INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2017; **111**:88–100
Published online 7 February 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nme.5455

SHORT COMMUNICATION

Bounds for decoupled design and analysis discretizations in topology optimization

D. K. Gupta, G. J. van der Veen, A. M. Aragón, M. Langelaar*,† and F. van Keulen

Structural Optimization and Mechanics, Faculty of 3mE, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands

SUMMARY

Topology optimization formulations using multiple design variables per finite element have been proposed to improve the design resolution. This paper discusses the relation between the number of design variables per element and the order of the elements used for analysis. We derive that beyond a maximum number of design variables, certain sets of material distributions cannot be discriminated based on the corresponding analysis results. This makes the design description inefficient and the solution of the optimization problem non-unique. To prevent this, we establish bounds for the maximum number of design variables that can be used to describe the material distribution for any given finite element scheme without introducing non-uniqueness. Copyright © 2016 The Authors. *International Journal for Numerical Methods in Engineering* Published by John Wiley & Sons Ltd

Received 6 July 2016; Revised 29 September 2016; Accepted 17 October 2016

KEY WORDS: topology design; finite element methods; structures

1. INTRODUCTION

Topology optimization (TO) is a numerical procedure that allows obtaining the optimal distribution of a given material within a domain, subject to certain constraints, such that the performance of the resulting structure is maximized. Since its introduction three decades ago [1], this approach has been used to solve a wide variety of problems in various academic and industrial fields. For the recent advancements in TO, please see [2–4]. The so-called density-based formulation is the most popular both in academia and in commercial applications. Conventionally, a single density design variable is associated with each finite element, typically of low order, in the analysis mesh.

Computational costs limit the application of TO in complex, high-resolution 3D problems. The largest part of these costs is associated with finite element analysis (FEA) and the associated sensitivity analysis. By increasing the resolution of the design field relative to the analysis mesh, attempts have been made to reduce this cost [5–9]. In these works, the design field is characterized by the distribution of design variables in the domain. Based on these design variables, the material density distribution is obtained for the whole domain using some interpolation scheme. While the design variables are used for the optimization, the interpolated density values are used for analysis purposes. For example, consider the 'Q4/n4/d9' element shown in Figure 1, that corresponds to a 4-node quadrilateral element containing 9 design points within the element. A design variable is

*Correspondence to: Matthijs Langelaar, Structural Optimization and Mechanics, Faculty of 3mE, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands.

†E-mail: M.Langelaar@tudelft.nl

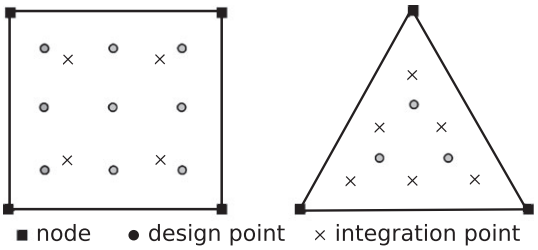


Figure 1. Q4 element (left) with 9 design points and 4 integration points (Q4/n4/d9) and a T3 element (right) with 3 design points and 6 integration points (T3/n6/d3).

associated to each design point. The values of these design variables are used to obtain the density values at 4 integration points where the stiffness matrix contributions are evaluated. Another example, a 'T3/n6/d3' element (Figure 1), refers to a 3-node triangle with 3 design points and 9 integration points. These elements are based on the conventions denoted in [5, 6] where similar examples such as Q4/n25/d25 and Q4/n25/d9 are presented.

Increasing the number of design points within an element surely increases the resolution of the material distribution without introducing additional degrees of freedom (DOFs) in the analysis. However, because analysis is an integral part of the TO process, one should make sure that the difference between the different high resolution designs obtained using such schemes can be observed in the FEA results. In other words, the question is whether the additional design freedom can be properly exploited in a particular FEA setting. If not, different designs may show similar performance, resulting in non-unique optima and unexpected convergence behavior. Numerical results related to a similar aspect have been reported in [10], where it is shown that varying the density resolution beyond a certain threshold can lead to unacceptable solutions. Moreover, introducing more design freedom than can be observed from the model only adds to the computational burden.

Note that for the examples mentioned earlier, design variables are characterized as points in the element domain. However, these may not necessarily be spatial points within an element, e.g. for a material density field expressed using polynomials, design variables can refer to the coefficients of such polynomials (as in Section 2). The study in this paper is not tailored towards a certain choice of design variables. Rather, the focus is on the number of design variables used for optimization.

In this paper, we study the relation between the number of design variables and the resolution of the analysis mesh used in TO. In this paper, linear elastostatic problems are considered, and the following aspects are covered:

- We first illustrate using examples that the number of useful design variables is bounded by a certain maximum value (Section 2).
- We show mathematically that the permissible number of design variables inside a finite element is bounded by the rank of the element matrix obtained from the discretization of the design domain (Section 3). The argument is further extended to the system level as well.
- We present a discussion on the choice of the number of design variables for some popular finite elements (Section 4). A brief discussion is provided in Section 5 on the issues related to examples that violate the conditions proposed in the present paper.

2. EXAMPLES

2.1. An analytical example

Consider a structural optimization problem involving a simple rod of length l and area of cross-section A. This rod is fixed at one end and loaded in axial tension by a force F acting at the other end. It is assumed that the Young's modulus E can be varied at any position along the rod. In line with classical density-based topology optimization, we define $E(x) = E_0\rho(x)$, where E_0 is a constant and $\rho(x)$ represents the material density distribution in the domain. Suppose the bar is modeled using a single finite element, and linear shape functions are used to interpolate the displacement field within the element. When using analytical integration, the stiffness matrix $\mathbf{K}$ can be expressed as

DOI: 10.1002/nme

$$\mathbf{K} = \frac{A}{l^2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\int_0^l E(x)dx = \frac{E_0 A}{l^2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\underbrace{\int_0^l \rho(x)dx}_{\bar{\rho}}. \tag{1}$$

In Equation (1), **K** depends on the material density distribution only in an average sense. For all choices of $\rho(x)$ that yield the same average density $\bar{\rho}$, the same stiffness matrix is obtained. As a result, the response of the structure for all those choices will be identical. In other words, the modeling of the physics shows independence to certain specific changes in $\rho(x)$. Next, we may choose a certain discretization to define $\rho(x)$ in terms of a finite number of design variables. For the simplest representation by a constant function $\rho(x) = \hat{\rho}$, a unique relation between K and $\hat{\rho}$ exists. However, for all higher-order representations, infinite combinations of design variable values exist that result in the same average density $\bar{\rho}$. This leads to loss of uniqueness of the optimal solution. A similar situation can arise in more complex, multidimensional topology optimization problems.

For a better understanding of the issue stated earlier, we analyze the system of linear equations obtained from the finite element discretization. Using Gaussian quadrature points x_i and weights w_i for integrating **K** in an isoparametric setting, the discretized system of equations for the bar can be expressed as

$$\mathbf{Ku} = \mathbf{f} \implies \frac{E_0 A}{l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\sum_{i=1}^{g} w_i \rho(x_i)\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}. \tag{2}$$

Here, **u** and **f** correspond to nodal DOFs and nodal loads, respectively. Suppose that **u** corresponding to a certain optimal design is known, then we would like to establish whether that design is unique.

Let us assume that we integrate the **K** matrix using a 2-point Gaussian quadrature scheme. With the given boundary conditions ($u_1 = 0$), the applied force ($f_2 = F$), and the Gaussian weights ($w_1 = w_2 = 0.5$), there exists only one linearly independent equation:

$$\frac{E_0 A}{2l}[\rho(x_1) + \rho(x_2)]u_2 = F, \tag{3}$$

where the two integration points x_1 and x_2 are $\left(1 \pm \frac{1}{\sqrt{3}}\right)\frac{l}{2}$. Now, assuming that F and u_2 are known, $\rho(x)$ is described by polynomials of different order (zeroth order, first order, *etc.*), where the coefficients of the polynomial terms are the design variables.

CASE I: $\rho(x) = a$, i.e. a constant density field. From Equation (3), we obtain

$$\frac{E_0 A}{2l}(a + a)\,u_2 = F \implies a = \frac{Fl}{E_0 A u_2}\ . \tag{4}$$

CASE II: $\rho(x) = ax + b$, i.e. a linearly varying density field. For this case, Equation (3) can be written as

$$\frac{E_0 A}{2l}[(ax_1 + b) + (ax_2 + b)]u_2 = F \implies (x_1 + x_2)a + 2b = \frac{2Fl}{E_0 A u_2}\ . \tag{5}$$

The order of the shape functions for both cases leads to a single equation, implying that not more than one design variable can be uniquely determined if u_2 and F are known. From the example presented earlier, we see that for *CASE I*, only one design variable is used, and the solution is unique. However, for *CASE II*, it can be observed from Equation (5) that there can be infinitely many combinations of a and b which can produce the same displacement at the end of the bar. Thus, for *CASE II*, the assumed $\rho(x)$ function cannot be solved uniquely. This still holds when restrictions are imposed on the design ($0 < \rho(x) \leqslant 1$). Similarly, it can be shown that for an even higher order representation of $\rho(x)$, the solution will always be non-unique. The reason is that the finite element discretization is unable to reflect the design changes.

DOI: 10.1002/nme

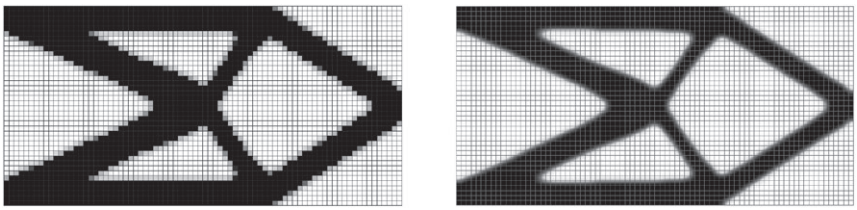
This simple example illustrates that there is an upper bound on the number of design variables that can be used for a given body discretized using finite elements above which the design is no longer unique. More design variables can be included; however, then the effect of certain changes in the design cannot be distinguished. Here, a single finite element with linear shape functions was used, and the strain-displacement matrix was constant. In Section 3, we generalize this concept of choosing the appropriate number of design variables to the case of shape functions of arbitrary order and systems of many elements.

2.2. Numerical example

Next, we focus on the compliance minimization of a Michell TO problem with aspect ratio 2:1. For details on the mathematical formulation of the optimization problem, refer to [11]. The design domain is discretized using 80×40 finite elements, and a maximum of 45% of the domain is allowed to be filled.

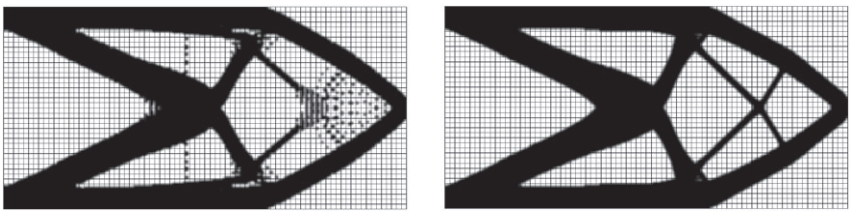
Figure 2(a) and (b) show two optimized designs obtained using 1 design point and 16 design points per element, respectively. A filter radius of $1.5h$, where h is the mesh size, is used for both cases. The approach used here is similar to that proposed in [5], and interested readers are referred to this work for details related to sensitivity analysis, projection scheme, *etc*. For both designs, bilinear shape functions are employed to interpolate the displacement field inside every element. It is observed that using a larger number of design points helps to achieve a higher resolution material description with less jagged boundaries. However, even with an increased design resolution, there is no improvement in the performance of the structure. This is due to the fact that the chosen finite element model (FEM) cannot accurately model such a material distribution. Also, because of the large number of design variables, the computational costs associated with sensitivity analysis, as well as the design update, increase significantly. Thus, given the fact that no improvement in performance is observed under the restriction imposed by the FEM, any attempt to use too many design points leads to an unnecessary increase in the computational cost of the problem without any gain in structural performance.

Figure 3 shows two cases with similar design resolution, but with different shape function orders. A total of 16 design points (4×4) is used per element for both cases. Here, the filter radius is chosen to be $0.4h$, and the objective is to represent fine structural features at relatively lower FEA costs.



(a) $d = 1, \mathcal{C} = 74.099$ (b) $d = 16, \mathcal{C} = 74.219$

Figure 2. Optimized designs for a Michell problem obtained on a mesh of 80×40 bilinear finite elements using a filter radius of $1.5h$.



(a) $p = 1, \mathcal{C} = 70.912$ (b) $p = 3, \mathcal{C} = 71.346$

Figure 3. Optimized designs for a Michell problem obtained on a mesh of 80×40 using 16 design points per element and a filter radius of $0.4h$.

DOI: 10.1002/nme

Figure 3(a) shows the optimized design obtained using bilinear shape functions, and based on the achieved objective value $\mathcal{C} = 70.912$, the structure appears to have reached a high quality optimum. However, small disconnected features can be clearly seen, based on which it can be questioned whether the reported performance is correct. Indeed when this design is evaluated with bicubic shape functions, the obtained objective value is $\mathcal{C} = 81.664$. The error in the objective value can be attributed to the incapability of the chosen FEM to accurately represent the material distribution. Figure 3(b) shows the optimized design obtained using bicubic shape functions and the performance of this structure compared with that in Figure 3(a) is approximately 12% better. Clearly, because of an increase in the order of the shape functions, the total number of DOFs is increased, thereby improving the capability of the FEM to handle this design resolution. The approach of increasing the polynomial interpolation order to represent more complex material descriptions is also found in the finite cell method. For more numerical examples on finite cell method-based TO that illustrate the interplay between design and analysis resolutions, see [10].

From the numerical cases studied earlier, it can be understood that the proper choice of the number of design points depends on the chosen FEM scheme. In the next section, we look into this aspect from a mathematical point of view and derive a relation between design resolution and FEM scheme.

3. CONDITION FOR UNIQUENESS OF THE DENSITY DISTRIBUTION

Optimization of a structural design in general can refer to finding the optimal distribution of material in a given domain as well as optimizing the material properties themselves at every point $\mathbf{x}$ in the domain. In the linear elasticity setting considered here, stresses and strains are related by generalized Hooke's law, $\boldsymbol{\sigma} = \mathbf{D}\boldsymbol{\epsilon}$. Here, $\mathbf{D}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ denote the constitutive matrix and stress and strain column vectors, respectively. The dependence of all these parameters on the position $\mathbf{x}$ is implied.

For a design representation in two (three) dimensions, the most general choices of design variables are six (21) entries of the $\mathbf{D}$ matrix and their spatial variation throughout the domain. However, our discretized system allows only the representation of a limited number of deformation modes, the combinations of which can represent all possible deformation states of the system. When choosing the number of design variables, a unique design representation should be obtained from these deformation modes. Depending on the type of optimization problem, a certain choice of design variables is made. For instance, in free material optimization of structures for stiffness maximization, the problem is reduced to solving only one design variable per material sampling point, namely the norm or the trace of the elasticity tensor [12]. It is shown analytically in [13] that the elements of the elasticity tensor can then be fully recovered from the optimal norm and the related displacements.

Traditionally, in TO one design variable per material sampling point is also assumed. However, the approach is different from the free material optimization discussed earlier. In TO, certain restrictions are typically imposed on the material properties throughout the domain, and $\mathbf{D}$ is expressed as $\mathbf{D} = \rho(\mathbf{x})\mathbf{D}_0$, where $\mathbf{D}_0$ is assumed to be constant for the whole material domain and ρ denotes the material density. Thus, at any point in the domain, only the material volume fraction (material density) is allowed to vary. The focus of this paper is on TO, and thus we tailor the foregoing discussion accordingly. Our objective is to mathematically relate the number of design variables to the number and type of elements used in the analysis. Further, the aim is to establish a threshold condition on the appropriate number of design variables that can be associated with any finite element as well as a system of elements in TO.

Assuming that the design variables represent the coefficients of a polynomial function, the density field inside an element can be defined as

$$\rho(\mathbf{x}) = s_1 + s_2 x + s_3 y + \ldots m \text{ terms} = \tilde{\mathbf{x}}^{\mathsf{T}}\mathbf{s}, \tag{6}$$

where, $\mathbf{s} = [s_1, s_2, \ldots, s_m]^{\mathsf{T}}$ represents the set of design variables and m the number of design variables used in the optimization problem. To determine the upper bound on m, here a linear elastostatic problem is considered. For any finite element, equilibrium can then be expressed as

DOI: 10.1002/nme

$$\mathbf{K}^{(e)}\mathbf{u}^{(e)} = \mathbf{f}^{(e)}, \tag{7}$$

where,

$$\mathbf{K}^{(e)} = \int_{\Omega_e} \mathbf{B}^\intercal \rho(\mathbf{x}) \mathbf{D}_0 \mathbf{B} \, d\Omega = \left(\int_{\Omega_e} \mathbf{B}^\intercal \mathbf{D}_0 \mathbf{B} \tilde{\mathbf{x}}^\intercal \, d\Omega \right) \mathbf{s}. \tag{8}$$

Here, $\mathbf{u}^{(e)}$ and $\mathbf{f}^{(e)}$ represent the kinematic nodal degrees of freedom and corresponding energetically conjugate nodal loads, respectively. The polynomial order of $\mathbf{K}^{(e)}$ depends on the choice of the density field. For a certain choice of $\rho(\mathbf{x})$ in Equation (6), the polynomial order of the corresponding $\mathbf{K}^{(e)}$ can be determined from Equation (8). Accordingly, an appropriate number of integration points can be chosen based on standard Gaussian quadrature.

Assume that $\mathbf{K}^{(e)}$ is now numerically integrated using g integration points, then $\mathbf{K}^{(e)} = \sum_{i=1}^{g} \rho(\mathbf{x}_i)\mathbf{K}_i = \sum_{i=1}^{g} \rho(\mathbf{x}_i)\mathbf{B}_i^\intercal \mathbf{D}_0 \mathbf{B}_i w_i$. It can further be expressed as

$$\mathbf{K}^{(e)} = \underset{(n \times cg)}{\begin{bmatrix} \mathbf{B}_1^\intercal & \mathbf{B}_2^\intercal & \dots & \mathbf{B}_g^\intercal \end{bmatrix}} \underbrace{\begin{bmatrix} \rho(\mathbf{x}_1) w_1 \mathbf{D}_0 & & & \\ & \rho(\mathbf{x}_2) w_2 \mathbf{D}_0 & & \\ & & \ddots & \\ & & & \rho(\mathbf{x}_g) w_g \mathbf{D}_0 \end{bmatrix}}_{\substack{\Theta_1 \\ (cg \times cg)}} \underset{(cg \times n)}{\begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_g \end{bmatrix}}. \tag{9}$$

Here, n represents the number of DOFs per element, and $\mathbf{K}^{(e)}$ is thus an $n \times n$ matrix. Because of the presence of rigid body modes, $\mathbf{K}^{(e)}$ is not a full-rank matrix. Let the rank of $\mathbf{K}^{(e)}$ be denoted by r. For a linear-elastic material, the stress components can be expressed as linear combinations of the strain components through the constitutive matrix. Because this constitutive relation is invertible and the strain energy density should be positive, $\mathbf{D}_0$ in Equation (9) is a full-rank and positive-definite matrix. The dimensions of $\mathbf{D}_0$ are assumed to be $c \times c$, where c is the number of independent stress components in the stress tensor. Because the rank of $\mathbf{D}_0$ is c, Θ_1 in Equation (9) is a positive-definite block-diagonal matrix with rank cg. Given that the rank of $\mathbf{K}^{(e)}$ is r, we obtain (proof in Appendix A)

$$\text{rank}\left(\begin{bmatrix} \mathbf{B}_1^\intercal & \mathbf{B}_2^\intercal & \dots & \mathbf{B}_g^\intercal \end{bmatrix}\right) = \text{rank}\left(\mathbf{K}^{(e)}\right) = r. \tag{10}$$

Based on numerical integration and rearrangement of terms, the system of linear equations in Equation (7) can be expressed in terms of $\mathbf{s}$ as

$$\underbrace{\left[\left(\sum_{i=1}^{g} \mathbf{K}_i\right)\mathbf{u}^{(e)} \;\; \left(\sum_{i=1}^{g} x_i \mathbf{K}_i\right)\mathbf{u}^{(e)} \;\; \left(\sum_{i=1}^{g} y_i \mathbf{K}_i\right)\mathbf{u}^{(e)} \;\; \dots \; m^{\text{th}}\text{term} \right]}_{\mathbf{A}^{(e)}} \mathbf{s} = \mathbf{f}^{(e)}. \tag{11}$$

For a choice of $\mathbf{u}^{(e)}$ and $\mathbf{f}^{(e)}$ such that $\mathbf{u}^{(e)}$ does not correspond to a rigid body mode, our goal is to determine the maximum possible value of m such that $\mathbf{s}$ has a unique solution. From Equation (11), it can be observed that m cannot be greater than the rank of $\mathbf{A}^{(e)}$, which we will determine next:

$$\mathbf{A}^{(e)} = \mathbf{P}\mathbf{Q}, \tag{12}$$

DOI: 10.1002/nme

where,

$$\mathbf{P} = \underset{(n\times cg)}{\begin{bmatrix}\mathbf{B}_1^{\mathsf{T}} & \mathbf{B}_2^{\mathsf{T}} & \dots & \mathbf{B}_g^{\mathsf{T}}\end{bmatrix}} \underbrace{\begin{bmatrix} w_1\mathbf{D}_0\mathbf{B}_1\mathbf{u}^{(e)} & & & \\ & w_2\mathbf{D}_0\mathbf{B}_2\mathbf{u}^{(e)} & & \\ & & \ddots & \\ & & & w_g\mathbf{D}_0\mathbf{B}_g\mathbf{u}^{(e)} \end{bmatrix}}_{\substack{\Theta_2 \\ (cg\times g)}}, \tag{13}$$

$$\mathbf{Q} = \begin{bmatrix} 1 & x_1 & \dots & (m^{\text{th}})_1 \\ 1 & x_2 & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_g & \dots & (m^{\text{th}})_g \end{bmatrix}. \tag{14}$$

To determine the rank of $\mathbf{A}^{(e)}$, the ranks of $\mathbf{P}$ and $\mathbf{Q}$ should be known.

Because Θ_2 is a block-diagonal matrix, $\text{rank}(\Theta_2) = g$. Using the fact that the rank of the product of two matrices cannot exceed the rank of the either factor (proof in [14]), the upper bound on the rank of $\mathbf{P}$ is given by the minimum of the number of deformation modes r and the number of integration points g, i.e.:

$$\text{rank}(\mathbf{P}) \leqslant \min(r, g). \tag{15}$$

To gain further insight into the rank of $\mathbf{P}$, it can be written as

$$\mathbf{P} = \begin{bmatrix}\mathbf{K}_1\mathbf{u}^{(e)} & \mathbf{K}_2\mathbf{u}^{(e)} & \dots & \mathbf{K}_g\mathbf{u}^{(e)}\end{bmatrix} \tag{16}$$

For $g \geqslant r$, we now look into the different possibilities where $\text{rank}(\mathbf{P}) < r$. Such instances need to be identified because for any such scenario, the maximum number of design variables that can be used without encountering non-uniqueness would be less than r.

CASE I: One of the possibilities is if $\mathbf{u}^{(e)}$ is such that no deformation is observed at one of the integration points, i.e. the respective product $\mathbf{B}_i\mathbf{u}^{(e)}$ is zero. Then, the respective column of the $\mathbf{P}$ matrix will become zero, decreasing its rank by 1. This scenario can even occur for more than one integration point at the same time. Suppose for certain $\mathbf{u}^{(e)}$, two columns of $\mathbf{P}$ matrix are equal to zero. This would imply that $\text{rank}(\mathbf{P})$ has to be strictly less than or equal to $g-2$. This can lead to a situation where $\text{rank}(\mathbf{P})$ becomes less than r. Such a scenario is restricted to certain choices of $\mathbf{u}^{(e)}$, and this problem is alleviated if an appropriate integration scheme is used. The number of integration points needed to accurately integrate $\mathbf{K}^{(e)}$ is generally higher compared with the number of design variables that can be used in a finite element. Thus, even if no deformation is observed at one of the integration points, it can still be observed on at least r other integration points ensuring that $\text{rank}(\mathbf{P}) = r$.

CASE II: Another possibility where $\text{rank}(\mathbf{P}) < r$ is when one or more columns of $\mathbf{P}$ can be expressed as linear combinations of other columns, and the total number of linearly independent columns becomes less than r. This would mean that there exist an $\boldsymbol{\alpha} \neq \mathbf{0}$ such that

$$\sum_{i=1}^{g} \alpha_i \mathbf{K}_i \mathbf{u}^{(e)} = \mathbf{0}. \tag{17}$$

Here, the $\mathbf{K}_i$ matrices differ only by their respective $\mathbf{B}$ matrices. In general, every element of $\mathbf{B}$ is a polynomial function expressed in terms of the spatial coordinates; thereby, it depends on the location of the point within the element at which $\mathbf{K}_i$ is evaluated. Thus, in general, no $\mathbf{K}_i$ matrix

DOI: 10.1002/nme

can be expressed as a linear combination of other such matrices which are evaluated at other sampling points in the element domain. This implies that Equation (17) can only be satisfied if $\mathbf{u}^{(e)}$ belongs to the nullspace of $\sum_{i=1}^{g} \alpha_i \mathbf{K}_i$. This can happen only for some specific cases, e.g. when it is a rigid body mode. However, for the vast majority of cases, it follows that the columns of $\mathbf{P}$ are linearly independent.

A case where linear dependence occurs for any deformation is encountered when linear functions are used to interpolate the state field within the element. Then the $\mathbf{B}$ matrix is constant and does not depend on the spatial variables. Because of this, all columns of $\mathbf{P}$ are equal, reducing its rank to 1. Some examples where this happens are 3-node triangles (T3) and 4-node tetrahedra (TET4). Thus, for the T3 and TET4 elements, not more than one unique design variable can be used per element. However, for cases where $\mathbf{D_0}$ varies throughout the element (*e.g.*, material which is functionally graded in Poisson ratio), the respective $\mathbf{K}_i$ matrices would differ. This will increase the number of linearly independent columns in $\mathbf{P}$.

In a generalized sense then, $\text{rank}(\mathbf{P}) = \min(r, g)$. Based on Equation (12) and the fact that $\text{rank}(\mathbf{A^{(e)}}) \leqslant \min(\text{rank}(P), \text{rank}(Q))$, we obtain that $m \not> r$. Thus, m in Equation (14) is substituted by r. Next, the structure of the $\mathbf{Q}$ matrix is studied and based on the discussion provided in Appendix B,

$$\text{rank}(\mathbf{A}^{(e)}) = r \ \ \forall \ \ g \geqslant r. \tag{18}$$

In general, if r design variables are used, then the number of integration points required in a Gaussian quadrature scheme to accurately integrate $\mathbf{K}^{(e)}$ is greater than or equal to r. The condition $g \geqslant r$ does not need to be explicitly imposed and can be ignored. This implies that for a finite element with r deformation modes, a maximum of r design variables should be used to represent the density field inside the element and going beyond this limit will lead to non-uniqueness.

Note that unlike the nodal state variables which are defined at the nodal points, design variables here characterize the density distribution within the element and do not possess any direct relationship with the immediate neighboring elements. Thus, even for a group of elements, the number of design variables allowed inside every element of the system cannot be greater than the rank of that element's coefficient matrix. Based on this, the following element-level upper bound can be stated:

'*For any finite element, the number of design variables used to define the element's internal material distribution that can be uniquely optimized by means of finite element analysis, cannot be greater than the rank of the element coefficient matrix*.'

The condition derived here needs to be satisfied for every element. In addition to this, a similar is proposed at the system level as well. A system of elements can itself be considered as a superelement with a larger number of DOFs. For such an element, a system level condition can be derived using a similar approach. At the system level, the following rule can be formulated:

'*For any structure discretized using finite elements, the total number of design variables used to define the material distribution in the whole domain that can be optimized to a unique solution based on the results of finite element analysis, cannot be greater than the rank of the global coefficient matrix obtained from the discretization of the structure*.'

Important to note here is that the DOFs at the vertices as well as on the faces (edges for 2D) of the element are shared with those of the adjacent elements. Because of this, the system level bound on average will be stricter compared with the element level bound. This aspect is further discussed in the next section using a simple example.

4. CHOICE OF THE NUMBER OF DESIGN VARIABLES

The allowed number of design points is element-dependent (Section 3). Table I shows the upper bound on the number of design points for various finite elements. In general, the upper bound m can be stated as

$$m = d_{el} - r_{el}, \tag{19}$$

DOI: 10.1002/nme

Table I. Maximum number of design points for some commonly used finite elements.

Element	DOFs per node	No. of design points
3-node triangle (T3)	2	1
6-node triangle (T6)	2	9
4-node quadrilateral (Q4)	2	5
8-node quadrilateral (Q8)	2	13
9-node quadrilateral (Q9)	2	15
4-node tetrahedron (TET4)	3	1
10-node tetrahedron (TET10)	3	24
8-node hexahedron (H8)	3	18
20-node hexahedron (H20)	3	54

DOFs, degrees of freedom.

where d_{el} and r_{el} are the total DOFs and the number of rigid body modes for the element, respectively.[‡] For example, for a Q4 element in 2D with 2 DOFs per node, $m = 2 \cdot 4 - 3 = 5$. However, as pointed out in Table I, for elements such as T3 and TET4, only one design variable should be used per element (as in Section 3). This is because these elements use linear shape functions which result into constant strain-displacement matrices for isotropic materials.

Other than the element-level bound, the number of design density points should be lower than the bound at system level as well. For example, let us assume a rectangular grid of $N \times N$ Q4 elements. In an average sense, the number of density points per element ($\tilde{m}$) for this system needs to satisfy the following criterion:

$$\tilde{m} \leqslant \frac{(N+1)^2 d - r_{sys}}{N^2}, \tag{20}$$

where d is the DOFs per node and r_{sys} is the number of zero energy modes for the system. For example, for a rectangular mesh of 3×3 Q4 elements with two DOFs per node, $\tilde{m}$ is approximately equal to 3, which means $\tilde{m}$ is less than m. For large values of N, however, the maximum value of $\tilde{m}$ is approximately equal to d. Thus, in an average sense, there should not be more design points per element than the DOFs per node. We observe that the system level bound here imposes a stricter limit over the possible design resolution than the element level criterion. Note that Equation (20) is valid only for 2D rectangular meshes with certain choice of elements. For other mesh geometries and element types, this bound needs to be reformulated accordingly.

Besides the number of DOFs, the choice of the number of design variables also depends on the number of integration points. In other words, while choosing a high order material distribution, one should make sure that the stiffness matrix $\mathbf{K}^{(e)}$ can be accurately integrated. For example, for a Q4 element, conventionally, a 2×2 Gauss quadrature rule is used to integrate $\mathbf{K}^{(e)}$. This scheme can integrate exactly up to a cubic polynomial in two variables. For an elementwise constant material density, the polynomial order of the integrand in Equation (8) is 2. Thus, the 2×2 scheme is sufficient to exactly integrate it. From Table I, it can be seen that for a Q4 element, up to five design variables can be used. With 5 design variables, a second order polynomial-based material distribution can be represented in two dimensions. With such a distribution, the order of the integrand in Equation (8) will be four and for such a case, a minimum of 3×3 quadrature points is needed. Similarly, it can be shown for a Q9 element, at least a 4×4 integration scheme is needed. A similar approach can be followed to determine the appropriate integration schemes for other elements as well. When the number of integration points g is chosen too low, not only is the solution likely to be inaccurate, but also the number of unique design variables can be affected (as determined by Equation (15)).

[‡]A similar result was obtained by Jog and Haber in a study on checkerboard instability in topology optimization [15].

DOI: 10.1002/nme

5. DISCUSSION AND CONCLUSION

The main drive behind using a large number of density points is to obtain a better representation of the material distribution while keeping the computational cost as low as possible. Let us take the case of the Q4/n25/d25 element [5]. Any deformed state obtained using 25 design variables for this element can also be obtained using just 5 design variables. Similarly, for a T6 element, 9 design variables are sufficient to describe all deformation states (Table I). An interesting inference is that for elements with constant strain-displacement matrix (*e.g.,* T3, TET4), elementwise constant density needs to be assumed. For more design variables, a high-dimensional nullspace is introduced which may hinder the convergence of topology optimization. Also, the desire to add more design density points will lead to building the stiffness matrix contributions at significantly more number of points (Table I), which will eventually lead to an unnecessary increase in the computational costs. Significant burden on sensitivity analysis will also be noticed for such cases.

The discussions throughout the paper have been tailored towards structural topology optimization problems, and in particular linear elastostatic problems. However, in a similar fashion, bounds on the number of design variables can be derived for other structural problems as well, e.g. harmonic excitation, eigenvalue problems, *etc*. As an example, for an undamped harmonic excitation problem, the system of linear equations is

$$\{\mathbf{K}(\boldsymbol{\rho}) - \omega_f^2 \mathbf{M}(\boldsymbol{\rho})\}\mathbf{u} = \mathbf{F}_0, \tag{21}$$

where $\mathbf{K}$ and $\mathbf{M}$ are the stiffness and mass matrices, respectively. The peak amplitude of the harmonic load and the excitation frequency are denoted by $\mathbf{F}_0$ and ω_f, respectively. This system of equations is very similar to that of the static problem presented in Equation (7). At the element level, the equilibrium equation is

$$\{\mathbf{K}_{el}(\boldsymbol{\rho}) - \omega^2 \mathbf{M}_{el}(\boldsymbol{\rho})\}\mathbf{u}_{el} = \mathbf{F}_{res}, \tag{22}$$

where $\mathbf{F}_{res}$ comprises the loads applied to the element because of external loads and neighboring elements. Following the approach presented in Section 3, it can be shown that the system level and element level bounds proposed in Section 3 hold for this problem as well.

For a generalized eigenvalue problem in structural dynamics, the system of equations is [16]

$$\{\mathbf{K}(\boldsymbol{\rho}) - \omega_i^2(\mathbf{M}(\boldsymbol{\rho}) + \mathbf{M}_0)\}\mathbf{u}_i = \mathbf{0}, \tag{23}$$

where ω_i and $\mathbf{u}_i$ denote eigenfrequency and the respective eigenvector and $\mathbf{M}_0$ is the density-independent added mass matrix (dead load) which ensures the existence of unique solution for this system. In [16], it is shown that for $\mathbf{M}_0 = \mathbf{0}$, the eigenvalues are invariant with respect to the scaling of the structure, thus leading to nonuniqueness. Equation (23) can be rearranged as

$$\{\mathbf{K}(\boldsymbol{\rho}) - \omega_i^2 \mathbf{M}(\boldsymbol{\rho})\}\mathbf{u}_i = \omega_i^2 \mathbf{M}_0 \mathbf{u}_i. \tag{24}$$

For a given set of ω_i and $\mathbf{u}_i$, Equation (24) refers to a system of linear equations similar to Equation (21), and for this system as well, the maximum number of design variables should not be greater than the rank of the system coefficient matrix $\mathbf{K}(\boldsymbol{\rho}) - \omega_i^2 \mathbf{M}(\boldsymbol{\rho})$. For $\omega_i = 0$, rank of the system matrix reduces to the number of elastic modes. If the rigid body modes are eliminated, the rare possibility of a rank deficient system is when ω_i and $\mathbf{u}_i$ are an eigenfrequency and an eigenvector, respectively for the matrix $\mathbf{M}_0$ also. However, in general, $\mathbf{K}(\boldsymbol{\rho}) - \omega_i^2 \mathbf{M}(\boldsymbol{\rho})$ is a full rank matrix. At element level, the system of equations for an eigenvalue problem are the same as that in Equation (22). For TO problems involving more than one eigenfrequency, a larger system of equations might be obtained which can lead to a more relaxed upper bound on the number of design variables. This aspect has not been thoroughly explored in this work and is a direction for future research.

In addition to structural optimization, the condition stated in Section 3 holds for other topology optimization problems as well, e.g. thermal, electrical, thermomechanical, *etc*. The argument in Section 3 is based on the fact that the design domain is discretized using finite elements. Thus, it would still be valid except that the size of the element as well as the system coefficient matrices may differ based on the number of DOFs per node for the chosen problem. In addition,

DOI: 10.1002/nme

for problems involving other physics than structural mechanics, deformation modes and rigid body modes may not be the suited terms. For such problems, the proposed condition can be interpreted in terms of the ranks of the element and system coefficient matrices.

In this paper, the density field is represented using polynomial functions. However, because of the fact that the bounds derived in this paper depend on the *number* of design variables and not the type of the function chosen, this does not limit the scope of this work. With the chosen polynomial representation, we show that there is no gain in using too many design variables, and the same conclusion holds for other density representations. Also the pointwise density design variables defined within each element, commonly used in TO, can be seen as a special sampling of the polynomial, continuous density field we have used for our analysis.

In some recent studies [5–8], high resolution designs are obtained for cases where the design resolution is allowed to be very high compared with the analysis resolution. The reason for the apparent quality of these designs can be attributed to the use of smoothing operations such as projection filters [5, 6], Shephard interpolation schemes [7, 8], *etc*. Though smoothing can suppress the non-uniqueness to some extent, the possibility of non-uniqueness still exists. While the proposed condition needs to be satisfied at element level as well as system level, a possibility exists to increase the density resolution inside the element. One could choose to use higher order shape functions which introduce extra DOFs providing more flexibility for the state field representation [17, 18]. However, this only shifts the upper bound on the order of density representation and going beyond this new threshold will bring back the issues of non-uniqueness.

To conclude, the number of design variables for a TO problem depends on the applied finite element and should not be chosen arbitrarily. Rather, a choice of the number of design points should be made in a way such that it satisfies the proposed element level as well as the system level conditions. As shown in this paper, going beyond the proposed threshold is neither computationally more efficient, nor does it contribute to any improvement in the performance of the structure.

APPENDIX A

Theorem : Given that $\mathbf{B}$ is a $p \times q$ matrix and $\mathbf{D}$ is a $p \times p$ positive definite matrix of full rank p,

$$\text{rank}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}) = \text{rank}(\mathbf{B}).$$

Proof
Based on the rank-nullity theorem, following can be stated:

$$\text{rank}(\mathbf{B}) + \text{nullity}(\mathbf{B}) = q = \text{rank}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}) + \text{nullity}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}). \tag{A.1}$$

Let $\mathbf{x}$ denote any vector in the nullspace of $\mathbf{B}$. Then,

$$\mathbf{B}\mathbf{x} = \mathbf{0} \implies \mathbf{B}^\intercal\mathbf{D}\mathbf{B}\mathbf{x} = \mathbf{0} \implies \mathbf{x} \in \text{nul}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}) \implies \text{nul}(\mathbf{B}) \subseteq \text{nul}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}). \tag{A.2}$$

Similarly, let $\mathbf{y}$ denote any vector in the nullspace of $\mathbf{B}^\intercal\mathbf{D}\mathbf{B}$. Then,

$$\begin{aligned} &\mathbf{B}^\intercal\mathbf{D}\mathbf{B}\mathbf{y} = \mathbf{0} \implies \mathbf{y}^\intercal\mathbf{B}^\intercal\mathbf{D}\mathbf{B}\mathbf{y} = \mathbf{0} \implies (\mathbf{B}\mathbf{y})^\intercal\mathbf{D}\mathbf{B}\mathbf{y} = \mathbf{0} \\ &\implies \mathbf{B}\mathbf{y} = \mathbf{0} \implies \mathbf{y} \in \text{nul}(\mathbf{B}) \implies \text{nul}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}) \subseteq \text{nul}(\mathbf{B}). \end{aligned} \tag{A.3}$$

From Equation (A.2) and (A.3),

$$\text{nul}(\mathbf{B}) = \text{nul}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}) \implies \text{nullity}(\mathbf{B}) = \text{nullity}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}). \tag{A.4}$$

Substituting Equation (A.4) in (A.1), following is obtained:

$$\text{rank}(\mathbf{B}^\intercal\mathbf{D}\mathbf{B}) = \text{rank}(\mathbf{B}). \tag{A.5}$$

□

DOI: 10.1002/nme

APPENDIX B

Let $\mathcal{N}$ be a set of all vectors that lie in the nullspace of $\mathbf{P}$ and $\boldsymbol{\psi} = [\psi_1 \ \ \psi_2 \ \ \dots \psi_g]^\mathsf{T}$ denote a vector such that

$$\mathbf{P}\boldsymbol{\psi} = \mathbf{0} \ \ \forall \ \ \boldsymbol{\psi} \in \mathcal{N}. \tag{B.1}$$

Next, let $\boldsymbol{\gamma} = [\gamma_1 \ \ \gamma_2 \ \ \dots \gamma_r]^\mathsf{T}$ be a vector in the nullspace of $\mathbf{A^{(e)}}$. Then

$$\mathbf{A^{(e)}}\boldsymbol{\gamma} = \mathbf{0} \implies \mathbf{PQ}\boldsymbol{\gamma} = \mathbf{0}. \tag{B.2}$$

Comparing Equations (B.1) and (B.2),

$$\mathbf{Q}\boldsymbol{\gamma} = \boldsymbol{\psi} \ \ \forall \ \ \boldsymbol{\psi} \in \mathcal{N}, \tag{B.3}$$

where,

$$\mathbf{Q} = \begin{bmatrix} 1 & x_1 & \dots & (r^{\text{th}})_1 \\ 1 & x_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_g & \dots & (r^{\text{th}})_g \end{bmatrix}. \tag{B.4}$$

As per the matrix structure, $\mathbf{Q}$ is an alternant matrix similar to a Vandermonde matrix [19] expressed in two variables. If x and y are independent variables and no two points coincide, the rows and columns of this matrix are linearly independent and $\text{rank}(\mathbf{Q}) = \min(r, g)$. Based on this, we have the following cases:

- For $g < r$: The vector $\boldsymbol{\gamma}$ has more than one solution which implies that in Equation (B.2), there exists a nullspace of $\mathbf{A}^{(e)}$, and the design set $\mathbf{s}$ cannot be solved uniquely.
- For $g = r$: $\text{rank}(\mathbf{P}) = g$ which means that the vector set $\mathcal{N} = \emptyset$. Thus, $\mathbf{A}^{(e)}$ does not have a nullspace and $\text{rank}(\mathbf{A}^{(e)}) = r$.
- For $g > r$: Equation (B.3) is an overdetermined but inconsistent system of linear equations, thus $\boldsymbol{\gamma}$ has no solution. This again implies that $\mathbf{A}^{(e)}$ does not have a nullspace and rank $(\mathbf{A}^{(e)}) = r$.

ACKNOWLEDGEMENTS

This work is part of the Industrial Partnership Programme (IPP) 'Computational sciences for energy research' of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). This research program is co-financed by Shell Global Solutions International B.V. The numerical examples presented in this paper are implemented using deal.II [20], an open-source C++ package for adaptive finite elements. We would like to thank the developers of this software.

REFERENCES

1. Bendsøe MP, Kikuchi N. Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering* 1988; **71**(2):197–224.
2. Sigmund O, Maute K. Topology optimization approaches: a comparative review. *Structural and Multidisciplinary Optimization* 2013; **48**:1031–1055.
3. van Dijk NP, Maute K, Langelaar M, van Keulen F. Level-set methods for structural topology optimization: a review. *Structural and Multidisciplinary Optimization* 2013; **48**(3):437–472.
4. Deaton JD, Grandhi RV. A survey of structural and multidisciplinary continuum topology optimization: post 2000). *Structural and Multidisciplinary Optimization* 2014; **49**:1–38.
5. Nguyen TH, Paulino GH, Song Junho, Le CH. A computational paradigm for multiresolution topology optimization (MTOP). *Structural and Multidisciplinary Optimization* 2010; **41**(4):525–539.
6. Nguyen TH, Paulino GH, Song J, Le CH. Improving multiresolution topology optimization via multiple discretizations. *International Journal for Numerical Methods in Engineering* 2012; **92**:507–530.

DOI: 10.1002/nme

7. Wang Y, He J, Luo Z, Kang Z. An adaptive method for high-resolution topology design. *Acta Mechanica Sinica* 2013; **29**(6):840–850.
8. Wang Y, He J, Kang Z. An adaptive refinement approach for topology optimization based on separated density field description. *Computers and Structures* 2013; **117**:10–22.
9. Park J, Sutradhar A. A multi-resolution method for 3D multi-material topology optimization. *Computer Methods in Applied Mechanics and Engineering* 2015; **285**:571–586.
10. Groen JP, Langelaar M, Sigmund O, Ruess M. Higher-order multi-resolution topology optimization using the finite cell method. *International Journal for Numerical Methods in Engineering (submitted)*:1–15. DOI: 10.1002/nme.5432, (to appear in print).
11. Rojas-Labanda S, Stolpe M. Benchmarking optimization solvers for structural topology optimization. *Structural and Multidisciplinary Optimization* 2015; **52**(3):527–547.
12. Zowe J, Kocvara M, Bendsøe MP. Free material optimization via mathematical programming. *Mathematical Programming* 1997; **79**:445–466.
13. Bendsøe MP, Guades JM, Haber RB, Pederson P, Taylor JE. An optimal model to predict optimal material properties in the context of optimal structural design. *Journal of Applied Mechanics* 1994; **61**:930–937.
14. Marcus M, Minc H. *Introduction to Linear Algebra*. Dover Publications: Mineola, New York, U.S.A, 1988.
15. Jog CS, Haber RB. Stability of finite element models for distributed-parameter optimization and topology design. *Computer methods in applied mechanics and engineering* 1996; **130**(3):203–226.
16. Achtziger W, Kocvara M. Structural topology optimization with eigen values. *SIAM Journal of Optimization* 2007; **18**(4):1129–1164.
17. Babuska I, Szabo BA, Katz IN. The p-version of the finite element method. *SIAM Journal on Numerical Analysis* 1989; **18**:515–545.
18. Strouboulis T, Copps K, Babuska I. The generalized finite element method. *Computer Methods in Applied Mechanics and Engineering* 2001; **190**:4081–4193.
19. Macon N, Spitzbart A. Inverses of Vandermonde matrices. *American Mathematical Monthly* 1958; **65**(2):95–100.
20. Bangerth W, Hartmann R, Kanschat G. deal.II – A general purpose object oriented finite element library. *ACM Transactions on Mathematical Software* 2007; **33**(4):24/1–24/27.