Reliability-Based Topology Optimization with Analytic Sensitivities

Patrick Ryan Clark

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
in
Aerospace Engineering

Mayuresh J. Patil, Chair
Robert A. Canfield
Rakesh K. Kapania

June 19, 2017
Blacksburg, Virginia

Keywords: Reliability-Based Topology Optimization, First-Order Reliability Method, Sensitivity Analysis

Reliability-Based Topology Optimization with Analytic Sensitivities

Patrick Ryan Clark

ABSTRACT (Academic)

Reliability-Based Design Optimization (RBDO) approaches often use the First-Order Reliability Method (FORM) to efficiently obtain an estimate of the reliability of a system. This approach treats the reliability analysis as a nested optimization problem, where the objective is to compute the most probable point (MPP) by minimizing the distance between the failure surface and the origin of a normalized random space. Numeric gradient calculation of the solution of this nested problem requires an additional solution of the FORM problem for each design variable, an approach which quickly becomes computationally intractable for large scale problems including Reliability-Based Topology Optimization (RBTO). In this thesis, an alternative analytic approach to the analysis and sensitivity of nested optima derived from the Lagrange Multiplier Theorem is explored. This approach leads to a system of nonlinear equations for the MPP analysis for any given set of design variables. Taking the derivative of these equations with respect to a design variable gives a linear system of equations in terms of the implicit sensitivities of the MPP to the design variable where the coefficients of the linear equations depend only on the current MPP. By solving this system, these sensitivities can be obtained without requiring addition solutions of the FORM problem. The proposed approach is demonstrated through several RBDO and RBTO problems.

Reliability-Based Topology Optimization with Analytic Sensitivities

Patrick Ryan Clark

ABSTRACT (General Audience)

It is a common practice when designing a system to apply safety factors to the critical failure load or event. These safety factors provide a buffer against failure due to the random or un-modeled behavior, which may lead the system to exceed these limits. However these safety factors are not directly related to the likelihood of a failure event occurring. If the safety factors are poorly chosen, the system may fail unexpectedly or it may have a design which is too conservative. Reliability-Based Design Optimization (RBDO) is an alternative approach which directly considers the likelihood of failure by incorporating a reliability analysis step such as the First-Order Reliability Method (FORM). The FORM analysis requires the solution of an optimization problem however, so implementing this approach into an RBDO routine creates a double-loop optimization structure. For large problems such as Reliability-Based Topology Optimization (RBTO), numeric sensitivity analysis becomes computationally intractable. In this thesis, a general approach to the sensitivity analysis of nested functions is developed from the Lagrange Multiplier Theorem and then applied to several Reliability-Based Design Optimization problems, including topology optimization. The proposed approach is computationally efficient, requiring only a single solution of the FORM problem each iteration.

I dedicate this work to Heather.

You have always been there for me and I will always be there for you.

Acknowledgements

I would like to acknowledge the contributions from my advisor, Dr. Mayuresh Patil. This thesis would not be possible without his advice and guidance. More importantly, through my regular discussions with Dr. Patil I learned the power of asking questions and the joy of seeking a deeper understanding.

I would also like to recognize the influences each member of my committee had on my academic career. I had the pleasure of studying under Dr. Robert Canfield multiple times throughout my undergraduate and graduate career, and it was his courses in design optimization and reliability-based structural design which first drew me to this field. Similarly, studying under Dr. Rakesh Kapania has given me an appreciation of the beauty of structural analysis.

Lastly, I would like to give special thanks to the Virginia Tech Center for the Enhancement of Engineering Diversity for supporting my graduate studies. Participating in your programs as an undergraduate was a life-changing experience, so I am grateful for the opportunity to give back. It truly has been a pleasure to serve on your team.

Table of Contents

List of Figures

List of Tables

Nomenclature

A	=	Linear density filter weighting matrix
C	=	Compliance
D	=	Finite element displacement vector
E	=	Elastic modulus
f	=	Objective function, nested objective function
f_X	=	Generic probability density function
F	=	Point force
F_X	=	Generic cumulative density function
FS	=	Factor of safety
g	=	Generic function, nested inequality constraint
G	=	Generic inequality constraint
h	=	Finite difference step size, nested equality constraint
H	=	Generic equality constraint
$H()$	=	Hessian matrix
i, j, k	=	Indices for vectors and matrices
I	=	Identity matrix
K	=	Finite element stiffness matrix
L	=	Lagrangian function
L_i	=	Locator matrix
p	=	SIMP penalization exponent
P_f	=	Probability of failure
$\bar{r}$	=	Filter radius
R	=	Finite element load vector
U	=	Standard normal random variable
U^*	=	Most probable point
v	=	Lagrange multiplier
w	=	Weighting factor
x	=	Generic design variable

x^*	=	Generic local minima
X	=	Generic random parameter
y	=	Generic nested variable
y^*	=	Generic nested minima
$\bar{z}$	=	Generic nested state variable
z^T	=	Vector which isolates a specified displacement
β	=	Reliability index
$\bar{\beta}$	=	Target reliability index
λ	=	Lagrange multiplier, adjoint vector
μ	=	Mean
ν	=	Poisson's ratio
ρ	=	Pseudo-density
σ	=	Standard deviation, stress
ϕ	=	Standard normal probability density function
Φ	=	Standard normal cumulative density function

1. Introduction

1.1 Background

In design optimization, a system or structure is reduced to a set of equations in terms of key but unknown characteristics, such as the width of a beam, the stiffness of a spring, or the resistance of an electrical component. After identifying performance metrics to be maximized or minimized as well as constraints on the design, a numerical procedure is then used to find the best feasible solution. Traditionally, this requires the assumption that all aspects of the problem are deterministic. That is, the behavior of every parameter or characteristic is known exactly and does not vary regardless of how many times the system is produced or observed. In reality, there are few systems which behave in this manner. Rather, most real systems have some degree of randomness in their configuration or response. For example, the material properties of a steel beam can vary locally due to changes in the microstructure, or vary on average between different production runs due to inexact control of the alloying agents. Similarly, the geometry of a machined part is commonly expected to fall within a range of values specified by a set of tolerances as a result of imperfect process control.

The common practice for addressing these uncertainties while retaining the deterministic analysis is to apply a factor of safety to the limit states of the system or make conservative assumptions. The result is a buffer against failure which covers all reasonably likely scenarios when chosen correctly. The question facing the engineer is then a question of what factor of safety is appropriate? While the factor of safety approach is simple to implement, it is only an indirect measure of the actual probability of failure of the system. Choosing too low of a factor of safety can result in a design which fails more frequently than desired while choosing too high results in a design which is inefficient. Typically, experience and extensive testing are required to determine appropriate safety factors, however this can be costly and time consuming.

Furthermore, even with the correct choice of factors of safety, the optimal design based on these assumptions is not optimal in terms of obtaining the best reliable solution.

An alternative approach is to formulate the optimization process in terms of the probabilities of failure, which has led to the fields of Reliability-Based Design Optimization (RBDO) and Robust Design Optimization (RDO). These methods directly consider the statistics of the system, with RBDO introducing probability as a constraint and RDO using statistical measures such as variance as a target for minimization or maximization. These approaches are appealing since the uncertainties present in the system can be directly quantified; however this is also the challenge of these methods: how can existing optimization tools be modified to consider parameters which no longer hold a fixed value, but instead may hold one of infinitely many values? Historically, a number of approaches have been used to perform this probabilistic analysis. These include simulation methods such as Monte Carlo simulation as well as approximation methods, including the First-Order Reliability Method. These methods typically require greater computational effort than deterministic analysis since each random parameter adds another dimension to the problem. The First-Order Reliability Method is particularly noteworthy, since this approach reformulates the probabilistic analysis into a deterministic optimization problem. However, when implemented into an RBDO or RDO problem the result is a nested structure where the FORM optimization problem has to be solved at every iteration of the design optimization problem. While this is computationally expensive by itself, the problem is compounded during the sensitivity analysis step, particularly if finite difference derivatives are used. In this case, the FORM problem needs to be solved once each iteration to get a baseline then re-solved for each design variable. Thus, the computational cost rapidly grows with the number of design variables.

Unfortunately, this limits the use of RBDO techniques for topology optimization, since these problems typically have thousands or more design variables. If an efficient sensitivity analysis procedure can be derived, then Reliability-Based Topology Optimization (RBTO) has the potential to combine the best of both of these methods. By itself, topology optimization is a powerful tool since the distribution of the material or another property is treated as unknown, thereby removing as many assumptions as possible from the analysis. This expands the design space, potentially leading to novel, efficient designs. If an efficient sensitivity analysis procedure can be found, then these designs can be constrained to satisfy a target reliability level, potentially leading to novel designs which are both efficient and reliable.

1.2 Overview of Research

The aim of this research is to develop a general procedure for calculating the gradients of the solution of a nested optimization problem to the design variables of the outer problem. Ideally, such a procedure would not require any further information than what is already computed as a part of the initial solution to the nested optimization problem, particularly if the nested optimum requires the solution of a finite element or other large-scale problem. The procedure should not require any additional solutions of the nested optimization problem either. After developing a generalized procedure, its implementation into a RBDO problem using the First-Order Reliability Method can be explored, including the efficient solution of Reliability-Based Topology Optimization problems.

The remainder of this thesis is as follows. First, a survey of existing research in the field of Reliability-Based Topology Optimization is presented. Particular focus is given to the first-order reliability method, the solid-isotropic material with penalization formulation for topology optimization, and existing approaches to the sensitivity analysis used to solve RBTO problems. An approach for the sensitivity analysis derived from the Lagrange Multiplier Theorem is then presented. Several example problems are then formulated and solved. These include a probabilistic variant of the classic 3-bar truss problem [2], the RBTO benchmark derived by Rozvany and Maute [1], and a probabilistic nodal displacement problem representative of topology optimization scenarios encountered in industry.

2. Literature Review

A common assumption made when designing or analyzing a system is that the configuration and behavior of the system can be expressed exactly by a set of defined parameters. This implies that there is perfect information about the system and that it will perform identically regardless of how many times it is produced or used. This assumption is at the core of most engineering analyses, ranging from simple hand calculations to large finite element problems. Few, if any, real systems are entirely deterministic however. Instead, there is usually some degree of randomness inherent in the system. This randomness is potentially derived from many different sources, which can include manufacturing imperfections, varying environmental conditions, or complex and chaotic interactions with other systems. The result is that there is a degree of uncertainty associated with any analysis using this assumption. When derived from random sources, the uncertainty is referred to as being aleatory and is typically irreducible. This differs from epistemic uncertainty, which is uncertainty which can be reduced by having better information.

2.1 Deterministic Design Using Factors of Safety

The simplest procedure for addressing uncertainties is to apply a factor of safety to the deterministic limiting conditions during the design process. The factor of safety is simply the ratio of the actual failure state to a more restrictive design failure state:

$$FS = \frac{g_{failure}}{g_{design}} \tag{1}$$

The result is a buffer between the design criteria and the actual limit state. Ideally, this buffer is sufficiently large enough to include all reasonably likely realizations of the system. This buffer also accounts for uncertainty which may be present in the actual limit state, such as how the yield

stress of a material is commonly reported as a statistical quantity with the expectation that samples may deviate by some amount. It is important to note that the factor of safety is only indirectly related to the random behavior of the system. While a factor of safety of 1.25 may be sufficient for one application, there is no guarantee that it will be sufficient for another. As an example of this, consider the axial bar shown below in Figure 1.
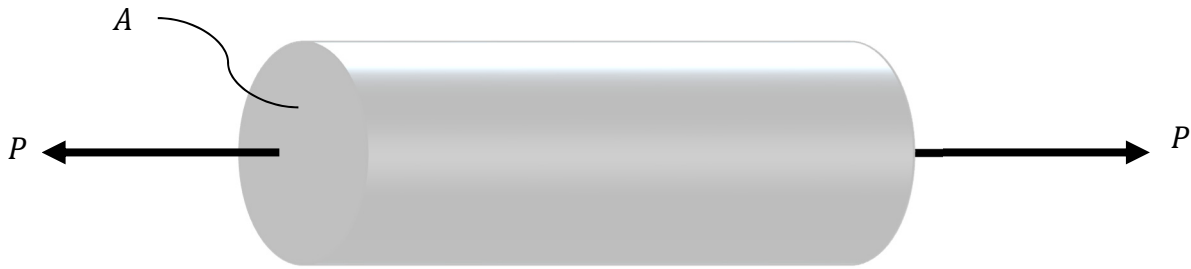
Figure 1: Axial bar with cross-sectional area A and applied load P

The stress in this axial bar is commonly known to be the ratio of the applied load P to the cross-sectional area of the bar A:

$$\sigma = \frac{P}{A} \tag{2}$$

A constraint on the stress can then be introduced to size the area of the cross-section. For example, it is common to require that the stress is less than a limiting value such as the yield stress or the critical buckling stress:

$$\sigma_{Design} = \frac{P}{A} \leq \sigma_{max} \tag{3}$$

A factor of safety can then be applied to the design to account for uncertainties in the applied load, limit stress, and the geometry of the bar:

$$FS \times \frac{P}{A} = \sigma_{max} \tag{4}$$

Equation (4) can then be used to choose an appropriate cross-sectional area given an assumed value for the load. In operation, the bar will not fail as long as the actual load falls within the buffer created by the factor of safety:

$$P_{actual} \leq FS \times P_{design} \tag{5}$$

It is important to note that while Equation (5) states a failure condition for the actual load, it does not provide any information about the reliability of the system, defined as the frequency with which the actual load will exceed this limit. If the actual load varies significantly from the design load, then there is the potential that failure will occur far more frequently than desired, even with the factor of safety. Similarly, the factor of safety may be providing too much of a buffer if the actual load does not frequently deviate from the design load, indicating that the cross-section area could be reduced further while satisfying the reliability requirements. Thus, appropriate choice of a factor of safety is essential. For a simple design problem with few sources of randomness it may be easy to select a factor of safety which balances the frequency of failure with the optimality of the design, however for complex systems the random behavior may be difficult to estimate without extensive testing.

2.2 Reliability Analysis

While deterministic analysis seeks to answer the questions of if or when failure will occur, reliability analysis instead analyzes the frequency with which failure occurs. This requires an understanding of the probabilistic nature of the system, including which parameters behave randomly and how this behavior can be modeled. By using probability principles, reliability analysis approaches are able to express the random behavior as a single quantity which is compatible with existing deterministic design tools. In some cases this can be achieved by reformulating the problem to include the probability distributions, however this approach is problem specific. More generally, the reliability analysis can be performed using either simulation approaches such as Monte Carlo Simulation or approximation approaches such as the First-Order Reliability Method. It is important to note that the reliability analysis methods discussed in this section are only suitable for addressing aleatory uncertainties. Other methods must be used for analyzing epistemic uncertainties.

2.2.1 Key Aspects of Probability Theory

Properties exhibiting random behavior are typically expressed using probability distributions. These distributions provide a mathematical relationship between a given a value of a random parameter and the frequency with which it is realized. Every continuous probability distribution can be expressed in one of two forms: a probability density function (PDF) or a cumulative density function (CDF). Probability density is the likelihood that a random parameter will fall between two values as the distance between the values approaches 0. When this is expressed as a continuous function, the result is the probability density function $f_X(X)$. Thus, given a probability density function for the random parameter X, the probability of X falling between two values a and b is the area under the curve between these values:

$$P_{a \leq X \leq b} = \int_a^b f_X(X)\, dX \tag{6}$$

The cumulative density F_X is the probability that a value less than X will be realized. Thus, the cumulative density is the area under a continuous probability density function from negative infinity to X:

$$F_X = \int_{-\infty}^X f_X(X)\, dX \tag{7}$$

If this is evaluated for every X, the result is the cumulative density function $F_X(X)$. Thus, the probability that a random parameter falls between a and b is simply the difference in the cumulative density functions evaluated at a and b:

$$P_{a \leq X \leq b} = F_X(b) - F_X(a) \tag{8}$$

Among the most commonly used distributions is the normal or Gaussian distribution. This distribution provides a good representation of many systems since the summation of a large number of random variables tends towards a normal distribution regardless of the types of distributions, a property commonly referred to as the Central Limit Theorem. The normal distribution is symmetric about its mean and can be completely described by its first two

statistical moments: its mean, μ, and the square root of its variance, the standard deviation σ. Given these two properties, the probability density function for a normal distribution is given by:

$$f_X(X) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2\right] \tag{9}$$

A related property of normal distributions is that the linear sum of normally distributed parameters is exactly normally distributed. For example, given a function $g(X_1, X_2, \ldots, X_n)$:

$$g(X_1, X_2, \ldots, X_n) = a_0 + \sum_{i=1}^{n} a_i X_i \tag{10}$$

The mean of g is given by Equation (11) and the standard deviation is given by Equation (12):

$$\mu_g = a_0 + \sum_{i=1}^{n} a_i \mu_i \tag{11}$$

$$\sigma_g = \sqrt{\sum_{i=1}^{n} (a_i \sigma_i)^2} \tag{12}$$

A normal distribution with zero mean and unit standard deviation is commonly referred to as a standard normal distribution, with a probability density function denoted ϕ and a cumulative density function denoted Φ.

When multiple random parameters are present, the likelihood of a specific combination being realized is given by a joint probability density function. This density function has the convenient property of being rotationally symmetric when all random parameters have uncorrelated standard normal distributions. That is, if a line is drawn from the origin in a space defined by orthogonal axes for each random parameter, the probability density along that line is only a function of the distance from the origin and will not vary regardless of how the line is rotated. Furthermore, if the distance along the line is parameterized into a single variable s then the probability density of s given by:

$$f_s(s) = \phi(s) \times \phi(0)^{n-1} \tag{13}$$

8
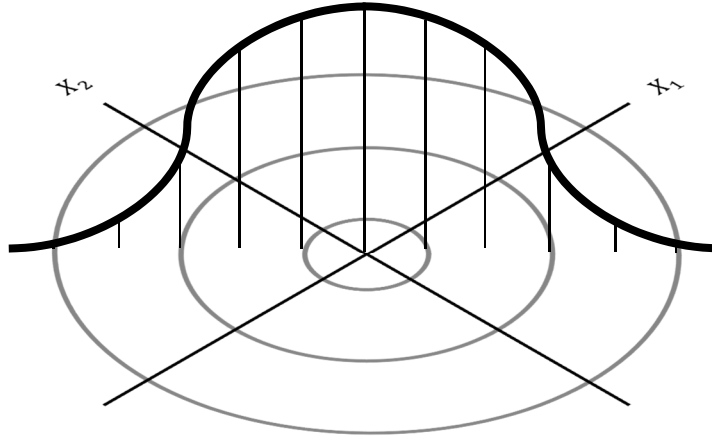
Figure 2: Rotationally symmetric joint probability density function

Where ϕ is the probability density function of a standard normal distribution and n is the number of random parameters. This implies that the probability density decreases as the distance from the origin increases. Thus, all points having a given probability density can be represented by a circle, sphere, or hypersphere in this space, depending on the number of random parameters. An example of this property for 2 random parameters is shown in Figure 2. Several reliability analysis approaches take advantage of this symmetry, as discussed in later sections.

2.2.2 *Monte-Carlo Simulation*

Probability of failure is the idea that if a system is observed an infinite number of times, a specific percentage of these realizations would not satisfy a given performance metric, such as a beam deflecting in excess of a given amount in 0.05% of all realizations. While it is impossible to observe a system an infinite number of times, it is reasonable to assume that a good estimate of the probability of failure can be obtained after sufficiently many observations. This is the basis of Monte-Carlo simulation and its derivatives: the probability distributions of each random parameter are used to generate a finite number of random system configurations which are then analyzed. The probability of failure is then estimated by the fraction of realizations which failed.

$$P_f = \frac{N_{fail}}{N_{total}} \tag{14}$$

The advantage of this approach is that the estimate of the probability of failure will converge to the actual value as the number of samples approaches infinity. However, this is also the drawback of these methods: many samples may be required to obtain a good estimate, especially as the number of random parameters increases or if the probability of failure is low. Additionally, Monte-Carlo analysis is generally not repeatable since a new set of samples may produce a different probability of failure as a result of the random nature of the sampling. Altogether, these limitations suggest that while Monte-Carlo simulations are a powerful tool for a standalone analysis, they should not be used for applications which require repeated analyses or the sensitivity of the probability of failure to small changes in the system.

2.2.3 Mean Value First-Order Second Moment Method (MVFOSM)

Simulation methods can be computationally expensive, particularly if the analysis of the system requires the solution of a finite element model. This has motivated the development of approximation methods which trade the accuracy of the probability estimate for improvements in the computation requirements. Among the simplest of these approaches is the mean value, first-order second moment method. This approach approximates a function of random variables using the Taylor series expansion about the mean value solution of the function given in Equation (15). In this thesis, the convention that positive values of $g(X)$ indicate that this constraint has been violated is used.

$$\tilde{g}(X) \approx g(\boldsymbol{\mu}) + \nabla g(\boldsymbol{\mu})^T (X - \boldsymbol{\mu}) \leq 0 \tag{15}$$

Since the above equation has the same form as Equation (10), the linear approximation of the random function has a normal distribution. Therefore, the mean and standard deviation are given by Equations (11) and (12), where a_0 is the linear approximation evaluated at the mean value point and the a_i are the gradient of $g(X)$ evaluated at the mean value point.

Using the linear expansion, the location of the failure point relative to the mean can then be estimated. It is common to express this as the reliability index, β, which is defined as the number of standard deviations between the mean value of $\tilde{g}(X)$ and $\tilde{g}(X) = 0$. Assuming that the mean value point is feasible, the reliability index is given by Equation (16).

$$\beta = -\frac{\mu_{\tilde{g}}}{\sigma_{\tilde{g}}} \tag{16}$$

The reliability index and the probability of failure are related since the probability distribution of the linear approximation is known to be a normal distribution. This relationship is depicted in Figure 3. Mathematically, this relationship is given by:

$$P_f = 1 - F_X(\mu_{\tilde{g}} + \beta\sigma_{\tilde{g}}) \tag{17}$$
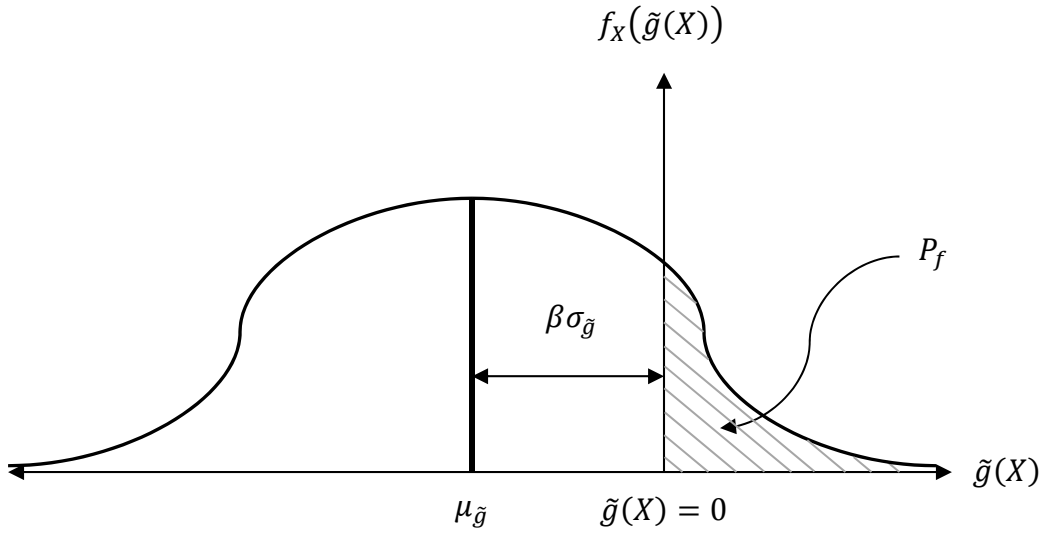
Figure 3: Relationship between the probability of failure and the MVFOSM reliability index

It is important to note that the equivalent probability of failure will differ from the actual probability of failure as a result of the linearization of the system. Additionally, the actual probability distribution may not be exactly normal, especially if $g(X)$ is nonlinear. More concerning however is the fact that the reliability index computed by MVFOSM is not invariant with respect to the problem formulation, as shown by *Choi et. al* example 4.2 [2]. That is, a different value of the reliability index may be computed for different algebraically equivalent expressions of $g(X)$. Because of this, the MVFOSM approach is typically not used by itself for reliability analysis. However, this approach is still commonly used to provide an initial estimate for more complex approaches, such as the First-Order Reliability Method.

2.2.4 First-Order Reliability Method (FORM)

The First-Order Reliability Method originally proposed by Hasofer and Lind [3] is an improvement to the MVFOSM approach which makes the reliability index invariant at the cost of increased computational effort. As discussed in Section 2.2.1, an uncorrelated standard multivariate normal distribution is rotationally symmetric about the origin of the random parameter space. Furthermore, the probability density decreases predictably as the distance from the origin increases. This suggests that the distance from the origin in this space can be a good proxy for the random behavior of the system, similar to the MVFOSM reliability index

Since many reliability problems do not have this joint probability density function, the first step of FORM is to map each random parameter onto a standard normal distribution. For a given point X_i on a normal probability distribution with mean μ_i and standard deviation σ_i, the equivalent point on a standard normal distribution U_i is given by:

$$U_i = \frac{X_i - \mu_i}{\sigma_i} \tag{18}$$

Next, the failure surface of the constraint can be mapped into this new space. The point on the failure surface which is closest to the origin will therefore have the highest joint probability density. This point is commonly referred to as the Most Probable Point (MPP) and is designated U^*. An equivalent to the MVFOSM reliability index can then be defined as the number of standard deviations between the origin and the MPP. Since the normalized space has unit standard deviation, this is equal to the distance between the origin and the MPP:

$$\beta = \sqrt{\sum_{i=1}^{n} U_i^{*2}} \tag{19}$$

The relationship between the standard normalized space, failure surface, MPP, and reliability index for a two-dimensional problem is visualized in Figure 4.

12
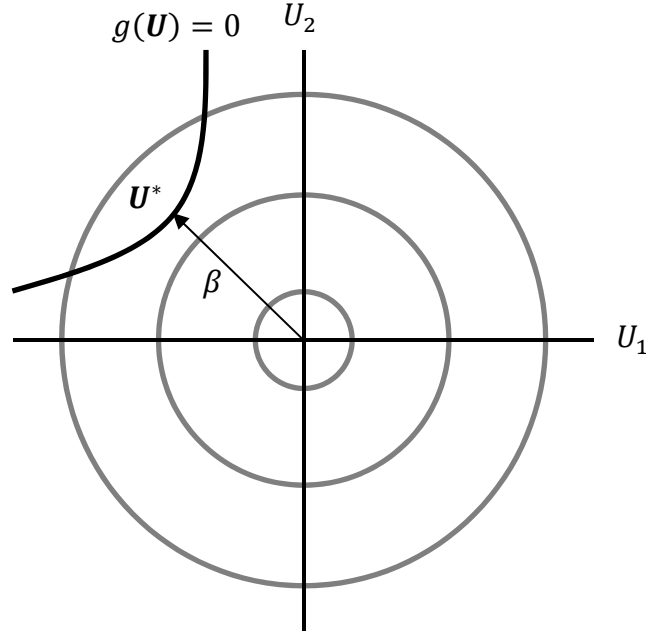
Figure 4: Projection of the failure surface into a random space with two dimensions

The fact that the MPP is a local minimum with respect to the distance from the origin implies that the constraint surface is orthogonal to the radial direction at the MPP. Therefore, linearization about the MPP will result in a line, plane, or hyperplane $\tilde{g}(U)$ which is also orthogonal to the radial direction. By considering the rotational symmetry of the joint probability density function, a relationship between the probability of failure and the reliability index can then be developed. The probability of failure can be approximated by integrating the region beyond the linearized failure surface, shown in Figure 5. While this calculation may be difficult when the radial direction is at an angle relative to the axes of the normalized space, the rotational symmetry of the probability density implies that the solution will be identical to that of a hyperplane orthogonal to one of the axes and located at the same distance from the origin. This rotation is also depicted in Figure 5. Since the equivalent failure surface is parallel to every other random axis, the integration of the uncorrelated multivariate standard normal distribution with respect to each of these random parameters is exactly 1. All that remains is the integration along the perpendicular axis, which is simply the integration of the standard normal probability density function. Thus, the relationship between the FORM reliability index and the probability density is given by:

$$P_f = 1 - \Phi(\beta) \tag{20}$$
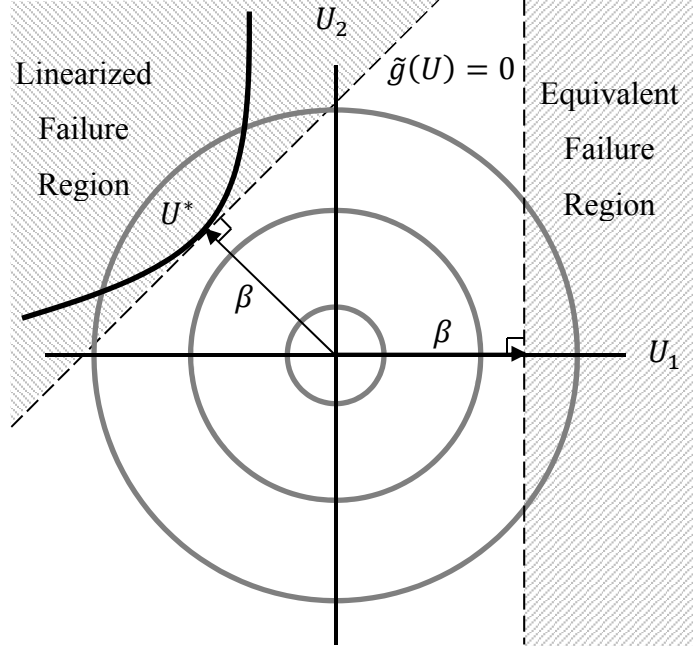
Figure 5: Rotation of the linearized failure surface

The challenge is that the MPP is typically not known beforehand, however it can be found by minimizing the reliability index with U constrained to lie on the failure surface:

$$\min_{U} \beta = \|U\|$$
$$\text{s. t. } g(U) = 0 \tag{21}$$

Thus, the First-Order Reliability Method is able to obtain an estimate of the random behavior of the system by solving a deterministic optimization problem. A commonly used approach for solving this minimization problem is the HL Iteration procedure originally developed by Hasofer and Lind [3]. This method is similar to the MVFOSM approach, except sensitivity factors are used to update the linear expansion point until it converges to the MPP. However, it has been observed that this procedure is not robust and will sometimes fail to converge to an optimum [4]. In general, any optimization procedure can be used to solve the FORM problem.

Another issue related to the optimization procedure is the choice of the initial point. While it is intuitive to use the mean value point as a starting point for the iteration procedure, the gradient at this point with respect to the standard normal random parameters, given by Equation (22), will be undefined since the reliability index of the mean value point is 0 by definition:

14

$$\frac{\partial \beta}{\partial U_i} = \frac{U_i}{\beta} \tag{22}$$

To avoid this issue, it is common to use the MVFOSM approach discussed in Section 2.2.3 to obtain an initial non-zero estimate of β. Additionally, it is important to note that FORM is only applicable to problems which exclusively have normally distributed variables, however extensions exist for non-normal distributions including the Hasofer Lind–Rackwitz Fiessler (HL-RF) method [5]. Lastly, the transformation given by Equation (18) indicates that there are actually two sets of variables used when solving the optimization problem: the optimization variables U_i and the analysis variables X_i. While it is possible to formulate the optimization process exclusively in terms of one or the other, this may not be convenient. If the transformation is used instead, the chain rule relationships given by Equations (23) and (24) can be used to obtain equivalent gradients:

$$\frac{\partial g}{\partial U_i} = \sigma_i \frac{\partial g}{\partial X_i} \tag{23}$$

$$\frac{\partial g}{\partial X_i} = \frac{1}{\sigma_i} \frac{\partial g}{\partial U_i} \tag{24}$$

Overall, FORM is a commonly used procedure for reliability analysis since it captures the random behavior of the system using a deterministic process. This approach is compatible with existing analysis tools since the process of estimating the MPP replaces each probability distribution with a single value. Additionally, the computational costs of the procedure scale well with the number of random variables in comparison to Monte Carlo simulation. Lastly, if the failure surface is linear or nearly-linear, the estimate of the probability of failure will be accurate.

2.2.5 Performance Metric Approach (PMA)

An alternative to the First-Order Reliability Method is the Performance Metric Approach [6] developed by Tu et al. Instead of finding the MPP, the PMA approach seeks to find the worst performing point for a specified reliability requirement, $\bar{\beta}$. This point can then be compared to a performance metric, such as an allowable stress or a maximum deflection. If the worst case point satisfies this metric, the system can be considered reliable. This approach is derived using the same principles as the First-Order Reliability Method, however the optimization problem is inverted, as shown in Equation (25).

$$\max_{U} g(U)$$
$$\text{s.t.} \, \bar{\beta} - \|U\| = 0 \tag{25}$$

Note that the maximization problem is a result of the sign convention used, where positive values of the constraint $g(U)$ are considered to be infeasible. While this method is not the focus of this thesis, it has been regularly used to solve Reliability-Based Topology Optimization problems, as discussed in Section 2.6. This approach also has many of the same shortcomings as the First-Order Reliability Method.

2.3 Reliability-Based Design Optimization (RBDO)

2.3.1 Double-Loop Approaches

The First-Order Reliability Method can easily be implemented in a design optimization scheme, however the result is a double-loop optimization problem. That is, for every iteration of the design optimization routine, the FORM optimization problem needs to be solved while treating the current design variables as fixed parameters. A general RBDO problem can be expressed as follows:

$$\min_{x} F(x)$$
$$\text{s.t.} \, G(x) = \bar{\beta} - \beta \leq 0 \tag{26}$$
$$H(x) = 0$$

16

Where x are the deterministic design variables, $G(x)$ is the reliability constraint, and $H(x)$ is a set of state equations such as the finite element equations. Additional deterministic constraints can be introduced to the problem as needed. The reliability constraint requires the computation of the reliability index each iteration of the design problem by solving the FORM problem:

$$\min_{U} \beta = \|U\|$$
$$\text{s.t.} \, g(U, x) = 0 \qquad\qquad (27)$$
$$h(U, x) = 0$$

Where $g(U, x)$ is the non-deterministic limit state and $h(U, x)$ is a set of state equations which may need to be satisfied while evaluating the constraint. A typical scheme is depicted below in Figure 6. While the Performance Metric Approach discussed in Section 2.2.5 is not the focus of this research, the double-loop structure of the problem would be identical.

The design problem is commonly solved using a gradient-based optimizer. While these approaches are efficient, they typically require good estimates of the sensitivities of the constraints to quickly converge. Since the reliability constraint is itself an optimization problem, this indicates that the sensitivities of the solution of the optimization problem are required. For FORM, these sensitivities are given by Equation (28).

$$\frac{\partial \beta}{\partial x_i} = \frac{1}{\beta} \sum_{j=1}^{n} U_j^* \frac{\partial U_j^*}{\partial x_i} \qquad\qquad (28)$$

Depending on the sensitivity analysis method used, additional solutions of the FORM problem may be required each iteration. This highlights the primary drawback of the double loop approach: while it is simple to implement, the repeated solutions of the nested optimization problem are computationally expensive.

RBDO

```
┌─────────────────┐
│  Initialize RBDO │
│     Problem      │
└─────────────────┘
```

Compute F, ∇F

Evaluate Reliability Constraint using FORM

Sensitivity Analysis of FORM

Update Design Variables

Convergence Check

Stop

FORM

Initialize FORM Problem

Compute β, $g(\boldsymbol{U})$ and $\nabla g(\boldsymbol{U})$

Update Random Variables

Convergence Check

Return β
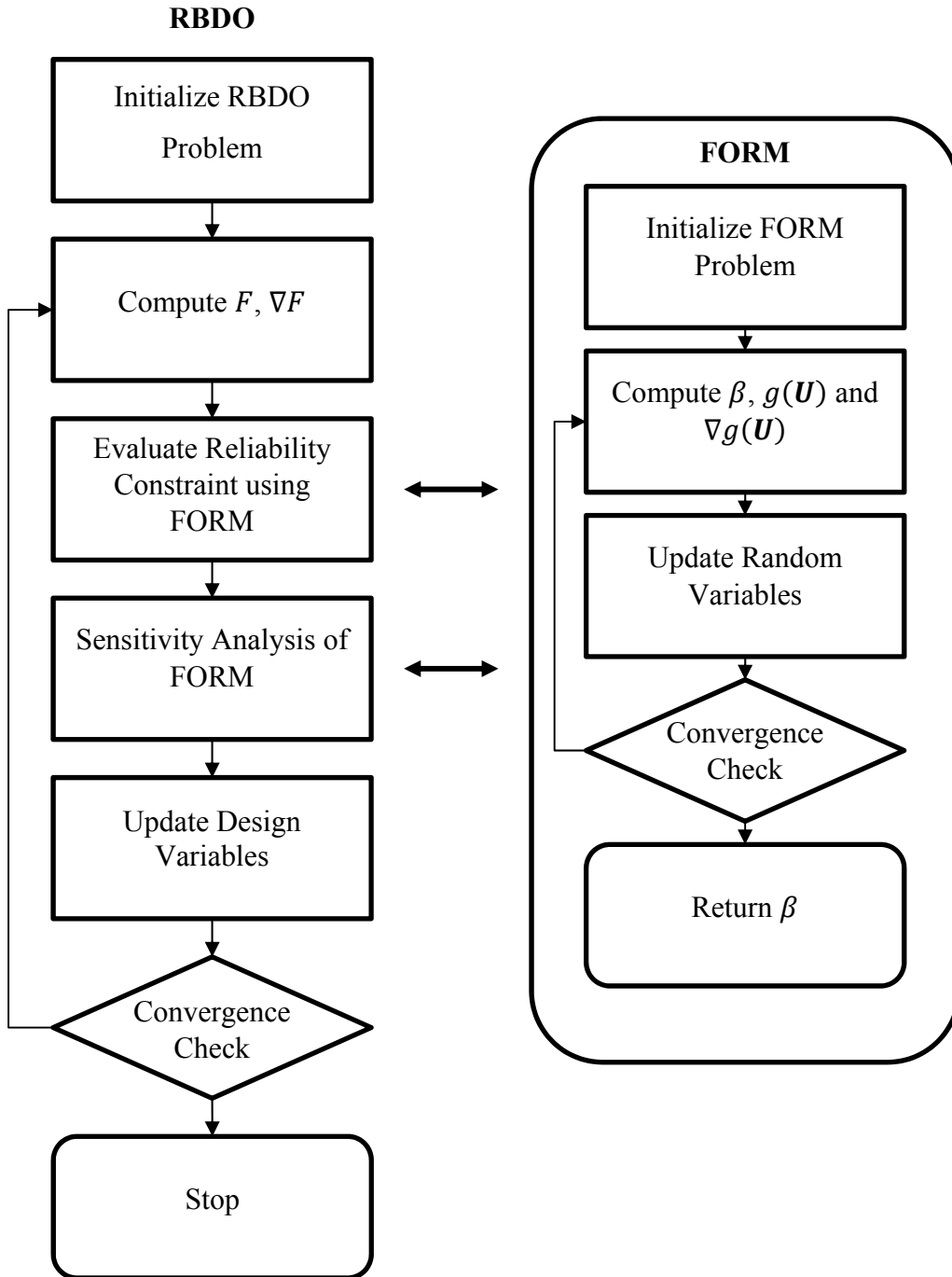
Figure 6: Typical RBDO procedure

2.3.2 Single-Loop and Decoupling Approaches

A number of alternative approaches to the double-loop problem have been proposed with the goal of reducing the computational effort required. While these methods are not the focus of this research, they will be briefly mentioned below since several have been used to solve Reliability-Based Topology Optimization problems.

As the name implies, single-loop approaches modify the reliability analysis problem so that it is coupled with the design optimization problem in a single loop. The design point and the MPP then converge simultaneously. Typically, these approaches are derived by considering the KKT conditions of the reliability analysis problems. Examples of these approaches include the Single Loop Single Vector (SLSV) algorithm developed by Chen *et al.* [7], the single loop methods for reliability index analysis and PMA developed by Silva *et al.* [8] and the Single Loop Approach (SLA) developed by Liang *et al.* [9].

Decoupling approaches extend this idea further by completely separating the reliability analysis and the deterministic design. For example, the Sequential Optimization and Reliability Assessment approach [10] expresses the RBDO problem as a series of alternating design optimization and reliability analysis steps. A shift parameter is computed after each reliability analysis in order to adjust the design optimization until the design converges to a point which satisfies the reliability constraint. Another approach is the Sequential Approximate Programming approach developed by Cheng *et al*. [11] This approach decomposes the optimization problem into a sequence of sub-problems with approximate objectives and constraints which are valid locally around the design point. A set of recurrence formulas in terms of the design point are used to compute the reliability index and MPP simultaneously, removing the need for a nested reliability analysis problem.

A comparison study of these methods was performed by Aoues and Chateauneuf [12]. This investigation concluded that the double loop methods are simplest to implement, however the decoupling approaches are generally more efficient and accurate. Additionally, the SLA approach is promising; having simplicity, efficiency, accuracy, and robustness.

2.4 Sensitivity Analysis

Many optimization schemes are gradient-based, using derivatives or sensitivities of the objective and constraint functions to iterate towards a local minimum. This includes schemes such as sequential quadratic programming (SQP), MATLAB's interior point algorithm [13], and the Method of Moving Asymptotes [14]. It may be possible to derive the sensitivities for simple functions, however in many engineering analyses the objective or constraint functions are sufficiently complex that it is more practical to use a numeric scheme to compute the sensitivities instead. Several of these schemes are detailed in the following subsections.

2.4.1 Finite Difference Method

The simplest approximation for a derivative is the finite difference approach. This approach comes from the definition of a derivative, which looks at the change in a function over an infinitesimally small step. Numeric procedures require a finite step size, so the derivative is instead approximated by the change in a function f over a small distance, h:

$$\frac{\partial f}{\partial x} \approx \frac{f(x+h)-f(x)}{h} \tag{29}$$

The finite difference method is simple to implement but has several drawbacks. First, error is introduced since the derivative is now the change over a non-infinitesimal distance. This error typically increases with the step size. While this would suggest that the smallest step size possible should be used, the precision limitations of computers can cause subtractive cancellation errors for extremely small distances. Thus, error is minimized at some middle ground which is problem dependent. A step size study is typically required to identify an appropriate value. The other disadvantage of this approach is that two function evaluations are required. For some functions the computation cost of the additional evaluation is trivial, however many functions require the evaluation of a large model so the computation cost of this method may be impractical.

2.4.2 Complex Step Method

The complex step approach is similar to the finite difference approach in that the derivative is approximated by using information from a nearby point. However, by taking a small step in the complex plane, the subtraction errors are avoided. Thus, step sizes as low as permitted by the precision of the computer can be used, giving a more accurate approximation of the derivative.

$$\frac{\partial f}{\partial x} \approx \frac{\mathrm{Im}[f(x+ih)]}{h} \tag{30}$$

Similar to the finite difference approach, this requires the evaluation of the function at a nearby point. This approach may also be more challenging to implement since the analysis must be compatible with complex numbers.

2.4.3 Direct and Adjoint Methods

In many cases, the function contains information from the solution of a large system of linear equations such as the displacements found by a finite element analysis. A derivative of this constraint therefore requires a derivative of the entire system of equations. Take for example the equilibrium equations for a finite element analysis:

$$\boldsymbol{KD} = \boldsymbol{R} \tag{31}$$

Where $\boldsymbol{K}$ is the stiffness matrix, $\boldsymbol{R}$ is the loads vector, and $\boldsymbol{D}$ is the vector of displacements, including a displacement of interest. A constraint on a given displacement D_i may then be written:

$$g = D_i - D_{max} \tag{32}$$

The derivative with respect to a given design variable is therefore:

$$\frac{\partial g}{\partial x} = \frac{\partial D_i}{\partial x} \tag{33}$$

Two approaches exist for efficiently computing these derivatives: the direct method and the adjoint method. In the direct method, the derivative of the displacement vector with respect to the design variables is found by taking the derivative of each side of the system, then solving for the vector of sensitivities:

$$\frac{\partial \boldsymbol{D}}{\partial x} = \boldsymbol{K}^{-1}\left(-\frac{\partial \boldsymbol{K}}{\partial x}\boldsymbol{D} + \frac{\partial \boldsymbol{R}}{\partial x}\right) \tag{34}$$

While this provides the sensitivity information required, the calculations have to be repeated for each design variable. The sensitivities of many other displacements are also computed even if they are not needed. Alternatively, the adjoint method can be used. First, the displacement is expressed as:

$$g = \boldsymbol{z}^T\boldsymbol{D} - D_{max} \tag{35}$$

Where $\boldsymbol{z}^T$ is a vector which is zero everywhere except for the ith term. The equilibrium equations can then be adjoined by multiplying them by an unknown vector $\boldsymbol{\lambda}$ which reduces the equations to a scalar:

$$g = \boldsymbol{z}^T\boldsymbol{D} - D_{max} + \boldsymbol{\lambda}^T(\boldsymbol{KD} - \boldsymbol{R}) \tag{36}$$

After taking the derivative, terms which multiply $\frac{\partial \boldsymbol{D}}{\partial x}$ can be grouped, as shown below:

$$\frac{\partial g}{\partial x} = (\boldsymbol{z}^T + \boldsymbol{\lambda}^T\boldsymbol{K})\frac{\partial \boldsymbol{D}}{\partial x} + \boldsymbol{\lambda}^T\left(\frac{\partial \boldsymbol{K}}{\partial x}\boldsymbol{D} - \frac{\partial \boldsymbol{R}}{\partial x}\right) \tag{37}$$

Through careful choice of $\boldsymbol{\lambda}$, $\frac{\partial \boldsymbol{D}}{\partial x}$ can be multiplied by 0, effectively removing it from the calculation. The resulting expression for the derivative of the constraint is:

$$\frac{\partial g}{\partial x} = -\boldsymbol{z}^T\boldsymbol{K}^{-1}\left(\frac{\partial \boldsymbol{K}}{\partial x}\boldsymbol{D} - \frac{\partial \boldsymbol{R}}{\partial x}\right) \tag{38}$$

Comparison of equations (34) and (38) shows that the expressions for the sensitivities are identical even though the approaches differed. Principles from each approach will be used later in this thesis.

22

2.4.4 First-Order Necessary Conditions

In addition to providing information which is useful for efficiently iterating towards an optimum, information about the sensitivities can also be used to determine if a local minimum or maximum has been found. If the optimization problem is constrained by a set of equality statements, then the Lagrange Multiplier Theorem can be used to assess optimality. First, the Lagrangian function is defined:

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} h_i(\boldsymbol{x})\lambda_i \tag{39}$$

Where $f(\boldsymbol{x})$ is the objective function, the $h_i(\boldsymbol{x})$ are equality constraints, and the λ_i are scalar multipliers for each constraint, commonly referred to as Lagrange multipliers. The Lagrange Multiplier Theorem states that at a local minimum $\boldsymbol{x}^*$, the derivatives of Lagrangian with respect to each design variable and each Lagrange multiplier are zero:

$$\nabla_x L(\boldsymbol{x}, \boldsymbol{\lambda})|_{x=x^*} = 0$$
$$\nabla_\lambda L(\boldsymbol{x}, \boldsymbol{\lambda})|_{x=x^*} = 0 \tag{40}$$

For simple problems, these conditions can be used to directly solve for the optimum point. This is rarely done in practice however due to the complexity of many engineering problems. Instead, these conditions are commonly used to assess the convergence of an optimization scheme and as the basis for penalty methods [15].

The Lagrange Multiplier Theorem is only applicable for problems with equality constraints, however the Karush-Kuhn-Tucker (KKT) conditions extend the principles of this theorem to inequality constraints. For a general optimization problem with objective function $f(\boldsymbol{x})$, inequality constraints $g_i(\boldsymbol{x})$, and equality constraints $h_j(\boldsymbol{x})$, the Lagrangian function is given by:

$$L(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} g_i(\boldsymbol{x})v_i + \sum_{j=1}^{n} h_j(\boldsymbol{x})\lambda_j \tag{41}$$

The KKT conditions can then be expressed in terms of the Lagrangian function. There are a total of four conditions. First, $\boldsymbol{x}^*$ must be feasible, indicating that it satisfies all constraints:

23

$$\nabla_v L(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\lambda})|_{x=x^*} \leq 0$$

$$\nabla_\lambda L(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\lambda})|_{x=x^*} = 0 \tag{42}$$

The second set of conditions is the first-order necessary conditions:

$$\nabla_x L(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\lambda})|_{x=x^*} = 0 \tag{43}$$

The third set of conditions is commonly referred to as the switching conditions. While equality constraints must be active, requiring $\boldsymbol{x}^*$ to lie on their failure surfaces, inequality constraints do not have to be active to be satisfied. This condition therefore adds slack, allowing each constraint to either be active, having a value of 0, or inactive, having a zero multiplier.

$$g_i(\boldsymbol{x}^*)v_i = 0 \tag{44}$$

The last set of conditions is the non-negativity conditions. These conditions limit the inequality constraint multipliers to have positive values. This requirement stems from the sign convention that positive constraint values are infeasible.

$$v_i \geq 0 \tag{45}$$

It is important to note that both the Lagrange Multiplier Theorem and the KKT conditions are only applicable if the problem is under-constrained and the constraint gradients are independent.

2.4.5 Sensitivity Analysis for RBDO Applications

Sensitivity analysis specific to RBDO applications has been considered by several authors. This includes an early investigation by Hohenbichler and Rackwitz [16] which showed that the derivative of the reliability index with respect to distribution parameters is fundamentally similar to the derivative with respect to design variables. This allows them to be grouped into a common vector τ and treated identically. A sensitivity expression is then derived using asymptotic approximations:

$$\frac{\partial \beta}{\partial \tau} = \frac{1}{\|\nabla_U g(\boldsymbol{U}^*, \tau_0)\|} \frac{\partial}{\partial \tau} g(\boldsymbol{U}^*, \tau_0) \tag{46}$$

24

A similar expression can also be derived using a perturbation procedure [11]. A small perturbation in β can be expressed as the result of a small change in the MPP times the sensitivity of β at the MPP:

$$\delta\beta = \nabla_U\beta\delta\boldsymbol{U}^* \tag{47}$$

Additionally, since the MPP is a local minimum, the KKT conditions given by Equation (42) will be satisfied, giving the relationship between the objective and constraint gradients:

$$\nabla_U\beta = -\lambda\nabla_U g(\boldsymbol{x},\boldsymbol{U}^*) \tag{48}$$

The non-deterministic constraint can then be perturbed by a small change in the design variables. If it is assumed that a new MPP will always be found then the equality constraint will still be satisfied:

$$\delta g(\boldsymbol{x},\boldsymbol{U}^*) = \nabla_x g(\boldsymbol{x},\boldsymbol{U}^*)\delta\boldsymbol{x} + \nabla_U g(\boldsymbol{x},\boldsymbol{U}^*)\delta\boldsymbol{U}^* = 0 \tag{49}$$

After combining the above equations and solving for the Lagrange Multiplier, the following expression is then obtained for the sensitivities of the reliability index to the design variables:

$$\nabla_x\beta = \frac{\nabla_x g(x,U^*)}{\|\nabla_U g(x,U^*)\|} \tag{50}$$

A more general expression was derived by Kwak and Lee[17] using a perturbation approach. This approach states that for any nested state equations $\boldsymbol{h}$ and nested inequalities $\boldsymbol{g}$, the sensitivity of the reliability index is given by:

$$\partial\beta = \lambda\frac{\partial h}{\partial x}\delta\boldsymbol{x} + v\frac{\partial g}{\partial x}\delta\boldsymbol{x} \tag{51}$$

Lastly, a general approach for double-loop optimization problems with inequality constraints is suggested by Haftka *et al.* [15]. Derivatives of the KKT conditions given by Equations (42) and (43) for a nested problem in terms of variables $\boldsymbol{y}$ can be taken with respect to a given outer variable x:

$$\frac{d}{dx}\nabla_y L(\boldsymbol{y}, \lambda)\bigg|_{y=y^*} \leq 0$$

$$\frac{d}{dx}\nabla_\lambda L(\boldsymbol{y}, \lambda)\bigg|_{y=y^*} = 0 \tag{52}$$

Assuming that the active constraints do not change, this expands into a system of equations in terms of the sensitivities of the nested solution to the design variables $\frac{dy^*}{dx}$:

$$(\boldsymbol{A} - \boldsymbol{Z})\frac{d\boldsymbol{y}^*}{dx} - \boldsymbol{N}\frac{d\lambda}{dx} + \frac{\partial}{\partial x}(\nabla_y f) - \left(\frac{\partial \boldsymbol{N}}{\partial x}\right)\lambda = 0$$

$$\boldsymbol{N}^T \frac{d\boldsymbol{y}^*}{dx} + \frac{\partial g}{\partial x} = 0 \tag{53}$$

Where $\boldsymbol{A}$ is the Hessian of the objective with respect to $\boldsymbol{y}$, $\boldsymbol{z}$ is a summation of the Hessians of the constraints with respect to $\boldsymbol{y}$ multiplied by their Lagrange multipliers, and $\boldsymbol{N}$ is a matrix of the derivatives of the constraints with respect to $\boldsymbol{y}$:

$$A_{ij} = \frac{\partial^2 f}{\partial y_i \partial y_j}$$

$$Z_{il} = \sum_j \frac{\partial^2 g_j}{\partial y_i \partial y_j}\lambda_j \tag{54}$$

$$N_{ij} = \frac{\partial g_j}{\partial y_i}$$

It is important to note that this approach has extra equations since the Lagrange multipliers will also be functions of the outer variables. Depending on the application, the sensitivities of the Lagrange multipliers may or may not be needed so the extra computational effort required is potentially a disadvantage. It is also important to note that second order information is required to use this approach.

2.5 Topology Optimization

Optimization problems within the field of structural optimization can broadly be categorized into three classes: sizing, shape, and topology optimization. Generally speaking, the difference between these approaches is the number of assumptions made. In sizing optimization, the geometry of the problem is well defined with the objective being to identify the best combination of dimensions. A typical sizing problem would be to determine the dimensions of an I-beam or the cross sectional areas of the members of a truss structure. In shape optimization, the general structural configuration is known however boundaries are allowed to move. For example, the coordinates of different joints in a truss structure might be treated as design variables. Lastly, in topology optimization the geometry is generally unknown, with only a set of boundary conditions being provided. The objective is then to identify the most efficient distribution of material.

A common structural topology optimization problem is to minimize the compliance of the structure C given a finite amount of material. This requires the solution of a finite element analysis problem each iteration:

$$KD = R \tag{55}$$

Where K is the global stiffness matrix, D is the unknown global displacements of the nodes, and R are the global loads applied at the nodes. The compliance is then defined as:

$$C = D^T KD \tag{56}$$

Compliance minimization is commonly used as the objective of a topology optimization analysis since it serves as a simple and efficient approximation for the behavior of the structure. Compliance is equivalent to the average strain energy density in the structure, so a lower value implies that the structure has not deformed as much, and is therefore stiffer. Derivatives of the compliance are also relatively simple to compute by considering the direct method described in Section 2.4.3.

An example of the solution to a typical compliance minimization problem is depicted in Figure 7. In this problem, commonly referred to as a Mitchell beam, a vertical load is applied at

Figure 7: Topology optimization solution for a Mitchell Beam

the center of a simply supported beam. It is common to model the Mitchell beam using a symmetry boundary condition to reduce the number of design variables needed. In Figure 7, the vertical force then is applied to the upper left corner while the structure is fixed at the lower right corner. The compliance of the structure is then minimized with the constraint that only 30% of the volume can be used. The result is a stiff, efficient structure with respect to the given loading condition. In practice, this topology would then be refined through further iterations of shape and sizing optimization.

2.5.1 Solid Isotropic Material with Penalization (SIMP)

One of the most intuitive approaches to topology optimization would be to simply turn on and off different regions until an optimum solution is found. This idea forms the basis for the Solid Isotropic Material with Penalization (SIMP) approach to the topology optimization of structures[18]. In this approach, the continuum is discretized into a mesh of many small blocks of material, each of which has a pseudo-density ρ_i which governs its behavior. This mesh is also used for finite element analysis so that these densities can be coupled with structural properties. Typically, the pseudo-density acts as a multiplier for the elastic modulus of the material. Thus, elements with zero density do not contribute to the stiffness of the structure, effectively acting like voids.

Ideally, each density would be permitted to have an integer value of either 0 or 1, corresponding to an element which contributes either no stiffness or its maximum stiffness. In

practice, this approach has several limitations. First, singularities can arise in the global stiffness matrix when zero stiffness is prescribed for an element, causing the finite element analysis to fail. To address this, a near-zero value is typically used for the lower bounds of the density instead. Second, discrete optimization is inefficient, especially when there are many design variables. A more efficient approach would be to use a gradient based-optimizer, however this requires continuous design variables. As a result, intermediate values of the density would be permissible. To avoid intermediate values, the SIMP approach replaces the elastic modulus of each element E_i with a power law, given in Equation (57):

$$E_i = E_0 \rho_i^p \tag{57}$$

Where E_0 is a nominal value of the elastic modulus, ρ_i is the pseudo-density of element i, and p is the penalization exponent. The penalization factor drives the optimizer away from intermediate densities by making them less efficient than full values, effectively emulating the discrete behavior with a continuous function. Note that this also requires the assumption that the material is isotropic. Since the elastic modulus ultimately is used in a finite element analysis, it is common practice to factor the density out of the element stiffness matrix:

$$\boldsymbol{k}_e = \rho_i^p \boldsymbol{k}_0 \tag{58}$$

This is particularly beneficial when a uniform mesh is used since every element will share the same unscaled stiffness matrix $\boldsymbol{k}_0$. Additionally, it is important to note that this approach does not make any assumptions about the type of finite element used. Thus, an identical approach can be used for truss problems, plate problems, 3D problems, and problems with higher-order elements.

2.5.2 *Level-Set Methods*

An alternative class of approaches to the topology optimization problem is the level-set methods [19]. These methods are not the focus of this research, however RBTO using these approaches has been explored in literature. In general, level-set methods differ from SIMP by replacing the densities of elements with a global scalar function $\varphi(\boldsymbol{x})$. The level-set curves of this function then define the boundaries of the structure. To iterate the design, a Hamilton-Jacobi equation is developed in terms of the scalar function with respect to space and a fictitious time t:

$$\frac{\partial \varphi(\boldsymbol{x})}{\partial t} + \nabla\varphi(\boldsymbol{x}, t)\frac{d\boldsymbol{x}}{dt} = 0 \tag{59}$$

This can then be expressed in terms of a speed vector of the level set surface, $\Gamma(\boldsymbol{x}, \varphi)$:

$$\frac{\partial \varphi(\boldsymbol{x})}{\partial t} = -\nabla\varphi(\mathbf{x})\Gamma(\boldsymbol{x}, \varphi) \tag{60}$$

The fictitious time is equivalent to the iteration of the design. Additionally, it has been shown that the iteration of this function is equivalent to the descent series of an optimization problem [19]. To simplify the problem, φ can be discretized across a mesh in a manner similar to a finite element problem, which also allows for finite element solutions to be easily incorporated into objectives or constraints. Overall, the level set method has the advantage that the boundaries of the structure can theoretically be defined by a continuous function instead of a set of discrete elements.

2.5.3 *Filtering*

A common issue encountered in topology optimization using the SIMP approach is checkerboarding, a behavior where the optimizer converges towards a solution where elements are only connected diagonally. The result is a grid of alternating zero and full density elements. An example of this behavior for a cantilever beam problem is depicted below in Figure 8. Checkerboarding is not a physically realistic solution however, but is instead an artifact of the finite element discretization since the corner nodes allow for the transfer of stiffness information between diagonal elements, even though the elements are otherwise discontinuous. Thus,
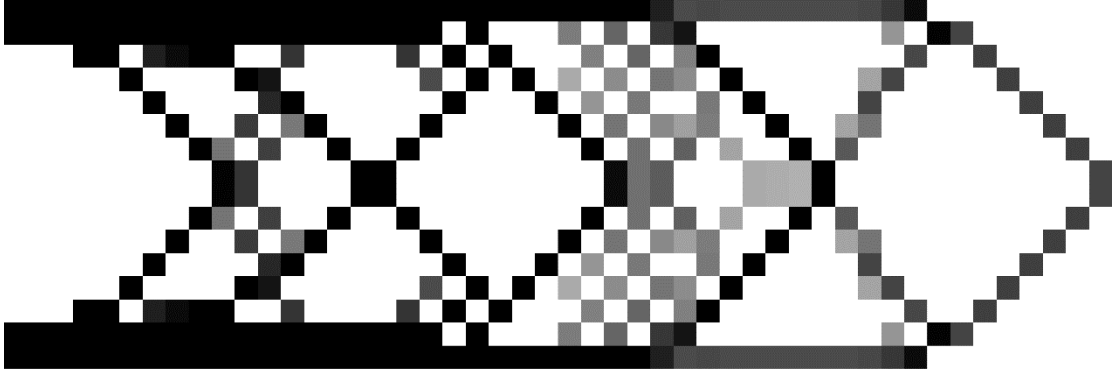
Figure 8: Cantilever beam problem exhibiting checkerboarding

adjacent elements are not required to efficiently fill in a region except when very high stiffness is needed.

Two of the most common approaches for preventing checkerboarding are density filtering and sensitivity filtering. Density filtering expresses the optimization problem as two sets of variables: the physical pseudo-densities ρ and the optimization variables x. The pseudo-densities are the values which are actually used in the finite element analysis, objective function, and constraint while the optimization variables are only used by the optimizer. These two sets of variables are related by expressing each pseudo-density as a weighted sum of the optimization variables [20]:

$$\rho_i = \sum_j w_{ij} x_j \tag{61}$$

A common set of weights are the conic weights. These weights vary linearly with the distance between the centers of elements i and j when this distance $d(i,j)$ is within a specified radius $\bar{r}$. The weights are zero outside of this radius. Additionally, the weights are normalized so that the sum of all weights for a given density i is 1. This prevents the physical densities from exceeding their upper bounds.

$$w_{ij} = \frac{\bar{r} - d(i,j)}{\sum_{k \in N_i}(\bar{r} - d(i,k))} \tag{62}$$

The weights can be computed when initializing the optimization process since they only depend on the geometry and finite element mesh. The weights can then be conveniently grouped into a weighting matrix **A** which relates the pseudo-density vector to the optimization variable vector.

31

$$\begin{Bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_i \\ \vdots \\ \rho_n \end{Bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdots & w_{1j} & \cdots & w_{1n} \\ w_{21} & w_{22} & w_{23} & \cdots & w_{2j} & \cdots & w_{2n} \\ w_{31} & w_{32} & w_{33} & \cdots & w_{3j} & \cdots & w_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & w_{i2} & w_{i3} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \cdots & w_{nj} & \ddots & w_{nn} \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{Bmatrix} \tag{63}$$

$$\{\rho\} = [A]\{x\} \tag{64}$$

One convenient property of this approach is that since this is a linear transformation, the chain rule can be used to relate the sensitivities with respect to the pseudo-densities to those with respect to the optimization variables. This has the advantage of simplifying the sensitivity analysis and preserving the KKT conditions. The sensitivity transformation is given below in Equation (65).

$$\left\{\frac{df}{dx_j}\right\} = [A]^T \left\{\frac{\partial f}{\partial \rho_i}\right\} \tag{65}$$

Overall, this approach avoids the checkerboarding behavior by specifying a maximum change in the density between elements which is a function of the filter radius. This also has the benefit of setting a minimum size for design features since SIMP penalizes intermediate values of the elastic modulus. This drives the optimizer towards features which are sufficiently large enough that a full value of the stiffness can be obtained for at least some elements.

Another commonly used approach is sensitivity filtering [21]. In this approach, both the analysis and the optimization procedure are expressed in terms of the pseudo-densities. The sensitivities of the objective function and/or the constraints however are then modified using Equation (66). Similar to the linear density filter, only elements within a radius $\bar{r}$ are allowed to influence the sensitivity of ρ_i. If an element is outside of this radius, then a value of 0 should be used instead:

$$\frac{\widehat{\partial f}}{\partial \rho_i} = \frac{1}{\rho_i \sum_{j=1}^n \bar{r} - d(i,j)} \sum_{j=1}^n (\bar{r} - d(i,j)) \rho_j \frac{\partial f}{\partial \rho_j} \tag{66}$$

This approach is simpler to implement, however the modified sensitivities cannot be used to assess the KKT conditions.

2.6 Reliability-Based Topology Optimization (RBTO)

Reliability-based topology optimization has been explored by multiple authors since 2003. A brief summary of these investigations is presented in this section, with a focus on the sensitivity analysis procedures used.

Among the earliest attempts to incorporate topology optimization into a RBDO routine was the effort by Kharmanda *et al.* [22]. This approach modified the 99 line topology optimization code developed by Sigmund [23] to include reliability analysis for a random load. Rather than use a double-loop implementation, it was proposed that the FORM problem is solved prior to each iteration in order to determine the MPP of the force. This force is then used in 99 line topology optimization routine to determine the next design. Sensitivity analysis is performed using finite difference derivatives. However, it has been observed that the implementation of this approach was flawed and resulted in a reliability problem with no physical significance [24].

Another early attempt at RBTO was made by Maute and Frangopol [25]. In this research, the deforming structure of a micro-elector-mechanical system is optimized using SIMP with PMA for the reliability analysis. The actuation force and spring stiffness are treated as random parameters. Sensitivity analysis is performed by computing the total derivative of the performance measure with respect to the design variables:

$$\frac{dg(\boldsymbol{U}(\boldsymbol{\rho}),\boldsymbol{\rho})}{d\boldsymbol{\rho}} = \frac{\partial g(\boldsymbol{U}(\boldsymbol{\rho}),\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} + \frac{\partial g(\boldsymbol{U}(\boldsymbol{\rho}),\boldsymbol{\rho})}{\partial \boldsymbol{U}} \frac{\partial \boldsymbol{U}}{\partial \boldsymbol{\rho}} \tag{67}$$

To simplify this calculation, the first order necessary conditions of the PMA problem are then considered:

$$\frac{\partial g(\boldsymbol{U}(\boldsymbol{\rho}),\boldsymbol{\rho})}{\partial \boldsymbol{U}} + \lambda \frac{\partial \beta}{\partial \boldsymbol{U}} = 0 \tag{68}$$

The authors argue that at the MPP, the reliability index is constant so the performance measure does not depend on U. Therefore, the sensitivity is simply:

$$\frac{dg(\boldsymbol{U}(\boldsymbol{\rho}),\boldsymbol{\rho})}{d\boldsymbol{\rho}} = \frac{\partial g(\boldsymbol{U}(\boldsymbol{\rho}),\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \tag{69}$$

Another effort at RBTO considered the minimization of compliance with a spatial varying and correlated elastic modulus represented by a marginally lognormal distribution [26]. It was observed that the tail of the compliance distribution approximates that of a lognormal distribution. Using the relationship between the reliability index and the probability of failure given by Equation (20) an equivalent relationship between the reliability index and statistical moments which are fitted to the tail is then derived. These statistical moments are then expressed in terms of the finite element analysis, allowing for analytic sensitivities of the reliability index with respect to the pseudo-densities to be derived. While this sensitivity analysis approach is efficient, it is very problem specific.

A benchmark problem for RBTO has been proposed and evaluated by Rozvany and Maute [1] using SIMP with PMA. This benchmark problem consists of a rectangular region which is clamped at the bottom, depicted below in Figure 9. A deterministic vertical force and a zero mean, nonzero standard deviation force is applied at the center of the upper surface. The objective is to minimize the volume given a constraint on the reliability of the structure.

The analytic solution for this structure is shown to be a 2-bar truss, with the angle between the bars related to the standard deviation of the horizontal force and the target reliability requirement. The RBTO solution will not exactly match this angle however as a result of the discretization of the region into finite elements. Sensitivity analysis was performed using the procedure from Maute and Frangopol [25], given by Equation (69).
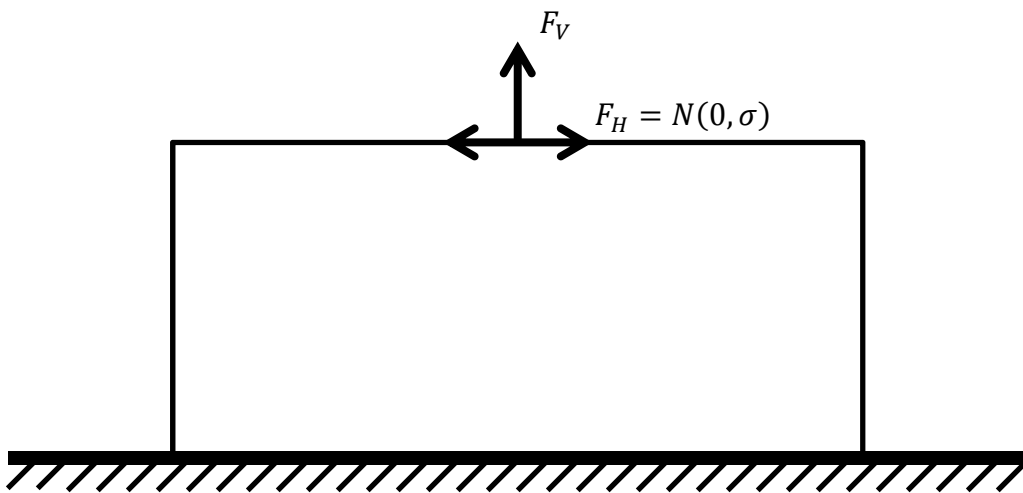

Figure 9: RBTO benchmark problem from Rozvany and Maute [1]

One of the issues related to RBDO in general is the difference between the reliability against a single type of failure and the reliability against failure when multiple modes are possible. This concept was explored in the context of RBTO by Silva *et al.* [8] The authors introduce the concepts of Component Reliability-Based Topology Optimization (CRBTO) and System Reliability Based Topology Optimization (SRBTO), with the difference being that CRBTO expresses each failure criteria as a separate constraint while SRBTO uses a single constraint to capture all possible failure criteria. Using a single loop method, the authors were able to solve a number of RBTO problems using each approach, demonstrating that the reliability of the system is lower when each constraint is considered separately. For the sensitivity analysis, the derivative of the reliability index with respect to a given design variable is approximated by:

$$\nabla_{\rho}\beta \approx \lambda\nabla_{\rho}g(\rho, U) \tag{70}$$

RBTO has also been applied to geometrically nonlinear structures including a plate with random elastic modulus and loads [27]. In this investigation, a double loop problem consisting of an outer sequential linear programming (SLP) SIMP problem was solved with a nested PMA analysis. Since PMA was used, the sensitivity of the nested problem only requires the derivatives of the performance metric with respect to each random variable. This is similar to the problem solved by Maute and Frangopol [25] summarized above. In this case, the performance metric was a displacement constraint, so the adjoint method (Section 2.4.3) was used to derive sensitivities of the displacement with respect to each design variable and the random loads. Interestingly, the authors also performed a study comparing the efficiency of this sensitivity analysis approach to the finite difference method (Section 2.4.1). This study found that the finite difference method required 6698s of computation time to compute the sensitivity of 100 design variables, which was approximately 1000 times the computational effort of the adjoint approach for a problem. This highlights the inefficiency of the finite difference approach for RBTO problems, especially since 100 design variables is a very small topology optimization problem. Lastly, a study was performed which compared RBTO to deterministic topology optimization with and without a factor of safety. Compared to the deterministic solution with no factor of safety, RBTO with a target reliability index of 3 reduced the probability of failure from 50% to 0.135% at the cost of 9.2% higher more material. A safety factor of 1.5 successfully reduced the

probability of failure to 0.00160%, however this required 13.9% more material than the deterministic solution. This showed that both approaches increase the safety of the design, however RBTO has the advantage of being able to target a specific probability of failure.

One of the challenges related to FORM is the inaccuracy due to the curvature of the constraint function. An investigation was performed into improving the accuracy of this approach using a procedure called the segmental multi-point linearization (SML) approach [28]. This approach replaces the single hyperplane approximation used by FORM with several hyperplanes fitted to the response surface in order to better capture the effects of changes in the curvature. A more accurate sensitivity expression was also derived for a general non-deterministic constraint using a perturbation procedure similar to Equation (49). The result is an expression similar to Equation (50), however a surface integral is required since no assumption is made about the linearity of the failure surface:

$$\nabla_x P_f = -\int_S \frac{\phi_n(U)}{\|\nabla_U\|} \nabla_x g \, dS \tag{71}$$

This generally cannot be evaluated exactly, however it can be approximated as a summation of weighted gradients of the constraint evaluated at different fitting points, similar to the SML procedure:

$$\nabla_x P_f \approx \sum_{j=1}^{n} w_j \nabla_x g_j \tag{72}$$

The weights are found by integrating Equation (71) across each hyperplane. Several ground structures problems are then solved using this approach.

An alternative sensitivity analysis procedure was used by Kim *et al.* to solve several RBTO problems using FORM and PMA [29]. For the first-order reliability method, the sensitivities are approximated by taking the derivative of the MVFOSM approximation for the reliability index given by Equation (16):

$$\frac{\partial \beta}{\partial x} = \frac{\partial}{\partial x} \frac{g(X)}{\sqrt{\sum_{j=1}^{n} \left(\frac{\partial g}{\partial X}\right)^2 \sigma_j^2}} \tag{73}$$

The sensitivity analysis used for PMA is simply the derivative of the constraint function with respect to the design variable, evaluated at the worst case point:

$$\frac{\partial g}{\partial x} = \frac{\partial}{\partial x} g(x, U)|_{U=U^*} \tag{74}$$

RBTO was then performed for a statically loaded cantilever plate with a random load, elastic modulus and thickness as well as for a beam subject to an eigenvalue constraint.

Several authors have also explored using decoupled algorithms for RBTO. This includes applying RBTO to large 3D problems including wing ribs and a tail cone structure using the SORA approach [30]. Another approach combined the Hybrid Cellular Automaton (HCA) algorithm with PMA [31]. HCA is an alternative topology optimization algorithm which uses both the density and strain energy to define the state of each element. This leads to an objective function of simultaneously minimizing these two parameters while constrained. Since these RBTO approaches used decoupled algorithms, the sensitivities of the MPP with respect to the design variables were not required.

3. Sensitivity of Nested Optima

3.1 General Derivation

A general double loop optimization problem can be expressed in terms of two sets of variables: the outer design optimization variables $x = x_1, x_2, \ldots, x_i, \ldots x_m$ and the nested design optimization variables $y = y_1, y_2, \ldots, y_j, \ldots y_n$. Since the solution of the nested optimization problem y^* is found using an algorithm, its relationship with x is assumed to be implicit. Additionally, y^* is potentially a function of all outer design variables:

$$y^* = \{y_1^*(x_1, x_2, \ldots, x_m), y_2^*(x_1, x_2, \ldots, x_m), \ldots, y_n^*(x_1, x_2, \ldots, x_m)\}^T \tag{75}$$

In this formulation, the outer loop is only a function of the outer design variables and the optimum of the nested problem:

$$\min_{x} F(x)$$
$$\text{s.t.} \ G_n(y^*(x), x) \leq 0 \tag{76}$$
$$G_o(x) \leq 0$$
$$H_o(x) = 0$$

Where $F(x)$ is the objective of the outer problem, $G_n(y^*(x), x)$ is a constraint which depends on the solution of the nested problem, and $G_o(x)$ and $H_o(x)$ are a set of deterministic inequality and equality constraints which need to be satisfied in the outer problem. Note that, $G_n(y^*(x), x)$ is expressed as an inequality constraint, however this approach is also applicable to equality constraints. The nested optimization problem is given by:

$$y^*(x) = \min_{y} f\big(x, y, \bar{z}(x, y)\big)$$

$$\text{s.t. } g_k\big(x, y, \bar{z}(x, y)\big) = 0, \, k = 1, 2, \ldots, p \tag{77}$$

$$\text{where } \bar{z}(x, y) \text{ is the solution to } h(x, y, \bar{z}) = 0$$

Where $f\big(x, y, \bar{z}(x, y)\big)$ is the nested objective function, $g\big(x, y, \bar{z}(x, y)\big)$ is the nested equality constraint(s), and $h\big(x, y, \bar{z}(x, y)\big)$ is a set of state equations which are solved to obtain the state variables $\bar{z}(x, y)$.

The sensitivities of the nested optima with respect to the outer variables $\frac{\partial y^*}{\partial x}$ may be required for some applications, including computing analytic sensitivities of $G_n(y^*(x), x)$. While these sensitivities could be obtained using a finite difference approach, it would require additional evaluations of the nested problem. Alternatively, analytic expressions for these sensitivities can be derived using the Lagrange Multiplier Theorem. As discussed in Section 2.4.4, at the optimum of the nested problem the gradients of the Lagrangian function will be zero with respect to all y and all Lagrange multipliers:

$$\nabla_y L(x, y, \bar{z}(x, y), \lambda)\big|_{y=y^*(x)} = 0$$

$$\nabla_\lambda L(x, y, \bar{z}(x, y), \lambda)\big|_{y=y^*(x)} = 0 \tag{78}$$

For a given x, the above equations can be solved for $y^*(x)$ and $\lambda^*(x)$. If it can be assumed that a local minimum of the nested problem can be found for any point in the neighborhood of x, then these equations will be satisfied everywhere in the neighborhood of x. Therefore, the derivative of these equations with respect to each design variable x_i can be used to compute the sensitivities of $y^*(x)$ and $\lambda^*(x)$ with respect to x. For a given x_i, this results in a new system of equations:

$$\frac{d}{dx_i} \nabla_y L(x, y, \bar{z}(x, y), \lambda)\big|_{y=y^*(x)} = 0$$

$$\frac{d}{dx_i} \nabla_\lambda L(x, y, \bar{z}(x, y), \lambda)\big|_{y=y^*(x)} = 0 \tag{79}$$

The definition of the Lagrangian function from Equation (39) can then be substituted in. Since the Lagrange multipliers are scalar multipliers for the constraints, the gradient of the Lagrangian

with respect to the Lagrange multiplier is simply the constraint functions. Therefore, the system of equations can be broken into two distinct subsets:

$$\left(\frac{d}{dx_i}\{\nabla_y f(x,y,\bar{z}(x,y))\} + \frac{d}{dx_i}[\nabla_y g(x,y,\bar{z}(x,y))]\{\lambda\} + [\nabla_y g(x,y,\bar{z}(x,y))]\left\{\frac{d\lambda}{dx_i}\right\}\right)\Big|_{y=y^*(x)} = \{0\}$$

$$\frac{d}{dx_i}\{g(x,y,\bar{z}(x,y))\}\Big|_{y=y^*(x)} = \{0\} \tag{80}$$

Where $\{\nabla_y f\}$ is a column vector of the derivatives of the nested objective function, $[\nabla_y g]$ is a $n \times p$ matrix of constraint derivatives, $\{\lambda\}$ is a column vector of the Lagrange multipliers, $\{d\lambda/dx_i\}$ is a column vector of the derivatives of the Lagrange multipliers with respect to x_i, and $\{g\}$ is a column vector of the constraints.

It is important to note that the constraints $g(x,y,\bar{z}(x,y))$ may be functions of the state variables $\bar{z}(x,y)$, implying that any derivatives of g with respect to x_i or y must consider the derivatives of $\bar{z}(x,y)$ with respect to x_i or y. For example, expanding the $[\nabla_y g(x,y,\bar{z}(x,y))]$ term from Equation (80) to explicitly include these derivatives yields the following expression:

$$\nabla_y g(x,y,\bar{z}(x,y)) = \frac{\partial g}{\partial y} + \frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{z}}{\partial y} \tag{81}$$

Taking the derivative again with respect to x_i will then require further expansion. Since $\bar{z}(x,y)$ is the solution to the state equations $h(x,y,\bar{z}(x,y))$, many of these derivatives can be efficiently computed using the direct and adjoint methods discussed in Section 2.4.3. For simplicity, it is assumed for the remainder of this section that any derivatives of g will implicitly include these calculations.

Each of the two subsets of equations can then be considered separately. The first set, containing $\nabla_y f$ and $\nabla_y g$, can be expanded using a total derivative since $y^*(x)$ is implicitly a function of x. In general, the total derivative expansion for a generic function $f(y(x), x)$ is given by:

$$\frac{d}{dx_i}[f(y(x),x)] = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{n}\frac{\partial f}{\partial y_j}\frac{\partial y_j}{\partial x_i} \tag{82}$$

After expanding the first subset of equations with the total derivative and simplifying, the resulting matrix equations are:

$$\left(\left[H(f) + \sum_{k=1}^{p} H(\boldsymbol{g}_k)\,\lambda_k\right]\left\{\frac{\partial \boldsymbol{y}^*}{\partial x_i}\right\} + \left\{\frac{\partial^2 f}{\partial x_i \partial \boldsymbol{y}}\right\} + \left[\frac{\partial^2 \boldsymbol{g}}{\partial x_i \partial \boldsymbol{y}}\right]\{\boldsymbol{\lambda}\} + \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}\right]\left\{\frac{d\boldsymbol{\lambda}}{dx_i}\right\}\right)\Bigg|_{\boldsymbol{y}=\boldsymbol{y}^*} = \{0\} \qquad (83)$$

Where $H(f)$ is the Hessian matrix of f with respect to $\boldsymbol{y}$, $H(\boldsymbol{g}_k)$ is the Hessian matrix of constraint g_k with respect to $\boldsymbol{y}$, $\left\{\frac{\partial \boldsymbol{y}^*}{\partial x_i}\right\}$ is the column vector of the sensitivities of each component of $\boldsymbol{y}^*$ with respect to x_i, $\left\{\frac{\partial^2 f}{\partial x_i \partial \boldsymbol{y}}\right\}$ is a column vector of mixed derivatives, and $\left[\frac{\partial^2 \boldsymbol{g}}{\partial x_i \partial \boldsymbol{y}}\right]$ which is equal to the derivatives of $\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}\right]$ with respect to x_i. This subset of equations contains the desired sensitivities $\frac{\partial \boldsymbol{y}^*}{\partial x_i}$, however it does not form a complete set of equations since the derivatives of the Lagrange multipliers with respect to x_i are also unknown. In total, there are n equations but $n + p$ unknowns in this subset.

A complete set of equations can be formed by considering the second subset of equations. Similar to before, the derivatives of the constraint equations need to be expanded using implicit differentiation. The result is given below in Equation (84).

$$\left(\left\{\frac{\partial \boldsymbol{g}}{\partial x_i}\right\} + \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}\right]^T \left\{\frac{\partial \boldsymbol{y}}{\partial x_i}\right\}\right)\Bigg|_{\boldsymbol{y}=\boldsymbol{y}^*} = \{0\} \qquad (84)$$

Where $\left\{\frac{\partial \boldsymbol{g}}{\partial x_i}\right\}$ is a column vector of the derivatives of $\{\boldsymbol{g}\}$ with respect to x_i. Similar to the first subset, there are more unknowns than equations. Specifically, this subset has p equations but n unknowns.

When combined, Equations (83) and (84) give a complete system of equations in terms of the unknown sensitivities of the MPP and Lagrange multipliers to a given outer optimization problem variable x_i. Conveniently, this system is entirely linear with respect to these sensitivities. The system is also typically small, having a size equal to the sum of the number of random parameters and equality constraints. The system can then be expressed in the form $Ax = b$ so that efficient linear solvers can be used, as shown below:

$$\begin{bmatrix} \left[H(f) + \sum_{k=1}^{p} H(g_k)\,\lambda_k\right] & \left[\dfrac{\partial g}{\partial y}\right] \\[2mm] \left[\dfrac{\partial g}{\partial y}\right]^{T} & [0] \end{bmatrix} \begin{Bmatrix} \left\{\dfrac{\partial y^{*}}{\partial x_i}\right\} \\[2mm] \left\{\dfrac{d\lambda}{dx_i}\right\} \end{Bmatrix} = \begin{Bmatrix} -\left\{\dfrac{\partial^{2} f}{\partial x_i \partial y}\right\} - \left[\dfrac{\partial^{2} g}{\partial x_i \partial y}\right]\{\lambda\} \\[2mm] -\left\{\dfrac{\partial g}{\partial x_i}\right\} \end{Bmatrix} \tag{85}$$

Where all information is evaluated at the current x and nested optima y^{*}. Another useful property of this system is that the left side does not contain any derivatives with respect to x_i. This implies that the left side only needs to be computed once for a given iteration of the outer problem. All sensitivities can then be obtained using a procedure analogous to the solution of a finite element system with multiple load cases. Therefore, the computational cost of this approach will scale linearly with respect to the number of outer optimization variables.

   This approach does have two drawbacks however. First, the sensitivities of the Lagrange multiplier to x_i must be computed as a byproduct of this analysis even though this information may not be needed. This adds an extra equation to the system for each Lagrange multiplier, increasing the computation costs. Additionally, the second derivatives of the nested objective function and constraints are required. It may be impractical or impossible to obtain these for some problems, though it should be noted that approximations may be available depending on the optimization scheme used to solve the nested problem. For example, the Sequential Quadratic Programming scheme requires the Hessian of the Lagrangian function for the nested problem, which can either be supplied or calculated during the optimization procedure using an update process [15]. In that case, the only additional information which needs to be computed is the mixed second derivatives of the objective and constraint functions.

   Overall, the result of this derivation is a procedure which can be used to obtain analytic sensitivities of the nested optima with respect to outer design variables if the nested problem contains only equality constraints. It should be noted that the result of this derivation is identical to the derivation from Haftka *et al*. for inequality constraints [15], discussed in Section 2.4.5. This is because the approach for inequality constraints requires the assumption that active constraints remain active and inactive constraints remain inactive. Thus, the active constraints behave like equality constraints for the sensitivity analysis only.

3.2 Application to the First-Order Reliability Method

As discussed in Section 2.3.1, a reliability-based design optimization routine which uses FORM has a double-loop structure, with the FORM optimization problem nested within a design optimization problem. The FORM problem, given by Equation (21), is a relatively straightforward optimization problem, having only a single equality constraint. Additionally, the sensitivity of the reliability index to each design variable may be expressed as the sum of the MPP times its sensitivities, as shown by Equation (28). Therefore, analytic sensitivities of the reliability index to the design variables can be obtained if the sensitivities of the MPP, $\frac{\partial U}{\partial x_i}$, are evaluated using the procedure described in Section 3.1.

Two simplifications can be made to the system of equations for the sensitivities given by Equation (85). First, FORM only has a single constraint so all constraint derivative matrices reduce to column vectors. Second, β is not explicitly a function of any design variables, so the mixed derivatives of β are all 0. The simplified system of equations is given in Equation (86):

$$
\begin{bmatrix} [H(\beta) + \lambda H(g)] & \left\{\frac{\partial g}{\partial U}\right\} \\ \left\{\frac{\partial g}{\partial U}\right\}^T & 0 \end{bmatrix} \left\{ \begin{matrix} \left\{\frac{\partial U^*}{\partial x_i}\right\} \\ \frac{d\lambda}{dx_i} \end{matrix} \right\} = - \left\{ \begin{matrix} \lambda\left\{\frac{\partial^2 g}{\partial x_i \partial U}\right\} \\ \frac{\partial g}{\partial x_i} \end{matrix} \right\}
\tag{86}
$$

Where all first and second derivatives are evaluated at the MPP and current design $\boldsymbol{x}$. This expression can be further simplified by considering the formulation of β, given by Equation (19). Since β can easily be expressed for any number of random parameters, the Hessian of β can be defined for any problem. Specifically, for a problem with n random parameters the second derivatives with respect to each are given by:

$$
H(\beta) = \frac{1}{\beta^3} \begin{bmatrix} \beta^2 - U_1^2 & -U_1 U_2 & \cdots & -U_1 U_n \\ -U_1 U_2 & \beta^2 - U_2^2 & \cdots & -U_2 U_n \\ \vdots & \vdots & \ddots & \vdots \\ -U_1 U_n & -U_2 U_n & \cdots & \beta^2 - U_n^2 \end{bmatrix}
\tag{87}
$$

Additionally, the Lagrange multiplier λ can be computed using the necessary conditions from Equation (40). Since an optimum has already been found, the only unknown in this system is λ so only one equation is required. For example, λ can be expressed in terms of the derivatives of the Lagrangian with respect to U_1:

$$\lambda = -\frac{U_1^*}{\beta}\left(\frac{\partial g}{\partial U_1^*}\right)^{-1} \tag{88}$$

Therefore, the only problem-specific information required is the derivatives and second derivatives of the constraint. For convenience, these derivatives can be computed in terms of the random parameters then converted to the equivalent standard normal form using the transformation given by Equation (23). It is also important to note that the sensitivities $\frac{\partial g}{\partial U}$ are typically already computed while solving the FORM problem.

A general procedure for obtaining analytic sensitivities of the reliability index each iteration of the design problem can then be constructed. This procedure consists of the following steps:

1. Solve the FORM problem to obtain β, U^*, and the sensitivities $\left\{\frac{\partial g}{\partial U}\right\}$
2. Compute the Lagrange multiplier for the equality constraint using Equation (88)
3. Evaluate the Hessians of β and g at the MPP, then formulate the left hand side of Equation (86)
4. For each x_i compute $\frac{\partial g}{\partial x_i}$ and $\left\{\frac{\partial^2 g}{\partial x_i \partial U}\right\}$, then formulate the right side of Equation (86)
5. Use a linear solver to obtain the sensitivities of the MPP to x_i
6. Evaluate the sensitivity of β to x_i using Equation (28)
7. Repeat steps 4-6 for every x_i

This procedure is validated in Section 4.1 using a 3-bar truss problem. It is then applied to several reliability-based topology optimization problems in the following sections.

4. Results and Discussion

4.1 Overview of the Analyses Performed

Reliability-based design optimization generates designs which are not only optimal, but also reliable, by introducing constraints on the non-deterministic behavior of the system. This comes at the cost of increased computational effort, particularly when a double loop approach is used. This cost can be minimized however if analytic sensitivities of the solution of the nested problem are available. A procedure for obtaining these was derived in Section 3.1 then extended to RBDO using the First-Order Reliability Method in Section 3.2. The remainder of this thesis is devoted to exploring applications for this procedure, including extensions to Reliability-Based Topology Optimization. First, a probabilistic variant of the classic 3-bar truss problem from Choi *et* al. [2] is used to validate the method. Then, three reliability-based topology optimization problems are considered. This includes the benchmark problem developed by Rozvany and Maute [1] which constrains the probabilistic compliance of a structure. Two new problems are then formulated and solved: a probabilistic compliance problem with multiple random loads and a deflection-constrained cantilever structure with a random load. A secondary objective of this research is to further explore the benefits of RBDO and RBTO. To achieve this, the results of these approaches are compared to deterministic equivalents when available.

4.2 RBDO of a 3-Bar Truss Structure

4.2.1 Problem Statement

The first problem considered is the probabilistic 3-bar truss from Choi *et al.* Section 7.3.3 [2]. The truss consists of 3 bar elements with unknown cross-sectional areas A_i, depicted below in Figure 10. The three bars are joined together at a single pin joint O, with the other end of each bar pinned at the locations shown. For simplicity, bars 1 and 3 are assumed to have the same unknown cross-sectional area. The structure is loaded with a force P applied at an angle θ at the pinned joint.
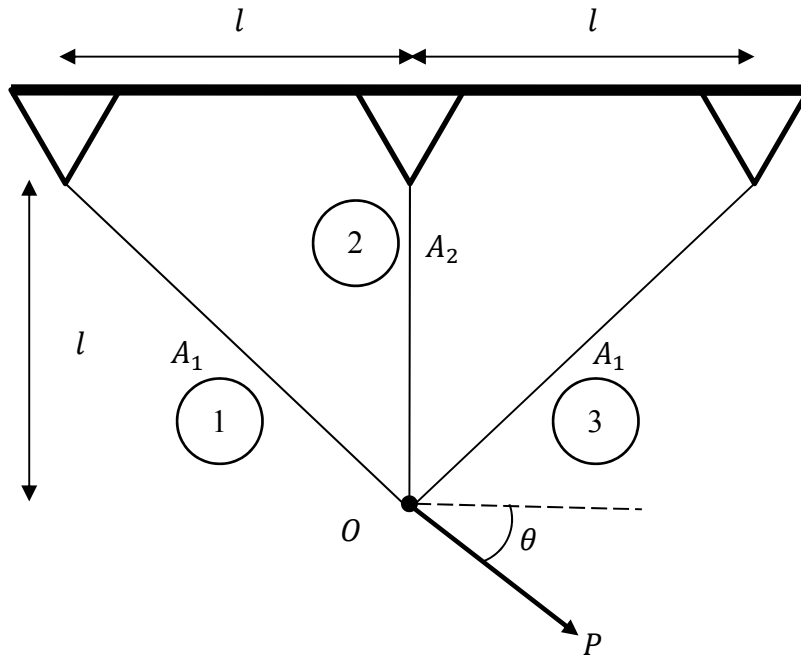
Figure 10: Probabilistic 3-bar truss problem

The objective of this optimization is to minimize the volume of the structure given constraints on the displacement of point O and the stress σ_i in each member. These stresses are given by Equations (89), (90), and (91) while the horizontal displacement of the joint u is given by Equation (92) and the vertical displacement v is given by Equation (93):

$$\sigma_1 = \frac{1}{\sqrt{2}}\left(\frac{P\cos\theta}{A_1} + \frac{P\sin\theta}{A_1 + \sqrt{2}A_2}\right) \tag{89}$$

$$\sigma_2 = \frac{\sqrt{2}P \sin\theta}{A_1 + \sqrt{2}A_2} \tag{90}$$

$$\sigma_3 = \frac{1}{\sqrt{2}}\left(\frac{P \sin\theta}{A_1 + \sqrt{2}A_2} - \frac{P \cos\theta}{A_1}\right) \tag{91}$$

$$u = \frac{\sqrt{2}lP \cos\theta}{A_1 E} \tag{92}$$

$$v = \frac{\sqrt{2}lP \sin\theta}{(A_1 + \sqrt{2}A_2)E} \tag{93}$$

Each stress and displacement is required to be less than a critical value, designated as $\sigma_{0,i}$, u_0, and v_0. The deterministic optimization problem can then be formulated by applying a factor of safety to each stress and displacement:

$$\min_{A} V = l\left(2\sqrt{2}A_1 + A_2\right)$$

Subject to:

$$g_1 = FS_1 \times \sigma_1 - \sigma_{0,1} \leq 0$$
$$g_2 = FS_2 \times \sigma_2 - \sigma_{0,2} \leq 0 \tag{94}$$
$$g_3 = -FS_3 \times \sigma_3 - \sigma_{0,1} \leq 0$$
$$g_4 = FS_4 \times u - u_0 \leq 0$$
$$g_5 = FS_5 \times v - v_0 \leq 0$$

In the probabilistic variant of the problem, the load, its direction, and the elastic modulus E of all bars is treated as normally distributed random parameters. As shown by Equations (89) through (93), all stresses and displacements are a function of at least 2 of these parameters. Therefore, each constraint g_i is replaced by a reliability constraint in an equivalent RBDO formulation, given below:

$$\min_{A} V = l\left(2\sqrt{2}A_1 + A_2\right)$$
$$\text{s.t. } g_i = \bar{\beta}_i - \beta_i \leq 0, \; i = 1{:}5 \tag{95}$$

Each reliability index is computed by solving a separate nested FORM problem:

$$\min_{U} \beta_i = \|U\|$$

$$\text{s.t. } g_i(U, A) = 0 \tag{96}$$

Where the $g_i(U, A)$ in Equation (96) are given by the constraints in the deterministic problem with a factor of safety of 1. In the following subsections, the sensitivity analysis procedure developed in Section 3.2 is demonstrated and validated by considering the stress in the first bar for a single iteration of the design. Then, several deterministic optimization and RBDO solutions are obtained and compared for this problem.

4.2.2 Sensitivity Analysis Demonstration and Validation

The procedure for computing the derivatives of the MPP with respect to each design variable, described in Section 3.2, can be validated by considering the reliability analysis of the stress in the first bar. This stress, given by Equation (89), is a function of two normally-distributed random parameters: the load P and its direction θ, as well as two design variables: A_1 and A_2. The values of the random and deterministic parameters used in the sensitivity validation are summarized below in Table 1.

Table 1: Summary of parameters for the sensitivity validation study

Parameter	Mean Value	Standard Deviation
A_1	1	N/A
A_2	$\sqrt{2}$	N/A
P	10	1
θ	45°	4.5°
$\sigma_{0,1}$	10	N/A

The first step of the sensitivity analysis is to solve the FORM problem for the current design. This was performed using the Hasofer-Lind iteration procedure [3] with a tolerance of 10^{-10} on the convergence of the reliability index and the MPP. The results of this analysis are listed in Table 2, including the Lagrange multiplier of the equality constraint.

48

Table 2: FORM solution for the reliability of the stress constraint for bar 1

Parameter	FORM Solution
β	4.5138
U_P^*	4.2071
U_θ^*	-1.6354
X_P^*	14.2072
X_θ^*	37.64°
λ	-1.3242

The next step of the analysis is to obtain all required first- and second-order derivatives. The equation for the stress is sufficiently simple that analytic expressions for these derivatives in terms of the standard normal random parameters can be obtained. For example, the derivatives of the stress constraint with respect to the design variables are:

$$\frac{\partial g_1}{\partial A_1} = \frac{1}{\sqrt{2}}(\mu_P + \sigma_P U_P)\left(-\frac{\cos(\mu_\theta + \sigma_\theta U_\theta)}{A_1^2} - \frac{\sin(\mu_\theta + \sigma_\theta U_\theta)}{\left(A_1 + \sqrt{2}A_2\right)^2}\right) \tag{97}$$

$$\frac{\partial g_1}{\partial A_2} = -\frac{(\mu_P + \sigma_P U_P)\sin(\mu_\theta + \sigma_\theta U_\theta)}{\left(A_1 + \sqrt{2}A_2\right)^2} \tag{98}$$

Similarly, the derivatives with respect to the standard normal random parameters are:

$$\frac{\partial g_1}{\partial U_P} = \frac{\sigma_P \cos\left(\mu_\theta + \sigma_\theta U_\theta\right)}{\sqrt{2}A_1} + \frac{\sigma_P \sin\left(\mu_\theta + \sigma_\theta U_\theta\right)}{\sqrt{2}A_1 + 2A_2} \tag{99}$$

$$\frac{\partial g_1}{\partial U_\theta} = \frac{\sigma_\theta\left(\mu_P + \sigma_P U_P\right)}{\sqrt{2}}\left(\frac{\cos\left(\mu_\theta + \sigma_\theta U_\theta\right)}{A_1 + \sqrt{2}A_2} - \frac{\sin\left(\mu_\theta + \sigma_\theta U_\theta\right)}{A_1}\right) \tag{100}$$

The second derivatives can then be computed. For simplicity, these derivatives are expressed in the vector and matrix forms required by Equation (86). It should be noted that the Hessian of the reliability index is not given below since it was previously derived in Equation (87).

$$H(g_1) = \begin{bmatrix} 0 & \frac{\sigma_P \sigma_\theta}{\sqrt{2}}\left(\frac{\cos(\mu_\theta+\sigma_\theta U_\theta)}{A_1+\sqrt{2}A_2} - \frac{\sin(\mu_\theta+\sigma_\theta U_\theta)}{A_1}\right) \\ \frac{\sigma_P \sigma_\theta}{\sqrt{2}}\left(\frac{\cos(\mu_\theta+\sigma_\theta U_\theta)}{A_1+\sqrt{2}A_2} - \frac{\sin(\mu_\theta+\sigma_\theta U_\theta)}{A_1}\right) & -\frac{\sigma_\theta^2(\mu_P+\sigma_P U_P)\left((A_1+\sqrt{2}A_2)\cos(\mu_\theta+\sigma_\theta U_\theta)+A_1\sin(\mu_\theta+\sigma_\theta U_\theta)\right)}{A_1(\sqrt{2}A_1+2A_2)} \end{bmatrix} \tag{101}$$

$$\left\{\frac{\partial^2 g}{\partial A_1 \partial U}\right\} = \frac{1}{\sqrt{2}}\left\{ \begin{array}{c} -\frac{\sigma_P \cos(\mu_\theta+\sigma_\theta U_\theta)}{A_1^2} - \frac{\sigma_P \sin(\mu_\theta+\sigma_\theta U_\theta)}{(A_1+\sqrt{2}A_2)^2} \\ \sigma_\theta(\mu_P + \sigma_P U_P)\left(-\frac{\cos(\mu_\theta+\sigma_\theta U_\theta)}{(A_1+\sqrt{2}A_2)^2} + \frac{\sin(\mu_\theta+\sigma_\theta U_\theta)}{A_1^2}\right) \end{array}\right\} \tag{102}$$

$$\left\{\frac{\partial^2 g}{\partial A_2 \partial U}\right\} = \frac{1}{(A_1+\sqrt{2}A_2)^2}\left\{ \begin{array}{c} -\sigma_P \sin(\mu_\theta + \sigma_\theta U_\theta) \\ -\sigma_\theta(\mu_P + \sigma_P U_P)\cos(\mu_\theta + \sigma_\theta U_\theta) \end{array}\right\} \tag{103}$$

Next, the systems of linear equations in terms of the unknown sensitivities of the MPP from Equation (86) can be formed. Since there are two design variables, there are two systems which need to be solved. Each system has a total of 3 equations since there are two random parameters and one equality constraint. After substituting in the data from Table 1 and the results of the FORM analysis from Table 2, these systems are:

$$\begin{bmatrix} 0.0291 & 0.1003 & 0.7039 \\ 0.1003 & 0.2741 & -0.2736 \\ 0.7039 & -0.2736 & 0 \end{bmatrix} \left\{\begin{array}{c} \frac{\partial U_P}{\partial A_1} \\ \frac{\partial U_\theta}{\partial A_1} \\ \frac{\partial \lambda}{\partial A_1} \end{array}\right\} = \left\{\begin{array}{c} -0.8050 \\ 0.5461 \\ 8.6366 \end{array}\right\} \tag{104}$$

$$\begin{bmatrix} 0.0291 & 0.1003 & 0.7039 \\ 0.1003 & 0.2741 & -0.2736 \\ 0.7039 & -0.2736 & 0 \end{bmatrix} \left\{\begin{array}{c} \frac{\partial U_P}{\partial A_2} \\ \frac{\partial U_\theta}{\partial A_2} \\ \frac{\partial \lambda}{\partial A_2} \end{array}\right\} = \left\{\begin{array}{c} -0.0899 \\ -0.1300 \\ 0.9641 \end{array}\right\} \tag{105}$$

Finally, these linear systems can be solved to obtain the sensitivities of the MPP with respect to the design variables. These results are summarized below in Table 3. For comparison, sensitivities obtained using the finite difference method with a step size of 10^{-8} and the complex step method with a step size of 10^{-20} are also included. These show very good agreement, thereby validating the sensitivities obtained using the Lagrange Multiplier Theorem approach.

Table 3: Comparison of MPP sensitivities found using three sensitivity analysis approaches

Sensitivity	Finite Difference	Complex Step	Lagrange Multiplier Theorem
$\frac{\partial U_P}{\partial A_1}$	11.0313226997505	11.0313226866851	11.0313226945795
$\frac{\partial U_P}{\partial A_2}$	1.02315595995606	1.02315571519760	1.02315571632285
$\frac{\partial U_\theta}{\partial A_1}$	-3.18715793756752	-3.18715791951515	-3.18715789920550
$\frac{\partial U_\theta}{\partial A_2}$	-0.891395535163042	-0.891395408720239	-0.891395405825368
$\frac{\partial \lambda}{\partial A_1}$	N/A	N/A	-1.14518334834460
$\frac{\partial \lambda}{\partial A_2}$	N/A	N/A	-0.0428924650647837

4.2.3 *Deterministic and Reliability-Based Design Optimization Results*

Having validated the sensitivity analysis procedure, this approach can then be used to solve a full RBDO problem. For comparison, deterministic optimization is also performed with an identical factor of safety applied to each constraint. Each analysis uses MATLAB's Sequential Quadratic Programming routine to control the design optimization process, with the problem configured using the parameters listed in Table 4.

Table 4: Summary of parameters for the deterministic optimization versus RBDO study

Parameter	Mean Value	Standard Deviation
A_1 (initial)	5in^2	N/A
A_2 (initial)	2in^2	N/A
A_{min}	0.001in^2	N/A
l	10in	N/A
P	30,000lb	4,500lb
θ	45°	7.5°
E	$3\times10^7\text{psi}$	$4.5\times10^6\text{psi}$
$\sigma_{0,1}$	5,000psi	N/A
$\sigma_{0,2}$	20,000psi	N/A
u_0	0.002in	N/A
v_0	0.002in	N/A

A total of 5 scenarios were considered. First, the deterministic optimization problem defined in Equation (94) was solved with a factor of safety of 1 applied to all constraints. Then, the deterministic optimization was repeated with a factor of safety of 1.5 applied to all constraints. Lastly, reliability-based design optimization was performed for three target reliability indices: $\bar{\beta} = 1.5$, 2.0, and 3.0. These analyses used the sensitivity analysis calculations described in the previous section for every iteration of the design optimization loop. Additionally, the probability of failure of each constraint as well as the system level probability of failure was then computed using Monte Carlo simulation with 10^6 random samples for all designs. The designs obtained using these 5 analyses are compared below in Table 5.

Table 5: Comparison of 3-bar truss designs using deterministic optimization and RBDO

Design Approach	A_1	A_2	Volume	Maximum Component P_f (MCS)	System P_f (MCS)
Det., $FS = 1$	5.000	1.768	159.1	0.4790	0.6103
Det., $FS = 1.5$	**7.500**	**2.652**	**238.6**	**0.0475**	**0.0478**
RBDO, $\bar{\beta} = 1.5$	7.257	0.1337	206.6	0.0618	0.1307
RBDO, $\bar{\beta} = 2.0$	**8.236**	**0.0010**	**233.1**	**0.0210**	**0.0399**
RBDO, $\bar{\beta} = 3.0$	10.795	0.0010	305.4	0.0012	0.0022

The advantages of using reliability-based design optimization are clearly demonstrated in Table 5. First, RBDO generated a novel design. While the deterministic solutions utilized all 3 bars, the reliability-based design approach found that the probabilistic performance of the structure was dominated by the diagonal bars. Therefore, the cross-sectional area of the middle bar was quickly driven to its lower bound as the target reliability index increased. More interestingly, RBDO found a design which was not only more reliable, but also more optimal with respect to volume than the deterministic designs. Specifically, comparison of the deterministic solution for $FS = 1.5$ and the RBDO solution for $\bar{\beta} = 2.0$ shows that RBDO produced a design which had a 2% lower volume but more importantly reduced the system probability of failure from 0.0478 to 0.0399, a 16% reduction. Note that RBDO only constraints the component probability of failure, which was reduced by 56%.

To explore the cause of this, the probability of failure for each constraint obtained using the Monte Carlo simulation above can be compared, as shown in Table 6. In the deterministic solution, the design was constrained by the stress in the first bar and the horizontal displacement of the joint. The likelihood of each constraint being violated was vastly different however, with the stress constraint having a probability of failure which was 2 orders of magnitude less than that of the displacement constraint. This indicates that while a factor of safety of 1.5 was sufficient for the horizontal displacement, this was too conservative for the stress constraint. The result is a design which is may be optimal with respect to deterministic criteria, but is suboptimal with respect to the stochastic performance of the system. In the RBDO solution, both of the displacement constraints were active with a FORM reliability index of $\beta = 2$ while all of the stress constraints were inactive. In comparison to the deterministic solution, the probability is more equally distributed between the constraints. In particular, the excess safety of the stress in bar 1 has been reduced in order to increase the reliability of other constraints. This highlights a key advantage of RBDO: excess component safety is redistributed in order to improve the overall system reliability.

Table 6: Comparison of constraint behavior for similar deterministic and RBDO designs

Design Approach	Probability of Constraint Violation (Bold Indicates Active Constraint)				
	$P(\sigma_1 > \sigma_{0,1})$	$P(\sigma_2 > \sigma_{0,2})$	$P(\sigma_3 < -\sigma_{0,1})$	$P(u > u_0)$	$P(v > v_0)$
Det., $FS = 1.5$	**0.000432**	<0.000001	<0.000001	**0.0475**	0.000758
RBDO, $\bar{\beta} = 2.0$	0.0050	<0.000001	<0.000001	**0.0210**	**0.0204**

It should be noted that two brief studies were performed to further investigate these results. First, a study was performed to verify that the deterministic approach was not converging to a different, less optimal minima. Specifically, the deterministic optimization procedure was repeated with the initial design point chosen to be the $\bar{\beta} = 2.0$ solution: $A = [8.236, 0.0010]$. Despite starting at the RBDO solution, the deterministic optimization procedure still converged to the design $A = [7.500, 2.652]$. Second, a brief study was performed to investigate why the probabilities of failure of the active constraints were lower than expected. Specifically, Equation (20) predicts a probability of failure for $\beta = 2.0$ of 0.0228, while the MCS solutions were 0.0210 and 0.0204. As shown by Figure 11, the probability distribution for the horizontal displacement

constraint obtained using Monte-Carlo simulation is not Gaussian. Since the probability density function of a linear sum of normally distributed parameters is exactly normal, the non-Gaussian distribution indicates that the failure surface is actually non-linear. Equation (20) assumes that the constraint is exactly linear with respect to the random parameters in the U-space however, based on a linear expansion using the slope at the MPP. Thus, any non-linearity in the failure surface will result in an error in the estimate of the probability of failure.
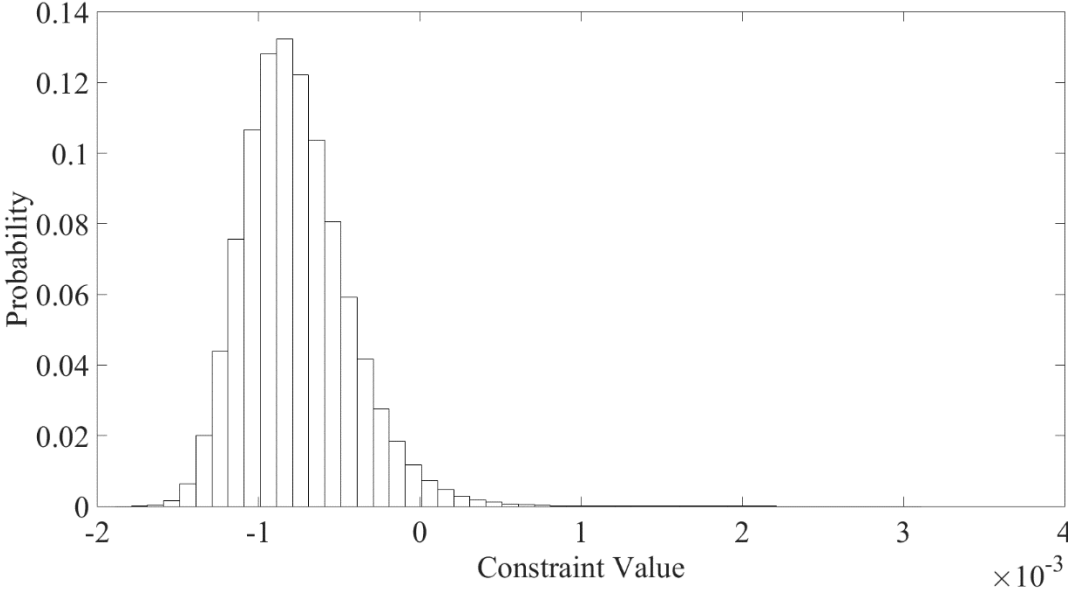

Figure 11: MCS distribution for the $\bar{\beta} = 2.0$ horizontal displacement constraint

4.3 Topology Optimization of a Benchmark Problem

4.3.1 Problem Statement

Having demonstrated that the sensitivity analysis procedure derived from the Lagrange Multiplier Theorem can be used to solve a double-loop reliability-based design optimization problem, this approach can be extended to reliability-based topology optimization. Fundamentally, the only difference between RBTO and RBDO is the size and scope of the problem. Because of the potentially thousands or more design variables however, RBTO requires every step of the optimization to be as computationally efficient as possible. Since the proposed procedure provides analytic sensitivities by solving a linear system of equations but without requiring additional solutions of the nested FORM problem, it has potential applications in the field of reliability-based topology optimization. These applications will be investigated by first solving the RBTO benchmark problem developed by Rozvany and Maute [1]. Since the benchmark problem has an analytic solution, it will also serve to validate the RBTO routine developed for this research.

The benchmark problem consists of a rectangular region of material of unknown topology. The bottom edge is fixed while a deterministic vertical load and random horizontal load are applied at the midpoint of the top edge. The horizontal load is normally distributed with a mean of 0 and a standard deviation σ_H. This configuration is depicted in Figure 9. Rozvany and Maute developed an analytic solution for the minimum weight topology which satisfies a constraint on the probability that the compliance does not exceed a critical value. This derivation did not require any assumptions about the shape of the structure. A brief summary of the results of their derivation follows.

Using an optimal layout theory approach, the minimum volume topology satisfying a constraint on the reliability of the compliance is a two bar truss which is symmetric about the vertical axis, depicted in Figure 12. The angle of the truss is related to the standard deviation of the horizontal force, σ_H, and the target reliability level, $\bar{\beta}$, through a multi-step process. First, a critical value of the horizontal load, H_0, can be defined:

$$F_{H,0} = \sigma_H \Phi^{-1} \left(\frac{1 + \Phi(\bar{\beta})}{2} \right) \tag{106}$$

55

The optimal angle α can then be expressed in terms of the angle between H_0 and the deterministic vertical load, F_V. This angle, B, is given by:

$$B = \tan^{-1} \frac{F_{H,0}}{F_V} \tag{107}$$

The optimal angle is then computed using:

$$\alpha = \tan^{-1} \sqrt{\left(\sqrt{\tan^4 B + 8\tan^2 B} - \tan^2 B\right)/4} \tag{108}$$
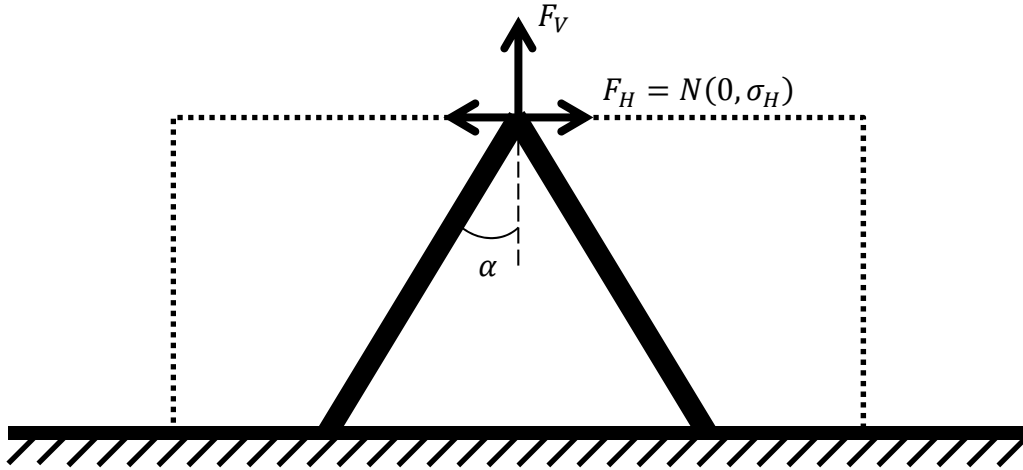
Figure 12: Analytic solution for the RBTO benchmark problem [1]

An equivalent RBTO problem can then be defined using the SIMP method for the topology optimization, discussed in Section 2.5.1, with the linear density filtering approach discussed in Section 2.5.3. This requires the discretization of the region into a finite element mesh. Since the design region is rectangular, it is convenient to use a structured grid of linear quadrilateral plane-stress elements. For simplicity, each element is assumed to be identically rectangular with a width of Δx, a height of Δy, and a thickness t. Thus, the structured mesh will be a grid of rows with n_x elements and columns with n_y elements, each with an associated pseudo-density ρ_i. This discretization is visualized in Figure 13.
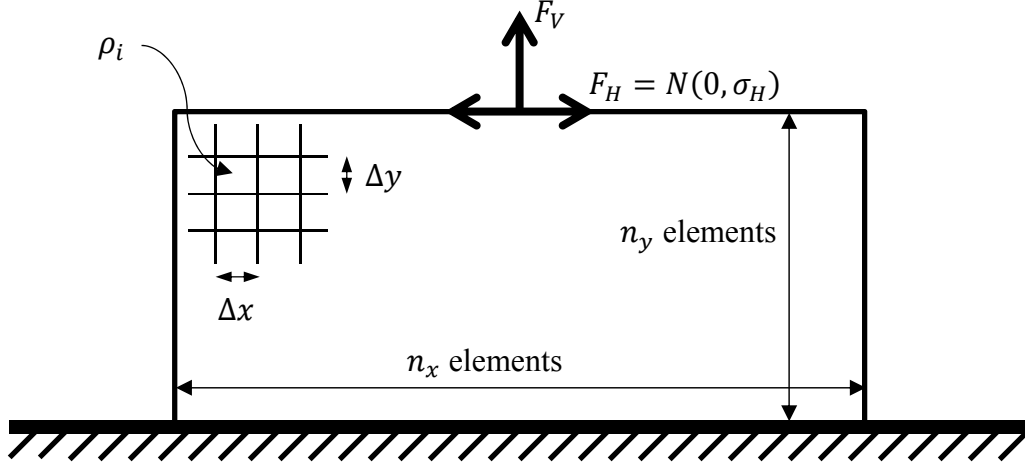
Figure 13: Discretization of the RBDO benchmark problem

The objective of the equivalent RBTO formulation is to minimize the weight of a structure of unknown topology. Since each element has identical geometry, this is equivalent to the minimization of the sum of the pseudo-densities. The topology is constrained to have a minimum reliability level $\bar{\beta}$ against the compliance exceeding a critical value, $\bar{C}$. Therefore, the RBTO problem is given by:

$$\min_{\boldsymbol{\rho}} W = \sum_{i=1}^{m} \rho_i$$
$$\text{s.t.} \, g = \bar{\beta} - \beta \leq 0 \tag{109}$$

Where β is computed by solving a nested FORM problem each iteration:

$$\min_{\boldsymbol{U}} \beta_i = \|\boldsymbol{U}\|$$
$$\text{s.t.} \, \boldsymbol{D}^T \boldsymbol{K} \boldsymbol{D} - \bar{C} = 0 \tag{110}$$
$$\boldsymbol{K}\boldsymbol{D} = \boldsymbol{R}$$

In the present study, RBTO is implemented in MATLAB using the Method of Moving Asymptotes [14] to control the design problem and MATLAB's Interior-Point algorithm [13] to solve the nested FORM problem each iteration. Both methods were chosen since their computational costs scale well for large numbers of design variables. Additionally, the Method of Moving Asymptotes is already a well-established approach for reliability-based topology

optimization, having been employed in several studies previously [1, 25, 26, 28]. The design problem is formulated using the unit-less values listed in Table 7. In general, these values are identical to those used by Rozvany and Maute [1], with the exception that the length to height ratio of the design region is reduced from 5:1 to 3:1. It is not expected that this will affect the results since Rozvany and Maute's solution has a considerable amount of void on each side of the truss structure. This results in a new mesh of 144×48 square plane-stress elements, requiring a total of 6,912 pseudo-densities. It should also be noted that similar to Rozvany and Maute, a Poisson's ratio of $v = 0$ is used to approximate the behavior of the bars in the analytic solution.

It is important to note that the nested problem requires multiple solutions of the finite element analysis, though the computational cost is offset slightly by the fact that the stiffness matrix does not need to be reassembled each iteration of the nested problem since the only random parameter is a force. To avoid repeated solutions of the FORM problem, and therefore minimize the number of finite element solutions required, the sensitivity analysis procedure proposed in Section 3.2 is used. Its specific implementation is discussed in the following section.

Table 7: Summary of unit-less parameters for the RBTO benchmark problem

Parameter	Mean Value	Standard Deviation
F_V	100	N/A
F_H	0	100/3
n_x	144	N/A
n_y	48	N/A
Δx	1/48	N/A
Δy	1/48	N/A
t	1	N/A
$\bar{r}$	1/24	N/A
$\bar{C}$	10^6	N/A
E_0	10	N/A
v	0	N/A
p	3	N/A

4.3.2 Sensitivity Analysis

Topology optimization problems can have thousands or more design variables. This necessitates analytic sensitivities, since finite differencing even a simple constraint can become computationally expensive if it has to be repeated many times over. In the present analysis, the sensitivities of the reliability index to the pseudo-densities are computed using Equation (28). This requires the sensitivities of the MPP to the pseudo-densities, which are obtained using the procedure described in Section 3.2. For efficiency, this procedure requires analytic expressions for several derivatives and second-derivatives. While the number of design variables suggests that obtaining expressions for each would be a futile effort, the coupling of the pseudo-densities with the finite element problem simplifies this considerably. Specifically, a given derivative can be expressed using a single equation for all pseudo-densities and/or the random loads by taking advantage of the direct and adjoint approaches discussed in Section 2.4.3.

The simplest example of this is the derivative of the compliance with respect to each pseudo-density. A complete derivation of this is presented below in order to demonstrate how these approaches can be applied. Starting with the definition of compliance from Equation (56):

$$C = \boldsymbol{D}^T \boldsymbol{K} \boldsymbol{D} \tag{111}$$

The derivative with respect to a given pseudo-density is given by:

$$\frac{\partial C}{\partial \rho_i} = \left(\frac{\partial \boldsymbol{D}}{\partial \rho_i}\right)^T \boldsymbol{R} \tag{112}$$

Equation (112) requires the derivative of the global displacement matrix with respect to a given pseudo-density. The direct method can be used to obtain these derivatives, which are given by Equation (34). After substituting this expression in and considering that the force vector $\boldsymbol{R}$ is not a function of the pseudo-densities, the following expression is obtained:

$$\frac{\partial C}{\partial \rho_i} = \left(-\boldsymbol{K}^{-1} \frac{\partial \boldsymbol{K}}{\partial \rho_i} \boldsymbol{D}\right)^T \boldsymbol{R} \tag{113}$$

After further simplification using various matrix properties, this reduces to:

59

$$\frac{\partial C}{\partial \rho_i} = -\boldsymbol{D}^T \left(\frac{\partial \boldsymbol{K}}{\partial \rho_i}\right)^T \boldsymbol{D} \tag{114}$$

As a last step, this can be further simplified by considering that the pseudo-density only multiplies a single element-level stiffness matrix, as shown in Equation (58). Therefore, the derivative of the global stiffness matrix reduces to the derivative of the element level stiffness matrix:

$$\frac{\partial \boldsymbol{K}}{\partial \rho_i} = p\rho_i^{P-1}\boldsymbol{L}_i^T \boldsymbol{k_0} \boldsymbol{L}_i \tag{115}$$

Where $\boldsymbol{L}_i$ is the locator matrix for element i. After substituting Equation (115) into Equation (114), the expression for the derivative of the compliance with respect to any pseudo-density is obtained:

$$\frac{\partial C}{\partial \rho_i} = -p\rho_i^{P-1}\boldsymbol{d}_i^T \boldsymbol{k_0} \boldsymbol{d}_i \tag{116}$$

Where $\boldsymbol{d}_i$ is a column matrix of the global displacements for the nodes of element i.

The remaining derivatives can be computed using similar procedures. A full derivation of these partial derivatives is included in Appendix A. The resulting expressions which are applicable for all pseudo-densities are given below. First, the derivative of the compliance with respect to each random load is given by:

$$\frac{\partial C}{\partial F_i} = 2d_{F_j} \tag{117}$$

Where d_{F_j} is the displacement of the nodal degree of freedom matching the location and direction of the random force F_j. Using the adjoint method, the mixed second derivative for any pseudo-density and random force combination reduces to the expression below:

$$\frac{\partial^2 C}{\partial \rho_i \partial F_j} = -2p\rho_i^{p-1}\{\boldsymbol{K}_j^{-1}\}^T \boldsymbol{k_0} \boldsymbol{d}_i \tag{118}$$

Where $\{K_j^{-1}\}$ is a column vector analogous to the element-level nodal displacement matrix d_i, except instead containing the terms from the jth column of K^{-1}. Lastly, the Hessian of the compliance with respect to any pair of random nodal forces is given by:

$$\frac{\partial^2 C}{\partial F_k \partial F_j} = 2K_{jk}^{-1} \tag{119}$$

Where K_{jk}^{-1} is the term of the inverse stiffness matrix whose global indices match those of the two random forces. It should be noted that several of these equations require the inversion of the stiffness matrix, which is typically a computationally expensive procedure. It should be possible to derive equivalent expressions for these derivatives which do not require this inversion, but instead use the solution of a system of linear equations.

The sensitivity analysis was validated by comparing the sensitivities of the reliability index obtained using the Lagrange Multiplier Theorem approach to those obtained using finite difference derivatives for a range of step sizes. To save on computation costs, a coarser problem with 30 horizontal elements and 10 vertical elements was used, giving a total of 300 design variables. Additionally, only the first 5 iterations of the RBTO routine were computed. Otherwise, the problem formulation was identical to the full scale problem.

The maximum difference between the analytic and finite difference sensitivities for each iteration and step size is summarized in Table 8. As expected, these results were step-size dependent, with a step of 10^{-7} providing the best agreement. At this step size, the derivatives agreed to within 1.52×10^{-6}. The maximum difference occurred in the second iteration, where the values of the analytic derivatives ranged from -1.022×10^{-5} in the void regions to -1.044 at the applied load. While the difference is within an order of magnitude of the smallest sensitivity, this is acceptable since the maximum sensitivity is 6 orders of magnitude greater. Since the sensitivity of the void regions is orders of magnitude less than the sensitivity elsewhere, these densities should still be driven to their lower bounds even if there is a 10% error in their values. It is also important to note that some discrepancy is expected as a result of the step size and subtractive errors associated with finite difference derivatives, discussed in Section 2.4.1. Overall, the good agreement between the analytic and finite difference sensitivities validates the analytic sensitivity analysis procedure.

Table 8: Comparison of finite difference and analytic sensitivities for the RBTO benchmark

Step Size	Maximum of $\left\| \frac{\partial \beta}{\partial \rho_{i_a}} - \frac{\partial \beta}{\partial \rho_{i_{fin}}} \right\|$ for each iteration					
	0	1	2	3	4	5
10^{-5}	7.30×10^{-5}	1.71×10^{-5}	6.83×10^{-6}	1.33×10^{-5}	1.34×10^{-5}	3.21×10^{-5}
10^{-6}	7.29×10^{-6}	1.69×10^{-6}	7.89×10^{-7}	1.41×10^{-6}	1.32×10^{-6}	3.18×10^{-6}
$\mathbf{10^{-7}}$	$\mathbf{7.19 \times 10^{-7}}$	$\mathbf{5.82 \times 10^{-7}}$	$\mathbf{1.52 \times 10^{-6}}$	$\mathbf{9.57 \times 10^{-7}}$	$\mathbf{1.52 \times 10^{-6}}$	$\mathbf{8.91 \times 10^{-7}}$
10^{-8}	1.36×10^{-6}	8.44×10^{-6}	1.21×10^{-5}	1.21×10^{-5}	9.43×10^{-6}	7.59×10^{-6}
10^{-9}	1.66×10^{-5}	5.90×10^{-5}	0.000126	0.000114	0.000116	0.000159
10^{-10}	0.000140	0.000801	0.001112	0.001100	0.001044	0.000984
10^{-11}	0.001678	0.007563	0.023594	0.013487	0.015372	0.008863
10^{-12}	0.016625	0.075536	0.209780	0.104333	0.107158	0.129979

To further explore the differences in the analytic and finite difference sensitivities, the variation in the sensitivities across the topology for each approach can be plotted. Since the sensitivities vary by multiple orders of magnitude across the topology, a lognormal scale is used. The analytic sensitivities for the 4th iteration of the design are mapped in Figure 14 while the finite difference sensitivities computed with a step size of 10^{-7} are mapped in Figure 15. Comparison of the two mappings shows very good agreement between the sensitivities across the topology. The only significant differences are in the upper left and right corners, however these regions also have the minimum values of the sensitivities. Therefore, the error in these regions is not expected to influence the design. Lastly, a map of the absolute differences in the sensitivities computed using the proposed analytic method and the finite difference method with a step size of 10^{-7} can be constructed, depicted in Figure 16. This shows that the errors were distributed randomly across the structure, with no correlation to whether the topology was a void or dense region.

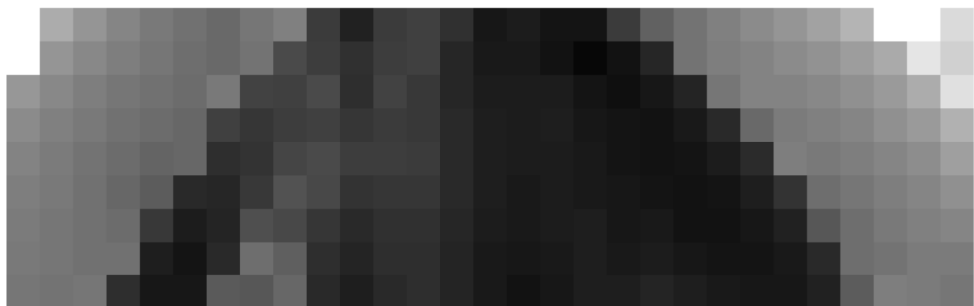**Figure 14:** 4th Iteration analytic sensitivities on a lognormal scale

Figure 15: 4th Iteration finite difference sensitivities for $h = 10^{-7}$ on a lognormal scale
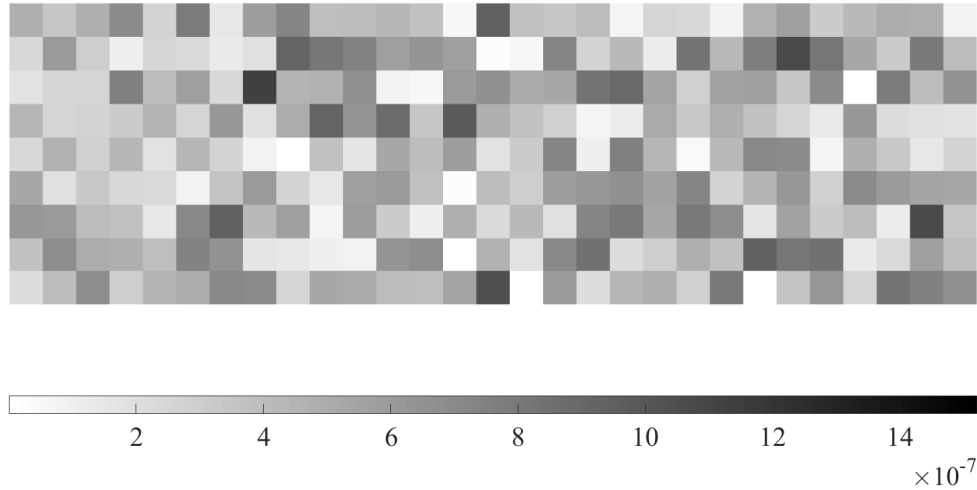
Figure 16: Absolute difference between the analytic and finite difference sensitivities

4.3.3 Solution versus Reliability Index

RBTO was initially performed using reduced-scale problems to verify that the routine was converging towards the analytic solution. Unexpectedly, these tests converged to asymmetric designs, such as the topology depicted in Figure 17. This topology was obtained after 1,000 iterations of a 60×20 grid of elements with a target reliability index of $\bar{\beta} = 3$. The optimization was terminated at this point because the number of iterations reached its cap, however as shown by Figure 18 the changes in the objective function were minimal at this point. Similarly, the topology had converged to a design which satisfied the reliability constraint, as shown by Figure 19. While these convergence histories appear to be smooth, in reality the optimization procedure was oscillating. This can be seen from the norm of the residuals of the KKT conditions plotted in Figure 20.
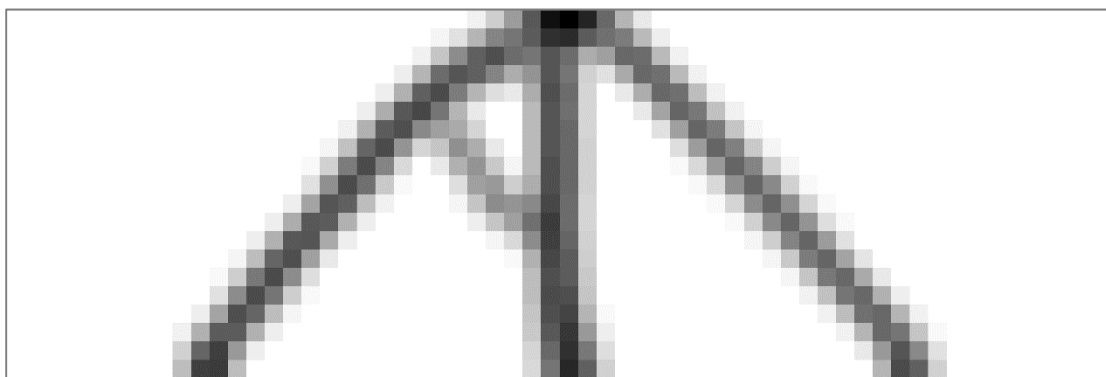

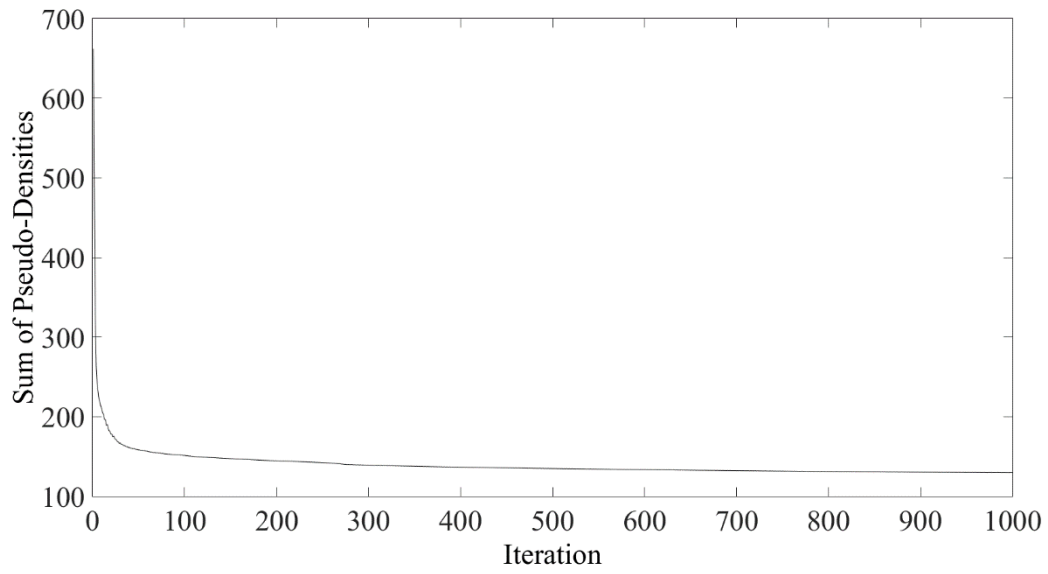**Figure 17:** RBTO benchmark topology after 1,000 iterations with $\bar{\beta} = 3$

64

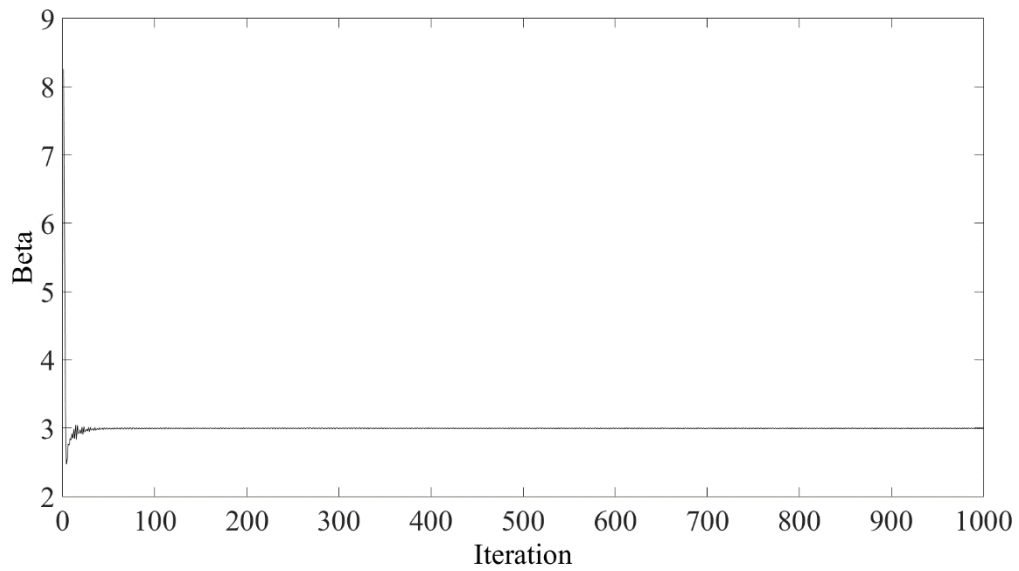**Figure 18:** Convergence of the weight of the RBTO benchmark topology with $\bar{\beta} = 3$

Figure 19: Convergence of β for the RBTO benchmark topology with $\bar{\beta} = 3$
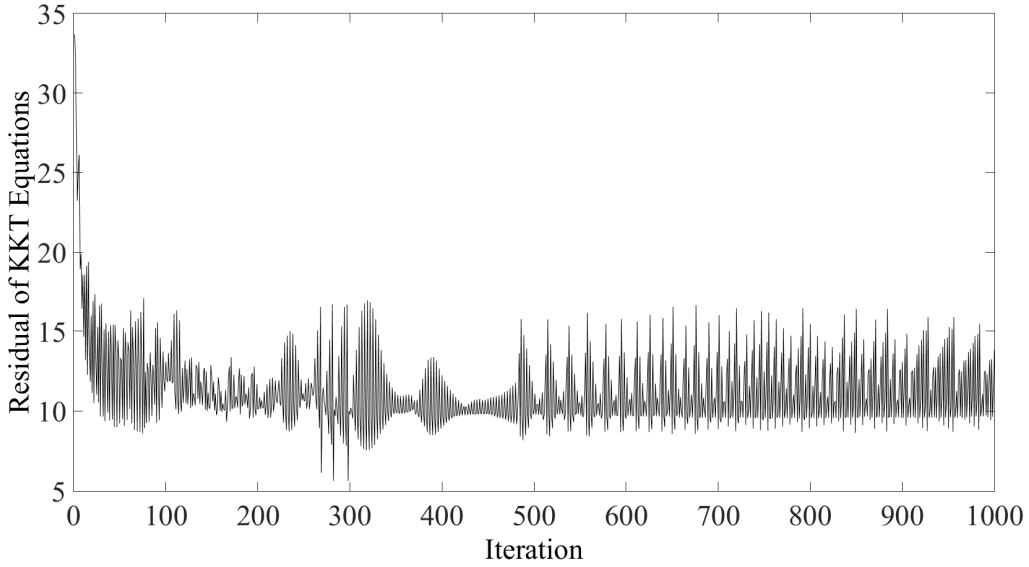
Figure 20: Norm of the residual of the KKT Equations for the RBTO benchmark with $\bar{\beta} = 3$

Upon further investigation, it was found that this oscillation is related to the symmetry of the problem. Specifically, the loads and boundary conditions are symmetric about the vertical midplane of the design region, as shown in Figure 13. This includes the random horizontal force, which is normally distributed with zero mean, implying that a given realization is as equally probable as the equivalent with the opposite sign. Furthermore, this implies that compliance is at a minimum at the mean value and will always increase regardless of whether the random force has a positive or negative realization. Thus, there are two MPPs in any given iteration, one associated with a positive value of the force and another associated with a negative value. For a topology which is symmetric about the midplane, such as the analytic solution or the evenly distributed densities used for the initial topology, these MPPs will hold the exact same value but opposite signs. There are also two distinct failure regions: one with all force realizations less than the negative MPP and the other with all force realizations greater than the positive MPP. Therefore, linear expansion about one MPP only captures one failure region, resulting in the reliability of the system being significantly underestimated. For a fully symmetric topology, the probability of failure computed using Equation (20) is exactly half that of the actual value.

66

The MPP is obtained by solving a minimization problem, as detailed in Section 2.4.4. Therefore, only a single MPP is computed each time the FORM subroutine is called. This occurs once per iteration of RBTO benchmark problem formulation, given by Equation (109), so only one of the two MPPs is considered when formulating the next iteration of the design. From the perspective of the optimizer, this is equivalent to optimizing for a one-sided random load in a given iteration, which encourages an asymmetric design. The challenge is that in the next iteration, the MPP may switch, as shown in Figure 21. If too big of a step was taken however, the topology may not return to a symmetric design after the new iteration, particularly if the move-limits are adjusted between iterations. This behavior can be seen by plotting the first 10 iterations of the asymmetric solution, depicted in Figure 22. Thus, the topology gradually moves towards an asymmetric solution which satisfies the reliability constraint, but is not optimal.
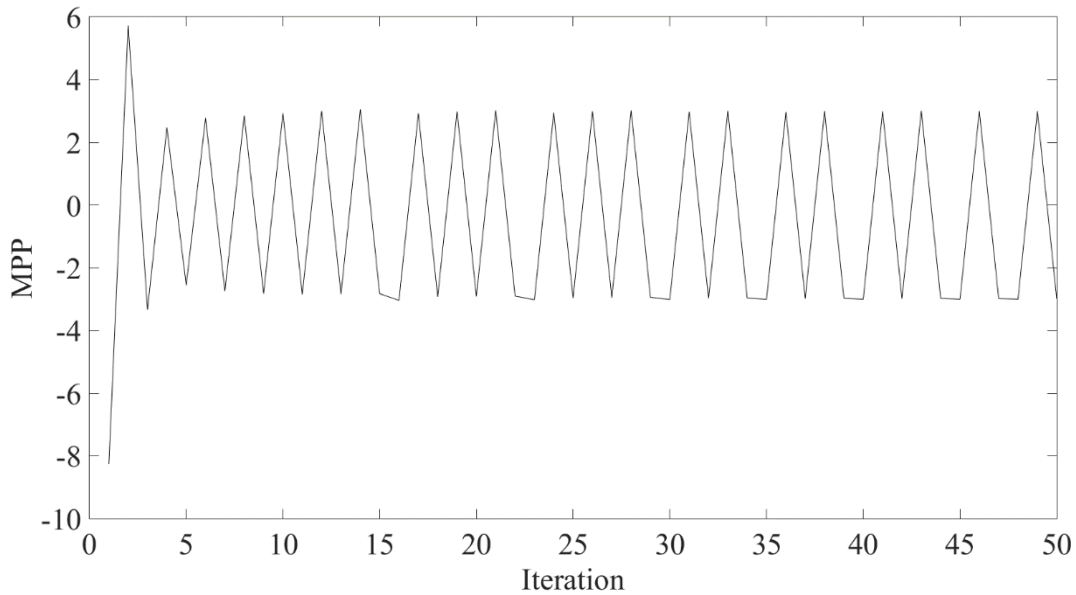
Figure 21: Oscillations of the MPP for the RBTO benchmark with $\bar{\beta} = 3$

67
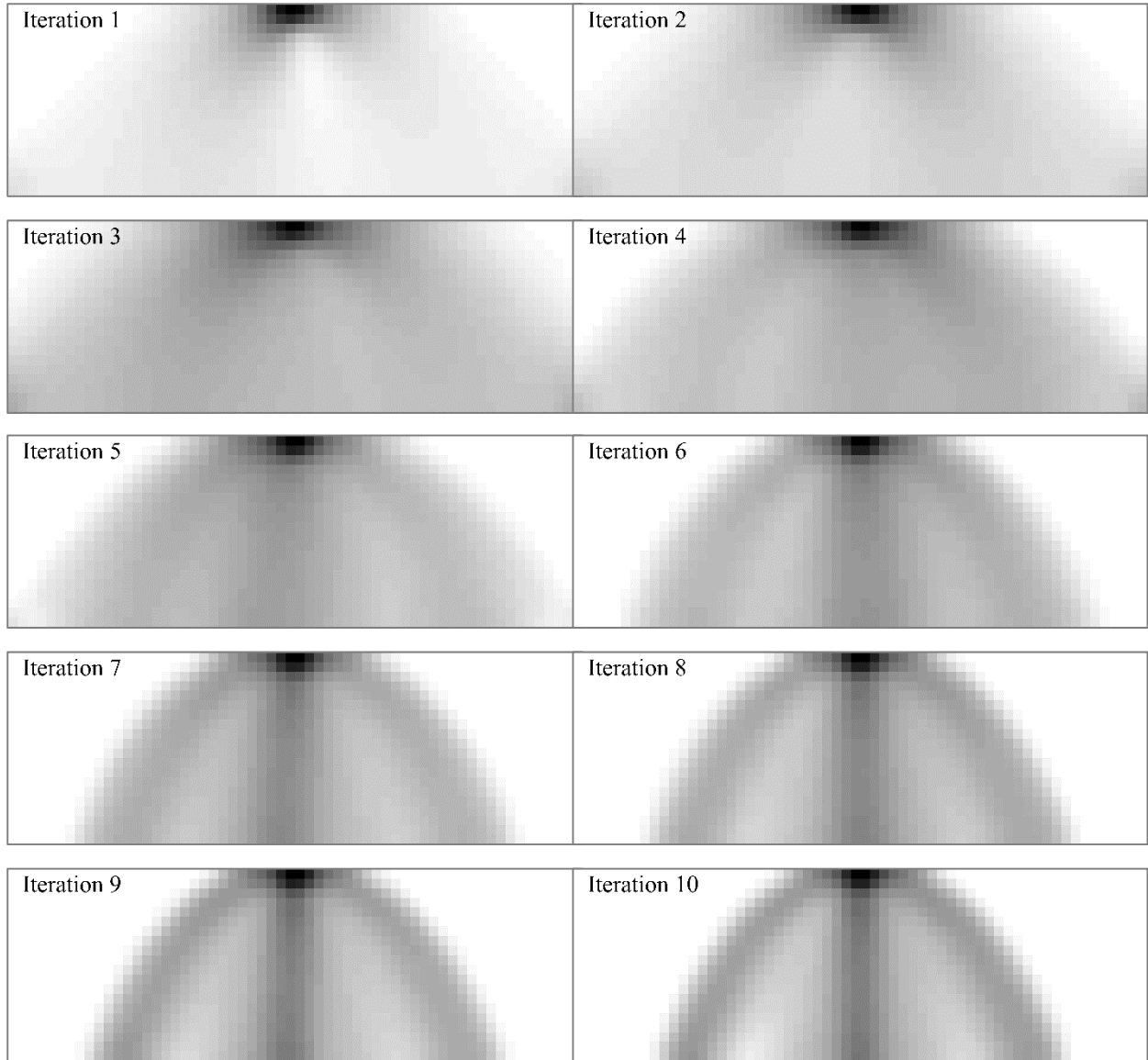
Figure 22: First 10 iterations of the asymmetric RBTO solution

To avoid this behavior, the problem formulation was modified to consider both MPPs by adding a second reliability constraint. Both reliability constraints use the formulation from Equation (110), however one constraint considers the reliability with respect to the positive MPP while the other considers the reliability with respect to the negative MPP. The new formulation for the design loop is given in Equation (120):

$$\min_{\rho} W = \sum_{i=1}^{m} \rho_i$$

$$\text{Subject to:}$$

$$g_1 = \bar{\beta} - \beta^+ \leq 0 \tag{120}$$
$$g_2 = \bar{\beta} - \beta^- \leq 0$$

Where β^+ is the reliability index for the positive MPP and β^- is the reliability index for the negative MPP. To ensure that the correct MPP was found, the initial analysis point for each FORM problem was chosen to be equal to $\bar{\beta}$ standard deviations away from the mean value of H in either the positive or negative direction. In general, this formulation encourages symmetry, however it requires twice the computational effort. It is important to note that the sensitivity analysis procedure detailed in Section 4.3.2 is still applicable, though it needs to be repeated twice. As shown in Figure 23, after modification a topology which was qualitatively similar to the analytic solution was obtained. Since the analytic solution assumes a continuum while RBTO is discretized, it was expected that there would be small discrepancies between the two results. Furthermore, repeating the analysis with the previously found asymmetric solution (Figure 17) as the initial design point with $\bar{\beta} = 3$ converges to the expected topology, as shown by Figure 24. This indicates that the asymmetric solution was not actually a local minimum. This also demonstrates that the modified approach is robust, since it was able to converge without assuming that the initial design is symmetric.
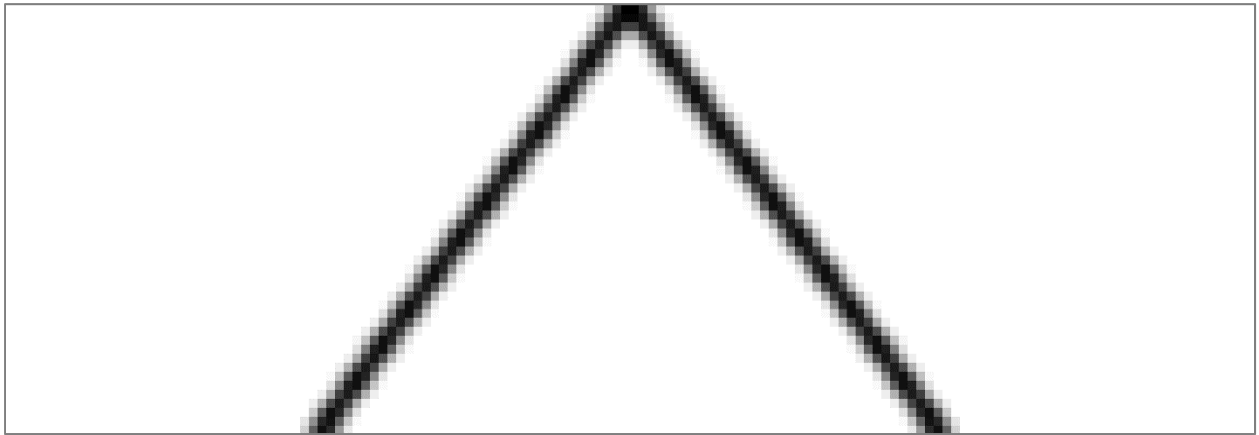
Figure 23: Topology obtained for $\bar{\beta} = 3$ after reformulation to consider multiple MPPs
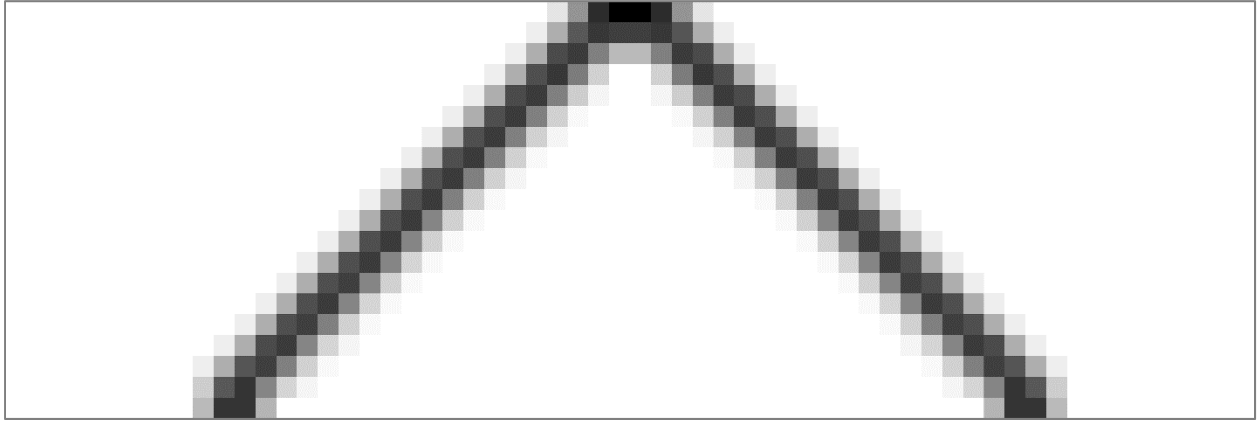
Figure 24: Topology found using the modified approach starting at the asymmetric solution

As shown by Equations (106) through (108), the truss angle is expected to increase as the required reliability level increases. To verify this behavior, the modified RBTO benchmark problem was solved for $\bar{\beta} = 1, 2, 3, 4, 5$, and 6. The resulting topologies are depicted in Figure 25. Qualitatively, these topologies demonstrated the expected behavior, with the trusses becoming wider as $\bar{\beta}$ is increased.
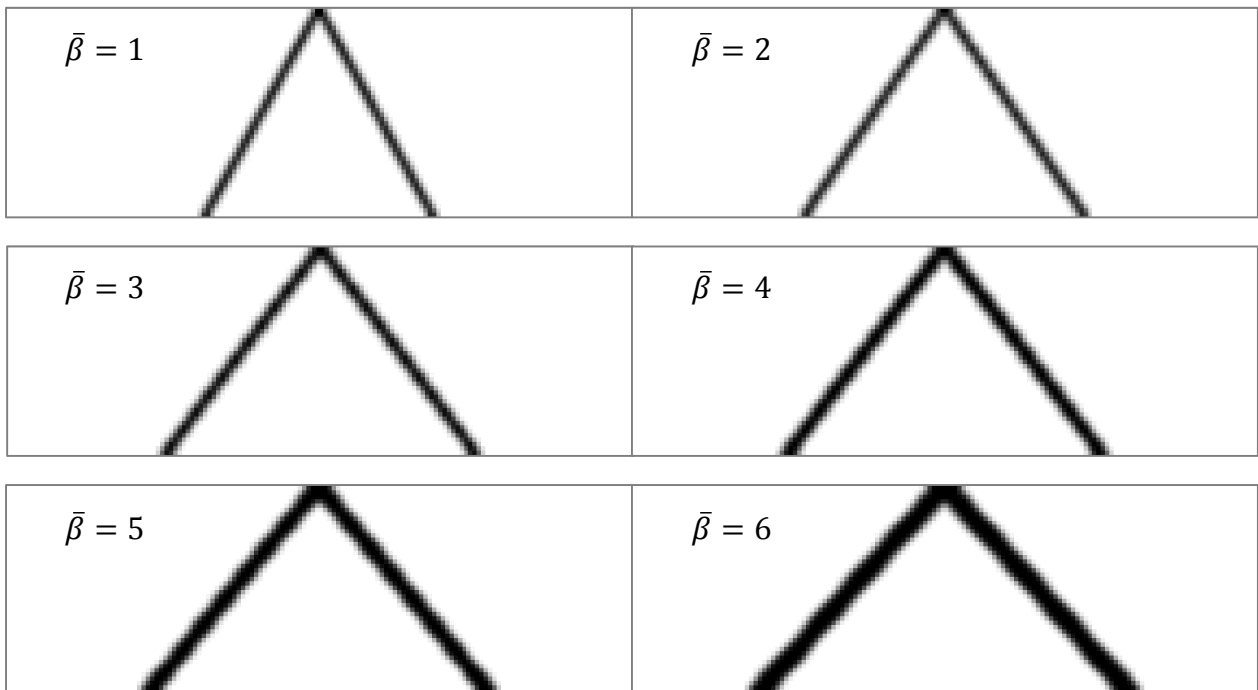
Figure 25: Solutions to the RBTO benchmark problem for different reliability indices
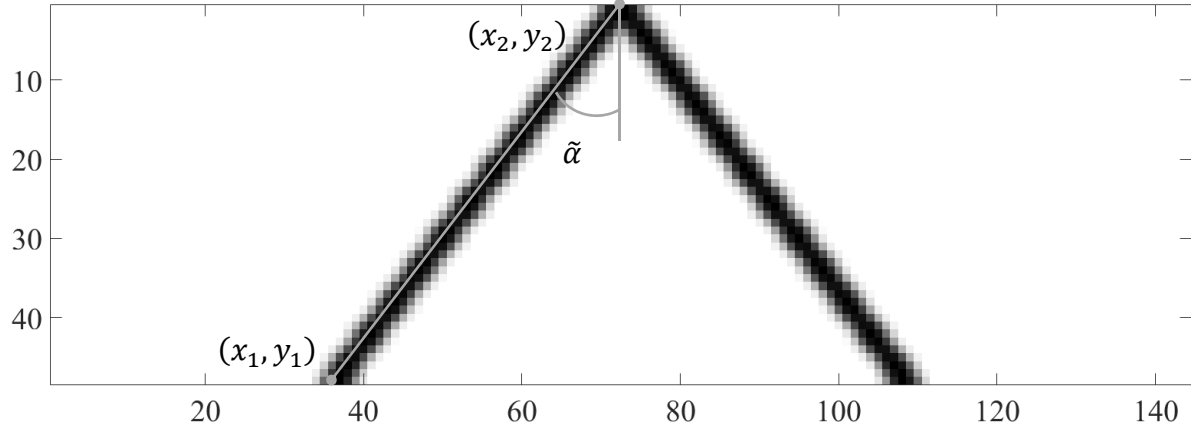
Figure 26: Estimation of $\tilde{\alpha}$ for the RBTO benchmark problem with $\bar{\beta} = 4$

To verify the present results , a rough estimate of the angle from the vertical, $\tilde{\alpha}$, was obtained for each topology by drawing an approximate line of best fit on one of the truss members. An example of this for the $\bar{\beta} = 4$ topology is depicted in Figure 26. Two points, (x_1, y_1) and (x_2, y_2), were then used to approximate the slope of the line. The slope of the line can then be related to the truss angle $\tilde{\alpha}$ using Equation (121):

$$\tilde{\alpha} \approx \tan^{-1} \frac{x_2 - x_1}{y_2 - y_1} \tag{121}$$

The approximate angles found for each topology using Equation (121) are compared to the exact solutions obtained using Equations (106) through (108) in Table 9. As expected, the approximated angles $\tilde{\alpha}$ differ by a small amount from the exact solutions α. This is likely a result of the discretization as well as errors resulting from the choice of the fitting points. The general trend versus $\bar{\beta}$ correlates well with the exact solution however, thus validating the RBTO routine.

71

Table 9: RBTO benchmark reliability indices versus approximated truss angle $\tilde{\alpha}$

$\bar{\beta}$	P_f	H_0	α	$\tilde{\alpha}$
1.0	0.1587	46.9870	27.9534	29.15692
2.0	0.0228	75.9202	32.6871	34.38034
3.0	0.00135	106.8385	35.8531	36.77328
4.0	$3.17 \times 10\text{-}5$	138.7035	38.0226	37.35959
5.0	$2.87 \times 10\text{-}7$	171.0673	39.5523	39.74620
6.0	$9.87 \times 10\text{-}10$	203.7188	40.6596	41.74805

As a last step, Monte Carlo simulation was performed for the $\tilde{\beta} = 2$ solution to verify that the topology has the required reliability level. Using Equation (20), the expected probability of failure for one of the MPPs is 0.0228. Therefore, the expected system probability of failure is 0.0455. Using Monte Carlo simulation of the optimal topology with 10,000 samples, a probability of failure of 0.0452 was obtained, confirming that the RBTO procedure converged to a reliable topology. The compliance constraint as a function of the horizontal load can also be mapped, as shown in Figure 27. This confirmed the symmetric behavior of the compliance and the presence of multiple MPPs. The probability density function for the compliance constraint can also be approximated using the samples, as shown in Figure 28.

Overall, this analysis proved that the proposed sensitivity analysis procedure can be incorporated into reliability-based topology optimization problems. In the process, an interesting behavior related to the symmetry of the reliability problem was observed. This required a modification to the problem formulation so that the reliability with respect to multiple MPPs was considered. It is important to note that the modification required knowledge of the existence of multiple MPPs *a priori* however. Therefore, this procedure is not universally applicable.
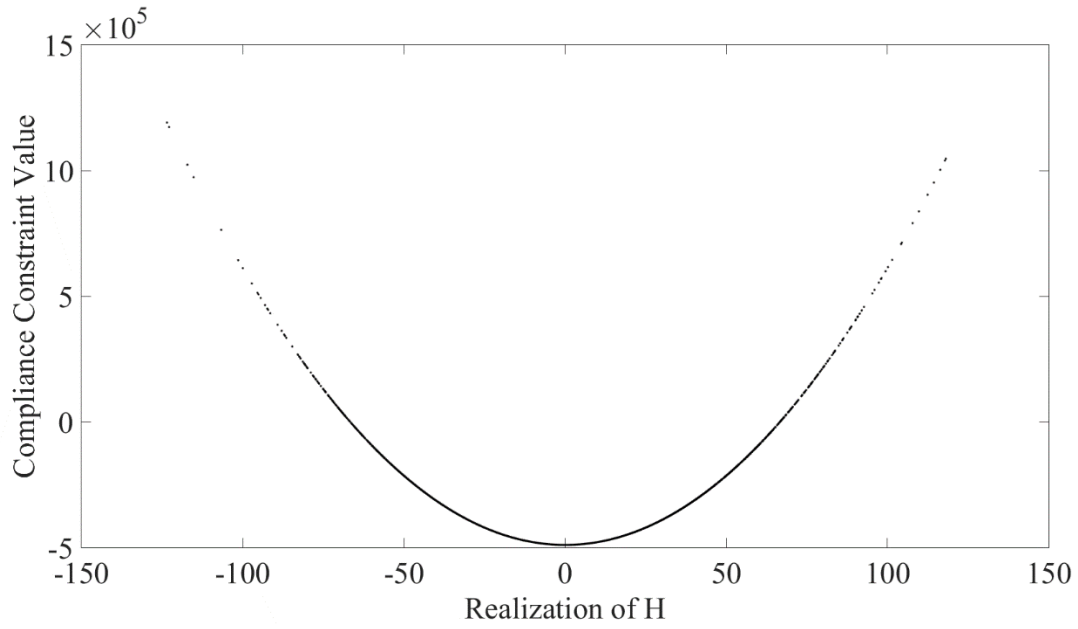
Figure 27: Symmetry of the compliance constraint for the RBTO benchmark with $\tilde{\beta} = 2$
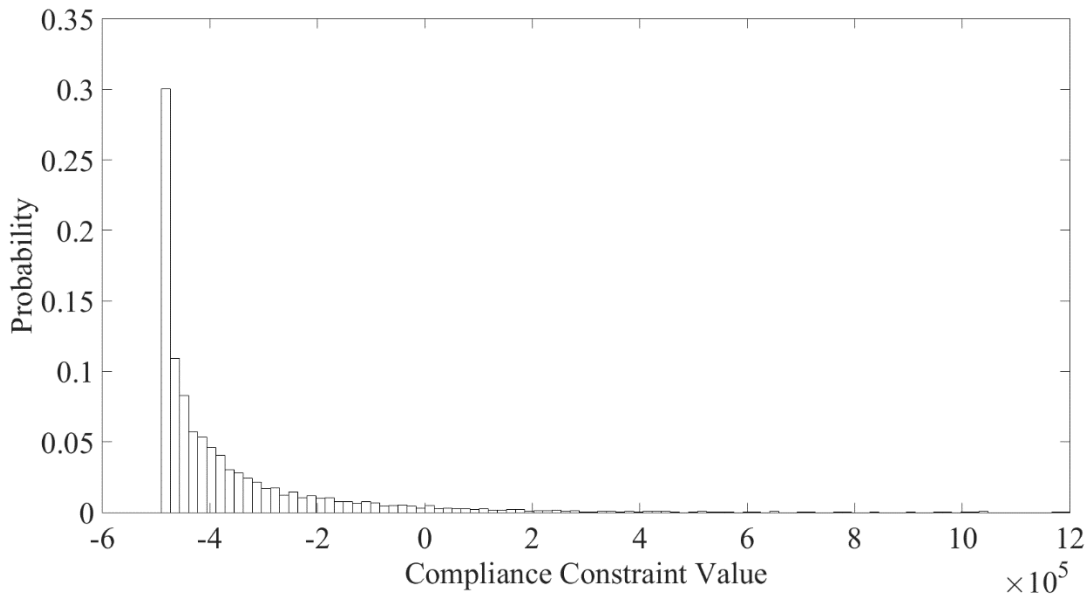
Figure 28: Probability density function for the compliance constraint obtained using MCS

4.4 Multiple Random Loads

4.4.1 Problem Statement

The sensitivity analysis procedure for the RBTO benchmark problem, derived in Section 4.3.2, is applicable for any number of random loads. To verify this, a new reliability-based topology optimization problem with two random forces was defined. Similar to the benchmark problem, the design region is a rectangular grid of linear quadrilateral plane-stress elements. The boundary conditions differ from the benchmark problem however. In the new formulation, the left edge of the region is fixed so that the plate is cantilevered. The two random loads F_1 and F_2 are applied vertically along the centerline of the region at ¾ of the width and the full width. This configuration is depicted in Figure 29.

While the boundary conditions are different, the formulation of the optimization problem is unchanged from the benchmark problem. For clarity, the design problem and nested FORM problem are repeated below. The RBTO problem is given by:

$$\min_{\rho} W = \sum_{i=1}^{m} \rho_i$$
$$\text{s.t.}\, g = \bar{\beta} - \beta \leq 0 \tag{122}$$

Where β is computed by solving a nested FORM problem each iteration:

$$\min_{U} \beta_i = \|U\|$$
$$\text{s.t.}\, D^T K D - \bar{C} = 0 \tag{123}$$
$$K D = R$$

The symmetry problems encountered in the RBTO benchmark are avoided by specifying nonzero-means and relatively small standard deviations. Therefore, while there may be solutions for the compliance constraint corresponding to values of the forces with the opposite sign, these realizations will be exceedingly unlikely and have little bearing on the reliability of the topology. The unit-less values used for all parameters in this analysis are summarized in Table 10.
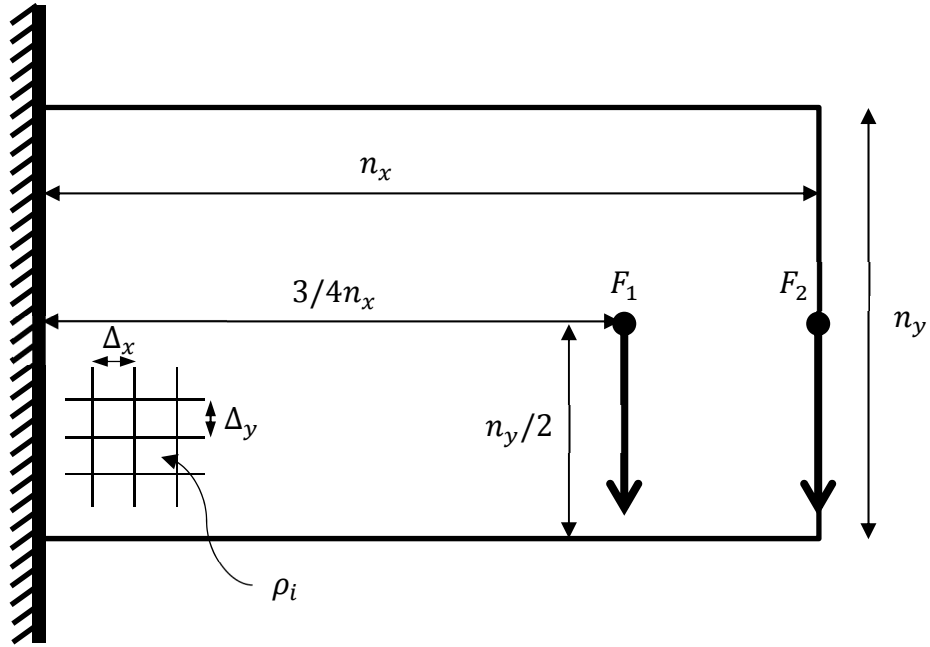
Figure 29: Problem formulation for RBTO with two random loads

Table 10: Summary of unit-less parameters for the multiple random loads RBTO problem

Parameter	Mean Value	Standard Deviation
F_1	200	10 (Case 1) or 20 (Case 2)
F_2	100	10
$\bar{\beta}$	3	N/A
n_x	144	N/A
n_y	48	N/A
Δx	1/48	N/A
Δy	1/48	N/A
t	1	N/A
$\bar{r}$	1/24	N/A
$\bar{C}$	2×10^7	N/A
E_0	1	N/A
v	0	N/A
p	3	N/A

A secondary objective of this analysis is to compare the topologies obtained using the deterministic and reliability-based approaches. As demonstrated in Section 4.2 with a 3-bar truss, RBDO can lead to novel designs which may be more reliable and/or optimal than those found using deterministic optimization, depending on how the deterministic problem is formulated. It is expected that reliability-based topology optimization will have similar benefits with respect to deterministic topology optimization. To investigate this, a deterministic formulation of the multi-load problem will also be solved using several different factors of safety then compared to the RBTO solution. The deterministic formulation will use the same boundary conditions and parameters as the RBTO formulation, with the random loads replaced by deterministic equivalents with magnitudes equal to the mean values of the random loads. The problem formulation for the deterministic topology optimization is given in Equation (124).

$$
\min_{\boldsymbol{\rho}} W = \sum_{i=1}^{m} \rho_i
$$
$$
\text{s.t.} \, g = \boldsymbol{D}^T \boldsymbol{K} \boldsymbol{D} - \frac{\bar{C}}{FS} \leq 0 \tag{124}
$$
$$
\boldsymbol{K} \boldsymbol{D} = \boldsymbol{R}
$$

Deterministic topology optimization was initially performed using the Method of Moving Asymptotes (MMA) [14], however it was observed that this approach tended to converge to infeasible solutions for this specific problem, regardless of how the compliance constraint was scaled. To avoid this behavior, MATLAB's implementation of the interior-point method [13] was used instead. This method was chosen since it incorporates a barrier function to stay in the feasible region of the design space. As a trade-off, the pseudo-densities will tend to be near but not exactly at their bounds. For example, a typical pseudo-density found at the root of the structure using MMA would be 1.000, while the interior-point method would converge to 0.9997 instead. Sensitivity analysis was performed using the expression for the sensitivities of the compliance to any pseudo-density given by Equation (116)

4.4.2 RBTO Solutions

The sensitivities were first verified using a reduced-scale problem. This problem used a 18x6 grid of elements, giving a total of 108 design variables. The finite difference approach was used to generate a baseline. A step size of 10^{-7} was chosen for this analysis since this step size provided the most accurate derivatives in the step-size study performed for the RBTO benchmark problem, detailed in Table 8. With this step size, the maximum difference between the finite difference and analytic sensitivities in the second iteration of the design was 8.77×10^{-5}, indicating that there was very good agreement between the two methods. This proves that the sensitivity equations derived in Section 4.3.2 for the RBTO benchmark problem are applicable for any compliance constrained formulation with random loading.

Having validated the sensitivity analysis procedure, RBTO was then performed for the deflection-limited cantilever structure with the full-scale mesh. Two cases were considered, each with a target reliability index of $\tilde{\beta} = 3$ but different standard deviations of the applied forces. In the first case, both forces had a standard deviation $\sigma_{F_1} = \sigma_{F_2} = 10$. The topology obtained using RBTO is depicted in Figure 30. This topology resembles a truss structure composed of relatively few but thick members which is symmetric about its horizontal mid-plane with a total weight of 2,305. Since there are 6,912 pseudo-densities in the full-scale problem, this indicates that only 33% of the material was used.

Figure 30: RBTO solution for two random loads with $\sigma_{F_1} = 10$ and $\sigma_{F_2} = 10$

For the second case, the standard deviation of F_1 was increased from 10 to 20, which is equivalent to 10% of the mean value of F_1. The standard deviation of F_2 remained unchanged. The resulting topology for $\tilde{\beta} = 3$ is depicted in Figure 31. This topology has a weight of 2,547 which is an increase of 10% over the RBTO solution for $\sigma_{F_1} = 10$. Interestingly, this topology consisted of many narrow members, unlike Case 1. This type of complex structure is typical of many topology optimization problems.
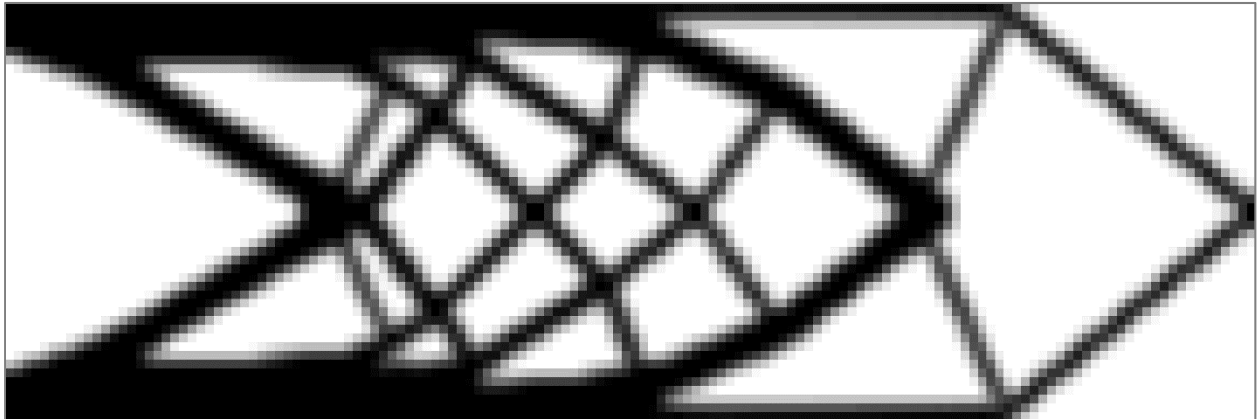
Figure 31: RBTO solution for two random loads with $\sigma_{F_1} = 20$ and $\sigma_{F_2} = 10$

4.4.3 Comparison to Deterministic Solutions for Various Factors of Safety

Deterministic topology optimization was performed for 3 different factors of safety: $FS = 1.25$, 1.375, and 1.5. The resulting topologies are depicted in Figure 32, Figure 33, and Figure 34 respectively. Comparison of the deterministic topologies shows that increasing the factor of safety altered the topology by adding more "X" intersections to the structure. A factor of safety of 1.25 had 2 distinct intersections while $FS = 1.5$ had 3. The topology for $FS = 1.375$ appears to be a transition, with a 3rd intersection of reduced size nested between two larger intersections. Interestingly, the topology for $FS = 1.375$ bears a strong resemblance to the RBTO solution for $\sigma_{F_1} = 10$ and $\sigma_{F_2} = 10$, depicted in Figure 30. This demonstrates that deterministic approaches can still produce reliable designs, however there is no guarantee of this unlike reliability-based approaches. It is also important to note that none of the deterministic topologies are similar to the RBTO solution with the increased standard deviation, depicted in Figure 31.

Figure 32: Deterministic topology for the multiple-load problem with $FS = 1.25$

Figure 33: Deterministic topology for the multiple-load problem with $FS = 1.375$

Figure 34: Deterministic topology for the multiple-load problem with $FS = 1.5$

Monte-Carlo simulation using 10,000 samples was then performed to compare the reliability of the deterministic topologies to that of the RBTO solutions. This analysis was performed twice for each design: once with $\sigma_{F_1} = \sigma_{F_2} = 10$ and then 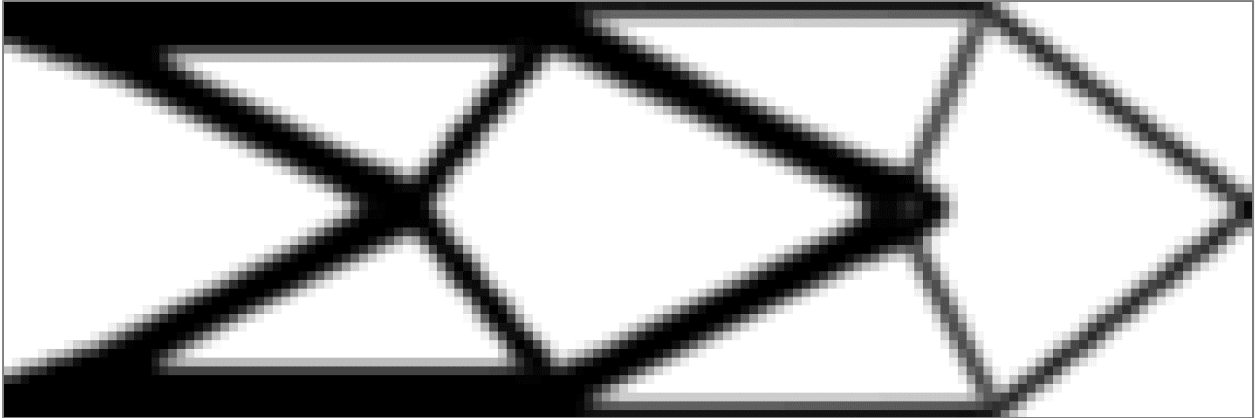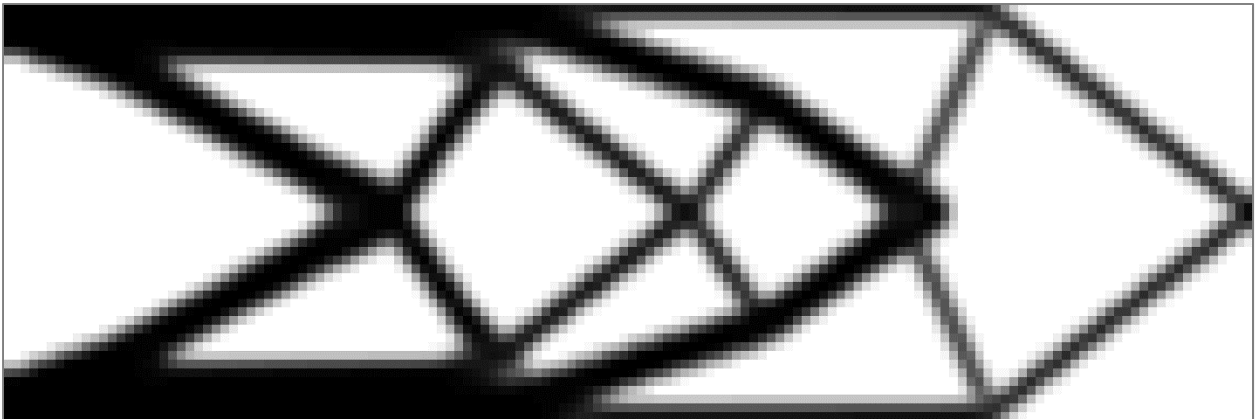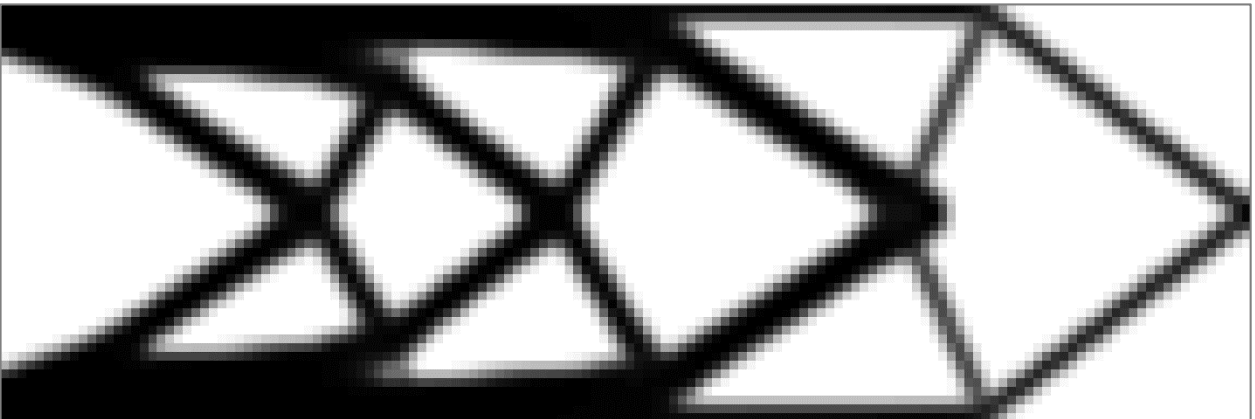again with $\sigma_{F_1} = 20$, $\sigma_{F_2} = 10$. The weight and probabilities of failure for the deterministic and reliability-based topologies are summarized in Table 11.

Table 11: Comparison of deterministic and reliability-based topologies for multiple loads

Design Approach	Weight	P_f (MCS), Case 1	P_f (MCS), Case 2	β (FORM), Case 1	β (FORM), Case 2
Det., $FS = 1.25$	2,146	0.0121	0.0479	2.2437	1.6591
Det., $FS = 1.375$	2,470	0.0006	0.0068	3.2760	2.4266
Det., $FS = 1.5$	2,609	<0.0001	0.0011	4.2631	3.1595
RBDO, $\bar{\beta} = 3.0$, Case 1	2,305	0.0013	N/A	3.0000	N/A
RBDO, $\bar{\beta} = 3.0$, Case 2	2,547	N/A	0.0013	N/A	3.0001

The FORM approximation of the probability of failure for $\bar{\beta} = 3.0$ computed using Equation (20) is 0.00135. As expected, both RBDO solutions had similar probabilities of failure. The deterministic solutions had a wide range of probabilities. Of note is that the deterministic topology for $FS = 1.5$ (Figure 34) had a similar performance to the RBDO Case 2 design (Figure 31) when $\sigma_{F_1} = 20$ and $\sigma_{F_2} = 10$ despite the dissimilar topologies. This demonstrates the ability of RBTO to find novel topologies. More importantly, by using reliability-based topology optimization it was possible to obtain this result after a single analysis instead of using a trial and error approach similar to that required to obtain the equivalent deterministic topology.

4.4.4 Comparison of Analytic and Finite Difference Computation Times

A final objective of the multiple random load analysis is to investigate the savings in computation time achieved using the proposed sensitivity analysis approach. To do this, the first iteration for the RBTO problem with multiple random loads was re-solved twice for several different mesh sizes. The first solution utilized the sensitivity analysis procedure developed in Section 3.2. The second solution utilized the finite difference approach, discussed in Section 2.4.1, which required re-solving the FORM problem once for every pseudo-density. It is

assumed that every MPP found when a pseudo-density is perturbed will be in the neighborhood of the original MPP, so for efficiency the finite difference derivatives used the baseline MPP as the initial point when solving the FORM problem. Both analyses used the parameters for Case 2, summarized in Table 10, with the exception of the number of horizontal and vertical elements. The meshes considered and the resulting computation times are summarized in Table 12. It is important to note that the computation time for each method includes the time required to solve the initial FORM problem and that parallelization of the finite difference derivatives was not considered.

Table 12: Analytic and finite difference computation times versus mesh density

Mesh	Number of Pseudo-Densities	Analytic Computation Time (s)	Finite Difference Computation Time (s)	Percent Reduction
30×10	300	4.872987	54.98623	91.14
36×12	432	7.111835	112.0820	93.65
42×14	588	10.01533	207.3676	95.17
48×16	768	15.53990	360.8196	95.69
54×18	972	17.55799	580.7918	96.98
60×20	1200	26.11555	916.7456	97.15

Comparison of the computation times shows that the analytic approach required a far lower computation time, even for very coarse meshes. For example, the time required to solve a single iteration of the RBTO algorithm for a 30×10 mesh was reduced by 50s by switching to analytic derivatives. This is equivalent to a reduction of over 90%. Furthermore, the difference in the computation costs rapidly increased as the number of pseudo-densities increased, with a 60×20 mesh requiring 890s less when using analytic derivatives. Assuming 26s per iteration, this indicates that over 35 iterations of the RBTO algorithm can be completed using the proposed sensitivity analysis procedure for every iteration of the finite difference algorithm given the 60×20 mesh. Clearly, the proposed sensitivity analysis procedure is more efficient.

To visualize the efficiency of the proposed sensitivity analysis procedure, the computation time as a function of the number of pseudo-densities can be plotted, as shown in Figure 35. This clearly depicts the rapid divergence in the computation cost between the

approaches. Specifically, the finite difference approach exhibits a non-linear increase in the computation cost which resembles a cubic function. This is unlike the analytic sensitivity procedure, which shows a more linear increase in the computation time with far slower growth. Thus, the analytic procedure is expected to become increasingly efficient as the size of the reliability-based topology optimization problem is increased. Overall, the number of pseudo-densities considered in this study is small compared to that of many topology optimization problems, yet there are already significant time savings achieved using the new approach. For these larger problems, the computation cost of finite difference derivatives is expected to be impractically large based off of the observed trends, however the computational challenge can be avoided by using the proposed sensitivity analysis procedure.
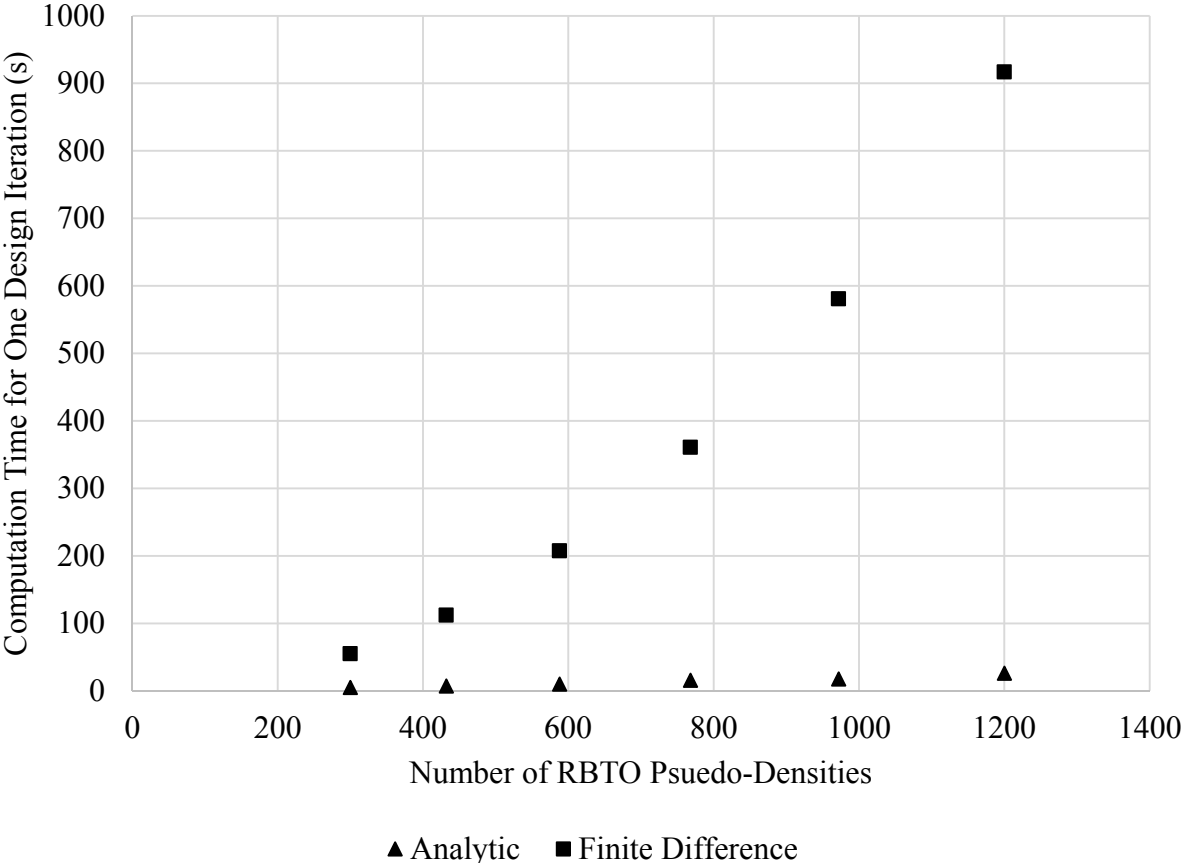

Figure 35: Computation time for one design iteration versus number of pseudo-densities

4.5 RBTO of a Deflection-Limited Cantilever Structure

4.5.1 Problem Statement

The reliability-based benchmark problem developed by Rozvany and Maute [1] is an ideal test case because of the simplicity of the volume objective and compliance constraint as well as the availability of an analytic solution. The benchmark problem does not represent a realistic implementation of reliability-based topology optimization however. Compliance is typically used as an objective rather than a constraint since it is equivalent to the average strain energy density of the structure. Thus, minimization of the compliance serves as an efficient proxy for the stiffness of the structure, since a structure with a lower strain energy has deformed less and is therefore stiffer. Introducing compliance as a constraint requires specifying a target value. The challenge is that it is not easy to correlate a value of the compliance to measurable properties such as the deflection of the structure. Therefore, a more useful problem would be to minimize the volume of a topology subject to a constraint on the displacement of a key location.

In this analysis, a cantilever plate structure will be constrained by the vertical displacement of the tip of the plate, designated point O. Specifically, the reliability against the vertical displacement at point O, d_O, exceeding a critical value d_{max} must be greater than a target reliability index of $\tilde{\beta} = 3$, which is equivalent to a probability of failure of 0.00135. There will be a normally-distributed vertical force F_1 applied at point O while the opposite edge of the plate is fixed. Similar to the previous examples, the design region will be modeled as a rectangular grid of linear quadrilateral plane-stress elements, each with an associated pseudo-density ρ_i. This problem formulation is depicted in Figure 36.

The objective of the deflection-limited RBTO problem is to minimize the weight of the structure given the constraint on the reliability of the tip displacement. The weight is equivalent to the sum of the pseudo-densities while the reliability can be assessed using the First-Order Reliability Method. In standard form, the RBTO problem is given by:

$$\min_{\rho} W = \sum_{i=1}^{m} \rho_i$$
$$\text{s. t. } g = \bar{\beta} - \beta \leq 0 \tag{125}$$

83

Where β is computed by solving a nested FORM problem each iteration:

$$\min_{\boldsymbol{U}} \beta_i = \|\boldsymbol{U}\|$$
$$\text{s.t.} \, \boldsymbol{z}_d^T \boldsymbol{D} - d_{max} = 0 \quad (126)$$
$$\boldsymbol{KD} = \boldsymbol{R}$$

The column vector $\boldsymbol{z}_d^T$ in Equation (126) is chosen to be everywhere zero except for the term with the same global index as d_O, which has a value of 1. Therefore, $\boldsymbol{z}_d^T$ isolates the displacement of interest from the global displacements vector. The nested constraint is expressed in this manner so that the adjoint method, detailed in Section 2.4.3, can be used to efficiently obtain the derivatives of d_O with respect to the pseudo-densities and the random load. Further details of the sensitivity analysis for the random displacement problem are presented in Section 4.5.2. A summary of the parameters used for this analysis is presented in Table 13. These parameters are similar to those used in the RBTO benchmark, so it is important to note that the values do not correlate to any particular unit system. However, this problem can easily be reformulated to consider realistic values for the material properties and loading.
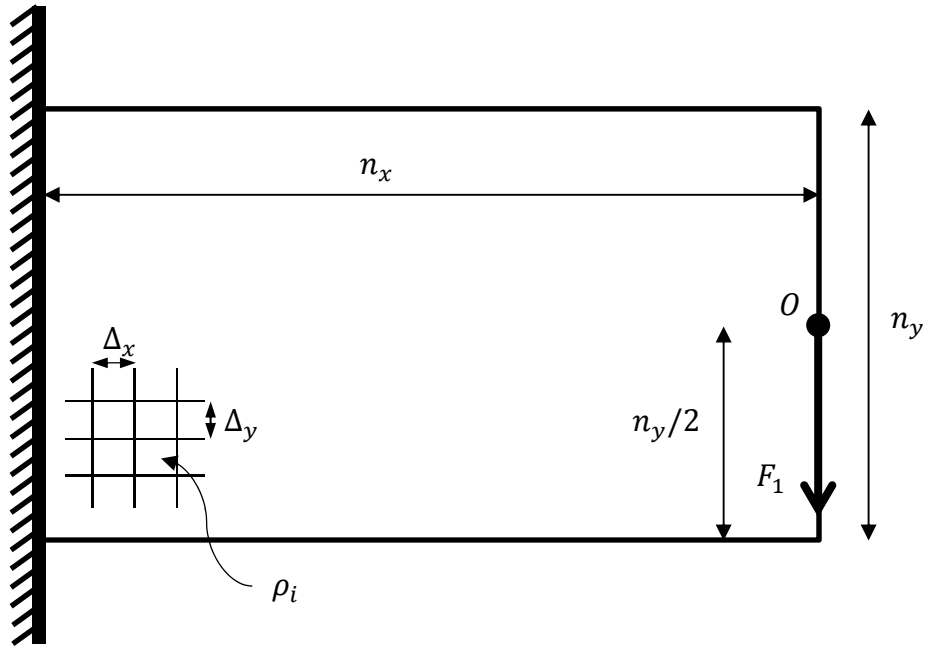

Figure 36: Problem formulation for the deflection-limited RBTO problem

Table 13: Summary of unit-less parameters for RBTO of a deflection-limited structure

Parameter	Mean Value	Standard Deviation
X_1	100	100/3
$\bar{\beta}$	3 or 4	N/A
n_x	144	N/A
n_y	48	N/A
Δx	1/48	N/A
Δy	1/48	N/A
t	1	N/A
$\bar{r}$	1/24	N/A
d_{max}	45,000	N/A
E_0	1	N/A
ν	0	N/A
p	3	N/A

4.5.2 Sensitivity Analysis

Reliability-based topology optimization requires efficient sensitivity analysis because repeated solutions of the nested FORM problem quickly become computationally intractable as the number of pseudo-densities increases. Similar to the RBTO benchmark problem, the sensitivity analysis procedure proposed in Section 3.2 is used in this analysis to obtain analytic sensitivities of the reliability index with respect to the pseudo-densities. This procedure requires expressions for the first and second derivatives of the displacement constraint from Equation (126) with respect the pseudo-densities and the random nodal force. The adjoint method will be used to obtain general expressions for these sensitivities which are applicable for any pseudo-density or random force. Only the results of these derivations are presented in this section, however a similar derivation is presented in Section 4.3.2. Full derivations for each partial derivative are also included in Appendix B.

The first sensitivity required is the derivative of the displacement with respect to each pseudo-density, which can be derived using the adjoint approach discussed in Section 2.4.3. The resulting expression for the derivative is given by Equation (127):

$$\frac{\partial g}{\partial \rho_i} = -p x_i^{p-1} \{K_d^{-1}\}^T k_0 u_i \tag{127}$$

Where $\{K_d^{-1}\}^T$ is a column vector analogous to the element nodal displacement matrix, except containing terms from the column of K^{-1} with the same index as the constrained displacement, d_O. The next derivative required is the sensitivity of the displacement constraint to the random loads. The adjoint method is also required to obtain a general expression for the sensitivity with respect to any random force, given below in Equation (128):

$$\frac{\partial g}{\partial F_j} = K_{dj}^{-1} \tag{128}$$

Where K_{dj}^{-1} is the term of the inverse stiffness matrix with the same global indices as the constrained displacement d_O and the random force F_j. The mixed derivative can then be computed by taking the derivative of Equation (127) with respect to a random force F_j. Similar to the other sensitivities, the adjoint method is required since Equation (127) is a function of the state variables, which in turn are function of the design variables and random parameters. The resulting expression is given in Equation (129).

$$\frac{\partial^2 g}{\partial \rho_i \partial F_j} = -p \rho_i^{p-1} \{K_d^{-1}\}^T k_0 \{K_j^{-1}\} \tag{129}$$

Where $\{K_j^{-1}\}$ is a column vector analogous to the element-level nodal displacement vector d_i, except instead containing the terms from the column of K^{-1} with the same global index as F_j. Lastly, every term in the Hessian of the displacement constraint with respect to any pair of random nodal forces is simply 0 since the first derivative of the displacement constraint, given in Equation (128), is a term from the inverse stiffness matrix, which is not a function of the random loads:

$$\frac{\partial^2 C}{\partial F_k \partial F_j} = 0 \tag{130}$$

The sensitivities obtained using the proposed procedure were validated by comparing the analytic sensitivities of the reliability index with respect to each design variable to sensitivities obtained using finite difference derivatives with a step size of 10^{-7}. A reduced scale problem with an 18×6 mesh was used to obtain these sensitivities. Of the 108 sensitivities computed, the maximum difference between the analytic and finite difference sensitivities computed for the second iteration was 1.856×10^{-6}. This discrepancy is small compared to the magnitude of the sensitivities, with the largest being 0.269, a full 5 orders of magnitude greater. Thus, the impact of the discrepancies on the optimization process would be negligible. It is important to note that this discrepancy is most likely from the subtractive cancellation and step size errors associated with the finite different method.

4.5.3 RBTO Solution

After validating the sensitivities, reliability-based topology optimization was performed on the full-scale mesh with target reliability indices of $\bar{\beta} = 3$ and $\bar{\beta} = 4$. The resulting topologies are depicted in Figure 37 and Figure 38 respectively. The topology optimized for $\bar{\beta} = 3$ has a weight of 2,833, corresponding to 41% of the material available while the topology optimized for $\bar{\beta} = 4$ is heavier, having a weight of 3,265, equivalent to 47% of the material available. Qualitatively, these topologies are similar to the topology obtained for the probabilistic compliance with two random loads, documented in Figure 30. While the loading was different for each problem, both topologies had a similar series of thick members intersecting each other in a manner resembling a truss structure. This demonstrates that compliance is a good approximation for the stiffness of a structure. Unlike the compliance based formulations however, the topologies obtained using the displacement formulation given by Equations (125) and (126) are optimized for an easily measurable property: the displacement of the tip of the cantilever structure. Overall, the solution of the displacement constrained problem demonstrated that the sensitivity analysis procedure proposed in Section 3.2 can be applied to a wide-range of reliability-based topology optimization problems.

Figure 37: Optimum topology for the deflection-limited cantilever plate problem with $\tilde{\beta} = 3$
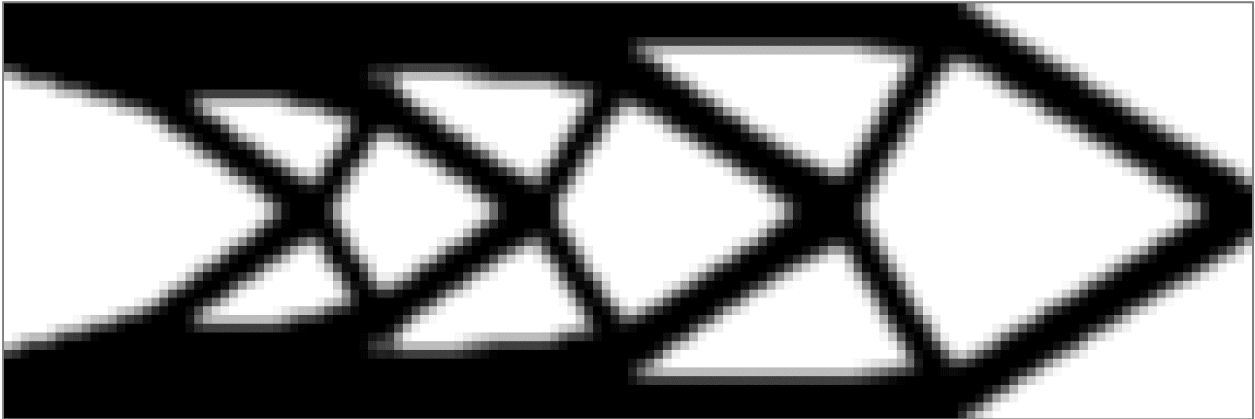

Figure 38: Optimum topology for the deflection-limited cantilever plate problem with $\tilde{\beta} = 4$

5. Conclusion

A sensitivity analysis method for nested optima derived from the Lagrange Multiplier Theorem was proposed in this thesis. This method does not require any additional solutions of the nested optimization problem, though as a trade-off second-order derivative information about the nested objective and its constraints is required to formulate and solve a system of linear equations for the sensitivities. For a computational problem with a large set of state equations, direct and adjoint sensitivity methods should be used to calculate these derivatives. When applied to the First-Order Reliability Method, the sensitivities computed using the proposed method can be used to obtain analytic derivatives of the reliability index with respect to each design variable. The computational effort required to calculate the sensitivities using the new approach is far less than that required to obtain finite difference derivatives of the reliability index, enabling the RBDO approach to be extended to large problems including topology optimization.

The proposed sensitivity analysis was validated through multiple examples. This includes a probabilistic variant of the classic 3-bar truss problem, where a novel design was obtained using the proposed approach. The sensitivity analysis approach was then extended to several reliability-based topology optimization problems, including the benchmark problem developed by Rozvany and Maute [1], where a unique issue related to the symmetry of the problem was identified. To address this, the RBTO procedure had to be modified to consider multiple Most Probable Points, though it should be noted that this required knowledge of the behavior of the probabilistic constraint at the optimum topology. Lastly, the flexibility of the proposed sensitivity analysis approach was demonstrated by solving for a compliance constrained topology with multiple random loads and a deflection-constrained topology optimization problem.

Overall, Reliability-Based Design Optimization can lead to novel designs which are more reliable and at a lower cost than those obtained using deterministic optimization. Reliability-Based Topology Optimization extends these principles further, allowing the practitioner to find improved designs by relaxing the number of assumptions. Existing deterministic design and topology optimization tools are not equipped to handle random parameters without modification however. For many efficient RBDO approaches, the required modifications to the deterministic optimization tool are extensive, thus discouraging the transition to the new design optimization philosophy. The First-Order Reliability Method is an alternative approach which is simple to implement, however the computational effort required is greater as a result of the double-loop structure. The sensitivity analysis procedure proposed in this thesis alleviates this issue, thus making the First-Order Reliability-Method a simple yet viable approach for RBDO, especially for large problems such as Reliability-Based Topology Optimization.

References

1. Rozvany, G. I. N., and Maute, K. "Analytical and numerical solutions for a reliability-based benchmark example," *Structural and Multidisciplinary Optimization* Vol. 43, No. 6, 2011, pp. 745-753.
   doi: 10.1007/s00158-011-0637-8

2. Choi, S.-K., Canfield, R. A., and Grandhi, R. V. *Reliability-based Structural Design*: Springer Verlag London Limited, 2007.

3. Hasofer, A. M., and Lind, N. C. "Exact and Invariant Second-Moment Code Format," *Journal of the Engineering Mechanics Division* Vol. 100, No. 1, 1974, pp. 111-121.

4. Liu, P.-L., and Der Kiureghian, A. "Optimization algorithms for structural reliability," *Structural Safety* Vol. 9, No. 3, 1991, pp. 161-177.
   doi: 10.1016/0167-4730(91)90041-7

5. Rackwitz, R., and Flessler, B. "Structural reliability under combined random load sequences," *Computers and Structures* Vol. 9, No. 5, 1978, pp. 489-494.
   doi: 10.1016/0045-7949(78)90046-9

6. Tu, J., Choi, K. K., and Park, Y. H. "A New Study on Reliability-Based Design Optimization," *Journal of Mechanical Design* Vol. 121, No. 4, 1999, p. 557.
   doi: 10.1115/1.2829499

7. Chen, X., Hasselman, T., and Neill, D. "Reliability based structural design optimization for practical applications," *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit, 38th, and AIAA/ASME/AHS Adaptive Structures Forum*. Kissimmee, FL.

8. Silva, M., Tortorelli, D. A., Norato, J. A., Ha, C., and Bae, H.-R. "Component and system reliability-based topology optimization using a single-loop method," *Structural and Multidisciplinary Optimization* Vol. 41, No. 1, 2010, pp. 87-106.
   doi: 10.1007/s00158-009-0401-5

9. Liang, J., Mourelatos, Z. P., and Nikolaidis, E. "A Single-Loop Approach for System Reliability-Based Design Optimization," *Journal of Mechanical Design* Vol. 129, No. 12, 2007, p. 1215.
   doi: 10.1115/1.2779884

10. Du, X. P., and Chen, W. "Sequential optimization and reliability assessment method for efficient probabilistic design," *JOURNAL OF MECHANICAL DESIGN* Vol. 126, No. 2, 2004, pp. 225-233.
    doi: 10.1115/1.1649968

11. Cheng, G., Xu, L., and Jiang, L. "A sequential approximate programming strategy for reliability-based structural optimization," *Computers and Structures* Vol. 84, No. 21, 2006, pp. 1353-1367.
    doi: 10.1016/j.compstruc.2006.03.006

12. Aoues, Y., and Chateauneuf, A. "Benchmark study of numerical methods for reliability-based design optimization," *Structural and Multidisciplinary Optimization* Vol. 41, No. 2, 2010, pp. 277-294.
    doi: 10.1007/s00158-009-0412-2

13.   "Constrained Nonlinear Optimization Algorithms." Vol. 2017, MathWorks.

14.   Svanberg, K. "The method of moving asymptotes—a new method for structural optimization," *International Journal for Numerical Methods in Engineering* Vol. 24, No. 2, 1987, pp. 359-373.
doi: 10.1002/nme.1620240207

15.   Haftka, R. T., Gürdal, Z., Kamat, M. P., and SpringerLink. *Elements of Structural Optimization*. Dordrecht: Springer Netherlands, 1990.

16.   Hohenbichler, M., and Rackwitz, R. "Sensitivity and importance measures in structural reliability," *Civil Engineering Systems* Vol. 3, No. 4, 1986, pp. 203-209.
doi: 10.1080/02630258608970445

17.   Kwak, B. M., and Lee, T. W. "Sensitivity analysis for reliability-based optimization using an AFOSM method," *Computers and Structures* Vol. 27, No. 3, 1987, pp. 399-406.
doi: 10.1016/0045-7949(87)90064-2

18.   Bendsøe, M. P. "Optimal shape design as a material distribution problem," *Structural Optimization* Vol. 1, No. 4, 1989, pp. 193-202.
doi: 10.1007/BF01650949

19.   Wang, M. Y., Wang, X., and Guo, D. "A level set method for structural topology optimization," *Computer Methods in Applied Mechanics and Engineering* Vol. 192, No. 1, 2003, pp. 227-246.
doi: 10.1016/S0045-7825(02)00559-5

20.   Bourdin, B. "Filters in topology optimization," *International Journal for Numerical Methods in Engineering* Vol. 50, No. 9, 2001, pp. 2143-2158.
doi: 10.1002/nme.116

21.   Sigmund, O., and Danmarks tekniske, u. *Design of material structures using topology optimization*, 1994.

22.   Kharmanda, G., Olhoff, N., Mohamed, A., and Lemaire, M. "Reliability-based topology optimization," *Structural and Multidisciplinary Optimization* Vol. 26, No. 5, 2004, pp. 295-307.
doi: 10.1007/s00158-003-0322-7

23.   Sigmund, O. "A 99 line topology optimization code written in Matlab," *Structural and Multidisciplinary Optimization* Vol. 21, No. 2, 2001, pp. 120-127.
doi: 10.1007/s001580050176

24.   Mozumder, C., Patel, N., Tillotson, D., Renaud, J., and Tovar, A. "An Investigation of Reliability-Based Topology Optimization."

25.   Maute, K., and Frangopol, D. M. "Reliability-based design of MEMS mechanisms by topology optimization," *Computers and Structures* Vol. 81, No. 8, 2003, pp. 813-824.
doi: 10.1016/S0045-7949(03)00008-7

26.   Jalalpour, M., and Tootkaboni, M. "An efficient approach to reliability-based topology optimization for continua under material uncertainty," *Structural and Multidisciplinary Optimization* Vol. 53, No. 4, 2016, pp. 759-772.
doi: 10.1007/s00158-015-1360-7

27.   Jung, H.-S., and Cho, S. "Reliability-based topology optimization of geometrically nonlinear structures with loading and material uncertainties," *Finite Elements in Analysis & Design* Vol. 41, No. 3, 2004, pp. 311-331.
doi: 10.1016/j.finel.2004.06.002

28. Liu, K., Paulino, G. H., and Gardoni, P. "Reliability-based topology optimization using a new method for sensitivity approximation - application to ground structures," *Structural and Multidisciplinary Optimization* Vol. 54, No. 3, 2016, pp. 553-571.
doi: 10.1007/s00158-016-1422-5

29. Kim, C., Wang, S., Bae, K.-r., Moon, H., and Choi, K. K. "Reliability-Based Topology Optimization with Uncertainties," *Journal of Mechanical Science and Technology* Vol. 20, No. 4, 2006, pp. 494-504.
doi: 10.1007/BF02916480

30. López, C., Baldomir, A., and Hernández, S. "Deterministic versus reliability-based topology optimization of aeronautical structures," *Structural and Multidisciplinary Optimization* Vol. 53, No. 4, 2016, pp. 907-921.
doi: 10.1007/s00158-015-1363-4

31. Patel, N., Renaud, J., Agarwal, H., and Tovar, A. "Reliability Based Topology Optimization Using the Hybrid Cellular Automaton Method."

Appendix A: Compliance Derivatives

The sensitivity analysis procedure for the FORM problem detailed in Section 3.2 requires expressions for several partial derivatives of the constraint function $g(x, U)$, as shown by Equation (86). These partial derivatives include the first derivatives of the constraint with respect to the design variables $\partial g/\partial x_i$ and the random parameters $\partial g/\partial U_j$. Expressions for the second derivatives of the constraint with respect to any pair of random parameters $\partial^2 g/\partial U_j \partial U_k$ as well as the mixed-second derivatives $\partial^2 g/\partial x_i \partial U_j$ are also required to use the proposed procedure. This appendix presents the derivations for each of these partial derivatives for the RBTO benchmark problem developed by Rozvany and Maute [1], presented in Section 4.3, as well as the multiple-load problem discussed in Section 4.4. Both of these examples are constrained by the reliability against the compliance exceeding a critical value, with the SIMP pseudo-densities ρ chosen as design variables and one or more applied nodal forces F being treated as random. In each example, the nested FORM constraint is given by:

$$g(\rho, F) = C(\rho, F) - \bar{C} = 0 \tag{131}$$

Where the compliance can be related to the finite element equations using:

$$C(\rho, F) = D(\rho, F)^T K(\rho) D(\rho, F) \tag{132}$$

As shown in the following subsections, it is possible to derive generalized expressions for the partial derivatives of the compliance which are applicable for any pseudo-density or random load.

94

A.1 Derivative with Respect to any Pseudo-Density

An expression for the partial derivative of the compliance with respect to any SIMP pseudo-density, $\partial C / \partial \rho_i$, can be obtained using the direct method, described in Section 2.4.3. The compliance can first be expressed as:

$$C = \boldsymbol{D}^T \boldsymbol{R} \tag{133}$$

The load vector is not a function of the pseudo-densities, however the displacement vector is. Therefore, the derivative with respect to a given pseudo-density ρ_i is:

$$\frac{\partial C}{\partial \rho_i} = \frac{\partial \boldsymbol{D}^T}{\partial \rho_i} \boldsymbol{R} \tag{134}$$

Using the direct method, the derivative of the displacement vector $\partial \boldsymbol{D}^T / \partial \rho_i$ can be obtained by taking the derivative of the finite element equations with respect to the pseudo-densities, given in Section 2.4.3 by Equation (34). After substituting Equation (34) in and considering that the loads vector is not a function of the pseudo-densities, the following expression for the compliance derivative is obtained:

$$\frac{\partial C}{\partial \rho_i} = \left(-\boldsymbol{K}^{-1} \frac{\partial \boldsymbol{K}}{\partial \rho_i} \boldsymbol{D} \right)^T \boldsymbol{R} \tag{135}$$

The above expression can be further simplified by considering the symmetry of the stiffness matrix. The resulting expression for the derivative of the compliance is:

$$\frac{\partial C}{\partial \rho_i} = -\boldsymbol{D}^T \left(\frac{\partial \boldsymbol{K}}{\partial \rho_i} \right)^T \boldsymbol{D} \tag{136}$$

Each pseudo-density is only a linear multiplier for a single element stiffness matrix. Therefore, the derivative of the global stiffness matrix with respect to a given pseudo-density is simply:

$$\frac{\partial \boldsymbol{K}}{\partial \rho_i} = p \rho_i^{P-1} \boldsymbol{L}_i^T \boldsymbol{k}_0 \boldsymbol{L}_i \tag{137}$$

Where L_i is the locator matrix for the element associated with pseudo-density ρ_i. After substituting Equation (137) into Equation (136) and simplifying further, the final expression for the derivative of the compliance with respect to any pseudo-density is obtained:

$$\frac{\partial C}{\partial \rho_i} = -p\rho_i^{P-1}\boldsymbol{d}_i^T \boldsymbol{k}_0 \boldsymbol{d}_i \tag{138}$$

Where $\boldsymbol{d}_i$ is the element level nodal displacement vector for the element influenced by ρ_i. Overall, this expression is applicable for any pseudo-density and makes no assumptions about the size or geometry of the finite element problem.

A.2 Derivative with Respect to any Random Load

The partial derivative of the compliance with respect to any random load F_j can be obtained using the direct method, detailed in Section 2.4.3. Since the random loads only appear in the load vector $\boldsymbol{R}$, only the displacement and load vector will be a function of the random loads. Therefore, the derivative of the compliance with respect to an arbitrary random load is:

$$\frac{\partial C}{\partial F_j} = \left(\frac{\partial \boldsymbol{D}}{\partial F_j}\right)^T \boldsymbol{K}\boldsymbol{D} + \boldsymbol{D}^T \boldsymbol{K}\frac{\partial \boldsymbol{D}}{\partial F_j} \tag{139}$$

The derivative of the displacement vector with respect to the random load can be obtained using the direct method. After taking derivatives of the finite element equations and rearranging, the following expression for $\partial \boldsymbol{D}/\partial F_j$ is obtained:

$$\frac{\partial \boldsymbol{D}}{\partial F_j} = \boldsymbol{K}^{-1}\frac{\partial \boldsymbol{R}}{\partial F_j} \tag{140}$$

The above expression can then be substituted into Equation (139). After simplification using the symmetry of the stiffness matrix, the following expression is obtained:

$$\frac{\partial C}{\partial F_j} = \left(\frac{\partial \boldsymbol{R}}{\partial F_j}\right)^T \boldsymbol{K}^{-1}\boldsymbol{K}\boldsymbol{D} + \boldsymbol{D}^T \boldsymbol{K}\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{R}}{\partial F_j} \tag{141}$$

The product of a matrix and its inverse is simply the identity matrix $\boldsymbol{I}$. Additionally, the two terms in Equation (141) are transposes of each other. Since each ultimately reduces to a scalar expression, the two terms are equal and the derivative reduces to:

$$\frac{\partial C}{\partial F_j} = 2 \left(\frac{\partial R}{\partial F_j} \right)^T \boldsymbol{ID} \tag{142}$$

The derivative of the load vector $\boldsymbol{R}$ with respect to the random force F_j can then be considered. Assuming that the random load is a point force applied in one of the principle global directions, the random load will only appear in a single term of the load vector. Therefore, the derivative of the load vector is everywhere zero with the exception of the term with the same index as F_j. To illustrate this, consider a load vector with a single random load:

$$\boldsymbol{R} = [0 \quad \ldots \quad F_j \quad \ldots \quad 0]^T \tag{143}$$

The derivative with respect to F_j is therefore:

$$\frac{\partial R}{\partial F_j} = [0 \quad \ldots \quad 1 \quad \ldots \quad 0]^T \tag{144}$$

Assuming that there are no scalar multipliers for the random load, the above expression can then be substituted into Equation (142). After simplifying, the final expression for the derivative of the compliance with respect to an arbitrary random load is obtained:

$$\frac{\partial C}{\partial F_j} = 2d_{F_j} \tag{145}$$

Where d_{F_j} is the term of the displacement vector $\boldsymbol{D}$ with the same global index as the random load F_j. Overall, this derivation makes no assumption about which random load is chosen, only that it is a point load aligned with one of the principle global directions. The derivation does not make any assumptions about the number of random loads either, so this expression is applicable for any problem with random point loads, regardless of the specific boundary conditions.

A.3 Second Derivative for any Pair of Random Loads

The second partial derivative with respect to any pair of random loads, $\partial^2 C / \partial F_j \partial F_k$, can be obtained by considering the first derivative of the compliance with respect to a random load F_j. Specifically, the derivative of Equation (142) with respect to the second random load F_k can be taken:

$$\frac{\partial^2 C}{\partial F_j \partial F_k} = \frac{\partial}{\partial F_k}\left(2\left(\frac{\partial R}{\partial F_j}\right)^T ID\right) \tag{146}$$

The above expression can then be expanded using the product rule:

$$\frac{\partial^2 C}{\partial F_j \partial F_k} = 2\left(\frac{\partial^2 R}{\partial F_k \partial F_j}\right)^T D + 2\left(\frac{\partial R}{\partial F_j}\right)^T \frac{\partial D}{\partial F_k} \tag{147}$$

Similar to the derivation of the first derivative with respect to any random load detailed in the previous section, the assumption that the random load is a point load aligned with one of the principle global directions is made. This assumption simplifies the first derivatives of the load vector R with respect to each random force, as shown by Equation (144). Since Equation (144) is not a function of any random loads, the second derivative of the load vector with respect to any pair of random point forces is simply 0:

$$\frac{\partial^2 R}{\partial F_k \partial F_j} = 0 \tag{148}$$

Therefore, Equation (147) reduces to:

$$\frac{\partial^2 C}{\partial F_j \partial F_k} = 2\left(\frac{\partial R}{\partial F_j}\right)^T \frac{\partial D}{\partial F_k} \tag{149}$$

The term $\partial R/\partial F_j$ was previously derived in Equation (144) and the term $\partial D/\partial F_k$ was previously derived in Equation (140). Substituting these in and simplifying gives the final expression for the second derivative of the compliance with respect to any pair of random loads:

$$\frac{\partial^2 C}{\partial F_j \partial F_k} = 2K_{jk}^{-1} \tag{150}$$

Where K_{jk}^{-1} is the term of the inverse stiffness matrix whose global indices match those of F_j and F_k. Since the stiffness matrix is symmetric, its inverse is as well. Therefore, K_{jk}^{-1} is equal to K_{kj}^{-1} and the order with which the derivatives are taken does not matter.

A.4 Mixed Derivative for any Density and Load Pair

A general expression for the mixed partial derivative $\partial^2 C / \partial \rho_i \partial F_j$ can be obtained by considering the first derivative of the compliance with respect to any random load, derived previously in Section A.2:

$$\frac{\partial C}{\partial F_j} = 2d_j \tag{151}$$

Where d_j is the term of the displacement vector $\boldsymbol{D}$ with the same global index as the random load F_j. An equivalent expression is:

$$\frac{\partial C}{\partial F_j} = 2\boldsymbol{z}_j^T \boldsymbol{D} \tag{152}$$

Where $\boldsymbol{z}_j^T$ is a vector which is everywhere zero except the term with the same index as d_j, which has a value of 1 instead. The derivative with respect to an arbitrary pseudo-density ρ_i can then be taken to obtain the mixed derivative using adjoint method, detailed in Section 2.4.3. After substituting in the results of this approach, given by Equation (38), the following expression is obtained:

$$\frac{\partial^2 C}{\partial \rho_i \partial F_j} = -2\boldsymbol{z}_j^T \boldsymbol{K}^{-1} \frac{\partial \boldsymbol{K}}{\partial \rho_i} \boldsymbol{D} \tag{153}$$

The derivative of the stiffness matrix with respect to the pseudo-density ρ_i was previously computed in Equation (137). Substituting Equation (137) into the above expression gives:

$$\frac{\partial^2 C}{\partial \rho_i \partial F_j} = -2p\rho_i^{P-1} \boldsymbol{z}_j^T \boldsymbol{K}^{-1} \boldsymbol{L}_i^T \boldsymbol{k}_0 \boldsymbol{L}_i \boldsymbol{D} \tag{154}$$

Since the inverse stiffness matrix is symmetric, this expression is equivalent to:

$$\frac{\partial^2 C}{\partial \rho_i \partial F_j} = -2p\rho_i^{P-1} \left(L_i \left(K^{-1} z_j \right) \right)^T k_0 d_i \tag{155}$$

The $K^{-1}z_j$ term creates a column vector equal to the column of K^{-1} with the same global index as d_j. Therefore, multiplying $K^{-1}z_j$ by the locator matrix L_i is analogous to obtaining the vector of nodal displacements for element i, d_i, from the global displacement vector D. The result is a reduced column vector containing terms from K^{-1} with the same column index as d_j and the same row indices as the degrees of freedom of element i, designated $\{K_j^{-1}\}$. Therefore, the mixed second derivative for any pseudo-density and load pair is:

$$\frac{\partial^2 C}{\partial \rho_i \partial F_j} = -2p\rho_i^{P-1} \{K_j^{-1}\}^T k_0 d_i \tag{156}$$

Similar to the derivations in the previous sub-sections, no assumptions are made about the specific problem geometry or the number and arrangement of the random loads.

Appendix B: Displacement Derivatives

The sensitivity analysis procedure for the FORM problem detailed in Section 3.2 requires expressions for several partial derivatives of the constraint function $g(x, U)$, as shown by Equation (86). These partial derivatives include the first derivatives of the constraint with respect to the design variables $\partial g / \partial x_i$ and the random parameters $\partial g / \partial U_j$. Expressions for the second derivatives of the constraint with respect to any pair of random parameters $\partial^2 g / \partial U_j \partial U_k$ as well as the mixed-second derivatives $\partial^2 g / \partial x_i \partial U_j$ are also required to use the proposed procedure. This appendix presents the derivations for each of the partial derivatives for the RBTO of a deflection limited cantilever structure example detailed in Section 4.5. This example is constrained by the reliability against a specified nodal displacement, d_O, exceeding a critical value. Additionally, the SIMP pseudo-densities ρ are chosen as design variables and one or more applied nodal forces F are treated as random. The nested displacement constraint is given below:

$$g(\rho, F) = d_O(\rho, F) - d_{max} = 0 \tag{157}$$

Alternatively, the displacement constraint can be formulated in terms of the global displacement matrix $D(\rho, F)$:

$$g(\rho, F) = z_O^T D(\rho, F) - d_{max} = 0 \tag{158}$$

The above formulation is chosen so that the adjoint method, described in Section 2.4.3, can be used to efficiently obtain the sensitivities of the finite element equations. As shown in the following subsections, it is possible to derive generalized expressions for the partial derivatives of the nodal displacement which are applicable for any pseudo-density or random load.

B.1 Derivative with Respect to any Pseudo-Density

The partial derivative of the displacement constraint with respect to any pseudo-density $\partial g/\partial \rho_i$ can be obtained using the adjoint method, discussed in Section 2.4.3. The displacement constraint for the RBTO problem, given above in Equation (158), has the same structure as the generic displacement constraint used in the adjoint method derivation in Section 2.4.3, given by Equation (35). Therefore, the first derivative of the constraint with respect to any pseudo-density is given by Equation (38):

$$\frac{\partial g}{\partial \rho_i} = -\boldsymbol{z}_d^T \boldsymbol{K}^{-1} \left(\frac{\partial \boldsymbol{K}}{\partial \rho_i} \boldsymbol{D} - \frac{\partial \boldsymbol{R}}{\partial \rho_i} \right) \tag{159}$$

Since the load vector $\boldsymbol{R}$ is not explicitly a function of the pseudo-densities, its derivative with respect to any pseudo-density is zero. Thus, the first derivative of the displacement constraint with respect to any pseudo-density reduces to:

$$\frac{\partial g}{\partial \rho_i} = -\boldsymbol{z}_d^T \boldsymbol{K}^{-1} \frac{\partial \boldsymbol{K}}{\partial \rho_i} \boldsymbol{D} \tag{160}$$

An expression for the derivative of the global stiffness matrix $\boldsymbol{K}$ with respect to any pseudo-density was previously derived in Equation (137). Substituting Equation (137) into the above expression gives:

$$\frac{\partial g}{\partial \rho_i} = -p\rho_i^{p-1} \boldsymbol{z}_d^T \boldsymbol{K}^{-1} \boldsymbol{L}_i^T \boldsymbol{k}_0 \boldsymbol{L}_i \boldsymbol{D} \tag{161}$$

An equivalent expression can be obtained by considering the symmetry of the inverse stiffness matrix:

$$\frac{\partial g}{\partial \rho_i} = -p\rho_i^{p-1} \left(\boldsymbol{L}_i (\boldsymbol{K}^{-1} \boldsymbol{z}_d) \right)^T \boldsymbol{k}_0 \boldsymbol{d}_i \tag{162}$$

The $\boldsymbol{K}^{-1}\boldsymbol{z}_d$ term creates a column vector equal to the column of $\boldsymbol{K}^{-1}$ with the same global index as d_O. Therefore, multiplying $\boldsymbol{K}^{-1}\boldsymbol{z}_d$ by the locator matrix $\boldsymbol{L}_i$ is analogous to obtaining the vector of nodal displacements for element i, $\boldsymbol{d}_i$, from the global displacement vector $\boldsymbol{D}$. The result is a reduced column vector containing terms from $\boldsymbol{K}^{-1}$ with the same column index as d_O

and the same row indices as the degrees of freedom of element i, designated $\{K_d^{-1}\}$. Therefore, the first derivative of the displacement constraint with respect to any pseudo-density is:

$$\frac{\partial g}{\partial \rho_i} = -p\rho_i^{p-1}\{K_d^{-1}\}^T k_0 d_i \tag{163}$$

The above expression is applicable for any pseudo-density ρ_i, regardless of the mesh or boundary conditions chosen.

B.2 Derivative with Respect to any Random Load

The partial derivative of the displacement constraint with respect to any random load F_j $\partial g / \partial F_j$ can be obtained using the adjoint approach, discussed in Section 2.4.3. Since the displacement constraint has the same form as Equation (35), the first derivative with respect to any random load is given by Equation (38). In terms of a random load F_j, the derivative is given by:

$$\frac{\partial g}{\partial F_j} = -z_d^T K^{-1}\left(\frac{\partial K}{\partial F_j}D - \frac{\partial R}{\partial F_j}\right) \tag{164}$$

The global stiffness matrix K is not explicitly a function of the random loads. Therefore, the above expression simplifies to:

$$\frac{\partial g}{\partial F_j} = z_d^T K^{-1}\frac{\partial R}{\partial F_j} \tag{165}$$

If it is assumed that each random load is a point force applied at one of the nodes and oriented in one of the global principle coordinate directions, then each F_j will only appear in a single term of the load vector R:

$$R = [0 \quad \dots \quad F_j \quad \dots \quad 0]^T \tag{166}$$

The derivative with respect to F_j is therefore:

$$\frac{\partial R}{\partial F_j} = [0 \quad \dots \quad 1 \quad \dots \quad 0]^T \tag{167}$$

Since $\boldsymbol{z}_d^T$ is zero for all terms except for the term with the same global index as d_O, the product $\boldsymbol{z}_d^T \boldsymbol{K}^{-1}$ reduces to a row vector equal to the row of $\boldsymbol{K}^{-1}$ with the same global index as d_O:

$$\frac{\partial g}{\partial F_j} = K_{dj}^{-1} \tag{168}$$

Where K_{dj}^{-1} is the term of the inverse stiffness matrix having the same row index as the displacement of interest d_O and the same column index as the random force F_j.

B.3 Second Derivative for any Pair of Random Loads

The second partial derivatives of the displacement constraint for any pair of random loads $\partial^2 g / \partial F_k \partial F_j$ can be obtained by considering the first derivative with respect to a random load F_j, given by Equation (168). Taking a derivative again with respect to the random load F_k yields the expressions

$$\frac{\partial^2 g}{\partial F_k \partial F_j} = \frac{\partial}{\partial F_k} K_{dj}^{-1} \tag{169}$$

Since the inverse stiffness matrix is not a function of any random load, all second derivatives of the displacement constraint with respect to the random loads are simply zero:

$$\frac{\partial^2 g}{\partial F_k \partial F_j} = 0 \tag{170}$$

B.4 Mixed Derivative for any Density and Load Pair

An expression for the mixed second partial derivatives with respect to any pseudo-density ρ_i and random load F_j pair $\partial^2 g / \partial F_j \partial \rho_i$ can be obtained by considering the first derivative with respect to any pseudo-density, given by Equation (163). The derivative with respect to a random load F_j can then be taken:

$$\frac{\partial^2 g}{\partial F_j \partial \rho_i} = -\frac{\partial}{\partial F_j}\left(p\rho_i^{p-1}\{\boldsymbol{K}_d^{-1}\}^T \boldsymbol{k}_0 \boldsymbol{d}_i\right) \tag{171}$$

The element stiffness matrix and the inverse stiffness matrix are not explicitly functions of any random loads. Therefore, these matrices can be treated as constants and the only derivatives needed are those of the displacement matrix with respect to the random load F_j:

$$\frac{\partial^2 g}{\partial F_j \partial \rho_i} = -p\rho_i^{p-1}\{K_d^{-1}\}^T k_0 \frac{\partial}{\partial F_j}(L_i D) \tag{172}$$

The derivative of the displacement vector with respect to a random load was previously derived in Section A.2. The resulting expression is given by Equation (140) and is repeated below:

$$\frac{\partial D}{\partial F_j} = K^{-1}\frac{\partial R}{\partial F_j} \tag{173}$$

After substituting the above expression into Equation (172), the mixed second derivative is:

$$\frac{\partial^2 g}{\partial F_j \partial \rho_i} = -p\rho_i^{p-1}\{K_d^{-1}\}^T k_0 L_i \left(K^{-1}\frac{\partial R}{\partial F_j}\right) \tag{174}$$

As shown by Equations (166) and (167), the derivative of the load vector R with respect to a given load F_j is a column vector which is zero everywhere except for the term corresponding to F_j, which has a value of 1. Therefore, multiplying the inverse stiffness matrix K^{-1} by this vector results in a column vector of the terms of K^{-1} with the same column index as F_j, designated K_j^{-1}. The resulting expression for the mixed second derivative is:

$$\frac{\partial^2 g}{\partial F_j \partial \rho_i} = -p\rho_i^{p-1}\{K_d^{-1}\}^T k_0 L_i K_j^{-1} \tag{175}$$

Lastly, multiplying K_j^{-1} by the locator matrix L_i is analogous to extracting the displacements for the nodes of element i from the global displacement vector D. The result is a reduced column vector containing the terms from the column of K^{-1} with the same index as F_j and the same row indices as the degrees of freedom of element i, designated $\{K_j^{-1}\}$. Thus, the mixed second derivative reduces to:

$$\frac{\partial^2 g}{\partial F_j \partial \rho_i} = -p\rho_i^{p-1}\{K_d^{-1}\}^T k_0\{K_j^{-1}\} \tag{176}$$

Overall, this expression is applicable for any pseudo-density ρ_i and random load F_j, with the difference being the specific terms from the inverse stiffness matrix used.