

RESEARCH ARTICLE

Open Access



Stress-constrained topology optimization using approximate reanalysis with on-the-fly reduced order modeling

Manyu Xiao^{1*}, Jun Ma¹, Dongcheng Lu¹, Balaji Raghavan² and Weihong Zhang³

*Correspondence:
manyuxiao@nwpu.edu.cn

¹Xi'an Key Laboratory of Scientific Computation and Applied Statistics, School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710072, China

²Laboratoire de Génie Civil et Génie Mécanique EA 3913, Institut National des Sciences Appliquées de Rennes, Rennes, France

³State IJR Center of Aerospace Design and Additive Manufacturing, Northwestern Polytechnical University, Xi'an 710072, China

Abstract

Most of the methods used today for handling local stress constraints in topology optimization, fail to directly address the non-self-adjointness of the stress-constrained topology optimization problem. This in turn could drastically raise the computational cost for an already large-scale problem. These problems involve both the equilibrium equations resulting from finite element analysis (FEA) in each iteration, as well as the adjoint equations from the sensitivity analysis of the stress constraints. In this work, we present a paradigm for large-scale stress-constrained topology optimization problems, where we build a multi-grid approach using an on-the-fly Reduced Order Model (ROM) and the p-norm aggregation function, in which the discrete reduced-order basis functions (modes) are adaptively constructed for adjoint problems. In addition to reducing the computational savings due to the ROM, we also address the computational cost of the ROM learning and updating phases. Both reduced-order bases are enriched according to the residual threshold of the corresponding linear systems, and the grid resolution is adaptively selected based on the relative error in approximating the objective function and constraint values during the iteration. The tests on 2D and 3D benchmark problems demonstrate improved performance with acceptable objective and constraint violation errors. Finally, we thoroughly investigate the influence of relevant stress constraint parameters such as the p norm factor, stress penalty factor, and the allowable stress value.

Keywords: Stress constraints, Topology optimization, PCA, ROM, Adjoint equations

Introduction

Mechanical stress is a critical factor that affects performance, service life, fatigue resistance and safety of structural components, and is inarguably an important design criterion. Traditional topology optimization formulations (such as the minimum compliance problem with a volume constraint [1–3]) typically do not consider stress constraints, which could result in the phenomenon of high stress concentrations, leading to a final "optimized" design that all too often fails to meet real engineering requirements, eventually requiring manual adjustment or shape optimization as a subsequent step.

In the past two decades, many researchers have worked on stress-constrained topology optimization [4,5] in order to ensure that the design results are closer to engineering reality. However, when compared with the stress-constraint-free problem, stress-constrained topology optimization design presents a unique set of difficulties, which may be broadly classified as follows:

1. The stress singularity phenomenon. This was first studied for the truss optimization problem. Cheng and Jiang [6,7] gave a mathematical explanation for the singularity in the topology optimization problem, attributing it to the discontinuity of the stress constraint. Duysinx and Bendsøe [8] linked stress singularity to the phenomenon of non-vanishing stresses as the design variables tended towards zero, in other words, a region with a very small relative density could end up with a non-zero strain which then gives rise to non-zero (or even very high) stress, despite the fact that it represents a hole (void area). Today, the singularity problem is typically resolved by adopting the-relaxation approach [7–9], QP-relaxation method [10] and stress penalization [11].
2. Large number of local stress constraints. Since stress is a local parameter, the greater the number of elements for the domain discretization, the more there will be stress evaluation points, resulting in a large number of local non-convex stress constraints, which in turn will lead to a steep rise in computational cost for the sensitivity analysis of local stress constraints [4,12]. To reduce the computational complexity, local stress constraints are usually transformed into a global stress constraint or some form of cluster-based stress constraints by using a condensation function that approximates the value of the maximum local function. Currently, the two most common types of condensation functions are the Kreisselmeier-Steinhauser function (K-S) [4,12–16] and the P-norm function [9,16–18], and they can merge a mass of local stresses into a single constraint formula.
3. Highly nonlinear nature of the stress constraint. Any change in density value of the adjacent region will significantly affect the stress level within some key regions [13,17] such as depressions and corners. This requires the optimization procedure to be able to effectively reduce or eliminate the stress concentration phenomenon, and that the solution algorithm must maintain numerical consistency with the optimization procedure in order to ensure stable convergence of the overall procedure [19,20].

In view of the three broad classes of problems associated with stress-based topology optimization listed above, researchers have put forward a variety of effective methods to deal with them.

However, with modern-day sophisticated numerical methods and the recent development of advanced manufacturing technologies, the performance requirements for mechanical parts in the aerospace industry have become increasingly demanding [21]. Modeling with highly refined finite element models typically results in a series of ultra-large-scale problems with degrees of freedom ranging from tens of millions to even billions. Given the presence of non-convex, non-linear stress constraints in large-scale topology optimization, the main hindrances then are the solution of large-scale linear systems during the FEA analysis (equilibrium equation) in each iteration, as well as the sensitivity analysis of stress constraint functions due to the non self-adjoint nature [22]. Stress coag-

ulation alone is then inadequate to improve the design, meaning that we need another approach [23].

One highly effective and proven solution is Approximate Reanalysis (AR) using a Reduced Order Model (ROM) technique [24–27]. Due to the non-self-adjointness of the problem, we need to solve two high-dimensional systems during each iteration of the optimization process, the first being the equilibrium equations from finite element analysis (FEA), the other being the adjoint equations for the sensitivity analysis of the stress constraints. To the authors' knowledge, virtually no work has been reported on coupling stress-constrained topology optimization with approximate reanalysis using a ROM.

In this work, we first apply two previously proposed ROM approaches to the stress-constrained topology optimization problem, the first being on-the-fly ROM using Principal Components Analysis (PCA) [28] and the other being multi-fidelity ROM using the multigrid method (MG) [29]. In addition, based on the proposed multi-fidelity ROM, we develop two improved strategies for further reductions in computational costs, the first one being the Adap-ROM where the residual threshold for the approximate solution of the equilibrium equations can be adaptively selected according to the relative error of the objective function value during the iteration; the second strategy being Bi-ROM where we simultaneously apply the ROM to both the equilibrium equations and adjoint equations (during the sensitivity analysis of the stress constraint). Instead of the classical Optimality Criteria (OC) method, the Method of Moving Asymptotes (MMA) [30] is used as the optimization algorithm in this work, and the P-norm function in [9] is adopted for coagulating the local stress constraints to reduce the computational cost involved. The stress singularity phenomenon is avoided by using stress penalization as in [11]. Recent years there are more papers about stress constraint [31–33].

This paper is organized in the following manner: in “[Theoretical formulation](#)”, the theoretical formulation is presented with a classical minimum compliance problem constrained by both stress and volume, followed by adjoint equation-based sensitivity analysis of both the objective function and the stress constraint function. “[Projection-based ROM approaches](#)” first introduces two projection-based ROM approaches including on-the-fly ROM construction using PCA and multi-fidelity ROM construction using the multigrid method, then combines both approaches into a single algorithm. In “[Two improved strategies](#)”, we propose two improved strategies (i.e. Adap-ROM and bi-ROM). “[Numerical tests](#)” details the numerical investigations using a benchmark 2D L-shaped beam example. In “[3D L-shaped beam](#)”, we first apply a variety of ROM methods to the 3D compliance minimization problem of L-shaped beam, and then analyze the impact of the p norm factor p , the stress penalty factor q and the allowable maximum stress value $\hat{\sigma}$ on the design results. The paper ends with discussion, concluding comments and recommendations for future work.

Theoretical formulation

Topology optimization model

The essence of topology optimization is mathematically determining the distribution of material in a geometrically fixed design domain under a set of constraints, such that the resulting structure achieves some optimal combination of mechanical performance objectives. Here, we consider the classical minimum compliance problem under stress

and volume constraints, expressed as follows:

$$\begin{aligned} \min_{\rho} c(\rho) &= \mathbf{f}^T \mathbf{u} = \mathbf{u}^T \mathbf{K} \mathbf{u} \\ \text{s.t.} \quad &\begin{cases} \mathbf{K}(\rho) \mathbf{u} = \mathbf{f} \\ v_f = \frac{\sum_{e=1}^{N_e} \rho_e V_e}{V_0} \leq \hat{v}_f \\ \sigma_{\max}^{vM} = \max(\sigma_e^{vM}) \leq \hat{\sigma} \\ 0 \leq \rho_e \leq 1, e = 1, \dots, N_e \end{cases} \end{aligned} \tag{1}$$

where, ρ is the vector consisting of the design variables (i.e. element (e) density), and N_e is the total number of design variables. $c(\rho)$ denotes the compliance of the structure (objective function), \mathbf{f} and \mathbf{u} are external load vector and displacement vector of the structure respectively, \mathbf{K} is the global stiffness matrix of the structure. V_e is the volume of element e and V_0 is the initial volume of the structure, occupying the entire design domain, v_f and \hat{v}_f is the material volume fraction and its allowable value respectively. σ_e^{vM} represents the von Mises stress for an arbitrary element, σ_e^{vM} and $\hat{\sigma}$ are the maximum von Mises stress of structure and its constraint limit value.

To suppress the generation of gray elements (intermediate values between 0 and 1) and obtain a clear black-and-white pattern, we use a modified solid isotropic material with penalization (SIMP) interpolation [31], where the stiffness is penalized by a power function relation between the Young’s modulus E_e and element relative density ρ_e , given by:

$$E_e(\rho_e) = E_{\min} + \rho_e^\omega (E_0 - E_{\min}), \tag{2}$$

where, $\omega > 1$ is the SIMP penalty factor. E_0 is the Young’s modulus of solid material while E_{\min} is a vanishingly small positive value to avoid numerical instabilities such as a singular global stiffness matrix. In this paper, ω assumes an empirical value of 3 while $E_{\min} = 10^{-9}$.

To alleviate the singular solution problem, in a manner similar to stiffness penalization in SIMP model, we penalize the stress vector in following scheme [12]:

$$\sigma_e = (\rho_e)^q \sigma_e^0 = (\rho_e)^q \mathbf{D}_0 \mathbf{B} \mathbf{u}_e, \tag{3}$$

where, $q < 1$ is the stress penalty factor, \mathbf{D}_0 is the constitutive matrix of solid materials and \mathbf{B} is strain-displacement matrix. $\sigma_e^0 = \mathbf{D}_0 \mathbf{B} \mathbf{u}_e$ is the stress vector (Voight notation) calculated at the geometric center of the FE element (e), and the σ_e^0 has a different form in 2D and 3D case, which is given by:

$$\sigma_e^0 = \begin{cases} (\sigma_{ex}^0, \sigma_{ey}^0, \tau_{exy}^0)^T & (2D \text{ case}) \\ (\sigma_{ex}^0, \sigma_{ey}^0, \sigma_{ez}^0, \tau_{exy}^0, \tau_{eyz}^0, \tau_{ezz}^0)^T & (3D \text{ case}). \end{cases} \tag{4}$$

The corresponding von Mises stress is then calculated as follows [34]:

$$\sigma_e^{vM} = \begin{cases} (\sigma_{ex}^2 + \sigma_{ey}^2 - \sigma_{ex}\sigma_{ey} + 2\tau_{exy}^2)^{\frac{1}{2}} & (2D \text{ case}) \\ (\sigma_{ex}^2 + \sigma_{ey}^2 + \sigma_{ez}^2 - \sigma_{ex}\sigma_{ey} - \sigma_{ey}\sigma_{ez} - \sigma_{ez}\sigma_{ex} + 3\tau_{eyz}^2 + 3\tau_{ezz}^2 + 3\tau_{ex}^2)^{\frac{1}{2}} & (3D \text{ case}). \end{cases} \tag{5}$$

In order to reduce the calculation cost caused by massive local stress constraints, we adopt the P-norm function [14] to coagulate the N_e local stress constraints in Eq. (1) so that to form a global stress constraint, given by:

$$\sigma^{PN} = \left(\sum_{e=1}^{N_e} (\sigma_e^{vM})^p \right)^{\frac{1}{p}}, \tag{6}$$

where, p may be interpreted as the P-norm factor. And it determines the difference between the original function and its approximate value. When larger values are used, more weight is given to peak stress. Theoretically, as the value of the parameter p approaches infinity, we have:

$$\lim_{p \rightarrow \infty} \sigma^{PN} = \max(\sigma_e^{vM}) = \sigma_{\max}^{vM}. \tag{7}$$

Thus, the global stress constraint in Eq. (1) is transformed into,

$$\sigma^{PN} \leq \hat{\sigma}. \tag{8}$$

However, it's impossible to choose an infinite value of p and global measure function σ^{PN} may not approximate σ_{\max}^{vM} very well. Therefore, we use following Eq. (8) to modify the constraint in Eq. (7),

$$\tilde{\sigma}^{PN} = \xi \sigma^{PN} \leq \hat{\sigma}, \tag{9}$$

where, adjusted parameter $\xi = \sigma_{\max}^{vM} / \sigma^{PN}$.

Generally, a high value of p could create problems with numerical accuracy since $(\sigma_e^{vM})^p$ could be very large. Actually when a high value of p is used, the peak stress is weighted more heavily, it often leads to oscillation or even divergence during the optimization iterations. The oscillation results from the high non-linearity of the global stress function. One work-around is to normalize the stress constraint, which is given by:

$$\tilde{\sigma}_{norm}^{PN} = \xi \left(\sum_{e=1}^{N_e} (\sigma_e^{vM})^p \right)^{\frac{1}{p}} \leq 1. \tag{10}$$

Sensitivity analysis

Using the adjoint method, the sensitivity of the objective function in Eq. (1) can be elegantly derived as follows:

$$\frac{\partial c}{\partial \rho_e} = -\mathbf{u}^T \frac{\partial \mathbf{K}}{\partial \rho_e} \mathbf{u}, \tag{11}$$

where, Eq. (10) is the well-known self-adjoint gradient as seen in [29].

Using the chain rule, we obtain the sensitivity of the stress constraint function $\tilde{\sigma}_{norm}^{PN}$

$$\frac{\partial \tilde{\sigma}_{norm}^{PN}}{\partial \rho_e} = \frac{\partial \tilde{\sigma}_{norm}^{PN}}{\partial \sigma_e^{vM}} \left(\frac{\partial \sigma_e^{vM}}{\partial \rho_e} \right)^T \frac{\partial \sigma_e}{\partial \rho_e}, \tag{12}$$

where, the three terms are respectively calculated as follows:

1. the derivative of normalized P-norm function with respect to von Mises stress.

$$\begin{aligned} \frac{\partial \tilde{\sigma}_{norm}^{PN}}{\partial \sigma_e^{vM}} &= \frac{\xi}{p \hat{\sigma}} \left(\sum_{e=1}^{N_e} \left(\frac{\sigma_e^{vM}}{\hat{\sigma}} \right)^p \right)^{\left(\frac{1}{p}-1\right)} \times p \left(\frac{\sigma_e^{vM}}{\hat{\sigma}} \right)^{p-1} \\ &= \frac{\xi}{\hat{\sigma}} \left(\sum_{e=1}^{N_e} \left(\frac{\sigma_e^{vM}}{\hat{\sigma}} \right)^p \right)^{\left(\frac{1}{p}-1\right)} \times \left(\frac{\sigma_e^{vM}}{\hat{\sigma}} \right)^{p-1}. \end{aligned} \tag{13}$$

2. The derivative of the von Mises stress with respect to stress components.

For the 2D case:

$$\begin{cases} \frac{\partial \sigma_e^{vM}}{\partial \sigma_{ex}} = \frac{1}{2\sigma_e^{vM}} (2\sigma_{ex} - \sigma_{ey}) \\ \frac{\partial \sigma_e^{vM}}{\partial \sigma_{ey}} = \frac{1}{2\sigma_e^{vM}} (2\sigma_{ey} - \sigma_{ex}) \\ \frac{\partial \sigma_e^{vM}}{\partial \sigma_{ez}} = \frac{2\tau_{exy}}{\sigma_e^{vM}}. \end{cases} \tag{14}$$

For the 3D case:

$$\begin{cases} \frac{\partial \sigma_e^{vM}}{\partial \sigma_{ex}} = \frac{1}{2\sigma_e^{vM}} (2\sigma_{ex} - \sigma_{ey} - \sigma_{ez}) \\ \frac{\partial \sigma_e^{vM}}{\partial \sigma_{ey}} = \frac{1}{2\sigma_e^{vM}} (2\sigma_{ey} - \sigma_{ex} - \sigma_{ez}) \\ \frac{\partial \sigma_e^{vM}}{\partial \sigma_{ez}} = \frac{1}{2\sigma_e^{vM}} (2\sigma_{ez} - \sigma_{ex} - \sigma_{ey}) \\ \frac{\partial \sigma_e^{vM}}{\partial \tau_{exy}} = \frac{3\tau_{exy}}{\sigma_e^{vM}}, \frac{\partial \sigma_e^{vM}}{\partial \tau_{eyz}} = \frac{3\tau_{eyz}}{\sigma_e^{vM}}, \frac{\partial \sigma_e^{vM}}{\partial \tau_{ezx}} = \frac{3\tau_{ezx}}{\sigma_e^{vM}}. \end{cases} \quad (15)$$

3. The derivative of the stress components with respect to the design variables:

$$\frac{\partial \sigma_e}{\partial \rho_e} = q\rho_e^{q-1} \mathbf{D}_0 \mathbf{B} \mathbf{u}_e + \rho_e^q \mathbf{D}_0 \mathbf{B} \frac{\partial \mathbf{u}_e}{\partial \rho_e}, \quad (16)$$

where, the term $\frac{\partial \mathbf{u}_e}{\partial \rho_e}$ is obtained from the equilibrium equations:

$$\mathbf{K} \mathbf{u} = \mathbf{f}. \quad (17)$$

By calculating the derivative of both sides of this equation, we have:

$$\frac{\partial \mathbf{K}}{\partial \rho_e} \mathbf{u}_e + \mathbf{K} \frac{\partial \mathbf{u}_e}{\partial \rho_e} = \mathbf{0}. \quad (18)$$

Giving,

$$\frac{\partial \mathbf{u}_e}{\partial \rho_e} = -\mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \rho_e} \mathbf{u}_e. \quad (19)$$

Substituting Eq. (18) into Eq. (15) and then Eq. (15) into Eq. (11) gives,

$$\frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \rho_e} = \frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \sigma_e^{vM}} \left(\frac{\partial \sigma_e^{vM}}{\partial \sigma_e} \right)^T \times \left(q\rho_e^{q-1} \mathbf{D}_0 \mathbf{B} \mathbf{u}_e - \rho_e^q \mathbf{D}_0 \mathbf{B} \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \rho_e} \mathbf{u}_e \right). \quad (20)$$

Here, an adjoint variable λ is defined as:

$$\lambda^T = \frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \sigma_e^{vM}} \left(\frac{\partial \sigma_e^{vM}}{\partial \sigma_e} \right)^T \mathbf{D}_0 \mathbf{B} \mathbf{K}^{-1}. \quad (21)$$

Thus, we obtain the adjoint equations:

$$\mathbf{K} \lambda = \frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \sigma_e^{vM}} \mathbf{B}^T \mathbf{D}_0^T \frac{\partial \sigma_e^{vM}}{\partial \sigma_e}. \quad (22)$$

Then, by solving Eq. (22), we obtain the adjoint variable λ which is inserted into Eq. (20) so that a final gradient of $\tilde{\sigma}_{\text{norm}}^{PN}$ can be expressed as,

$$\frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \rho_e} = \frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \sigma_e^{vM}} \left(\frac{\partial \sigma_e^{vM}}{\partial \sigma_e} \right)^T q\rho_e^{q-1} \mathbf{D}_0 \mathbf{B} \mathbf{u}_e - \rho_e^q \lambda^T \frac{\partial \mathbf{K}}{\partial \rho_e} \mathbf{u}_e, \quad (23)$$

where, the terms $\partial \tilde{\sigma}_{\text{norm}}^{PN} / \partial \sigma_e^{vM}$ and $\partial \sigma_e^{vM} / \partial \sigma_e$ are determined by Eq. (13) and Eq. (14) (or Eq. (15)) respectively. And as we know that,

$$\frac{\partial \mathbf{K}}{\partial \rho_e} = \omega \rho_e^{\omega-1} \tilde{E}_e \mathbf{K}^0. \quad (24)$$

So we can substitute Eq. (24) into equation Eq. (23) to get

$$\frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \rho_e} = \frac{\partial \tilde{\sigma}_{\text{norm}}^{PN}}{\partial \sigma_e^{vM}} \left(\frac{\partial \sigma_e^{vM}}{\partial \sigma_e} \right)^T q\rho_e^{q-1} \mathbf{D}_0 \mathbf{B} \mathbf{u}_e - \omega \rho_e^{q+\omega-1} \lambda^T \tilde{E}_e \mathbf{K}^0 \mathbf{u}_e. \quad (25)$$

Projection-based ROM approaches

On-the-fly ROM construction using PCA

For large-scale topology optimization problems, the main computational workload stems from the numerical solution of the equilibrium equation at each iteration:

$$\mathbf{K}\mathbf{u} = \mathbf{f}, \quad (26)$$

where, the computation of this high-precision full-field solution involves the inversion of a high-dimensional linear systems that can consist of up to millions or billions of degrees of freedom (DOF) [30]. Since the coefficient matrix \mathbf{K} is typically ill-conditioned (due to the SIMP interpolation) and also has a large number of DOFs, traditional methods depending on iteration methods like PCG may still be inadequate to significantly reduce the computational cost.

To alleviate this solution process, the authors call on a previously-proposed method called on-the-fly ROM construction using PCA [28]. The gist of this approach is mapping the displacement field solution of the large-scale problem Eq. (26) to a low-dimensional space spanned by an appropriately obtained orthonormal basis Φ ($\Phi^T \Phi = \mathbf{I}$) that is constructed on-the-fly and updated by implementing PCA for a set of stored solution snapshots obtained during the previous iterations. Here, we briefly present the salient features of the approach.

Given a snapshot matrix consisting of N_b displacement solutions at previous N_b successive iterations,

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{N_b}] \in \mathbb{R}^{N \times N_b}, \quad (27)$$

by zero-mean processing for matrix \mathbf{U} , we obtain its centered version $\bar{\mathbf{U}}$,

$$\bar{\mathbf{U}} = [\mathbf{u}_1 - \bar{\mathbf{u}}, \mathbf{u}_2 - \bar{\mathbf{u}}, \dots, \mathbf{u}_{N_b} - \bar{\mathbf{u}}], \bar{\mathbf{u}} = \frac{1}{N_b} \sum_{i=1}^{N_b} \mathbf{u}_i, \quad (28)$$

then performing singular value decomposition (SVD) on the centered matrix $\bar{\mathbf{U}}$, we have,

$$\bar{\mathbf{U}} = \Psi_{N \times N_b} \Sigma_{N_b \times N_b} \mathbf{V}_{N_b \times N_b}^T, \quad (29)$$

the reduced basis Φ is obtained by truncating the first $M \ll N_b$ columns corresponding to the largest N_b singular values in matrix Ψ which is composed of N_b left singular vectors $\phi_i, i = 1, \dots, N_b$,

$$\Phi \leftarrow [\phi_1, \phi_2, \dots, \phi_M], \quad (30)$$

usually, the value of M (i.e. the size of reduced basis ϕ), is determined by following energy rule:

$$1 - \frac{\sum_{i=1}^M \kappa_i^2}{\sum_{i=1}^{N_b} \kappa_i^2} \leq \varepsilon_{trunc}, \quad (31)$$

where, κ_i denotes the i th largest singular value of centered matrix $\bar{\mathbf{U}}$ and ε_{trunc} is a user-defined truncation error (in this paper 1%).

Now, by mapping the displacement field \mathbf{u} on to the low-dimensional space spanned by the reduced basis Φ constructed above, we have the following ROM expression,

$$\mathbf{u} \approx \tilde{\mathbf{u}} = \Phi \boldsymbol{\alpha} + \bar{\mathbf{u}}. \quad (32)$$

Substituting the approximate solution $\tilde{\mathbf{u}}$ into Eq. (26) and then projecting the equations onto the basis, we get a reduced system with a much smaller dimensionality,

$$\Phi^T \mathbf{K} \Phi \boldsymbol{\alpha} = \Phi^T (\mathbf{f} - \mathbf{K} \bar{\mathbf{u}}), \quad (33)$$

the high computational cost of solving full-scale system Eq. (26) may then be greatly reduced through a significantly less expensive solution of Eq. (33). Finally, a linear combination of projection coefficient α as in Eq. (32) is used to recover the full-field solution.

The approximate accuracy of $\tilde{\mathbf{u}}$ is monitored during the optimization via the residual equilibrium error which is given by,

$$e_{rb} = \frac{\|\mathbf{K}\tilde{\mathbf{u}} - \mathbf{f}\|}{\|\mathbf{f}\|}, \tag{34}$$

when the force residual of a certain optimization loop iteration (after at least N_b iterations have been completed so as to be able to obtain the basis) is lower than a pre-specified threshold \hat{e}_{rb} , we accept the reduced-basis solution in place of the full-field solution. Once the force residual exceeds the tolerance \hat{e}_{rb} , we are forced to recompute a full-field solution of Eq. (26) by the direct method, and then use this new solution to replace the “oldest” stored displacement vector in current snapshot matrix. In order to reduce the storage requirements for large scale problems, recalculating PCA basis for a single snapshot update, the incremental PCA can be considered as shown in the reference [35]. In this way, the original basis is updated and improved using PCA for the next loop iteration, which forms the so-called on-the-fly construction and updating of the ROM.

Here two points have to be noted. Firstly, the construction of the snapshot matrix, requiring the solution of N_b linear systems, may be done by an iterative solver (see [28]), avoiding full-size \mathbf{K} construction. Secondly, the projection of the linear problem in Eq. (33) also does not require the assembly of the global stiffness matrix \mathbf{K} , as the reduced matrix $\Phi^T \mathbf{K} \Phi$ and the RHS term $\Phi^T \mathbf{K}$ may be assembled directly by projecting individual element contributions.

Multi-fidelity ROM construction using multigrid

Topology optimization with ROM-based reanalysis (as seen in “On-the-fly ROM construction using PCA”) still suffers from a significant computational burden, as we need to repeatedly update the reduced basis to obtain a “sufficiently good” prediction of the displacement field when the accuracy of the ROM is insufficient. The new snapshot is then treated as a full-field solution calculated on a fine fixed FE mesh (i.e. single-fidelity), which should be effectively computed in some way. So, based on the above mentioned method, [26] proposed an approach of multi-fidelity ROM construction using the multigrid method, where the original equilibrium equations are solved on grids of various resolutions, and then quickly obtain a multi-fidelity snapshot solution for constructing and updating the reduced basis. Next, A typical “V” cycle MG iteration of recursive form is given in Algorithm 1 as follows ($\ell = 1, 2, \dots, n - 1$):

Algorithm 1 MG ($\mathbf{K}^{(\ell)}, \mathbf{f}^{(\ell)}, \mathbf{u}^{(\ell)}, \ell$)

```

1: procedure INITIALIZATION: ( $\mathbf{u}^{(\ell)}, \{\mathbf{K}^{(\ell+1)}\}, \{\mathbf{P}^{\ell+1}\}, \{\mathbf{W}_\ell\}$ )
2: if  $\ell + 1 = n$  then
3:    $\mathbf{u}^{(\ell+1)} \leftarrow (\mathbf{K}^{(\ell+1)})^{-1} \mathbf{f}^{(\ell+1)}$ 
4: else
5:    $\mathbf{K}^\ell \mathbf{u}^\ell = \mathbf{f}^\ell$ :  $\mathbf{u}^{(\ell)} \leftarrow \mathbf{W}_\ell (\mathbf{K}^{(\ell)}, \mathbf{f}^{(\ell)}, \mathbf{u}^{(\ell)})$  ▷ Pre-smoothing, apply  $\beta_1 (\beta_1 \leq 3)$ 
6:    $\mathbf{f}^{(\ell+1)} = (\mathbf{P}_{i+1}^i)^T (\mathbf{f}^{(\ell)} - \mathbf{K}^{(\ell)} \mathbf{u}^{(\ell)})$ ,  $\mathbf{u}_{\ell+1} = 0$ 
7:    $\mathbf{u}^{(\ell+1)} \leftarrow MG((\mathbf{K}^{(\ell+1)}, \mathbf{f}^{(\ell+1)}, \mathbf{u}^{(\ell+1)}, \ell + 1))$  ▷ Recursive iteration
8:    $\mathbf{u}^{(\ell)} \leftarrow \mathbf{u}^{(\ell)} + \mathbf{P}_{i+1}^i \mathbf{u}^{(\ell+1)}$  ▷ Coarse-grid correction
9:    $\mathbf{K}^\ell \mathbf{u}^\ell = \mathbf{f}^\ell$ :  $\mathbf{u}^{(\ell)} \leftarrow \mathbf{W}_\ell (\mathbf{K}^{(\ell)}, \mathbf{f}^{(\ell)}, \mathbf{u}^{(\ell)})$  ▷ Post-smoothing, apply  $\beta_2 (\beta_2 \leq 3)$ 

```

In this paper, we set the level number $n = 2$ and $\beta_1 = \beta_2 = 3$, i.e. a simple two-grid “V” cycle. And we choose the FE displacement solution obtained from the previous main optimization loop as the initial guess \mathbf{u}_ℓ , details of the coarse grid operators $\mathbf{K}^{(\ell+1)} = (\mathbf{P}_{\ell+1}^\ell)^\top \mathbf{K}^\ell \mathbf{P}_{\ell+1}^\ell$ and prolongation operators can be seen in [24]. The smooth operators use the PCG method. Algorithm 1 starts from $MG(\mathbf{K}^{(1)}, \mathbf{f}^{(1)}, \mathbf{u}^{(1)}, 1)$, where $\mathbf{K}^{(1)} = \mathbf{K}, \mathbf{f}^{(1)} = \mathbf{f}$. For the sake of accuracy, we repeat the above “V” cycle loop twice (initial guess for the next loop is selected as the MG solution of the previous loop)

The algorithms for both approaches (introduced in “Projection-based ROM approaches”) are combined in Algorithm 2, where SF-ROM stands for on-the-fly ROM construction using PCA using single-fidelity (SF) snapshots calculated on a fixed FE mesh, and MF-ROM for multi-fidelity ROM construction using multigrids where multi-fidelity (MF) snapshots are obtained by MG iteration. Full-order model (FOM) represents the full-field solution in Eq. (26) obtained using either the direct method (for SF-ROM) or the MG method (for MF-ROM).

Algorithm 2 SF-ROM and MF-ROM

```

1: procedure INITIALIZATION:(compute global stiffness matrix  $\mathbf{K}_i$  at  $i_{th}$  iteration loop, given button,
   set  $i = 1$ )
2: if  $1 \leq i \leq N_b$  then
3:    $\mathbf{u}_i \xleftarrow{FOM \text{ by direct method}} \mathbf{K}_i \mathbf{u}_i = \mathbf{f}$ 
4:   add  $\mathbf{u}_i$  to snapshot matrix  $\mathbf{U}$  and compute its centered version  $\bar{\mathbf{U}}$ 
5:   if  $i = N_b$  then
6:      $\Phi \leftarrow \text{SVD}(\bar{\mathbf{U}})$ 
7:   else
8:      $\alpha \leftarrow \Phi^\top \mathbf{K}_i \Phi \alpha = \Phi^\top (\mathbf{f} - \mathbf{K}_i \bar{\mathbf{u}})$ 
9:      $e_{rb} \leftarrow \frac{\|\mathbf{K}_i \bar{\mathbf{u}} - \mathbf{f}\|}{\|\mathbf{f}\|}$ 
10:    if  $e_{rb} > \hat{e}_{rb}$  then
11:      remove the oldest snapshot from  $\mathbf{U}$ 
12:      if button=1 then
13:         $\mathbf{u}_i \xleftarrow{FOM \text{ by direct method}} \mathbf{K}_i \mathbf{u}_i = \mathbf{f}$  ▷ SF-ROM
14:      else
15:         $\mathbf{u}_i \xleftarrow{FOM \text{ by MG method}} \mathbf{K}_i \mathbf{u}_i = \mathbf{f}$  ▷ MF-ROM
16:      add  $\mathbf{u}_i$  to snapshot matrix  $\mathbf{U}$  and compute its centered version  $\bar{\mathbf{U}}$ 
17:       $\Phi \leftarrow \text{SVD}(\bar{\mathbf{U}})$ 
18:    else
19:       $\bar{\mathbf{u}}_i \xleftarrow{ROM} \Phi \alpha + \bar{\mathbf{u}}$ 
20:    Update the cyclic variable :  $i = i + 1$  Back to line 1

```

Two improved strategies

Adap-ROM

Whether we use the SF-ROM or the MF-ROM, a pivotal ROM parameter, the force residual threshold \hat{e}_{rb} must always be predefined by users. Sometimes it’s hard to give an appropriate value of \hat{e}_{rb} as it clearly involves a trade-off between calling the ROM solution and the FOM solution during the optimization procedure. On the one hand, if we use a smaller value, then the optimal objective may have better accuracy but would require more CPU time as the FOM would ostensibly be called more frequently, on the other hand, if we use a bigger value, then the optimal objective may have poorer accuracy but would require less CPU time as the ROM would tend to be called more frequently. Therefore, in this paper, we propose an adaptive selection strategy for the value of force residual threshold which is then chosen according to the relative error variation of objective function value during the iteration loop.

$$e_c = \frac{\|c_{i-1} - c_{i-2}\|}{\|c_{i-2}\|} (i \geq 3), \tag{35}$$

c_{i-1}, c_{i-2} are the objective function values at two consecutive iterations $i - 1_{th}$ and $i - 2_{th}$, here the relative errors gradually decreasing because of the convergence behavior of the compliance minimization problem. Details of the strategy are summarized in the following Algorithm 3.

Algorithm 3 Adaptive parameter selection of force residual threshold

```

1: procedure INPUT:( $c_{i-1}, c_{i-2}, \hat{e}_{rb}^0$ )           ▷ Two Objective function values  $c_{i-1}, c_{i-2}$ 
2:                                     ▷ Initial force residual threshold  $\hat{e}_{rb}^0$ 
3: procedure INPUT:( $\theta, \hat{e}_c^0, \varsigma$ )           ▷ Scale factor  $\theta$ , the relative error threshold  $\hat{e}_c^0$ 
4:                                     ▷ Decay factor  $\varsigma$ 
5: if  $e_c > \hat{e}_c^0$  then
6:    $\hat{e}_{rb} = 0$                                      ▷ No ROM
7:   if  $\varsigma \hat{e}_c^0 < e_c \leq \hat{e}_c^0$  then
8:      $\hat{e}_{rb} = \hat{e}_{rb}^0$ 
9:   else
10:    set  $\hat{e}_{rb} = \theta \hat{e}_{rb}^0$                        ▷ Update the  $\hat{e}_{rb}$ 

```

In this paper, we set $\theta > 1$, so that the force residual threshold changes from a relatively smaller value towards a bigger value as the relative error e_c reduces. More specially, the adaptive selection of \hat{e}_{rb} is roughly divided into three stages: In the first stage, e_c is relatively higher and so the FOM rather than ROM is called on to maintain the stability of the optimization procedure. In the second stage, the variation of the objective function is less, so we begin using the ROM with a smaller value of \hat{e}_{rb} . Finally in the third stage, e_c is small enough (meaning that objective is sufficiently converged) therefore a larger value \hat{e}_{rb} of can be used, thus the ROM solution is more readily accepted. Note that the code in Algorithm 3 may be directly inserted between lines 7 and 8 in Algorithm 2 for obtaining a new version of ROM using adaptive parameter selection for the residual threshold.

Bi-ROM

Recall that in “[Introduction](#)”, we have mentioned that stress-constrained topology optimization problems are generally non-self-adjoint, and will yield an additional adjoint equation (see (Eq. (22)) derived from sensitivity analysis of the stress constraint function, that needs to be solved. Taking advantage of the fact that the adjoint equation is of the same scale as the FE equilibrium equations, it would seem logical to apply the ROM idea given in Algorithm 2 to the adjoint equations, giving:

$$\lambda = \Xi \mu + \bar{\lambda}, \tag{36}$$

where, Ξ is the reduced basis, μ represents the projection coefficients and $\bar{\lambda}$ is the mean snapshot for the set of adjoint solutions. An important feature here is that the construction of the ROM for the adjoint equation and the equilibrium equations in this paper are completely independent, which means the snapshot number and residual threshold of the ROM for the adjoint equations could be completely different from those for the equilibrium equations.

Numerical tests

In this section, we first apply the two ROM methods that were introduced in “[Projection-based ROM approaches](#)”: SF-ROM and MF-ROM to a 2D benchmark stress-constrained compliance minimization problem. Next, using MF-ROM, we demonstrate the effectiveness of the two improved strategies that were presented in “[Two improved strategies](#)”. In order to better understand the impact of the parameters associated with

the stress constraints on optimal design, we use different values for the p norm factor p , the stress penalty factor q and the allowable maximum stress $\hat{\sigma}$ for numerical experiments using a benchmark 3D L-shaped beam. Unless otherwise stated, the variables and geometric parameters are dimensionless for all test examples in this paper. Pre-defined parameters: Young's modulus $E_0 = 1$ for the solid material with a lower limit $E_{\min} = 10^{-9}$ for void material, Poisson's ratio $\nu = 0.3$. The penalty factor of stiffness $\omega = 3$, density filter $r_{\min} = 2.5$ and the allowable volume fraction $\hat{v}_f = 0.2$. Since the objective function (here compliance) is conveniently represented using a separable convex approximation by the Method of Moving Asymptotes (MMA), which is easily adapted to different types of topology optimization problems, we retain the MMA as the optimizer for iteration loop. The optimization iterations stop either when the convergence condition is satisfied or once the maximum allowable number of iterations is reached. Here, the convergence criterion is defined as follows:

$$\frac{\|\rho_{i-1} - \rho_i\|_2}{\|\rho_i\|_2} \leq Tol, \quad (37)$$

where, $\|\cdot\|_2$ denotes the binary norm of a vector. i denotes the current iteration step, and Tol is the allowable convergence error, in this paper, $Tol = 10^{-4}$.


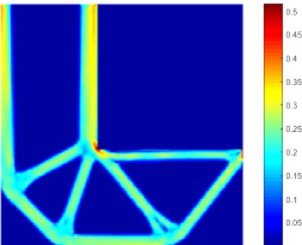

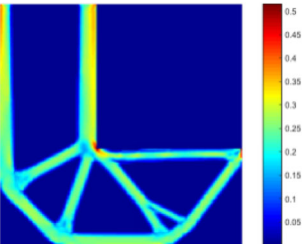

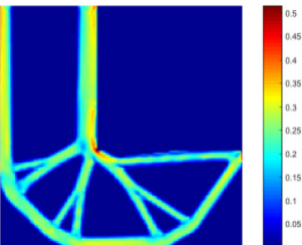

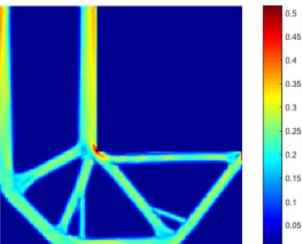
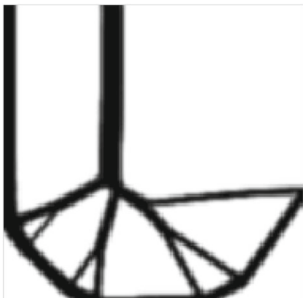
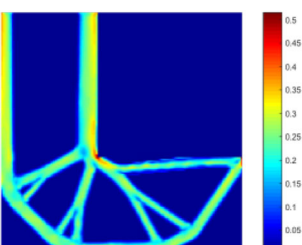
2D L-shaped beam

Consider the classical optimization problem of the 2D L-shaped beam [4,5,8,10], the design domain and boundary conditions are illustrated in Fig. 1. The structure's top face on the left side is fixed and an external load $f = 1$ is applied on the top right corner A . We use 16384 bilinear quadrilateral elements to discretize the design domain, each element having unit volume and thickness. To avoid stress concentration at the loading node A , the concentrated load is evenly distributed over the 6 adjacent nodes between node A and B , as shown in Fig. 1. For this 2D case, we set the p norm factor $p = 8$, the stress penalty factor $q = 0.8$ and the allowable maximum stress. The maximum iteration number is set to 200. Note that for all test cases shown hereafter, the iterations were halted due to the maximum allowable number of optimization loop iterations having been attained.

Table 1 shows the optimal topology and corresponding stress distribution obtained using the reference method (without ROM), as well as the SF-ROM method and the MF-ROM method using a fixed snapshot number N_b and various values for the force residual threshold $\hat{\epsilon}_{rb}$. Figures 2 and 3 compare the iteration histories for the convergence of the compliance and stress (von Mises stress σ_{\max}^{vM} and P-norm stress σ^{PN}) between the SF-ROM and MF-ROM methods.

From Table 1 a macroscopic view leads us to the premature conclusion that the optimal topology using SF-ROM and MF-ROM seem basically consistent with that using the reference method; but on closer inspection more structural branches occur and the arc of reentry of the L-shaped beam is more distinct when using a bigger value of $\hat{\epsilon}_{rb}$. And it should be noted that for both SF-ROM and MF-ROM, stress concentration still exists at the location of the applied force. Figure 2 shows that the objective optimization is converged regardless of whether or not we use SF-ROM or MF-ROM. Figure 3 shows that stress constraints using both the SF-ROM and MF-ROM are active and gradually satisfied during the overall optimization procedure. Furthermore, due to the constraint violation shown in Fig. 3, the convergence history of the objective function is non-monotonous as shown in Fig. 2.

Table 1 Comparison of optimal topology and stress distribution for 2D L-shaped beam obtained using reference method(No ROM) and SF-ROM and MF-ROM under various $\hat{\epsilon}_{rb} = 0.1, 0.5$ and fixed $N_b = 4$

Method	Optimal topology	Stress distribution
Reference		
SF-ROM $N_b = 4, \hat{\epsilon}_{rb} = 0.1$		
SF-ROM $N_b = 4, \hat{\epsilon}_{rb} = 0.5$		
MF-ROM $N_b = 4, \hat{\epsilon}_{rb} = 0.1$		
MF-ROM $N_b = 4, \hat{\epsilon}_{rb} = 0.5$		

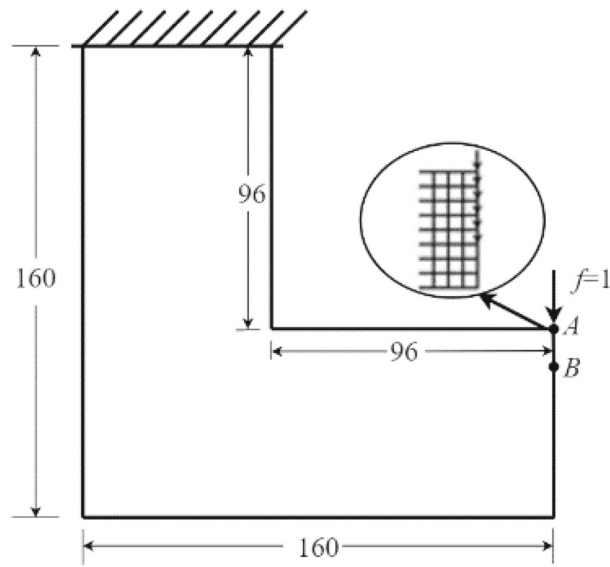


Fig. 1 Design domain and boundary conditions of L-shaped beam

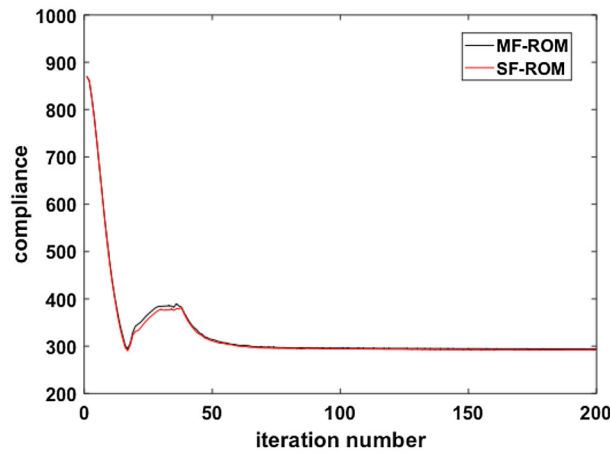


Fig. 2 Comparison of compliance iterative history between SF-ROM and MF-ROM using and

The above design results are summarized in Table 2, where the relative error of the optimal compliance (c^*) characterizes the approximate accuracy of the objective function value using ROM,

$$error = \frac{|c_{refe}^* - c_{rom}^*|}{|c_{refe}^*|}, \tag{38}$$

where, c_{refe}^* and c_{rom}^* represent the optimal compliance obtained by using the reference method and the ROM method respectively.

From Table 2, as we increase the value of $\hat{\epsilon}_{rb}$, the CPU clock time using both SF-ROM and MF-ROM show a significant decrease compared to the reference method, which is easy to explain : the cheaper reduced-basis solution (ROM) is called more frequently, thus the computational efficiency of the overall optimization routine is improved. However, the optimal compliance using both ROM methods has a very poor accuracy if we use a larger value of the force residual threshold ($\hat{\epsilon}_{rb} = 0.5$). This indicates that we must make

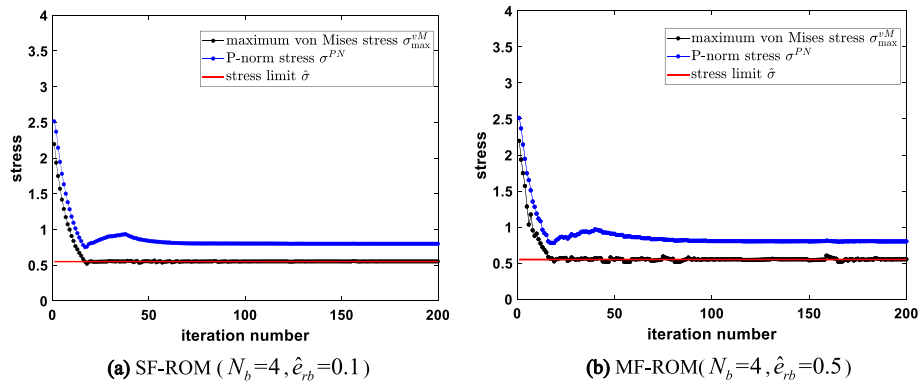


Fig. 3 Comparison of stress iterative history between SF-ROM (a) and MF-ROM (b)

Table 2 Comparison of design results between SF-ROM and MF-ROM (after 200 iterations)

Method	N_b	\hat{e}_{rb}	Calls to ROM	CPU running time (/s)	Optimal compliance (c^*)	Relative error of c^* (%)
Reference	–	–	0/200	1884.26	291.48	–
SF-ROM	4	0.1	135/200	1624.22	292.3	0.28
		0.5	164/200	1355.04	301.5	3.44
MF-ROM	4	0.1	129/200	1541.80	294.19	0.93
		0.5	164/200	1296.32	300.11	2.96

Table 3 Comparison of design results between MF-ROM and Adap-ROM (after 200 iterations)

Method	N_b	\hat{e}_{rb}	Calls to ROM	CPU running time (/s)	Optimal compliance (c^*)	Relative error of c^* (%)
Reference	–	–	0/200	1884.26	291.48	–
MF-ROM	4	0.05	105/200	1677.06	291.10	0.13
		0.5	164/200	1296.32	300.11	2.96
Adap-ROM	4	0 → 0.05 → 0.5	116/200	1652.98	291.68	0.07

a trade-off between efficiency and accuracy. Furthermore, by comparing the performance of the MF-ROM with the SF-ROM, we find that MF-ROM can further drive down the computational cost with a minor loss of objective accuracy when using the same set of values of N_b and \hat{e}_{rb} .

Next, we use the first improved strategy: adaptive parameter selection of the force residual threshold \hat{e}_{rb} to modify the MF-ROM method in “Projection-based ROM approaches”. Some parameters need to be predefined: initial force residual threshold $\hat{e}_{rb}^0 = 0.05$ and its scale factor $\theta = 10$, initial limit of objective relative error (see Eq. (32)) $\hat{e}_{rb}^0 = 2 \times 10^{-3}$ and its decay factor $\zeta = 10^{-1}$. The optimized design results are summarized in Table 3 and Fig. 4 shows the optimal topology using the MF-ROM and Adap-ROM as well as the reference method.

From Table 3, we see that topology optimization using Adap-ROM has a clear advantage in terms of accuracy compared to the MF-ROM method using a bigger force residual ($\hat{e}_{rb} = 0.5$), while requiring less computational time compared to the reference method, as well as the MF-ROM method using a smaller $\hat{e}_{rb} = 0.05$ (as expected), which shows that this improved method allows a good compromise between accuracy and speed. From Fig. 4 we see that the optimal topology obtained using Adap-ROM is closer to that obtained using the reference method, compared to the optimal topology obtained using MF-ROM,

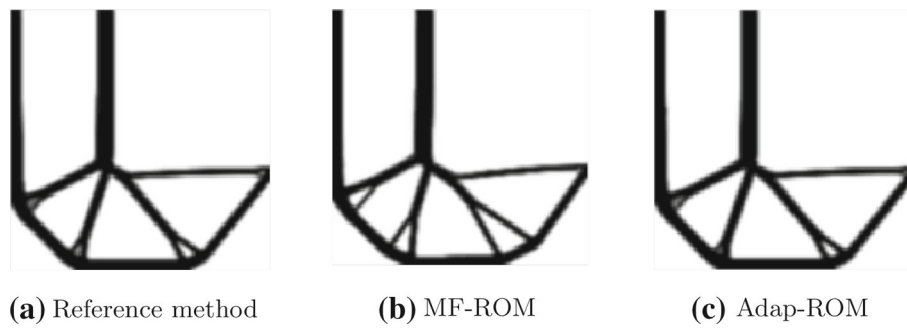


Fig. 4 Comparison of optimal topology

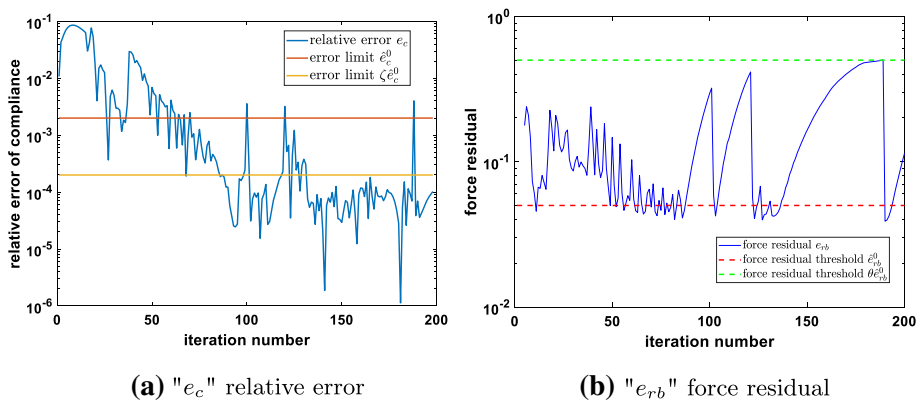


Fig. 5 Change history of using Adap-ROM

Table 4 Comparison of design results between MF-ROM and Bi-ROM (after 200 iterations)

Method	N_b	e_{rb}	Calls to ROM	CPU running time (/s)	Optimal compliance (c^*)	Relative error of c^* (%)
Reference	–	–	0/200	1884.26	291.48	–
MF-ROM	4	0.1	129/200	1541.80	294.19	0.93
Bi-ROM	4	0.1	132/200	1338.92	293.88	0.82
	6	0.05	148/200	1338.92	293.88	0.82

which is again expected. Next we give the change history of relative error e_c generated by using Adap-ROM in Fig. 5a, which allows us to decide when to use different force residual thresholds. And Fig. 5b gives the evolutionary curve of the force residual e_{rb} .

Then, for the same 2D benchmark stress-constrained topology optimization problem, we use another improved strategy Bi-ROM to modify the MF-ROM method in “[Projection-based ROM approaches](#)”. For the equilibrium equations (Eq. (23)), we set $N_b = 4$ and $e_{rb} = 0.1$, for adjoint equations (Eq. (21)), we set $N_b = 6$ and $e_{rb} = 0.05$. The design results are summarized in Table 4 and Fig. 6 shows the optimal topology obtained using MF-ROM and Bi-ROM as well as reference method.

From Table 4 we can see that Bi-ROM is less time consuming compared to the MF-ROM and the accuracy of the optimized compliance is better. Figure 6 shows that the Bi-ROM appears to yield a clearer optimal topology than the MF-ROM method.

Figure 7 shows the iteration history for the force residual e_{rb} of the equilibrium equations as well as the adjoint equations, using the Bi-ROM method.

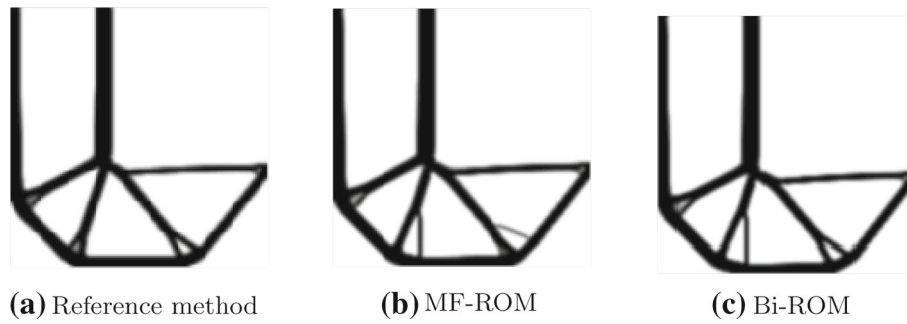


Fig. 6 Comparison of optimal topology

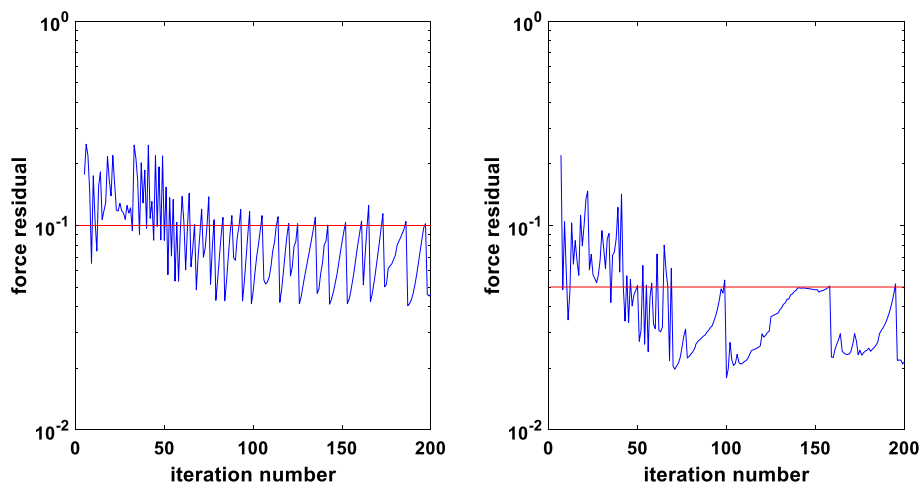


Fig. 7 Evolutionary curve of force residual e_{rb} using Bi-ROM method for equilibrium equations(left) and adjoint equations(right)

A conclusion may be directly drawn from Fig. 8 that for the Bi-ROM, the ROM solution of the adjoint equations is more readily “acceptable” than that of the equilibrium equations, even for smaller values of e_{rb} . This is because the residual evolutionary curve of equilibrium equations frequently oscillated above and below the threshold line(red) while the residual evolutionary curve of equilibrium equations basically stays below the threshold line as the iteration progresses.

Finally, we combine the Adap-ROM and Bi-ROM methods together (denoted by AdapBi-ROM) to further improve the MF-ROM method. Some parameters are pre-defined like before: initial limit of objective function relative error $\hat{e}_c^0 = 2 \times 10^{-3}$ decay factor $\zeta = 10^{-1}$, for the equilibrium equations: $N_b = 4$, initial force residual threshold $\hat{e}_c^0 = 0.05$ with scale factor $\theta = 10$, for adjoint equations $N_b = 6$, initial force residual threshold $\hat{e}_c^0 = 0.01$ and its scale factor $\theta = 10$. The optimized design results are summarized in Table 5 as follows.

From Table 5 we can see, the combined strategy AdapBi-ROM shows better performance than any one of the individual strategies(Adap-ROM or Bi-ROM) from a holistic perspective since this combined method can not only significantly reduce the computation time but also ensure a high accuracy of the value of objective function.

Table 5 Comparison of design results among AdapBi-ROM, Adap-ROM and Bi-ROM (after 200 iterations)

Method	N_b	$\hat{\epsilon}_{rb}$	Calls to ROM	CPU running time (/s)	Optimal compliance (c^*)	Relative error of c^* (%)
reference	–	–	0/200	1884.26	291.48	–
Adap-ROM	4	0 \rightarrow 0.05 \rightarrow 0.5	116/200	1652.98	291.68	0.07
Bi-ROM	4	0.5	167/200	1294.97	299.09	2.61
	6	0.1	153/200	1294.97	299.09	2.61
AdapBi-ROM	4	0 \rightarrow 0.05 \rightarrow 0.5	112/200	1352.96	291.52	0.01
	6	0 \rightarrow 0.01 \rightarrow 0.1	96/200	1352.96	291.52	0.01

Discussion

The multi-fidelity reduced-order-model (MF-ROM) approach clearly enhances the performance of the topology optimization procedure with a superior reduction of computational effort and CPU time compared with single-fidelity reduced-order-model (SF-ROM) for all the test cases studied, noting that a small loss of accuracy of the objective function is expected when using MF-ROM. By using the first proposed improved strategy (Adap-ROM) of adaptively selecting the force residual threshold ($\hat{\epsilon}_{rb}$) in MF-ROM, we found a trade-off can be made between optimization efficiency and objective accuracy. We also note that the second proposed improved strategy (Bi-ROM) of applying MF-ROM method to both adjoint equations derived from sensitivity analysis of stress constraint function and equilibrium equations in FE analysis can reduce the CPU running time in a more significant manner. In addition, even better accuracy of objective function can be obtained compared with the original MF-ROM method (for this 2D case at least), this may be because using a ROM for the adjoint equations corrects the convergence of the objective function. Finally, we combined both strategies and this can obtain better optimized results because there is a significant improvement in efficiency with a higher accuracy of approximation for the objective function.

3D L-shaped beam

In the previous section, we investigated the application of reanalysis using ROM in stress-constrained topology optimization through a 2D numerical example. In this section, we apply the previously described ROM methods to the 3D compliance minimization problem of the L-shaped beam. To better understand the influence of the stress constraint on the optimization results, we then consider making an in depth analysis of some of the parameters associated with the stress constraint based on the AdapBi-ROM method. The L-shaped beam is shown in Fig. 8, where the structure's top surface is fixed and a vertically downward load $f = 1$ is applied along the top right side.

We use standard hexahedral elements to discretize the design domain with each element having unit volume and thickness. The maximum number of iterations is set to 300.

ROM methods on 3D L-shaped beam

Here, we set $L = 80$ and $W = 50$, thus the design domain is discretized by a total of 32768 elements. 5 ROM methods were tested for this 3D case, The corresponding design results are given in Table 6.

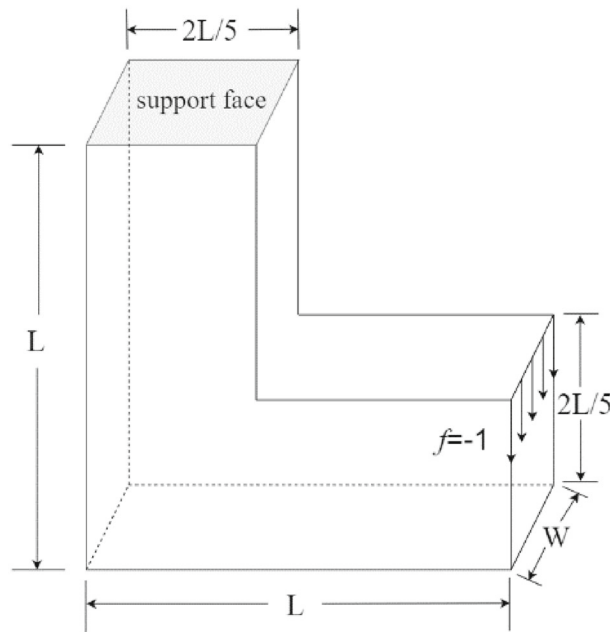


Fig. 8 Design domain and boundary conditions of 3D L-shaped beam

Table 6 Comparison of design results using several ROM methods (after 200 iterations)

Method	N_b	\hat{e}_{rb}	Calls to ROM	CPU running time (/s)	Optimal compliance (c^*)	Relative error of c^* (%)
Reference	–	–	0/300	23406.28	45.02	–
SF-ROM	4	0.05	224/300	21388.52	45.02	0
MF-ROM	4	0.05	225/300	19682.72	45.02	0
	4	0.5	260/300	19578.27	45.07	0.11
Adap-ROM	4	0 → 0.05 → 0.5	233/300	19655.47	45.01	0.02
Bi-ROM	4	0.5	260/300	18663.07	45.09	0.16
	6	0.1	281/300	18663.07	45.09	0.16
AdapBi-ROM	4	0 → 0.05 → 0.5	233/300	18896.49	45.00	0.04
	6	0 → 0.01 → 0.1	254/300	18896.49	45.00	0.04

From Table 6 we see that the combined AdapBi-ROM method behaves well both as far as computational time saving as well as final objective accuracy are concerned, as expected, in the 3D large-scale test case. Next, we will use this method as the ROM approach of choice to conduct a parameter study.

Parametric study

For this, we set $L = 50$ and $W = 6$ in Fig. 8, thus the design domain is discretized by a total of 9600 elements. Some parameters are predefined as follows: initial limit of objective relative error $\hat{e}_c^0 = 5 \times 10^{-3}$ and its decay factor $\zeta = 10^{-2}$, for equilibrium equations $N_b = 4$; initial force residual threshold $\hat{e}_c^0 = 0.04$ and its scale factor $\theta = 20$, for adjoint equations $N_b = 6$; initial force residual threshold $\hat{e}_c^0 = 0.01$ and its scale factor $\theta = 50$.

Table 7 Optimization results of AdapBi-ROM and reference method using varying p (after 300 iterations)

p	Calls to ROM		CPU running time (/s)		Optimal compliance (c^*)	
	Reference	AdapBi-ROM	Reference	AdapBi-ROM	Reference	AdapBi-ROM
2	0	120/300	2269.06	2097.43	80.04	79.92
	0	34/300				
4	0	192/300	2401.96	2020.23	78.73	78.79
	0	85/300				
6	0	207/300	2354.69	1768.36	78.81	78.86
	0	141/300				
8	0	202/300	2277.97	1953.88	78.38	78.63
	0	130/300				
12	0	201/300	2209.22	1871.71	79.59	78.45
	0	147/300				
20	0	227/300	2159.00	1560.70	80.33	80.30
	0	255/300				
30	0	231/300	2180.82	1558.75	80.30	80.29
	0	252/300				

Influence of p norm factor p

Here, we vary the p norm factor from 2 to 30, and while fixing the stress penalty factor at $q = 0.8$ and the allowable maximum stress at $\hat{\sigma}$. Table 7 gives the design results.

It can be seen from Table 7 that the CPU runtime trend for both the reference method as well as the combined AdapBi-ROM method is not monotonic. More computation time is spent when using a smaller p value (like 2 and 4) while less time is spent when using a larger p value (like 20 and 30) for these two methods. However, both methods still have some differences in terms of the changing trend of the CPU runtime, where the CPU time of the reference method goes down while that of the AdapBi-ROM method goes up, when we vary the p value from 6 to 8.

Interestingly, we find that the optimal compliance obtained using both methods has the same trend of first dropping and then increasing as the value of p increases, which means that we can likely obtain the smallest optimal compliance (meaning the “stiffest” possible structure) by using a mid-range value for p . In general, to better balance CPU time and structural performance, the p value should neither be too large nor too small.

Figures 9, 10 and 11 show the optimal topology and stress iterative history as well as the volume fraction iteration history for different values of the P-norm p norm factor with the AdapBi-ROM method.

We see from Fig. 9 that the optimal topologies obtained using different values of p based on AdapBi-ROM method are visually different when using $p = 4$ and $p = 6$ (or $p = 12$ and $p = 20$).

From Fig. 10, we can see that the stress value during the entire procedure satisfy its constraint limit progressively faster with a larger value of p . Meanwhile, it seems that the larger the value of p , the smaller the value of the final maximum von Mises stress σ_{\max}^{VM} . For $p = 20$ and $p = 30$, the curves are below the stress limit within 50 iterations.

Fig. 11 shows that the volume fraction during the iterations can rapidly, and in a stable manner, satisfy its respective constraint limiting value, provided $p > 2$.

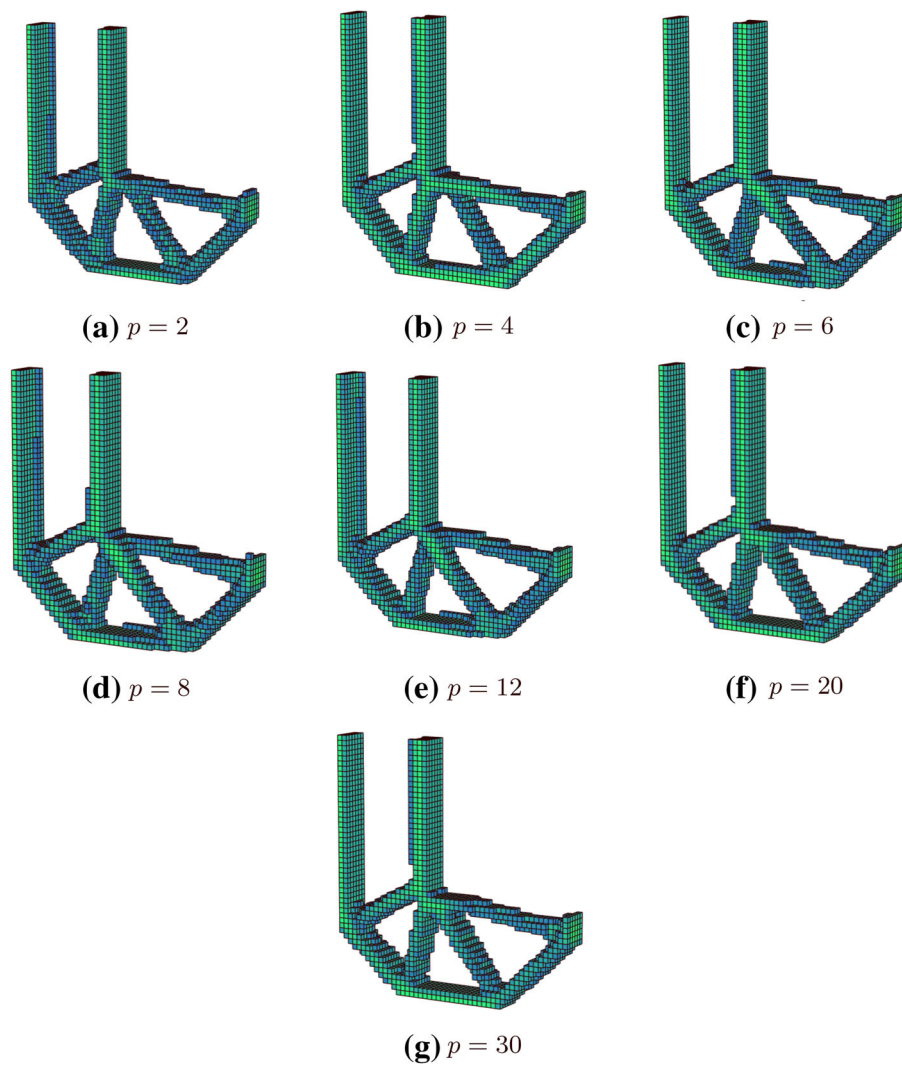


Fig. 9 Comparison of optimal topology using different value of p based on AdapBi-ROM method

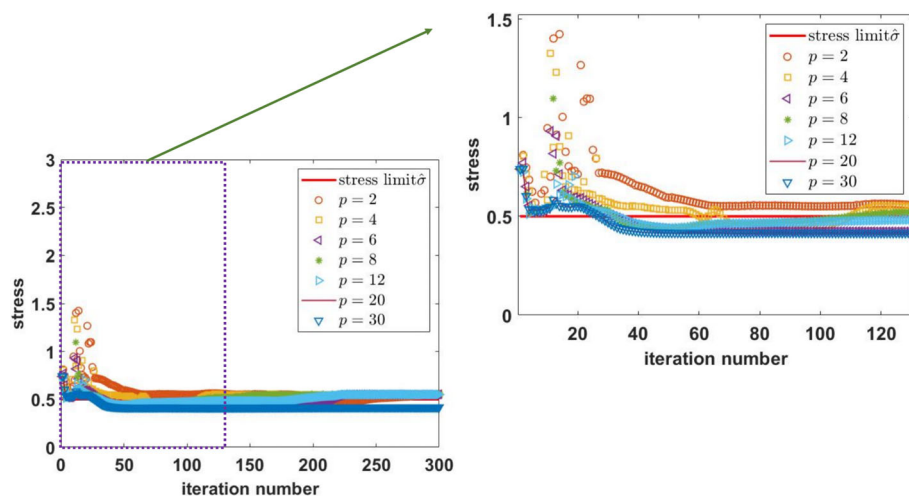


Fig. 10 Comparison of stress iterative history using different value of p based on AdapBi-ROM method

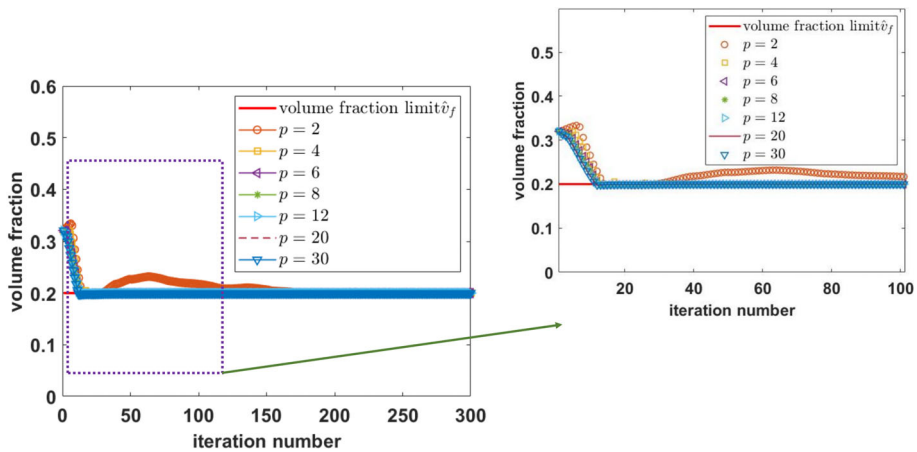


Fig. 11 Comparison of volume fraction iterative history using different value of p based on AdapBi-ROM method

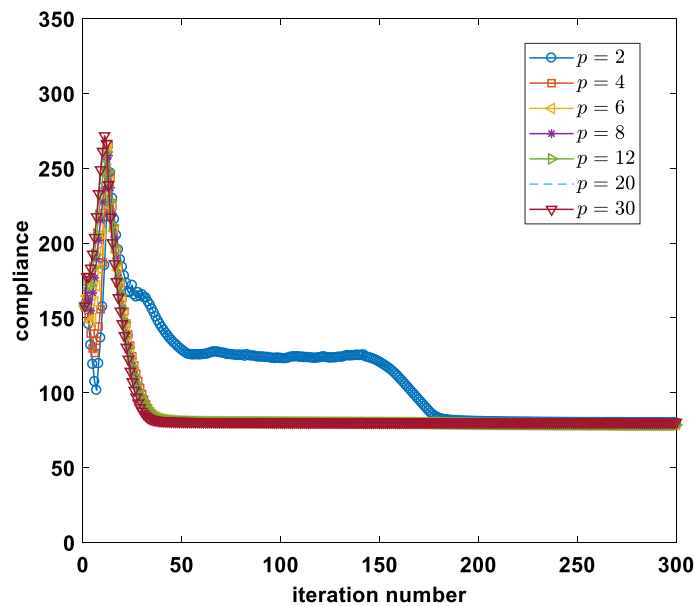


Fig. 12 Comparison of compliance iterative history using different values of p based on AdapBi-ROM method

Next, we give a comparison of the compliance (objective function) iteration history using different values of p based on AdapBi-ROM method in Fig. 12 where it appears that the larger the p value, the faster the rate of convergence of the optimization.

Investigation of varying stress penalty factor q

We vary the stress penalty factor q while fixing the p norm factor $p = 8$ and the allowable maximum stress at. Table 8 gives the design results.

It can be seen from Table 8 that the trends for CPU time and optimal compliance for both the reference method and the AdapBi-ROM method are exactly the same and not monotonic as the q value decreases. Therefore, the ROM retains the same impact of the stress penalty factor parameter as for the reference case. Interestingly, the CPU time has a general tendency to first increase and then decrease while the optimal compliance has the opposite tendency, implying a potential conflict between CPU time and optimal

Table 8 Optimization results of AdapBi-ROM method and reference method using varying q (after 300 iterations)

q	Calls to ROM		CPU running time (/s)		Optimal compliance (c^*)	
	Reference	AdapBi-ROM	Reference	AdapBi-ROM	Reference	AdapBi-ROM
0.9	0	228/300	2150.16	1584.64	80.00	79.98
	0	236/300				
0.8	0	202/300	2277.97	1953.88	78.38	78.63
	0	130/300				
0.5	0	146/300	2180.45	1825.66	79.90	79.74
	0	151/300				
0.3	0	198/300	2232.93	1884.87	80.76	80.76
	0	128/300				
0.1	0	226/300	2197.40	1698.84	80.26	80.27
	0	172/300				

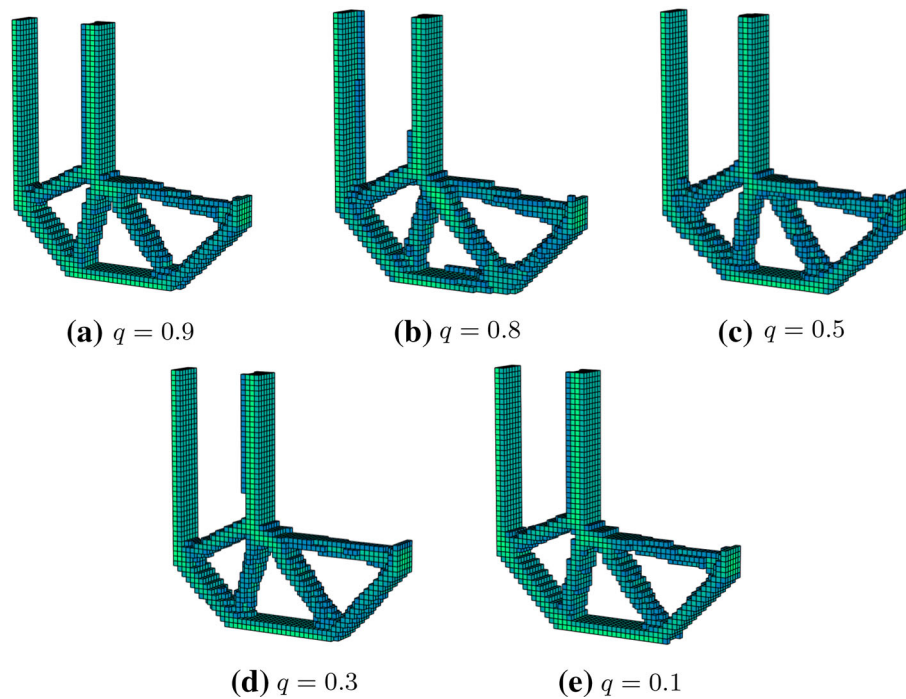


Fig. 13 Comparison of optimal topology using different value of q based on AdapBi-ROM method

compliance for several q values. This means that q needs to be properly determined to better balance the computational efforts and structural performance.

Just like in “[Influence of p norm factor p](#)”, we give the optimal topology and stress iteration history as well as the volume fraction iteration history for different values of the stress penalty factor, using the AdapBi-ROM method, as shown in Figs. 13, 14 and 15, respectively.

From Fig. 13 we can see that the optimal topologies using different values of q based on AdapBi-ROM method are visually distinguishable from each other, which can be seen from the test case with $q = 0.5$ and $q = 0.8$.

From Fig. 14 we note that the stress constraint is not satisfied when using relatively smaller values of $q(0.1 - 0.3)$ which means the stress penalization is not sufficient. In

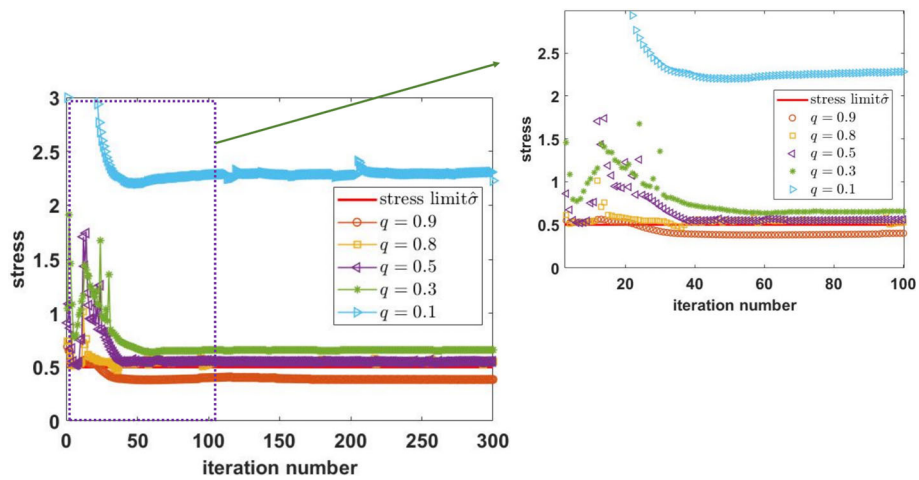


Fig. 14 Comparison of stress iterative history using different value of q based on AdapBi-ROM method

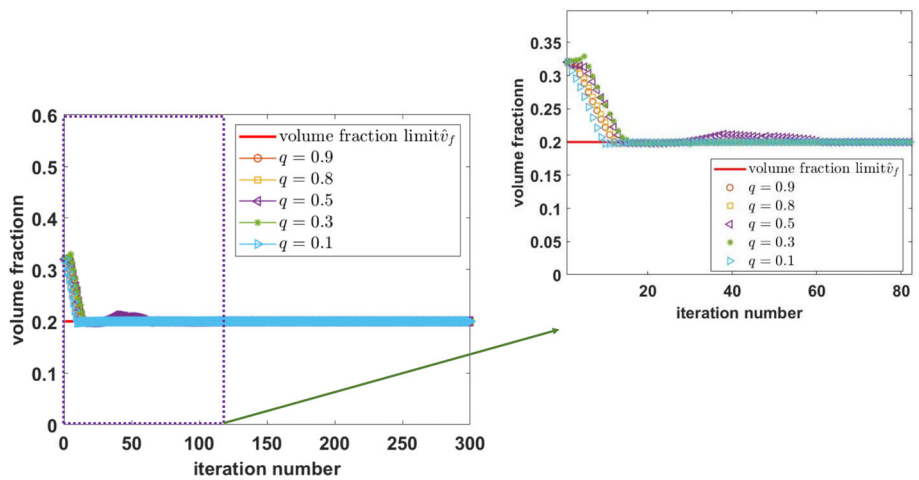


Fig. 15 Comparison of volume fraction iterative history using different value of q based on AdapBi-ROM method

addition, using a larger value of q can cause the maximum von Mises stress to satisfy its constraint limiting value as soon as possible during the procedure. In the case of $q = 0.9$, the curves end up below the stress limit after about 25 iterations.

Figure 15 shows that the volume fraction during the iterations can rapidly and in a stable manner satisfy the constraint for all 5 different values of q tested here.

Next, we compare the compliance iteration histories using different values of q based on the AdapBi-ROM method in Fig. 16 where different convergence rates of optimization can be obtained by varying q , but no monotonicity is observed as q increases. It seems plausible that a midrange value of q would lead to a relative lower convergence rate as shown below.

Varying the allowable maximum stress

In this case, we vary the allowable maximum stress $\hat{\sigma}$ while simultaneously fixing the p norm factor p at 8 and the stress penalty factor q at 0.8. Table 9 gives the design results.

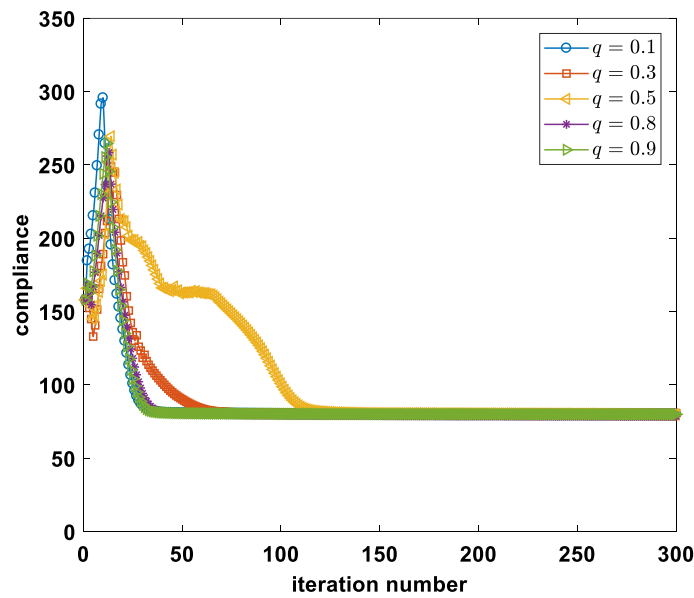


Fig. 16 Comparison of compliance iterative history using different values of q based on AdapBi-ROM method

Table 9 Optimization results of AdapBi-ROM method and reference method using varying $\hat{\sigma}$ (after 300 iterations)

$\hat{\sigma}$	Calls to ROM		CPU running time (/s)		Optimal compliance (c^*)	
	Reference	AdapBi-ROM	Reference	AdapBi-ROM	Reference	AdapBi-ROM
0.15	0	112/300	2413.22	2237.27	65.20	65.09
	0	35/300				
0.3	0	96/300	2408.48	2250.57	96.30	98.64
	0	26/300				
0.45	0	136/300	2664.67	2332.58	78.58	78.88
	0	103/300				
0.65	0	229/300	2240.31	1566.90	80.18	80.18
	0	251/300				
0.75	0	227/300	2171.39	1573.41	80.00	80.00
	0	248/300				

From Table 9 we see that the trends for the CPU time are the same for both the reference as well as the AdapBi-ROM methods, first increasing and then decreasing as the $\hat{\sigma}$ value increases. A similar trend is seen for the optimal compliance for both methods.

Just like in the previous investigations, we give the optimal topology and stress iteration history as well as the volume fraction iteration history for different values of the allowable maximum stress based on the AdapBi-ROM method, as shown in Figs. 17, 18 and 19, respectively.

From Fig. 17 we can see that the optimal topologies using different values of $\hat{\sigma}$ based on the AdapBi-ROM method are clearly very different from each other. When using smaller values of $\hat{\sigma}$ ($\hat{\sigma} = 0.15, \hat{\sigma} = 0.3$), the final design structural branches are discontinuous and in a “fractured” state which means that the optimized design has not yet converged sufficiently despite having completed maximum iterations. Interestingly, we find a clear curvature forming at the corner of the L-shaped beam by using these two smaller values

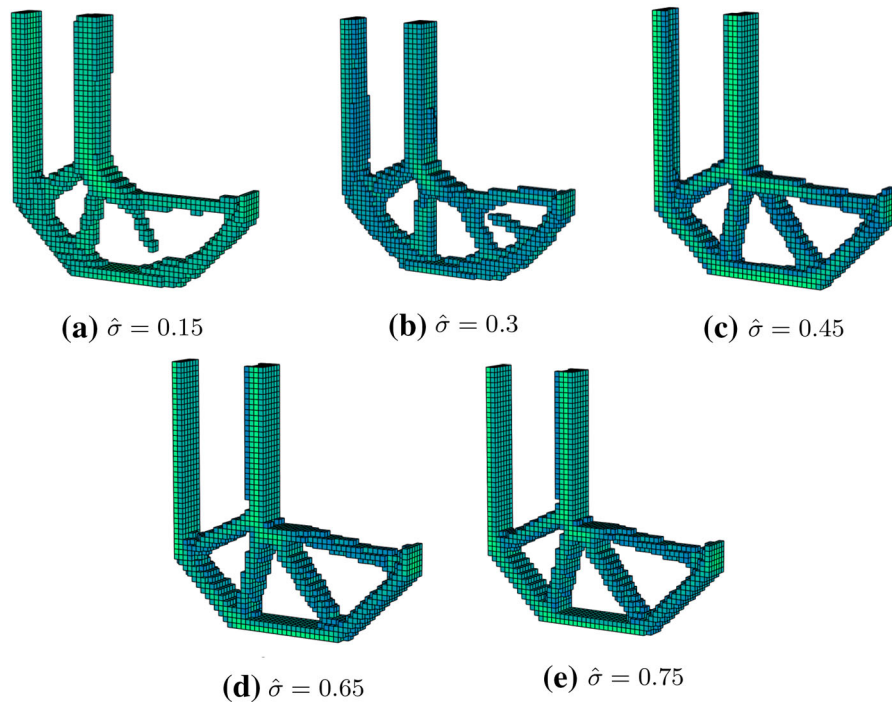


Fig. 17 Comparison of optimal topology using different value of $\hat{\sigma}$ based on AdapBi-ROM method

of $\hat{\sigma}$. As for other values of $\hat{\sigma}$, the optimal topologies are also different from each other (see the case of $\hat{\sigma} = 0.45$ and $\hat{\sigma} = 0.65$).

From Fig. 18 we can see that the stress constraint is not satisfied in the case of the smallest value of $\hat{\sigma}$ ($\hat{\sigma} = 0.15$). Also the maximum von Mises stress constraint is increasingly easier to satisfy during the procedure as $\hat{\sigma}$ increases, but we note that an overtly large value of $\hat{\sigma}$ could potentially render the stress constraint inactive (as seen in the case of $\hat{\sigma} = 0.65$ and $\hat{\sigma} = 0.75$).

From Fig. 19 we can see that the volume constraint is not satisfied for smaller values of $\hat{\sigma} = 0.15$ or $\hat{\sigma} = 0.3$ despite having been satisfied for larger values of $\hat{\sigma}$.

Finally, we compare the compliance iteration history using different values of $\hat{\sigma}$ and the AdapBi-ROM method in Fig. 20, where the compliance clearly has different convergence rate depending on the values of $\hat{\sigma}$ and a relatively large value of $\hat{\sigma}$ could potentially accelerate convergence.

Conclusions and perspectives

In this paper, we first presented two novel ROM-based Approximate Reanalysis approaches: on-the-fly ROM using Principal Components Analysis (SF-ROM) and multi-fidelity ROM using multigrid method (MF-ROM) and applied them to the non-self-adjoint stress-constrained topology optimization problem, and assessed their performance in terms of efficiency and reduction in computation time, using a 2D benchmark problem. Furthermore, we improved the second approach with two improved strategies: a so-called Adap-ROM method where the residual threshold of the equilibrium equations is adaptively selected according to the relative error of the objective function value during the iteration; and the so-called Bi-ROM method where the ROM is simultaneously applied

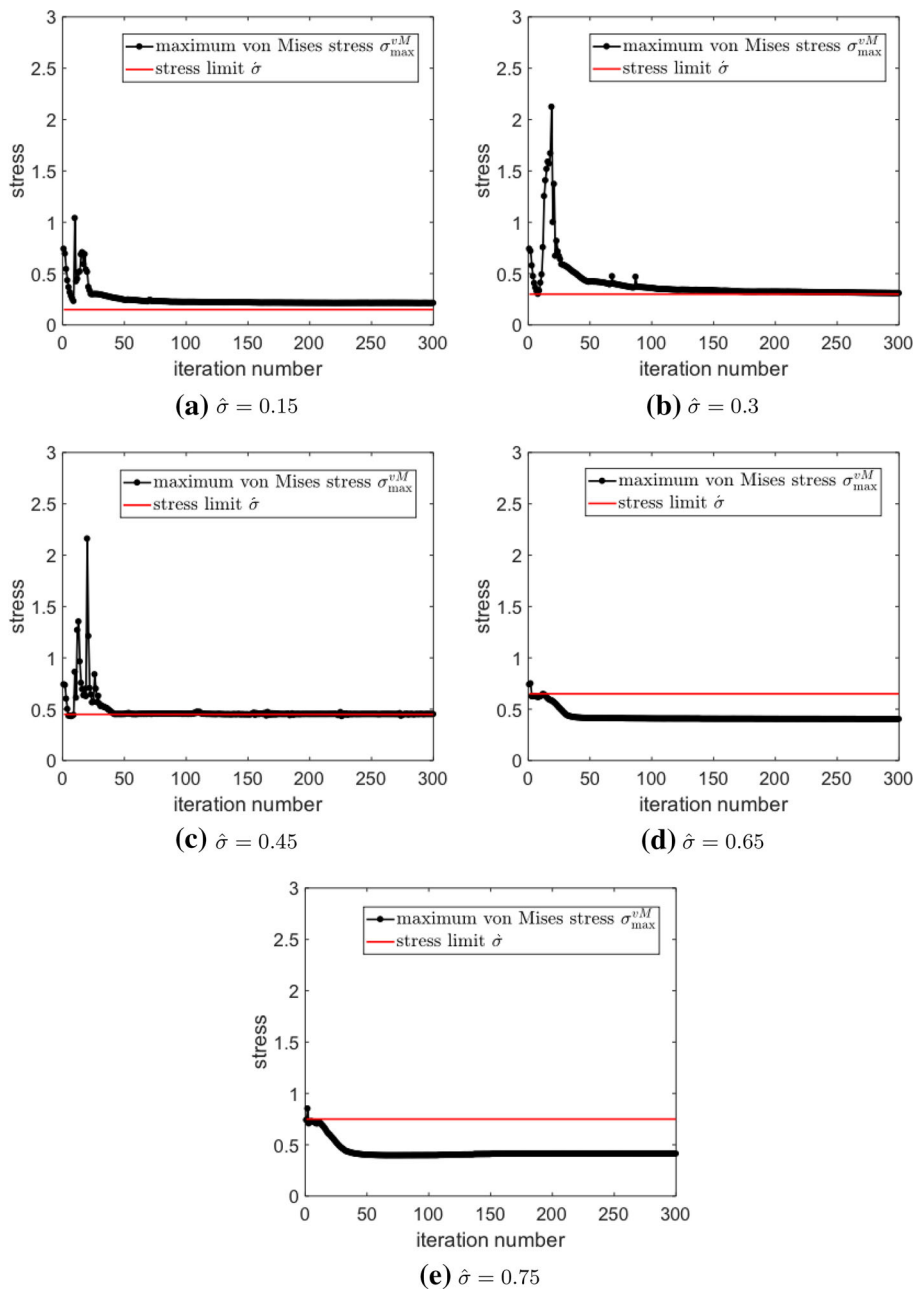


Fig. 18 Comparison of stress iterative history using different value of $\hat{\sigma}$ based on AdapBi-ROM method

to both the equilibrium equations as well as the adjoint equations obtained from the sensitivity analysis of the stress constraint. The first improved strategy achieves a balance between efficiency and accuracy, while the second improved strategy can further reduce the computational effort needed. To amplify the effect of the ROM, we combined both strategies for this 2D numerical example showing a good balance between efficiency and accuracy.

We then tested various versions of the ROM methods in this paper through a 3D L-shaped beam. Then using the combined version of Adap-ROM and Bi-ROM, we studied various parameters associated with the stress constraint, including the P-norm factor, the

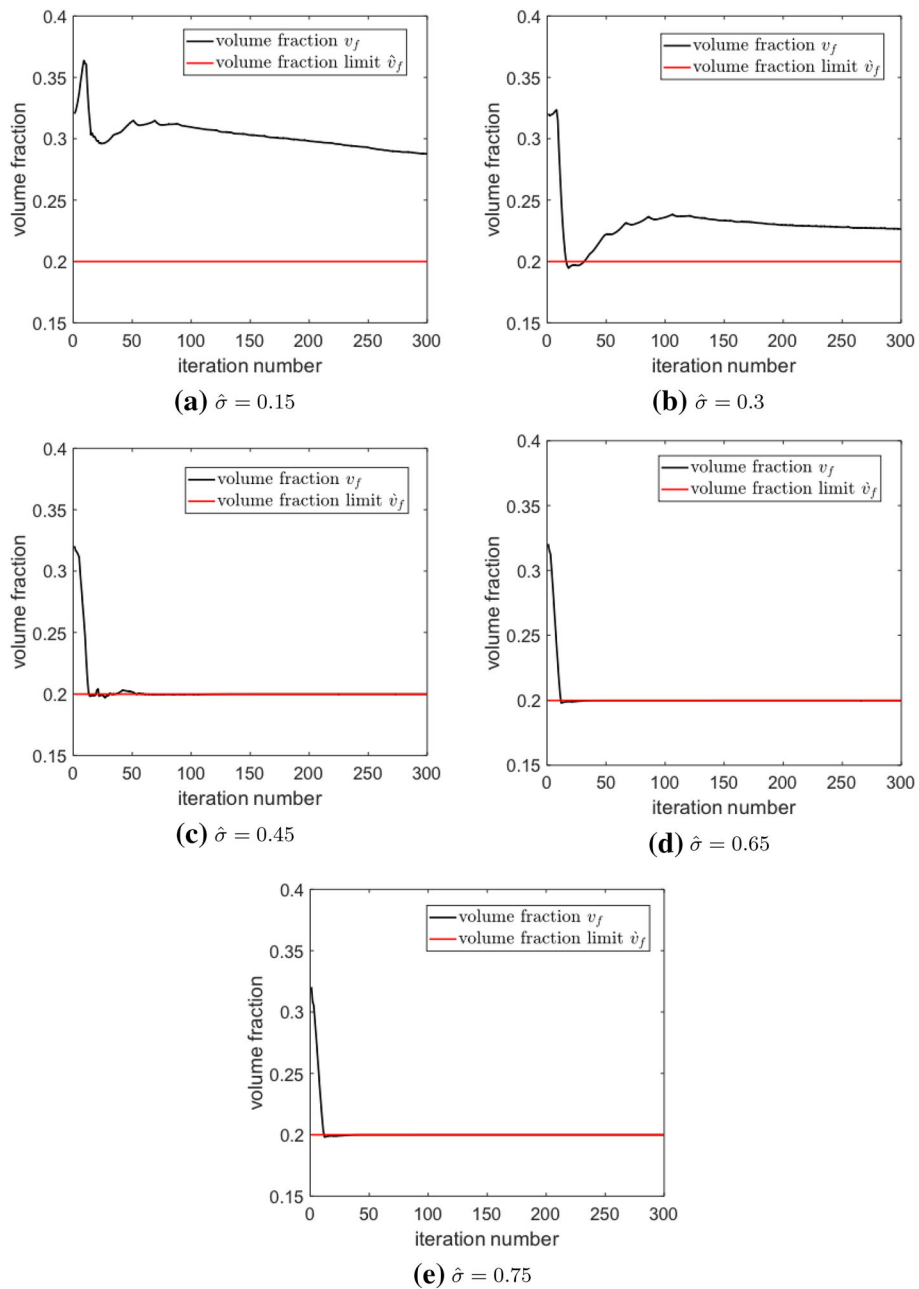


Fig. 19 Comparison of volume fraction iterative history using different value of $\hat{\sigma}$ based on AdapBi-ROM method

stress penalty factor as well as the allowable maximum stress value, in order to better understand the influence of the stress constraint on the ROM-based optimization results. We showed that the aforementioned parameters have a significant effect on the final optimized results (efficiency, structural performance etc), and therefore a judicious selection of these key parameters is required.

Furthermore, from the point of view of storage, the amount of storage required for the proposed on-the-fly ROM approach is much smaller compared to that needed for the full-field problem with finite element simulation calculation during the optimization

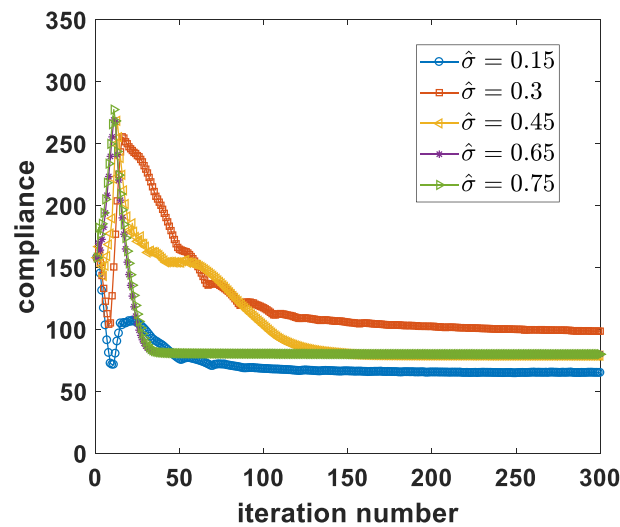


Fig. 20 Comparison of compliance iterative history using different values of $\hat{\sigma}$ based on AdapBi-ROM method

process. The proposed method only needs to store a small dimension snapshot matrix of small dimensionality to obtain the reduced basis. The only thing that needs to be done sequentially is the on-the-fly updating of the reduced basis using the residual errors, in order to ensure the quality of approximation.

A logical extension of this work would focus on efficient large-scale dynamic topology optimization. In addition, other non-self-adjoint problems besides the numerical examples presented in this paper need to be addressed.

Acknowledgements

The first author expresses her sincere gratitude to the associate professor Kai Long at North China Electric Power University for his help in testing the codes in its initial stages.

Author contributions

MX, WZ developed the main idea of the two novel approaches and introduced JM, and DL to implement benchmark test case, analyze numerical results and prepare the concept manuscript. BR provided his feedback on the manuscript. All authors read and approved the final manuscript.

Authors' information

Dr. Manyu Xiao, Associate professor, School of Mathematics and Statistics of Northwestern Polytechnical University, Xi'an, Shaanxi, People's Republic of China. Mr. Jun Ma, Dongcheng Lu, Master students, School of Mathematics and Statistics of Northwestern Polytechnical University, Xi'an, Shaanxi, People's Republic of China. Dr. Balaji Raghavan, Associate professor, Laboratoire de Génie Civil et Génie Mécanique EA 3913, Institut National des Sciences Appliquées de Rennes, Rennes, France. Prof. Weihong Zhang, Academician of the Chinese Academy of Sciences, State IJR Center of Aerospace Design and Additive Manufacturing, Northwestern Polytechnical University, Xi'an, 710072, People's Republic of China.

Funding

This research study was supported by the National Natural Science Foundation of China (Grant No. 12032018, Grant No. 11620101002 and Grant No. 11972166) and the Natural Science Foundation of Shaanxi Province, China (Grant No. 2021JM-043).

Availability of data materials

The data sets and models used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 27 November 2021 Accepted: 20 July 2022

Published online: 06 September 2022

References

- Sigmund O. A 99 line topology optimization code written in matlab. *Structural and Multidisciplinary Optimization*. 2001;21(2):120–7.
- Andreassen E, Clausen A, Schevenels M, Sigmund O. Efficient topology optimization in matlab using 88 lines of code. *Structural and Multidisciplinary Optimization*. 2011;43(1):1–16.
- Rietz A. Sufficiency of a finite exponent in simp (power law) methods. *Structural and Multidisciplinary Optimization*. 2001;21(2):159–63.
- E. Holmberg, T. Bo, and A. Klarbring. Stress constrained topology optimization. *Springer-Verlag*, 48(1), 2013.
- Le C, Norato J, Bruns T. Stress-based topology optimization for continua. *Structural and Multidisciplinary Optimization*. 2010;41(4):605–20.
- Cheng GD, Guo X. ϵ -relaxed approach in structural topology optimization. *Structural optimization*. 1997;13(4):258–66.
- Cheng GD, Jiang Z. Study on topology optimization with stress constraints. *Engineering Optimization*. 1992;20(2):129–48.
- P. Duysinx and M. P. Bendsøe. Topology optimization of continuum structures with local stress constraints. *International Journal for Numerical Methods in Engineering*, 1998.
- P. Duysinx and O. Sigmund. *New developments in handling stress constraints in optimal material distribution*, pages 1501–1509. AIAA, 2008.
- M. Bruggi and P. Venini. A mixed fem approach to stress constrained topology optimization. *International Journal for Numerical Methods in Engineering*, 73:1693 – 1714, 03 2008.
- Bruggi M. On an alternative approach to stress constraints relaxation in topology optimization. *Structural and Multidisciplinary Optimization*. 2008;36(2):125–41.
- Verbart A, Langelaar M, Keulen FV. A unified aggregation and relaxation approach for stress-constrained topology optimization. *Structural & Multidisciplinary Optimization*. 2017;55(2):1–17.
- París J, Navarrina F, Colominas I, Casteleiro M. Topology optimization of continuum structures with local and global stress constraints. *Structural and Multidisciplinary Optimization*. 2009;39(4):419–37.
- Yang RJ, Chen CJ. Stress-based topology optimization *Structural optimization*. 1996;12(2):98–105.
- J. París, F. Navarrina, I. Colominas, and M. Casteleiro. Block aggregation of stress constraints in topology optimization of structures. *Advances in Engineering Software*, 41(3):433–441, 2010. *Advances in optimum engineering design*.
- Yang DX, Liu HL, Zhang WS, et al. Stress-constrained topology optimization based on maximum stress measures. *Computers & Structures*. 2018;198:23–39.
- J. H. Rong, T. T. Xiao, and L. H. Yu et al. Continuum structural topological optimizations with stress constraints based on an active constraint technique. *International Journal for Numerical Methods in Engineering*, 108(4), 2016.
- P. Duysinx, L. Van Miegroet, E. Lemaire, and O. Brûls et al. Topology and generalized shape optimization: Why stress constraints are so important? <https://doi.org/10.1051/ijsmdo/2008034>, 2, 12 2008.
- Wang MY, Li L. Shape equilibrium constraint: a strategy for stress-constrained structural topology optimization. *Structural and Multidisciplinary Optimization*. 2013;47(3):335–52.
- Wang S, Sturler ED, Paulino GH. Large-scale topology optimization using preconditioned krylov subspace methods with recycling. *International Journal for Numerical Methods in Engineering*. 2007;69(12):2441–68.
- Mitjana F, Caferri S, Bugarin F, Segonds S, Castanie F, Duysinx P. Topological gradient in structural optimization under stress and buckling constraints. *Applied Mathematics and Computation*. 2021;409: 126032.
- Collet M, Noël L, Bruggi M, Duysinx P. Topology optimization for microstructural design under stress constraints. *Structural and Multidisciplinary Optimization*. 2018;58(6):2677–95.
- Noël L, Duysinx P. Shape optimization of microstructural designs subject to local stress constraints within an xfm-level set framework. *Structural and Multidisciplinary Optimization*. 2017;55(6):2323–38.
- Amir O, Stolpe M, Sigmund O. Efficient use of iterative solvers in nested topology optimization. *Structural and Multidisciplinary Optimization*. 2010;42(1):55–72.
- C. Gogu. Improving the efficiency of large scale topology optimization through on-the-fly reduced order model construction. *International Journal for Numerical Methods in Engineering*, 2015.
- Y. Choi, G. Oxberry, D. White, T. Kirchdoerfer. Accelerating topology optimization using reduced order models. In *The World Congress of Structural and Multidisciplinary Optimization*, May 20–24. Beijing. China. 2019;04:2019.
- Amir O, Bendsøe MP, Sigmund O. Approximate reanalysis in topology optimization. *International Journal for Numerical Methods in Engineering*. 2010;78(12):1474–91.
- Xiao M, Lu D, Breitkopf P, Raghavan B, Dutta S, Zhang W. On-the-fly model reduction for large-scale structural topology optimization using principal components analysis. *Structural and Multidisciplinary Optimization*. 2020;62(1):209–30.
- M. Xiao, D. Lu, P. Breitkopf, B. Raghavan, and S. Dutta. Multi-grid reduced-order topology optimization. *Structural and Multidisciplinary Optimization*, 2020.
- Svanberg K. The method of moving asymptotes—a new method for structural optimization. *International Journal for Numerical Methods in Engineering*. 1987;24(2):359–73.
- Bendsøe MP. Optimal shape design as a material distribution problem. *Structural optimization*. 1989;1(4):193–202.
- Andreassen E, Clausen A, Schevenels M, et al. Efficient topology optimization in matlab using 88 lines of code. *Structural and Multidisciplinary Optimization*. 2011;43(1):1–16.
- Aage N, Andreassen E, Lazarov B, Sigmund O. Giga-voxel computational morphogenesis for structural design. *Nature*. 2017;550(7674):84–6.

34. Wang D, Lee J, Holland K, Bibby T, Beaudoin S, Cale T. Von mises stress in chemical-mechanical polishing processes. *Journal of The Electrochemical Society*. 1997;144(3):1121–7.
35. Phalippou P, Bouabdallah S, Breikopf P, Villon P, Zarroug M. 'on-the-fly' snapshots selection for proper orthogonal decomposition with application to nonlinear dynamics. *Computer Methods in Applied Mechanics and Engineering*. 2020;367: 113120.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.