National Technical University of Athens

School of Civil Engineering

Institute of Structural analysis and Antiseismic Research

Master thesis

Reduced Order Modeling Methods in Topology Optimization

Rallis Athanasios

Supervisor Professor: Nikos Lagaros

Interdisciplinary Postgraduate Programme: COMPUTATIONAL MECHANICS

Reduced Order Modeling Methods in Topology Optimization

RALLIS ATHANASIOS

Civil Engineering

Interdisciplinary Postgraduate Programme: COMPUTATIONAL MECHANICS

NATIONAL TECHNICAL UNIVERSITY OF ATHENS

February 2019

Summary

This thesis examines efficient solution procedures for the structural analysis problem within topology optimization. The research is motivated by the observation that when the nested approach to structural optimization is applied, most of the computational effort is invested in repeated solutions of the analysis equations. For demonstrative purposes, the discussion is limited to topology optimization of linear problems within the field of structural mechanics.

The main focus of the thesis is on the utilization of various approximations to the solution of the analysis problem, where the underlying model corresponds to linear elasticity. For computational environments that enable the direct solution of large linear equation systems using matrix factorization, we utilize efficient procedures based on model order reduction via reduced basis.

These approaches are tested on two- and three-dimensional topology optimization problems of minimum compliance design. The topologies generated by the approximate procedures are practically identical to those obtained by the standard approach.

The thesis starts with the presentation of Structural optimization, where emphasis is given in topology optimization and the implementation of SIMP approach. We also make a brief presentation of direct and iterative solution methods for linear structural analysis. In the following chapters of the thesis the concept of model order reduction is introduced, the approximate procedures that are utilized are described and finally the results from the numerical tests are presented.

Contents

1. Introduction

The presented thesis deals with efficient solution procedures for the structural analysis problem within topology optimization. Over the last two decades topology optimization has undergone a rapid period of growth both in industry and academia spurred by a large number of theoretical, practical and algorithmic developments. This growth has resulted in dealing with large scale problems that can involve millions of degrees of freedom. In topology optimization problems the nested approach is frequently applied meaning optimization is performed in the design variables only while the equilibrium equations are solved separately. Such case is the topology optimization problem of minimization of the mean compliance of a static structure. It aims at finding the stiffest structure that fits into a given domain, satisfies external loads and boundary conditions and has a prescribed volume. Independently of the strategy adopted for solving such a problem, the most expensive step is the solution of the system of equations that describes the equilibrium conditions of the structure. Under the assumption that the structure is formed by an elastic material and is subject to small displacements, a single linear system has to be solved at each iteration of the optimization algorithm, to determine the nodal displacements that are used to evaluate the objective function. Because the stiffness matrix of the structure is symmetric and positive definite, this linear system is usually solved by means of the Cholesky factorization, despite its high cost in the large-scale setting. This motivates the search for efficient approaches aimed at reducing the computational effort invested in the analysis. Moreover, by applying efficient procedures the solution of larger and more complex models compared to standard procedures are available.

In order to mitigate this cost, the use of reduced order modeling approaches, also known as reduced basis models, have been proposed. By projection of the system of equations that describes the equilibrium conditions of the structure on a reduced basis have proved to be efficient methods for achieving drastic dimensionality and computational cost reductions. The main idea behind this concept is to construct a so called reduced basis, and then solve the problem projected on this low dimensional basis with drastically reduced computational cost.

In this thesis we address structural topology optimization problems in which the underlying analysis model is linear. For linear problems, the proposed procedures are based on constructing an appropriate reduced basis in which the equilibrium equations are being projected and a good approximation of the solution is obtained. The efficiency of such proposed approaches are examined and construction of the reduced basis via Proper Orthogonal Decomposition is proposed.

The thesis is organized as follows.

In Chapter 2, we present the theoretical background of structural optimization, with particular reference to the method of topology optimization and its formulations using the modified SIMP approach. Emphasis is given on the problem formulations, objective functions and sensitivity analysis procedures considered in the various test cases that are examined. Chapter 3 introduces the concept of Model Order Reduction and we briefly discuss various approximate procedures for linear structural analysis in topology optimization. Chapter 4 describes the background of the Solution Methods that are used in the various test cases examined. Chapter 5 includes a summary of the results from the numerical tests derived from each solution method for different parameters and the efficiency and the accuracy these methods are examined.

2. Structural -Topology Optimization

In this chapter we present an overview of what is structural optimization, the general mathematical concepts used to formulate the structural topology, the theory and the implementation issues of topology optimization and finally a few words about the solution of the equilibrium equation $KU = f$ that arises from linear structural analysis.

2.1 Introduction to Structural optimization

To formulate the structural optimization problem, an objective function, design variables and state variables needs to be introduced as described in [5]. The objective function f represents an objective that could either be minimized or maximized. Such an objective could be the volume or the stiffness of a structure. Moreover, the structural design domain and the state variables associated to the objective function has to be defined. The design variables x describes the design of the structure, it could represent for example the geometry. The state variables y describes the structural response which may represent stress, strain or displacement. The state variables depend on the design variables $y(x)$. The objective function is subjected to the design and state variable constraints to steer the optimization to a sought solution.

$$\begin{cases} \min\limits_{x} & f\big(x, y(x)\big) \\ subject\ to & \begin{cases} design\ constraint\ on\ x \\ state\ constraint\ on\ y(x) \\ equilibrium\ constraint \end{cases} \end{cases} \tag{2.1}$$

A state function $g(y)$ that represents the state variables can be introduced, for example a displacement in a certain direction. This state function can be incorporated as a constraint to the optimization procedure, where it is usually formulated such that $g(y) \leq 0$. Consider the case where $g(y)$ is represented by a displacement vector $g(u(x))$ in a discrete finite element problem. To establish the state function, this requires that nodal displacement are solved for

$$u(x) = K(x)^{-1} f(x) \tag{2.2}$$

where K is the global stiffness matrix and f is the global load vector. This means that the optimization task can be expressed in a so-called nested formulation where the equilibrium constraint is taken care of by the state function formulation

$$\begin{cases} \min\limits_{x} & f(x) \\ subject\ to & g(u(x)) \leq 0 \end{cases} \tag{2.3}$$

The optimization task presented in equation (2.1) is called simultaneous formulation in comparison. Equation 2.3 is usually solved by evaluating derivatives of f and g with respect to x. In this context, x will represent a geometrical feature. Based on what geometrical feature that is parametrized, the structural optimization problem can be classified into:

Size optimization: the design variable x, represents a structural thickness such as a distributed thickness or a cross-sectional area of a truss model that can be varied. The optimal thickness distribution typically minimizes (or maximizes) a physical quantity such as the mean compliance (strain energy) or the deflection, while the equilibrium and other constraints on the state and design variables have to be fulfilled. The state function may then be the deflection of the structure

Shape optimization: the design variable x, represents the boundary of the state equation. In this case, the boundary of the considered domain x could vary such that some physical quantity is minimized.

Topology optimization: the design variable x, represents the connectivity of the domain. It involves the determination of features such as the number and location ad shape of holes and the connectivity of the design domain.

In the present study we make use of the topology optimization and thus in the following section we will only describe the structural problem of topology optimization.

2.2 Topology optimization

Topology optimization is a mathematical method that optimizes material layout within a given design space, for a given set of loads, boundary conditions and constraints with the goal of maximizing the performance of the system. We seek an optimal placement of material points where the reference domain is partitioned into void and solid elements by a finite element discretization.

2.2.1 Material interpolation

In the design of topology of a structure we are interested in determination of the optimal placement of a given isotropic material in space, i.e. we should determine which points of space should be material points and which should remain void. That is, we think of the geometric representation of a structure as similar to a black-white rendering of an image. In discrete form this then corresponds to a black-white raster representation of the geometry, with "pixels" (or "voxels") given by the finite element discretization.

In mathematical terms we seek an optimal subset $\Omega_{mat} \subset \Omega$ where Ω is the available design domain. The design variable x is now represented by the density vector ρ containing elemental densities ρ_e. The local stiffness tensor $\boldsymbol{E}$ can be formulated by incorporating $\boldsymbol{\rho}$ as a integer formulation:

$$E(\rho) = \rho E^0$$
$$\rho_e = \begin{cases} 1 \; if \; e \in \Omega_{mat} \\ 0 \; if \; e \in \Omega \backslash \Omega_{mat} \end{cases} \tag{2.4}$$

and a volume constraint:

$$\int_{\Omega} \rho d\Omega = Vol(\Omega_{mat}) \leq V \tag{2.5}$$

where V is the volume of the initial design domain. When $\rho_e = 1$ we consider an element to be filled whereas an element with $\rho_e = 0$ is considered to be a void element. To use a gradient based solution strategy for the optimization problem, the integer problem described in (2.4) needs to be formulated as a continuous function so that the density function can take values between 0 and 1 [3]. The most common method to relax the integer problem is the so-called “power-law approach” or Solid Isotropic Material with Penalization (SIMP) method. Here, material properties are assumed constant within each element used to discretize the design domain and the variables are the element relative densities. The material properties are modelled as the relative material density raised to some power times the material properties of solid material. The density function is then written as

$$E = \rho^p E^0, \qquad \rho \in [\rho_{min}, 1], \qquad p > 1 \tag{2.6}$$

where p is the penalizing factor that penalizes elements with intermediate densities to approach 0 or 1, ρ_{min} is the lower density value limit to avoid singularities. Thus,

the penalization is achieved without introducing any explicit penalization scheme. For materials with Poisson ratio ν = 0.3, it is recommended in [4] to use p ≥ 3. To ensure existence of solutions, the power-law approach must be combined with filtering techniques.

2.2.2 The checkerboard problem

Checkerboarding refers to the problem where optimization results shows elements which are alternating solid and void in a checkerboard like pattern. It was earlier believed that these regions represented some optimal microstructure design but proved to be due to poor stiffness representation using finite elements [6]. To avoid the formation of checkerboard patterns some sort of restriction on the resulting design must be introduced. This can be done by applying a filtering technique. An illustration of the checkerboard problem for a two dimensional problem is shown in figure 2.1 produced by the MATLAB code described in [2]. It is presented in Figure 2.1, where it can be seen that the checkerboard pattern occurs in Figure 2.1a. Looking at Figure 2.1b, where a filter to mitigate the checkerboard phenomenon is applied, it can be seen that the material points are placed more homogeneously. Furthermore, higher order elements and mesh refinement could also mitigate the checkerboard problem.

Figure 2.1: Illustration of the consequence by applying a filter (right case) or not (left case).

2.2.3 Problem formulation

The optimization problem formulated in a nested formulation in equation (2.3) is now written as:

$$\begin{cases} \min\limits_{x} & f(\rho) \\ subject\ to & \begin{cases} 0 \leq \rho \leq 1 \\ state\ function\ constraint \\ Manufacturing\ constraints \end{cases} \end{cases} \tag{2.7}$$

when considering topology optimization using the SIMP interpolation method, ρ is a vector containing the element densities. Two common objectives to be minimized are the compliance (C) and the volume (V). A state function constraint may be a displacement in a certain direction.

Minimize compliance

A possibility to maximize the global stiffness of a structure is to minimize its compliance. The compliance is therefore defined as the equivalent strain energy of the FE solution which yields higher stiffness when minimized. The compliance is defined as

$$C(\rho) = f^T u \tag{2.8}$$

where u is the solution of the equilibrium equation

$$K(\rho)u = f \tag{2.9}$$

and $K(\rho)$ is

$$K(\rho) = \sum_{e=1}^{nel} \rho_\varepsilon^p K_e^0 \tag{2.10}$$

K_e^0 is the elemental stiffness matrix with the initial stiffness tensor E^0. To prevent the optimized structure from ending up with the full design volume as a result when searching for its maximum structural stiffness, we need to impose a volume constraint. If a gradient based approach is used, derivatives with respect to $C(\rho)$ are evaluated.

Minimize volume

Another possibility is to minimize the volume.

$$V(\rho) = \sum_{e=1}^{nel} \rho_e^p V_e^0 \tag{2.11}$$

where V_e^0 is the initial volume. To prevent the optimization from minimizing all material, we need for example to impose a constraint for maximum displacement or effective stress. The optimization task is carried out with respect to the objective function and constraints. However, if the objective function is formulated with respect to volume or weight, derivatives are evaluated with respect to the constraints.

If a gradient based solution method is used, the derivatives are evaluated with respect to the constraint instead of the objective. For example, if a displacement vector is imposed, the so-called state derivatives with respect to $u(\rho)$ are evaluated.

2.2.4 Solution method

Two common solutions methods for topology optimization are the Optimality Criteria (OC) and the Method of Moving Asymptotes (MMA), both of which are described more in detail in [3].
The MMA is similar to other mathematical programming algorithms such as Sequential Linear Programming (SLP) and Sequential Quadratic Programming (SQP) to solve non-linear optimization problems in the sense that they also uses sequences with sub-problems which are approximations of the original problem. For MMA, these sub-problems are constructed by gradient information, furthermore these approximations are assumed to be convex.
The OC method uses the method of Lagrange multipliers to establish the optimization task where compliance is minimized under a volume constraint. This defines an update scheme for the design densities ρ. For simple compliance optimization problems, the OC may be faster but for more complicated problems involving several load cases and constraints, the MMA gives better convergence.

2.3 Topology optimization using SIMP approach

2.3.1 SIMP approach

In this thesis, we make use of the SIMP approach in compliance minimization problems. The design domain is descretized by square finite elements and a density-based approach to topology optimization is followed [13]; i.e. each element e is assigned a density x_e that determines its Young's modulus E_e:

$$E_e(x_e) = E_{min} + x_e^p(E_0 - E_{min}), \qquad x_e \in [0,1] \tag{2.12}$$

where E_0 is the stiffness of the material (nominal Young's modulus of the material), E_{min} is a very small stiffness assigned to void regions in order to prevent the stiffness

matrix from becoming singular, and p is a penalization factor (typically $p = 3$) introduced to ensure black-and-white solutions.
The topology optimization problem is then expressed as a compliance minimization problem, i.e. the optimization formulation seeks to find the density distribution over all the elements that minimizes the work done by the external forces under prescribed loadings, boundary conditions and material volume fraction to be used. The mathematical formulation of the optimization problem reads as follows:

$$\min_{x} \quad c(x) = U^T U = \sum_{e=1}^{N} E_e(x_e) u_e^T k_0 u_e$$

$$\text{Subject to:} \qquad V(x)/V_0 = f \tag{2.13}$$

$$KU = F$$

$$0 \leq x \leq 1$$

where $\boldsymbol{c}$ is the compliance, $\boldsymbol{U}$ and $\boldsymbol{F}$ are the global displacement and force vectors, respectively, $\boldsymbol{K}$ is the global stiffness matrix, u_e is the element displacement vector, k_0 is the element stiffness matrix for an element with unit Young's modulus, $\boldsymbol{x}$ is the vector of design variables (i.e. the element densities), N is the number of elements used to discretize the design domain, $V(x)$and V_0 are the material volume and design domain volume, respectively, and *f* is the prescribed volume fraction.

2.3.2 Optimality criteria method

The optimization problem (2.13) is solved by means of a standard optimality criteria method. A heuristic updating scheme is followed

$$x_e^{new} = \begin{cases} \max(0, x_e - m) & if\ x_e B_e^{\eta} \leq \max(0, x_e - m) \\ \min(1, x_e + m) & if\ x_e B_e^{\eta} \geq \max(0, x_e - m) \\ x_e B_e^{\eta} & otherwise \end{cases} \tag{2.14}$$

where *m* is a positive move limit, $\eta\ (= 1/2)$ is a numerical damping coefficient, and B_e is obtained from the optimality condition as:

$$B_e = \frac{-\frac{\partial c}{\partial x_e}}{\lambda \frac{\partial V}{\partial x_e}} \tag{2.15}$$

where the Lagrangian multiplier λ must be chosen so that the volume constraint is satisfied; the appropriate value can be found by means of a bisection algorithm. The sensitivities of the objective function c and the material volume V with respect to the element densities x_e are given by:

$$\frac{\partial c}{\partial x_e} = -p x_e^{p-1}(E_0 - E_{min}) u_e^T k_0 u \tag{2.16}$$

$$\frac{\partial V}{\partial x_e} = 1 \tag{2.17}$$

Equation (2.17) is based on the assumption that each element has unit volume.

2.3.3 Filtering

In order to ensure existence of solutions to the topology optimization problem and to avoid the formation of checkerboard patterns [14][15][16], some restriction on the design is imposed. A common approach is the application of a filter to either the sensitivities or the densities.

The sensitivity filter modifies the sensitivities $\partial c / \partial x_e$ as follows:

$$\frac{\widehat{\partial c}}{\partial x_e} = \frac{1}{\max(\gamma, x_e) \sum_{i \in N_e} H_{ei}} \sum_{i \in N_e} H_{ei} x_i \frac{\partial c}{\partial x_i} \tag{2.18}$$

where N_e is the set of elements i for which the center-to center distance $\Delta(e,i)$ to element e is smaller than the filter radius $r_{\min}$ and H_{ei} is a weight factor defined as:

$$H_{ei} = \max(0, r_{min} - \Delta(e,i)) \tag{2.19}$$

The term γ ($= 10^{-3}$) in (2.18) is a small positive number introduced in order to avoid division by zero. This is a difference to the classical SIMP approach where the term γ is not required, because the density variables cannot became zero.

The density filter transforms the original densities x_e as follows:

$$\hat{x}_e = \frac{1}{\sum_{i \in N_e} H_e} \sum_{i \in N_e} H_{ei} x_i \tag{2.14}$$

In this thesis a density filter is used [11] [12].

The original densities x_e are referred to as the design variables. The filtered densities $\tilde{x}_e$ are referred to as the physical densities. This terminology is used to stress the fact that the application of a density filter causes the original densities x_e to loose their physical meaning. One should therefore always present the filtered density field $\tilde{x}_e$ rather than the original density field x_e as the solution to the optimization problem [17].

The sensitivities with respect to the design variables x_j are obtained by means of the chain rule:

$$\frac{\partial \psi}{\partial x_j} = \sum_{e \in N_j} \frac{\partial \psi}{\partial \tilde{x}_e} \frac{\partial \tilde{x}_e}{\partial x_j} = \sum_{e \in N_i} \frac{1}{\sum_{i \in N_e} H_{ei}} H_{je} \frac{\partial \psi}{\partial \tilde{x}_e} \tag{2.14}$$

where the function ψ represents the objective function c

As was mentioned earlier for topology optimization of large scale structures the bulk of the computational cost comes from the requirement to compute at each step of the optimization the solution of the equilibrium equations:

$$KU = F \tag{2.14}$$

Computing this solution for large scale problems involves the inversion of a very large system of equations that can consist of up to millions of degrees of freedom.

2.4 Linear structural analysis

The system of algebraic equations to be solved in a SIMP problem of topology optimization and generally in any linear static finite element analysis (FEA) is

$$Ku = F \tag{2.15}$$

where K is the global stiffness matrix, u is the unknown displacements vector and F is the external load vector. The stiffness matrix Kin such cases is symmetric, positive definite and sparse. Having these properties K can be stored in memory in a very compact manner by exploiting symmetry and sparsity. The solution of (2.15) is obtained by employing either a direct or an iterative equation solver. In general, direct solvers are more robust and are preferred when the factorized form of K can be stored in memory. This is the case for small and medium scale 2-D FE problems. For 3-D models, K usually has a relatively large bandwidth so that iterative solvers are more appropriate due to their low memory requirements.

2.4.1 Direct solution methods

As far as direct solution methods are concerned they are algorithms based on Gauss elimination. Due to the symmetry and positive definiteness of the stiffness matrix, it can be decomposed using the Cholesky factorization

$$K = U^T U \tag{2.14}$$

where U is an upper triangular matrix. Then, we can obtain the vector of displacements u by forward and backward substitutions

$$U^T v = f$$

$$Uu = v$$

In structural optimization, a sequence of analysis equations of the form (2.15) is generated and should be solved. In the nested approach to topology optimization, the overall computational effort is typically dominated by the cost of solving the analysis equations.

The relatively high cost of matrix factorization in large-scale problems, in particular in three-dimensional FEA, motivates the development of efficient procedures that avoid repeated factorizations.

2.4.2 Iterative solution methods

Iterative methods for solving large sparse linear systems have been gaining popularity over direct methods. In earlier times, iterative methods were usually developed for particular applications and their performance depended on the actual problem parameters. Nowadays, various general-purpose iterative solvers are available, among which the family of Krylov subspace solvers is applied most extensively. For 3-D models and parallel high performance computers, Krylov iterative solvers are much more efficient than direct solvers. Therefore in the context of reducing computational effort in topology optimization, it is essential to address the use of such solvers for solving the structural analysis equations. Among the family of Krylov subspace solvers, the most appropriate method for solving symmetric positive definite systems such as (2.15) is the conjugate gradient (CG) method. The rate of convergence depends on the condition number of the system matrix K, therefore it is necessary to use effective preconditioning in order to achieve fast convergence. Demonstrated with symmetric preconditioning, this means that in practice CG will be applied to solve

$$\tilde{K}\tilde{u} = \tilde{f}$$

where

$$\tilde{K} = M^{-T} K M^{-1}$$

$$\tilde{u} = Mu$$

$$\tilde{f} = M^{-T} f$$

The preconditioner M can be, for example, an incomplete factor of K so that the eigenvalue distribution of $\tilde{K}$ is much better than that of K.

3. MODEL ORDER REDUCTION

Abstract With the use of current computing technology and advanced algorithms we are able to cope with more complex problems than we could in previous years. Nevertheless, there is a need for model order reduction in order to address even more complex problems. In this chapter we give a detail aspect of what model order reduction is and how it can be implemented in linear systems.

3.1 Introduction

Prior to the recent appearance of powerful digital computers, it was necessary to construct models of physical behaviors that took advantage of existing analytical techniques or which involved numerical calculations with small numbers of degrees of freedom. Now, partial differential equations, representative of complex physics that were previously unobtainable, can be discretized and integrated with numerical algorithms implemented on massive parallel supercomputers. The simulation of physical behaviors in even three space dimensions has become relatively commonplace. This has led in today's technological world, the computational simulation of a model to be generally accepted as an equal discipline with theory and real experiment. Physical experiments leads to theories which are constantly validated through the performance of more experiments, and with the use of theory we can produce predictions by performing virtual experiments.

Computer simulations are now performed routinely for many physical, chemical and other processes, and virtual design environments have been set up for a variety of problem classes in order to ease the work of designers and engineers. In this way, new products can be designed faster, more reliably, and without having to make costly prototypes.

There is an increasing demand for realistic simulations of complex problems which in turn increases the demand for more efficient mathematical models in the area of computational science and engineering. Realistic simulations imply that the errors of the virtual models should be small, and that different aspects of the product must be taken into account. To do so more care must be taken in the numerical treatment.

An important factor in enabling the complex simulations carried out today is the increase in computational power. This increase in computational power goes hand-in-hand with developments in numerical algorithms. Iterative solution techniques for linear systems are mainly responsible for this speed-up in algorithms. Important contributions in this area are the conjugate gradient (CG) method, preconditioned conjugate gradient (PCG) methods and multigrid methods. The combination of computational power with the use of advanced numerical algorithms offers a great

deal of speed-up that result in the continuous upgrade of computational science. The solution to many problems, that once where unsolvable, is now derived routinely.

The developments described above also have a counter side. The increased power of computers and algorithms reduces the need to develop smart, sophisticated solution methods that make use of properties of the underlying systems. For example, whereas in the 1960s and 1970s one often had to construct special basis functions to solve certain problems, this can be avoided nowadays by using brute force methods using grids that are refined in the right places. The question arises whether we could use the knowledge generated by these very accurate, but time-consuming, simulations to generate the special basis functions that would have constituted the scientific approach a few decades ago. This is a promising idea, as many phenomena are described very well by a few dominant modes.

3.2 The Concept of Model Order Reduction

There are several definitions of model order reduction, and it depends on the context which one is preferred. Originally, MOR was developed in the area of systems and control theory, which studies properties of dynamical systems in application for reducing their complexity, while preserving their input-output behavior as much as possible. Nowadays, model order reduction is a field of research, both in systems and control theory and in numerical analysis. This has a very healthy effect on MOR as a whole, bringing together different techniques and different points of view, pushing the field forward rapidly.

What do we mean by model order reduction? As was mentioned in the previous sections, we need to deal with models that may contain many equations and variables (10^5 – 10^9). To do so certain simplification of the models needs to be made. Such simplification is needed in order to perform simulations within an acceptable amount of time and limited storage capacity, but with reliable outcome. In some cases, we would even like to have on-line predictions of the behavior with acceptable computational speed, in order to be able to perform optimizations of processes and products.

Model Order Reduction tries to capture as quickly as possible the essential features of a structure. This means that in an early stage of the process, the most basic properties of the original model must already be present in the smaller approximation. At a certain moment the process of reduction is stopped. At that point all necessary properties of the original model must be captured with sufficient precision. All of this has to be done automatically

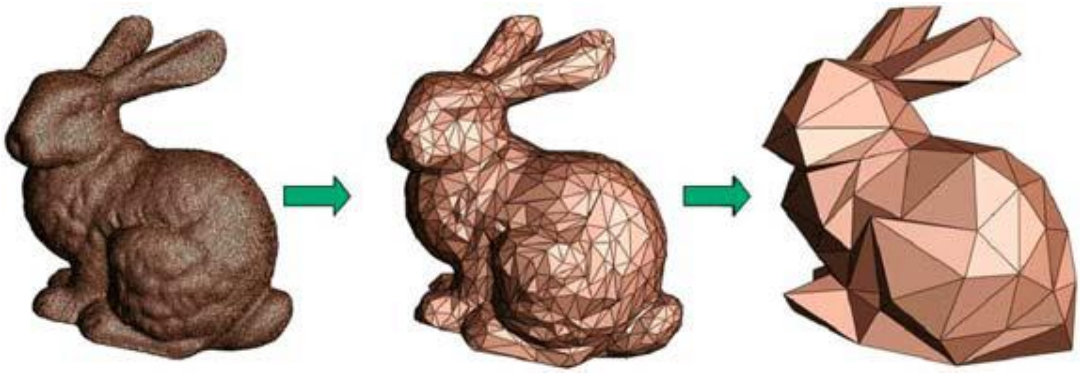


Figure 3.1 Graphical illustration of model order reuction

Figure 3.1 illustrates the concept in a graphical easy-to-understand way, demonstrating that sometimes very little information is needed to describe a model. This example with pictures of the Stanford Bunny shows that, even with only a few facets, the rabbit can still be recognized as such. Although this example was constructed for an entirely different purpose, and does not contain any reference to the way model order reduction is performed mathematically, it can be used to explain what model order reduction is about.

3.3 Model order reduction via Reduced basis method

The model order reduction methodology of reduced basis (RB) techniques offers efficient treatment to finite element schemes by providing both approximate solution procedures and efficient error estimates.

Many real world problems can be modelled by parametrized partial differential equations, where a parameter vector $\boldsymbol{\mu}$ characterizes the system in terms of material, geometry or control parameters. In simulation of general PDEs, repeated computation runs for varying parameters $\boldsymbol{\mu}$ and fast simulation response can be required. Such scenarios occur in design optimization, optimal control with PDE-constraints, online simulation, parameter identification or state estimation. Model order reduction techniques must be applied to satisfy these time demands. The methodology of reduced basis (RB) techniques aims at efficient treatment in such cases by providing both an approximate solution procedure and efficient error estimates.

The main goal of RB-methods is to provide a reduced simulation scheme, in which for a newly given parameter vector $\boldsymbol{\mu}$ a function $u_N(\mu)$ is determined as an approximation to the unknown detailed solution $u_H(\mu)$. In addition or alternatively to the field-variable itself, approximate computation of derived outputs can be desired, i.e. determination of $s_N(\mu) \approx s(u_H(\mu))$ for some output functional s. The underlying reduced simulation scheme is based on a Galerkin projection of the

detailed simulation onto a low-dimensional space $W_N \subset W_H$, the reduced basis space of dimension N. This space is deliberately determined such that it captures the solution variety under parameter variations. This basis-construction is frequently based on so called snapshots, i.e. detailed simulation results $u_H(\mu_i)$ for stationary problems. In addition to the reduced simulation scheme, a second main focus of RB-methods is to provide rigorous a-posteriori error bounds for the error between the reduced to the detailed solution or output. In addition to the reduced simulation scheme, a second main focus of RB-methods is to provide rigorous a-posteriori error bounds for the error between the reduced to the detailed solution or output. These error bounds are in general based on effective computation of residual norms during the reduced simulation and are quite tight. A further important aspect in RB-methods is to make the reduced simulation useful in a multi-query context [34].

3.4 Computational drawbacks of the solution of the full order model in linear systems

Inverses of matrices play a role in formulas for important model order reduction problems. However, in numerical computations the explicit inversion of a matrix has to be avoided if possible. In most situations, the inverse of a matrix is used to denote an operator acting on a vector or another matrix. We consider the evaluation of the vector *y* defined as

$$y = A^{-1}b$$

Determining the explicit inverse of *A* is expensive. The classical way is to compute an *LU* factorization first and then solve linear systems with *L* and *U* for all canonical basis vectors. After A^{-1} has been obtained, it has to be multiplied with b in order to obtain y.

The vector *y* can be obtained in a much cheaper way. We see that y has to be solved from

$$Ay = b$$

Again, we use the LU factorization and first solve z from $Lz = b$, followed by the solution from $Uy = z$. These two solution steps require as much computation as the sole multiplication of A^{-1} and b, and we have avoided to solve the *n* systems with the canonical basis vectors as right-hand sides. Likewise, always when the matrix A^{-1} occurs as an operator acting on some vector or some other matrix, then the expression can be numerically evaluated as we have demonstrated for y.

For sparse matrices, computational differences may be even much more dramatic. In relevant cases, A^{-1} may be dense, while *L* and *U* are sparse. For instance, if *A* is a positive definite tridiagonal matrix, then solving $Ax = b$ via *LU* decomposition requires only in the order of *n* arithmetic computations. The bottom line is: *avoid explicit inversion*.

Direct computation of *y* via *LU* requires, for dense matrices, in total in the order of *n*3 floating point operations, which may pose practical limits for large values of *n*. This is a major motivation for Model Order Reduction: to save computational costs and computer memory storage.

Eigenvalue computations lead to similar observations. The standard way to compute non-trivial λ and *x*, satisfying

$$Ax = \lambda x$$

requires the diagonalization of the matrix *A*. The state of the art technique for this is to first transform A to a more efficient to handle form by orthogonal transformations. In particular, A is transformed by a finite number of Householder or Givens transformations to upper Hessenberg form H: $QTAQ = H$. The eigenvalues of A are equal to those of A and the eigenvalues of *H* are computed with the QR-method. The operations in QR can efficiently be done on Hessenberg matrices. The eigenvalue computation is essentially an iterative process, but QR converges so fast that in practice the complete method (reduction to upper Hessenberg plus QR) is viewed as a direct technique. The whole process requires a few times *n*3 arithmetic operations, and again this poses practical problems if n is large.

3.5 Reduced basis modeling for linear systems

Many numerical simulations in the engineering domain involve solving a partial differential equations problem. After space (or time) discretization, the problem often involves a large system of equations of the form

$$K(u;\mu) = F \tag{3.1}$$

with $u \in \mathbb{R}^n$ the unknown state variables and $\mu \in \mathbb{R}^p$ a set of p parameters of interest (material parameters, time...) so that K: $\mathbb{R}^n \times \mathbb{R}^p$, n being the number of state variables. Let us assume that K is such that given any value of the set of parameters $\boldsymbol{\mu}$ a unique solution $u = u(\mu)$ exists.

Model order reduction aims at significantly decreasing the computational burden associated with the inversion of system (3.1). With the utilization of reduced basis approach we aim at reducing the number of state variables of the model by

projection on a certain basis. Accordingly, an approximation of the solution is sought in a subspace V of dimension *m* (with usually *m*<<*n*), while enforcing the residual to be orthogonal to the same sub-space V. Typically, V is defined by a so called *reduced-basis* $\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_m\}$
The initial problem of Eq. 3.1 is rewritten projected onto the reduced basis:

$$\Phi^T K(\Phi\alpha, \mu) = \Phi^T F \tag{3.2}$$

where $\boldsymbol{\alpha}$ are the reduced state variables, that is the coefficients of vector $\boldsymbol{u}$ expressed in the reduced basis $\boldsymbol{\Phi}$.

If K is linear with respect to its first variable $\boldsymbol{u}$, the problem of Eq. 3.1 can be written as:

$$K(\mu)u = F \tag{3.3}$$

In structural mechanics $K(\mu)$ is the stiffness matrix, which usually depends on some parameters of interest $\boldsymbol{\mu}$, that can be material properties, $\boldsymbol{u}$ is the displacement vector and $\boldsymbol{F}$ the vector of the forces.
Similarly then, the projected problem of Eq. 3.3 can be written as:

$$\Phi^T K(\mu)\Phi\alpha = \Phi^T F \tag{3.4}$$

At this point it is important to realize that Eq. 3.4 is equivalent to a reduced order model of the initial problem of Eq. 3.3. Indeed, solving the problem of Eq. 3.3 typically involves the inversion of a large system of equations of size n, the size of the stiffness matrix $K(\mu)$, which for large scale problems can easily reach hundreds of thousands. On the other hand solving the reduced order model of Eq. 3.4 involves the inversion of a much smaller system of equations of size *m*, the size of the projected stiffness matrix , which is equal to the dimensionality of the reduced basis *m* (typically *m*<<*n*, since *m* does usually not exceed a few dozen). Solving this reduced order model leads directly to $\boldsymbol{\alpha}$, the coefficients of the solution in the reduced basis.
The problem projected onto the reduced basis thus yields an approximate solution whose accuracy can be quantified by measuring the following residual:

$$e_{rb}^2 = \frac{\|\Phi^T K(\mu)\Phi\alpha - F\|^2}{\|F\|^2} \tag{3.5}$$

The subspace V on which the problem is projected, or more precisely one of its basis Φ, is not specified and a variety of different choices are available for this projection.

For example, eigenmodes of the operators have been used to reduce numerically the size of the problems for applications in dynamics. These approaches are known as modal analysis, Craig-Bampton [27] In [18],[26] the reuse of Krylov subspaces generated during a Krylov iterative solvers was used as a reduced basis. The generalized modes of variables separation techniques (like Proper Generalized decomposition) can also be used in such a context [25].

The proper orthogonal decomposition (POD) can also be a way to build a relevant reduced basis in the context of reduced order modeling by projection [18]-[20]. Indeed, POD [21] (also known as Karhunen Loeve decomposition [22],[23] or principal component analysis [24]) is an approach which consists in constructing a reduced basis from a set of solution, called snapshots. Mathematically, the extraction of the reduced basis from the snapshots is done by singular value decomposition. Generally, the snapshots are the results of full simulations on a set of points.

Note that classical POD-type approaches do not generally go all the way to solving or even formulating the reduced order model of Eq. 3.4. Instead POD is often used only as a dimensionality reduction approach. Solutions $\boldsymbol{u}$ that were already calculated by solving the full size problem (Eq. 3) are projected on the POD-reduced basis and expressed in terms of their basis coefficients $\boldsymbol{\alpha}$ thus allowing to express initial solutions $\boldsymbol{u}$ with a drastically reduced dimensionality. Note however that these basis coefficients $\boldsymbol{\alpha}$ can also be obtained by solving the reduced order model of Equation (3.4), that is by solving the initial problem projected on the reduced basis. This option is not typically used in POD because solving the reduced order model for solutions that are already in the reduced basis (thus for which the full solution is already available) has no interest. However the reduced order model has an interest for obtaining the basis coefficients $\boldsymbol{\alpha}$ at a new point which is not part of the reduced basis. In this case, the reduced order solution is only an approximation and the approximation error can be assessed by the metric provided in Equation (3.5).

The reduced basis modeling approach has two major assets: dimensionality reduction by the use of the basis coefficients and computational time reduction for approximating new solutions by the use of the reduced basis model.

4. Solution methods

This section presents the theoretical background of the solution methods that are used for model reduction of the equilibrium equations that arise from the application of topology optimization using SIMP approach on linear structure analysis. These methods, as we mentioned earlier, project the state equations onto the subspace spanned by a reduced basis of small dimension.

4.1. Model Reduction via Proper Orthogonal Decomposition

4.1.1. Introduction

Proper orthogonal decomposition (POD) is a powerful and elegant method for data analysis aimed at obtaining low-dimensional approximate descriptions of a high-dimensional process. The POD provides a basis for the modal decomposition of an ensemble of functions, such as data obtained through experiments or numerical simulations. Its properties suggest that it is the preferred basis to use in various applications. The basis functions it yields are commonly called empirical eigenfunctions, empirical basis functions, empirical orthogonal functions, proper orthogonal modes, or basis vectors. The most striking feature of the POD is its optimality: it provides the most efficient way of capturing the dominant components of an infinite-dimensional process with only a finite number of "modes", and often surprisingly few "modes"[1][28]. In recent years, there have been many reported applications of the POD methods in engineering fields. The POD has been used in various disciplines including random variables, image processing, signal analysis, data compression, process identification and control in chemical engineering, oceanography, etc. [28]. In the bulk of these applications, the POD is used to analyze experimental data with the objective of extracting dominant features. The POD has been used to obtain approximate, low-dimensional descriptions of turbulent fluid flows [28], structural vibrations and chaotic dynamical systems, and more recently, microelectromechanical systems (MEMS).

4.1.2 Construction of the POD basis by Singular value decomposition

The main idea of the POD is to find a set of ordered orthonormal basis vectors in a subspace (without loss of generality, denoting the subspace as $\mathbb{R}^m$) where a random vector takes its values, such that the samples in the sample space can be expressed optimally using the selected first l basis vectors.
The mean square error can be used as a measure for the optimal problem, i.e.,

$$E\{\|x - x(l)\|^2\} \leq E\{\|x - \hat{x}(l)\|^2\} \tag{4.1}$$

where $x(l)$ is the approximate expression of a random vector x using the first l basis vectors of the undetermined set of orthonormal basis vectors, and $\hat{x}(l)$ is the approximate expression of x using arbitrary l basis vectors in $\mathbb{R}^m$.
The problem can be stated as follows.
Assume that $x \in \mathbb{R}^m$ is a random vector and $\{\varphi_i\}_{i=1}^m$ is a set of arbitrary orthonormal basis vectors: then x can be expressed as

$$x = \sum_{i=1}^{m} y_i \varphi_\iota = \Phi y \tag{4.2}$$

where

$$y_i = \varphi_i^T x \quad (i = 1,2, \dots, m)$$

$$y = (y_1, y_2, \dots, y_m)^T$$

$$\Phi = [\varphi_1, \varphi_2, \dots, \varphi_m]$$

The objective of the POD is to find a set of basis vectors that satisfies the following extreme value problem:

$$\begin{aligned} &\min_{\varphi_i} e^2(l) = E\{\|x - x(l)\|^2\} \\ &\text{s.t} \qquad \varphi_i^T \varphi_j = \delta_{\iota j} \qquad i, j = 1,2, \dots, m \end{aligned} \tag{4.3}$$

where

$$x(l) = \textstyle\sum_{i=1}^{l} y_i \, \varphi_i (l \leq m).$$

A method to realize the POD is the Singular Value Decomposition (SVD) which was established for real-square matrices in the 1870s by Beltrami and Jordan, for complex square matrices in 1902 by Autonne, and for general rectangular matrices in 1939 by Eckart and Young. SVD uses the singular-value decomposition to find the basis vectors satisfying the POD requirement in the sample space. The process for realizing the POD by using the SVD is stated as follows.

We are interested in reducing the linear system of Equations

$$Kx = f \tag{4.4}$$

Since POD tries to find a good approximation in the state, we are not interested in the output behavior. POD finds a projection matrix W to minimize $\|x - W\hat{x}\|$ for a given norm $\|.\|$ at least approximately. POD is a Galerkin projection method, which means we only have to find one matrix W. The method we describe is sometimes referred to as the method of snapshots. Within some communities the general underlying concept of POD is also called principal component analysis (PCA) or Karhuenen–Loeve decomposition. A set of N snapshots, samples are collected and collocated column-wise in a matrix Y. So The snapshot matrix Y has the following form $Y = [x_1, x_2, \dots, x_N]$.
Given samples, snapshots $x_1, x_2, \dots, x_N$ of any kind (they could be solutions or outcomes of a process in general), the method extracts a basis $u_1, u_2, \dots, u_l$ that solves the following minimization problem:

$$\min_{u_1,\dots,u_l} \sum_{k=1}^{N} \left\| x_k - \sum_{i=1}^{l} \langle x_k, u_i \rangle u_i \right\| \quad \text{s.t.} \quad \langle u_i, u_j \rangle = \delta_{ij} \tag{4.5}$$

The solution to this problem is directly connected to the singular value decomposition (SVD) of the matrix Y.

Singular value decomposition
Every matrix $A \in \mathbb{R}^{n\,x\,m}$ can be decomposed into a product of three matrices

$$A = U\Sigma V^T \in \mathbb{R}^{n\,x\,m} \tag{4.6}$$

where U and V are unitary (orthogonal) and $\Sigma = diag(\sigma_1, \dots, \sigma_2) \in \mathbb{R}^{n\,x\,m}$ is diagonal with nonnegative diagonal entries called *real singular values*: $\sigma_i = \sqrt{\lambda_i(A^T A)} \geq \sigma_{i+1}$, and the 2-induced norm of A is σ_1. The left singular vectors are $U = (u_1\ u_2\ \dots\ u_n)$, $UU^T = I_n$ and the right singular vectors are $V = (v_1\ v_2\ \dots\ v_n)$,. $VV^T = I_m$.The columns of U and V together with σ_i can be obtained by solving the eigenvalue problems

$$AA^T U_i = (\sigma_i)^2 U_i \tag{4.7}$$

And

$$A^T A V_i = (\sigma_i)^2 V_i \tag{4.8}$$

Finally, the *dyadic decomposition* of A is

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_n u_n v_n^T \tag{4.9}$$

Given the SVD of $Y = U\Sigma V^T$ the solution to (4.5) for the standard Euclidean inner product and Euclidean norm is obtained by the first l left singular vectors which are the first l columns of U. Now that the singular vectors u_i of the snapshot matrix $Y = [x_1, x_2, \dots, x_N]$ are available we have that

$$x \approx \sum_{i=1}^{l} \langle x, u_i \rangle u_i = \sum_{i=1}^{l} \hat{x}_i u_i = W\hat{x} \tag{4.10}$$

where W is given by the matrix $W = [u_1, u_2, \dots, u_N]$. We are not able to compute $\|x - W\hat{x}\|$ for every parameter values, but we minimize

$$\sum_k \|x_k - W\hat{x}_k\| \tag{4.11}$$

If the system is furthermore dependent on a parameter, this is typically extended by creating snapshots for several parameter values. All the snapshots will be put into one large matrix Y and the rest can be done as described above.

4.1.3 POD on topology optimization

In this section an approach is proposed for constructing the reduced basis in topology optimization with the use of POD using calculations of the state variables resulting from the high fidelity model.

The snapshot matrix Y mentioned earlier will contain the state variables of N solutions of the full system of Eq (4.4). By applying SVD on the matrix Y, the reduced basis W containing the N left singular vectors is formed. We then project the equilibrium equations on the basis W

$$W^T K W a = W^T F \tag{4.12}$$

and calculate the reduced state variables a, that is the coefficients of vector x expressed in the reduced basis W. From the coefficients a we obtain the approximate solution $\hat{x} = Wa$ whose accuracy is measured based on a threshold on the value of the residuals Eq (3.5). If the approximate solution does not satisfy this threshold a new solution of the high fidelity model is obtained, the snapshot matrix Y is enriched with this solution and at the same time we discard the oldest solution of the matrix Y. We are ready now to apply SVD on the snapshot matrix and obtain a new reduced basis W. The basic idea is that using previously calculated solutions to construction the reduced basis ensures relatively quick convergence whereas forming the reduced basis with randomly generated designs will not converge here due to the very high dimensionality of the problem (up to millions of input design variables). By doing so we keep the size of the reduced basis constant throughout the optimization procedure and thus make the whole procedure as efficient and fast as possible. The Algorithm POD is presented below.

Algorithm POD: The code below replaces the line " Ku=F" in a usual topology optimization algorithm.

```
if iteration = 1
    x ⟵ solution of Kx=F
    Add x to matrix Y
else if 1 < iteration ≤ N
    x ⟵ solution of Kx=F
    Add x to matrix Y
    if iteration=N
        Perform thin SVD on matrix Y (Y = UΣV^T)
        W⟵U
    end if
else
    α ⟵ solution of (W^T K W)α=W^T F
    e_rb ⟵ norm(K W α - F) / norm(F)
    if e_rb > ε
        Remove oldest basis vector from matrix Y
        x ⟵ solution of Kx=F
        Add u to matrix Y
        Perform thin SVD on matrix Y (Y = UΣV^T)
        W⟵U
    else
        x ⟵ Wα
    end if
end if
```

In the above algorithm lines 1 to 10 correspond to the initialization of the reduced basis. The reduced basis is calculated for the first time in line 9. This occurs when N solutions of the high fidelity model are acquired and SVD is performed on the snapshot matrix Y (line 8). After the computation of W, the approximate solution and the error of the residual are calculated (line 12 and 13 respectively). If the error is higher than the threshold ε then a new solution of the full model is calculates, the snapshot matrix is updated and finally a new reduced basis is computed (lines 14 - 19). If the error is lower than the threshold ε then the approximate solution is used in place of the exact solution. The *algorithm POD* has two parameters that are defined by the user. These are the reduced basis size N and the value of the threshold. The effect of the choice of these parameters will be investigated in the next chapter of this paper.

4.2 Model order reduction On The Fly

Christian Gogu in his paper [29] proposed an approach for constructing the reduced basis *on the fly* using previous calculations of the state variables which are orthogonalized and normalized by the Gram_Shimdt procedure. The redused basis is thus adaptively constructed and enriched, based on the convergence behavior of the topology optimization.

4.2.1 Basic idea of constructing the Reduced Basis on the fly

In a classical topology optimization process at iteration i, i displacement vectors have already been calculated (one at each iteration) by inverting the full equilibrium equations of Eq. 3.1. Making use of these i displacement vectors is the main idea, so as to form a reduced basis Φ which will be used for calculating the displacement vector at the next iteration. This means that the displacement vector, at iteration $i+1$ will be calculated using the reduced order model of Eq. 3.4, which calculates the reduced state variables at the current iteration $i+1$. So by solving the equilibrium equations projected on the subspace generated by the i previously calculated displacement vectors will give the approximate displacement vector at iteration $i+1$. At iteration $i+2$ a new approximation of the displacement vector at this iteration can still be calculated using the reduced order model with the same reduced basis Φ as before. This process can be applied until the approximate solution using the reduced order model is no longer sufficiently accurate, based for example on a threshold on the value of the residuals (Eq. 3.5).
Once the reduced order solution no longer satisfies this threshold a new full solution (inversion of the full system of Eq. 3) is computed again, solution which is also used

for enriching the reduced basis. To enrich the existing reduced basis $\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_m\}$ the newly calculated solution U_{new} is orthogonalized as shown in Eq. 4.2 (Gram-Schimdt procedure), normalized, then added to the basis.

$$\widetilde{\Phi}_{\iota+1} = U_{new} - \sum_{j=1}^{i} \langle U_{new}, \Phi_j \rangle \Phi_j \tag{4.13}$$

Where $\langle \bullet, \bullet \rangle$ denotes the L^2 scalar product.

The reduced basis uses solutions from previous iterations, and any new solution is only added to the basis after orthonormalization by the Gram Schimdt procedure. This is because only the component of the solution which is orthogonal to the already existing basis vectors is needed to be added to the basis.

The basic idea is that due to the very high dimensionality of optimization problems, reduced basis that are formed with randomly designs will not converge. But when using previously calculated solutions to construct the reduced basis ensures relatively quick convergence.

4.2.2 Construction of the Reduced Basis on the fly in topology optimization

The proposed iterative approach works with a fixed, user defined, reduced basis size, denoted N_b. The reduced basis is initialized using the full equilibrium solutions of first N_b iterations of the topology optimization process. At the first N_b iterations the full equilibrium equations of Eq. 3.1 are solved and the corresponding solution displacement vector obtained. These N_b solution vectors form the initial subspace of the reduced basis. In order to obtain the basis vectors of this subspace, the N_b solution vectors need to be orthonormalized. This is achieved by normalizing the first solution vector U_1, orthogonalize the subsequent basis vectors by the Gram Schmidt orthogonalization and finally normalize them as shown:

$$\Phi_1 = \frac{U_1}{\|U_1\|} \tag{4.14}$$

$$\widetilde{\Phi}_{i+1} = U_{i+1} - \textstyle\sum_{j=1}^{i} \langle U_{i+1}, \Phi_j \rangle \Phi_j \qquad \text{for} \qquad i = 1, \dots, N_b - 1 \tag{4.15}$$

$$\Phi_{i+1} = \frac{\widetilde{\Phi}_{i+1}}{\|\widetilde{\Phi}_{i+1}\|} \qquad \text{for} \qquad i = 1, \dots, N_b - 1 \tag{4.16}$$

After the first N_b iterations the reduced basis is initialized and the reduced basis solutions obtained via Eq. 3.5 will be used in the next topology optimization iterations, replacing the call to the full solution of the equilibrium equation (Eq. 3.4). In order to maintain appropriate convergence of the topology optimization, the quality of the reduced basis approximation at each iteration needs to be checked. This is done by calculating the value of the residuals (Eq. 3.5). If the value of the residual e_{rb} is lower than a user defined threshold ε, then the reduced basis solution is considered to be accurate enough to be used. If this criterion is verified, the approximate solution will replace the exact solution for the current optimization iteration of the topology optimization process (e.g. for calculating sensitivities, objective function, etc). Otherwise, if $e_{rb} > \varepsilon$ the reduced basis solution is not considered accurate enough and a new full solution is calculated by inverting the full equilibrium equations (Eq.3.4). The reduced basis is then updated using this new full solution vector. In order to maintain the size of the reduced basis the oldest vector in the reduced basis is discarded. Finally, to add the new solution vector to the reduced basis, it has to be is orthogonalized (Eq. 4.13) and normalized. The same process can then start over.

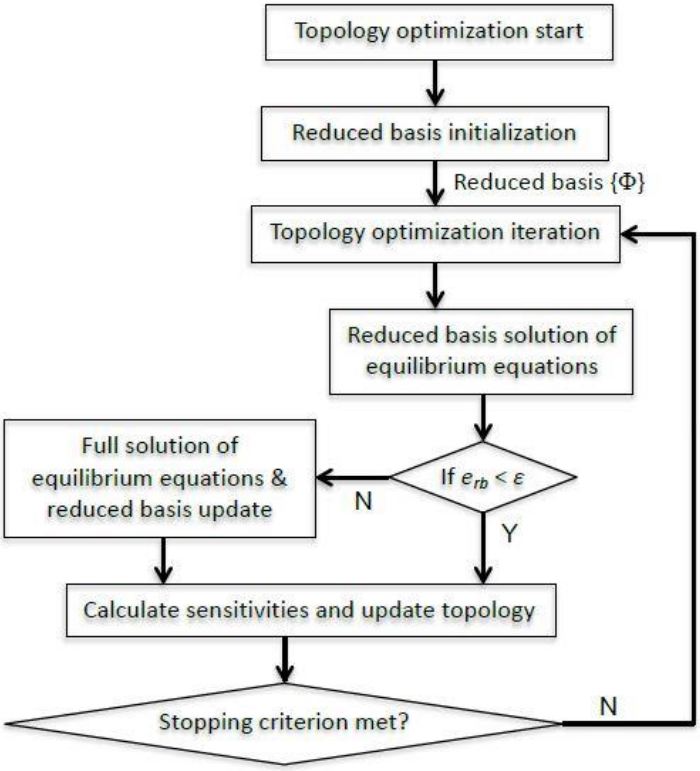


Figure 2: Flowchart of the reduced basis topology optimization process using on the *on the fly reduced basis construction*

The advantage of the proposed approach over the usual topology optimization process is that after the reduced basis initialization, at each iteration of the optimization process the reduced basis solution of the equilibrium equations is first

calculated. If the residuals criterion is verified then the approximate reduced basis solution is used at this iteration in place of the exact solution, meaning that significant computational savings can be achieved since obtaining the reduced basis solution is much quicker than obtaining the full solution.

From an implementation point of view the proposed approach is quite simple, since one can take the topology optimization code of his preference and just replace the line of the code that calculates the solution of the equilibrium equation (typically this line calculates the solution of $KU = F$) with the *Algorithm on the fly* presented below.

Algorithm on the fly: The code below replaces the line " KU=F" in a usual topology optimization algorithm.

```
 1: if iteration = 1
 2:     U ⟵ solution of KU=F
 3:     Add U / norm(U) to matrix Φ
 4: else if 1 < iteration ≤ N_b
 5:     U ⟵ solution of KU=F
 6:     U_ortho ⟵ U - Φ(Φ^T U)
 7:     Add U_ortho / norm(U_ortho) to matrix Φ
 8: else
 9:     α ⟵ solution of (Φ^T K Φ)α=Φ^T F
10:     e_rb ⟵ norm(K Φ α - F) / norm(F)
11:     if e_rb > ε
12:         Remove oldest basis vector from matrix Φ
13:         U ⟵ solution of KU=F
14:         U_ortho ⟵ U - Φ(Φ^T U)
15:         Add U_ortho / norm(U_ortho) to matrix Φ
16:     else
17:         U ⟵ Φα
18:     end if
19: end if
```

In algorithm on the fly, lines 1-7 correspond to the reduced basis initialization. From iteration 1 to *Nb* of the topology optimization, the full solutions of the equilibrium equation $KU = F$ are calculated. Note that here, *iteration* denotes the number of the iteration (or cycle) of the topology optimization routine and that the matrix of the basis vectors Φ is the empty matrix before the first iteration. It is also recalled that "$A \leftarrow B$" means assign to variable *A* the value of expression *B*.

Once the reduced basis initialized, the reduced basis solution is first calculated (line 9). The residual is calculated on line 10. If it is below the threshold ε, then the reduced basis solution is used in place of the exact solution (line 17) for the remaining calculations (sensitivity, objective function) of the current topology

optimization iteration. If the residual is higher than the threshold then the full solution is calculated (line 13) and the reduced basis updated.
In usual topology optimization codes, the line that is being replaced here by Algorithm 1 provides the exact displacement vector solution *U* satisfying the equilibrium equations. Algorithm 1 also provides a displacement vector *U* as an output. This displacement vector is sometimes the exact solution and sometimes only the reduced basis solution, if the reduced basis approximation is found to be acceptable in terms of accuracy. Since the reduced basis solution is much cheaper to evaluate than the full solution, the frequency with which only the reduced basis solutions can be used will determine the computational savings of the proposed method.
The Algorithm on the fly has only two parameters that are defined by the user, the size of the reduced basis N_b and the threshold ε on the residuals. The effect of the choice of these parameters will be investigated in chapter 5 of this paper.

4.3 Approximate Reanalysis by Combined Approximations

Abstract. As we have already mentioned, in the nested approach to structural optimization, most of the computational effort is invested in the solution of finite element analysis equations. In the study of O. Amir [30] the integration of an approximate reanalysis procedure into the framework of topology optimization of continuum structures is investigated.

4.3.1 Introduction

The aim of using approximate reanalysis is to reduce the computational effort involved in repeated solutions of the equilibrium equations, which for large problems will dominate the computational cost of the whole process. In the study of O. Amir [30], approximate reanalysis is performed following the Combined Approximation (CA) approach for linear static reanalysis, originally proposed by Kirsch [31] for linear static reanalysis. The main feature of CA is the utilization of a local series expansion that form a reduced basis. By using CA for repeated structural analysis, the number of required factorizations of the stiffness matrix is reduced, thus removing a significant portion of the computational effort. It has been shown that the CA method is theoretically equivalent to the preconditioned conjugate gradient method [32]. Therefore, available results from one method, such as convergence criteria and error bounds, can also be used in the other method
Amir in his study examined the integration of linear static reanalysis by CA into the widely used procedures for topology optimization using a distribution of material approach. The investigation was focused on two aspects — accuracy, meaning the

possibility to obtain the same topological designs as would be generated by a standard procedure; and efficiency, meaning that the approximate procedure should offer significant savings in computation time when compared with a standard procedure. In the next Sections we will review the CA approach for structural reanalysis and present the way CA method for Approximate reanalysis in topology optimization is formulated.

4.3.2 Reanalysis by CA

In this section we will present the formulation of linear reanalysis by CA.

The linear system of equilibrium equations resulting from a finite element (FE) discretization of the computational domain corresponding to a certain iterative step of the optimization process is

$$Ku = f \tag{4.17}$$

where $\boldsymbol{K}$ is the stiffness matrix, $\boldsymbol{u}$ is the unknown displacement vector, and $\boldsymbol{f}$ is the external force vector. We assume that $\boldsymbol{f}$ does not depend on the optimization and remains constant throughout the whole process. Instead of solving the full system of equations, it is possible to find an efficient approximation of the solution $\tilde{u}$ that will be accurate enough for the purpose of optimization

$$\tilde{u} \approx u \tag{4.18}$$

$$\mathrm{K\tilde{u}} = \mathrm{f} \tag{4.19}$$

The approximate solution according to the CA approach is obtained as follows. We denote $\boldsymbol{K_0}$ as the stiffness matrix corresponding to a certain previous optimization step and given in its factorized form. The equation system (4.17) can be rewritten as

$$(\mathrm{K}_0 + \Delta\mathrm{K})\mathrm{u} = \mathrm{f} \tag{4.20}$$

and hence

$$\mathrm{K}_0\mathrm{u} = \mathrm{f} - \Delta\mathrm{K}\mathrm{u} \tag{4.21}$$

Here, $\boldsymbol{\Delta K}$ represents the matrix of changes in stiffness due to changes in the values of the design variables. Defining the following recurrence relation

$$\mathrm{K_0 u^k = f - \Delta K u^{k-1}} \tag{4.22}$$

leads to the so-called binomial series expansion

$$u = (I - B + B^2 - B^3 + \cdots)u_1 \tag{4.23}$$

Where

$$u_1 = K_0^{-1} f \tag{4.24}$$

$$B = K_0^{-1} \Delta K \tag{4.25}$$

$$u_i = -B u_{i-1} \tag{4.26}$$

It is important to note that the first term u_1 is already known from a previous optimization step and the following terms u_i can be easily computed by forward and backward substitutions based on the available factorization of $\mathbf{K_0}$. The main feature of CA is the utilization of the series terms from (4.23) as basis vectors which compose the reduced basis. Considering only the first *s* series terms, the approximate solution can now be expressed as

$$\tilde{u} = y_1 u_1 + y_2 u_2 + \cdots + y_s u_s = R_B y \tag{4.27}$$

where R_B is an *n*×*s* matrix (redused basis matrix) containing the basis vectors $u_1, u_2, \ldots, u_s$ and y is a vector of s unknowns. Replacing u in (4.17) with $\tilde{u}$ from (4.27) and premultiplying both sides by $\mathrm{R_B^T}$

$$\mathrm{R_B^T K R_B y = R_B^T f} \tag{4.28}$$

This ends up in a reduced system of equations, with s equations instead of n

$$\mathrm{K_R y = f_R} \tag{4.29}$$

Where

$$\boldsymbol{K_R} = \mathrm{R_B^T K R_B} \tag{4.30}$$

$$\mathrm{f_R} = \mathrm{R_B^T f} \tag{4.31}$$

It is important to point out that the computational benefit of this procedure is due to the fact that the basis vectors are easily generated by solving a linear system where the stiffness matrix is already given in its factorized form. The computational cost of the reanalysis procedure is much smaller than the cost of a complete new analysis, which is dominated by the cost of a matrix factorization.

4.3.3 CA method for Approximate reanalysis in topology optimization

We will now present the implementation of the CA method for approximate reanalysis into a topology optimization procedure. The issue will be addressed for minimum compliance problems, aiming at finding the optimal distribution of elastic material in a certain domain so that the stiffest structure is obtained.
When applying the so-called nested approach, where optimization is performed in the design variables only and where the equilibrium equations are treated as function calls, the optimization problem is formulated as follows:

$$\begin{aligned} \min_{\rho} \quad & c(\rho) = f^T u \\ \text{s.t:} \quad & \textstyle\sum_{e=1}^{N} u_e \rho_e \leq V \\ & 0 < \rho_{min} \leq \rho_e \leq 1\,, \qquad e = 1, \dots, N \\ \text{with:} \quad & K(\rho) = f \end{aligned} \tag{4.32}$$

where $\mathbf{K}(\boldsymbol{\rho})$ is the stiffness matrix whose entries depend on the design variables $\boldsymbol{\rho}$. When using the SIMP interpolation and element-wise constant densities scheme, $\mathbf{K}(\boldsymbol{\rho})$ can be rewritten as a sum over all elements

$$K(\rho) = \sum_{e=1}^{N} \rho_e K_e \tag{4.33}$$

where K_e is a standard element stiffness matrix referring to an element density equal to 1.
When applying approximate reanalysis, this formulation accurately represents the problem only for optimization cycles in which the full equation system is solved to

satisfy the structural equilibrium. For each optimization cycle in which an approximate reanalysis is performed by CA, an appropriate optimization problem is formulated in terms of the approximate solution of the equilibrium equations, $\tilde{u}$. The objective function can be expressed as follows:

$$c(\rho) = f^T u \approx \tilde{u}^T K(\rho)\tilde{u} \tag{4.34}$$

$$\tilde{u} = R_B y \tag{4.35}$$

$$c(\rho) = y^T R_B^T K(\rho) R_B y \tag{4.36}$$

When performing approximate reanalysis within a certain optimization cycle, a set of $\boldsymbol{s}$ basis vectors is generated and utilized according to the CA procedure. The size of this set is determined in the reanalysis stage and is then fixed for the sensitivity analysis. In this case the optimization problem will have the form
It can be seen that the nested equations are the reduced reanalysis system of Equation (4.28) and the equation systems governing the generation of the basis vectors, based on the available factorization of the stiffness matrix $K_0(\rho_0)$ corresponding to a previous design cycle. In this nested problem, the matrix R_B contains the basis vectors

$$\mathbf{R}_\mathrm{B} = [\mathrm{u}_1, \dots, \mathrm{u}_\mathrm{s}] \tag{4.38}$$

and the matrix $\mathbf{K}(\rho)$ is split into two parts, the first corresponding to a previous factorization and the second to the changes in stiffness due to changes in the design

$$K(\rho) = K_0(\rho_0) + \Delta K(\rho, \rho_0) \tag{4.39}$$

When using the SIMP interpolation and element-wise constant densities scheme, both matrices can be rewritten as a sum over all elements

$$K_0(\rho_0) = \sum_{e=1}^{N} \rho_{e,[0]}^p K_e \tag{4.40}$$

$$\Delta K(\rho, \rho_0) = \sum_{e=1}^{N} (\rho_e^p - \rho_{e,[0]}^p) K_e \tag{4.41}$$

The optimization problem is formulated in terms of the original basis vectors $\mathbf{R}_{\mathrm{B}}$ and the corresponding reduced solution y since this simplifies the sensitivity analysis.

Computational considerations

Solving an optimization problem in which the actual stiffness matrix is factorized only once within a certain number of design iterations can be seen as solving a series of 'short' optimization problems, each one of them a valid optimization problem on its own. Each 'short' optimization problem begins with a matrix factorization and therefore an accurate design iteration; then it continues with a sequence of approximate design iterations based on an approximate reanalysis.

The main goal of using approximate reanalysis is to achieve an accurate result efficiently. When considering the accuracy of the proposed procedure, the somewhat limited scope of the 'short' problems should be taken into account. On the one hand, it is clear that an approximate procedure cannot accommodate extremely large changes in stiffness, hence these should be bounded in some way. This results in the use of an outdated physical model, so that in many cases the approximate problem cannot reach the true optimum and will converge to a higher value (assuming minimization is considered). On the other hand, in certain cases the approximate solution is not accurate enough and the optimization process leads to unreliable results, sometimes with a better objective value than the final optimum. Owing to these shortcomings, a certain frequency of updating the factorized matrix should be defined if we seek a stable convergence leading to the same optimum as found by solving the full problem. Considering the efficiency of the procedure, matrix factorizations are the most expensive part so that the number of updates should be minimized. Therefore, the key for achieving an accurate result efficiently is choosing the right time to stop a sequence of reanalyses and perform a new factorization. Possible options for controlling the procedure could be:

- State a fixed frequency of matrix factorizations, which will be performed regardless of the convergence of the 'short' problem.

- Perform a new matrix factorization when the current design variable vector $\boldsymbol{\rho}_{[k]}$ is significantly different from the design variable vector corresponding to the factorized matrix, $\boldsymbol{\rho}_{[0]}$. This can be done by examining the angle between the two vectors.

- Perform a new matrix factorization when the 'short' problem reaches a certain convergence criterion. Such a criterion could be, for example, the

relative change in the value of the objective function within the 'short' problem.

In addition to the frequency of matrix factorizations, the size of the reduced basis also has a direct impact on the accuracy and efficiency of the procedure. Therefore, it is important to state a certain criterion that will determine the sufficient number of basis vectors. In this study, we examine the relative magnitude of the residual forces (the error due to approximation) for this purpose. If the relative magnitude of this residual (measured by the ratio between the Euclidean norms of the residual and the external force vector) exceeds a certain permitted tolerance value, another basis vector is generated, until the residual satisfies the criterion or the predefined maximum number of basis vectors is reached.

4.4 Sensitivity analysis of the proposed Reduced Basis methods

4.4.1 Sensitivity adjustment

The approaches that were just presented have the advantage of being simple and straight forward to implement. Nevertheless they make the assumption that the exact solution can be replaced by the reduced basis solution at some iterations of the topology optimization process without negatively affecting the convergence or the solution of the topology optimization. This assumption can be reasonable depending on the choice of the user defined parameters. Nevertheless in the present subsection a different approach, is presented, taking into account the fact that only an approximate solution is being sometimes used.

4.4.2 Sensitivity adjustment for POD and on the fly approach

In order to update the topology at each iteration, the topology optimization process uses the sensitivity of the objective function (compliance) with respect to density variations. The classical expression of the sensitivity is:

$$\frac{\partial c}{\partial \rho_e} = -U^T \frac{\partial K}{\partial \rho_e} U \tag{4.42}$$

When using the reduced order models presented in the previous subsections, the displacement vector *U* is sometimes an approximation (the reduced basis solution $U_{RB} = \Phi\alpha$), this expression of the sensitivity is thus only approximate. For this reason we seek to correct the value of the sensitivity by using the adjoint method [3] to take into account the approximation error.
We define the following modified objective function, by adding additional terms that are equal to zero:

$$c(\rho_e) = a^T\Phi^T K\Phi\alpha - 2\tilde{\alpha}^T(\Phi^T K\Phi\alpha - \Phi^T F) - \sum_{i=1}^{N_b} \lambda_i^T (K_i U_i - F) \qquad (4.43)$$

Where $\tilde{\alpha}$, λ_i (with $i = 1, \dots, N_b$) are adjoint variables and K_i and $U_i$are the stiffness matrix and displacement fields corresponding to the i[th] basis vector.
The modified expression of the sensitivity is obtained by differentiating Eq. (4.43) with respect to the elemental density:

$$\begin{aligned}\frac{\partial c}{\partial \rho_e} = {} & 2\frac{\partial a^T}{\partial \rho_e}\Phi^T K\Phi\alpha + 2a^T\frac{\partial \Phi^T}{\partial \rho_e}K\Phi\alpha + 2\tilde{\alpha}^T\frac{\partial \Phi^T}{\partial \rho_e}F \\ & - 2\tilde{\alpha}^T\frac{\partial \Phi^T}{\partial \rho_e}K\Phi\alpha - 2\tilde{\alpha}^T\Phi^T\frac{\partial K}{\partial \rho_e}\Phi\alpha - 2\tilde{\alpha}^T\Phi^T K\frac{\partial \Phi}{\partial \rho_e}\alpha \\ & - 2\tilde{\alpha}^T\Phi^T K\Phi\frac{\partial \alpha}{\partial \rho_e} - \sum_{i=1}^{N_b}\lambda_i^T\frac{\partial K_\iota}{\partial \rho_e}U_i - \sum_{i=1}^{N_b}\lambda_i^T K_\iota\frac{\partial U_i}{\partial \rho_e}\end{aligned} \qquad (4.44)$$

We chose the adjoint variable $\tilde{\alpha} = \alpha$ such as to eliminate derivatives of the solution vectors. Furthermore, we can define the variable of the residual vectors as expressed in Eq. 13.

$$\Delta F = F - K\Phi\alpha \qquad (4.45)$$

Moreover by applying the equality $2a^T\frac{\partial \Phi^T}{\partial \rho_e}\Delta F = \sum_{i=1}^{N_b} 2\alpha_i\frac{\partial {U_i}^T}{\partial \rho_e}\Delta F$ we obtain the following simplified expression of the sensitivity:

$$\frac{\partial c}{\partial \rho_e} = -a^T\Phi^T\frac{\partial K}{\partial \rho_e}\Phi\alpha - \sum_{i=1}^{N_b}\lambda_i^T\frac{\partial K_\iota}{\partial \rho_e}U_i - \sum_{i=1}^{N_b}\frac{\partial {U_i}^T}{\partial \rho_e}(K_\iota\lambda_i - 2a_i\Delta F) \qquad (4.46)$$

In order to eliminate the derivatives of the displacement vectors we define the adjoint variables λ_i as the solutions of Eq. 15.

$$K_\iota \lambda_i = 2a_i \Delta F \qquad i = 1, \dots, N_b \tag{4.47}$$

Accordingly we obtain the final expression of the corrected sensitivity:

$$\frac{\partial c}{\partial \rho_e} = -a^T \Phi^T \frac{\partial K}{\partial \rho_e} \Phi \alpha - \sum_{i=1}^{N_b} \lambda_i^T \frac{\partial K_\iota}{\partial \rho_e} U_i \tag{4.48}$$

We note that the expression of the corrected sensitivities includes two terms. The first one is the sensitivity calculated on the approximate solution. This is the only term considered in the sensitivity of the POD and on the fly approach described earlier. The second term is the adjustment term, which corrects the sensitivity by taking into account that an approximate, reduced basis solution was used instead of the true solution of the problem. This correction has the potential to allow higher residuals for the approximate reduced basis solution because the sensitivity is corrected to account for this residual.

When correcting the sensitivity we have the advantage of accounting for the approximation error. This means that fewer full simulations are likely to be required. On the downside this approach is more computationally expensive since it requires to calculate the adjoint variables. Even though this additional cost is decreased by storing the factorized stiffness matrices, required for the adjoint variables calculation, the number of adjoint variables calculations is large (N_b times the number of reduced basis iterations). This means that the approach may only be efficient for very large systems for which the cost of factorizing the stiffness matrix is large compared to the other costs, including that of backsubstitution, given the factorized stiffness matrix.

4.4.3 Sensitivity adjustment for CA approach

As was foretold when a full analysis is performed within a certain optimization cycle, the sensitivity is obtained using the adjoint method leading to the well-known expression

$$\frac{\partial c}{\partial \rho_e} = -u^T \frac{\partial K}{\partial \rho_e} u \tag{4.49}$$

Within the optimization cycles, in which approximate reanalysis is performed, the reduced solution in terms of the original basis vectors is obtained according to Equation.4.35 and then the corresponding sensitivity analysis can be performed by the adjoint method. Introducing adjoint variables $\tilde{y}$, λ_i (*i* =1*, . . . , s, s* being the number of basis vectors used in the reanalysis), the following modified objective function is obtained by adding zero terms:

$$c(\rho_e) = y^T R_B^T K R_B y - 2\tilde{y}^T (R_B^T K R_B y - R_B^T f) - \lambda_i^T (K_0 u_1 - f) - \sum_{i=2}^{s} \lambda_i^T (K_0 u_i + \Delta K u_{i-1}) \tag{4.50}$$

When differentiating the objective function with respect to a certain design variable, derivatives of the adjoint variables are obviously eliminated and the derivative will be

$$\begin{aligned} \frac{\partial c}{\partial \rho_e} = {} & 2\frac{\partial y^T}{\partial \rho_e} R_B^T K R_B y + 2y^T \frac{\partial R_B^T}{\partial \rho_e} K R_B y + y^T R_B^T \frac{\partial K}{\partial \rho_e} R_B y \\ & + 2\tilde{y}^T \frac{\partial R_B^T}{\partial \rho_e} f - 2\tilde{y}^T \frac{\partial R_B^T}{\partial \rho_e} K R_B y - 2\tilde{y}^T R_B^T \frac{\partial K}{\partial \rho_e} R_B y \\ & - 2\tilde{y}^T R_B^T K \frac{\partial R_B}{\partial \rho_e} y - 2\tilde{y}^T R_B^T K R_B \frac{\partial y}{\partial \rho_e} - \lambda_i^T \frac{\partial K_0}{\partial \rho_e} u_1 \\ & - \lambda_i^T K_0 \frac{\partial u_1}{\partial \rho_e} \\ & - \sum_{i=2}^{s} \lambda_i^T \left[\frac{\partial K_0}{\partial \rho_e} u_i + K_0 \frac{\partial u_i}{\partial \rho_e} + \frac{\partial \Delta K}{\partial \rho_e} u_{i-1} \right. \\ & \left. + \Delta K \frac{\partial u_{i-1}}{\partial \rho_e} \right] \end{aligned} \tag{4.51}$$

As is the case for the standard (full) problem, the adjoint variable vector should be chosen such that the derivatives of the solution vector are eliminated, is simply the solution vector itself, meaning that for the approximate problem

$$\tilde{y} \equiv y \tag{4.52}$$

By defining the vector of residual forces due to approximation errors

$$\Delta f = f - K R_B y \tag{4.53}$$

and applying the following equalities:

$$2y^T \frac{\partial R_B^T}{\partial \rho_e} \Delta f = \sum_{i=1}^{s} 2y_i \frac{\partial u_i^T}{\partial \rho_e} \Delta f \tag{4.54}$$

$$\frac{\partial K_0}{\partial \rho_e} = 0 \tag{4.55}$$

$$\frac{\partial \Delta K}{\partial \rho_e} = \frac{\partial K}{\partial \rho_e} \tag{4.56}$$

the following expression for the sensitivity is obtained as:

$$\begin{aligned} \frac{\partial c}{\partial \rho_e} = & -y^T R_B^T \frac{\partial K}{\partial \rho_e} R_B y - \sum_{i=2}^{s} \lambda_i^T \frac{\partial \Delta K}{\partial \rho_e} u_{i-1} \\ & - \sum_{i=2}^{s-1} \frac{\partial u_i^T}{\partial \rho_e} [K_0 \lambda_i + \Delta K \lambda_{i+1} - 2 y_i \Delta f] \\ & - \frac{\partial u_s^T}{\partial \rho_e} [K_0 \lambda_s - 2 y_s \Delta f] \end{aligned} \tag{4.57}$$

It can be seen that the derivatives of the basis vectors can be eliminated by solving the adjoint problems, where $i = 1, \ldots, s-1$:

$$K_0 \lambda_s = 2 y_s \Delta f \tag{4.58}$$

$$K_0 \lambda_i = 2 y_i \Delta f - \Delta K \lambda_{i+1} \tag{4.59}$$

Therefore, the final expression for the sensitivity is

$$\frac{\partial \mathrm{c}}{\partial \rho_\mathrm{e}} = -\mathrm{y}^\mathrm{T} \mathrm{R}_\mathrm{B}^\mathrm{T} \frac{\partial \mathrm{K}}{\partial \rho_\mathrm{e}} \mathrm{R}_\mathrm{B} \mathrm{y} - \sum_{\mathrm{i}=2}^{\mathrm{s}} \lambda_\mathrm{i}^\mathrm{T} \frac{\partial \Delta \mathrm{K}}{\partial \rho_\mathrm{e}} \mathrm{u}_{\mathrm{i}-1} \tag{4.60}$$

It can be seen that if there are no approximation errors, then the first term is identical to the sensitivity in the standard problem and the second term is eliminated. Once approximation errors occur, they are accounted for in the sensitivity analysis, meaning that relatively large inaccuracies could be tolerated. The number of adjoint problems to be solved is $s-1$ since the first basis vector does not depend on changes in the design; hence, the additional computational cost is very low considering the availability of the factorized stiffness matrix.

5. Numerical Tests

Several numerical results are presented in this chapter, and the effectiveness of the three approximate approaches described earlier is investigated. The addressed problems are minimum compliance problems in two and three dimensions. From the results we can see that quite accurate results can be obtained when using the approximate procedures and promising computational savings are achieved due to the reduction in the number of the required factorizations of the stiffness matrix.

5.1. Description of topology optimization problem number 1

To start with, we examine the impact of the proposed approximate methods in a classical minimum compliance benchmark problem, the MBB-beam in 2D. The boundary conditions of the beam are represented in Figure 5. Due to symmetry conditions only the half of the structure is modeled. The main task is to find the optimal material distribution inside the beam for minimum compliance given an upper bound on the material volume fraction.

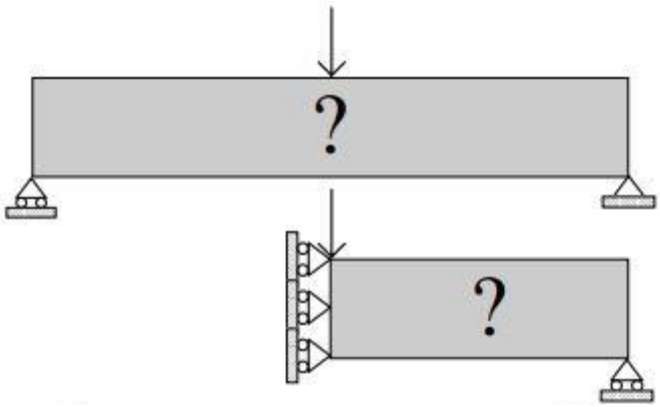


Figure 5.1: MBB-beam problem. Top: full design domain, bottom: half design domain, with symmetry boundary conditions

As far as the value of the parameters of the problem are considered we chose:

Nominal Young's modulus	:	$E_{nominal} = 1$
Minimum Young's modulus	:	$E_{min} = 1 * 10 - 9$
Poisson's ratio	:	$\nu = 0.3$
Penalization factor	:	$p = 3$
Maximum allowable volume fraction	:	$volfrac = 0.5$
Density filter radius	:	$rmin = 1.5$

The optimization iterations are stopped either when the density variation within any of the elements is less than 1% or when the total number of iterations has reached

the maximum number of allowable iterations, whichever criterion is met first. The code implementation used was adapted from Andreassen et al. [2].

5.1.2 Results using the three proposed approximation methods

To examine the performance of the described approaches we made use of two different mesh sizes, a low density and a high density mesh of a 2D MBB beam Problem. The low density mesh consists of 150 x 50 elements and the high density mesh is of 600 x 200 elements, that is 15.402 dofs and 241.602 dofs respectively.
The study focuses on the impact the user define parameters have in the accuracy and efficiency of the optimization problem. For this reason different values of the parameters ε (residuals threshold) and N_b (size of the reduced basis) are taken under consideration.

In Table 1 we present the results of the topology optimization for the 2D MBB beam problem we obtained from the low density mesh of 150 x 50 elements. The first line of the table provides the results obtained by running the topology optimization of the high fidelity model. The proceeding lines provide the results of the topology optimization obtained when the reduced order model approaches are applied. In order to examine the efficiency me took under consideration the time speedup of each approximate approach, which is the CPU time of the full order model topology optimization solution divided by the CPU time of the reduced order model approach and the results are presented in the third column. The forth column provides the total number of iterations in the topology optimization routine. This is either the number of iterations until the convergence criterion was met or until the total number of available iterations (400 here) was reached. The fifth column provides the number of full solutions of the equilibrium equations that had to be computed throughout the optimization. To examine the accuracy of each approximate approach we calculated the relative error in the objective function at the optimum for the approximate solution obtained using the reduced order models, compared to the solution of the high fidelity model(last column). This error (given in %) is expressed in Eq. (5.1).

$$err = \frac{c_{approx} - c_{exact}}{c_{exact}} \times 100 \tag{5.1}$$

Where c_{exact} represents the value of the objective function (compliance) at the optimal design.

Table 5.1. Results obtained from the different approximations approaches for different parameters for a 150 x 50 mesh density.

Residuals threshold,ε	Size of reduced basis, N_b	Time speedup	Number of iterations	Number of full solutions	Error in objective function (%)
High Fidelity Model					
-	-	1	400	400	0
POD					
0.1	4	1.39	400	99	-0,00757
0.1	10	1.31	400	97	-0,01262
0.1	40	1.21	400	103	-0,00929
On the fly					
0.1	4	1.19	400	116	-0,02202
0.1	10	1.14	400	117	-0,02076
0.1	40	1.11	400	107	-0,01293
AR					
0.1	4	1.11	400	100	0,14435
0.1	10	0.9	400	160	0,05904
POD					
0.05	4	1.32	400	115	-0,01060
0.05	10	1.17	400	114	-0,01091
0.05	40	1.16	400	118	-0,00934
On the fly					
0.05	4	1.18	400	143	-0,01025
0.05	10	1.14	400	143	-0,01005
0.05	40	1.11	400	131	-0,01071
AR					
0.05	4	1,15	400	94	0,04159
0.05	10	1,00	400	182	-0,00904
POD					
0.01	4	1.21	400	173	-0,00439
0.01	10	1.11	400	167	-0,00288
0.01	40	1.05	400	167	-0,00318
On the fly					
0.01	4	1.04	400	263	-0,00379
0.01	10	1.03	400	263	-0,00389
0.01	40	0.99	400	230	-0,00434
AR					
0.01	4	1,09	400	110	0,06429
0.01	10	0.94	400	177	0,04068

Table 5.1 gives us much information about the efficiency and accuracy of the proposed approximate models. First of all reduced order models give good solutions. The accuracy of these solutions is not affected from the size Nb of the reduced basis and so we can conclude that relatively small values for the reduced basis size *Nb* are already sufficient to obtain good solutions. In other words the size has very little impact on the error in the final value of the objective function. As far as the efficiency of the ROMs is concerned, it is obvious that the size of the reduced basis has a clear impact on computational cost (time speedup), with larger bases requiring more CPU time. A size of 4 appears enough to obtain good accuracy and topology optimization results (Fig 5.2). Figure 5.2 provides the obtained topologies for the high fidelity model as well as for the proposed reduced order modeling approaches when using the user defined parameters ε = 0.05 and N_b = 4. The images of Figure 5.2 are direct plots of the density value of each pixel, without any other post-processing. For all the different values examined of the user defined parameters the topology results were similar which makes the difference between the approximate solutions and the solution of the full order model undistinguishable to the naked eye.

From Table 1 we can also see that the residuals threshold ε affects both the efficiency and the accuracy of the approximate approaches. With lower values of ε the efficiency of the ROMs is decreased significantly whereas the accuracy gains are minor. The value of the residuals threshold has such an impact in the efficiency of a ROM because it implicitly defines the number of the full solutions of the equilibrium equations need to be calculated. More specifically at the beginning of the optimization only full solutions of the equilibrium equations are calculated until the residuals go below the residuals threshold and the first reduced order solution is taken as a sufficient solution. So a lower threshold means more full solutions of the equilibrium equations at the beginning of the optimization. From then on, time to time the residuals exceed once again the threshold, thus requesting a new full solution, which will be used to update the reduced basis and thus guarantee that the reduced basis solutions will again be acceptable over a few iterations. This will be done more often for a lower value of the residuals threshold.

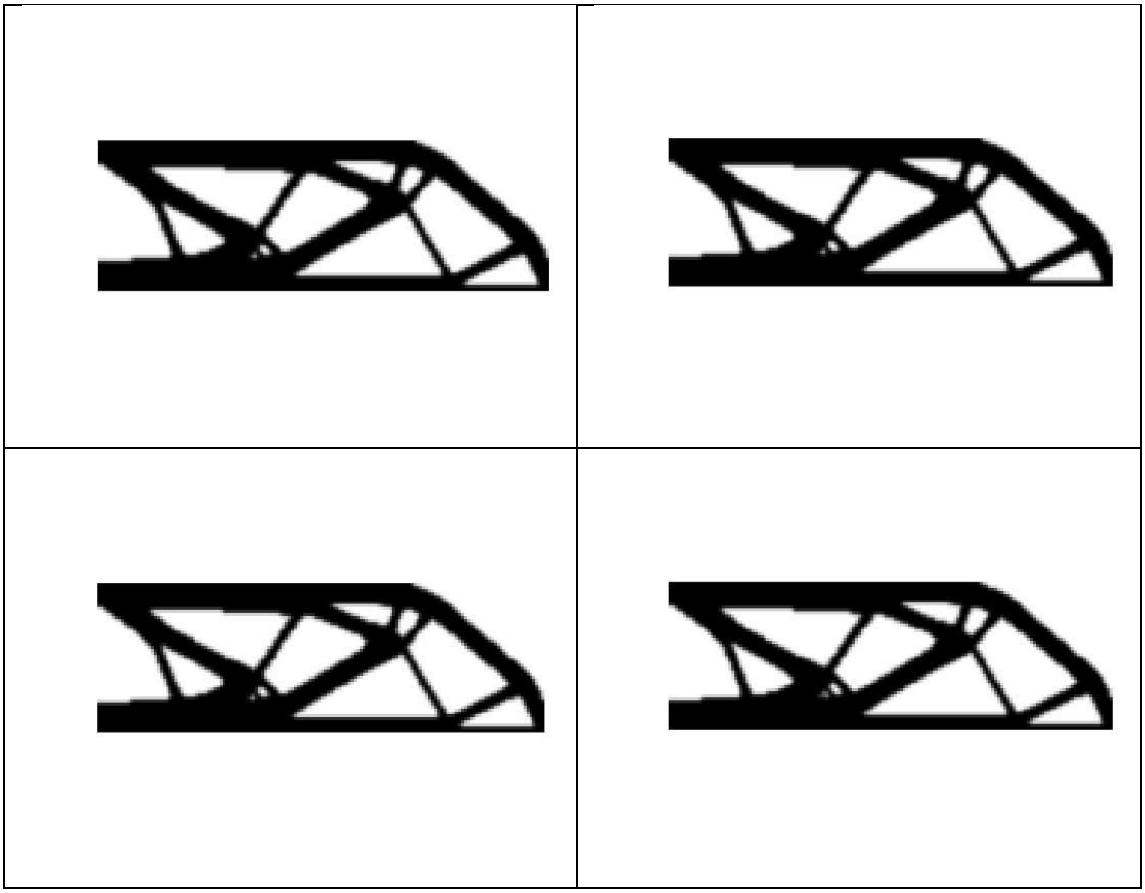


Figure 5.2: Optimal topologies with a 150x50 mesh size for : a) the high fidelity topology optimization(upper left), b) the On the fly approach with $\boldsymbol{\varepsilon = 0.05}$ and $\boldsymbol{N_b = 4}$ (upper right), c) the POD approach with $\boldsymbol{\varepsilon = 0.05}$ and $\boldsymbol{N_b = 4}$ (lower left) d) the CA approach with $\boldsymbol{\varepsilon = 0.05}$ and $\boldsymbol{N_b = 4}$ (lower right).

We can see that the maximum speedup is 1.39 when using the approximate models in the above optimization problem. We expect to achieve higher speedups in larger scale problems. So a similar investigation was carried out using a high density mesh of 600 x 200 elements (241602 dofs). The user define parameters that are examined are ε=0,1 e=0.05 and Nb=4, Nb=10.

The computational results are provided in Table 2 and the topologies obtained are presented in Figure 5.3.

Table 5.2 Relative computational cost associated to the different approximations approaches for different parameters for a **600 x200** mesh density.

Residuals threshold,ε	Size of reduced basis, N_b	Time speedup	Number of iterations	Number of full solutions	Error in objective function (%)
High Fidelity Model					
-	-	1	400	400	0
POD					

0.1	4	1,40	400	131	-0,02738
0.1	10	1,37	400	125	-0,02436
On the fly					
0.1	4	1,36	400	150	-0,01933
0.1	10	1,31	400	152	-0,05422
CA					
0.1	4	1,26	400	97	-0,16073
0.1	10	1,08	400	90	-0.00463
POD					
0.05	4	1,30	400	171	-0,0082
0.05	10	1,28	400	165	-0,00825
On the fly					
0.05	4	1,27	400	188	-0,01132
0.05	10	1,25	400	188	-0,01007
CA					
0.05	4	1,24	400	107	-0,05901
0.05	10	1,06	400	77	-0,00221

From analysis of Table 5.2 we can make the following comments regarding the choice of the user defined parameters N_b and ε.

We can now make a safer conclusion that the reduced basis size N_b has an impact on the efficiency but hardy any on the accuracy of the final result. It appears to be more beneficial to use a low reduced basis size. As we see a reduced basis size of 10 compared with that of size 4, degrades the efficiency of the problem The problem is that a too small reduced basis (i.e. size of one or two) may be counterproductive even in terms of efficiency since the residual threshold will be exceeded more often with such a small reduced basis, thus requiring more full simulations. So going much lower than 4 is not recommended since the efficiency gains would be negligible and for some problems the small size could be problematic. A size of four seems a safe starting point.

As for the residuals threshold ε it appears again to have a bigger effect in the efficiency than in the accuracy of the approximate approaches. Lower values of the threshold imply a decrease in the efficiency. Making a choice of value of the residual threshold for a given problem seems to be more difficult than choosing the basis size, because the threshold is much more problem dependent. In the MBB Beam topology problem a value of 0.1 for the threshold ε gives sufficiently accurate results.

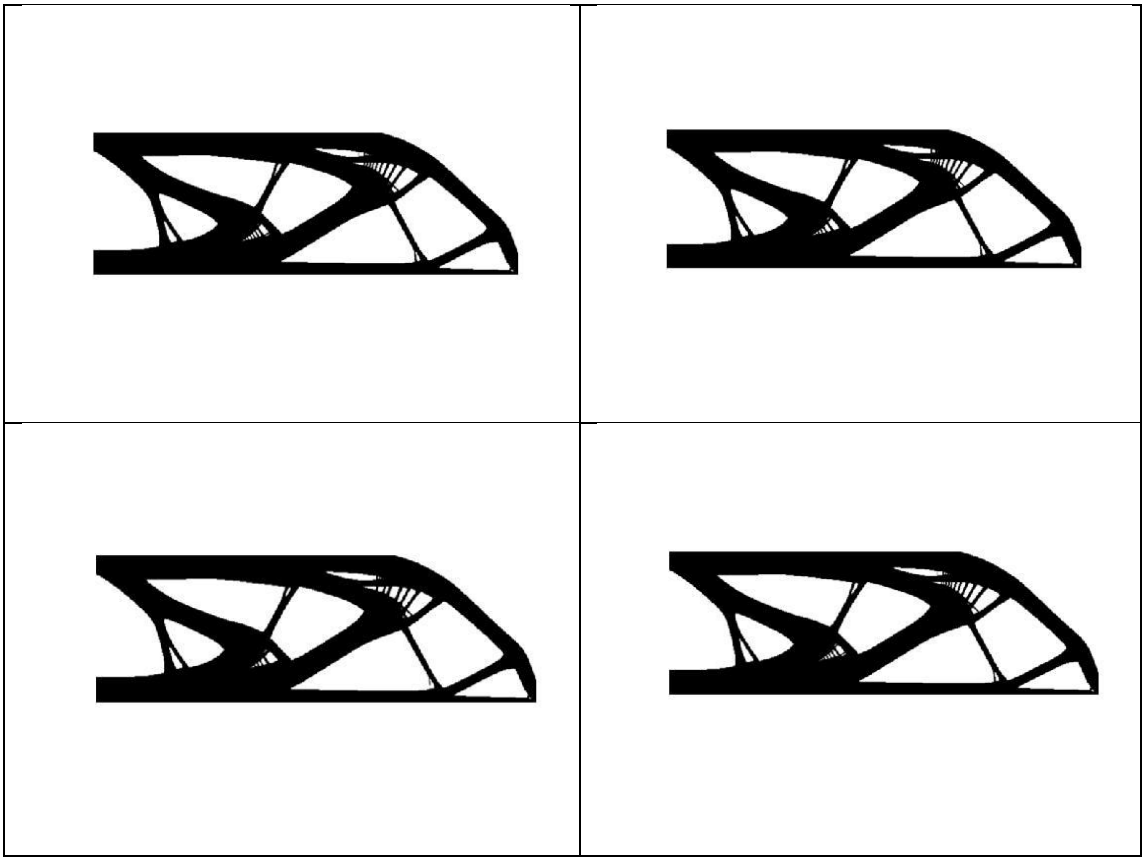


Figure 5.3: Optimal topologies with a 600x200 mesh size for : a) the high fidelity topology optimization(upper left), b) the On the fly approach with $\boldsymbol{\varepsilon = 0.05}$ and $\boldsymbol{N_b = 4}$ (upper right), c) the POD approach with $\boldsymbol{\varepsilon = 0.05}$ and $\boldsymbol{N_b = 4}$ (lower left) d) the CA approach with $\boldsymbol{\varepsilon = 0.05}$ and $\boldsymbol{N_b = 4}$ (lower right).

In the topologies in Figure 5.3 occur some very small topological details, which are slightly different between the various solutions. A way to mitigate these fine topological details is to increase the filter radius.

5.2 Description of topology optimization problem number 2

To further examine how the performance of the described approaches scale up with the size of the problems for this topology optimization problem we consider a classical minimum compliance benchmark problem of a 3D Cantilever beam Fig. (5.4). The code implementation used was adapted from Liu and Tovar [9].

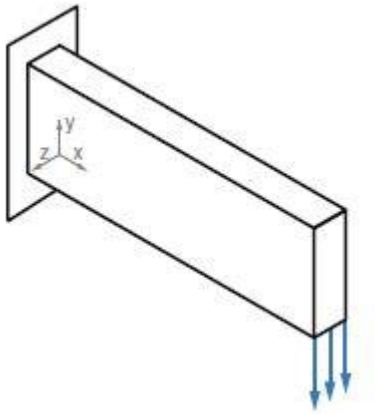


Figure 5.3: 3D Cantilever beam problem

The general parameters used for topology optimization problem of the 3D Cantilever beam are:

Nominal Young's modulus	:	$E_{nominal} = 1$
Minimum Young's modulus	:	$E_{min} = 1 * 10 - 9$
Poisson's ratio	:	$\nu = 0.3$
Penalization factor	:	$p = 3$
Maximum allowable volume fraction	:	$volfrac = 0.2$
Density filter radius of	:	$rmin = 1.5$

5.2.1 Results using the three proposed approximation methods

To examine the performance of the approximate approaches we made use of two different mesh sizes, a low density and a high density mesh of a 3D Cantilever beam Problem. The low density mesh consists of 60x20x4 elements (19215 dofs) and the

high density mesh is of 100x40x8 elements (111807 dofs). The maximum number of iterations is set to 300.
We start by investigating the results for different values of the user define parameters in the 60 x 20 x 4 mesh density which are presented in Table 5.3.

Table 5.3. Relative computational cost associated to the different approximations approaches for different parameters for a 60 x 20 x 4 mesh density.

Residuals threshold,ε	Size of reduced basis, N_b	Time speedup	Number of iterations	Number of full solutions	Error in objective function (%)
High Fidelity Model					
-	-	1	300	300	0
POD					
0.1	4	1,28	300	231	-0,03347
0.1	10	1,28	300	218	-0,01763
On the fly					
0.1	4	1,17	300	272	-0,01569
0.1	10	1,19	300	262	-0,0261
CA					
0.1	4	1,29	300	48	0,012443
0.1	10	1,14	300	51	-0,01763
POD					
0.05	4	1.21	300	146	-0.01722
0.05	10	1.24	300	137	-0.0105
On the fly					
0.05	4	1.11	300	199	-0.01756
0.05	10	1.22	300	197	-0.01529
CA					
0.05	4	1.28	300	50	0.004179
0.05	10	1.15	300	54	-0.00075

As we can see the maximum speed-up is value is 1.29 and only 48 factorizations of the stiffness matrix were needed. Also the optimized topologies of the different approximate approaches have no difference compared with the topology of the high fidelity model is distinguished Fig. (5.3).

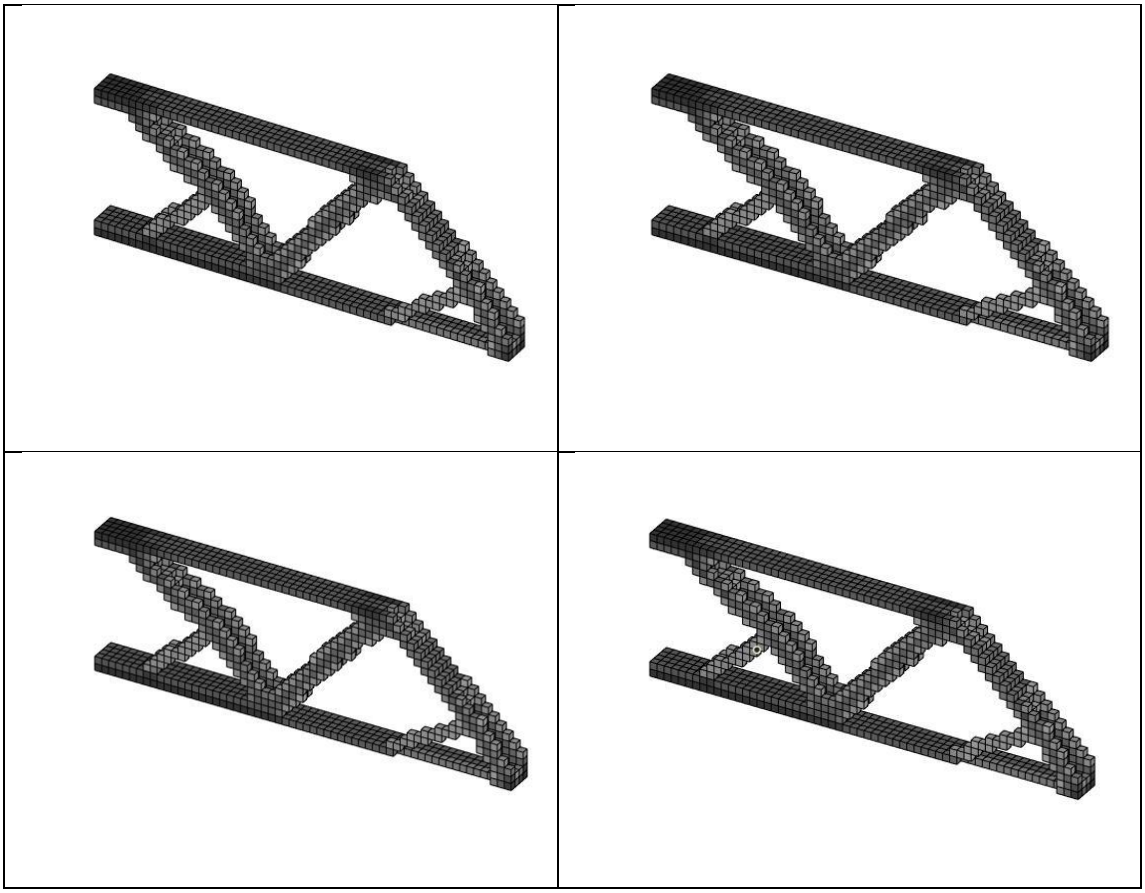


Figure 5.3: Optimal topologies with a 60 x 20 x 4 mesh size of a 3D cantilever beam problem of a) the high fidelity model (upper left), b) the POD approach with ε = 0.05 and *Nb* = 4 (upper right), c) the On the fly approach with ε = 0.05 and *Nb* = 4 (lower left), d) the CA approach with ε = 0.05 and *Nb* = 4(lower right).

In order to further investigate how the computational savings of the proposed methods scale up with the size of the problems we run the topology optimization for the same 3D beam topology optimization problem with a 100x40x8 mesh. The same parameters are used as before. In table 5.2 we present the numerical results.

Table 5.4 Numerical results for the 3D problem with a100x40x8 mesh density

Residuals threshold,ε	Size of reduced basis, N_b	Time speedup	Number of iterations	Number of full solutions	Error in objective function (%)
-	-	1	200	200	0
POD					
0.05	4	1.62	200	86	-0.00873
On the fly					
0.05	4	1.42	200	109	-0.01222
CA					
0.05	4	2.01	200	38	-0.00734

Comparing the results obtained from the two different meshes we can see that the computational savings continue to grow with the size of the problem. In the high density mesh a speed-up of 2.01 occurs whereas in the low density mesh the maximum speedup is 1.29.

In figure 5.4 the optimized topologies of the different approximate approaches are presented. No difference compared with the topology of the high fidelity model is distinguished.

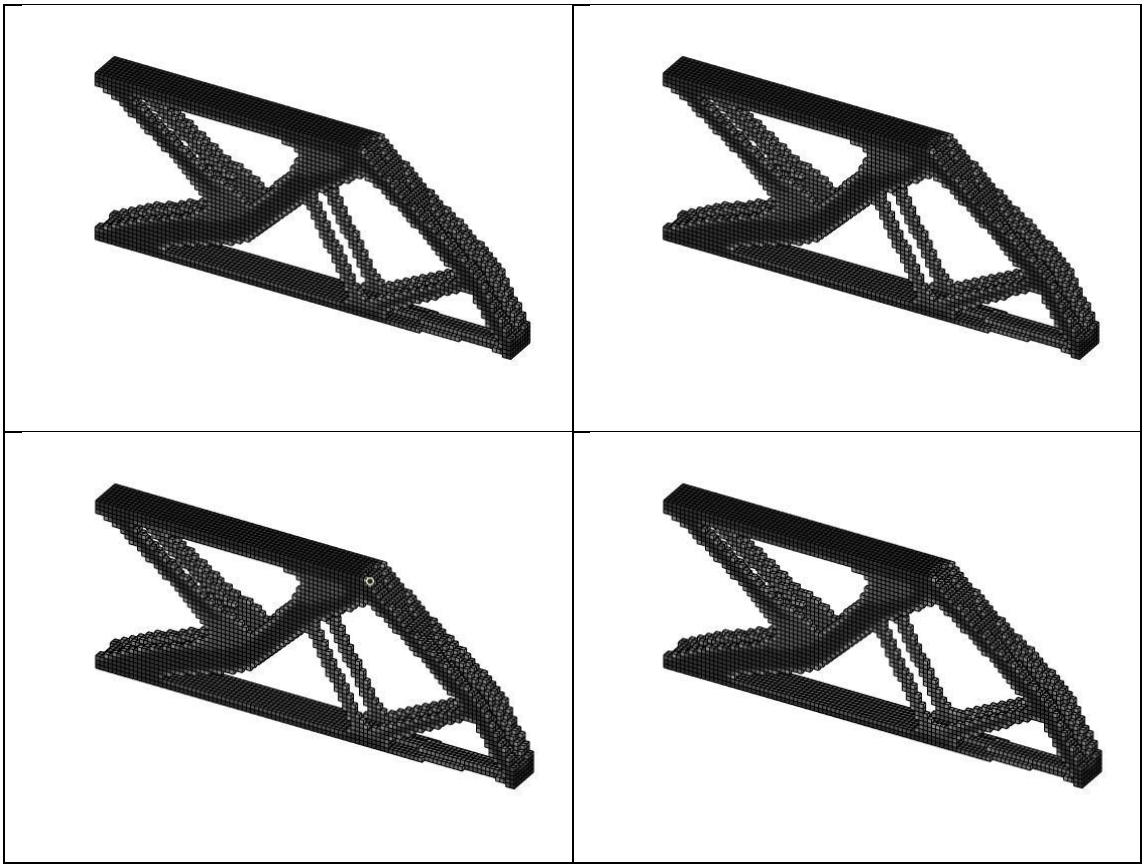


Figure 5.4: Optimal topologies with a 100x40x8 mesh size of a 3D cantilever beam problem of a) the high fidelity model (upper left), b) the POD approach with ε = 0.05 and Nb = 4 (upper right), c) the On the fly approach with ε = 0.05 and Nb = 4 (lower left), d) the CA approach with ε = 0.05 and Nb = 4(lower right).

5.3 Description of topology optimization problem number 3

In this topology optimization problem a more complex 3D problem, known as the 3D wheel problem, is considered to further investigate the scaling of the computational cost savings for the approximate approaches described. Figure 5.5 illustrates the boundary conditions of this problem. The solid is simply supported at the four corners and a downward force is applied at the center of its bottom face. A mesh of 40x20x40 is considered in this test case which accounts for 105.903 degrees of freedom in the finite element model. The code implementation used was adapted from Liu and Tovar [9].

The general parameters used for topology optimization problem of the 3D Cantilever beam are:

Nominal Young's modulus	:	$E_{nominal} = 1$
Minimum Young's modulus	:	$E_{min} = 1 * 10 - 9$
Poisson's ratio	:	$\nu = 0.3$
Penalization factor	:	$p = 3$
Maximum allowable volume fraction	:	$volfrac = 0.2$
Density filter radius of	:	$rmin = 1.5$

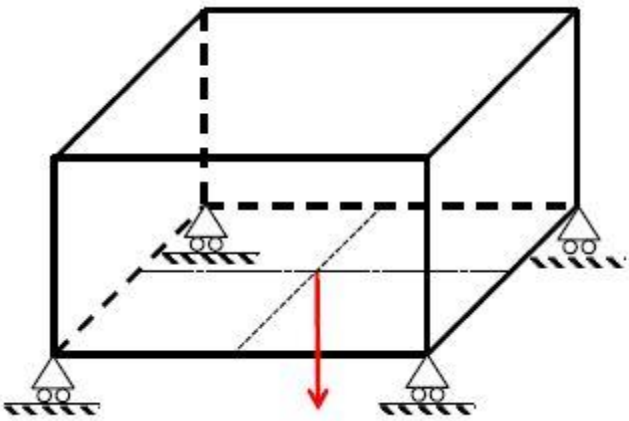


Figure 5.5: Boundary conditions for the 3D wheel problem

5.3.1 Results using the three proposed approximation methods

The numerical results of the proposed reduced order models with a reduced basis size of $N_b = 4$ and a residuals threshold of $\varepsilon = 0.05$ are presented in Table 5.5

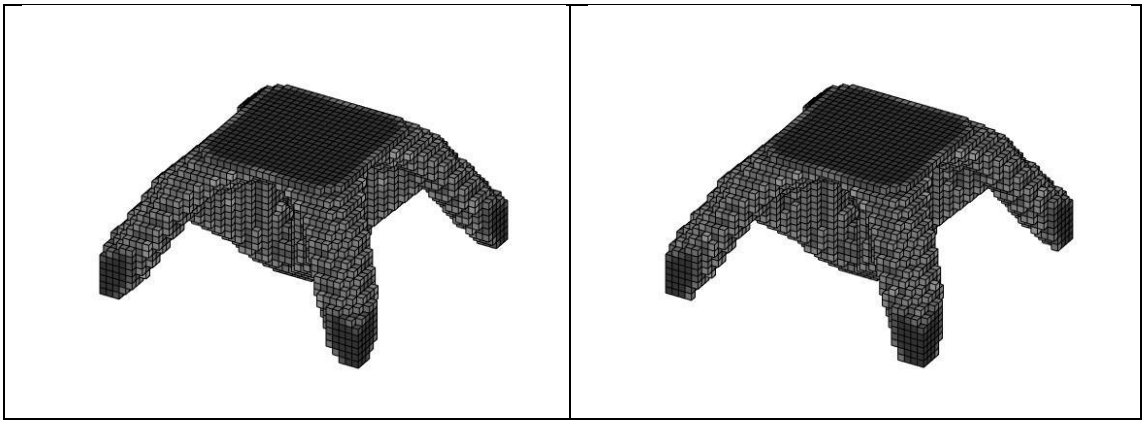
and the results of the reference topology optimization problems are presented in Figure 5.6. The maximum number of iterations in this analysis is set to 178.

Table 5.6. Numerical results for the 3D problem with a 40 x 20 x 40 mesh

Residuals threshold,ε	Size of reduced basis, N_b	Time speedup	Number of iterations	Number of full solutions	Error in objective function (%)
High Fidelity Model					
-	-	1	178	178	0
POD					
0.05	4	5.3	178	13	-0.13711
On the fly					
0.05	4	5.03	178	15	-0.10016
CA					
0.05	4	5.03	148	18	-0.04215

In this topology optimization problem the approximate approaches give an accurate optimized topology of up to 5.3 times faster than the full order analysis. This is because only 13 out of 178 factorizations of the stiffness matrix are required for the solution via the POD approach.

In figure 5.4 the optimized topologies of the different approximate approaches are presented. No significant difference compared with the topology of the high fidelity model is distinguished.



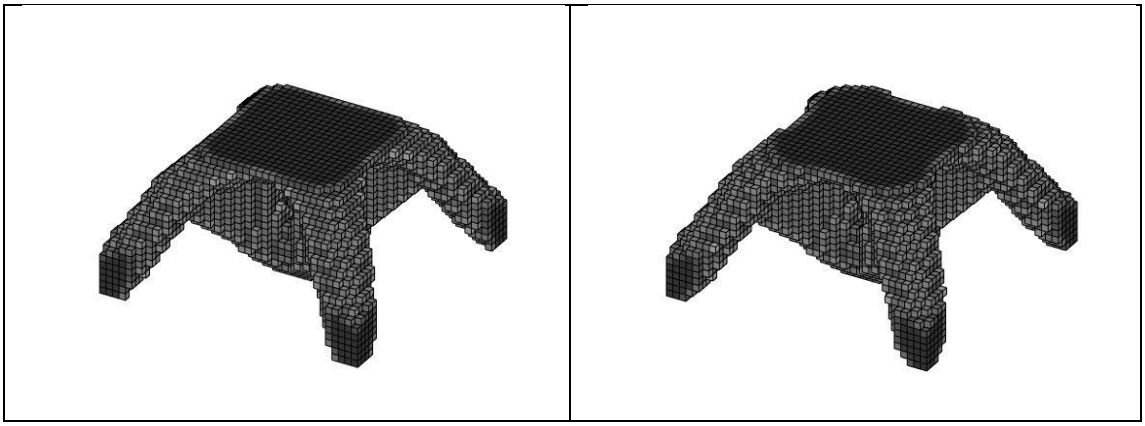


Figure 5.4: Optimal topologies with a 40x20x40 mesh size of a 3D cantilever beam problem of a) the high fidelity model (upper left), b) the POD approach with ε = 0.05 and Nb = 4 (upper right), c) the On the fly approach with ε = 0.05 and Nb = 4 (lower left), d) the CA approach with ε = 0.05 and Nb = 4(lower right).

6. Conclusions

In this thesis we examined the impact of reduced order models in topology optimization problems. More precisely the reduced basis approaches of POD, On the fly and Approximate Reanalysis were described and utilized in minimum compliance structural topology optimization problems. Different test cases were taken under consideration and conclusions of the efficiency and accuracy of the reduced order models for different user defined parameters were derived. The numerical results showed that for a small size reduced basis large speed-ups over topology optimization without reduced order modeling are obtained. Speedups up to a factor of 5 were achieved and this speedup can be increased with the size of the problem.

As a final conclusion we can say that the larger the problem and the more time consuming a single solution of the equilibrium equations will be, the higher the efficiency gains allowed by calculating the full solutions at only some iterations while using quick reduced order models at the other.

References

[1] A. CHATTERJEE 2000 *Current Science* 78, 808}817. An introduction to the proper orthogonal decomposition.

[2] E. Andreassen et al*. Efficient topology optimization in MATLAB using 88 lines of code*. English. *Structural and Multidisciplinary Optimization* 43.1 (2011), 1-16.

[3] M. P. Bendsoe and O. Sigmund. *Topology optimization : theory, methods and applications*. English. 2.;2; Berlin: Springer-Verlag, 2003;2004;

[4] M. P. Bendsoe and O. Sigmund. Material interpolation schemes in topology optimization. English. *Archive of Applied Mechanics* 69.9 (1999), 635-654.

[5] P. W. Christensen et al. An introduction to structural optimization. English. Vol. 153. Dordrecht: Springer, 2009;2008;

[6] O. Sigmund and J. Petersson. Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima. English. *Structural Optimization* 16.1 (1998), 68-75.

[7] K. Svanberg. *The method of moving asymptotes|a new method for structural optimization.* English. *International Journal for Numerical Methods in Engineering* 24.2 (1987), 359-373.

[8] Robin Larson *Methodology for Topology and Shape Optimization: Application to a Rear Lower Control Arm* 2016.

[9] Kai Liu and Andres Tovar *An efficient 3D topology optimization code written inMatlab* 50 (2014), 1175-1196.

[10] Wilhelmus H. A. Schilders and Henk A. van den Vorst Model Order Reduction: Theory, Research Aspects and Applications.

[11] Bruns TE, Tortorelli DA. “Topology optimization of non-linear elastic structures and compliant mechanisms”. *Comput Methods Appl Mech Eng*, 190:3443–3459, 2001.

[12] Bourdin B, “Filters in topology optimization.” *Int J Num Meth Eng*, 50:2143–2158, 2001.

[13] Bendsøe M (1989) Optimal shape design as a material distributionproblem. Struct Optim 1:193–202.

[14] Díaz A, Sigmund O (1995) Checkerboard patterns in layout optimization. Struct Optim 10(1):40–45.

[15] Jog C, Haber R (1996) Stability of finite element models for distributed-parameter optimization and topology design. Comput Methods Appl Mech Eng 130(3–4):203–226.

[16] Sigmund O, Petersson J (1998) Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. Struct Optim 16(1): 68–75.

[17] Sigmund O (2007) Morphology-based black and white filters for topology optimization. Struct Multidisc Optim 33(4–5):401–424.

[18] P. Kerfriden, P. Gosselet, S. Adhikaric, and S.P.A. Bordas. “Bridging proper orthogonal decomposition methods and augmented newton–krylov algorithms : An adaptive model order reduction for highly nonlinear mechanical problems.” *Comput. Meth. Appl. Mech*. *Eng.*, 200(5–8) :850–866, 2011.

[19] P. Krysl, S. Lall and J.E. Marsen. “Dimensional model reduction in non-linear finite element dynamics of solids and structures.” *International Journal for Numerical Methods in Engineering*, 51:479:504, 2001

[20] D. Ryckelynck, “A priori hypereduction method : an adaptive approach”, *Journal of Computational Physics*, 202(1):346 - 366, 2005.

[21] K. Kunisch and S. Volkwein. “Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics.” *SIAM Journal on Numerical Analysis*, 40(2) :492–515, 2002.

[22] K. Karhunen. “Uber lineare methoden für wahrscheinigkeitsrechnung.” *Annales of Academic Science Fennicae Series A1 Mathematical Physics*, 37:3-79, 1943.

[23] M.M. Loeve. *Probability Theory*. Van Nostrand, Princeton, NY. 1955.

[24] H. Hotteling. “Analysis of Complex of statistical variables into principal component.” *Journal of Educational Psychology*, 24:417--441, 1933.

[25] P. Ladeveze, J.-C. Passieux, and D. Neron. “The LATIN multiscale computational method and the proper generalized decomposition.” *Comput. Methods Appl. Mech. Eng.*, 199(21) :1287—1296, 2009.

[26] Gosselet, P., “Méthode de décomposition de domaines et méthodes d’accélération pour les problèmes multichamps en mécanique non-linéaire”, PhD thesis, Université Paris 6, 2003.

[27] Craig R.R., M.C.C Bampton, “Coupling of substructures for Dynamic Analyses”, *AIAA Journal*, 6(7):1313-1319, 1968.

[28] P. HOLMES, J. L. LUMLEY and G. BERKOOZ 1996 ‘H*urbulence*, *Coherent Structures*, *Dynamical Systems and Symmetry*. Cambridge: Cambridge University Press.

[29] Christian Gogu, “*Improving the efficiency of large scale topology optimization through on the fly reduced order model construction”,* Université de Toulouse.

[30] Oded Amir, Martin P.Bendsoe and Ole Sigmund, “*Approximate reanalysis in topology optimization*”, Int. J. Numer. Meth. Engng., 78:1474-1491, 2009

[31] Kirsch U. Reduced basis approximations of structural displacements for optimal design. *AIAA Journal* 1991; **29**:1751–1758.

[32] Kirsch U, Kocvara M, Zowe J. Accurate reanalysis of structures by a preconditioned conjugate gradient method. *International Journal for Numerical Methods in Engineering* 2002; **55**:233–251.

[33] A.T. Patera and G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, Version 1.0, Copyright MIT 2006,

[34] Bernard Haasdonk and Mario Ohlberger *Reduced basis method for finite volume approximations of parametrized linear evolution equations* 2008; 42:277-302.