Consiglio Nazionale delle Ricerche
Istituto di Matematica Applicata
e Tecnologie Informatiche
"Enrico Magenes"

REPORT SERIES

Andrea Montanino, Gianluca Alaimo, Ettore Lanzarone

A gradient-based optimization method with functional principal component analysis for efficient structural topology optimization

20-02

IMATI REPORT Series

Nr. 20-02
30 June, 2020

Managing Editor
Michela Spagnuolo

Editorial Office
Istituto di Matematica Applicata e Tecnologie Informatiche "*E. Magenes*"
Consiglio Nazionale delle Ricerche
Via Ferrata, 5/a
27100 PAVIA (Italy)
Email: reports@imati.cnr.it
http://www.imati.cnr.it

IMATI Report Series Nr. 20-02
30 June, 2020

A gradient-based optimization method with functional principal component analysis for efficient structural topology optimization

Andrea Montanino, Gianluca Alaimo, Ettore Lanzarone

Andrea Montanino
Dipartimento di Strutture per l'Ingegneria e l'Architettura
Università degli Studi di Napoli Federico II, Napoli, Italy
Email: andrea.montanino@unina.it

Gianluca Alaimo
Dipartimento di Ingegneria Civile e Architettura - Sezione Strutture e Materiali
Università di Pavia, Pavia, Italy
Email: gianluca.alaimo@unipv.it

Ettore Lanzarone
Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes"
Consiglio Nazionale delle Ricerche, Milano, Italy
Email: ettore.lanzarone@cnr.it

Abstract.

Structural Topology Optimization (STO) is usually treated as a constrained minimization problem, which is iteratively addressed by solving the equilibrium equations for the problem under consideration. To reduce the computational effort, several reduced basis approaches that solve the equilibrium equations in a reduced space have been proposed. In this work, we apply Functional Principal Component Analysis (FPCA) to generate the reduced basis, and we couple FPCA with a gradient-based optimization method for the first time in the literature. The proposed algorithm has been tested on a large STO problem with 4.8 million degrees of freedom. Results show that the proposed algorithm achieves significant computational time savings with negligible loss of accuracy. Indeed, the density maps obtained with the proposed algorithm capture the larger features of maps obtained without reduced basis, but in significantly lower computational times, and are associated with similar values of the minimized compliance.

Keywords: *Structural topology optimization, Gradient-based approach, Functional principal component analysis, Computational efficiency*

A GRADIENT-BASED OPTIMIZATION METHOD WITH FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS FOR EFFICIENT STRUCTURAL TOPOLOGY OPTIMIZATION

ANDREA MONTANINO[1] , GIANLUCA ALAIMO[2] , AND ETTORE LANZARONE[3]

[1]Department of Structures for Engineering and Architecture,
University of Naples "Federico II", Naples, Italy

[2]Department of Civil Engineering and Architecture,
University of Pavia, Pavia, Italy

[3]Institute for Applied Mathematics and Information Technologies (IMATI),
National Research Council of Italy (CNR), Milan, Italy

Abstract. Structural Topology Optimization (STO) is usually treated as a constrained minimization problem, which is iteratively addressed by solving the equilibrium equations for the problem under consideration. To reduce the computational effort, several reduced basis approaches that solve the equilibrium equations in a reduced space have been proposed. In this work, we apply Functional Principal Component Analysis (FPCA) to generate the reduced basis, and we couple FPCA with a gradient-based optimization method for the first time in the literature. The proposed algorithm has been tested on a large STO problem with 4.8 million degrees of freedom. Results show that the proposed algorithm achieves significant computational time savings with negligible loss of accuracy. Indeed, the density maps obtained with the proposed algorithm capture the larger features of maps obtained without reduced basis, but in significantly lower computational times, and are associated with similar values of the minimized compliance.

Key words. Structural topology optimization, Gradient-based approach, Functional principal component analysis, Computational efficiency

1. Introduction. The need for optimized solutions in structural applications has increased over the years, and has become nowadays fundamental because of the limited availability of commodities, environmental impacts, increasingly stringent industrial time-to-market requests, and emerging manufacturing processes as additive manufacturing. In this framework, Structural Topology Optimization (STO) aims at obtaining optimized performance from a structure while satisfying some functional constraints, e.g., bounds on total mass or stresses. In particular, we focus on STO, whose goal is to produce optimized structures by determining an optimal mass or volume distribution in a given design domain. Unlike other alternatives as shape and size optimization, which deal with predefined configurations, STO can virtually produce any mass distribution within the design domain.

In the literature, STO is treated as a constrained minimization problem, for which several approaches and numerical schemes have been proposed [11, 13, 8]. Almost all the proposed approaches are iterative, which means that the optimized mass distribution is determined by repeatedly performing structural Finite Element Analyses (FEA) that involve the solution of the equilibrium equations for the problem under consideration. On the one hand, iterative approaches are very effective because they allow to obtain optimized solutions in a variety of situations without requiring additional assumptions. On the other hand, however, their bottleneck is the computational effort, as they require solving FEA a large number of times. For example, in a minimum compliance problem, up to 97% of the total computational time could be spent for numerically solving the equilibrium equations [1].

From these considerations, it becomes evident the need for reducing the computational time to solve the equilibrium equations in STO, which is a key requirement for large problems and three-dimensional applications. Reduced basis approaches represent a valid solution to such an issue, because they allow to reduce the dimensionality of the structural problem, i.e., the number of equilibrium equations to solve and, therefore, the related computational effort.

However, reduced basis approaches have been exploited in a very limited number of recent works. Gogu [7] constructed reduced basis on-the-fly by using the previously generated solutions of the equilibrium equations, and the reduced basis is adaptively enriched based on the STO convergence. Ferro *et al.* [6] applied the Principal Component Analysis (PCA) to the density map, obtaining an efficient numerical scheme for STO. Xiao *et al.* [15] proposed a Proper Orthogonal Decomposition (POD) scheme, related to PCA, to construct a reduced basis for any FEA solution in terms of displacements, to be used during the optimization; then, the reduced basis is adaptively updated on-the-fly according to an error metric. Less recently, Wang *et al.* [14] proposed a method based on the Krylov subspaces, while Amir *et al.* [2] proposed the construction of a reduced order model using the combined approximations method. Finally, Yoon [16] used eigenmodes along with Ritz vectors for the construction of a reduced model in dynamic STO.

In this work, we propose and test a reduced basis method that relies on Functional PCA (FPCA). FPCA for FEA is quite new in STO, i.e., it has only been exploited for the variable thickness sheet design problem in combination with a Simulated Annealing optimization heuristic [1]; moreover, it was initially developed for uncertainty quantification [4]. In this work, we extend the use of FPCA in STO by applying it to the class of gradient-based optimization methods, which are the most commonly used. We demonstrate that significant savings in STO computational times can be achieved with negligible loss of accuracy.

More specifically, we employ FPCA along with a gradient-based approach, considering the code proposed by Sigmund [10] and his gradient-based approach as a starting point for incorporating FPCA reduced basis. At each STO iteration, we convert the FEA problem into a *reduced* one by projecting onto a smaller space by means of a data-driven FPCA reduced basis. The reduced problem is solved and, if error measures associated with the solution are below given thresholds, the reduced solution is accepted; otherwise, the original FEA problem is solved and the complete solution is used to update the FPCA basis.

The remainder of the paper is structured as follows: the proposed gradient-based optimization method with FPCA is presented in Sect. 2; the test case used to validate our algorithm is described in Sect. 3, while the results are reported in Sect. 4; finally, discussions and conclusion of the work are drawn in Sect. 5.

2. Gradient-based optimization with FPCA. In this Section, we present the addressed STO problem (Sect. 2.1), the gradient-based optimization of [10] (Sect. 2.2), and the novel contribution included in the optimization, i.e., the FPCA to generate the reduced basis (Sect. 2.3). Finally, we detail the overall algorithm in Sect. 2.4. Without loss of generality, we present the approach for two-dimensional cases; from a conceptual point of view, the approach is the same also for three-dimensional cases while, from an operational point of view, the extension is very straightforward.

2.1. Structural topology problem. We address a STO problem in which a given mass is distributed over the domain Ω in order to minimize the structural

compliance

$$c = \int_{\Omega} \sigma \cdot \varepsilon \, d\Omega$$

where ε is the strain tensor, $\sigma = \mathbb{D}\varepsilon$ the stress tensor, and $\mathbb{D}$ the fourth-order linear elastic tensor. We assume that $\mathbb{D}$ depends on the mass density distribution as $\mathbb{D} = \rho^p \, \tilde{\mathbb{D}}$, where $\rho \in [\rho_{min}, 1]$ is the mass density, $p \geq 1$ a penalty coefficient, $\tilde{\mathbb{D}}$ the fourth order elastic tensor of the isotropic homogeneous bulk material, and ρ_{min} a positive value close to 0.

Finally, ε is defined in terms of the displacement field $\mathbf{u}$ as $\varepsilon = \frac{1}{2}\left(\nabla \mathbf{u} + \nabla \mathbf{u}^T\right)$, where superscript ^{T} denotes transposition.

The problem is discretized following a classical FEA approach, using four-node elements with bilinear displacement functions. Accordingly, the compliance c is expressed in the form

$$c = \widehat{\mathbf{u}}^T \mathbf{f} = \sum_{i=1}^{N_{el}} \widehat{\mathbf{u}}_i^T \mathbf{f}_i = \sum_{i=1}^{N_{el}} \widehat{\mathbf{u}}_i^T \mathbf{K}_i \widehat{\mathbf{u}}_i \tag{2.1}$$

where N_{el} is the number of elements in the discretized domain, $\widehat{\mathbf{u}}_i$ the column vector of the nodal displacements of element i, $\mathbf{f}_i$ the column vector of the nodal forces of element i, and $\mathbf{K}_i$ the stiffness matrix of element i, while $\widehat{\mathbf{u}}$ and $\mathbf{f}$ are the assembled (over all elements) displacement vector and the applied force vector, respectively. Assuming a uniform mass density ρ_i over each element i, $\mathbf{K}_i$ depends on ρ_i in the form $\mathbf{K}_i = \rho_i^p \tilde{\mathbf{K}}_i$, where $\tilde{\mathbf{K}}_i$ is the stiffness matrix of element i analogous to $\tilde{\mathbb{D}}$, which corresponds to a homogeneous and isotropic material.

The minimization of c, expressed as in (2.1), is subject to mass (volume) constraint expressed in the form

$$\begin{aligned} &\sum_{i=1}^{N_{el}} \rho_i V_i = \rho^* V \\ &\rho_i \in [\rho_{min}, 1] \qquad \forall i = 1, \ldots, N_{el} \end{aligned}$$

where ρ^* is the imposed mean density that uniquely defines the overall mass, and V_i and V are the volume of element i and the overall domain volume, respectively.

Finally, the equilibrium equations are expressed in compact form as

$$\mathbf{K}\widehat{\mathbf{u}} = \mathbf{f} \tag{2.2}$$

where $\mathbf{K}$ is the assembled stiffness matrix.

2.2. Gradient-based optimization. The gradient-based minimization is an iterative procedure in which, at each iteration, the new solution in terms of ρ_i in all elements i is obtained following an optimality criterion, as proposed in [3] and applied in [10].

The optimality criterion is based on the sensitivity of c with respect to the densities ρ_i in the form

$$\frac{\partial c}{\partial \rho_i} = \frac{\partial \widehat{\mathbf{u}}^T}{\partial \rho_i} \mathbf{f} \tag{2.3}$$

However, there is no explicit formula that links the displacements to the density distribution. For this reason, the computation of derivatives (2.3) is performed by adding an adjoint term to c, in the form

$$c = \widehat{\mathbf{u}}^T\mathbf{f} = \mathbf{f}^T\widehat{\mathbf{u}} = \mathbf{f}^T\widehat{\mathbf{u}} - \widetilde{\mathbf{u}}^T(\mathbf{K}\widehat{\mathbf{u}} - \mathbf{f}) \tag{2.4}$$

where $\widetilde{\mathbf{u}}$ is an arbitrary term, and the difference in parentheses is equal to 0 thanks to (2.2). The gradient of c is then expressed, for each component, as

$$\frac{\partial c}{\partial \rho_i} = (\mathbf{f}^T - \widetilde{\mathbf{u}}^T\mathbf{K})\frac{\partial \widehat{\mathbf{u}}}{\partial \rho_i} - \widetilde{\mathbf{u}}^T\frac{\partial \mathbf{K}}{\partial \rho_i}\widehat{\mathbf{u}} \tag{2.5}$$

which is simplified if the arbitrary term $\widetilde{\mathbf{u}}$ is chosen equal to the actual displacement field $\widehat{\mathbf{u}}$, turning into

$$\frac{\partial c}{\partial \rho_i} = -\widetilde{\mathbf{u}}^T\frac{\partial \mathbf{K}}{\partial \rho_i}\widehat{\mathbf{u}} \tag{2.6}$$

Considering the expression of the stiffness matrix $\mathbf{K}$ and that ρ_i is constant over each element i, the gradient of the compliance can be computed element-wise and expressed as

$$\frac{\partial c}{\partial \rho_i} = -p\widehat{\mathbf{u}}_i^T\rho_i^{(p-1)}\mathbf{K}_i\widehat{\mathbf{u}}_i \tag{2.7}$$

Let us remark that (2.7) is valid when $\widehat{\mathbf{u}}$ represents the exact solution of the equilibrium equations. This is of particular importance when considering a solution $\widehat{\mathbf{u}}_{rec}$ reconstructed from the reduced space, which can be slightly different from the exact solution $\widehat{\mathbf{u}}$; in this case, the complete expression (2.5) must be considered, or the problem must be adapted in order to use (2.7). In this work, as detailed in Sects. 2.3 and 2.4, we consider the second alternative and we solve the problem with a modified force $\mathbf{f}_{curr}$ in such a way that $\mathbf{K}\widehat{\mathbf{u}} = \mathbf{f}_{curr}$.

With (2.7), it is possible to iteratively update ρ_i to minimize the compliance c step by step. In particular, the new density distribution can be computed with the following heuristic approach proposed by Bendsøe and Sigmund [3]

$$\rho_i^N = \begin{cases} \max(\rho_{min}, \rho_i - m) \\ \quad \text{if} \quad \rho_i\alpha_i^\eta \leq \max(\rho_{min}, \rho_i - m) \\ \rho_i\alpha_i^\eta \\ \quad \text{if} \quad \max(\rho_{min}, \rho_i - m) \leq \rho_i\alpha_i^\eta \leq \min(1, \rho_i + m) \\ \min(1, \rho_i + m) \\ \quad \text{if} \quad \min(1, \rho_i + m) \leq \rho_{min}\alpha_i^\eta \end{cases} \tag{2.8}$$

where ρ_i^N denotes the new density of element i, m is a positive coefficient that limits density variation, $\eta = \frac{1}{2}$ is a damping coefficient, and α_i is given by the optimality condition as

$$\alpha_i = -\frac{1}{\kappa}\frac{\partial c/\partial \rho_i}{\partial V/\partial \rho_i} \tag{2.9}$$

where κ is a Lagrangian multiplier found through a bisectioning algorithm [10].

2.3. FPCA for reduced model. The idea behind the use of FCPA in STO is to solve the equilibrium equations (2.2) in a reduced space for as many gradient-based iterations as possible. Such a reduced space, whose dimension is given by a limited number of principal displacement components, is generated by applying FPCA to a set of previously obtained solutions [1, 4].

To this aim, let us consider M displacement solutions $\widehat{\mathbf{u}}(\boldsymbol{\rho}_m)$ of the equilibrium equations (with $m = 1, \ldots, M$) corresponding to several realizations $\boldsymbol{\rho}_m$ of $\boldsymbol{\rho}$, where $\boldsymbol{\rho}$ denotes the vector including all ρ_i.

The subspace spanned by these M solutions represents an approximation of the entire solution space $\mathcal{U}(\Theta) = \{\mathbf{u}(\boldsymbol{\rho}), \boldsymbol{\rho} \in \Theta\}$, where $\Theta = [\rho_{min}, 1]^{N_{el}}$ gathers all possible solutions of the structural problem. By applying FPCA to the subspace, we obtain the set of the K orthonormal basis functions (with $1 \leq K \leq M$) that optimally approximate $\mathcal{U}(\Theta)$ in terms of total explained variance.

These basis functions can be obtained by solving an eigenvalue problem that involves the covariance operator. Let us denote by $\{v_m(\mathbf{x}), m = 1, \ldots, M\}$ a set of M functions with cross-sectional mean $\frac{1}{M}\sum_{m=1}^{M} v_m(\mathbf{x})$ equal to 0 for each $\mathbf{x}$. The corresponding empirical covariance function is

$$\mathbb{V}(\mathbf{s}, \mathbf{t}) = \frac{1}{M}\sum_{m=1}^{M} v_m(\mathbf{s}) v_m(\mathbf{t})$$

and the covariance operator is

$$\mathbf{V}\xi : \xi(\cdot) \mapsto \int_{\Omega} \mathbb{V}(\cdot, \mathbf{t})\xi(\mathbf{t}) d\mathbf{t}$$

Based on that, each basis function solves the following eigenproblem:

$$\mathbf{V}\xi(\mathbf{x}) = \gamma\xi(\mathbf{x}) \tag{2.10}$$

Operatively, we start from M displacement solutions $\widehat{\mathbf{u}}(\boldsymbol{\rho}_1), \ldots, \widehat{\mathbf{u}}(\boldsymbol{\rho}_M)$ corresponding to different $\boldsymbol{\rho}_m$. We build a $M \times 2L$ matrix $\mathbb{U} = [\widehat{\mathbf{u}}(\boldsymbol{\rho}_1), \ldots, \widehat{\mathbf{u}}(\boldsymbol{\rho}_M)]^T$, where L is the number of nodes in the mesh and $2L$ the number of degrees of freedom.

The buffer matrix $\mathbb{U}$ is then split into two parts: the average part $\mathbb{U}_{av}$ and the residual part $\mathbb{U}_{cov} = \mathbb{U} - \mathbb{U}_{av}$ that stores the fluctuations of the solution with respect to the average.

The orthonormal basis is obtained with the functional principal components of $\mathbb{U}_{cov}$. We use the method of snapshots as in [1, 4], which is valid if $2L > M$ as in our case. Accordingly, the basis of rank K is obtained as follows:

1. Define a $2L \times M$ matrix $\mathbb{Y} = (\mathbb{U}_{cov}\, \mathbb{M}^{1/2})^T$ with

$$\mathbb{M} = \int_{\Omega} \phi_i \phi_j d\Omega$$

   where ϕ_k are the finite element shape functions, and $\mathbb{M}$ is a $2L \times 2L$ matrix that can also be intended as a mass matrix.
2. Solve the eigenvalue problem

$$\mathbb{Y}^T\mathbb{Y}\mathbf{v}_k = \lambda_k \mathbf{v}_k \qquad k = 1, \ldots, K$$

   where the eigenvalues λ_k are in descending order along k.

3. Choose the value of K by retaining all eigenfunctions that correspond to an eigenvalue $\lambda_k > \Lambda\,\lambda_1$, where λ_1 is the maximum eigenvalue and Λ assumes very small values [5].

The projection of (2.2) onto the reduced space of dimension K is carried out using the $2L \times K$ matrix $\mathbb{U}_{rb} = [\xi_1, \ldots, \xi_K]$, where

$$\xi_k = \frac{1}{\sqrt{\lambda_k}} \mathbb{U}_{cov}^T \mathbf{v}_k$$

Let now consider the problem $\mathbf{K}\widehat{\mathbf{u}} = \mathbf{f}_{curr}$, where $\mathbf{f}_{curr}$ is the modified force that allows to exploit (2.7). With the above projection, we may solve the following *reduced problem*, instead of the complete one, in the subspace spanned by the principal components:

$$\mathbf{K}_{red}\,\widehat{\mathbf{u}}_{rs} = \mathbf{f}_{red} \tag{2.11}$$

with

$$\mathbf{K}_{red} = \mathbb{U}_{rb}^T\,\mathbf{K}\,\mathbb{U}_{rb} \tag{2.12}$$

$$\mathbf{f}_{red} = \mathbb{U}_{rb}^T\,(\mathbf{f} - \mathbf{K}\bar{\mathbf{u}}) \tag{2.13}$$

where $\widehat{\mathbf{u}}_{rs}$ denotes the solution of the problem in the reduced space and

$$\bar{\mathbf{u}} = \frac{1}{M} \sum_{m=1}^{M} \widehat{\mathbf{u}}_i$$

is the average of the solutions stored in buffer $\mathbb{U}$.

After solving (2.11) in the reduced space, the displacement solution is reconstructed as

$$\widehat{\mathbf{u}}_{rec} = \bar{\mathbf{u}} + \mathbb{U}_{rb}\,\widehat{\mathbf{u}}_{rs} \tag{2.14}$$

where subscript *rec* indicates that the solution is *reconstructed* starting from $\widehat{\mathbf{u}}_{rs}$.

Though $\widehat{\mathbf{u}}_{rec}$ is computationally cheaper to obtain, it is only an approximation of the actual $\widehat{\mathbf{u}}$. Therefore, we must decide whether to accept $\widehat{\mathbf{u}}_{rec}$ or not based on the approximation error and the quality of the solution.

The approximation error is evaluated by computing the errors e_1 and e_2 between the reconstructed force $\mathbf{f}_{rec} = \mathbf{K}\widehat{\mathbf{u}}_{rec}$ and $\mathbf{f}_{curr}$ and between $\mathbf{f}_{rec}$ and $\mathbf{f}$, respectively:

$$e_1 = (\mathbf{f}_{rec} - \mathbf{f}_{curr})^T \mathbb{M}_\rho\,(\mathbf{f}_{rec} - \mathbf{f}_{curr})$$

$$e_2 = (\mathbf{f}_{rec} - \mathbf{f})^T \mathbb{M}_\rho\,(\mathbf{f}_{rec} - \mathbf{f})$$

The global mass matrix $\mathbb{M}_\rho$ is introduced to weigh less or exclude the elements with lower density ρ_i from the computation of the errors. It is computed as the assembly of elementary mass matrices, obtained for each element as

$$\mathbb{M}_{\rho_k} = \rho_k \int_{\Omega_k} \phi_i \phi_j d\Omega \tag{2.15}$$

Finally, the compliance c generally decreases with the iterations in gradient-based approaches [10]. Therefore, to avoid too large deviations from this decreasing behavior, a reduced solution is also evaluated in terms of the difference δ_c between the compliance c obtained in the current iteration and that at the previous one.

Algorithm 1 Gradient-based optimization with FPCA

Require: Problem parameters ($\mathbf{f}$; ρ^*; ρ_{min}; p; r).
Require: Algorithm parameters (M_{min}; M_{max}; Λ; τ_1; τ_2; τ_3).
$M = 0$; $\mathbf{f}_{curr} \leftarrow \mathbf{f}$; $\boldsymbol{\rho} \leftarrow$ homogeneous with constant ρ^*;
for $g = 1$ **to** G **do**
  $\mathbf{K} \leftarrow$ computed with $\boldsymbol{\rho}$;
  if $g \leq M_{min}$ **then**
    $\widehat{\mathbf{u}} \leftarrow$ computed through complete problem with $\mathbf{K}$ and $\mathbf{f}$;
    $M \leftarrow M + 1$; $\mathbb{U}_{sol} \leftarrow \mathbb{U}_{sol} \cup \{\widehat{\mathbf{u}}\}$;
    if $g = M_{min}$ **then**
      $\mathbb{U}_{cov} \leftarrow$ computed with $\mathbb{U}_{sol}$;
    end if
    $c(g) \leftarrow$ computed with $\widehat{\mathbf{u}}$ and $\boldsymbol{\rho}$; $dc \leftarrow$ computed with $\widehat{\mathbf{u}}$ and $\boldsymbol{\rho}$;
  else
    $Admissible(g) \leftarrow 0$; $Accepted(g) \leftarrow 0$;
    $\mathbf{K}_{red} \leftarrow$ projection of $\mathbf{K}$ with $\mathbb{U}_{rb}$; $\mathbf{f}_{red} \leftarrow$ projection of $\mathbf{f}_{curr}$ with $\mathbb{U}_{rb}$;
    $\widehat{\mathbf{u}}_{rs} \leftarrow$ computed through reduced problem with $\mathbf{K}_{red}$ and $\mathbf{f}_{red}$;
    $\widehat{\mathbf{u}}_{rec} \leftarrow$ inverse projection of $\widehat{\mathbf{u}}_{rs}$ with $\mathbb{U}_{rb}$;
    $\mathbf{f}_{rec} \leftarrow \mathbf{K}\widehat{\mathbf{u}}_{rec}$;
    $e_1 \leftarrow (\mathbf{f}_{rec} - \mathbf{f}_{curr})^T \mathbb{M}_\rho \, (\mathbf{f}_{rec} - \mathbf{f}_{curr})$; $e_2 \leftarrow (\mathbf{f}_{rec} - \mathbf{f})^T \mathbb{M}_\rho \, (\mathbf{f}_{rec} - \mathbf{f})$;
    if $e_1 < \tau_1$ **and** $e_2 < \tau_2$ **then**
      $c(g) \leftarrow$ computed with $\widehat{\mathbf{u}}_{rec}$ and $\boldsymbol{\rho}$; $dc \leftarrow$ computed with $\widehat{\mathbf{u}}_{rec}$ and $\boldsymbol{\rho}$;
      $Admissible(g) \leftarrow 1$;
    end if
    if $Admissible(g) = 1$ **and** $c(g) - c(g-1) < \tau_3$ **then**
      $\mathbf{f}_{curr} \leftarrow \mathbf{f}_{rec}$;
      $Accepted(g) \leftarrow 1$;
    else
      $\mathbf{f}_{curr} \leftarrow \mathbf{f}$;
      $\widehat{\mathbf{u}} \leftarrow$ computed through complete problem with $\mathbf{K}$ and $\mathbf{f}_{curr}$;
      $M \leftarrow M + 1$;
      if $M \leq M_{max}$ **then**
        $\mathbb{U}_{sol} \leftarrow \mathbb{U}_{sol} \cup \{\widehat{\mathbf{u}}\}$; $\mathbb{U}_{cov} \leftarrow$ computed with $\mathbb{U}_{sol}$;
      else
        $\mathbb{U}_{sol} \leftarrow \mathbb{U}_{sol} \cup \{\widehat{\mathbf{u}}\} \setminus \{oldest\ solution \in \mathbb{U}_{sol}\}$;
        $\mathbb{U}_{cov} \leftarrow$ computed with $\mathbb{U}_{sol}$;
      end if
      $c(g) \leftarrow$ computed with $\widehat{\mathbf{u}}$ and $\boldsymbol{\rho}$; $dc \leftarrow$ computed with $\widehat{\mathbf{u}}$ and $\boldsymbol{\rho}$;
    end if
  end if
  $dc \leftarrow$ filtered with mesh-independency filter from dc with r (see [10]);
  $\boldsymbol{\rho} \leftarrow$ new field based on dc, ρ^*, ρ_{min} and p;
  if $g \geq M_{min}$ **and** $Accepted(g) = 0$ **then**
    $\mathbb{M} \leftarrow$ mass with $\boldsymbol{\rho}$; $\mathbb{M}_\rho \leftarrow$ global mass with $\boldsymbol{\rho}$;
    $\mathbb{U}_{rb} \leftarrow$ projection matrix computed with $\mathbb{U}_{cov}$ and Λ;
  end if
end for
Ensure: $\boldsymbol{\rho}$; $c(g)$ $\forall g = 1, \ldots, G$.

If $e_1 < \tau_1$, $e_2 < \tau_2$ and $\delta_c < \tau_3$, where τ_1, τ_2 and τ_3 are fixed positive thresholds, the reconstructed solution $\mathbf{u}_{rec}$ is accepted. Otherwise, it is discarded and the complete problem (2.2) is solved.

2.4. Overall integrated algorithm. The algorithm reflects the iterative structure of the gradient-based optimization in [10], enriched by the proposed reduced basis approach based on FPCA. The basic structure of each iteration g (with $g = 1, \ldots, G$) consists of solving equilibrium equations for the current density solution $\boldsymbol{\rho}_g$, and computing the new solution $\boldsymbol{\rho}_{g+1}$ according to (2.8).

The buffer $\mathbb{U}$ contains a number M of solutions between M_{min} and M_{max}. This basic structure of the gradient-based optimization, which consists of solving the complete problem (2.2), is repeated for the first M_{min} iterations while considering the nominal force $\mathbf{f}$ of the problem and storing the displacement solutions in $\mathbb{U}$. At iteration $g = M_{min}$, the matrix $\mathbb{U}_{rb}$ is also built from $\mathbb{U}$.

Starting from iteration $g = M_{min} + 1$, the problem is projected onto the reduced space exploiting $\mathbb{U}_{rb}$, and the equilibrium equations are solved in the reduced space according to (2.11)-(2.14).

As mentioned, the reconstructed solution $\widehat{\mathbf{u}}_{rec}$ is accepted if $e_1 < \tau_1$, $e_2 < \tau_2$ and $\delta_c < \tau_3$. In this case, the reconstructed force $\mathbf{f}_{rec}$ is also kept as the force $\mathbf{f}_{curr}$ to solve the problem at the next iteration.

If $\widehat{\mathbf{u}}_{rec}$ is not accepted, the complete problem (2.2) is solved while considering the nominal force $\mathbf{f}$. At the same time, the force $\mathbf{f}_{curr}$ to solve the problem at the next iteration is reset equal to $\mathbf{f}$. Moreover, the new complete solution is included in $\mathbb{U}$; if the number of already stored solutions is equal to M_{max} the oldest solution is also removed. The new $\mathbb{U}_{cov}$ is then computed with the updated $\mathbb{U}$.

At the end of each iteration g, the new density map $\boldsymbol{\rho}_{g+1}$ is generated according to the updating scheme (2.8), either considering the solution $\widehat{\mathbf{u}}$ of the complete problem (in the first M_{min} iterations or when the reduced solution is rejected) or the reconstructed solution $\widehat{\mathbf{u}}_{rec}$ (when the reduced solution is accepted starting from iteration $g = M_{min} + 1$).

A detailed pseudo-code of the algorithm is reported in Alg. 1, while the full code written in MATLAB language is available at `www.mi.imati.cnr.it/ettore/FPCA`.

The *dynamic* management of the buffer dimension M starting from M_{min} allows to calculate, and possibly accept, reduced solutions even in the very first iterations without waiting for the first M_{max} complete solutions. In fact, accepting reduced solutions even in the very first iterations is a key requirement for increasing computational efficiency. At the same time, the upper limit M_{max} allows to replace the oldest solution with the new one, thus giving preference to more recent displacement solutions that are associated with closer density distributions than the older ones.

As mentioned, due to the approximation induced by the reconstructed displacement solution $\widehat{\mathbf{u}}_{rec}$, the difference in parentheses in (2.4) is not exactly null when $\widehat{\mathbf{u}} \leftarrow \widehat{\mathbf{u}}_{rec}$, and expression (2.7) is not exact in relation to problem (2.2). In the literature, such an expression has been modified by adding additional terms that compensate the error, as in [7], but this approach requires additional computational effort at each iteration. In this work, we deal with the issue by replacing the nominal force $\mathbf{f}$ with $\mathbf{f}_{curr}$. In this way, expression (2.7) becomes correct for the problem $\mathbf{K}\widehat{\mathbf{u}} = \mathbf{f}_{curr}$ without the need to calculate additional terms. Obviously, the problem with $\mathbf{f}_{curr}$ is representative of (2.2) if the distance between $\mathbf{f}_{curr}$ and $\mathbf{f}$ is limited, as verified by the conditions $e_1 < \tau_1$ and $e_2 < \tau_2$ together.

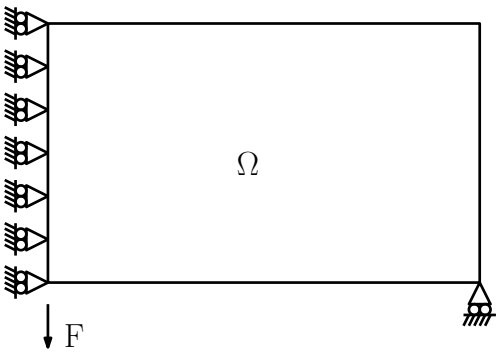

FIG. 3.1. *Structural scheme of the considered MBB beam with domain* Ω*, boundary conditions, symmetry boundary conditions on the left, and applied force* $\mathbf{F}$.

3. Experimental validation. We compare our algorithm with the case in which the reduced basis approach is not exploited and all gradient-based iterations are performed with the complete problem, as in [10].

Analyses have been conducted on the well-known Messerschmitt-B̈olkow-Blohm (MBB) beam problem. It consists of a simply-supported rectangular domain of size $2n_x \times n_y$ elements, with a vertical force $\mathbf{F}$ downward applied in the midpoint of the lower side (force module equal to 1). Thanks to the symmetry, only half of the domain is studied, as represented in Fig. 3.1. We consider $n_x = 2000$ and $n_y = 1200$ elements in the analyses, which define a computationally demanding problem with more than 4.8 million degrees of freedom.

Moreover, the following parameter values are assumed:

- penalty $p = 3$;
- minimum density $\rho_{min} = 0.001$;
- number M of solutions in the buffer between $M_{min} = 2$ and $M_{max} = 20$;
- parameter Λ to include eigenfunctions equal to 10^{-12};
- error threshold $\tau_1 = 10^{-4}$;
- error threshold $\tau_2 = 4\tau_1$;

Experiments have been conducted by considering the following three factors and related levels:

- mean density ρ^* for mass constraint: $\{0.3, 0.5, 0.7\}$;
- radius for filter r: $\{5, 10, 15\}$ nodes;
- increment threshold τ_3: $\{2, 4, 8\}$;

In addition to these experiments, different M_{min} and M_{max} values are tested for $\rho^* = 0.5$, $r = 5$ and $\tau_3 = 2$.

Each setting of ρ^* and r is also solved with no reduced basis to get a reference solution for comparison. In particular, a number of iterations $G = 100$ is assumed to get the reference compliance $\tilde{c}$ of the setting without reduced basis. This number G is large enough to guarantee convergence; indeed, we also tested a greater numbers of iterations, up to $G = 1000$, and we observed very limited variations of $\tilde{c}$, lower than

ρ^*	r	τ_3	Reference $\tilde{c}$ from complete	CPU time $[10^5 s]$ from complete with $G = 100$	To reach $\tilde{c}$ plus 5%		To reach $\tilde{c}$ plus 10%	
					CPU time $[10^5 s]$	% reduction	CPU time $[10^5 s]$	% reduction
0.3	5	2	94.41	4.70	1.14	38.29	0.83	55.27
		4			1.16	37.25	0.84	54.92
		8			0.98	47.41	0.77	58.48
	10	2	94.89	4.92	1.21	9.72	1.00	25.47
		4			—	—	—	—
		8			1.05	21.75	0.74	45.25
	15	2	96.09	5.14	—	—	—	—
		4			—	—	—	—
		8			—	—	—	—
0.5	5	2	65.02	5.12	0.93	22.77	0.81	32.78
		4			0.63	47.76	0.51	57.81
		8			0.86	28.11	0.64	46.24
	10	2	65.26	4.95	—	—	0.49	43.57
		4			0.61	30.80	0.55	37.00
		8			0.61	30.83	0.45	49.23
	15	2	65.71	5.20	0.70	14.94	0.47	43.21
		4			—	—	0.54	35.03
		8			0.77	7.17	0.48	41.89
0.7	5	2	53.39	5.04	0.43	52.14	0.33	63.17
		4			0.64	28.72	0.28	68.77
		8			0.51	43.60	0.33	62.86
	10	2	53.57	5.12	0.48	43.54	0.36	57.16
		4			0.69	17.67	0.30	64.23
		8			—	—	—	—
	15	2	53.76	5.34	0.48	38.02	0.36	53.14
		4			0.75	36.35	0.30	60.85
		8			—	—	—	—

TABLE 4.1

Results from the experiments for each combination of ρ_i, r and τ_3 with $M_{min} = 2$ and $M_{max} = 20$: reference $\tilde{c}$ with $G = 100$ iterations without reduced basis and corresponding CPU time; time to reach the reference $\tilde{c}$ plus 5% and 10% with the proposed algorithm and percentage reduction with respect to the computational time to reach the same value without reduced basis. "—" denotes that $\tilde{c}$ plus both 5% or 10% was not reached by our algorithm in $G = 100$ iterations.

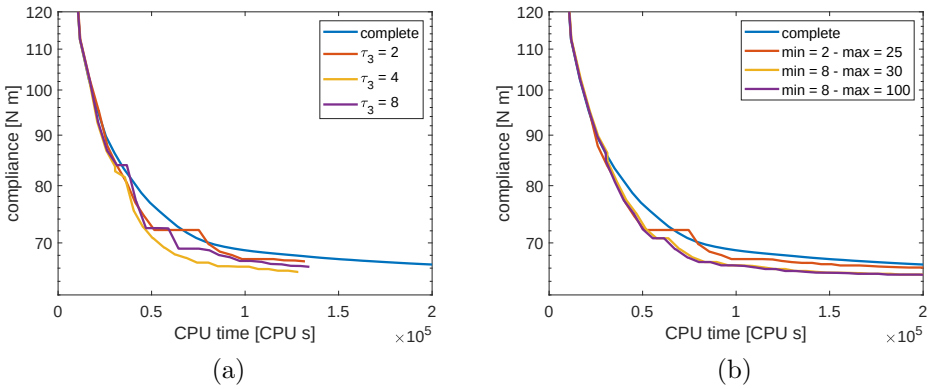

FIG. 4.1. *Best known solution c in function of the CPU time for $\rho^* = 0.5$, $r = 5$, $M_{min} = 2$ and $M_{max} = 20$ under different τ_3 values (a). Best known solution c in function of the CPU time for $\rho^* = 0.5$, $r = 5$ and $\tau_3 = 2$ under different M_{min} and M_{max} values (b).*

2%, between iteration $g = 100$ and iteration $g = 1000$.

4. Results. Tab. 4.1 shows the reference $\tilde{c}$ obtained with no reduced basis with $G = 100$ iterations for each setting of ρ^* and r, and the corresponding computational time. Moreover, it shows the CPU time to reach the reference $\tilde{c}$ plus 5% and 10% for each τ_3, and the percentage reduction with respect to the CPU time to reach the same compliance without reduced basis.

The results show that the proposed algorithm is almost always able to find similar values of compliance c with respect to the reference $\tilde{c}$ obtained with no reduced basis by the original method of Sigmund [10], for at least one value of τ_3. Moreover, a significant CPU time saving is always obtained. A reduction up to 52% and equal to at least 25% in most cases is observed when considering the reference $\tilde{c}$ plus 5%, and the reduction is even higher when considering $\tilde{c}$ plus 10%, being up to 68%. These higher reductions show that the biggest gain is obtained in the first iterations, where the gradient-based optimization gives higher c reduction per iteration. Thus, the algorithm proved to be effective in these iterations, and this confirms the effectiveness of the dynamic buffer management to start using reduced problems early.

It is worth remarking that the reduced solutions are rejected because the condition $\delta_c < \tau_3$ is not respected, while the conditions $e_1 < \tau_1$ and $e_2 < \tau_2$ are always respected even with the low imposed thresholds. Moreover, as observed in Tab. 4.1, the computational gain is not strictly correlated to the value of τ_3, due to the fact that the time gain is affected by competing factors. On the one hand, smaller τ_3 values prevent instability of the whole algorithm and the need of restarting from higher values of c; on the other hand, this greater stability is paid for with a higher number of complete FEA iterations, which are clearly more expensive in terms of CPU time.

Fig. 4.1(a) shows the evolution of the best known compliance c as a function of the CPU time for $\rho^* = 0.5$ and $r = 5$, under the different values of τ_3 and for the case with no reduced basis. We may observe a faster reduction of the objective function c over time when adopting our algorithm with FPCA reduced basis, which corresponds to the shorter CPU times reported in Tab. 4.1 to get $\tilde{c}$ plus 5% and 10%. Even when the objective function increases, due to a bad representation of the solution with few bases (and the best known c remains constant over a certain time period), our algorithm is still able to redirect to acceptable values of c, which are even better

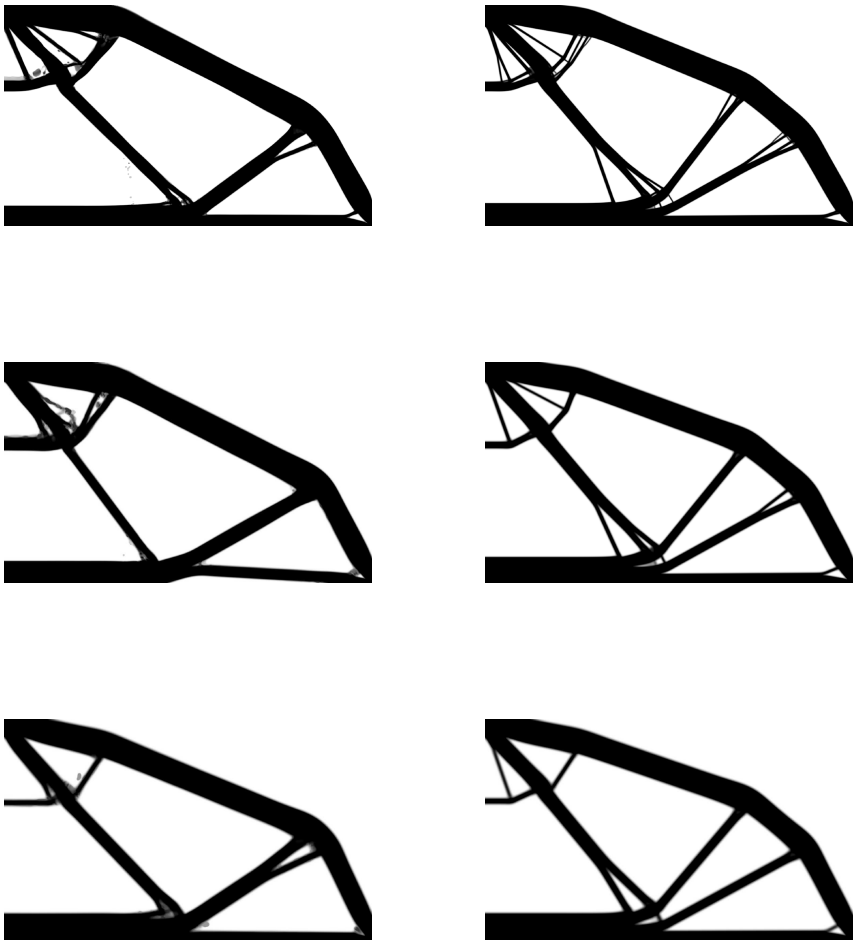

FIG. 4.2. *Maps for $\rho^* = 0.3$ from the proposed algorithm with $\tau_3 = 2$ when $\tilde{c}$ plus 5% is reached (left column) and from the case with no reduced basis at 100 iterations (right column); $r = 5$ (upper row), $r = 10$ (central row) and $r = 15$ (lower row).*

than the case with no reduced basis at the same CPU time.

Figs. 4.2-4.4 show the density maps obtained with the proposed algorithm when $\tilde{c}$ plus 5% is reached and the corresponding maps from the case with no reduced basis after $G = 100$ iterations. The maps are reported for each setting of ρ^* and r, and for $\tau_3 = 2$. When $\tilde{c}$ plus 5% is not reached within $G = 100$ iterations by our algorithm for $\tau_3 = 2$ (i.e., with $\{\rho^*, r\}$ equal to $\{0.3, 15\}$ and $\{0.5, 10\}$), the map obtained from the algorithm at iteration $g = 100$ is reported. The maps from our algorithm show to capture the larger features of the solution. Moreover, although topologically different for small details compared with the maps obtained with no reduced basis, our maps correspond to similar values of compliance c.

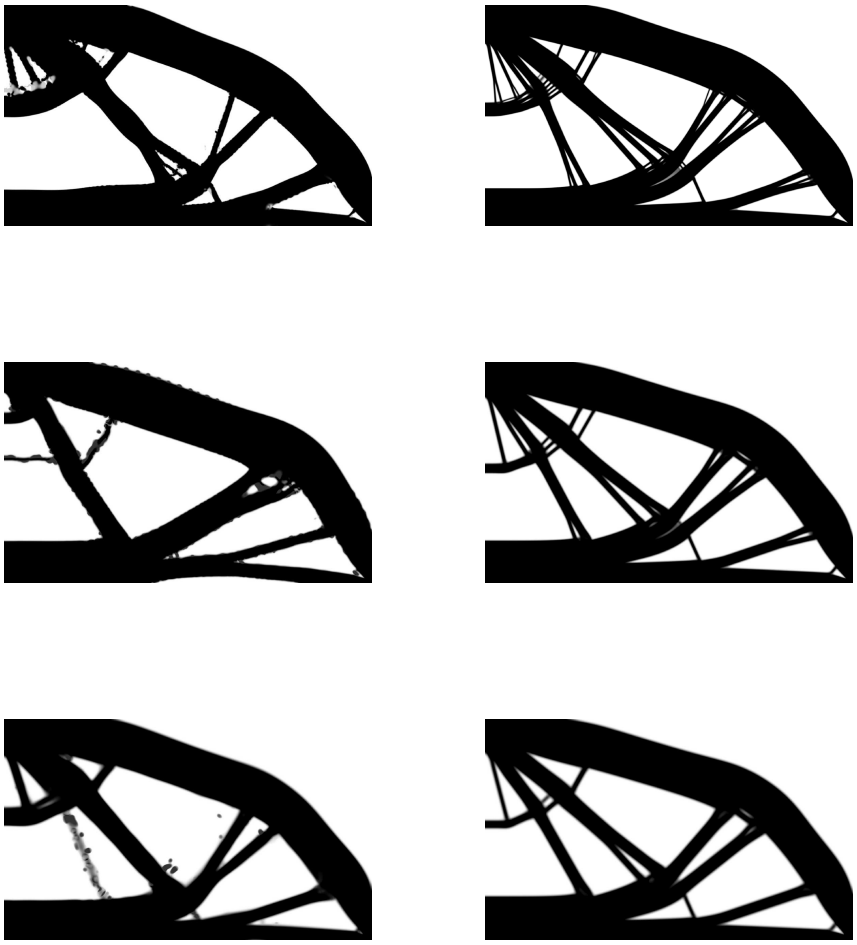

FIG. 4.3. *Maps for $\rho^* = 0.5$ from the proposed algorithm with $\tau_3 = 2$ when $\tilde{c}$ plus 5% is reached (left column) and from the case with no reduced basis at 100 iterations (right column); $r = 5$ (upper row), $r = 10$ (central row) and $r = 15$ (lower row).*

4.1. Impact of buffer dimension M. Additional experiments are run to analyze the impact of the M_{min} and M_{max}. Fig. 4.1(b) shows the best known compliance c as a function of the CPU time, under different M_{min} and M_{max} values, for the case with $\rho^* = 0.5$, $r = 5$ and $\tau_3 = 2$. As longer tracts with a constant best known c correspond to solutions with unstable c over the iterations, we may notice that the solution is far more stable when using a higher buffer size in terms of both M_{min} and M_{max}. In particular, a higher M_{min} value avoids the destabilization of the solution in the very first iterations of the algorithm (red line in the figure), but at the price of a higher computational cost during the initial phase when the buffer starts to collect the solutions to build up the first reduced basis.

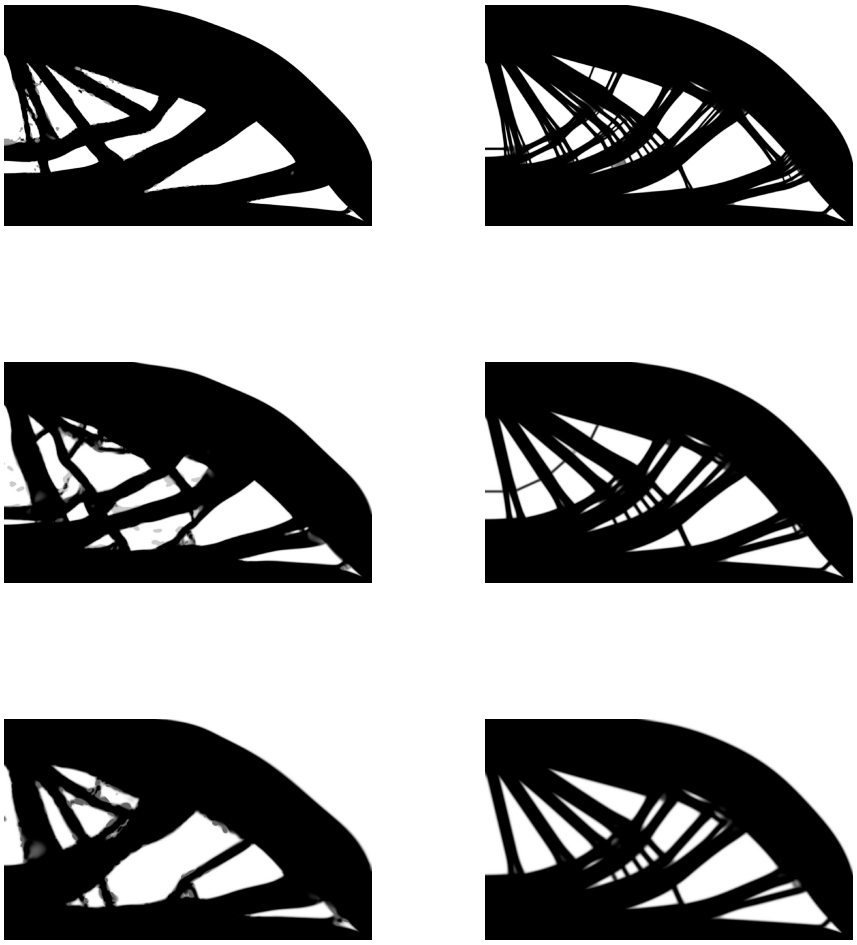

FIG. 4.4. *Maps for $\rho^* = 0.7$ from the proposed algorithm with $\tau_3 = 2$ when $\tilde{c}$ plus 5% is reached (left column) and from the case with no reduced basis at 100 iterations (right column); $r = 5$ (upper row), $r = 10$ (central row) and $r = 15$ (lower row).*

5. Discussions and conclusion. In this work, we couple a gradient-based optimization for STO with the FPCA method to generate reduced basis, onto which the equilibrium equations are projected at several optimization iterations. Unlike similar works available in the literature, we exploit FPCA rather than PCA or POD, following our previous work [1].

The proposed algorithm has been tested on a large problem with more than 4.8 million degrees of freedom, much larger than the problems considered to test similar approaches. Results confirm that the reduced FEA problem can be exploited in most of the optimization iterations, guaranteeing limited errors below τ_1, τ_2 and τ_3. This allows to significantly reduce the STO computational times.

In particular, our density maps with a gap of 5% or 10% with respect to the

reference $\tilde{c}$ are topologically close to the ones at convergence when the gradient-based STO is performed without reduced basis. But, exploiting reduced basis through FPCA, these solutions are obtained in a drastically lower CPU time. Moreover, in addition to being close, the maps are completely equivalent from a practical point of view. In fact, before being used for producing artefacts, the geometries obtained by STO are almost always subjected to "cleaning" post-processing operations, such as the removal of elements having a density outside a specific range. In this way, the maps of our gradient-based optimization with FPCA and the maps obtained without FPCA turn out to be practically equivalent for manufacturing purposes.

Moreover we remark that, while the uniqueness of the STO optimal solution has been proved for $p = 1$ [9], ll-posedness of the problem for higher p values (e.g., $p = 3$ used in this work) is well-known [3, 12]. Consequently, rather than simply being an approximation, the maps obtained using FPCA could represent optimal alternative solutions. In fact, by modifying the problem by means of $\mathbf{f}_{curr}$, we do not follow the same sequence of solutions followed without reduced basis, and this allows the procedure to move to other regions where different minima can be found. Obviously, we stopped the iterations when we reached a predefined gap, to pursue time savings, but the evolution could lead to a different solution even in the subsequent iterations.

Our strategy to include $\mathbf{f}_{curr}$ has avoided additional calculations to compute corrective terms to adapt relation 2.6, as carried out in [7]. However, in this way, the reduced problems generate solutions with an increase of c in a few iterations. But, at the same time, the compliance c is reduced so much in the next iteration with a complete FEA solution to provide significant gain in the overall computing time. Our future work will consist in defining the best thresholds τ_1, τ_2 and τ_3 in order to optimize this trade-off. Furthermore, we will compare the proposed strategy with f_{curr} with different alternatives in which corrective terms are added in (2.6).

From an application point of view, we will apply the proposed algorithm to several practical cases. For example, we will consider the computational gain when performing a pre-screening of different layouts, such as orientations or alternative features, while limiting a finer analysis without reduced basis to the selected layout only. At the same time, we can exploit the computational gain to explore different initializations of the optimization procedure in a reduced time.

References.

[1] Gianluca Alaimo, Ferdinando Auricchio, Ilaria Bianchini, and Ettore Lanzarone, *Applying functional principal components to structural topology optimization*, International Journal for Numerical Methods in Engineering, 115 (2018), pp. 189–208.

[2] Oded Amir, Martin P Bendsøe, and Ole Sigmund, *Approximate reanalysis in topology optimization*, International Journal for Numerical Methods in Engineering, 78 (2009), pp. 1474–1491.

[3] Martin P Bendsøe and Ole Sigmund, *Optimization of structural topology, shape, and material*, vol. 414, Springer, 1995.

[4] Ilaria Bianchini, Raffaele Argiento, Ferdinando Auricchio, and Ettore Lanzarone, *Efficient uncertainty quantification in stochastic finite element analysis based on functional principal components*, Computational Mechanics, 56 (2015), pp. 533–549.

[5] Jean-Luc Dulong, Frédéric Druesne, and Pierre Villon, *A model reduction approach for real-time part deformation with nonlinear mechanical be-*

havior, International Journal on Interactive Design and Manufacturing, 1 (2007), p. 229.

[6] Nicola Ferro, Stefano Micheletti, and Simona Perotto, *Pod-assisted strategies for structural topology optimization*, Computers & Mathematics with Applications, 77 (2019), pp. 2804–2820.

[7] Christian Gogu, *Improving the efficiency of large scale topology optimization through on-the-fly reduced order model construction*, International Journal for Numerical Methods in Engineering, 101 (2015), pp. 281–304.

[8] David J Munk, Gareth A Vio, and Grant P Steven, *Topology and shape optimization methods using evolutionary algorithms: a review*, Structural and Multidisciplinary Optimization, 52 (2015), pp. 613–631.

[9] Joakim Petersson, *Some convergence results in perimeter-controlled topology optimization*, Computer Methods in Applied Mechanics and Engineering, 171 (1999), pp. 123–140.

[10] Ole Sigmund, *A 99 line topology optimization code written in matlab*, Structural and multidisciplinary optimization, 21 (2001), pp. 120–127.

[11] Ole Sigmund and Kurt Maute, *Topology optimization approaches*, Structural and Multidisciplinary Optimization, 48 (2013), pp. 1031–1055.

[12] Andrés Tovar and Kapil Khandelwal, *Uniqueness in linear and nonlinear topology optimization and approximate solutions*, in Proceedings of the 2nd International Conference on Engineering Optimization (EngOpt 2010), Lisbon, Portugal, 2010.

[13] Nico P van Dijk, Kurt Maute, Matthijs Langelaar, and Fred Van Keulen, *Level-set methods for structural topology optimization: a review*, Structural and Multidisciplinary Optimization, 48 (2013), pp. 437–472.

[14] Shun Wang, Eric de Sturler, and Glaucio H Paulino, *Large-scale topology optimization using preconditioned krylov subspace methods with recycling*, International journal for numerical methods in engineering, 69 (2007), pp. 2441–2468.

[15] Manyu Xiao, Dongcheng Lu, Piotr Breitkopf, Balaji Raghavan, Subhrajit Dutta, and Weihong Zhang, *On-the-fly model reduction for large-scale structural topology optimization using principal components analysis*, Structural and Multidisciplinary Optimization, (2020), pp. 1–22.

[16] Gil Ho Yoon, *Structural topology optimization for frequency response problem using model reduction schemes*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 1744–1763.

Recent titles from the IMATI-REPORT Series:

2019

19-01: *Feature curves, Hough transform, characteristic elements*, C. Romanengo, S. Biasotti, B. Falcidieno.

19-02: *On expansions and nodes for sparse grid collocation of lognormal elliptic PDEs,* O.G. Ernst, B. Sprungk, L. Tamellini.

19-03: *Augmented Reality in manufacturing engineering*, B. Bonino, F. Giannini, M. Monti.

19-04: *Parametric shape optimization for combined additive-subtractive manufacturing*, C. Altenhofen, M. Attene, O. Barrowclough, M. Chiumenti, M. Livesu, F. Marini, M. Martinelli, V. Skytt, L. Tamellini.

19-05: *Architecture of a client-server platform for IMATI cultural heritage applications*, A. Repetto, C.E. Catalano, M. Spagnuolo

2020

20-01: *Compressive isogeometric analysis*, A. Brugiapaglia, L. Tamellini, M. Tani.

20-02: *A gradient-based optimization method with functional principal component analysis for efficient structural topology optimization*, A. Montanino, G. Alaimo, E. Lanzarone

Istituto di Matematica Applicata e Tecnologie Informatiche ``*Enrico Magenes*", CNR
Via Ferrata 5/a, 27100, Pavia, Italy
Genova Section: Via dei Marini, 6, 16149 Genova, Italy • **Milano Section:** Via E. Bassini, 15, 20133 Milano, Italy
http://www.imati.cnr.it/