2} A_{ij} - V_{2,i}^{(m-1)} \\ &= \sum_{j=1}^{N_2} A_{ij} \left(1 - f_{ij}^{(m-1)}\right) \end{aligned} \tag{32}$$

iii. Distribute $\Delta V_{2,i}^{(m-1)}$ equally over the entire unconstrained area A_i in the vertical region bounded by $(i-1)^{\text{th}}$ and i^{th} vertical gridlines to obtain the new surface $f_{ij}^{(m_2)}$ after correcting the errors in the conditions of Equation (26):

$$
\begin{aligned}
A_i &= \sum_{j=1}^{N_2} A_{ij}\delta_{ij} \\
f_{ij}^{(m_2)} &= f_{ij}^{(m-1)} + \delta_{ij}\,\frac{\Delta V_{2,i}^{(m-1)}}{A_i} \\
f_{ij}^{(m_2)} &= f_{ij}^{(m-1)} + \delta_{ij}\cdot\frac{\sum_{j=1}^{N_2} A_{ij}\left(1-f_{ij}^{(m-1)}\right)}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}}
\end{aligned}
\tag{33}
$$

7) Step 5 yields a surface $f_{ij}^{(m_1)}$ obtained by adjusting the excess volumes in all regions between the horizontal gridlines according to Equation (25). Step 6 yields $f_{ij}^{(m_2)}$ obtained by adjusting the excess volumes in all regions between the vertical gridlines according to Equation (25). The final surface $f_{U,V}^{(m)}(u,v)$ of m^{th} iteration is calculated by averaging the solutions obtained in Step 5 and Step 6:

$$
\begin{aligned}
f_{ij}^{(m)} &= \frac{f_{ij}^{(m_1)} + f_{ij}^{(m_2)}}{2} \\
f_{ij}^{(m)} &= \max\left(0\,,\; f_{ij}^{(m-1)} + \frac{1}{2}\delta_{ij}\cdot\left\{\frac{\sum_{i=1}^{N_1} A_{ij}\left(1-f_{ij}^{(m-1)}\right)}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij}\left(1-f_{ij}^{(m-1)}\right)}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}}\right\}\right)
\end{aligned}
\tag{34}
$$

Since the value of any probability density function can never be less than zero, Equation (23) must always be enforced while evaluating the $f_{ij}^{(m)}$ using Equation (34).

8) Repeat Steps 5-7 till convergence is achieved. The convergence is reached when all the conditions in Equations (23) – (26) are satisfied.

2.5 Other solutions

The algorithm presented in Section 2.5 attempts to calculate the function $f_{U,V}(u,v)$ at $N_1 \times N_2$ centroid points given by f_{ij} where $i = 1 \ldots N_1$; $j = 1 \ldots N_2$. The Equations (25) - (26) represent a total of $N_1 + N_2$ equations. Suppose that a total of *K* grid-objects belong to the constrained region R. That gives us a total of $N_1 + N_2 + K$ conditions to solve for $N_1 \times N_2$ variables. We can see that a unique solution for $f_{U,V}(u,v)$ will exist only if $K = N_1 \cdot N_2 - (N_1 + N_2)$. Such a condition entails that a large proportion of the sample space be constrained by the region R. To see this more clearly, let $N_1 = N_2 = 100$. Now, for a unique solution, $K = 10^4 - 20$. This implies that essentially 98% of the total U-V space must be constrained. For most practical problems, the number of variables is far greater than the total number of conditions available. Hence, many candidate solutions for $f_{U,V}(u,v)$ might exist. The algorithm presented in this paper arrives at one possible solution for $f_{U,V}(u,v)$. One way to determine alternate solutions is to distribute the excess volumes $\Delta V_{1,i}^{(m-1)}$ and $\Delta V_{2,i}^{(m-1)}$ in Steps 5(i) and 6(i) not equally among the unconstrained areas along the respective horizontal or vertical regions, but to use a different distribution rule. Different 'distribution

rule' could yield different solution surfaces. It will be interesting to see how they would differ in their entropies.

3. APPLICATION EXAMPLES

3.1 Simple cases

In this section, we discuss examples to illustrate the application of algorithm described in the previous section. Firstly, three different cases are considered so that $f_{\Omega_1}(\omega_1)$ and $f_{\Omega_2}(\omega_2)$ are described by various kinds of distributions:

- Ω_1 and Ω_2 are uniformly distributed ; $E[\Omega_2]$ is 10% higher than $E[\Omega_1]$
- Ω_1 and Ω_2 are log-normally distributed ; $E[\Omega_2]$ is 10% higher than $E[\Omega_1]$
- Ω_1 and Ω_2 are Gaussian; $E[\Omega_2]$ is 2% higher than $E[\Omega_1]$

The properties of $f_{\Omega_1}(\omega_1)$ and $f_{\Omega_2}(\omega_2)$ for each of the three cases are listed in Table 1. Ω_1 and Ω_2 can be sampled from their associated copula, $f_{U,V}(u,v)$ which implements the constraint $P(\Omega_1 > \Omega_2) = 0$. The iterative algorithm described earlier may be used to obtain a solution for $f_{U,V}(u,v)$. The complete description of the joint distribution surface $f_{U,V}(u,v)$ can be used to determine the conditional distribution of *V* given *U* - $f_{V|U}(v \mid u)$ which enables

us to generate sample sets (u,v) using any of the well-known sampling techniques like "Inverse CDF" method (Gentle, 2003). Consequently, the sample sets (Ω_1,Ω_2) can be obtained from (u,v) sample sets using the transformations defined in Equation (9).

The results for the three cases are described in Figures 5-7. Each figure has 4 items (a)-(c) which are described as the following: a) a 3-D plot of the associated copula $f_{U,V}(u,v)$; b) a scatter plot of the sampled sets; c) Comparison between simulated and expected PDFs. It can be seen that in all cases each sampled Ω_2 is always greater than the corresponding Ω_1 sample. The simulated distributions in each case were found to be close to the expected distributions. Furthermore, it has been verified in all the three cases that the statistics of the individual variables Ω_1 and Ω_2 pass the Kolmogorov-Smirnov (K-S) test at 3% significance level.

3.2 Application to realistic cases

We now extend the methodology developed in this paper for sampling of sets of multiple random variables in a realistic engineering system with order constraints. Consider an 8-DOF system as shown in Figure 8. The natural frequencies of the structure, ω_i (i=1,2,…8) are 1.0, 1.8, 3.7, 4.4, 6.0, 6.4, 7.4 and 7.6 Hz respectively. It can be observed that the frequency sets (ω_3, ω_4) and (ω_5, ω_6, ω_7, ω_8) are closely-spaced with overlapping individual probability density functions. Hence, the simulation of random sample sets for the uncoupled frequencies cannot be achieved by independent sampling. For any structure with k natural frequencies,

the associated sampling constraint can be expressed as: $\Omega_1 < \Omega_2 < ... < \Omega_k$. In order to sample *k*-dimensional vectors with the constraint that $\Omega_1 < \Omega_2 < ... < \Omega_k$, we use the following sequential procedure:

1. Generate samples for (Ω_1, Ω_2)

2. For each $i = 3,4...k$, generate the samples for Ω_i given the samples for Ω_{i-1} by evaluating the joint probability density function $f_{\Omega_{i-1},\Omega_i}(\omega_{i-1}, \omega_i)$ that considers the constraint that $P(\Omega_{i-1} > \Omega_i) = 0$.

Using this sequential procedure, the sample sets for each of the 8 natural frequencies of the structure in Figure 8 have been generated. It has been observed that each sampled set of natural frequencies were always ordered. Figure 9 compares the probability density functions of sampled frequencies with the expected uniform distributions for structure's frequencies.

4. EXTENSION TO OTHER CONSTRAINTS

The iterative methods described in this paper can be extended to other kinds of constraints between Ω_1 and Ω_2 other than $P(\Omega_1 > \Omega_2) = 0$ discussed so far. Let any real-valued inequality constraint function ψ between two random variables Ω_1 and Ω_2 be given such that:

$$P[\Omega_2 < \psi(\Omega_1)] = 0 \tag{35}$$

Using the transformations in the Equation (4), the above constraint is expressed in *U*-*V* space as

$$\begin{aligned} &P[v < g(u)] = 0 \\ &g(u) = F_{\Omega_2}\left(\psi\left(F_{\Omega_1}{}^{-1}(u)\right)\right) \forall u \in [0,1] \end{aligned} \tag{36}$$

Following the coordinate transformations of section 2.2, any generic constraint function ψ results in a constrained region **R** in the *U*-*V* space. For example, it was seen that the constraint $P(\Omega_2 < \Omega_1) = 0$ i.e. $\psi(\Omega_1) = \Omega_1$ results in a constrained region **R** in the *U*-*V* space as shown in Figure 2. Now, suppose that a constraint function ψ results in a constrained region **R** in the *U*-*V* space as shown in Figure 10. Furthermore, if we define $h(u,v)$ as:

$$\begin{aligned} h(u,v) &= 1 \; if \; (u,v) \in \mathbf{R} \\ &= 0 \; if \; (u,v) \notin \mathbf{R} \end{aligned} \tag{37}$$

It can be concluded using a argument similar to that presented in Section 2.2 that a solution for $f_{U,V}(u,v)$ subject to the constraints of Equation (36) exists only if $h(u,v)$ satisfies the following conditions:

$$\int_0^1 h(u,v)dv < 1 \;\forall u \in [0,1]$$
$$\int_0^1 h(u,v)du < 1 \;\forall v \in [0,1] \tag{38}$$

where $h(u,v)$ is defined according to Equation (37). The conditions in Equation (38) simply indicate that the constrained region **R** cannot span the entire domain of either *U* or *V*.

4.1 Example to illustrate sampling random variables with a nonlinear constraint

Let Ω_1 and Ω_2 both be Uniformly distributed random variables over the interval $[0.1]$. Assume that the nonlinear constraint between Ω_1 and Ω_2 be described as

$$\psi \equiv P\left[(\Omega_1 - 0.5)^2 + (\Omega_2 - 0.5)^2 < 0.2^2\right] = 0\,;\, \Omega_1, \Omega_2 \sim U(0,1) \tag{39}$$

Accordingly, the corresponding Region **R** is a circle in the *U*-*V* transformed coordinates:

$$R \equiv \left\{(u,v) : (u-0.5)^2 + (v-0.5)^2 < 0.2^2;\, u,v \in [0.1]\right\} \tag{40}$$

Since the conditions in Equation (40) are satisfied by region **R**, a solution for $f_{U,V}(u,v)$ can be found. Figure 11 shows the solution surface $f_{U,V}(u,v)$ for this problem using the same algorithm of Section 2.5.

5. CONCLUSION

The paper addresses the problem of sampling closely-spaced random variables subject to inequality constraints. The central idea of the paper is determine a joint probability distribution between the two random variables Ω_1 and Ω_2, $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ such that the constraint $P(\Omega_1 > \Omega_2) = 0$ is always satisfied. The solution for the joint distribution $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ becomes easier by transforming the variable space Ω_1-Ω_2 to their associated copula space U - V. Such a transformation is always one-to-one. Since, the constraint boundary can be non-linear in the associated copula space, obtaining an analytical solution for the joint distribution between U and V, $f_{U,V}(u,v)$ is not always straight-forward; therefore, a simple iterative numerical procedure is presented to obtain the copula $f_{U,V}(u,v)$. Moreover, a numerical description of the joint distribution is required for sampling the random variables anyway. Application of the procedure is demonstrated for 3 different cases:

- Ω_1 and Ω_2 are uniformly distributed ; $E[\Omega_2]$ is 10% higher than $E[\Omega_1]$
- Ω_1 and Ω_2 are log-normally distributed ; $E[\Omega_2]$ is 10% higher than $E[\Omega_1]$
- Ω_1 and Ω_2 are Gaussian; $E[\Omega_2]$ is 2% higher than $E[\Omega_1]$

It can be seen that in all the 3 cases, the constraint is always satisfied while the marginal distributions are close to what are initially assumed. A detailed discussion is presented on the necessary and sufficient conditions on the distributions $f_{\Omega_1}(\omega_1)$ and

$f_{\Omega_2}(\omega_2)$ required for a solution $f_{\Omega_1,\Omega_2}(\omega_1,\omega_2)$ to exist. If in some cases, the distributions do not strictly meet the stipulated conditions, certain approximations are suggested to enable the sampling. It has been observed that these approximations do not affect the accuracy of the procedure in any appreciable manner. Even though a simple monotonic inequality constraint of the type $\Omega_1 < \Omega_2$ is considered in most of the discussion, it is also illustrated that the same logic of the numerical algorithm can be extended to any generalized constraint space as long as the constrained region meets the necessary and sufficient conditions. An example is illustrated for the case of sampling random variables whose non-linear constraint region is circular in shape.

In conclusion, the numerical procedure can be used to sample closely-spaced random variables described by any continuous probability density functions and subject to any kind of non-linear constraints as long as certain conditions required for a solution to exist are satisfied. It must be noted that the conditions arise from the basic probability principles but not as a limitation of the numerical procedure presented.

ACKNOWLEDGEMENTS

This research was partially supported by the Center for Nuclear Power Plant Structures, Equipment and Piping at North Carolina State University. Resources for the Center come from sponsoring industrial organizations, National Science Foundation, the Civil Engineering Department, and College of Engineering in the University. The authors wish to convey

sincere thanks to all the reviewers for their valuable comments during the initial drafts of this manuscript.

REFERENCES

Adhikari S., "Random Eigenvalue problems revisited," *Sādhanā,* Vol. 31, Part 4, pp. 293–314, 2006.

Adhikari S. and Friswell M.I., "Random Matrix eigenvalue problems in structural dynamics," Int. J. Numer. Meth. Engng ; 69:562-591, 2007.

Clinton J.F., Bradford S.C., Heaton T.H. and Favela J., "The observed wander of the natural frequencies in a structure", Bulletin of the Seismological Society of America, Vol. 96, No. 1, 237-257, February 2006.

Collins J.D. and Thomson W.T., "The eigenvalue problem for structural systems with statistical properties," AIAA J 7(4): 642-648, 1969.

Fang K.-T., Li R., and Sudjianto A., "Design and Modeling for Computer Experiments," Chapman & Hall/CRC, 2006.

Gao W., "Natural frequency and mode shape analysis of structures with uncertainty," Mechanical systems and signal processing, 21, 24-39, 2007.

Gentle J.E., "Random Number Generation and Monte Carlo Methods," Springer, 2003.

Hamilton C.W. and Hadijan A.H., "Probabilistic frequency variations of structure-soil systems," Nuclear Engineering and Design, Vol 38, 303-322, 1976.

Hart G.C., "Eigenvalue uncertainties in stressed structures," J. Engrg. Mech. Div., ASCE, 99(EM3), 481-494, 1973.

Melchers, R.E., "Search-based Importance Sampling, " *Structural Safety* v 9, n 2, p 117-128, Dec 1990.

Mignolet M.P. and Soize C., "Nonparametric stochastic modeling of linear systems with prescribed variance of several natural frequencies," Probabilistic Engineering Mechanics, 23, 267-278, 2008.

Modares M., Mullen R.L. and Muhanna R.L., "Natural Frequencies of a Structure with Bounded Uncertainty," Journal of Engineering Mechanics, Vol. 132, No. 12, December 1, 2006.

Petelet M., Iooss B., Asserin O., Loredo A., "Latin hypercube sampling with inequality constraints," *Advances in Statistical Analysis*, submitted. Available at: http://hal.archives-ouvertes.fr/docs/00/52/00/21/PDF/AStA10_iooss_rev2.pdf, 2010.

Sim J., Qiu Z., Wang X., "Modal analysis of structures with uncertain-but-bounded parameters via interval analysis," Journal of sound and vibration, 303, 29-45, 2007.

Sobczyk K. and Trebicki J., "Maximum entropy principle in stochastic dynamics," Probabilistic Engineering Mechanics, Vol. 5, No. 3, 1990.

Soize C., "Transient responses of dynamical systems with random uncertainties," Probabilistic Engineering Mechanics, Vol. 16, Issue 4, 363-372, 2001.

Song D., Chen S., Qiu Z.,"Stochastic sensitivity anlaysis of eigenvalues and eigenvectors," Comput. Struct. 54(5), 891-896, 1995.

Soong T.T. and Bogandoff J.L., "On the Natural Frequencies of a disordered Linear Chain of N Degrees of Freedom," International Journal of Mechanical Sciences, Vol.5, No.3, 237-265, 1963.

Tadinada S.K., "Consideration of Uncertainties in Seismic Analysis of Non-classically damped coupled systems," Unpub. Master's Thesis, North Carolina State University, 2009.

U.S. Nuclear Regulatory Commission (USNRC), Regulatory Guide, Revision 1.122 February, 1978.

APPENDIX

APPENDIX-A: CONVERGENCE ANALYSIS FOR THE ALGORITHM

We now examine the conditions under which the proposed algorithm of Section 2.5 converges to a solution for $f_{U,V}(u,v)$. Let us consider the solution after m^{th} iteration as given by Equation (28). Initially, we assume that $f_{ij}^{(m)} \geq 0$ so that the expression for $f_{ij}^{(m)}$ is given as:

$$f_{ij}^{(m)} = f_{ij}^{(m-1)} + \frac{1}{2}\delta_{ij} \cdot \left\{ \frac{\sum_{i=1}^{N_1} A_{ij}\left(1 - f_{ij}^{(m-1)}\right)}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij}\left(1 - f_{ij}^{(m-1)}\right)}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \right\} \tag{A-1}$$

If we define the incremental change $\varepsilon_{ij}^{(m)}$ in the value of f_{ij} in the m^{th} iteration as

$$\varepsilon_{ij}^{(m)} = f_{ij}^{(m)} - f_{ij}^{(m-1)} \tag{A-2}$$

Equations (A-1) and (A-2) gives us

$$\varepsilon_{ij}^{(m)} = \frac{1}{2}\delta_{ij} \cdot \left\{ \frac{\sum_{i=1}^{N_1} A_{ij}\left(1 - f_{ij}^{(m-1)}\right)}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij}\left(1 - f_{ij}^{(m-1)}\right)}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \right\} \tag{A-3}$$

Since $f_{ij}^{(0)} = 0$ (according to Step 4 of the algorithm in Section 2.5), Equation (A-3) gives us the value of $\varepsilon_{ij}^{(1)}$ as

$$\varepsilon_{ij}^{(1)} = \frac{1}{2}\delta_{ij} \cdot \left\{ \frac{\sum_{i=1}^{N_1} A_{ij}}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij}}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \right\} = \frac{1}{2}\delta_{ij} \cdot \left(\theta_i^{(0)} + \gamma_j^{(0)}\right) \tag{A-4}$$

where $\theta_j^{(0)}$ and $\gamma_j^{(0)}$ are positive constants given by the following expressions:

$$\theta_i^{(0)} = \frac{\sum_{j=1}^{N_2} A_{ij}}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \;\; ; \;\; \gamma_j^{(0)} = \frac{\sum_{i=1}^{N_1} A_{ij}}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} \tag{A-5}$$

Furthermore, $f_{ij}^{(0)} = 0$ implies that we can express the value of $f_{ij}^{(m)}$ as sum of all increments up to the m^{th} iteration i.e. $f_{ij}^{(m)} = \sum_{t=1}^{m} \varepsilon_{ij}^{(t)}$. Using this condition in Equation (A-3) enables us to express $\varepsilon_{ij}^{(m)}$ in terms of $\varepsilon_{ij}^{(m-1)}$.

$$\varepsilon_{ij}^{(m)} = \frac{1}{2}\delta_{ij} \cdot \left\{ \frac{\sum_{i=1}^{N_1} A_{ij}\left(1 - \sum_{t=1}^{m-1} \varepsilon_{ij}^{(t)}\right)}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij}\left(1 - \sum_{t=1}^{m-1} \varepsilon_{ij}^{(t)}\right)}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \right\} \forall\ m > 1 \tag{A-6}$$

$$\varepsilon_{ij}^{(m)} = \varepsilon_{ij}^{(m-1)} - \frac{1}{2}\delta_{ij} \cdot \frac{\sum_{i=1}^{N_1} A_{ij} \cdot \varepsilon_{ij}^{(m-1)}}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij} \cdot \varepsilon_{ij}^{(m-1)}}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \;\; \forall\, m > 1 \tag{A-7}$$

In order to conclude that the algorithm convergences, we need to prove that $\varepsilon_{ij}^{(m)}$ gets smaller and smaller as the number of iterations *m* increase i.e. $\lim_{m\to\infty} \varepsilon_{ij}^{(m)} \to 0$.

Let us now examine the expression for $\varepsilon_{ij}^{(2)}$ which can be written from $\varepsilon_{ij}^{(1)}$ using Equation (A-7):

$$\varepsilon_{ij}^{(2)} = \frac{1}{2}\delta_{ij}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right) - \frac{1}{2}\delta_{ij}\cdot\left\{\frac{\sum_{i=1}^{N_1} A_{ij}\cdot\frac{1}{2}\delta_{ij}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} + \frac{\sum_{j=1}^{N_2} A_{ij}\cdot\frac{1}{2}\delta_{ij}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}}\right\} \quad \text{(A-8)}$$

$$\varepsilon_{ij}^{(2)} = \frac{1}{2}\delta_{ij}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right) - \frac{1}{2}\delta_{ij}\cdot\left\{\sum_{i=1}^{N_1}\left(\frac{A_{ij}\delta_{ij}}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}}\right)\frac{1}{2}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right) + \sum_{j=1}^{N_2}\left(\frac{A_{ij}\delta_{ij}}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}}\right)\frac{1}{2}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)\right\} \quad \text{(A-9)}$$

Defining λ_{ij} and μ_{ij} as:

$$\lambda_{ij} = \frac{A_{ij}\delta_{ij}}{\sum_{i=1}^{N_1} A_{ij}\delta_{ij}} \geq 0 \;\; ; \;\; \mu_{ij} = \frac{A_{ij}\delta_{ij}}{\sum_{j=1}^{N_2} A_{ij}\delta_{ij}} \geq 0 \tag{A-10}$$

Substituting Equation (A-8) into Equation (A-9) gives us:

$$\varepsilon_{ij}^{(2)} = \delta_{ij} \cdot \frac{\left(\theta_i^{(0)} + \gamma_j^{(0)}\right)}{2} - \frac{1}{2}\delta_{ij} \cdot \left\{ \sum_{i=1}^{N_1} \lambda_{ij} \frac{\left(\theta_i^{(0)} + \gamma_j^{(0)}\right)}{2} + \sum_{j=1}^{N_2} \mu_{ij} \frac{\left(\theta_i^{(0)} + \gamma_j^{(0)}\right)}{2} \right\} \tag{A-11}$$

Rewriting Equation (A-11) as

$$\varepsilon_{ij}^{(2)} = \delta_{ij} \cdot \frac{\left(\theta_i^{(0)} + \gamma_j^{(0)}\right)}{2} - \frac{1}{2}\delta_{ij} \mathrm{H}\left(\frac{\theta_i^{(0)} + \gamma_j^{(0)}}{2} \right) \tag{A-12}$$

in which the function $\mathrm{H}\left(\theta_i^{(0)} + \gamma_j^{(0)}\right)$ is given by:

$$\mathrm{H}\left(\theta_i^{(0)} + \gamma_j^{(0)}\right) = \sum_{i=1}^{N_1} \lambda_{ij} \left(\theta_i^{(0)} + \gamma_j^{(0)}\right) + \sum_{j=1}^{N_2} \mu_{ij} \left(\theta_i^{(0)} + \gamma_j^{(0)}\right) \tag{A-13}$$

From the definitions of λ_{ij} and μ_{ij}, we have

$$\sum_{i=1}^{N_1} \lambda_{ij} = 1 \;\; ; \;\; \sum_{j=1}^{N_2} \mu_{ij} = 1 \tag{A-14}$$

Equation (A-12) can be used to simplify the function $\mathrm{H}\left(\theta_i^{(0)} + \gamma_j^{(0)}\right)$ as

$$\mathrm{H}\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)=\sum_{i=1}^{N_1}\lambda_{ij}\theta_i^{(0)}+\gamma_j^{(0)}\sum_{i=1}^{N_1}\lambda_{ij}+\theta_i^{(0)}\sum_{j=1}^{N_2}\mu_{ij}+\sum_{j=1}^{N_2}\mu_{ij}\gamma_j^{(0)}$$
$$\mathrm{H}\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)=\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)+\left(\theta_i^{(1)}+\gamma_j^{(1)}\right) \quad \text{(A-15)}$$

where $\theta_i^{(1)}$ and $\gamma_i^{(1)}$ are defined as follows:

$$\theta_i^{(1)}=\sum_{j=1}^{N_2}\mu_{ij}\gamma_j^{(0)}\ ;\ \gamma_j^{(1)}=\sum_{j=1}^{N_2}\lambda_{ij}\theta_j^{(0)} \quad \text{(A-16)}$$

Substituting Equation (A-14) into Equation (A-15) yields

$$\varepsilon_{ij}^{(2)}=\frac{1}{2}\delta_{ij}\cdot\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)-\frac{1}{2}\delta_{ij}\left\{\frac{1}{2}\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)+\frac{1}{2}\left(\theta_i^{(1)}+\gamma_j^{(1)}\right)\right\}$$
$$\varepsilon_{ij}^{(2)}=\delta_{ij}\cdot\left[\frac{1}{4}\left(\theta_i^{(0)}+\gamma_j^{(0)}\right)+\frac{1}{4}\left(\theta_i^{(1)}+\gamma_j^{(1)}\right)\right] \quad \text{(A-17)}$$

Based on Equation (A-14), we can express $\varepsilon_{ij}^{(m)}$ as

$$\varepsilon_{ij}^{(m)}=\delta_{ij}\left(\varepsilon_{ij}^{(m-1)}-\frac{1}{2}\mathrm{H}\left(\varepsilon_{ij}^{(m-1)}\right)\right) \quad \text{(A-18)}$$

Also, the expression for the function H of Equation (A-15) can be written more generally as:

$$\mathrm{H}\left(\theta_i^{(k)}+\gamma_j^{(k)}\right)=\left(\theta_i^{(k)}+\gamma_j^{(k)}\right)+\left(\theta_i^{(k+1)}+\gamma_j^{(k+1)}\right)$$
$$\text{where } \theta_i^{(k+1)}=\sum_{j=1}^{N_2}\mu_{ij}\gamma_j^{(k)}\ ;\ \gamma_j^{(k+1)}=\sum_{j=1}^{N_2}\lambda_{ij}\theta_j^{(k)} \quad \text{(A-19)}$$

A close examination of expressions for $\varepsilon_{ij}^{(1)}$ and $\varepsilon_{ij}^{(2)}$ from Equations (A-4) and (A-17) indicates that the $\varepsilon_{ij}^{(m)}$ can be written in the form:

$$\varepsilon_{ij}^{(m)} = \sum_{k=1}^{m} \alpha_{m,k} \left(\theta_i^{(k-1)} + \gamma_j^{(k-1)} \right) \tag{A-20}$$

where $\alpha_{m,k}$ are constants. Substituting Equation (A-20) into Equation (A-18) gives us

$$\varepsilon_{ij}^{(m)} = \sum_{k=1}^{m-1} \alpha_{m-1,k} \left(\theta_i^{(k-1)} + \gamma_j^{(k-1)} \right) - \frac{1}{2} \delta_{ij} \cdot \mathrm{H}\left(\sum_{k=1}^{m-1} \alpha_{m-1,k} \left(\theta_i^{(k-1)} + \gamma_j^{(k-1)} \right) \right) \tag{A-21}$$

Using Equation (A-19), Equation (A-21) simplifies as

$$\begin{aligned} \varepsilon_{ij}^{(m)} = \delta_{ij} \Bigg\{ & \frac{1}{2} \alpha_{m-1,1} \left(\theta_i^{(0)} + \gamma_i^{(0)} \right) + \sum_{k=1}^{m-2} \frac{1}{2} \left(\alpha_{m-1,k} - \alpha_{m-1,k-1} \right) \left(\theta_i^{(k)} + \gamma_j^{(k)} \right) \\ & - \frac{1}{2} \alpha_{m-1,m-1} \left(\theta_i^{(m-1)} + \gamma_j^{(m-1)} \right) \Bigg\}, \end{aligned} \tag{A-22}$$

Comparing coefficients of $\left(\theta_i^{(k)} + \gamma_j^{(k)} \right)$ in the expressions for $\varepsilon_{ij}^{(m)}$ between Equation (A-20) and Equation (A-22) can enable us to write expressions for $\alpha_{m,k}$ as

$$\begin{aligned} \alpha_{m,1} &= \frac{1}{2} \alpha_{m-1,1} \\ \alpha_{m,k} &= \frac{1}{2} \left(\alpha_{m-1,k} - \alpha_{m-1,k-1} \right) \text{ for } k = 2,3 \ldots m-1 \\ \alpha_{m,m} &= -\frac{1}{2} \alpha_{m-1,m-1} \end{aligned} \tag{A-23}$$

Equations (A-4) and (A-17) already give us the following coefficients:

$$\begin{aligned} \alpha_{1,1} &= \frac{1}{2} \\ \alpha_{2,1} &= \frac{1}{4} \\ \alpha_{2,2} &= -\frac{1}{4} \end{aligned} \tag{A-24}$$

Using the coefficients in Equation (A-24), one can evaluate the values of all $\alpha_{m,k}$ for all values of *m* and *k*<=*m* using the progression rules given in Equation (A-23). More generally, it can proved that

$$\alpha_{m,k} = (-1)^{k-1}\binom{m}{k}\cdot\frac{k}{m}\cdot\frac{1}{2^m} \quad for\ k = 1,2\ldots m \tag{A-25}$$

which gives us

$$\varepsilon_{ij}^{(m)} = \sum_{k=1}^{m}(-1)^{k-1}\binom{m}{k}\cdot\frac{k}{m}\cdot\frac{1}{2^m}\left(\theta_i^{(k-1)} + \gamma_j^{(k-1)}\right) \tag{A-26}$$

Furthermore, one can prove that

$$\lim_{m\to\infty}\alpha_{m,k} = \lim_{m\to\infty}(-1)^{k-1}\binom{m}{k}\cdot\frac{k}{m}\cdot\frac{1}{2^m} \to 0 \quad for\ k = 1,2\ldots m \tag{A-27}$$

Now, we can write $\theta_i^{(0)} \le \max_{i=1\ldots N_1}\theta_i^{(0)}$ and $\gamma_j^{(0)} \le \max_{j=1\ldots N_2}\gamma_j^{(0)}$ which implies that:

$$\theta_i^{(0)} + \gamma_j^{(0)} \le \max_{i=1\ldots N_1} \theta_i^{(0)} + \max_{j=1\ldots N_2} \gamma_j^{(0)} \tag{A-28}$$

Since $\theta_i^{(1)} + \gamma_i^{(1)} = \sum_{j=1}^{N_2} \mu_{ij}\gamma_j^{(0)} + \sum_{j=1}^{N_2} \lambda_{ij}\theta_j^{(0)}$ and $0 \le \lambda_{ij}, \mu_{ij} \le 1$, we can write

$$\begin{aligned} &\theta_i^{(1)} + \gamma_j^{(1)} \le \max_{i=1\ldots N_1} \theta_i^{(0)} + \max_{j=1\ldots N_2} \gamma_j^{(0)} \\ &\ldots \\ &\ldots \\ &\theta_i^{(k)} + \gamma_i^{(k)} \le \max_{i=1\ldots N_1} \theta_i^{(0)} + \max_{j=1\ldots N_2} \gamma_j^{(0)} < \infty \;\forall\; k = 0,1... \end{aligned} \tag{A-29}$$

This means all $\theta_i^{(k)} + \gamma_i^{(k)}$ are always finite and greater than zero. With Equations (A-27) and (A-29), for a large value of *m*,

$$\lim_{m\to\infty} \varepsilon_{ij}^{(m)} = \lim_{m\to\infty} \sum_{k=1}^{m} \alpha_{m,k} \left(\theta_i^{(k-1)} + \gamma_i^{(k-1)}\right) \to 0 \tag{A-30}$$

Since the increments $\varepsilon_{ij}^{(m)}$ at any $(i, j)^{th}$ grid-point tend to be negligible with increase in iterations, it can be argued that the algorithm will eventually converge to the solution.

To illustrate the convergence of the algorithm, we consider a criterion $E(m)$ based on square-errors between the governing Equations (19) and (20).

$$E(m) = \sum_{j=1}^{N_2} \left(\sum_{i=1}^{N_1} f_{ij} \cdot A_{ij} - \sum_{i=1}^{N_1} A_{ij} \right)^2 + \sum_{i=1}^{N_1} \left(\sum_{j=1}^{N_2} f_{ij} \cdot A_{ij} - \sum_{j=1}^{N_2} A_{ij} \right)^2 \tag{A-31}$$

Figure A-1 plots $E(m)$ vs. *m* for the example in Case 1 of Section 3.1. It can be seen that the convergence is quite rapid and takes less than 10 iterations for even reasonable accuracy.

The convergence analysis so far considers that $f_{ij}^{(m)} \geq 0$ for all values of *m*. However, since there is no restriction on the value of the increments $\varepsilon_{ij}^{(m)}$, it is certainly possible in some cases that the value of $f_{ij}^{(m)}$ could be negative. This can particularly be the case when the random variables are closely-spaced or if the constrained region **R** contributes to a large fraction of the total sample space. However, $f_{ij}^{(m)}$ is the value of a probability density and must always be non-negative as expressed in Equation (17). To account for this, Equation (28) of the algorithm uses a "maximum" function to ensure that the minimum possible value for $f_{ij}^{(m)}$ is zero. Considering a "maximum" condition in the proof of convergence of the algorithm is quite intricate and is not covered in this paper. To show that the algorithm still converges, we take the example in Case 3 of Section 3.1 where the means of the Gaussian random variables are just 2% apart. Figure A-2 plots $E(m)$ vs. *m* for this case. It can be seen for the case of very closely-spaced random variables, $E(m)$ is monotonically decreasing with increasing values of *m*. However, the rate of convergence is slower compared to the example Case 1 of Section 3.1.

Table 1: Application Example Cases

Case #	$f_{\Omega_1}(\omega_1)$		$f_{\Omega_2}(\omega_2)$	
	Distribution Type	Parameters	Distribution Type	Parameters
1	Uniform (Between ±15% of $\mathrm{E}[\Omega_1]$)	$\mathrm{E}[\Omega_1]=10$	Uniform (Between ±15% of $\mathrm{E}[\Omega_2]$)	$\mathrm{E}[\Omega_2]=11$
2	Lognormal	$\mathrm{E}[\Omega_1]=10$ ln*N* (2.2915, 0.1492)	Lognormal	$\mathrm{E}[\Omega_2]=11$ ln*N* (2.3868, 0.1492)
3	Normal	$\mathrm{E}[\Omega_1]=10$ COV = 0.15	Normal	$\mathrm{E}[\Omega_2]=10.2$ COV = 0.15

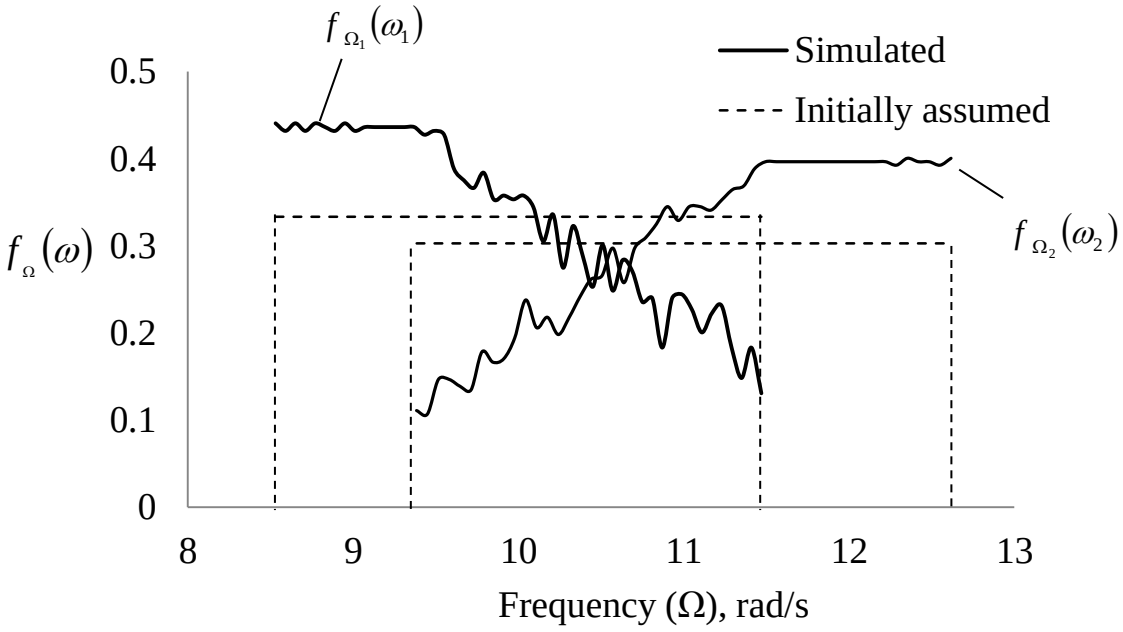


Figure 1: Probability Density Functions of Ω_1 and Ω_2

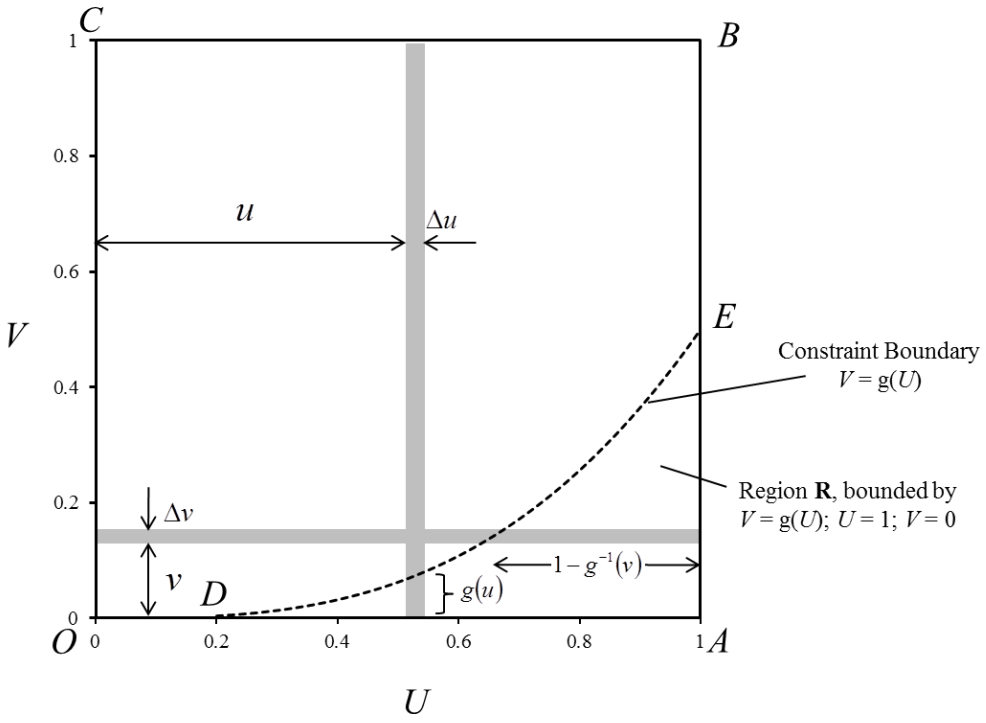


Figure 2: Sample Space for $f_{U,V}(u,v)$

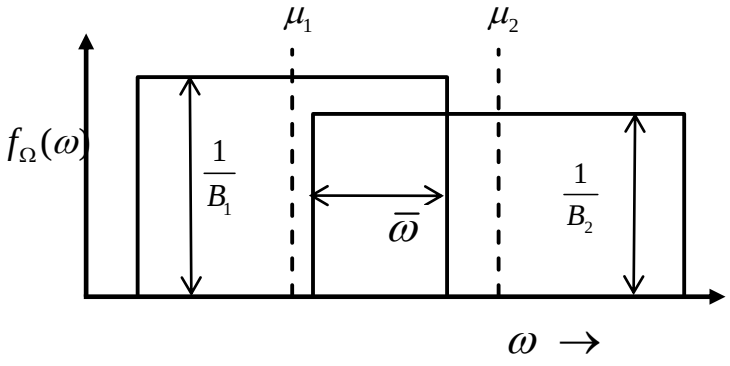


Figure 3: Uniformly- Distributed Probability Density Functions of Ω_1 and Ω_2

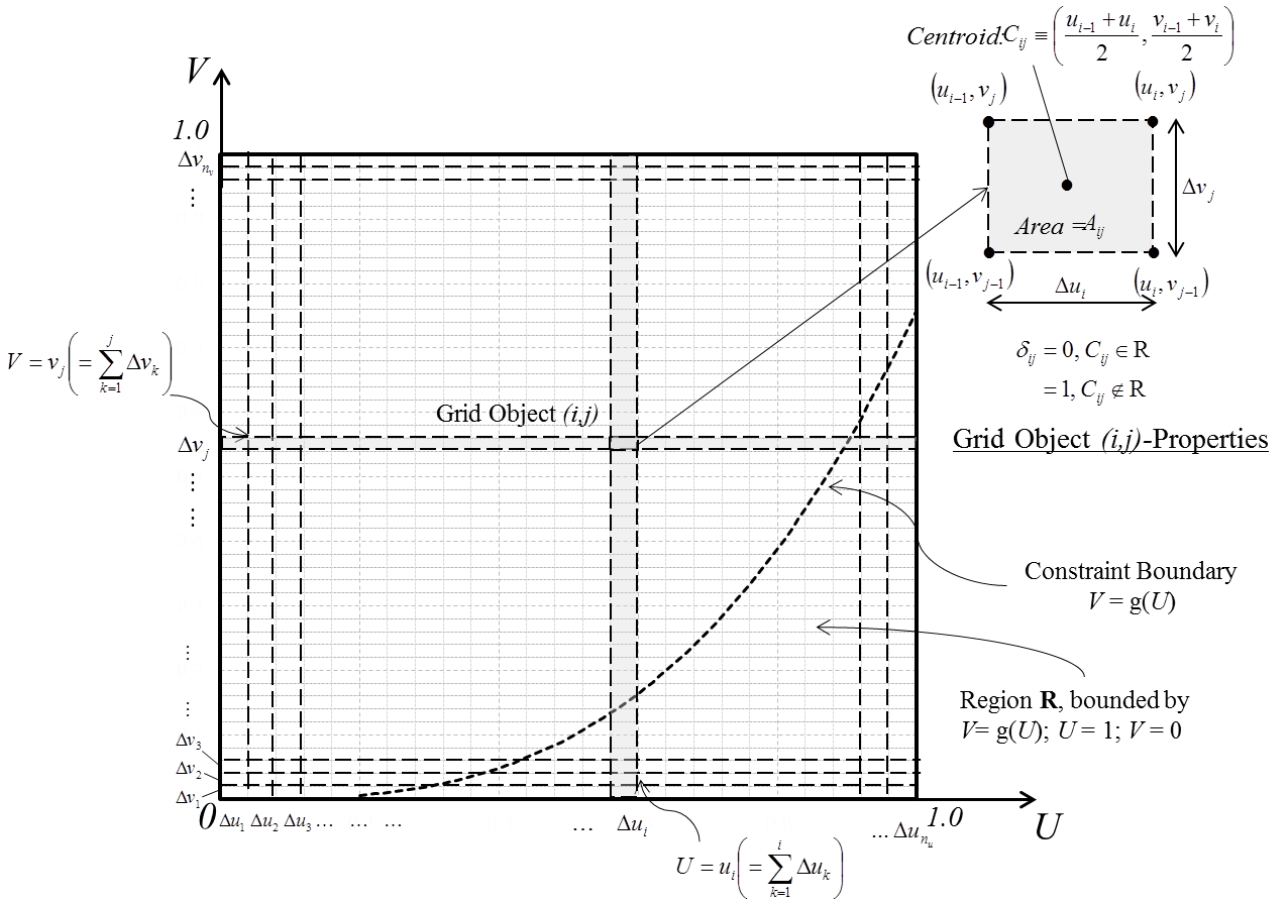


Figure 4: Sample space of *U-V* Discretized as a Grid and the Properties of Constituent Grid Objects

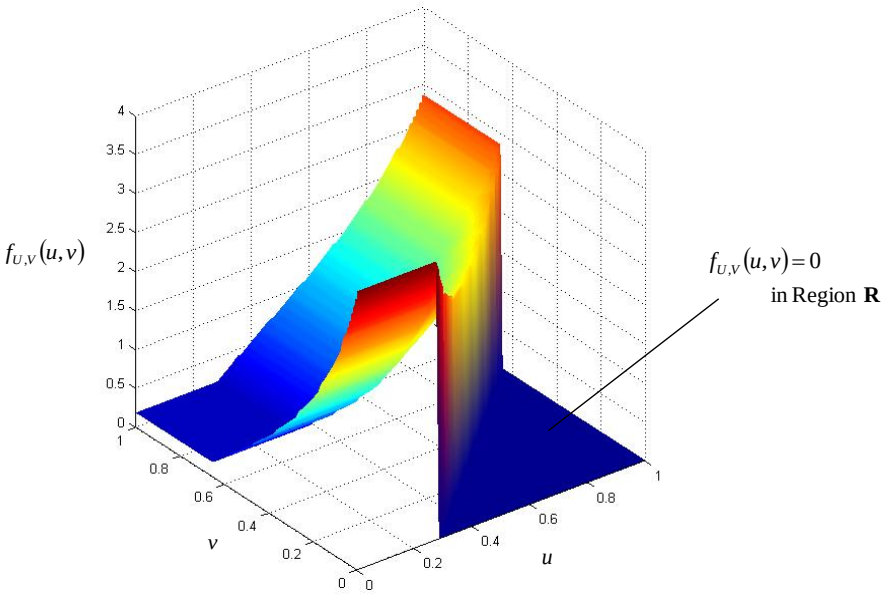


Figure 5(a): Plot of $f_{U,V}(u,v)$ for Example Case 1

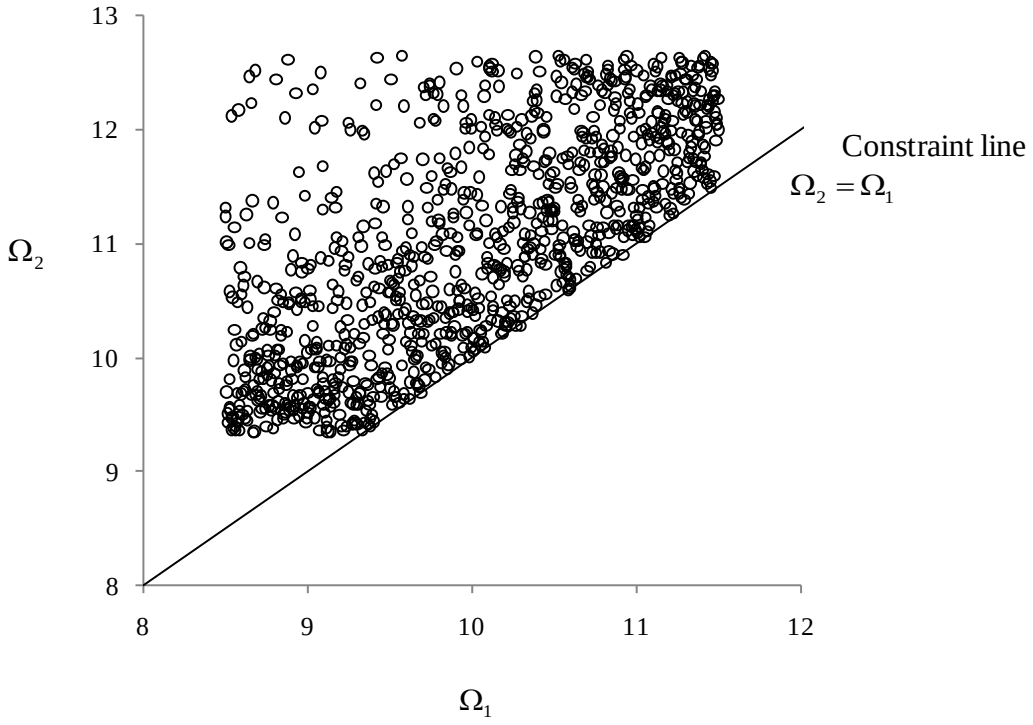


Figure 5(b): Scatter plot of 1000 Samples (ω_1, ω_2) as Simulated (Case 1)

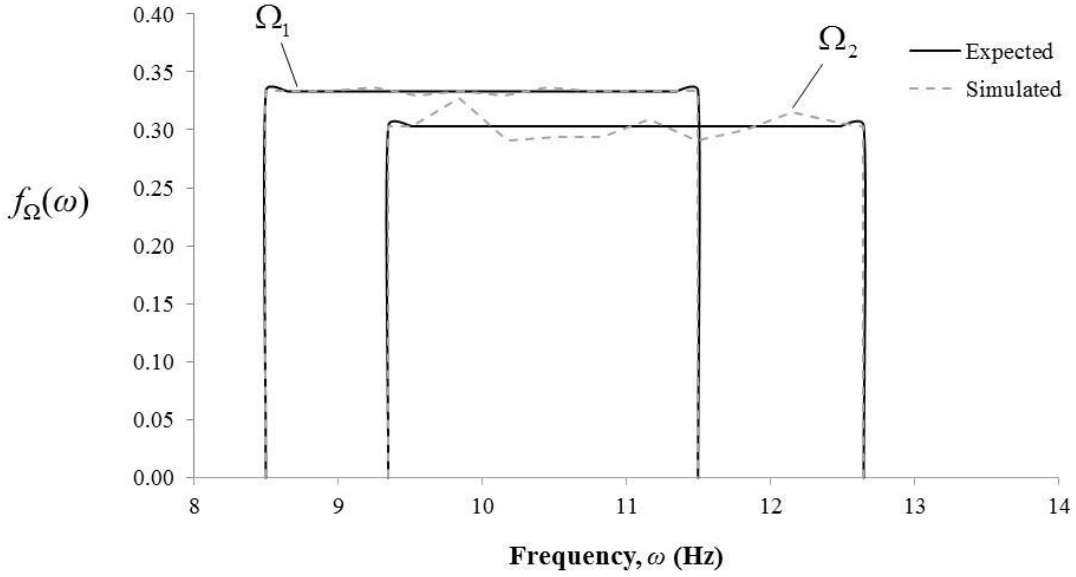


Figure 5(c): Comparison of Sampled Vs. Expected PDFs (Case 1)

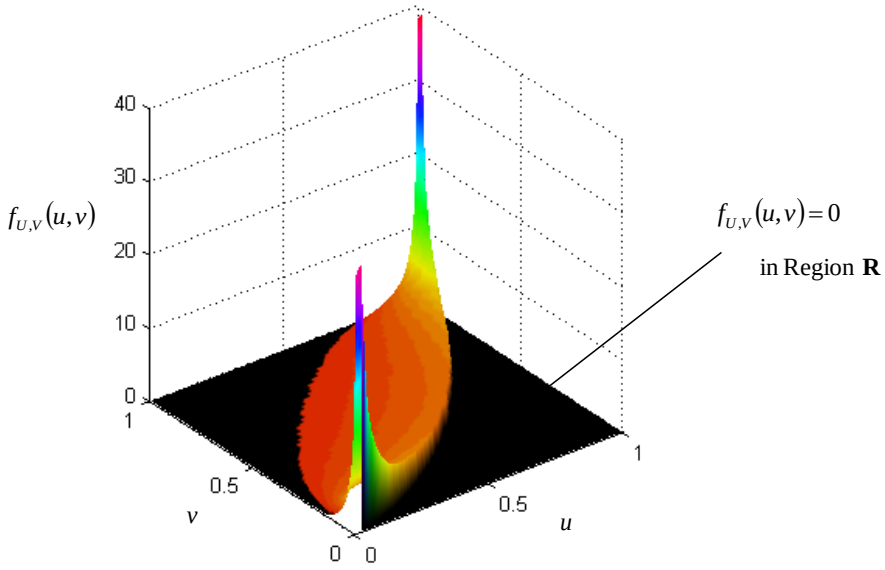


Figure 6(a): Plot of $f_{U,V}(u,v)$ for Example Case 2

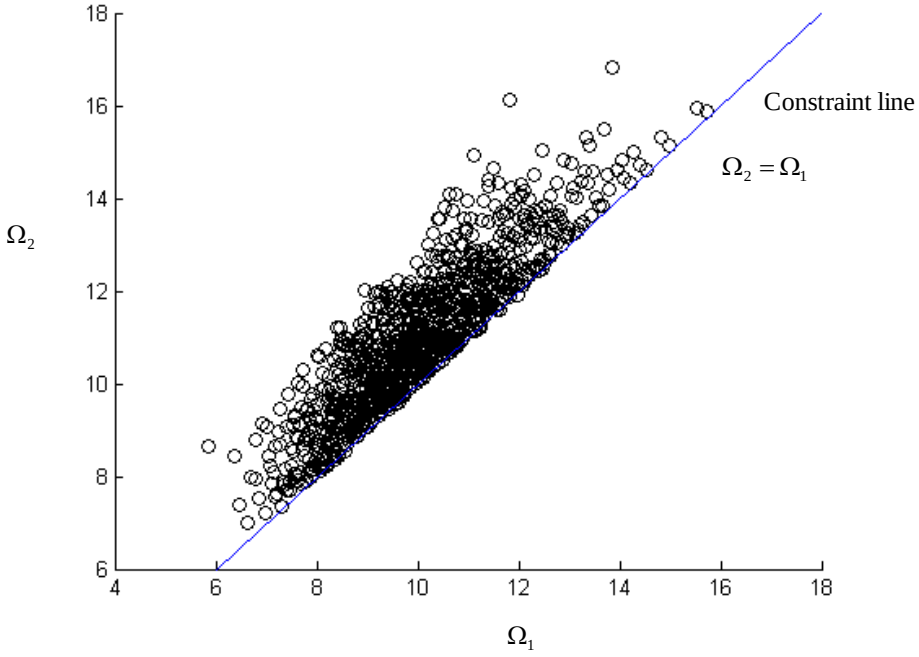


Figure 6(b): Scatter Plot of 1000 Samples (ω_1, ω_2) as Simulated (Example Case 2)

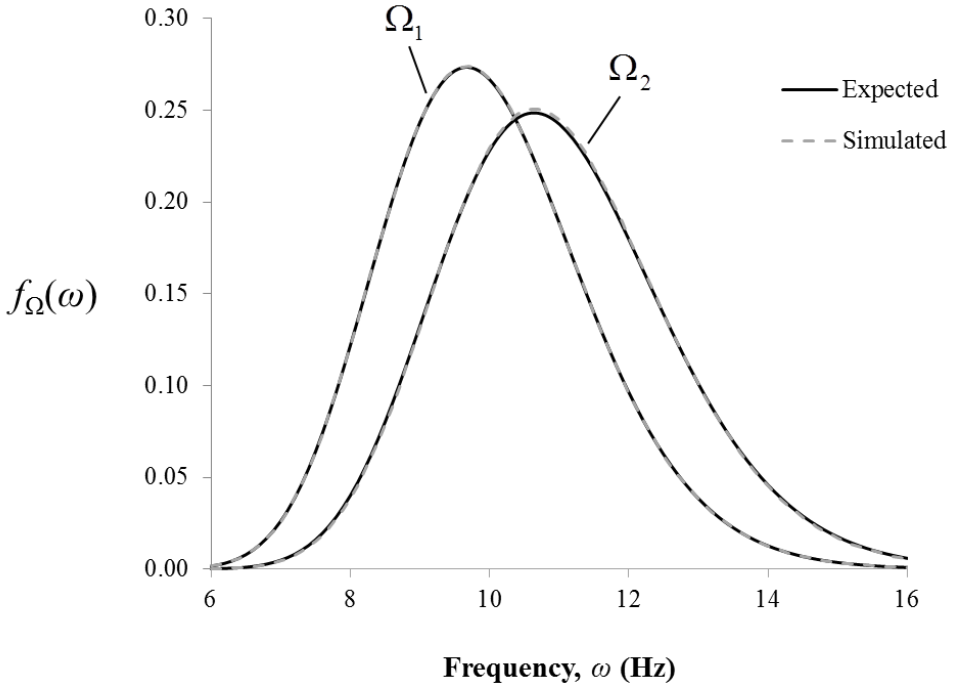


Figure 6(c): Comparison of Sampled Vs. Expected PDFs (Case 2)

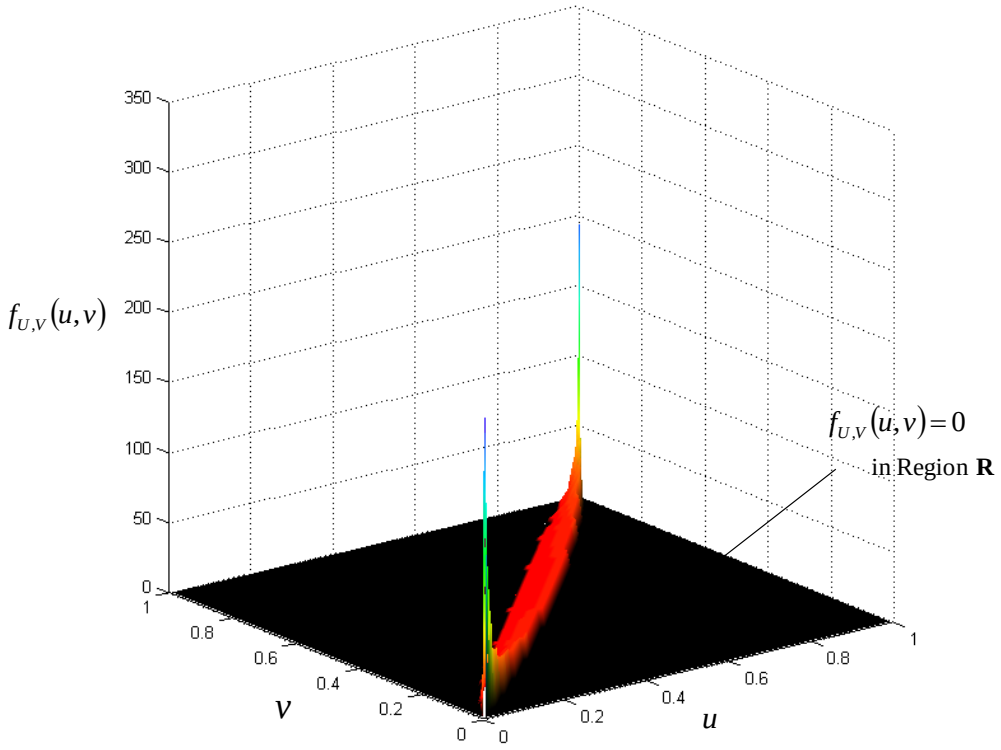


Figure 7(a): Plot of $f_{U,V}(u,v)$ for Example Case 3

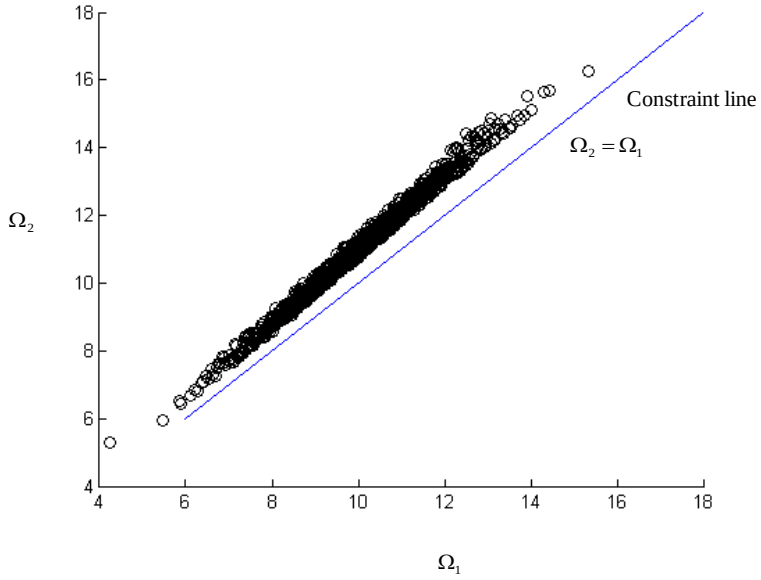


Figure 7(b): Scatter Plot of 1000 Samples (ω_1, ω_2) as Simulated (Example Case 3)

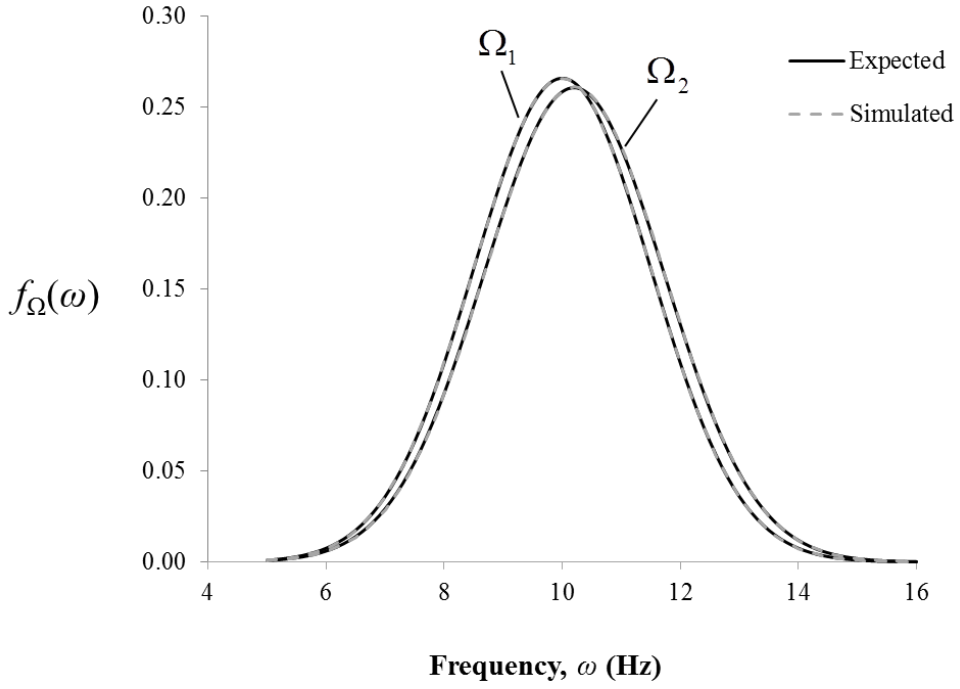


Figure 7(c): Comparison of Sampled Vs. Expected PDFs (Case 3)

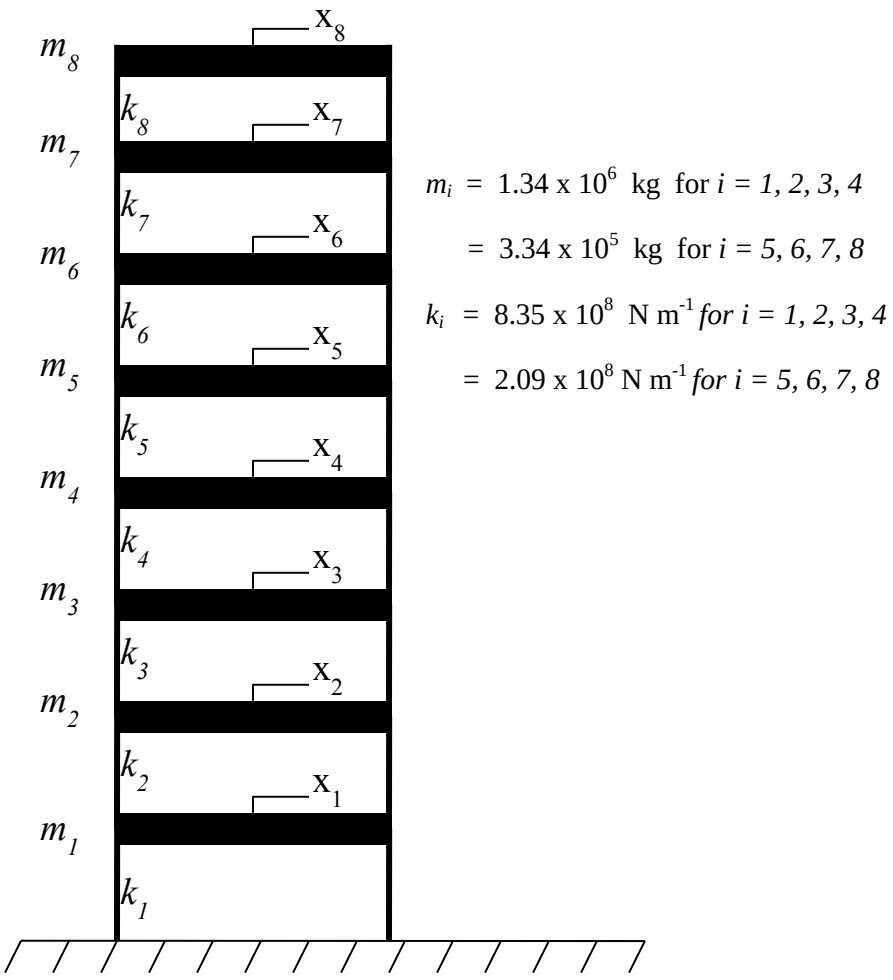


Figure 8: An 8 Storey Structural System

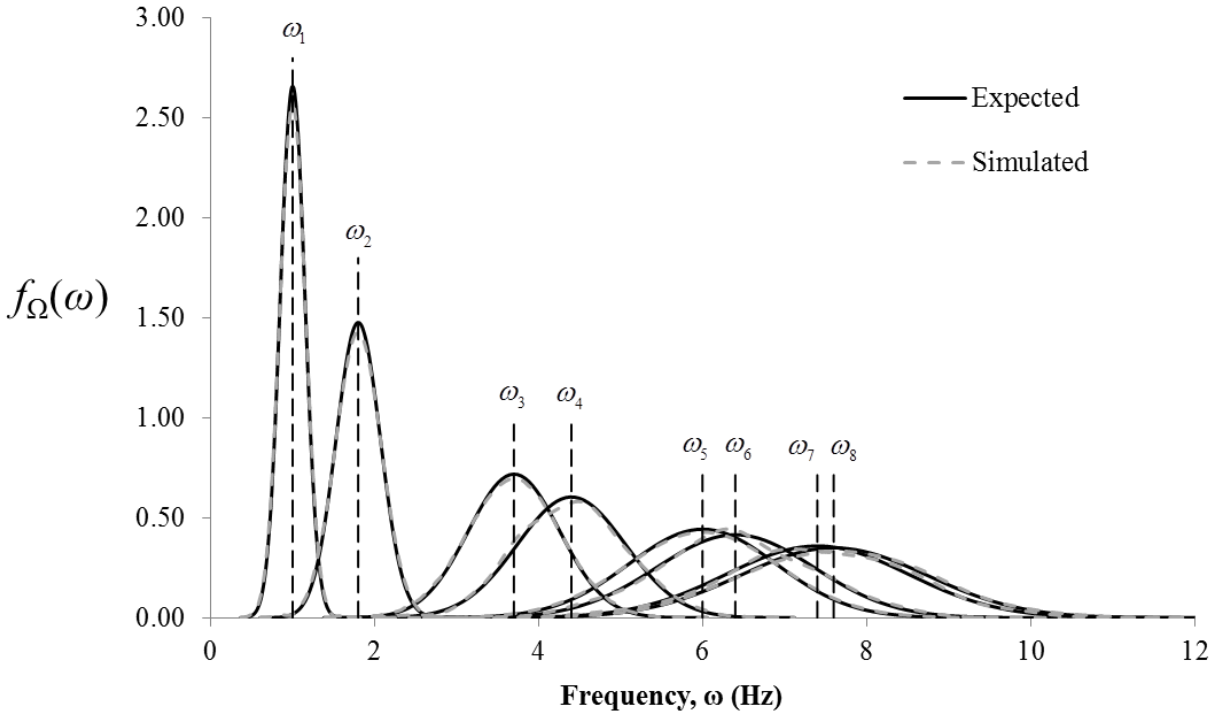


Figure 9: Comparison of Simulated and Expected PDFs of Natural Frequencies of the 8-DOF Structure

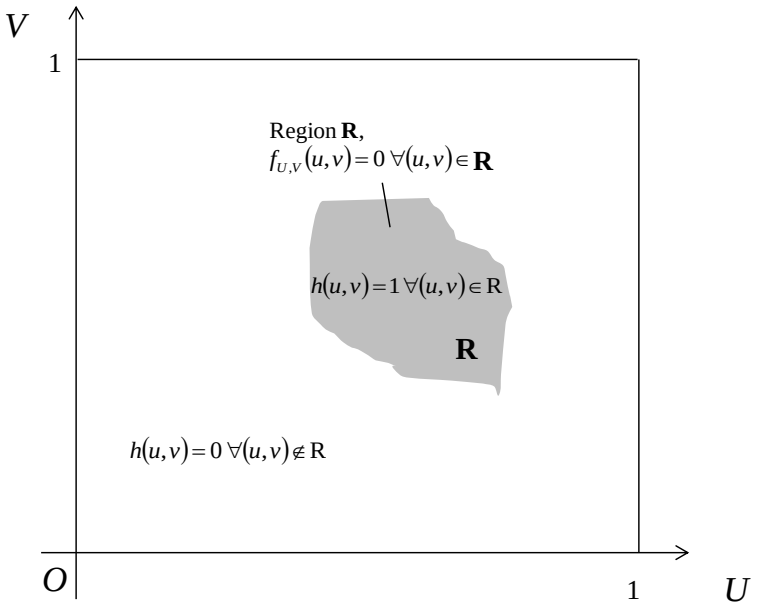


Figure 10: Application to Other Constraints

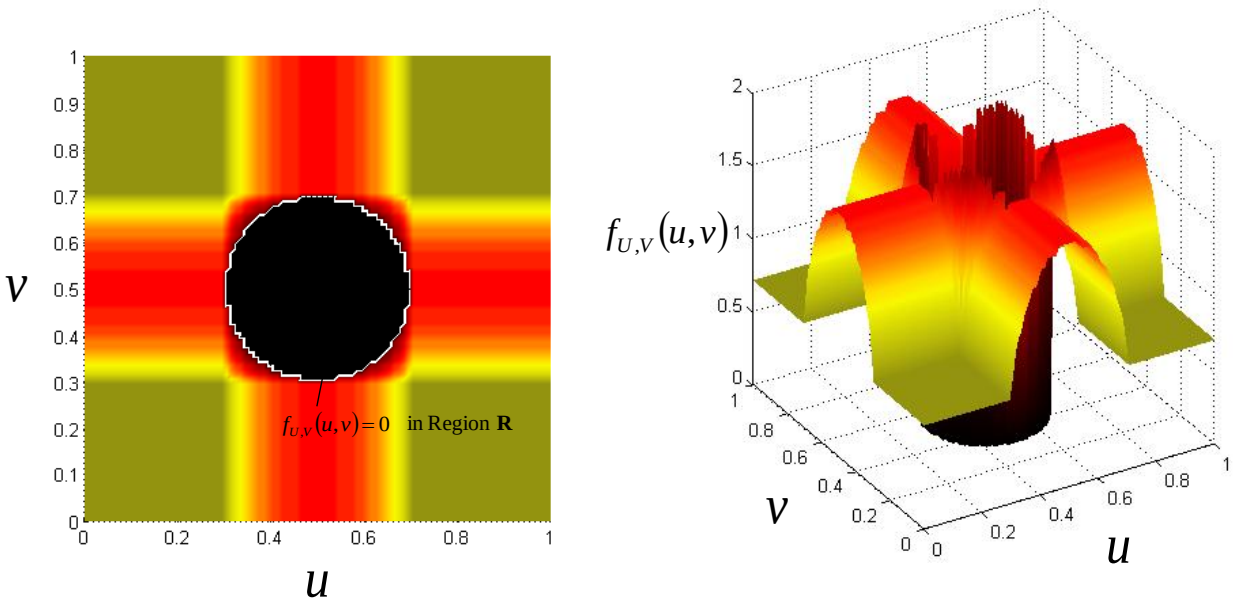


Figure 11: $f_{U,V}(u,v)$ if Region **R** is a circle

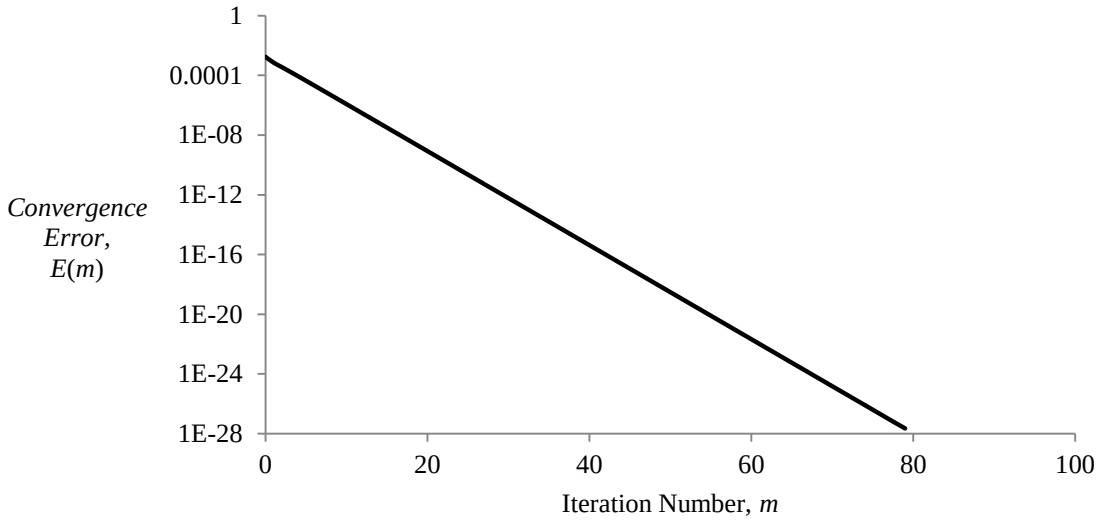


Figure A-1: Convergence of the Algorithm in Example Case 1

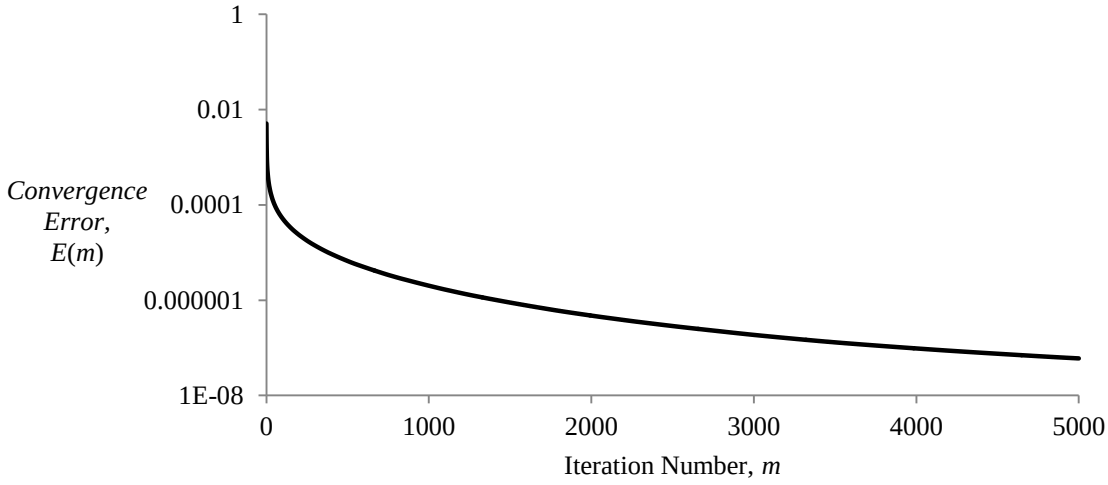


Figure A-2: Convergence of the Algorithm in Example Case 3

PART V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

FOR FUTURE RESEARCH

1. SUMMARY AND CONCLUSIONS

This dissertation presents a comprehensive statistical framework for evaluating system-level structural fragilities as well as updating them based on the availability of additional fragility data. This thesis is organized primarily into three manuscripts that focus on describing the following key aspects of the proposed framework:

- Estimation of the structural fragility estimates to desired limits of accuracy with a high degree of confidence.
- Reduction in the total number of nonlinear time history analyses required for a specified degree of accuracy in the fragility curve.
- Robustness to handle uncertainties in complex structures where interdependence among various physical variables must be accounted for.

The summary and conclusions of each topic are given below.

1.1 A Bayesian framework for estimating and updating seismic fragility of structures

In the USNRC's Probabilistic Seismic Risk Assessment (PSRA) framework for nuclear power plants, the risk of damage in a structural system is described by a "fragility curve". A fragility curve gives the probability that the structure could exceed a specified limit-state at a particular level of earthquake loading intensity like Peak Ground Acceleration (PGA). For real-life systems with a large number of degrees of freedom, using conventional reliability methods to estimate the risk may result in large errors due to non-standard design

characteristics. In such cases, one realistic way of estimating fragility of the system is to use an experimentally-validated Finite-Element (FE) model and perform multiple dynamic analyses by considering uncertainties in material, modeling and loading variables. But, this Monte-Carlo simulation based approach to risk estimation has one important drawback - it requires a large number of simulations to be performed for accurate estimates of risk. This can be a drawback in the case of large FE models which often take significant computational time even for a single analysis. Therefore, a trade-off is often required between the desired accuracy of the risk estimates and the total number of analyses that can be performed in a reasonable time.

To address this issue, a statistical framework based on Bayesian inference was developed to minimize the total computational effort required for determining fragility functions in real-life structural systems with a specified degree of accuracy. The proposed Bayesian framework has three important features:

1. The framework provides a consistent way to incorporate different kinds of seismic risk/fragility data and dynamically update the structural fragility of engineering structures as and when data becomes available.

2. The framework enables quick convergence by reducing the confidence intervals even with limited amount of data. Not only that, the framework can be used to continually monitor the current accuracy of fragility estimates and also forecast how much more data might be required in order to achieve desired accuracy in seismic fragility curves.

3. The framework combines optimization techniques with the ability of the Bayesian fragility model to forecast the required amount of data for desired accuracy to form an efficient planning tool that is capable of minimizing the total computational cost in collecting the data itself.

This framework was illustrated using the example of a box-shaped reinforced concrete structural shear wall in a nuclear reactor. A comprehensive experimentally-validated nonlinear FE model for the system created in ABAQUS takes about several hours of computation time for a single dynamic analysis. It is demonstrated that the Bayesian framework and planning methodology is instrumental in reducing the total number of nonlinear analyses by nearly 40% for updating fragility of the RC box shear-wall. It must be noted that the suggested methodology can be extended to many different planning or resource allocation problems in engineering risk analyses where the objective is to minimize the cost of simulations or data collection as long as the convergence or acceptance criteria is specified.

1.2 Seismic fragility of large-scale piping systems using equivalent elastic time-history analyses

A key part of the Bayesian framework summarized in Section 1.1 is the availability of the prior information or some simplified approach to obtain preliminary estimates of seismic risk. Modeling and generalizing the behavior of nonstructural systems such as piping is not straightforward due to factors like mass interaction with the primary system, non-classical

damping properties, significant higher mode effects etc. Additionally, numerous experiments indicate that although large piping systems behave essentially elastic, their seismic fragility is dominated by presence of localized nonlinearity at locations of joints and fixtures. Since the overall structure behaves predominantly elastic and that nonlinearity is restricted to small regions, equivalent linearization of the nonlinear components in the structure seems like an attractive possibility to estimate the fragility of the system using only elastic time-history simulations. Extending any of the existing equivalent linearization techniques to large MDOF systems is computationally intensive and in most cases not a straight-forward implementation. Moreover, most equivalent linearization techniques aim to minimize the error in mean response between the systems whereas in seismic fragility analyses, it is the peak responses that determine the risk. To address this issue, a concept called equivalent elastic limit state is introduced which proposes to disregard the localized nonlinearity and instead modify the failure capacity (limit-state) of the component such that the damage probabilities evaluate to be close to that of the actual nonlinear system. By studying a large number of representative nonstructural systems which are essentially linear but characterized by localized fragile nonlinearities, a comprehensive model is developed to predict the value of the modified failure capacity or "equivalent elastic limit state" of the linearized member.

By studying a large number of representative nonstructural systems which are essentially linear but characterized by localized nonlinearities, we proposed a model that uses elastic analysis results from an equivalent linearized system to obtain "prior" estimates which

can be input into the Bayesian framework. The efficacy of this procedure was demonstrated by considering the real-life floor piping model with fragile 2" threaded T-joint connections. It is demonstrated the methodology based on prior estimates using equivalent elastic analyses in conjunction with Bayesian updating by incorporating nonlinear simulation data is an efficient approach to minimize the total computational effort needed in evaluating seismic fragility of critical nonlinear connections in large piping systems.

1.3 Simulation of constrained, closely-spaced random variables in engineering risk analyses

In evaluating fragility using probabilistic simulations, each dynamic analysis is instantiated by a unique set of randomly-sampled input values for all the material, modeling as well as loading variables. Therefore, modeling of input variables is an important step in setting up a Monte-Carlo simulation. For instance, one or more physical variables of the model could be inter-related or correlated with each other and it is imperative to address these constraints while sampling for Monte-Carlo methods. The motivation for this research was focused on sampling sets of natural frequencies of a structure with large degrees of freedom. In nuclear power plant structural systems, the U.S. Nuclear Regulatory Commission (USNRC) and ASME recommend a ±15% uncertainty in building or primary system natural frequencies that is uniformly distributed over the entire range. When the individual PDFs of closely-spaced natural frequencies overlap, generating sample sets by assuming them to be independent random variables can lead to incorrect sets of sampled frequencies in the sense

that the frequencies do not remain ordered. However, any set of natural frequencies of an uncoupled primary or secondary system should all be in an ascending order. Rejecting the disordered samples result in individual density functions that are significantly different from the distributions initially assumed for sampling of each frequency. Thus, any sampling scheme must take into account the above constraints. Thus, the problem of sampling closely-spaced random variables subject to constraints must be taken into account while setting up any Monte-Carlo simulation as ignoring the interdependencies between constrained input parameters can lead to simulation of unrealistic structural systems. A numerical iterative procedure based on fundamental probability approaches is suggested in order to sample closely-spaced random variables subject to any kind of non-linear constraints. The numerical procedure is versatile and can be used for sampling random variables described by most probability distributions and typical inequality constraints.

2. RECOMMENDATIONS FOR FUTURE RESEARCH

The research presented in this dissertation is a first step to develop solutions and tools to various issues that arise in the modeling, planning, set-up and execution of probabilistic simulations in the context of evaluating seismic risk on engineering structures. Based on the experience gained in conducting this research, a few recommendations for future work in this area are summarized below.

- **Epistemic Uncertainties:** The Bayesian framework can be more robust by including the capability to incorporate epistemic uncertainties i.e. explicit consideration of reliability / confidence level of the new data source along with the fragility data being used to update the current estimates of fragility.
- **Uncertainty Propagation:** A key factor in determining the total computational effort in conducting probabilistic seismic simulations is the nature of uncertainties in various parameters of Finite Element (FE) or engineering analysis model as well as the randomness in the loading on the system. It is important for us to recognize as far as possible the nature of uncertainty present in every parameter of the model and also the effect of that uncertainty on the overall reliability of the design. Moreover, trying to account for uncertainty in every parameter of the model can be computationally limiting as Monte-Carlo methods generally require intensive simulations for convergence. Typically, the properties of the random variables used to describe various uncertain parameters are obtained from existing real-life data, past studies etc. and sometimes insufficient or inconclusive data can also be constraint in the modeling uncertainties. In order to minimize the computational complexity, understanding the relative significance of uncertainties in various parameters within the same model can help us identify the most influential random variables and subsequently, simplify the probabilistic model.
- **Model Selection:** Each estimate of risk using computational models depends on two important factors: (1) the robustness of FE models describing the physical behavior of

the structure (example: simple shear beam model or a dense FE mesh made up of 3D shell elements); (2) amount of experimental / real-life data that was used to validate and test the FE model. If the FE model is too robust with many features, it could be computationally intensive for performing a large number of simulations. If it's too simplified, it could lead to high bias although it could demand much lesser computing resources. Developing rigorous statistical criteria in order to identify the right level of detail described in the probabilistic FE model corresponding to the specific risk criteria of interest can be very instrumental in minimizing the total computational effort while maintaining the adequate accuracy in seismic fragility predictions.

- **Validation of Equivalent Elastic Limit State Concept:** Additional verification and validation is needed to check the robustness and applicability of the equivalent elastic limit state concept with fragility analyses for different kinds of representative real-life systems with localized fragile nonlinearities.