Structural and Multidisciplinary Optimization
https://doi.org/10.1007/s00158-020-02649-6

REVIEW PAPER

A review on feature-mapping methods for structural optimization

Fabian Wein[1] · **Peter D. Dunning**[2] · **Julián A. Norato**[3]

Received: 23 October 2019 / Revised: 1 June 2020 / Accepted: 4 June 2020

Abstract

In this review we identify a new category of methods for implementing and solving structural optimization problems that has emerged over the last 20 years, which we propose to call *feature-mapping methods*. The two defining aspects of these methods are that the design is parameterized by a high-level geometric description and that features are mapped onto a non-body-fitted mesh for analysis. One motivation for using these methods is to gain better control over the geometry to, for example, facilitate imposing direct constraints on geometric features, while avoiding issues with re-meshing. The review starts by providing some key definitions and then examines the ingredients that these methods use to map geometric features onto a fixed mesh. One of these ingredients corresponds to the mechanism for mapping the geometry of a single feature onto a fixed analysis grid, from which an ersatz material or an immersed-boundary approach is used for the analysis. For the former case, which we refer to as the pseudo-density approach, a test problem is formulated to investigate aspects of the material interpolation, boundary smoothing, and numerical integration. We also review other ingredients of feature-mapping techniques, including approaches for combining features (which are required to perform topology optimization) and methods for imposing a minimum separation distance among features. A literature review of feature-mapping methods is provided for shape optimization, combined feature/free-form optimization, and topology optimization. Finally, we discuss potential future research directions for feature-mapping methods.

Keywords Structural optimization · Fixed-grid · High-level design · Feature-mapping

1 Introduction

Structural optimization methods can be classified into size, shape, and topology optimization. Size optimization modifies dimensions of the structure such as the cross section of truss members or the point-wise thickness of plates. Shape optimization modifies the boundaries of the structure, but without altering its topology, i.e., without adding or removing holes. Topology optimization can simultaneously change the shape of the structure and its connectivity.

A key aspect of these methods is the mechanism they employ to update the analysis model upon design changes. Some methods deform the analysis mesh when the design changes. Most topology optimization methods, and some shape optimization methods, use a mesh that does not conform to the boundaries of the structure. Density-based methods, which are the most prevalent topology optimization techniques, employ a pixel/voxel representation of the design, typically based on the analysis grid. Level-set methods, which can be used for both shape and topology optimization, use the zero level set of a function to define the structural boundaries. Density-based and level-set methods endow the optimizer with substantial freedom, rendering organic, free-form designs.

In recent years, new methods have been developed to implement and solve structural optimization problems that are motivated by obtaining designs that have some desired "high-level" geometric features (which we will define later in this article) without the need to re-create a

Responsible Editor: Ming Zhou

✉ Julián A. Norato
julian.norato@uconn.edu

Fabian Wein
fabian.wein@fau.de

Peter D. Dunning
peter.dunning@abdn.ac.uk

[1] Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

[2] University of Aberdeen, Aberdeen, UK

[3] University of Connecticut, Storrs, USA

Published online: 03 August 2020

body-fitted mesh upon design changes. These methods have been largely motivated by the need to embed primitive-shaped components in free-form designs, to design structures made of stock material, to control certain dimensions of the structure, and ultimately, to provide a geometric representation that is directly understood by computer-aided design (CAD) systems. In addition, these methods may represent designs with a low number of variables, which may be beneficial to, for example, the use of gradient-free optimizers. These methods build on aspects of existing techniques in (density-based) topology optimization, shape optimization and level-set methods. Despite their similarities, these methods originated independently and hence do not describe themselves as part of a common approach within structural optimization. Consequently there is no commonly used label for these new methods.

We propose the term *feature-mapping*, which is defined as a method that uses a high-level geometric feature parameterization that is mapped onto a non-body-fitted (and typically fixed) mesh for analysis (see Section 2.4 for the definition).

We note that density-based and level-set methods have successfully incorporated techniques to introduce some geometric constraints that are driven by manufacturing considerations (so-called manufacturing constraints) (cf., for instance, the reviews by Liu and Ma 2016; Liu et al. 2018b and the references therein). However, these techniques with low-level geometric parameterizations cannot readily impose constraints on high-level geometric features, and thus, they are not considered in this review.

We recognize that some methods covered in this review have possibilities that other methods do not. However, from the perspective of numerical implementation, all methods considered have two things in common: a high-level geometric description and mapping to a non-body-fitted mesh for analysis. Thus, the scope and focus of this review is on the numerical implementation (e.g., the feature-mapping), regardless of what the method is named, or whether it has other possibilities—although these are discussed in detail in Sections 6 to 8. Furthermore, we also recognize that some methods covered in this review could, or have been implemented without using a non-body-fitted mesh, e.g., they use the same high-level geometric description, but use a fitted mesh or the boundary element method (for example). However, this review is from the perspective of numerical implementation and the inclusion of a method here does not exclude it from being implemented in different ways, or belonging to a different categorization.

This review article is structured as follows. Section 2 provides some key definitions. Section 3 reviews methods for mapping a single feature to a fixed-grid, including pseudo-density and immersed-boundary methods. For pseudo-density methods we also use a test case to investigate material interpolation, boundary smoothing and numerical integration. Methods for combining features are reviewed in Section 4. Some feature-mapping methods employ separation constraints, which are reviewed in Section 5. We then give a literature review of feature-mapping methods for the following: shape optimization (Section 6), hybrid methods (which combine feature-mapping with free-form topology optimization; Section 7), and topology optimization (Section 8). Finally, we discuss potential future research directions for feature-mapping methods in Section 9.

2 Definitions and key components

2.1 High-level geometric features

In this paper, we define a geometric feature as a geometric solid with a high-level parameterization. A geometric solid is here understood as a closed regular set of points, i.e., a set that equals the closure of its interior (cf. Shapiro 2002). Physically, we consider the feature can either be a solid component or a hole in a solid component. By high-level parameters, we refer to those with a direct spatial dimension associated with the feature's size, position, or orientation. Examples of these parameters are the radius of a fillet, the thickness of a plate, or the location of a primitive (e.g., a bar or circle). Notably, these high-level parameters are the ones often employed to represent solids in CAD systems. The advantage of having these dimensions as direct design variables is that they simplify enforcing the presence of these features and controlling their dimensions, as opposed to the indirect and more verbose low-level representations of solids, such as those that are pixel or voxel-based.

2.2 Design region, body-fitted mesh, and fixed-grid

The design region corresponds to the sole region of space where material can be placed. A mesh is a spatial partition of the design region for the purpose of computing a numerical approximation of the structural response. A body-fitted mesh is one that conforms to the boundaries of the structure for any given design, i.e., a mesh in which no element boundary is cut by a structural boundary. While some changes in the shape of the structure can be accommodated by re-positioning the nodes in a body-fitted mesh, significant shape changes and all topological changes require the mesh to be entirely re-created in at least some portion of the structure. These changes can be involved and computationally expensive, particularly

for practical mesh sizes; and in the case of quadrilateral and hexahedral element meshes, this re-meshing cannot be robustly automated.

The opposite approach is that of non-body-fitted meshes, in which the element boundaries need not follow the structural boundaries. This is the prevalent approach in density-based topology optimization techniques, as it greatly facilitates accommodating design changes in the analysis. Limitations or additional effort in interpreting design boundaries are accepted in this approach. If the mesh remains fixed during the optimization process, it is often referred to as a *fixed-grid*. The term "mesh" is perhaps more adequate to describe this spatial discretization in that it does not necessarily convey that the partition is structured (as in, for example, a "10×10 grid"). However, the term "fixed-grid" is widely used in the structural optimization and computational mechanics literature to refer to the same concept and hence we adopt it here. Accordingly, when used alone (i.e., "grid" or "mesh"), we use these terms interchangeably.

The majority of the works reviewed in this manuscript use a fixed-grid approach (with the exception of some adaptive resolution methods referenced in Section 8.6), therefore we henceforth only refer to feature-mapping to a fixed-grid.

2.3 Explicit and implicit geometric representations

For reasons of clarity we first define the terms *explicit* and *implicit* with respect to the geometric representation of solid objects. During the course of our review, we encountered inconsistencies in the way these terms are used in the structural optimization community. In particular, the term explicit seems to be used in several instances whenever the geometric representation employs high-level parameters, regardless of the actual representation mechanism. An explicit representation is one where points on the solid (or its boundary) are generated by a rule, whereas an implicit representation is one where a rule provides a test as to whether or not a point belongs to the solid (Shapiro 2002). For example, an explicit representation of a disc of radius R centered at the origin is given by $f(t, r) = \{r\ \cos(t), r\ \sin(t)\}$, where values of the parameters $0 \leq t \leq 2\pi$ and $0 \leq r \leq R$ generate points within the disc. On the other hand, an implicit representation of the same disc is given by $f(\mathbf{x}) = \{1 \text{ if } \|\mathbf{x}\|_2 \leq R, 0 \text{ otherwise}\}$.

Density-based and classical level-set methods are implicit. In the discretized representation of density-based methods, given a point $\mathbf{x}$, the element-constant pseudo-density of the element that contains $\mathbf{x}$ determines if $\mathbf{x}$ is outside or inside the solid. This point classification test is not "sharp," however, since in most density-based methods the density is a relaxed continuous variable. Level-set methods are implicit by definition, since the value of the level-set function at $\mathbf{x}$ determines if $\mathbf{x}$ is inside or outside the structure. In level-set methods, a sharp representation of the boundary is available for any design throughout the optimization. This is also true when a diffuse boundary is used for the analysis, as in ersatz material methods (see Sigmund and Maute 2013). The fact that free-form density and level-set methods use implicit representations of the geometry is not fortuitous, since (a) implicit representations more easily accommodate topological changes than explicit ones (Shapiro 2002), and (b) the mesh can be used to parameterize the implicit representations, which facilitates coupling with fixed-grid analysis techniques, leading to efficient, robust methods to solve the governing equation and compute the design sensitivities.

We consider level-set methods as those that directly represent the design using an implicit function, independent of the design update approach, as discussed by van Dijk et al. (2013). It should be noted that some feature-mapping methods reviewed in this paper also utilize implicit functions in their formulation. Thus, they could be viewed as level-set methods. However, our aim is to emphasize the ability of feature-mapping methods to control high-level geometric features, which is not a property of level-set methods in general.

2.4 Feature-mapping

We define feature-mapping methods as those that capture high-level geometric features in their design parameterization and that map those features onto a non-body-fitted mesh to perform the analysis. Note that the high-level geometric description can be either explicit or implicit.

3 Geometry mapping to fixed-grid

There are currently two main approaches to mapping high-level geometric features onto a fixed-grid for analysis: pseudo-density-based mapping and immersed-boundary

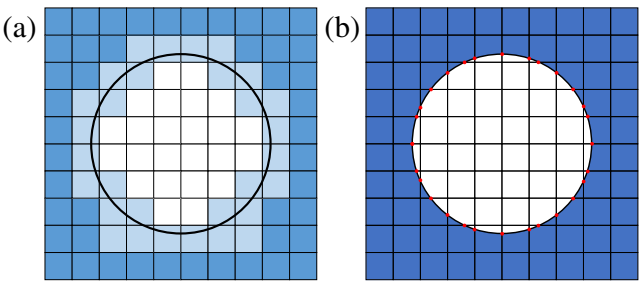


Fig. 1 Fixed-grid mapping methods: (a) pseudo-density, and (b) immersed-boundary based

mapping, as illustrated in Fig. 1. Both approaches utilize a fixed-grid for the analysis, thus circumventing the need for re-meshing during optimization.

Generally speaking, the purpose of fixed-grid analysis methods is to replace volume integrals evaluated over the structural domain ω with integrals over a domain $\Omega \supseteq \omega$ that encompasses the structure

$$\int_{\omega} f(\mathbf{x})\,\mathrm{d}\,\mathbf{x} = \int_{\Omega} \chi_{\omega}(\mathbf{x})\, f(\mathbf{x})\,\mathrm{d}\,\mathbf{x}, \tag{1}$$

where χ_{ω} is the *characteristic* or *indicator* function defined as

$$\chi_{\omega}(\mathbf{x}) := \begin{cases} 1 & \text{if } \mathbf{x} \in \omega \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

and f is the domain integrand (for example, the virtual strain energy density in elasticity).

Feature-mapping techniques that employ element-constant pseudo-densities accomplish this by replacing the corresponding element volume integral over Ω_e as

$$\int_{\Omega_e} \chi_{\omega}(\mathbf{x})\, f(\mathbf{x})\,\mathrm{d}\,\mathbf{x} \approx \rho_e \int_{\Omega_e} f(\mathbf{x})\,\mathrm{d}\,\mathbf{x}, \tag{3}$$

with the element-constant pseudo-density

$$\rho_e = \frac{1}{|\Omega_e|} \int_{\Omega_e} \chi_{\omega}(\mathbf{x}). \tag{4}$$

Further, ρ_e might be subject to an interpolation function μ (see Section 3.1.1). Consequently, the element stiffness matrix is: $\mathbf{K}_e = \mu(\rho_e)\,\mathbf{K}_e^0$, where $\mathbf{K}_e^0$ is the "fully-solid" element matrix. This approach is also known as the *ersatz material* approach. In feature-mapping methods, ρ_e depends explicitly and (ideally) smoothly on the high-level parameterization of the geometric features. This mapping typically leads to elements of intermediate pseudo-density near the feature boundary (see Figs. 1(a) and 2). If the mapping is differentiable, the chain rule can be readily used to obtain sensitivities with respect to the high-level geometric parameters, as we explain in Section 3.1.5.

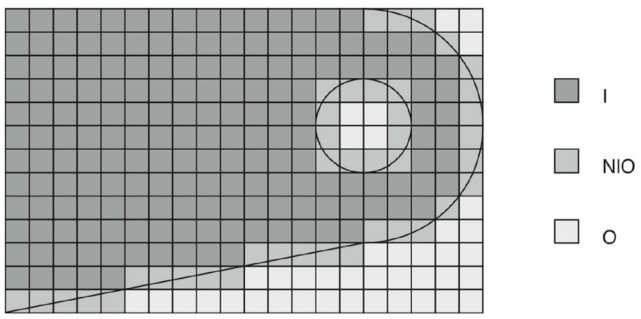


Fig. 2 Fixed-grid mapping in García-Ruíz and Steven (1999) with I (inside), O (outside) and NIO (neither inside nor outside) elements

The immersed-boundary approach employs techniques widely used in finite element analysis, such as the extended finite element method (XFEM), to capture sharp interfaces on a fixed-grid. In other words, the element volume integrals are evaluated as

$$\int_{\Omega_e} \chi_{\omega}(\mathbf{x})\, f(\mathbf{x})\,\mathrm{d}\,\mathbf{x} = \int_{\Omega_e \cap \omega} f(\mathbf{x})\,\mathrm{d}\,\mathbf{x}. \tag{5}$$

These techniques have the advantage over pseudo-density approaches that there are no "gray regions," which require some assumption on their material properties. For the same reason, immersed-boundary methods (in principle) render more accurate analysis solutions. These advantages come at the expense of challenges on, e.g., numerical evaluation of integrals in elements cut by the structure boundary and sensitivity calculation. These challenges and proposed solutions from the literature are discussed in Section 3.2.

In the remainder of this section some important aspects of these approaches are discussed and examined using test case examples, in relation to their application to feature-mapping methods.

3.1 Element-constant pseudo-density

Feature-mapping techniques that use pseudo-densities essentially differ in the way they compute ρ_e to be used in (3).

An often used approach in earlier feature-mapping methods is to compute the element pseudo-density ρ_e as an approximation of the volume fraction, defined as the portion of the element that intersects the feature, $|\omega \cap \Omega_e|/|\Omega_e|$. By making simplifying assumptions about the shape of the intersected region, it is possible to use simple expressions to approximate this volume fraction.

García-Ruíz and Steven (1999) is the earliest publication where the volume fraction approach on fixed-grids is used in the context of feature-mapping methods (see Fig. 2). Elements are classified as completely inside (I), completely outside (O), or neither inside nor outside (NIO) the structure, i.e., if the element is cut by the boundary of the geometric feature. I and O elements have pseudo-densities of 1 and $\rho_{\min}$, respectively (with $0 < \rho_{\min} \ll 1$ a small bound to prevent an ill-posed analysis). The pseudo-density of NIO elements is computed as the volume fraction of the portion of the element that intersects the feature (albeit few details are provided as to how this fraction is computed), and there is no further interpolation, i.e., $\mu(\rho_e) = \rho_e$ (see Sections 3.1.1 and 3.1.2 for discussion on the limitations of the element volume fraction approach).

As we will detail in Section 4, all pseudo-density techniques employ an implicit geometric representation of the feature; even when the high-level parametric representation

is explicit (e.g., using a B-spline), it is first converted to an implicit representation ϕ_ω satisfying the properties

$$\begin{cases} \phi_\omega(\mathbf{x}) > 0, & \mathbf{x} \in \omega \\ \phi_\omega(\mathbf{x}) = 0, & \mathbf{x} \in \partial\omega \\ \phi_\omega(\mathbf{x}) < 0, & \mathbf{x} \notin \omega, \end{cases} \tag{6}$$

where $\mathbf{x}$ is a point in the fixed-grid design domain and $\partial\omega$ is the feature boundary. Note that there is no convention in the structural optimization literature on the meaning of the implicit function sign. In (6), we define positive values as being inside the feature. Also, the implicit function may be defined as the signed-distance function, $d(\mathbf{x})$, where the magnitude is computed as the shortest distance from the point $\mathbf{x}$ to the feature boundary $\partial\omega$ and the sign is the same as $\phi_\omega(\mathbf{x})$. The distance can be computed by different approaches, which are discussed in Section 3.1.7.

With the implicit function ϕ_ω representing the feature, the Heaviside function is:

$$H\left(\phi_\omega(\mathbf{x})\right) = \begin{cases} 1 \text{ if } \phi_\omega(\mathbf{x}) \geq 0 \\ 0 \text{ if } \phi_\omega(\mathbf{x}) < 0. \end{cases} \tag{7}$$

Clearly, H can directly replace χ_ω in the left-hand side of (3). The volume fraction approach directly uses H. However, H (and χ_ω) is not differentiable, and so it is often replaced by a smooth approximation $\widetilde{H}$ (see Section 3.1.4 for a detailed discussion of smooth boundary modeling functions). One could consider the function values of $\widetilde{H}$ as a continuous pseudo-density field.

The element pseudo-density is then found by integrating the Heaviside function, or continuous pseudo-density field, over the element volume as

$$\rho_e = \frac{1}{|\Omega_e|} \int_{\Omega_e} \widetilde{H}(\phi_\omega(\mathbf{x}))\, \mathrm{d}\, x, \tag{8}$$

which may be evaluated directly, or approximated by numerical integration in the form of a weighted sum

$$\rho_e = \sum_i^{N_{\text{ip}}} w_i\ \widetilde{H}(\phi_\omega(\mathbf{x}_i)). \tag{9}$$

The process of mapping conceptually consists of firstly generating the continuous pseudo-density field and then its integration (see Section 3.1.6 for more information on numerical integration).

Thus, feature-mapping using pseudo-densities requires a choice of several key ingredients: (1) the type of material interpolation function, $\mu(\rho_e)$; (2) the form of the Heaviside (or smooth boundary) function, $\widetilde{H}$; (3) the form of the implicit function, $\phi_\omega(\mathbf{x})$ (signed distance or otherwise); and (4) the integration method used to evaluate (8). These ingredients are now discussed in detail, using test cases to highlight the effect of certain choices.

3.1.1 Material interpretation of pseudo-density

Pseudo-density-based feature-mapping approaches not only inherit the advantages of density-based topology optimization in terms of the easy analysis and sensitivity computation but also inherit one of its challenges, which is how to interpret material properties for intermediate values of the pseudo-density. This is dictated by the form of the function $\mu(\rho_e)$ in (3).

In the pioneering work on topology optimization by Bendsøe and Kikuchi (1988), the stiffness properties of porous material, in between solid and void, were determined by mathematical homogenization of a periodic structure. This work showed that, due to its unfavorable stiffness-to-porosity ratio, intermediate material was barely used. This led to the famous power law approximation ρ^p introduced in Bendsøe (1989), now known as the Solid Isotropic Material with Penalization (SIMP) model, to replace the homogenization step (see also Rozvany et al. 1992). In Bendsøe and Sigmund (1999) it was shown that the power law

$$\mu_{\text{PL}}(\rho) = \rho^p \tag{10}$$

with exponent $p = 3$ never overestimates the maximal physical stiffness-to-porosity relationship of isotropic material given by the upper Hashin-Shtrikman bounds (see Fig. 3). Homogenization and the Hashin-Shtrikman bounds show that the relative stiffness of porous isotropic material is below its volume fraction. In Fig. 3 two further graphs are given: the linear interpolation

$$\mu_{\text{lin}}(\rho) = \rho \tag{11}$$

corresponds to the case of interpreting the volume fraction of an element covered by a geometry directly as pseudo-density, and the RAMP (Rational Approximation of Material Properties) (Stolpe and Svanberg 2001) interpolation

$$\mu_{\text{RAMP}}(\rho) = \rho/(1 + q(1 - \rho)) \tag{12}$$

is shown for $q = 1$ and discussed further below. Both linear and RAMP (with $q = 1$) interpolations are above the upper Hashin-Shtrikman bounds and therefore overestimate the stiffness of intermediate material in a non-physical way (particularly the linear interpolation). Consequently, we want to emphasize that the term *penalization* in SIMP does not indicate a mathematical trick to prevent intermediate material in the optimal design, but a realistic and physical modeling of stiffness-to-porosity relationship.

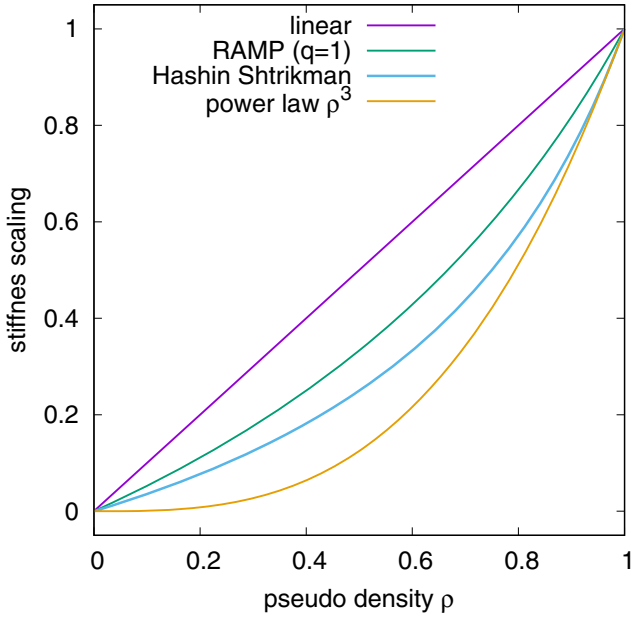


Fig. 3 Different material interpolation functions (see Section 3.1.1). Only the Hashin-Shtrikman bound and the common SIMP power law satisfy physical limits for an isotropic material with intermediate density

3.1.2 Test problem to investigate the effect of intermediate material

A test problem is introduced in Fig. 4 to investigate the effect of intermediate material modeling, particularly in the context of feature-mapping methods. The vertical bar is subject to a continuous horizontal movement with position s. The width of the bar is four elements. According to Fig. 2 we assign a pseudo-density $\rho = 1$ for elements fully contained in the bar (I), a very small value to elements fully outside the bar (O), and for the partially covered elements (NIO) a density value corresponding to the covered element volume fraction (which is the same as using the exact Heaviside in (3)).

Note that if NIO elements along the left-hand boundary of the bar have pseudo-density ρ, then NIO elements along the right-hand boundary have a pseudo-density $1 - \rho$, since the width of the bar is a multiple of the element size.

The element pseudo-density ρ_e is then interpolated using the functions shown in Fig. 3. We refer to $\mu(\rho)$ as the *physical pseudo-density*, since, according to (3), this is the element-constant material property scaling in the finite element analysis. The upper Hashin-Shtrikman bounds are given for a Poisson's ratio of 0.3 as $\mu(\rho) = \rho/(3 - 2\rho)$ (see Bendsøe and Sigmund 1999). To measure the grayness of the boundary elements we introduce

$$g(\mu(\rho_e)) = 4\,\mu(\rho_e)\,(1 - \mu(\rho_e)), \tag{13}$$

where $\mu(\rho_e) = 0.5$ results in the highest grayness value of 1.0.

The compliance for the test problem in Fig. 4 is evaluated for the linear material interpolation (material stiffness directly proportional to the covered element fraction) in Fig. 5a and additionally for all material interpolation functions from Fig. 3 in Fig. 5(c). The linear material interpolation shows improved (lower) compliance when the bar edge is positioned between elements, giving a high grayness value (see Fig. 5(a)). Peaks of poorer (high) compliance are seen when the bar is aligned with element edges, resulting in no gray boundary elements. Using the upper Hashin-Shtrikman bounds for material interpolation shows an increased compliance when gray elements are involved, reflecting the inefficient stiffness of porous structures in reality. This effect is amplified for the classical SIMP power law. Realistic compliance values are obtained only when the bar edges align with element boundaries, as there are no intermediate, gray densities.

These three material interpolation functions show that the compliance is non-monotonic with respect to the bar position s. In an optimization problem, the linear interpolation function will likely favor intermediate bar positions, while the Hashin-Shtrikman bounds and the power law will likely favor bar positions aligned to mesh elements - hence the problem becomes somewhat mesh-dependent. We note that this effect is caused only by the process of mapping the feature onto the fixed-grid using pseudo-densities, since analysis of this example with a conforming mesh would render a monotonic compliance curve. Interestingly, the RAMP interpolation function with parameter $q = 1$, which is $\mu(\rho) = \frac{\rho}{2-\rho}$, exhibits an almost monotonic compliance with respect to the design change for this test problem.

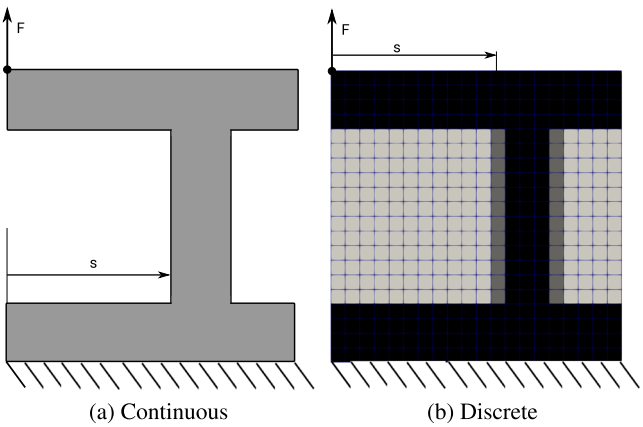
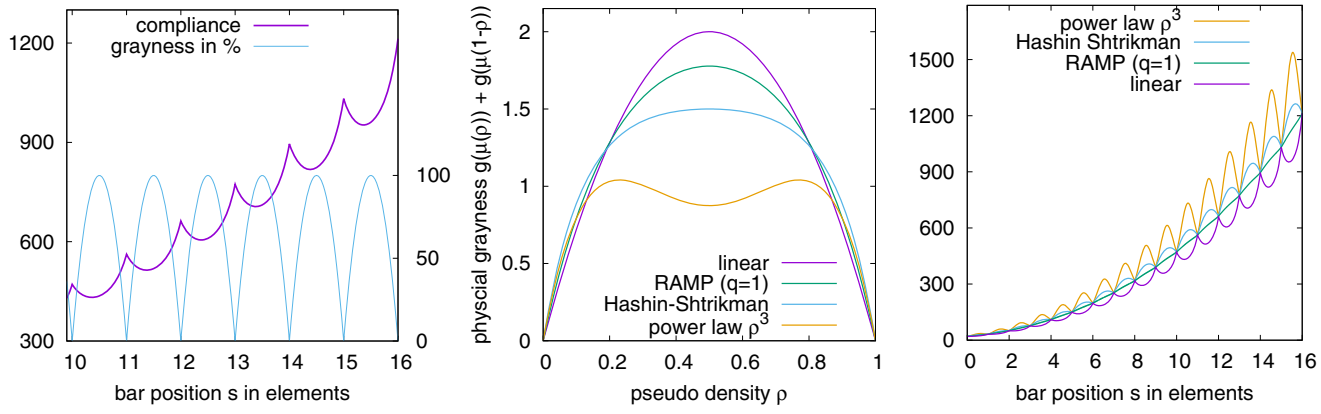


Fig. 4 H-shape test problem where the horizontal position of the vertical bar is given by the variable s. The force vector is applied on the upper left corner: (a) continuous setting; (b) discretized problem



(a) Linear (b) Grayness (c) Compliance

Fig. 5 (a) Compliance and grayness for the benchmark from Fig. 4 with linear material interpolation (where stiffness is scaled by the volume fraction covered by the bar); (b) sum of grayness values of the physical density $g(\mu(\rho))$ (13) for left and right edges of the vertical bar, corresponding to moving s by one element in Fig. 4 for different interpolation functions; and (c) their corresponding structural compliance values. See Section 3.1.1 for a detailed discussion

The authors are only aware of one work in the literature where this interpolation function is used in feature-mapping methods (Zhang et al. 2017a). However, RAMP is not introduced with the purpose of avoiding mesh dependency, but to favor reintroduction of geometric features during optimization.

In addition to non-monotonicity, Fig. 5(c) reveals another aspect of the feature-mapping: the compliance is non-smooth, as it exhibits "kinks" whenever the vertical boundaries of the moving bar coincide with element boundaries. This non-smoothness is present even for the seemingly smoother RAMP interpolation.

3.1.3 Principal boundary modeling approaches

Mapping a feature to a fixed analysis grid requires modeling the boundary when the design is not exactly aligned to the mesh. However, as shown in the previous section, this may lead to non-monotonicity, non-smoothness and mesh dependency. The results in Section 3.1.1 assume the boundary is modeled by an exact Heaviside function (7). In this section, we investigate the effect of using a smoothed Heaviside, or boundary smoothing, on the test problem.

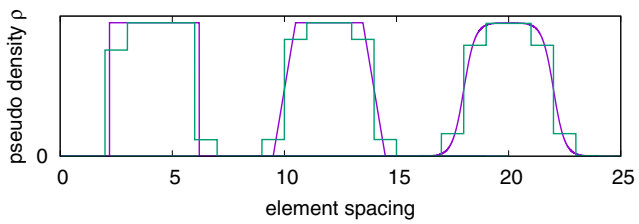


Fig. 6 1D boundary modeling (in magenta): exact Heaviside function (left), piecewise linear (14) (center) and tanh smoothing (15) (right) and their element-constant pseudo-density values (in green)

First, we consider a 1D model of the bar cross section in the test case example of Fig. 4. This feature is modeled by assigning a pseudo-density $\rho = 1$ to any point inside the feature and a very small value $\rho_{\min}$ to points outside the feature (similar to the characteristic function as defined in (2)). In Fig. 6, three approaches to model the transition between material and void across the boundary are shown. The first is an exact Heaviside function, (7), as in Section 3.1.1; this function is discontinuous, a fact we denote as the function being C^{-1}. The other functions are as follows: a continuous, but non-differentiable piecewise linear function, which is C^0

$$\widetilde{H}_{\text{lin}}(d(\mathbf{x}), h) := \begin{cases} \rho_{\min} & \text{if } d(\mathbf{x}) < -h \\ (1-\rho_{\min})\,\frac{d(\mathbf{x})}{2h} + \frac{1+\rho_{\min}}{2} & \text{if } |d(\mathbf{x})| \le h \\ 1 & \text{if } d(\mathbf{x}) > h, \end{cases} \tag{14}$$

where h defines the size of the transition zone between material and void, and a tanh-like function, which is C^∞, as used in Wein and Stingl (2018)

$$\widetilde{H}_{\tanh}(d(\mathbf{x}), \beta) = (1-\rho_{\min})\left(1 - \frac{1}{e^{\beta\, d(\mathbf{x})} + 1}\right) + \rho_{\min}, \tag{15}$$

where β is a parameter that controls the size of the transition zone. The function is also known as sigmoid function. Note that we have chosen the signed-distance function ($d(\mathbf{x})$) as the implicit function in the above equations. For further discussion on how the choice of implicit function affects boundary mapping, see Section 3.1.4.

To obtain the element-constant pseudo-densities for the fixed analysis grid we evaluate (8). For simplicity, we

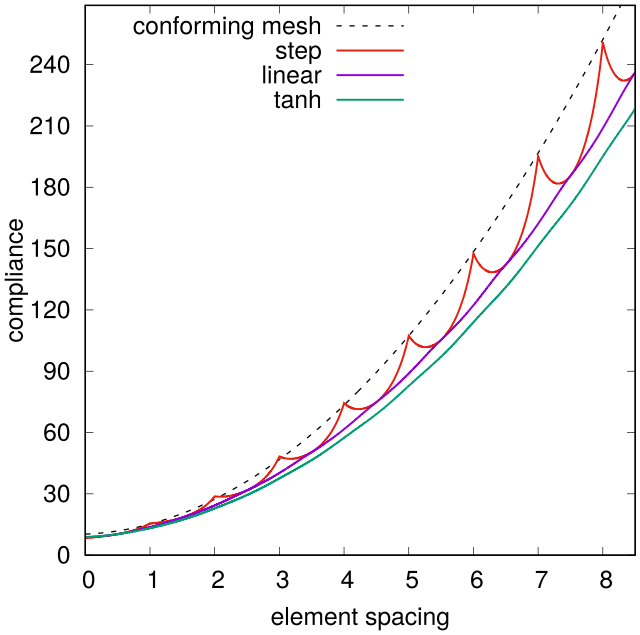


Fig. 7 Compliance values for the boundary modeling functions in Fig. 6 applied to the test problem in Fig. 4. Due to the linear density-to-stiffness interpolation the compliance is underestimated (see Fig. 5(c)). The compliance for a conforming mesh is plotted for reference

assume the linear material interpolation model from Fig. 3. The corresponding element-constant pseudo-densities are shown in Fig. 6. Again, the element grayness values depends on the alignment of the boundary modeling function with respect to the mesh (see Fig. 5(b)). However, the sum over all elements attains a constant value for the piecewise linear and tanh functions (see Fig. 12). Applying these functions to the test problem, we get the compliance values shown in Fig. 7. The result for the exact Heaviside function has already been given in Fig. 5(c); the piecewise linear and tanh functions result in visually smooth compliance functions that are artificially good (low). The smoothness results from the boundary smoothing, while the artificially low compliance results from the linear material interpolation.

Upon evaluation of (8), the resulting pseudo-density $\rho_e(s)$ and hence the compliance become C^0 for the exact Heaviside function, C^1 for the piecewise linear function and stays C^∞ for the tanh function, with respect to a change in s. For optimization we require C^1, hence the piecewise linear boundary modeling function is sufficient. For more discussion on the effect of numerical integration, see Section 3.1.6.

3.1.4 Further smooth boundary modeling approaches

In this section, we review further smooth boundary modeling approaches used in feature-mapping methods. Typically they are piecewise defined with a transitioning zone controlled by the parameter h, and transition function σ_{trans} as

$$\widetilde{H}(\phi_\omega(\mathbf{x}), h) := \begin{cases} \rho_{\min} & \text{if } \phi_\omega(\mathbf{x}) < -h \\ \sigma_{\text{trans}}(\phi_\omega(\mathbf{x}), h) & \text{if } |\phi_\omega(\mathbf{x})| \le h \\ 1 & \text{if } \phi_\omega(\mathbf{x}) > h. \end{cases} \tag{16}$$

For example, the transition function for the piecewise linear boundary model (14) is

$$\sigma_{\text{lin}} = (1 - \rho_{\min}) \frac{\phi(\mathbf{x})}{2h} + \frac{1 + \rho_{\min}}{2}. \tag{17}$$

A common choice of transition function is based on a spline representation as a cubic function with zero slope on both sides of the transition zone, which is used by, e.g., Zhang et al. (2016c) and Dunning (2018)

$$\sigma_{\text{poly}} = \frac{3\,(1 - \rho_{\min})}{4} \left(\frac{\phi(\mathbf{x})}{h} - \frac{\phi(\mathbf{x})^3}{3\,h^3} \right) + \frac{1 + \rho_{\min}}{2}. \tag{18}$$

Another choice is based on a trigonometrical function:

$$\sigma_{\cos} = \frac{(1 - \rho_{\min})}{2} \cos\left(\left(\frac{\phi(\mathbf{x})}{2h} - \frac{1}{2} \right) \pi \right) + \frac{1 + \rho_{\min}}{2}. \tag{19}$$

Note that the tanh-like function (15) has no finite transition zone.

In this section, the more general implicit function, $\phi(\mathbf{x})$, is used instead of the signed-distance function, $d(\mathbf{x})$, as not all feature-mapping methods use a signed-distance function. If a signed-distance function is used, then the magnitude of the implicit function spatial gradient is 1, $||\nabla d|| = 1$, and the width of the transition zone between solid and void is defined as follows: $w = 2h$. In Section 3.1.6 we also introduce the discrete element transition zone w_b. If $||\nabla\phi|| < 1$, the transition zone will be stretched, with $w > 2h$. Conversely, if $||\nabla\phi|| > 1$, then the transition zone will be compressed, with $w < 2h$. This issue is discussed and investigated by Zhou et al. (2016), where it is argued that a signed-distance function should be used to avoid issues caused by a varying spatial gradient of ϕ around the feature boundary, as this influences the accuracy of the structural response and gradient computation (see also Section 3.1.6). For further discussion on computing the signed-distance function, see Section 3.1.7.

Section 3.1.2 shows the limitations of obtaining the element pseudo-density ρ_e as volume fraction of a grid cell e covered by a non-smoothed feature. A particular issue demonstrated in Norato et al. (2004) is that the volume fraction calculation becomes non-differentiable if a portion with non-zero measure of the feature boundary coincides with the element boundary—for instance, in the test problem of Fig. 4 when the sides of the vertical bar align with element boundaries. This problem can be readily circumvented by using a circular (2D) or spherical (3D)

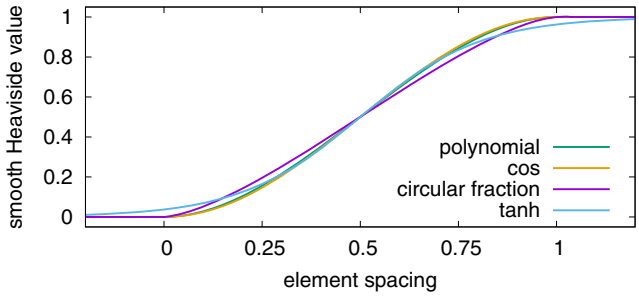


Fig. 8 Smooth boundary modeling by polynomial (18) (almost identical to cos (19)) and volume fraction with circular sampling window (20) with transition zone $2h$ equal to one element width. The tanh-based function (15) is plotted with $\beta = 6.5$. All functions are plotted assuming a signed-distance function

sampling window (instead of the element itself) to compute the volume fraction, and by linearizing the boundary of the feature within the sampling window (cf. Norato et al. 2015). This leads to a closed formula for the transition function, which is given for 2D as

$$\sigma_{\text{circ}} = \frac{1}{\pi}\left[\cos^{-1}\left(\frac{-d(\mathbf{x})}{h}\right) + \frac{d(\mathbf{x})}{h}\sqrt{1-\left(\frac{d(\mathbf{x})}{h}\right)^2}\right] \tag{20}$$

within the framework of (16). However, note that numerical integration in (8) is not used in this case, as (20) is derived from an exact analytical integration of the volume fraction of the linearized feature boundary within the circular sampling window and thus only requires the signed-distance information from the element center to the feature boundary.

A comparison of boundary smoothing functions is shown in Fig. 8.

3.1.5 Sensitivity analysis

One of the appealing features of the element pseudo-density approach in feature-mapping methods is that, as in density-based topology optimization, the computation of design sensitivities is much simpler than for approaches that must compute boundary sensitivities (as in some level-set methods). Moreover, as we will show in this section, the computation of sensitivities is closely connected to that of density-based methods.

Sensitivity analysis in density-based topology optimization is well established. It can be readily performed on the discretized algebraic system resulting from a finite element analysis for a wide range of functions, and even multiphysics problems fit one of the known generalized derivations (see Bendsøe and Sigmund 2003).

We briefly review sensitivity analysis for standard density-based topology optimization and consider the easy static case, where the finite element system matrix $\mathbf{K}$ depends explicitly on the vector of element pseudo-densities $\boldsymbol{\rho}$, the state solution $\mathbf{u}$ depends only implicitly on $\boldsymbol{\rho}$ and the boundary conditions are assumed to be design-independent. The system of linear equations arising from the finite element discretization reads

$$\mathbf{K}(\boldsymbol{\rho})\,\mathbf{u}(\boldsymbol{\rho}) = \mathbf{f}. \tag{21}$$

Using adjoint differentiation (e.g., Tröltzsch 2010), the sensitivity of a function $J(\boldsymbol{\rho}, \mathbf{u}(\boldsymbol{\rho}))$ with respect to an element pseudo-density can be written as

$$\frac{\mathrm{d}\,J}{\mathrm{d}\,\rho_e} = \frac{\partial J}{\partial \rho_e} + \boldsymbol{\lambda}(\boldsymbol{\rho})^\top \frac{\partial K_e}{\partial \rho_e}\,\mathbf{u}(\boldsymbol{\rho}). \tag{22}$$

The partial derivative $\frac{\partial J}{\partial \rho_e}$ is for many functions zero, $\frac{\partial \mathbf{K}_e}{\partial \rho_e}$ is trivial to obtain and $\boldsymbol{\lambda}$ solves the adjoint problem

$$\mathbf{K}(\boldsymbol{\rho})\,\boldsymbol{\lambda}(\boldsymbol{\rho}) = -\left(\frac{\partial J}{\partial \mathbf{u}}\right)^\top. \tag{23}$$

Notably, the adjoint solution $\boldsymbol{\lambda}$ is independent of the design parameterization, because the pseudo-load in (23) does not depend explicitly on the design variables. Consequently, the adjoint solution needed for feature-mapping methods is the same as the one obtained for other topology optimization techniques.

Once the adjoint solution is computed, feature-mapping methods with pseudo-densities only need to compute the derivative of the boundary mapping function to obtain derivatives of pseudo-densities with respect to the high-level design parameters, s_j. The final derivative is obtained by the chain rule as

$$\frac{\mathrm{d}\,J}{\mathrm{d}\,s_j} = \sum_e^{N_\mathrm{e}} \left\{\left[\frac{\partial J}{\partial \rho_e} + \frac{\partial \mu}{\partial \rho_e}\boldsymbol{\lambda}(\boldsymbol{\rho})^\top \mathbf{K}_e\,\mathbf{u}(\boldsymbol{\rho})\right]\frac{\partial \rho_e}{\partial s_j}\right\}. \tag{24}$$

Note that the boundary modeling function (16) is constant outside of the transition region, hence $\frac{\partial \widetilde{H}}{\partial s_j} = 0$ in the void region (i.e., $\widetilde{H} = \rho_{\min}$) and the solid region (i.e., $\widetilde{H} = 1$). It thus follows from (9) that $\frac{\partial \rho_e}{\partial s_j}$ is non-zero only in regions with intermediate pseudo-density values, namely in the gray regions around the boundaries of the structure (see Fig. 9). The choice of width h of the smoothing functions in Section 3.1.4 controls the amount of information collected from (22).

3.1.6 Numerical integration of the boundary mapping function

In principle, density-based feature-mapping requires the element-constant pseudo-density to be found by integrating the smoothed Heaviside, or boundary mapping function (8). In the test case example from Fig. 4, the vertical feature is aligned with the fixed-grid. This effectively makes the volume integral of the boundary modeling function 1D.

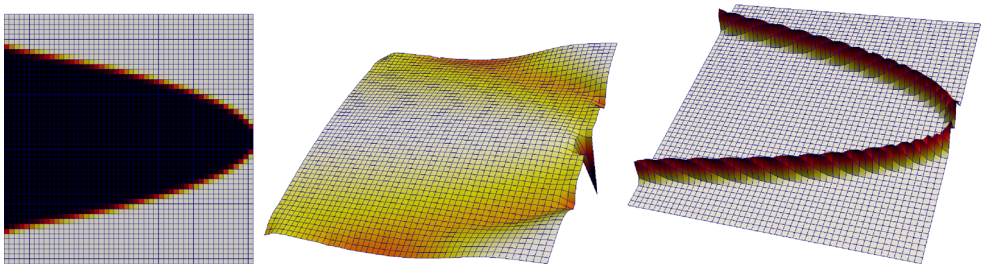


Fig. 9 For a smoothly mapped shape (left) the compliance sensitivity with respect to pseudo-density (22) is shown in the center. The element-wise summand of the sensitivity with respect to shape variables (24) is non-zero only where the pseudo-density has intermediate values (right). From Wein and Stingl (2018)

Thus, analytical integration is reasonably straightforward and is used to generate the results above.

However, analytical integration can become involved in two and three dimensions. Therefore, many methods compute the pseudo-density by numerical integration as a weighted sum via (9).

The boundary modeling function (16) influences the choice of quadrature rule and the number of sampling points. We now examine the effect of the number of sampling points when using Newton-Cotes formulae to evaluate (9). Note that zero-degree quadrature corresponds to midpoint integration (i.e., the value of the function at the element center) and first-degree quadrature corresponds to the trapezoidal rule (i.e., the average of function values at the corner positions of the element). The number of integration points for the Newton-Cotes formula with degree deg is $N_{\text{ip}} = (\text{deg} + 1)^{\text{dim}}$.

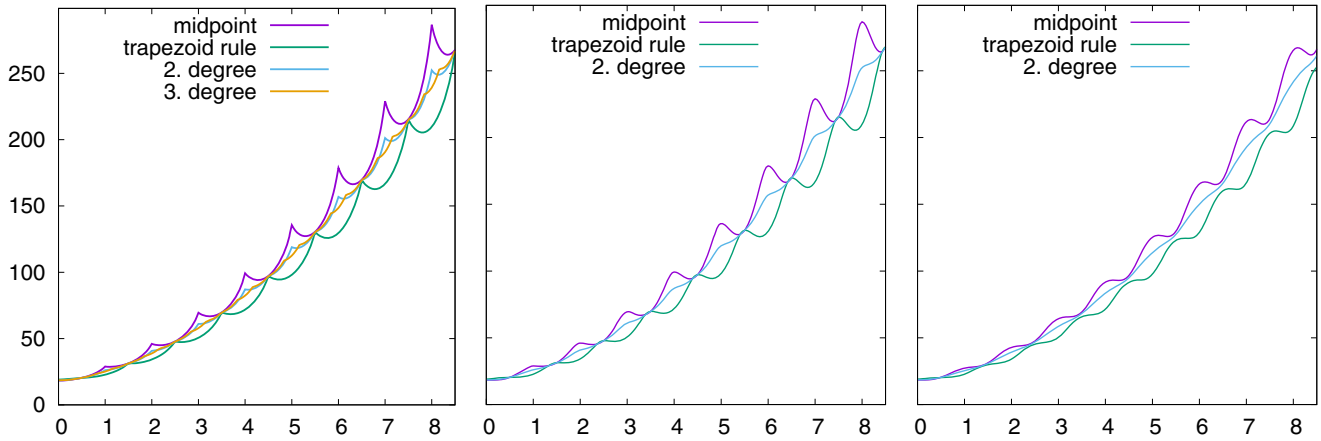
The investigation uses the piecewise linear (14) and polynomial (18) boundary modeling functions with transition zone w of one element. The tanh-like function (15) is also included, with a similar maximal slope. The signed-distance implicit function is used. It is clear that element-wise numerical integration of the density function does not increase the regularity with respect to the shape variables. In particular the piecewise linear boundary modeling function (17) stays non-continuous differentiable. Thus, this combination is not suitable for gradient based optimization, but we feel it worth including in the discussion.

Figure 10 clearly shows that, for the test problem, the smoothing effect shown with analytical integration in Fig. 7 is lost when the number of sampling points in numerical integration is too low. Although, it should be noted that the test case is selected to reveal extreme response.

When using numerical integration, the term $\frac{\partial \rho_e}{\partial s}$ in (24) is found by

$$\frac{\partial \rho_e}{\partial s} = \sum_i^{N_{\text{ip}}} w_i \, \frac{\partial \widetilde{H}(\mathbf{x}_i)}{\partial s}. \tag{25}$$

We note that the multi-resolution approach in Liu et al. (2018) effectively also performs higher order numerical integration. The test problem is now used to investigate the effect of the number of sampling points in numerical integration on (25). We consider the integral of the first element from the left and vary the bar position (see the upper row in Fig. 11). We also introduce the element transition zone, w_ρ, which is defined as the number of elements across



(a) linear (17) (b) polynomial (18) (c) tanh (15)

Fig. 10 We perform the similar experiment as in Fig. 7 but this time with numerical integration via low order closed Newton-Cotes formulae (see Fig. 6). On the x-axis, the positional variable s is varied in element units, the y-axis shows the compliance. (a) Linear (17). (b) Polynomial (18). (c) tanh (15)

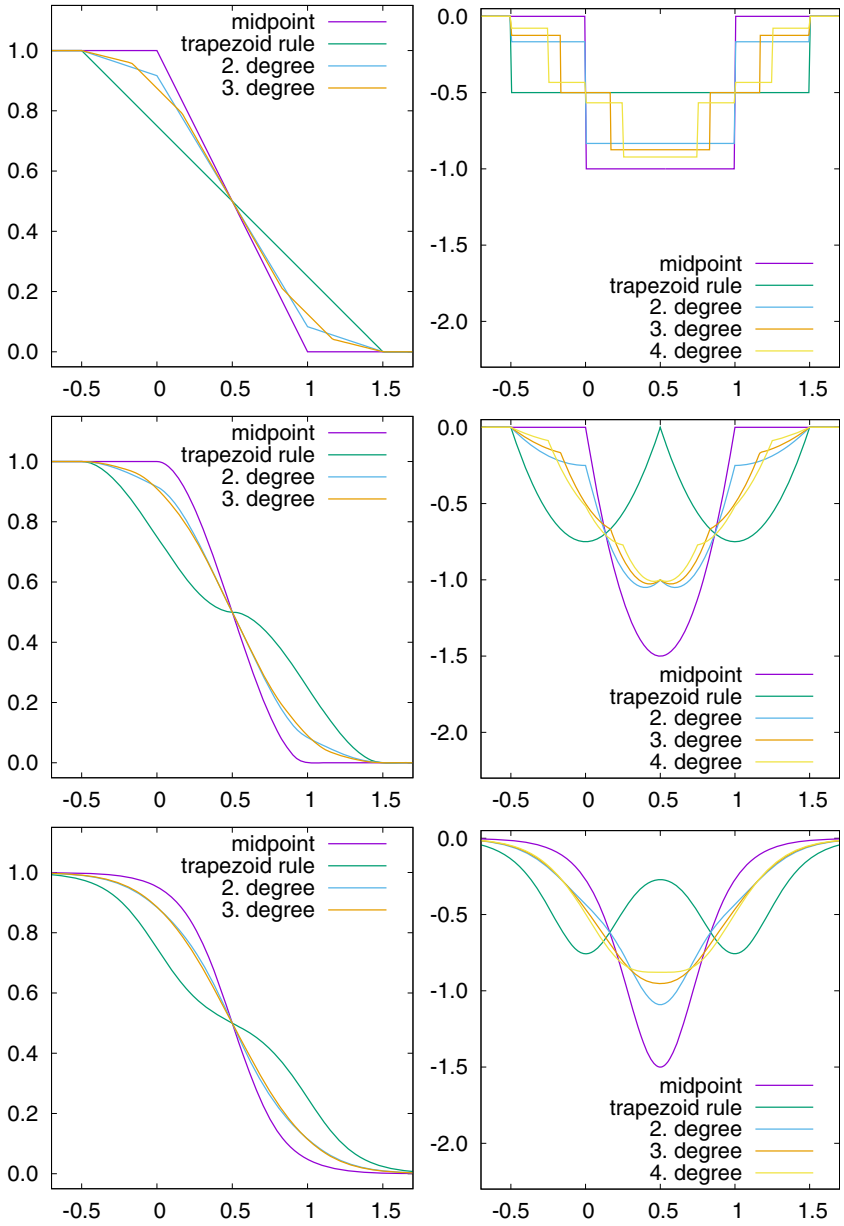


Fig. 11 For the first element $[0:1]$ we perform numerical integration (9) by closed Newton-Cotes in the left column (the y-axis shows the integrated pseudo-density ρ) and numerically evaluate the shape sensitivity in the right columns (the y-axis shows $\frac{\partial \rho}{\partial s}$). The x-axis gives the positional variable s. The boundary modeling functions are linear (first row), polynomial (second row) and tanh (last row)

the boundary with intermediate density ($\rho_{\min} < \rho_e < 1$) multiplied by the element edge length, l_{el}.

For the linear boundary modeling function (top row in Fig. 11) with midpoint integration at $\mathbf{x}_0$, $\rho_e(\mathbf{x}_0)$ varies from 1 to 0 with $w_\rho = w$. With trapezoidal rule, averaging the boundary modeling function values at the left and right node of element 0, $\rho_e(\mathbf{x}_0)$ "sees" the boundary a half element earlier and a half element longer, and the element transition zone is as follows: $w_\rho = w + l_{\text{el}}$.

The polynomial function (18), shown in Fig. 8, has zero slope at the end of its transition zone. Positioning the shape in the center of the element, a variation of the position has low impact when the boundary modeling function is only sampled at the ends of the transition zone by trapezoidal rule (see the center row in Fig. 11). The tanh-like function (15) shows similar behavior, but less pronounced.

The transition zone parameter w for the smoothing function and number of sampling points in numerical integration are correlated. Enlarging the transition zone allows for a lower degree of numerical integration. Generally the density transition zone is $w \leq w_\rho \leq w + l_{\text{el}}$.

In the following, we extend the transition zone w from one element to up to four elements. We use polynomial smoothing with midpoint integration. The left figure in Fig. 12 shows that already a doubled transition zone of two elements results in a significantly more monotonous compliance function over the parameter. However, due to the linear material interpolation (see Fig. 5), the compliance

value becomes artificially good due to a more blurred boundary. The grayness measured by (13) is shown in the center image, revealing again that midpoint integration with transition zone of one element is not sufficient. Note that a wider grayness zone w_ρ allows to include more information from the ersatz material sensitivity. Finally, we apply the RAMP material interpolation, discussed in Section 3.1.2; the right figure in Fig. 12 shows that it helps compensate for the non-smoothness resulting from inaccurate numerical integration.

3.1.7 Computing the signed distance

Feature-mapping methods often employ a signed-distance implicit function when using pseudo-density mapping, as this maintains the transition zone, as discussed in Section 3.1.4.

For some explicit geometry descriptions, the signed distance can be easily computed using an analytical expression. For example, the distance to the edge of a circular, or spherical feature can be directly computed from the feature parameters (center coordinates and radius). Features described by offset surfaces (Norato et al. 2015; Zhang et al. 2016a), whose boundary is defined as the set of all points equidistant to a medial line segment or surface, readily provide direct expressions for the signed distance in terms of the design parameters. Also, for a design feature aligned to one principal axis, the signed distance is easily computed. We use this approach in the test problem in Section 3.1.2, where the design variable is simply the position of the left edge of the vertical feature. Wein and Stingl (2018) use a simplified spline model of horizontal or vertical features, where the distance is obtained in the same way (see Fig. 22(b)), although the distance in their model is not necessary the exact distance to the feature.

To obtain the signed distance for more complex explicit geometry descriptions, a popular method is to compute an equivalent implicit function, which is also a signed distance (i.e., a signed-distance level-set function). This can be achieved using schemes popular with level-set methods, such as the fast-marching method (Adalsteinsson and Sethian 1999), or iteratively solving a Hamilton-Jacobi equation in pseudo-time t

$$\frac{\partial d(\mathbf{x})}{\partial t} + \text{sign}(\phi(\mathbf{x}))\,(||\nabla d(\mathbf{x})|| - 1) = 0. \tag{26}$$

The alternative is to directly compute the shortest distance from a point to the boundary of the feature, as done by Norato (2018) when using explicit features defined by supershapes.

These methods can also be used to compute the signed distance for other implicit geometry descriptions. However, these methods can be computationally expensive, especially if required each time the design changes. Thus, for implicit geometry representations, Zhou et al. (2016) propose using a first-order Taylor approximation of the signed-distance function in the form of

$$d(\mathbf{x}) \approx \frac{\phi(\mathbf{x})}{||\nabla\phi(\mathbf{x})||}. \tag{27}$$

3.2 XFEM approaches

The alternative to pseudo-density mapping is to use an immersed-boundary method. The main challenge of mapping geometry onto a fixed-grid is that the boundary does not align with the fixed-grid elements. Immersed-boundary methods resolve this by introducing extra terms that model discontinuities within elements, while preserving the sharpness of the geometric interfaces. The eXtended Finite Element Method (XFEM) is a popular immersed-boundary approach that has been utilized by several feature-mapping methods.

XFEM approaches model discontinuities by adding enrichment functions and additional degrees of freedom to nodes around the discontinuity. It was originally developed to model crack propagation without re-meshing (Belytschko and Black 1999; Moës et al. 1999). XFEM can also model discontinuities between different materials, or material and void, within an element. Thus, XFEM can be used to model the material discontinuities created by mapping features onto a fixed-grid. The literature on XFEM is vast (Belytschko et al. 2009; Yazid et al. 2009) and an in-depth review is not the focus of this paper. Instead we focus on relevant methods and issues encountered when using XFEM for feature-mapping methods.

There are two types of discontinuity that are considered in feature-mapping methods: material-void (or a strong discontinuity) and material-material (or a weak discontinuity). In general, three components are required to implement an XFEM scheme for material discontinuity: (1) enrichment strategy, (2) interface conditions, and (3) numerical integration.

3.2.1 The simple scheme

For the strong discontinuity case of a material-void interface, a simple scheme may be used, whereby a Heaviside enrichment is applied to the primary field (e.g., displacement or temperature) within the element

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{n} H(\phi(\mathbf{x}))N_i(\mathbf{x})\mathbf{u}_i, \tag{28}$$

where $\mathbf{u}(\mathbf{x})$ is the physical field at point $\mathbf{x}$ within an element with n nodes and shape functions $N_i(\mathbf{x})$, and $\phi(\mathbf{x})$ and $H(\mathbf{x})$ are implicit and Heaviside functions, as defined in Section 3.1.

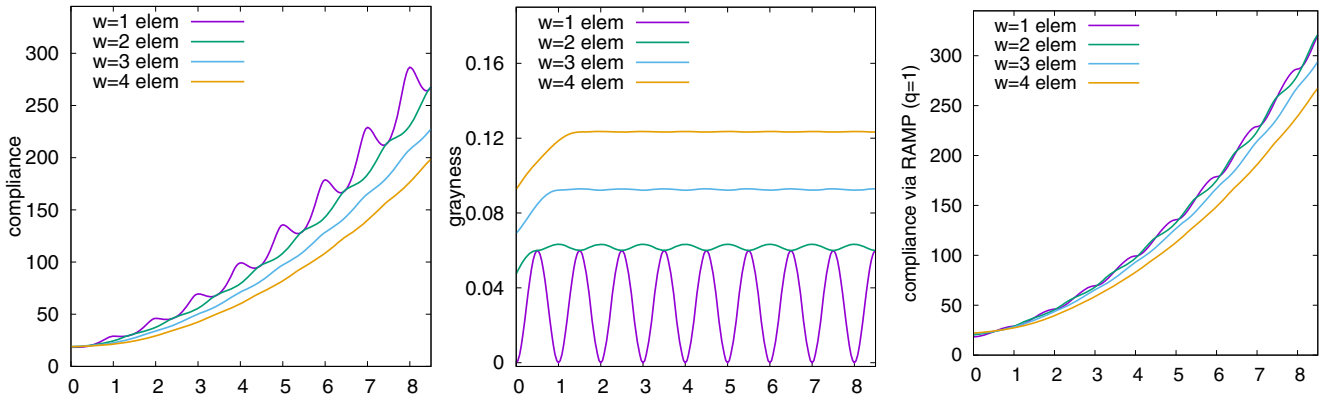


Fig. 12 We vary the polynomial smoothing (center in Fig. 10) by increasing the transition zone $w = 2h$ from one to four elements, each time with midpoint integration. In the right figure RAMP interpolation (12) is applied. On the x-axis, the positional variable s is varied in element units

In this scenario, if the boundary is traction-free then there are no interface conditions and no additional degrees of freedom are required (Villanueva and Maute 2014). This leads to a simple scheme, where element matrices are computed by numerical integration over the material domain. This is usually achieved by automatically subdividing the material domain into triangular sub-cells (e.g., using Delaunay triangulation) and using quadrature rules over each sub-cell. However, integration schemes without quadrature sub-cells have also been used (Li et al. 2012).

3.2.2 Numerical aspects

The simple scheme for strong discontinuities has been utilized in several feature-mapping methods (Li et al. 2012; Zhou and Wang 2013; Liu et al. 2014), its appeal being simplicity of implementation and ability to capture the sharp interface at the material-void boundary. However, there are several issues, or pitfalls, that can be encountered when using the simple scheme. These issues are discussed in the following along with potential solutions from the literature. In addition, the weak material-material discontinuity requires a more complex treatment.

During topology optimization, situations could occur where the design contains a material "island," completely surrounded by void material and disconnected from the main structure. This causes the global system matrix to become singular, leading to numerical problems in solving the discretized governing equations. This can occur in feature-mapping methods if a solid component is mapped onto the fixed-grid, but does not overlap any other part of the solid region. A common remedy is to fill the void region with a fictitious weak material, which has properties several orders of magnitude lower than the real solid material (Wei et al. 2010). If the fictitious material is sufficiently weak, then the simple Heaviside enrichment scheme can still be used, as the error in ignoring the interface condition is small (Wei et al. 2010). An alternative was proposed by Makhija and Maute (2014), where each node is attached to a fictitious point in space by a soft spring. The advantages of this approach are that elements completely in the void phase are not assembled into the global matrix, reducing computational effort, and it avoids spurious load transfer through the void regions.

The simple scheme is only valid if the smallest geometric detail is larger than 2 elements (Villanueva and Maute 2014). However, situations may occur during optimization when this is not true, potentially leading to interpolation error of the geometry and non-physical coupling between disconnected material phases (when the width of void feature is smaller than an element—see Fig. 13). This issue was demonstrated by Makhija and Maute (2014) using a 1D bar example, where a non-zero reaction force was obtained when a gap in the bar was less than the element edge length. To address this issue, Makhija and Maute (2014) proposed a generalized Heaviside enrichment strategy, based on the work of Hansbo and Hansbo (2004) and Terada et al. (2003), which captures the physical discontinuity by adding enrichment functions and additional degrees of freedom, depending on the order of discontinuity around a node.

It should be noted that the issue of non-physical coupling within an element is more common in free-form topology optimization methods, compared with feature-mapping

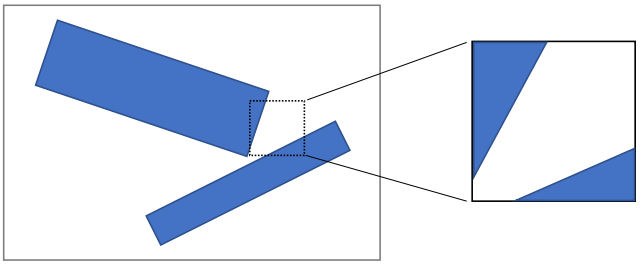


Fig. 13 Potential non-physical coupling in XFEM

methods, as these methods have some high-level control of the geometry. However, non-physical coupling can still occur in feature-mapping methods if solid components are close, such that the gap between them is less than the size of an element, as shown in Fig. 13. This type of non-physical coupling can also occur in pseudo-density mapping methods. However, a similar treatment has not been developed, possibly because pseudo-density methods do not aim to create a sharp interface in the analysis and therefore, this numerical artifact is not seen as an issue.

Modeling the weak discontinuity created by a material-material interface requires a more advanced XFEM scheme. Several authors have used an enrichment function proposed by Moës et al. (2003), which is a C^0-continuous enrichment function that inherently satisfies continuity in the primal solution (e.g., displacement, temperature) at an interface with a weak discontinuity in material properties. Alternatively, the generalized Heaviside enrichment strategy proposed by Makhija and Maute (2014) can be used to model both strong and weak discontinuities. However, an additional constraint is required to enforce continuity across the interface for a weak discontinuity. This can be achieved using a scheme such as the stabilized Lagrange multiplier method, or Nitsche's method.

A further issue that affects both simple and more advanced XFEM schemes is the ill-conditioning of global system matrices due to very small integration regions, compared with the element size. This leads to convergence issues for nonlinear problems and iterative linear solvers (Lang et al. 2014). A standard solution is to use a preconditioner to improve the condition number. For example, Lang et al. (2014) introduced a simple and efficient geometric preconditioner for the generalized Heaviside enrichment scheme. It only requires knowledge of nodal basis functions and the interface geometry, so it can be computed before assembling the system matrix. This method proved effective at reducing the condition number, while maintaining accuracy.

Small integration regions can also affect the accuracy of results at the interface. For example, Van Miegroet and Duysinx (2007) showed large errors in stress when the integration region of an element on the solid-void boundary was small. They discussed several possible remedies, including: removing elements with small solid parts, moving the closest mesh node or moving the boundary to eliminate the small area, post-processing to remove stresses in elements with small solid areas, or computing stresses using a smoothing scheme.

Finally, Sharma et al. (2017) showed that the generalized Heaviside enrichment strategy proposed by Makhija and Maute (2014) produces a smooth, non-oscillatory response function as the design changes. Thus, XFEM methods do not appear to produce a non-smooth response function when modeling the boundary using a sharp step function, in contrast to some of the pseudo-density material interpolation schemes (as shown in Section 3.1.1). However, it was also shown that the shape sensitivity for the XFEM scheme can be oscillatory and that oscillations decreased with mesh refinement. Thus, it was concluded that the oscillations were mainly caused by the accuracy of mapping the geometry to the fixed-grid elements for numerical integration. Note that pseudo-density schemes can also produce a smooth, non-oscillatory response function, if implemented correctly (see discussion in Section 3.1).

3.2.3 Sensitivity analysis

Sensitivity analysis for XFEM is generally more difficult compared with the pseudo-density approach. This is mainly because XFEM uses a more complex procedure to compute element matrices that involves numerical integration over material sub-domains. In the literature there are three main approaches to computing sensitivities when using XFEM in optimization.

The first approach is to differentiate, then discretize. This avoids computing derivatives of the change in integration regions as the interface moves, as sensitivities are derived from the continuum equations. A common example of this approach is to use shape sensitivities (Zhou and Wang 2013; Liu et al. 2014; Wang et al. 2014b). However, it is well-known that convergence issues may occur due to the discretization error and often some form of regularization or smoothing is required (van Dijk et al. 2013; Liu et al. 2014).

The second approach is to discretize, then differentiate, but with a semi-analytical approach. The idea is to compute the derivative of the element matrices with respect to the design variables using the finite difference method. This derivative term is then inserted into the analytically derived sensitivity formula. Thus, sensitivities are consistent with the numerical discretization, but the semi-analytical approach avoids explicitly computing derivatives with respect to changes in the integration sub-domains. The finite difference approach is reasonably efficient, as it is only performed for elements that contain an interface and does not require assembling and solving a system of equations. This approach has proved effective and has been used in several feature-mapping methods (Van Miegroet and Duysinx 2007; Sharma et al. 2017).

However, the finite difference scheme should ensure that the design variable perturbation does not cause a change in element status, e.g., an element containing a material-void interface does not become either fully void, or fully solid. This causes problems in the derivative computation as it changes the number of degrees of freedom (Zhang et al. 2012; Noël et al. 2016). Several methods have been proposed to avoid this issue. One method is to perform both forward and backward finite differences and check if

either cause a status change. If neither cause a change, then the central difference is used. However, if the forward (or backward) difference causes a status change then only the backward (or forward) difference is used. Another method is to perturb the interface such that the finite difference perturbation cannot cause a status change (Sharma et al. 2017). Alternatively, the finite difference perturbation step can be reduced to a magnitude that avoids a status change, although if the step magnitude is too small, numerical round-off errors can occur.

The third approach is to discretize first and then differentiate using a full analytical approach, without finite differencing. The challenge is to compute the analytical derivative for the change in the integration sub-domains with respect to the design variables. Zhang et al. (2012) developed an analytical derivative for a material-material interface, when the geometry is represented by nodal implicit function values. Noël et al. (2016) and Najafi et al. (2015) proposed schemes utilizing a velocity field to efficiently compute the analytical derivatives. These fully analytical schemes are more complex and difficult to implement than the semi-analytical scheme, but are more efficient and avoid the status change issue.

4 Combination of features

The foregoing section describes the approaches that existing techniques use to map individual geometric features onto the fixed analysis mesh. To be able to modify the topology of the structure, it is also necessary to combine these features. This is one of the key ingredients of performing topology optimization with high-level geometric features, and has received considerable attention in recent years. In this section we specifically consider the combination of closed regular sets (solids or holes). Unless otherwise stated and for brevity, whenever we refer to combination of solids we also refer to combination of holes.

The combination of solids in all approaches corresponds in effect to Boolean operations between solids. Just like other aspects in this review, it is possible to categorize approaches that combine solids in different ways. The main criterion we use to categorize combination methods is whether the combination occurs before or after mapping to the fixed analysis mesh.

4.1 Smooth combination functions

Many feature-mapping methods utilize smooth combination functions so that derivatives with respect to the high-level geometric parameters are continuous. For example, the non-differentiable Boolean union of multiple solids represented by implicit functions f_i corresponds to their maximum, i.e., $\bigcup_i f_i = \max_i(f_i(\mathbf{x}))$ (Shapiro 2002). Common choices for differentiable smooth approximations $\widetilde{\max}_i(f_i(\mathbf{x}))$ include the well-known Kreisselmeier-Steinhauser (K-S) function

$$\widetilde{\max_i}^{KS}(f_i) := \frac{1}{p}\ln\sum_i^N(\exp(pf_i)) \tag{29}$$

and the p-norm

$$\widetilde{\max_i}^{p}(f_i) := \left(\sum_i^N f_i^p\right)^{\frac{1}{p}}, \tag{30}$$

where N is the number of values and p is a parameter that controls the sharpness and accuracy of the approximation (a larger p results in a more accurate estimate of the true maximum and a sharper function). The work by Coniglio et al. (2019) makes a comparison of several smooth maximum functions for the combination of geometric features.

Another type of functions used to perform Boolean operations is R-functions (cf., Shapiro 2007). The R-conjunction corresponds to the logical AND, whereas the R-disjunction corresponds to logical OR. Compositions of these two fundamental R-functions can be used to construct any Boolean expression. There are several forms of R-functions; for example, Chen et al. (2007) use the following definition

$$\begin{aligned} f_1 \cap f_2 &:= f_1 + f_2 - \sqrt{f_1^2 + f_2^2}, \\ f_1 \cup f_2 &:= f_1 + f_2 + \sqrt{f_1^2 + f_2^2}, \end{aligned} \tag{31}$$

which is differentiable everywhere except at $f_1 = f_2 = 0$. It can be seen that $f_1 \cap f_2$ is positive if and only if both f_1 and f_2 are positive, whereas $f_1 \cup f_2$ is positive if either f_1 or f_2 is positive. For example, assume ϕ_1 and ϕ_2 are implicit functions for two solid features and ϕ_v the implicit function of a void feature. The combined implicit function can then be defined as

$$\phi(\mathbf{x}) := (\phi_1(\mathbf{x}) \cup \phi_2(\mathbf{x})) \cap \phi_v(\mathbf{x}). \tag{32}$$

4.2 Combine-then-map approaches

Since a geometric representation of the solids is available that is independent of the analysis mesh, a natural approach is to combine the solids directly using their geometric representation, and then map the combined solid onto the analysis mesh, as described in Section 3.

4.2.1 Implicit geometric representations

As mentioned in the previous section, the Boolean union or intersection of features represented by implicit functions

can be attained by computing their maximum or minimum, respectively. This is a strategy that has been used to combine both solids (Cheng et al. 2006; Zhou and Wang 2013; Guo et al. 2014; Zhang et al. 2016c) and holes (Cheng et al. 2006; Chen et al. 2007; Mei et al. 2008; Wang et al. 2012; Zhang et al. 2017b).

This combination approach is illustrated in Fig. 14, where three rectangular bars are modeled using hyperellipse implicit functions (as used in, e.g., Guo et al. 2014; Zhang et al. 2016c). Figure 14(b)–(d) show contour plots of ϕ_i for each of the three bars. All contour and fringe plots in this section are produced using a grid of 48×48 square elements. The Boolean union of the implicit functions for these three bars, as given by the true maximum function, is shown in Fig. 14(e). Note that smooth maximum functions could also be used. The combined implicit function is subsequently mapped onto the analysis mesh using a pseudo-density or immersed-boundary approach, as discussed in Section 3. A combined implicit function subject to a smooth Heaviside (16) is shown in Fig. 14(f) with the element-constant pseudo-densities shown in Fig. 14(g) (which were obtained by the method used by Zhang et al. 2016c).

This approach is notable for the simplicity of the Boolean operations. The simplicity is inherited from the implicit geometric representation, since in general it is much easier to perform Boolean operations with implicit, rather than with explicit geometric representations. Also, this combination approach is readily extended to 3-dimensional problems (e.g., Liu and Ma 2015; Zhang et al. 2016c).

Topological changes using this combination approach occur in one of three ways: (1) as solids move and overlap, the connectivity of the structure may change (holes may appear or disappear); (2) if a solid is engulfed inside another solid, it has no effect in the analysis due to the maximum operation, and thus the engulfed solid is effectively removed from the design; and (3) if one or more dimensions of a solid become sufficiently small, the effect of the solid on the analysis is negligible.

4.2.2 Explicit geometric representations

When the original geometric representation is explicit, two approaches have been employed to combine solids. The first approach consists of performing the Boolean union directly on the explicit representation, and then converting the resulting design into an implicit geometric representation before mapping to the analysis mesh. This strategy is illustrated in Fig. 15, where the three rectangles of the previous example are modeled using cubic B-splines with eight control points per side (see Fig. 15(a)). A combination technique used in this case (Lee et al. 2007; Seo et al. 2010; Zhang et al. 2017d; Gai et al. 2020) consists of deleting from the current design those control points that lie in the overlapping region between the bars, so that the union of

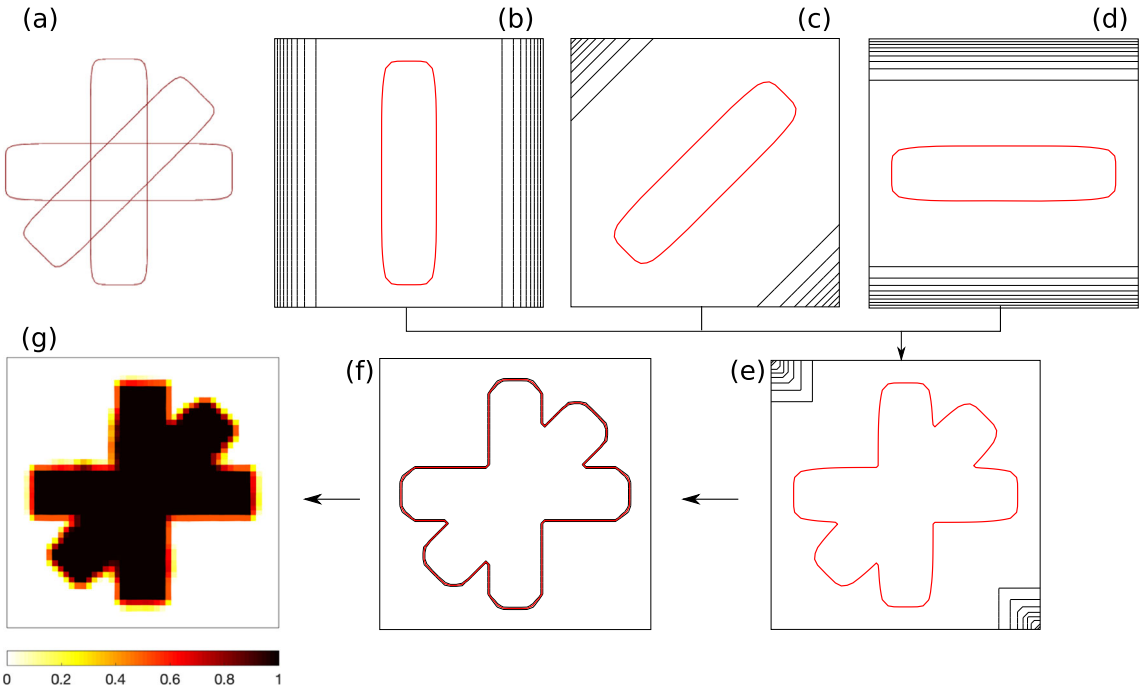


Fig. 14 Combination by union of implicit functions-then-mapping: (a) rectangular bars modeled with hyperellipses; contour plots of (b)–(d) implicit functions ϕ_i, (e) union of implicit functions (Φ) via maximum function, and (f) smoothed Heaviside of Φ; and (g) pseudo-density used for analysis. Thick red line corresponds to zero level set for implicit functions, 0.5 level set for Heaviside (here and in following figures)

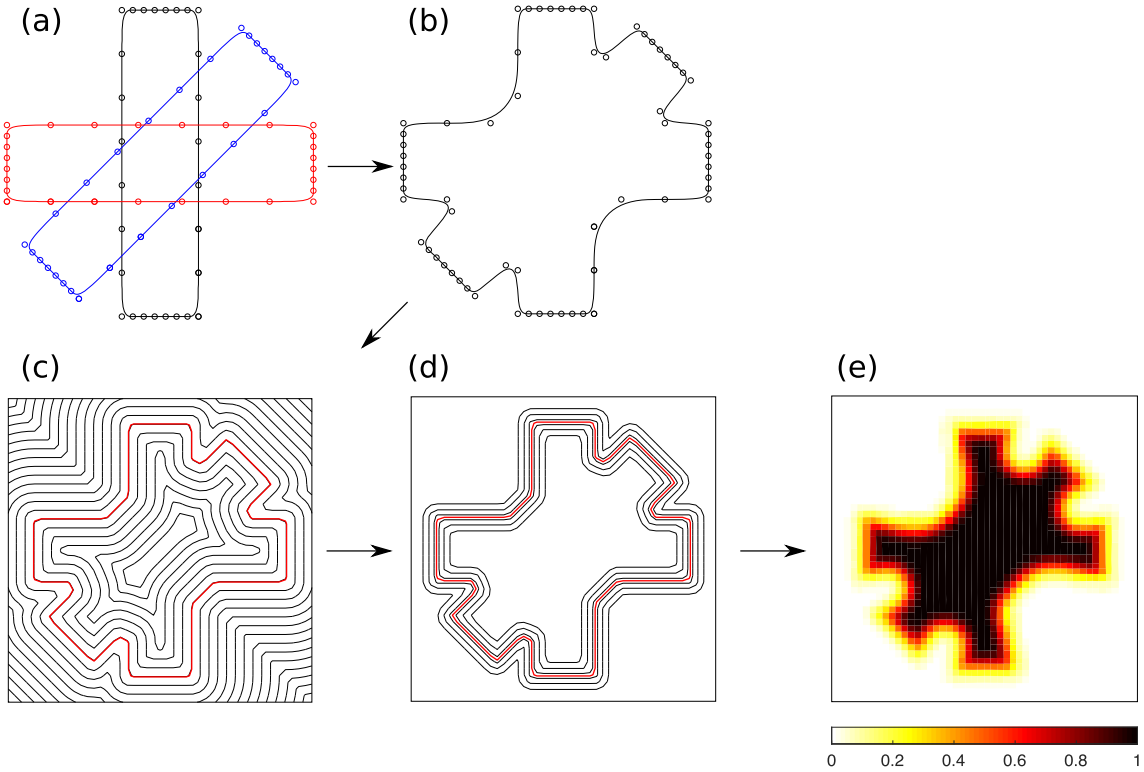


Fig. 15 Combination by union of explicit functions-then-mapping: (a) rectangular bars modeled with cubic B-splines, (b) union by deleting B-spline control points in the overlapping region, (c) signed-distance function of combined B-spline, (d) smoothed Heaviside of signed-distance function, and (e) pseudo-density used for analysis

the primitives is given by a single B-spline made of the remaining control points, as shown in Fig. 15(b). In Gai et al. (2020), cusps arising from merged holes are removed by replacing the merged hole with the curve formed by the trajectory followed by the touching points of a rolling circle. These works all consider B-spline-shaped holes; however, here we consider solid rectangles for consistency with the examples given for the other strategies.

After combining the solids, the explicit representation is transformed to an implicit representation, namely by computing the signed distance to the combined B-spline, as shown in Fig. 15(c). An exact or smooth Heaviside approximation, such as the one presented in the preceding section, is then applied to the signed-distance function, at which point the mapping to the analysis can be completed in the different ways discussed in Section 3, i.e., using pseudo-densities or an immersed-boundary method.

This combination approach presents several challenges. In the particular case of B-splines, it is possible for the control points to be placed such that the B-spline can present self-intersections, which requires placing bounds on the positions of the control points (Lee et al. 2007), or employing special parameterizations of the B-spline (Zhang et al. 2017d). In both cases, it is necessary to determine the correct order of control points in the combined B-spline to generate the correct shape. Another, perhaps more pernicious challenge, is the potential lack of differentiability introduced by the control-point deletion approach. Suppose two solids overlap, and a control point of one of them lies exactly on the boundary of the other. Small positive and negative rotations of any of the primitives will cause that control point to be deleted or retained in the combined B-spline. Therefore, the combined B-spline may look appreciably different in both cases, which means the structural response will not be differentiable with respect to the orientation angle of either solid. To obtain an accurate union of the B-splines it is of course possible to introduce control points at the intersections with multiple knots to capture the sharp corners. However, this introduces additional challenges in the optimization, as the number of design variables (i.e., the positions of the control points) would increase with the additional control points. A third challenge lies in the computational cost incurred in translating the explicit representation to an implicit representation, e.g., the computation of the signed-distance field. Although there exist computational strategies to do this efficiently (see Section 3.1.7), it still adds computational cost compared with directly using implicit

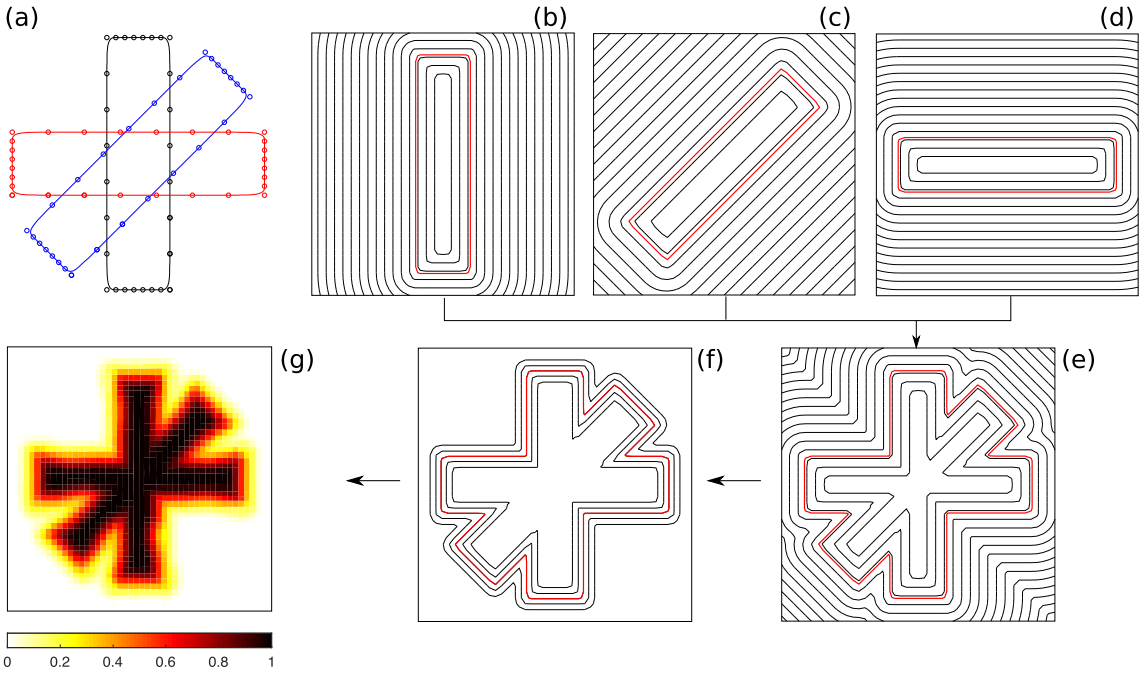


Fig. 16 Combination by conversion of explicit to implicit representation-then-mapping: (a) rectangular bars modeled with cubic B-splines, (b)–(d) signed distance of individual bars, (e) union of implicit signed-distance functions via maximum function, (f) smoothed Heaviside of implicit union, and (g) pseudo-density used for analysis

geometric representations. Finally, while it is possible to perform Boolean operations of explicit representations of 3-dimensional solids, the aforementioned challenges are more difficult to solve for 3-dimensional problems.

The second strategy to combine features with explicit representations is to first convert the explicit representation of each solid to an implicit representation, and then perform the combination of the individual implicit representations as in the previous section (Zhang et al. 2017e). This strategy is depicted in Fig. 16, where each B-spline is first converted to an implicit representation (a signed-distance function), and then the combination of bars is achieved via the true maximum—in the present case—of the implicit functions. This approach circumvents the problems arising from the deletion of control points and greatly facilitates the combination of primitives. However, there is additional computational cost, as a separate signed-distance function must be computed for each solid. This strategy is arguably similar to the map-then-combine approach described in the next section, because in this method the individual implicit functions are computed on the fixed-grid prior to the combination.

4.3 Map-then-combine approaches

The alternative to the combination strategies described in the previous section, is to first map each individual solid to the analysis mesh and then combine the ensuing mapped variables, such as pseudo-densities, Heaviside function values, or even material property values. The combination could be done element-wise (e.g., element-constant pseudo-densities), or at integration points (e.g., when using (9)).

All existing map-then-combine approaches require an implicit geometry description to achieve the mapping. Thus, map-then-combine approaches that start with an explicit geometry description first convert each feature to an implicit function before mapping the individual implicit functions to the analysis grid. This is the same as the first step in Fig. 16(a–d), where the geometry of each solid is described by a B-spline, which are then converted to implicit signed-distance functions. Therefore, the remainder of this section describes map-then-combine methods starting from an implicit geometry description.

4.3.1 Property interpolation for hybrid approaches

Some hybrid approaches described in Section 7 combine a free pseudo-density field (as in density-based methods) with features using an extended material interpolation function, which interpolates between the solid-void pseudo-density field and solid features, e.g., Qian and Ananthasuresh (2004) and Wang et al. (2014b), or holes, e.g., Kang and Wang (2013). For example, if E_s is the Young's modulus of the solid phase of the free pseudo-density field ρ, and E_c modulus of the embedded solid component (assuming isotropic materials), an interpolation of the

Young's modulus may be given by (cf., Wang et al. 2014b)

$$E(\mathbf{x}) = \widetilde{H}(\phi(\mathbf{x}))E_c + \left[1 - \widetilde{H}(\phi(\mathbf{x}))\right]\left[\rho^p(\mathbf{x})\right]E_s, \tag{33}$$

where ϕ is an implicit representation of the embedded solid primitive, $\widetilde{H}$ is a smooth approximation of the Heaviside function and p is the SIMP penalization power. Note that in (33), $\mathbf{x}$ is usually the element center, which is equivalent to mid-point integration (see Section 3.1.6). Further approaches are discussed in Section 7.1.

4.3.2 Combining Heaviside functions

Combination of multiple features can be achieved by combining Heaviside function values of each mapped feature using a maximum function. The single, combined Heaviside function can then be used to compute element pseudo-densities, or in immersed-boundary methods (as discussed in Section 3). This approach was used by Wein and Stingl (2018) to obtain element pseudo-densities by numerical integration via

$$\rho_e = \sum_j^{N_{\text{ip}}} w_j \, \widetilde{\max_i} \, \widetilde{H}_i(\mathbf{x}), \tag{34}$$

where the combination is done at the integration points. In (34) a smooth maximum and smooth Heaviside are used, thus making sensitivity analysis straightforward.

It is interesting to note that in the case of midpoint integration and using (9), (34) simplifies to

$$\rho_e = \widetilde{\max_i} \, \rho_e^i, \tag{35}$$

which is effectively the same as combining mapped element pseudo-density values (see Section 4.3.3). This highlights the close connection between the Heaviside function and pseudo-densities in feature-mapping.

4.3.3 Combining pseudo-density values

Another map-then-combine approach is to compute element pseudo-density values with respect to each feature, as described in Section 3.1, and then perform the combination using a true or smooth maximum function. This approach also allows for an additional control of the combined features, by introducing variables that penalize the mapped densities of each solid feature separately. These are called size variables, which are penalized in the spirit of SIMP, so that a zero value indicates the solid has no effect on the analysis and thus can be removed from the design, whereas a value of unity indicates the solid must be retained (Norato et al. 2015; Zhang et al. 2016a).

Without consideration for the aforementioned penalized size variable (that is, assuming all bars have a size variable of unity), the pseudo-density map-then-combine strategy

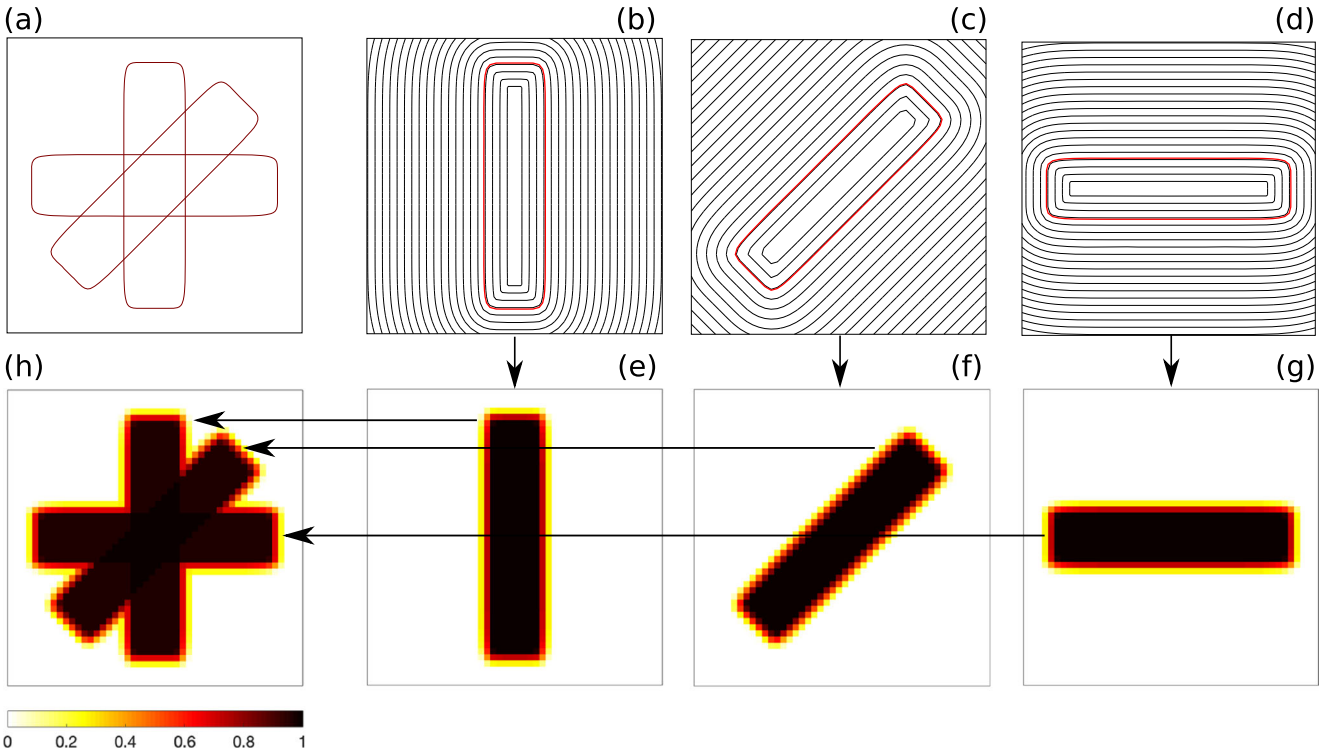


Fig. 17 Map-then-combine by union of implicit functions: (a) bars modeled with hyperellipses, (b)–(d) signed-distance field for each bar, (e)–(g) corresponding mapped pseudo-density for each bar, and (h) union of mapped densities using a smooth approximation of the maximum

is depicted in Fig. 17. Figure 17(a) shows the three bars modeled as hyperellipses as before; Fig. 17(b)–(d) are the signed-distance fields corresponding to these surfaces; Fig. 17(e)–(g) show the mapped pseudo-densities for each bar, computed at each element of the mesh using (20); and Fig. 17(h) shows the Boolean union of bars, obtained using a smooth approximation of the maximum function (here the p-norm (30)).

The combination of solids when a penalized size variable is used is as follows. First, element pseudo-densities are computed. The effective density at element e for bar i is subsequently computed as

$$\widehat{\rho}_e^i = (\alpha^i)^p \rho_e^i, \tag{36}$$

where $\alpha^{(i)}$ is the size variable corresponding to solid i and p is the penalization power. We note that if $\alpha^{(i)} = 0$, then the effective density at element e for bar i is zero, hence this solid has no effect on the material properties at element e; this is true for every element for which $\rho_e \neq 0$, and so making the size variable zero effectively removes the solid from the design. The combination of the solids is subsequently obtained via, for example, a smooth maximum of the effective densities as

$$\rho_e = \widetilde{\max_i} \widehat{\rho}_e^i. \tag{37}$$

Figure 18 shows the combined density after the union of all three bars when the diagonal bar has different values of its size variable α (while the other two bars have a size variable of unity). Clearly, as its size variable nears zero, the effect of the diagonal bar on the combined density vanishes.

Topological changes using this approach can occur in different ways. As in the methods of Section 4.2, when solids "move" in the optimization, holes can be created or disappear. We also note that some approaches (e.g., Norato et al. 2015; Zhang et al. 2016a) have the limitation that a solid cannot be removed by collapsing its dimensions, because for the sensitivities to be always well defined, it is necessary that the size of the sample window used to compute the volume fraction is smaller than the dimensions of the solid. The other removal mechanism, as aforementioned, is to make the size variable of the solid zero.

4.4 Local minima

The combination of geometric features may lead to unfavorable local minima. To illustrate this, we consider the example shown in Fig. 19. Four bars are modeled with hyperellipses. Three of the bars are fixed, and another one is rotated about the left edge midpoint by changing h. For $h \in \{0, L/2, L\}$ the moving bar entirely overlaps with one of the fixed bars. The design region is meshed with square bilinear elements with a relatively fine mesh. A binary pseudo-density mapping is used, where the element pseudo-density is either $\rho_{\min}$ or 1 depending on whether the element centroid is outside or inside of a bar, respectively. The combination is performed using a map-then-combine approach with a true maximum of the pseudo-density values.

Figure 19 shows the compliance as a function of h/L. The actual magnitude of the compliance is not important; what is important is the presence of two distinct local minima, one of which ($h/L \approx 0.43$) is clearly worse than the other ($h/L \approx 0.79$). Therefore, if a gradient-based optimizer is used and the initial design has $h < L/2$, the optimizer will most likely converge to the poor local minimum. Thus, the more compact design representation used by feature-mapping methods (as opposed to the verbose representation used by density-based and level-set methods) is more prone to falling into unfavorable local minima depending on the initial design. Although all topology optimization techniques are dependent on the initial design (e.g., as shown by Yan et al. (2018) for density-based topology optimization of heat conduction structures), this dependency is more pronounced in feature-mapping techniques, as noted in Norato et al. (2015) and Zhang et al. (2016a). We note that this has nothing to do with the particular feature-mapping technique used, but with the more restrictive geometric representation.

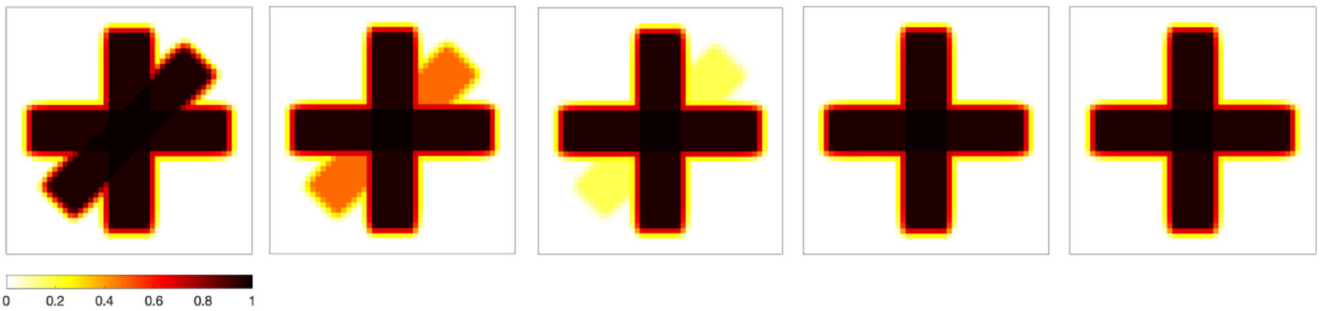
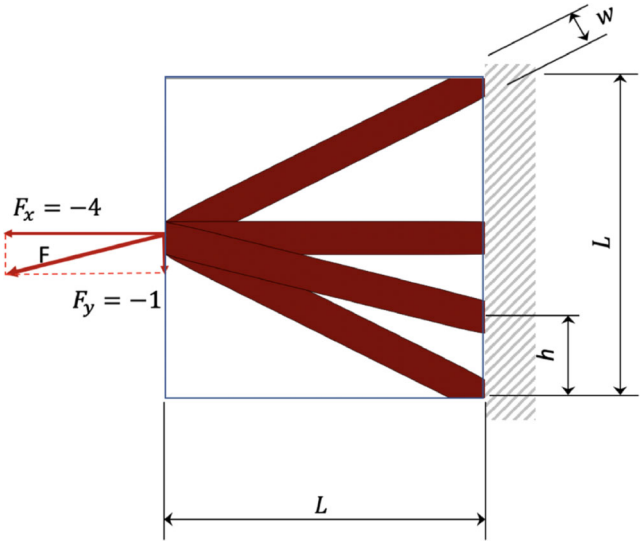
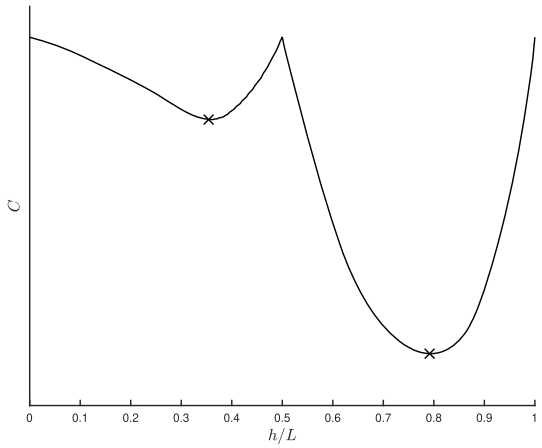


Fig. 18 Three-bar example, where the diagonal bar has a size variable of $\alpha = 1.0, 0.8, 0.6, 0.1$ and 0, from left to right. A power value of $p = 3$ is used in (36) in all cases



(a) Setup



(b) Compliance

Fig. 19 Example of effect of combination on local minima. Top: setup. Bottom: compliance. The crosses mark the locations of the minima

5 Separation constraints

Separation constraints are high-level geometric constraints that specify a minimum distance between solid components (or holes). When the minimum distance is set to zero, separation constraints are often called non-overlap constraints, as component overlap is prevented. These constraints can also be used to prevent components from leaving the design domain. Several techniques have been proposed to enforce this type of constraint in feature-mapping methods.

The earliest method is the finite circle method (FCM). The main idea is to approximate the shape of each component with a number of circles (Fig. 20). Separation constraints can then be posed as simple geometric constraints on the minimum distance between circle centers. Qian and Ananthasuresh (2004) used a single circle for each component and Zhang and Zhu (2006) extended the idea to use multiple circles to approximate the shape of each component.

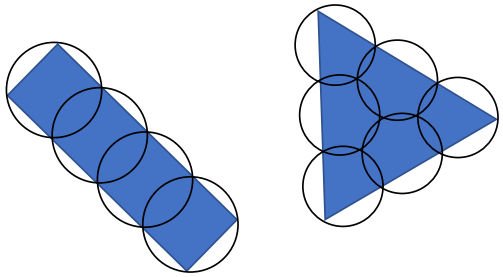


Fig. 20 Finite circle method

The main benefits of FCM are the simple definition of the constraints, which are continuous and differentiable. However, if there is a large number of components, then a large number of separation constraints is required, although most are usually inactive at the optimum. For example, even if only one circle is used to approximate each component, then $N(N-1)/2$ constraints are required for N components. Also, component shapes are only approximated by circles, so separation constraints may not be able to reach their lower bound in some situations (due to circles covering a larger volume than the actual component). If high accuracy in the separation constraint is required, then more circles can be used for each component, which adds more constraints. A large number of constraints may affect the efficiency of the optimization (Zhang et al. 2011), although the number of constraints can be reduced by using different sized circles. Xia et al. (2012b) also showed that to prevent components leaving the design domain using FCM, only a small circle at each corner of the component is required.

To use FCM more efficiently with a large number of components and constraints, Gao et al. (2015) used constraint aggregation to combine all finite circle separation constraints into a single constraint function. An adaptive Kreisselmeier-Steinhauser (K-S) function is used, where the aggregation parameter is automatically determined to ensure the accuracy of the aggregation function. Another approach was proposed by Zhu et al. (2017) where the finite circle separation constraints are added to the objective using a combination of exterior and interior penalty methods.

A further limitation of the standard FCM is that it does not automatically adapt to components that are changing in size or shape. This is a challenging problem if the constraints are to remain continuous and differentiable throughout the optimization. However, Zhang et al. (2012) showed how this can be achieved for elliptically shaped components by linking the location and radius of each circle to the parameters of the elliptical shape.

An alternative to the FCM, which is suitable for feature-mapping methods where solid and/or void features are represented by an implicit function, was first proposed in the master's thesis of Shan (2008) and later refined by Kang and Wang (2013). The idea is to compare the integral of the solid region represented by the combination of implicit functions with the known volume of solid components. (The same idea also applies to void components.) If the integral is less than the known volume, then there must be some overlap of components, or part of a component has left the design domain (Fig. 21). This observation is used to formulate a single, differentiable constraint that prevents overlap for arbitrary shaped components and also prevents components leaving the design domain.

The integral method has also been extended for minimum distance separation constraints (Zhang et al. 2015). To achieve this, each component is represented by a signed-distance implicit function. The signed-distance information is used to construct "virtual" components, whose boundaries are offset by half the minimum distance constraint value. The original separation constraint formulation is then applied to the "virtual" components to ensure a minimum distance between components.

Zhang et al. (2015) introduced an approach using the structural skeleton, which is also suitable for methods where features are represented by implicit functions. The structural skeleton is defined as the set of interior points that have at least two closest boundary points (this is also called the medial axis). To enforce a minimum distance separation constraint between two components, a signed-distance function is constructed for the combined implicit function of both components. A skeleton is then constructed to identify all points that are equidistant from both components. Finally, an explicit constraint is imposed on the minimum signed-distance value of all points belonging to the skeleton. This method can also be used to set maximum separation constraints. However, it requires construction of the signed-distance functions in each iteration and the number of constraints is linked to the number of components. Furthermore, the formulation is not differentiable and derivatives are approximated using finite differences and by assuming that the skeleton does not change when components move.

For methods based on pseudo-densities, Zhang and Norato (2017) proposed a simple method using a map-then-combine approach. First the pseudo-densities for each component are mapped onto the fixed-grid. These are then combined using simple summation to create an auxiliary pseudo-density field. If any density value in the auxiliary field is greater than 1, then there must be some component overlap. Thus, an aggregated constraint function can be defined that ensures the maximum value of the auxiliary pseudo-density field is unity and hence prevent component overlap. This idea can easily be extended to provide a minimum separation constraint, by uniformly enlarging the size of the components by half the minimum separation distance before computing the auxiliary pseudo-density field.

The FCM- and integral-type methods for preventing component overlap can also be used to prevent components leaving the design domain. The integral approach achieves this without any modification to the original method, as it automatically detects when any part of a component has left the design domain. The FCM approach needs additional distance constraints to prevent components leaving the domain. This is straightforward for convex design domains (Zhu et al. 2008), but non-convex domains present difficulties in defining continuous differentiable distance constraints. An approach for pseudo-density methods is to use a layer of ghost points that lie a short distance outside the design domain (Zhang et al. 2018a). The pseudo-densities at ghost points are then computed and if any value is non-zero, then a component must be outside the domain. This idea is used to create an aggregated constraint function that ensures the maximum pseudo-density value at all ghost points is zero. This approach can be used for both convex and non-convex design domains without modification.

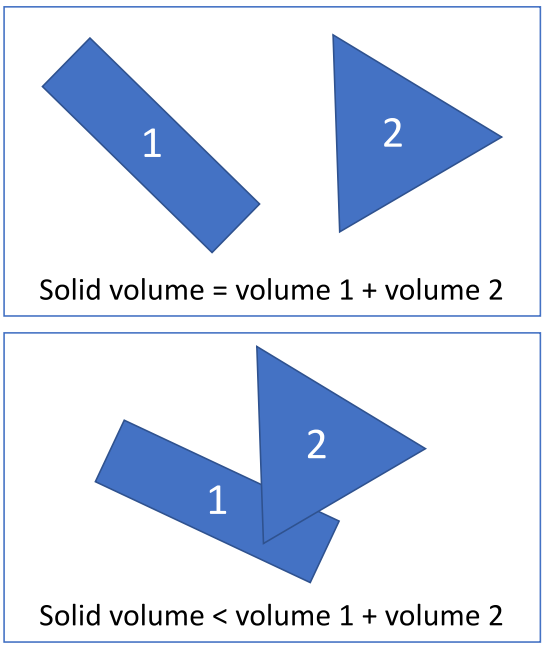


Fig. 21 Integral method for preventing component overlap

6 Feature-mapping methods for shape optimization

In this section we discuss the application of feature-mapping to solve shape optimization problems, which is essentially a

classical shape optimization approach with the design mapped to a fixed-grid. To this end we start with a brief description of what we consider as *classical shape optimization*.

6.1 Classical shape optimization

In classical shape optimization, only the structural part is discretized using finite elements. The structural interface is exactly modeled, which is for some applications an essential feature. Void regions are not part of the finite element analysis, which can significantly reduce the computational effort.

As a consequence, the boundary mesh nodes are moved during optimization. To maintain mesh quality for accurate analysis, mesh smoothing and/or re-meshing is necessary. However, re-meshing can become rather involved and if the quality of the finite element approximation is insufficient, there is the potential for the mesh to be optimized for numerical artifacts, similar to the checkerboard effect in density-based topology optimization.

Classical shape optimization has been successfully applied over several decades. However, the mathematical and technical realization is rather involved, especially compared with density-based topology optimization. The mathematical approach is often formulated in an infinite dimensional setting (see e.g., Sokolowski and Zolesio 1992 or Haslinger and Mäkinen 2003) and differentiable mesh generation must be provided (Haslinger and Mäkinen 2003). The gradient information is based on the shape gradient.

Shape optimization is performed with a wide range of different parameterizations. These can be categorized as either boundary-node-based parameterization, or higher order forms of design parameterization. We begin with the first variant where each surface node is a design variable. This is called the independent node movement approach (Imam 1982) or parameter free shape optimization. This provides a large space of admissible shapes, but it comes with its own challenges in terms of regularization and feature size control (see, e.g., Le et al. 2011). Closing and creation of holes are generally difficult to achieve, or even impossible. During the optimization process, insertion or deletion of boundary nodes may be necessary, as well as re-meshing. This generally prevents the use of first-order mathematical programming algorithms. As a consequence, constraint functions need to be handled indirectly. Furthermore, no rigorous convergence criteria are available.

Aside from the parameterization of boundary nodes there are many variants of higher order parameterizations established in shape optimization (see Haftka and Grandhi 1986 for an early survey). Conveniently, this corresponds to the construction of geometries by spline functions in computer-aided design (CAD). Here, the mapping from the design parameters onto the boundary nodes is differentiable and thus allows gradient-based optimization (see Braibant and Fleury 1984). The separate meshing of the geometry can be alternatively handled by isogeometric shape optimization (see Wall et al. 2008). Provided a differentiable parameter-to-boundary mapping, the shape sensitivity of an arbitrary parameterization can be obtained from the nodal shape gradient by the chain rule.

6.2 Using pseudo-density feature-mapping

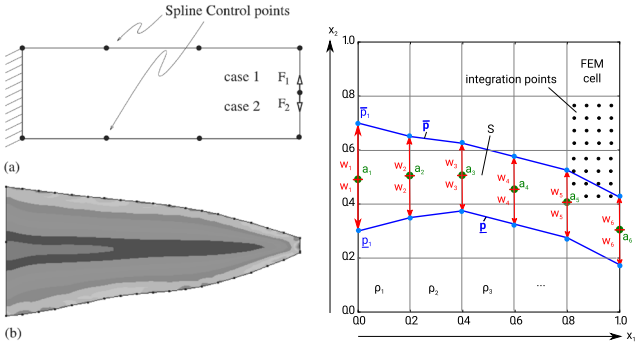
Some major advantages of shape optimization are the crisp boundary description and the wide and versatile range of design parameterizations. However, the modeling and technical realization is often quite involved. Density-based topology optimization is known for its elegant and easy modeling, standard approaches for sensitivity analysis and straightforward technical realization, e.g., no re-meshing is necessary. On the other side, the extremely rich design space allows only an implicit design description and is, for some applications, difficult to control. Enforced by standard regularization approaches, one has to deal with a more or less blurred interface description which is anyway rasterized by the fixed analysis mesh.

It is now a natural approach to combine features of both worlds, density-based topology optimization and parameterized shape optimization. This can be seen from two points of view. One can perform standard parameterized shape optimization, but skip the re-meshing by using a fixed analysis grid, leading to an inexact boundary modeling. The same picture is given from the point of view of density-based topology optimization, where a higher level design parameterization is mapped to the pseudo-density field, allowing much closer control of the design and, as such, intentionally sacrificing the rich design space of density-based topology optimization. Under certain circumstances, this approach can be seen as a simplified shape optimization where, aside from the design to density mapping, the technical implementation and sensitivity follows that of density-based topology optimization.

In principle, any higher order classical shape parameterization can be combined with feature-mapping, by simply mapping the boundary transition to a fixed-grid using pseudo-densities (see Section 3.1).

In Garcia and Gonzalez (2004) B-splines are optimized with a small number of control points (see Fig. 22). The design parameters are the location of the control points. The optimization in Garcia and Gonzalez (2004) is performed without sensitivity analysis using a gradient-free method.

The first rigorous high-level geometry-to-density mapping was demonstrated in Norato et al. (2004), where boundary smoothing is discussed and the convergence of the rasterization of an explicitly given primitive is shown theoretically and numerically. Optimization results based on



(a) B-spline (b) first order spline

Fig. 22 (a) B-spline parameterization in Garcia and Gonzalez (2004); (b) piecewise linear spline with mesh aligned control points in Wein and Stingl (2018)

implicit shapes (via radial basis functions) are also shown, but the approach is not discussed in detail.

In Wein and Stingl (2018), splines of order 1 (piecewise linear functions) are used, where the control points are aligned to the analysis mesh. The design space is rather restricted, as either the y- or x-component of the control points are design variables for horizontal or vertical structures, respectively. Additionally, the width is parameterized by thickness variables. This facilitates, but also requires, the control of first and second spatial derivatives of the design variables in the form of slope constraints and curvature constraints, which are both directly defined from the design variables. Also, as the strips are primarily aligned with one of the axes (see Fig. 22), the computation of an approximate signed-distance function is easily obtained and is differentiable with respect to the control points.

An interesting variant of the first-order spline representation is used in Kasolis et al. (2012). Here the angle ϑ_i / l between line segments of length l expressing the boundary Γ_d is the design variable (see Fig. 23). The application is the optimization of an acoustic horn. The line segments are larger than the mesh element size and no regularization is necessary. The structure is of constant thickness and fixed on the left side. A change in the angle ϑ_i leads to a rigid body movement of all line segments to the right. The design to pseudo-density mapping is differentiable.

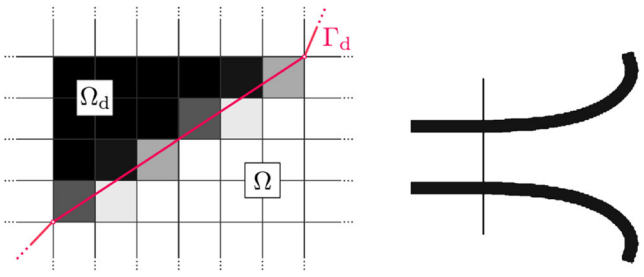


Fig. 23 Optimization of the angle between line segments in Kasolis et al. (2012). The right figure shows an optimized acoustic horn in 2D

6.3 Using immersed-boundary feature-mapping

Combining immersed-boundary feature-mapping with higher order classical shape parameterization has also been demonstrated. Again, in principle, any higher-order parameterization can be combined with feature-mapping using an immersed-boundary method, such as XFEM.

Van Miegroet and Duysinx (2007) use XFEM and a fixed-grid to optimize the shape of a 2D fillet for stress minimization. The fillet is parameterized by an implicit geometry description, such as a super-circle, or super-ellipse. Noël et al. (2016) use XFEM and feature-mapping for shape optimization of bi-material structures, such as those with elliptical stiff and soft inclusions. The same framework is also used to optimize the shape of inclusions when designing micro-structural material layout to minimize the maximum stress (Noël and Duysinx 2017).

The interface-enriched generalized FEM (IGFEM) is used by Najafi et al. (2015) for shape optimization of bi-material structural and thermal problems, where an explicit geometry parameterization is used, e.g., circular and elliptical shaped inclusions. XFEM and IGFEM both use a fixed-grid and enrichment functions. The main difference is that the additional degrees of freedom in XFEM enrichment are added to the fixed-grid nodes, whereas in IGFEM they are added to points where the interface intersects a fixed-grid element edge.

7 Hybrid feature-mapping/free-form methods for topology optimization

Hybrid methods are defined as those that combine feature optimization with free-form topology optimization, where free-form means topology optimization without any high-level feature control (such as conventional density-based and level-set methods). Features can be solid or void and feature design variables can include position, orientation, size, shape, and/or number. Historically, the development of hybrid methods began before pure feature-mapping methods for topology optimization ("pure" meaning without any free-form optimization). Some techniques developed for hybrid methods were later exploited in the development of pure feature-mapping methods.

7.1 Combining free-form with features

A key technical challenge of hybrid methods is to combine the properties of features with the properties of the free-form structure in a single analysis model. Ideally, the

method should enable gradients of all design variables to be computed, so that an efficient gradient-based optimization approach can be used to solve the problem. This is achieved by combining feature and free-form geometry using either a map-then-combine, or combine-then-map approach (as discussed in Section 4).

Qian and Ananthasuresh (2004) used an implicit Gaussian, or peak function to map the stiffness of rectangular solid features onto the analysis mesh via

$$\hat{E}_i(x, y) = E_i \exp\left(-\frac{(\Delta\tilde{x}_i)^\eta}{\sigma_{xi}^2} - \frac{(\Delta\tilde{y}_i)^\eta}{\sigma_{yi}^2}\right) \tag{38}$$

with

$$\begin{pmatrix} \Delta\tilde{x}_i \\ \Delta\tilde{y}_i \end{pmatrix} = \begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix} \begin{pmatrix} x - x_i \\ y - y_i \end{pmatrix}, \tag{39}$$

where $\hat{E}_i(x, y)$ is the stiffness of the feature at location (x, y); E_i is the stiffness of the feature material, x_i and y_i are the location design variables; and θ_i is the orientation. If the exponent $\eta >= 4$, then the feature is approximately rectangular and the σ values control the dimensions.

The peak function was originally developed for multi-material topology optimization. Here, it is used to combine the stiffness of the features with the stiffness of free-form structure using a simple summation as

$$E(x, y) = \hat{E}_0(x, y) + \sum_i^N \hat{E}_i(x, y), \tag{40}$$

where N is the number of components and $\hat{E}_0$ is the stiffness of the free-form structure, which is defined by element-wise density design variables and penalized using a similar peak function. This map-then-combine approach allows straightforward computation of sensitivities, as the effective stiffness of an element in the fixed mesh is a linear combination of free-form and feature design variable values. The multi-material formulation also allows for features and free-form structure to have different properties. However, the implicit geometry representation in (38) is limited by the shapes it can represent.

Several methods define features using implicit functions with more geometric freedom. A popular choice is to define the boundary of a feature as the zero level set of an implicit scalar function (6).

Chen et al. (2007) represent both the features and free-form structure using implicit functions. The benefit of this approach is that the free-form structure can be combined with arbitrary features using a combine-then-map approach with Boolean operations, resulting in a single implicit function that describes the overall structure. However, Boolean operations cannot be differentiated. Thus, smooth approximations of Boolean operations are used instead, such as smooth R-functions (see Section 4.1). However, this combine-then-map approach makes solid features become indistinguishable from the free-form structure in the analysis model. Thus, solid features are limited to have the same material properties as the free-form structure. R-functions have also been used in other hybrid methods to create complex features by combining primitives represented by implicit functions (Xia et al. 2013).

If features are defined using implicit functions then a smoothed Heaviside function, e.g., (16), can be used to map feature properties onto a fixed-grid using pseudo-densities (Xia et al. 2013; Zhang et al. 2015). This method is discussed in more detail in Section 4.2.1. Using the smoothed Heaviside also allows for features with different stiffness values to be combined in the analysis model. For example, Xia et al. (2013) combined implicit feature representations with an element density-based free-form structure representation and used the following map-then-combine approach to define the Young's modulus within each element of the fixed mesh

$$E_e = \rho^p \left(E_0 + \sum_{i=1}^N \widetilde{H}(\phi_i)(E_i - E_0)\right). \tag{41}$$

This formulation works for void features where $E_i \approx 0$, although it does not inherently prevent element density variables beneath solid features from becoming zero, potentially leading to changes in the shape or topology of a feature. However, this can be avoided in the optimum for the minimization of compliance problem (subject to a volume constraint on the free-form structure) by excluding elements associated with solid features from the volume calculation. Pollini and Amir (2020) used a similar approach when using geometric features to specify different material properties within an interface region in a structural assembly.

It is also possible to use an exact Heaviside function when mapping implicit void features onto the fixed mesh (Shan 2008; Kang and Wang 2013). The free-form structure is parameterized using element-wise density variables and the pseudo-density within an element, centered at x, is simply computed by

$$\rho(\mathbf{x}) = \tilde{\rho}(\mathbf{x})H(\phi_\omega(\mathbf{x})), \tag{42}$$

where ρ is the physical density used to formulate the finite element stiffness matrices and $\tilde{\rho}$ is obtained using a density filter of the free-form density design variables. The inherent discontinuity of the exact Heaviside function does not allow explicit derivative calculation for feature location and orientation. Thus, shape derivatives (obtained from the analytical pre-discretized governing equations) are used to update the location and orientation variables. These shape derivatives are then used to obtain velocity components for feature location and orientation, which are then used with the Hamilton-Jacobi equation to update the implicit functions that define the void

features. However, using this approach does not allow for feature design variables (location and orientation) to be combined with free-form structure density variables in the same nonlinear programming optimization problem. Thus, to achieve simultaneous optimization of free-form structure and feature location/orientation, the feature velocity components are included as design variables when using a gradient-based optimizer.

Li et al. (2017) introduced the stiffness spreading method to combine properties of explicit geometry features and the free-form structure. The method starts with a fitted finite element mesh of the component whose shape does not change, but it only moves and rotates. To combine this with the free-form structure, defined on a fixed mesh, the component stiffness matrix is transformed into an equivalent matrix that has the same dimensions and degrees of freedom as the fixed mesh. The free-form structure and transformed feature matrices are then simply summed for analysis. Thus, this can be considered a map-then-combine approach. The transform essentially provides a linear map between a solid feature and fixed-mesh degrees of freedom. This could be achieved by a local transform, where feature nodes are simply mapped to the closest fixed-mesh nodes (i.e., those associated with a single element). However, this creates a discontinuity in the derivatives of the feature location and orientation. Thus, a non-local transform is used, where the connection is smoothed, or spread over several nodes in the fixed mesh within a predefined radius.

Multi-point constraints (MPCs) have also been used to connect explicit solid features to the fixed mesh used to parameterize the free-form structure (Zhu et al. 2015; Gao et al. 2015). The solid features are meshed independently and connected to the free-form structure fixed mesh at a number of predefined locations that represent bolts or rivets. This approach is in contrast to other methods that assume a perfect bonding at the interface between solid features and free-form structure. The use of MPC connections has also been extended to a multi-frame problem, where component locations and free-form structure for a number of frames are simultaneously optimized with the location of the frames in a larger design space (Zhu et al. 2017). MPCs are also used to connect frames to the free-form structure in the larger design space.

The above methods all aim to map feature properties onto a fixed analysis mesh in the spirit of feature-mapping methods as defined in this review. The alternative approach of re-meshing has also been used in hybrid methods. Solid features of fixed size and shape are discretized using a fitted mesh, whereas a fixed regular mesh is used to parameterize the free-form structure. The two meshes are combined into a single fitted mesh for analysis. This is achieved by first placing feature meshes over the fixed mesh. Elements in the fixed mesh that are partly covered by a feature are then divided, or re-meshed, to create a single conforming mesh for analysis (Zhu et al. 2008).

This local re-meshing at the boundary leads to changes in the free-form mesh whenever a feature moves. Thus, the usual element-wise density parameterization for the free-form structure cannot be used. Instead, the concept of density points is introduced, where the density at a point influences several elements within a local domain (Zhu et al. 2008).

Re-meshing approaches add computational complexity and cost (especially in 3D) compared with the fixed-grid mapping methods. Also, semi-analytical derivatives of feature design variables are required, where the derivative of the global stiffness matrix with respect to a small perturbation in the variables is computed using finite differences (Zhang et al. 2011). This process also adds significant computational cost, as the mesh is perturbed and stiffness matrices recomputed for each feature design variable.

Techniques have been proposed to reduce the computational cost of semi-analytical derivatives. Xia et al. (2012b) used a superelement formulation to reduce the size of feature stiffness matrices. The same authors also proposed using a smoothed Heaviside approximation of the boundary to facilitate full analytical derivatives (Xia et al. 2012a). The idea is to compute stiffness matrix derivatives based on the change in material properties as a feature moves. The smoothed Heaviside approximation is required to smooth the sharp discontinuity in material properties. However, re-meshing is still performed when features move.

One benefit of re-meshing is that the interface between solid features and free-form support structure is explicitly modeled. This could be important if the bonding between solid features and free-form support structure is included in the analysis as a nonlinear interface behavior. A recent example of this was presented by Liu and Kang (2018), where a cohesive zone interface model was used and the dissipated interface energy added to a compliance objective function. It was observed that features move to positions where the free-form support structure is under compression, as this helps minimize the dissipated interface energy.

XFEM is an alternative to re-meshing that can also capture the explicit interface (see Section 3.2). Hybrid methods that use XFEM also use implicit functions to describe features, as this provides a natural way to identify elements in the fixed mesh that contain an interface between the free-form structure and a feature. Furthermore, analytical derivatives of feature location/orientation design variables can be obtained using the chain rule (Zhang et al. 2012). XFEM has been combined with density-based (Zhang et al. 2012; Wang et al. 2014b), level-set-based (Zhou and Wang 2013) and parameterized level-set (Liu et al. 2014) methods for simultaneous optimization of features and free-form structure.

In summary, various methods have been proposed to combine properties of features and free-form structure in hybrid methods. Implicit and explicit features have been combined with density and level-set-based free-form structure. A key difference in the methods is whether they follow a combine-then-map or map-then-combine approach, as only the latter allows different material properties for solid features and free-form structure.

7.2 Feature design variables

The original motivation for developing hybrid methods was to simultaneously optimize the location and orientation of a number of solid features (with predefined shapes) and a free-form structure to support the features (Qian and Ananthasuresh 2004). The solid features represent components to be embedded within a free-form structure and their properties (such as stiffness) contribute to the overall performance of the structure. Thus, by simultaneously optimizing the free-form structure and position/orientation of solid features, an overall optimum can be found.

This idea also applies to void features of fixed size, which could, for example, represent necessary cutouts for systems to pass through the structure. Again, by simultaneously optimizing the position/orientation of void features and the free-form structure, an overall optimum can be found.

Beyond features of fixed shape and size, hybrid methods have been proposed that can also simultaneously optimize the size and/or shape of features. Cai and Zhang (2015) demonstrated the simultaneous optimization of circular hole radius and support structure. The hole had a fixed location and the solution was trivial, as the optimal radius is simply the lower bound. The method presented by Zhou and Wang (2013) can simultaneously optimize the location, orientation and shape of features. Furthermore, the method can be used to optimize only the shape of selected feature edges, thus providing several levels of control over feature geometry. This is achieved by modeling the free-form and feature geometries using implicit functions. These are then combined by multiplying Heaviside functions to give an indicator function

$$\chi_\omega(\mathbf{x}) = \prod_i^M H(\phi_i(\mathbf{x})), \tag{43}$$

where M is the total number of implicit functions (i.e., one for the free-form structure and $M-1$ features). The indicator function is then used to map the combined structure onto the fixed mesh for analysis (i.e., a combine-then-map approach). Note that this formulation assumes solid features and free-form structure have the same material, although the framework does allow for regions with different material properties.

The method presented by Liu et al. (2014) also uses implicit functions to represent both free-form and feature geometries, which are combined together using R-functions (also a combine-then-map approach). The features considered are primitives that can also change shape and size. For example, the semi-major and semi-minor axes of an elliptical void feature were simultaneously optimized with the free-form structure. In addition, Liu et al. (2014) introduced the idea of a tolerance zone, which is a fixed-width region of material around a void feature. This was achieved by ensuring the presence of free-form structure within a certain region around void feature edges.

Lin et al. (2015) proposed a hybrid method that can optimize the size, shape, and number of features. Features are defined explicitly using splines and the free-form topology is parameterized using a pseudo-density type approach. A criterion based on the topological derivative is used to create new features, such as holes and inclusions.

The hybrid method proposed by Pollini and Amir (2020) uses a piecewise linear feature to define an interface region between parts of a structural assembly. The geometric variables are simply the horizontal positions of points on the line (with vertical positions fixed). This enables simultaneous optimization of the free-form geometry (via an element-wise density field) and the interface region in a structural assembly, while taking account the different material properties and geometric restrictions within the interface region.

7.3 Optimization strategy

Most hybrid methods aim to simultaneously optimize the features and free-form structure. However, it has long been recognized that the hybrid problem formulation with movable features is non-convex and the solution is often highly dependent on the starting position of the features (Qian and Ananthasuresh 2004; Zhu et al. 2015; Cheng et al. 2018; Liu and Kang 2018).

A simple method to overcome this is to solve the problem several times with different starting positions of the features, although this adds computational cost. Another approach is to initially optimize just feature positions/orientations. Then, after a set number of iterations, or after the feature-only problem converges, both features and free-form structure are optimized simultaneously (Kang and Wang 2013; Wang et al. 2014b; Li et al. 2017). This strategy may provide a better design, as features can get stuck near their starting positions if the simultaneous approach is used from the start.

A variant of this approach is to fix feature designs and positions after the initial optimization (Bakhtiarinejad et al. 2017). This results in a sequential optimization approach that could be repeated several times, i.e., optimize the

feature design and then the free-form structure alternately, until convergence (Xia et al. 2012b). An advantage of the sequential approach, is that different optimization strategies can be applied to feature design and free-form structure.

7.4 Applications

Most hybrid methods have been applied to maximize the stiffness (minimize compliance) of the combined feature and free-form structure. Thus, solid features are placed to best utilize their stiffness, whereas void features are placed to least disrupt the free-form structure from obtaining the stiffest design.

Developments for other applications include the design of compliant smart structures with embedded movable actuators (Wang et al. 2014b), where the actuator features provide an input force and the free-form structure is used to design the compliant part of the mechanism. By simultaneously optimizing embedded piezoelectric actuator placement and the compliant mechanism design, an overall optimum can be found. Cai and Zhang (2015) used features and free-form structure defined by implicit functions and combined by R-functions to optimize structures with stress constraints. Finally, there have also been applications for heat transfer problems (Li et al. 2017; Cheng et al. 2018), where the position of solid features are optimized to take advantage of their thermal conductivity and/or heat flux boundary conditions.

Currently, there are few applications of hybrid methods beyond designing stiff structures. Thus, there is potential for future research into other application areas, such as electromagnetic and electrical-mechanical problems, fluid flow and fluid-structure interactions problems and problems involving nonlinear mechanics.

8 Feature-mapping methods for topology optimization

In this section we discuss feature-mapping methods for topology optimization, where the structure is exclusively defined by the combination of high-level parametric descriptions of voids in a solid design region, or of solids in a void design region. For brevity in the discussion, we will refer to the former as holes and the latter as components. We note that topology optimization using only solid components has also been called the adaptive ground structure approach by Guo et al. (2014).

The choice between using components or holes for topology optimization may seem arbitrary. However, if components are used, then each component may have different functionalities and/or different physical models could be used for each component, as discussed by Guo et al. (2014).

As a historical note, a mention is due to the bubble method (Eschenauer et al. 1994), which describes holes in a solid region using B-splines and is, in fact, one of the first topology optimization techniques. However, this method does not fit into the family of methods covered in this review, as it employs a conforming mesh and thus requires re-meshing upon design changes, and also lacks a mechanism to merge holes.

We also note that the methods discussed in this section can be shown to be essentially equivalent to density-based or level-set techniques under certain conditions—for instance, by having a geometric component per element in the mesh such that the component size does not exceed the element size, as discussed in Liu et al. (2018). However, exploring this relationship is outside of the scope of this review.

8.1 Combine-then-map methods

As described in Section 4, feature-mapping methods for topology optimization render structures exclusively made of the combination of geometric features. Examples of these methods are shown in Fig. 24. A pioneering work using combine-then-map was introduced in Cheng et al. (2006) and Mei et al. (2008).[1] This method formulates and successfully demonstrates all of the hallmark features of methods in this category, as well as other features that have not been explored elsewhere at the time of writing this review, such as the primitives interpolation strategy described in Section 8.3. To our surprise, this contribution went unnoticed in many subsequent works; only until recently have some works started citing it. We hope our review aids in restoring credit to this foundational contribution. In this method, holes are combined using (non-smooth) R-functions.

The first use of combine-then-map with primitive-shaped solid components using gradient-based optimization is by Guo et al. (2014), in implementing the method of moving morphable components (MMC). This work has many derivatives and it discusses the merits of this approach. At the time of writing this review, is the most cited topology optimization approach using feature-mapping. It uses a true maximum function to combine components into a single implicit function. The majority of combine-then-map implementations with implicit representations use a non-differentiable maximum. Exceptions are the works by Zhou et al. (2016), Sharma (2017), and Du et al. (2019), which use a smooth Kreisselmeier-Steinhauser (K-S) approximation to combine the signed-distance representations; the works by Liu et al. (2018) and Du et al. (2020), which employ a K-S function to combine the implicit functions of the individual

[1] The latter work seems to be a more detailed journal version of the former, hence from hereon we only reference the latter.

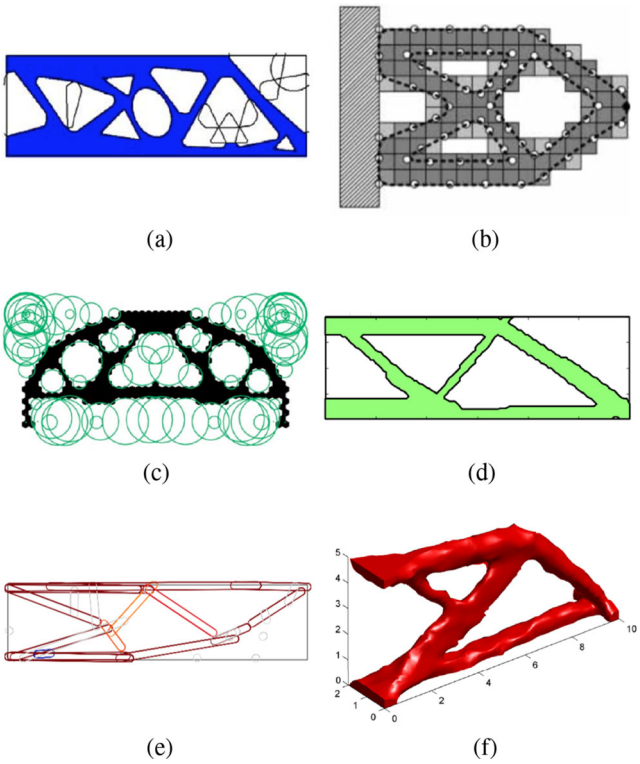


Fig. 24 Examples of different feature-mapping methods for topology optimization: (a) feature-based method (Mei et al. 2008), (b) smooth boundary method (Lee and Kwak 2008), (c) circular shaped masks method (Saxena 2011), (d) MMC method (Guo et al. 2014), (e) geometry projection method (Norato et al. 2015) and (f) MMV method (Zhang et al. 2017b)

components; and the works by Zhang et al. (2016b) and Xie et al. (2018), which use smooth R-functions. The work by Sharma (2017) smooths out the combined level-set function with an anisotropic filter to produce smooth boundaries, although it indicates that, for the examples presented, such filtering is not necessary to produce good results.

The works by Lee et al. (2007) and Lee and Kwak (2008) and Kim et al. (2008) are the first to employ an explicit representation, namely B-splines. Holes are merged by deleting control points, as outlined in Section 4.2.2. Wang and Yang (2009) is perhaps the first work to design structures made exclusively of solid components, consisting of wide curves modeled using Bézier curves. The pseudo-density mapping in this method is not differentiable, as described in Section 8.4; thus, this method uses a genetic algorithm for the optimization.

In the material mask overlay strategy (MMOS) of Saxena (2011), holes are represented as the union of primitive-shaped regions (the "masks"), such as circles, rectangles and ellipses. The mapping to the analysis mesh uses a small or unity pseudo-density depending on whether an element's centroid is inside or outside any of the masks, respectively. As this mapping is not differentiable, a gradient-free hill-climbing method is used for the optimization. Wang et al. (2012) introduced a differentiable version of MMOS for the design of photonic waveguides that combines a free-form density field with circular holes. The boundary of the holes is smoothed as described in Section 3.1.4 and holes are combined using a product of the smooth Heavisides, similarly to (43). In the examples provided in this work, however, the circular holes do not overlap. Hoang and Jang (2017) present a modification of this differentiable MMOS approach for designs with solid components.

8.2 Map-then-combine methods

The most common method that embodies the map-then-combine strategy is the geometry projection method of Bell et al. (2012) and Norato et al. (2015). This method employs a volume fraction calculation using a circular (in 2D) or spherical (in 3D) sample window (cf. (20)) to obtain a pseudo-density for each of the components. To combine features, the preliminary work by Bell et al. (2012) used the minimum signed distance to any of the components, which is equivalent to a true maximum. The work by Norato et al. (2015) introduced the size variable that allows the complete removal of a component via penalization (cf. (36)). It also uses a p-norm (cf. (30)) as a smooth approximation of the maximum to perform the combination of the penalized pseudo-densities.

An interesting recent contribution is the generalized geometry projection method of Coniglio et al. (2019). This method shows that the mappings to an ersatz material in the geometry projection method (Norato et al. 2015), the MMC method (specifically, the implementation of Zhang et al. (2016c) that uses an ersatz material for the analysis, discussed in Section 8.4), and the moving nodes approach discussed later in this section, can be seen as special cases of the numerical integration of the smooth Heaviside (a view that is in line with the approach of Wein and Stingl (2018)).

A mention is due to the works of Guest and Zhu (2012) and Ha and Guest (2014), which remove holes or produce components of a fixed shape and size (e.g., circular), respectively, by using projection filters with the given shape on a grid of density variables. In these methods, the design is described by a density field instead of a high-level parametric description of the holes or components, hence they do not fall under our categorization of feature-mapping techniques. The shape of the hole or component determines the neighborhood of elements that contribute to the filtered density in each element.

Another mention is due to the moving nodes approach (MNA) of Overvelde (2012). In this method, the geometry is described by mass nodes with variable positions within

the design region. This method employs an element-free Galerkin (EFG) method for the analysis (which is a meshless method). A combined density for the ersatz material is obtained via a filter operation that uses the same kernel function as the one used for the EFG analysis. This filter plays both the role of combining the mass nodes and mapping to the pseudo-density field for the ersatz analysis. However, we do not consider this method to fall in the category of feature-mapping techniques, since the mass nodes do not have any geometric features in the sense defined in this paper.

We also note that techniques that use geometric representations such as B-splines to represent a density field (e.g., Qian 2013) or a level-set function (e.g., Wang et al. 2019a) are also excluded from our review, as these parameterizations do not directly represent high-level geometric features.

8.3 Geometric representations

Most of the combine-then-map implementations that use upfront implicit representations use hyperellipses to represent various primitives. The original MMC implementation and most of its derivatives employ a direct hyperellipse representation to model 2D bars (Guo et al. 2014), and 3D bars and plates (Zhang et al. 2017c; Sharma 2017). Zhang et al. (2016c) and Guo et al. (2016) modify the hyperellipse equation to produce curved bars with linearly, quadratically or sinusoidally varying width. Other methods, notably Mei et al. (2008) and Zhou et al. (2016), use a hyperellipse representation as well, but they convert it to a signed-distance representation, which, as mentioned in Section 3.1.4, can lead to better convergence. The geometry projection method in Norato (2018) models 2D solid components using supershapes, which correspond to a generalization of the superellipse formula that has variable symmetry, and thus can approximate a wide range of primitive shapes with a single equation. As in all geometry projection implementations, a signed distance to each supershape is computed to obtain the mapped pseudo-density.

Lee et al. (2007) and Lee and Kwak (2008) and Kim et al. (2008) modeled holes with B-splines. The 2D solid components in Wang and Yang (2009) are modeled using wide Bézier curves, with design variables being the control point locations. The curves connect support and loading points and a constraint in the optimization is required to prevent curve self-intersections. 2D (Zhang et al. 2017d; Du et al. 2019; Du et al. 2020; Gai et al. 2020; Xie et al. 2020a) and 3D (Zhang et al. 2017b; Xie et al. 2020a; Xie et al. 2020b) B-splines have been used to model holes in the Moving Morphable Voids (MMV) method, which is a variant of the MMC method. Control points inside the intersection are made inactive, and the resulting combined B-spline is converted to a signed-distance function prior to mapping to the analysis.

An implicit geometry representation common in geometry projection implementations is offset surfaces (cf. Bloomenthal and Wyvill 1990), in which the component boundary is given by the set of points equidistant to a medial axis or surface. For instance, a 2D bar is given by a rectangle with semicircular ends (Norato et al. 2015; Smith and Norato 2020), a 3D bar by a cylinder with hemispherical ends (Watts and Tortorelli 2017; Kazemi et al. 2018, 2019; Smith and Norato 2020), and a 3D plate as a cuboid with semi-cylindrical edges and quarter-sphere corners (Zhang et al. 2016a). Offset bars have also been used in conjunction with the MMC method in Deng and Chen (2016), Deng et al. (2019), and Deng and To (2020) and the geometry projection method (Tahhan 2019) to produce designs where bars are connected at all times by sharing endpoints of their medial axes; in Hoang and Jang (2017) to impose a minimum thickness in the structure (see Section 8.5); and in Hoang et al. (2020a) to design coated structures. Offset plates with a single curvature radius are modeled in Zhang et al. (2018a). The offset surface representation is combined with a free-form density field in Zhang et al. (2016a) to produce designs made of plates with free-form holes.

An interesting geometric representation is employed in Mei et al. (2008), whereby each hole is represented as a weighted sum of prescribed geometric primitives (e.g., circles and triangles) that are implicitly represented using signed-distance functions. A constraint is added to the optimization to penalize the primitive weights so that the holes converge to being, for example, either a pure circle or a pure triangle. A similar constraint is used in Norato (2018) to penalize supershapes parameters so that the optimal design is exclusively made of, e.g., rectangles and ellipses. Mei et al. (2008) also use shape matching techniques to take the resulting designs and consolidate them into fewer primitives, after which an additional optimization stage is performed.

8.4 Analysis approaches

As mentioned in Section 4.2, combine-then-map approaches can perform the analysis using either immersed-boundary or pseudo-density techniques. Map-then-combine approaches, on the other hand, have only used pseudo-density techniques, although they could in principle use immersed-boundary approaches, as discussed in Section 4.3.2.

Combine-then-map approaches that utilize an implicit representation and pseudo-densities for the analysis (cf. Mei et al. 2008; Zhang et al. 2016c) compute a smoothed Heaviside of the combined implicit function (see Section 3.1.4). Those that use an immersed-boundary approach, have used the XFEM method (cf. Guo et al. 2014; Sharma 2017).

Other approaches use pseudo-densities and isogeometric analysis (IGA) (Hou et al. 2017; Xie et al. 2018).

Different analysis techniques are used in combine-then-map approaches that use a direct explicit representation (i.e., without conversion to an implicit representation such as a signed distance). In Lee et al. (2007) and Lee and Kwak (2008) and Kim et al. (2008), the analysis uses a pseudo-density that is computed by determining the intersections of the B-spline holes with the element boundaries, replacing the intersected boundary with a straight line, and using a volume fraction approach, as in García-Ruíz and Steven (1999). The works by Seo et al. (2010) and the recent MMV method of Gai et al. (2020) are similar to these methods; however, they use IGA to decompose elements that intersect the structural boundary into multiple cells to perform the integration. The work by Zhang et al. (2017g) uses the weighted B-spline finite cell method for the analysis. The technique with wide curves of Wang and Yang (2009) uses an element pseudo-density, which is calculated by successively subdividing elements that intersect the structural boundary, and computing the area ratio of the sub-cells that lie inside the curve (or of overlapping curves). In this approach, a non-gradient-based optimizer is used.

An interesting analysis technique is presented in the MMV method of Zhang et al. (2017b), whereby elements in the void regions are removed from the analysis at each iteration to decrease the size of the analysis problem. These elements may be reintroduced in subsequent iterations if they become non-void.

A mention is due to the recent works of Zhang et al. (2020b) and Zhang et al. (2020c), which combine holes in 2D and shell structures using IGA. However, unlike some of the previously mentioned approaches that use IGA merely to perform he numerical integration in elements that intersect the hole boundaries, these methods actually partition the element boundaries into new elements to produce a body-fitted mesh, and therefor, they do not fit our definition of feature-mapping techniques.

8.5 Complexity and minimum size control

We start this section with a brief discussion on complexity control. By complexity, we here mean the topological genus of the structure, i.e., its number of holes—the higher the genus, the more complex the structure. One of the well-known characteristics of most density topology optimization techniques is that, in the absence of a control mechanism (e.g., filtering, or a perimeter or slope constraint), the design is usually mesh-dependent. This is a byproduct of the fact that in these methods the design representation is tied to the analysis mesh. An alternative in these techniques is to use a grid of design variables that is independent of the mesh (cf. Nguyen et al. 2010).

Feature-mapping methods in general do not suffer from this mesh dependency because the representation of the components or holes is entirely independent of the mesh. Moreover, if lower bounds are imposed on the dimensions of the components, it is of course not possible for the optimization to obtain smaller feature sizes. Therefore, as long as the total number of components is kept to a maximum, the designs produced by these methods exhibit similar complexity and member size. This does not entirely mean, however, that different mesh sizes produce the exact same design due to several reasons. First, the mesh resolution affects the accuracy of the geometry mapping and the analysis solution and consequently the sensitivities, hence different mesh resolutions can lead to different local minima. Furthermore, in the case of pseudo-density approaches, mesh alignment (in conjunction with the interpolation approach) may also cause small differences, as noted in Section 3.1.2. Second, even if a lower bound on the component dimensions guarantees a minimum feature size, when components intersect or "touch" they can produce regions whose size is smaller (and possibly zero, i.e., a point intersection) than the desired minimum size.

A minimum size for individual components can be readily attained in various geometric representations by directly imposing bounds on the design parameters. For instance, when hyperellipses or supershapes are used to represent components, the width and length of the component can be directly controlled by bounding the hyperellipse radii (see Zhang et al. 2016b; Norato 2018); in the case of bars modeled using offset surfaces, the bar width can also be bounded in this way (see Smith and Norato 2020). When the component dimensions to be bound do not directly correspond to design parameters, then it is often straightforward to add constraints that are functions of the design parameters, as in the case of length constraints for bars represented by offset surfaces (see Tahhan 2019).

In addition to imposing bounds on the geometric parameters to ensure a minimum component size, the work in Sharma (2017) adds a penalty to the objective function to prevent the onset of plates whose medial surface area is smaller than a specified threshold. The motivation to impose this penalty is that small components with zero shape sensitivities may disconnect from the structure and cause oscillations in the optimization.

Several approaches have been proposed to enforce a minimum size for the combined structure in feature-mapping methods for topology optimization. In Zhang et al. (2016b), a constraint is added to impose a minimum distance between any two members that intersect. Since

the constraint is applied only to components that overlap, it appears to be non-differentiable. Hoang and Jang (2017) employ bars modeled with offset surfaces (i.e., rectangles with semicircular ends) and effectively control the minimum thickness by introducing two constraints: one that ensures a minimum volume in a mask that covers the rectangular portion of the bar, and another that limits the amount of intermediate material in the semicircular ends of the bars. The work by Niu and Wadbro (2019), which uses rectangular bars and the MMC method, adds a penalty term to the objective function to enforce a minimum size. This term consists of a smoothed Heaviside of the minimum distance between the center lines of two bars for all pairs of bars; this term is multiplied by another that consists of the angle difference between the two bars (raised to a power), so that the orientation between bars that are closer to being parallel is penalized, but those that are closer to being perpendicular are not. While the foregoing methods are effective, they are associated with specific geometric representations (i.e., bars), and a minimum size control that works for combined structures with components of any shape is currently missing. The recent work by Wang et al. (2019b) constrains the minimum distance between the medial axes of nearby or intersecting components to be less than a prescribed value, together with a constraint on the minimum thickness of the components.

A note is worth making with regard to intersections between components, which has an effect on the minimum size at intersections. For pseudo-density approaches, it is possible that components that should intersect in the optimal design are not fully connected and there is a small gap between them. As discussed in Norato (2018), this is due to the fact that some smooth approximations of the maximum, such as the p-norm of (30), approximate the true maximum from above, and thus they may render artificially high pseudo-densities, which increases the stiffness of small gaps between components. To circumvent this problem, several works have proposed smooth saturation functions that cap the combined pseudo-density to 1 (cf. Overvelde 2012; Coniglio et al. 2019; Deng and To 2020).

The onset of gaps between components could also be associated with linear volume-to-stiffness interpolations. The recent work of Smith and Norato (2020) notes that the penalization of (36) has a potential problem, namely that it penalizes the size variable but not the projected density; therefore, in the event that the size variable is 1, elements along the boundary of a component that have intermediate values of the projected density will result in a linear volume-to-stiffness relation; this, as noted extensively in Section 3.1.2, is incorrect as it leads to unrealistically stiff material. To circumvent this, Smith and Norato (2020) propose penalizing both terms, and they note this leads to less instances of gaps between components.

8.6 Adaptive resolution

Several feature-mapping methods for topology optimization have advanced different mechanisms to adaptively improve the resolution of the mapping. The work by Liu et al. (2018) employs a coarse finite element mesh for the analysis, and a fine mesh of sampling windows for the feature-mapping. Therefore, several sampling windows are contained in a finite element. This technique renders a higher resolution of the geometry. The numerical integration to compute the element stiffness matrix sums over the contributions of all the pseudo-density mapping elements. A similar scheme using the MMV method and IGA for the analysis is presented in Du et al. (2020).

The recent work in Zhang et al. (2020a) introduces an adaptive refinement technique for the geometry projection method. The refinement indicator is given by the projected pseudo-density, so that the refined mesh is finer along component boundaries and coarser in solid and void regions. The refinement corresponds to h-refinement of quadrilateral (in 2D) and hexahedral (in 3D) elements with hanging nodes. The refinement and coarsening is done at every iteration, and the method is demonstrated for meshes with several million elements. The recent MMV implementation by Xie et al. (2020a) employs hierarchical B-splines for adaptive refinement using IGA for the analysis. The refinement indicator corresponds to a band of values around the zero value of the combined implicit function and so, as in the foregoing work, it refines boundaries. This method does not coarsen previously refined elements, and it performs a full optimization for each level of refinement. We note that these two methods are the exception to using a fixed analysis grid, as the mesh is obviously modified throughout the optimization; nevertheless, they still employ feature-mapping to a non-body-fitted mesh.

8.7 Design space modification

As discussed in Section 4.4, feature-mapping methods for topology optimization are more prone to falling into poor local minima than free-form methods. Therefore, the choice of initial design, as well as any mechanisms to adaptively modify it, are important aspects in these methods.

One strategy employed by several of the techniques that design holes in a solid region is to adaptively introduce holes at locations that attain the most negative value of the topological derivative of the compliance. This strategy, which has been used by level-set techniques (cf. Sigmund and Maute 2013 and references therein), was introduced in feature-mapping methods by Mei et al. (2008) and later used by other methods (e.g., Lee et al. 2007; Lee and Kwak 2008; Kim et al. 2008). To our knowledge, this strategy has not been used yet to introduce solid components

into a void domain. In addition to adaptively introducing holes, the work of Mei et al. (2008) also removes holes that do not intersect any solid region to simplify the design. The generative design method of Li et al. (2019) adaptively adds components to generate tree-like structures for area-to-point conduction problems. It performs a sequence of optimizations using the MMC method; at the end of each optimization run, it adds or removes components at the ends of existing branches depending on whether the leaf-side thickness of the component exceeds or falls short of prescribed thresholds, respectively.

To prevent entrapment in poor local minima, the MMC method of Zhang et al. (2017f) introduces a function that smoothly approximates the number of "effective" bars in the design. Two or more bars are considered to form a single effective component if they overlap (near) co-linearly . The function examines the intersection between the bars: if the intersected area is greater than some specified threshold, and if the angle between the bars does not nearly equal zero or π, they are considered separate components. A constraint is then imposed on the number of effective components.

A "bootstrapping" strategy to produce good initial designs for 2D problems with bars is presented in Weiss et al. (2018), which uses density-based topology optimization to produce an initial design. This result is first converted to a 0-1 design by thresholding, and then, it is skeletonized to produce the medial axis of the structure. The nodes of this skeleton are connected with straight bars to produce an initial design for the MMC method. A similar approach is presented in Lian et al. (2020).

In Zhang and Norato (2018), the gradient-based tunneling method is used in conjunction with the geometry projection method to go from one local minimum to a better one. In essence, after converging to a local minimum (the optimization phase), this technique adds a term to the objective function that makes the current minimum a pole of the modified objective, and uses this function to find another design with an equal or lower objective (the tunneling phase). If successful, a new optimization phase is started with the original function.

Finally, taking advantage of the reduced number of design variables in feature-mapping methods, and to attempt to prevent entrapment in poor local minima, recent works have used statistical techniques such as: an evolutionary strategy (Bujny et al. 2018), Bayesian optimization (Sharpe et al. 2018), and machine learning methods such as support vector regression and K-nearest neighbors (Lei et al. 2019) to perform the optimization. Raponi et al. (2019) also used a Kriging-based surrogate model to improve the convergence rate and optimal designs when using a statistical optimization technique.

8.8 Geometric constraints

One of the most appealing aspects of feature-mapping methods is the ability to impose geometric constraints. These may be motivated by, for example, wanting to make the structure out of stock material (such as bars or plates), or facilitating its manufacturing. The simplest of these constraints, as noted in Section 8.5, is to enforce bounds on the component geometric parameters. For instance, it is possible to obtain a structure that is made of bars with the same cross section (cf. Norato et al. 2015), or to impose upper bounds on the dimensions of rectangular plates to account for commercial availability (Zhang et al. 2016a). Although conceptually simple, enforcing these requirements in free-form methods is difficult.

Surprisingly, however, there have only been a few works to date that incorporate other geometric constraints. A common requirement, particularly for structures made of stock material, is to ensure that the components lie entirely within the design region, as otherwise a component that is only partially inside may require cuts (for example, through the thickness of a plate) that are impractical to manufacture. For rectangular and cuboid design regions in 2D and 3D, respectively, this requirement is easily satisfied by imposing bounds on the design variables that determine the positions of the components. However, for other design region shapes, particularly those that are not convex, using bounds or simple constraints on the design parameters does not work. The geometry projection method in Zhang et al. (2018a) proposes a way to address this by creating a layer of points slightly outside of the boundary of the design region, and imposing a constraint in the optimization that the maximum projected pseudo-density in any of these points is zero. By imposing the constraint on the pseudo-densities, this technique works for components of any shape. An alternative to address this issue is to employ the no-overlap techniques discussed in Section 5, and require that the geometric components do not overlap the exterior of the design region.

Another type of geometric constraint aims to control the orientation of components. In the MMC implementation of Guo et al. (2017), a constraint is imposed on the orientation of bars in structures fabricated via additive manufacturing to ensure their angle with respect to the print direction is smaller than the overhang angle. In the same work, holes on a solid structure are modeled using B-splines, and a constraint is imposed on the control points positions to ensure the boundaries of the holes do not exceed the overhang angle. In the latter approach, however, another constraint is also imposed to prevent holes from merging in order to completely avoid violations of the overhang angle requirement, thus this method does not change the

topology of the design. The recent MMC implementation by Xian and Rosen (2020) imposes an angle constraint between bars to avoid overhangs in additive manufacturing. In the work by Wein and Stingl (2018), described in Section 6.2 and shown in Fig. 22(b), a maximum overhang on the strip boundaries is readily enforced by imposing constraints on the relative position of boundary points (which are themselves a function of the control points and the strip semi-widths). Similarly, using constraints on the positions of the control points and on the semi-widths, this method enforces slope and curvature constraints on the overall shape of the strip.

In the geometry projection technique of Smith and Norato (2019b), a constraint is imposed on the minimum angle between any two bars (modeled as offset surfaces) to ease manufacturing. The form of the angle constraint is somewhat similar to that of Niu and Wadbro (2019) in that it multiplies a term that penalizes the angles between two bars with another that measures the distance between their medial axes. However, the angle term penalizes angles that are smaller than a prescribed value (instead of the angle difference), and the distance term is used to impose the angle constraint only between bars that are closer than a prescribed value. Moreover, the penalty term is multiplied by the size variables of both bars, so that no angle constraint is imposed when one or both of the bars have a zero size variable.

Feature-mapping methods often produce designs where components—particularly bars—intersect such that they have a "long" overlap, i.e., they are close to each other and they are near parallel. If the structure is fabricated with stock material, these overlaps make fabrication difficult. The aforementioned techniques to impose an angle constraint between bars are very effective in preventing this situation. Smith and Norato (2019b) define a no-overlap region in the bar, corresponding to the bar minus its circular ends. A constraint is then imposed that the sum of the projected pseudo-densities for each individual bar on the overlapping region is at most unity. In effect, these techniques ensure that a bar intersects another only at its ends. However, it may be intersected by other bars anywhere along its length. Another technique to prevent long overlaps is the one proposed by Watts and Tortorelli (2017) for design of periodic lattices, where a constraint is imposed that ensures the largest distance from the endpoint of a bar's axis to the closest endpoint of another bar is at most the bar's half width. This ensures a bar only intersects another at its ends but not along the bars; thus, it is more restrictive than the foregoing techniques.

The geometry projection technique of Kazemi et al. (2020), which designs multi-material lattices made of cylindrical struts, introduces a no-cut constraint to ensure struts are not partially cut by the boundaries of the periodic unit cell or by the material symmetry planes. This renders designs more amenable to manufacturing as they only produce struts that are whole. This constraint ensures the difference between the values of the volume of an individual strut computed from its geometric parameters and from its projected pseudo-density on the reference region (i.e., the region that is reflected to enforce material symmetries) is negligible. A smooth maximum function is used to render a single constraint for all the components. This idea is similar in spirit to the one used in Kang and Wang (2013) to prevent the overlap of embedded components in a free-form structure.

Symmetry is another geometric constraint considered by existing methods. Due to the more restrictive design representation, feature-mapping methods may render designs that are not exactly symmetric even if the design region shape and the boundary conditions are such that a symmetric design is expected with free-form methods. In the case of compliance minimization with a volume constraint, this behavior was noted in Norato et al. (2015) for components of fixed width. This finding bears resemblance to the known fact that optimal truss designs can also be asymmetric not only for discrete 0-1 trusses (Stolpe 2010, 2016) but also for trusses whose element stiffness matrices are penalized as in density-based topology optimization schemes (Stolpe 2010). The works of Stolpe (2010), Rozvany (2011), Guo et al. (2012), and Guo et al. (2013) examine the issue of symmetry in layout of trusses and may provide clues to the lack of symmetry in feature-mapping methods for topology optimization.

To ensure a symmetric design, the geometry projection techniques in Watts and Tortorelli (2017), Kazemi et al. (2018), and Kazemi et al. (2020) use a simple strategy in which the point where the pseudo-density is computed is reflected with respect to the symmetry plane before computing the signed distance to the components. This strategy can be applied to any number of symmetry planes, as demonstrated in Watts and Tortorelli (2017) and Kazemi et al. (2020) to design periodic truss lattices whose homogenized properties exhibit desired symmetries. It can also be applied to any component shape. A similar strategy to produce symmetric and periodic structures is employed in the MMC implementation of Xie et al. (2020b). In Wein and Stingl (2018), the control points that define the strips are reflected with respect to symmetry planes to obtain, e.g., square symmetric structures.

8.9 Applications

Not surprisingly, most works in feature-mapping methods for topology optimization consider minimization of compliance with a volume constraint. Recently, some works have incorporated other structural responses and physical

regimes. Some of the applications mentioned in this section are shown in Fig. 25.

One problem feature-mapping methods have been used for is the design of linear compliant mechanisms (Deng and Chen 2016; Guo et al. 2016; Zhang et al. 2016b; Hoang and Jang 2017), whereby the objective function is typically related to maximizing displacements at certain locations. The mechanisms to impose a minimum size in Zhang et al. (2016b) and Hoang and Jang (2017) are effective in preventing single-node hinges often seen in compliant mechanism design using free-form methods. Another application in elastostatic structures is an MMV method for design-dependent loading (Zhou et al. 2019). Stress constraints are incorporated in the geometry projection technique of Zhang et al. (2017a) and in the MMV method of Zhang et al. (2018b); and stresses were minimized in the former work and in the MMC method of Takalloozadeh and Yoon (2017), which uses the topological derivative to

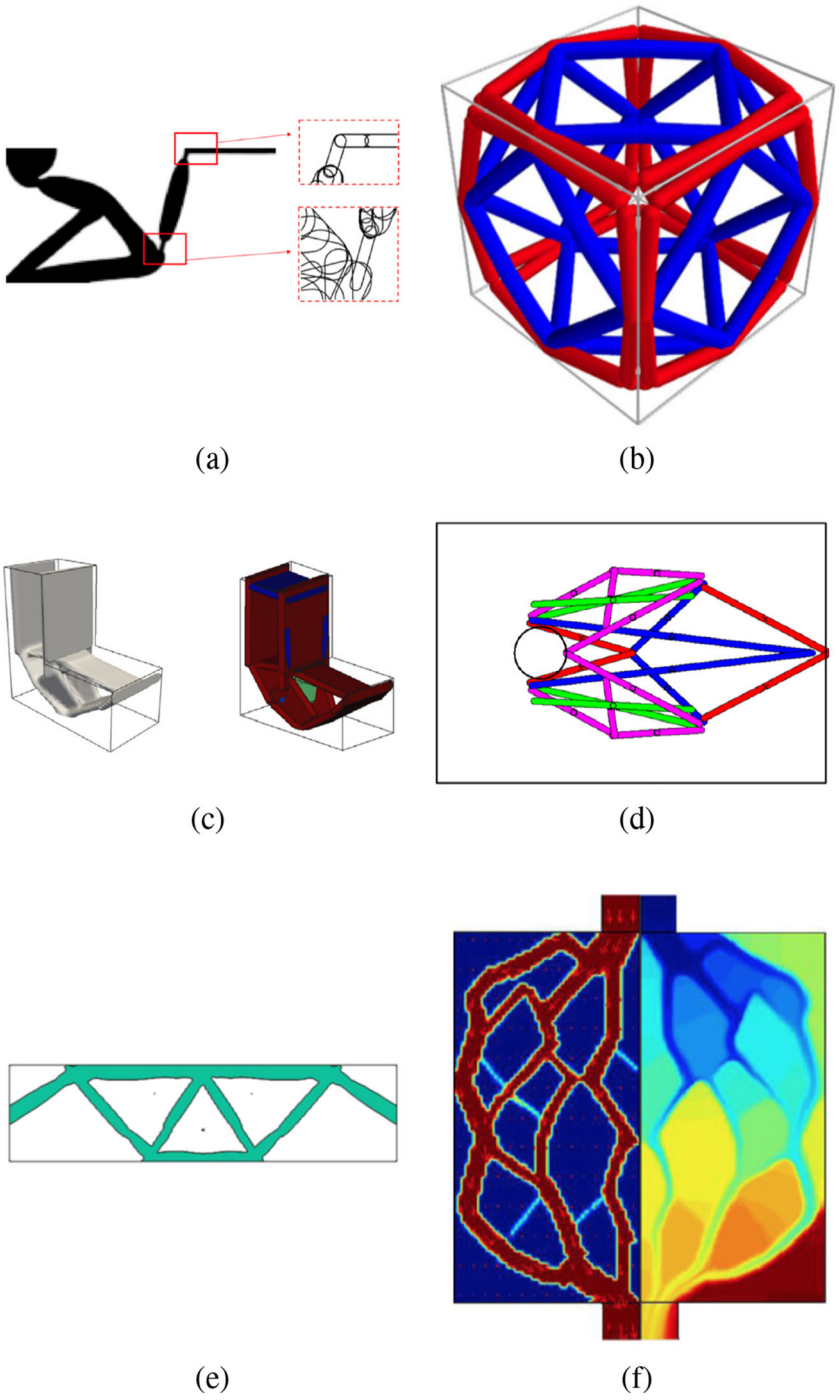


Fig. 25 Examples of different feature-mapping applications in topology optimization: (a) compliant mechanism design (Hoang and Jang 2017), (b) material design (Watts and Tortorelli 2017), (c) stress constraints (Zhang et al. 2017a), (d) multi-material structures (Kazemi et al. 2018), (e) geometric nonlinearities (Zhu et al. 2018), and (f) thermal-fluid problems (Yu et al. 2019)

compute the design sensitivities of the stress function. The recent work in Cui et al. (2020) employs MMC to design plate structures subject to out-of-plane loading, with the plate modeled using Kirchhoff plate elements.

In the area of material design, the geometry projection methods of Watts and Tortorelli (2017) and Kazemi et al. (2020) are applied to design periodic truss latices with desired material symmetries, to maximize the effective bulk modulus, or to obtain negative Poisson's ratio.

The design of multi-material structures, where each component is made from one material from a set of available materials, is another recent application of these methods. The first work to accomplish this is the aforementioned work of Watts and Tortorelli (2017). By adapting the density-based multi-material technique of Sigmund and Torquato (1997) to the geometry projection framework, their method simultaneously designs the layout of the bars within the unit cell and selects the best material for each component. The MMC method of Zhang et al. (2018d) designs structures with components made of multiple materials. However, the material for each component is assigned a priori and does not change during the optimization, and only the component layout is optimized. Kazemi et al. (2018) simultaneously design the topology and material choice of multi-material structures by adapting the discrete material optimization (DMO) technique of Stegmann and Lund (2005) to the geometry projection method. This method is extended in Kazemi et al. (2020) to the design of multi-material lattices.

Hoang et al. (2020b) present a method to design 2D components made of a porous material. The shape and topology of the macro component is given by combining bars; the honeycomb corresponds to a solid matrix with circular holes in a honeycomb tiling with designable hole radius and wall thickness. Two pseudo-density fields are computed—one for the macro component and one for the honeycomb—which are combined to produce the ersatz material.

The MMC method of Zhu et al. (2018) minimizes the compliance of geometrically nonlinear structures. It uses a neo-Hookean material along with the material interpolation of Wang et al. (2014a), which models low-density elements with a linear material and thus stabilizes the analysis. The asymptotes in the optimizer—namely, the method of moving asymptotes (MMA) of Svanberg (1987)—are adaptively updated to enforce a more conservative approximation of the optimization functions when the nonlinear finite element analysis has difficulty converging. The MMV approach in Xue et al. (2019) designs minimal-compliance structures and compliant mechanisms with finite deformations. To circumvent the convergence issues brought upon by low-density elements, this method adaptively removes elements inside the holes whose pseudo-density is lower than a prescribed threshold.

An interesting application of feature-mapping methods is the design of particular types of structures that are typically manufactured using stock material, such as plates or rods. These include the layout design of reinforcing ribs made of plates, including the geometry projection technique of Zhang and Norato (2017) and the MMC method of Zhang et al. (2018c). The latter work incorporates buckling constraints. The recent method of Bai and Zuo (2020) employs MMC to design structures made of hollow components, such as commonly used structural shapes. A subset of these works is devoted to the design of aircraft wingbox design: the method of Smith and Norato (2019a), which aims to find an optimal layout of ribs for the wingbox, and the method of Li et al. (2018), in which the layout of ribs is fixed, and MMV is used to find the optimal design of holes in the ribs to minimize fuel sloshing. A geometry projection technique is employed in Coniglio (2019) for the design of an aircraft engine pylon. The recent work in Chu et al. (2019) uses MMC to design periodic truss cores of sandwich panels with minimal compliance. The method of Hoang et al. (2020a) designs structures with one or two uniform-thickness coatings and an infill, each made of a homogeneous isotropic material. We note that the methods of Bai and Zuo (2020) and Hoang et al. (2020a) use not only Boolean unions of geometric components but also Boolean subtractions.

Another application area with growing interest is dynamics. For instance, several recent works have been devoted to design of flexible multi-body systems, such as the MMC techniques of Sun et al. (2018b, c) to design a variable-length structure, and Sun et al. (2018a) to design a component with large motion and large deformation. MMC is used to design a rotating thin plate to maximize its first eigenfrequency, or maximize the gap between two consecutive eigenfrequencies in Sun et al. (2019), and to remove its internal resonances in Sun (2020). In Xie et al. (2019), MMC is used to design the layout of damping patches on a vibrating plate to minimize its average kinetic energy over a frequency range. Wormser et al. (2017) design periodic lattice structures that act as phononic band gaps—a problem that is very hard to solve with free-form topology optimization techniques, which render disconnected structures that cannot be manufactured. The MMC method has also been combined with gradient-free optimization methods to optimize structures for crash-worthiness (Bujny et al. 2018; Raponi et al. 2019).

In the realm of heat transfer, combine-then-map methods have been applied to the design of thermo-elastic structures. Takalloozadeh and Yoon (2017) use the MMC method to design minimum compliance structures subject to a

uniform temperature change. Sharma (2017) minimizes the mismatch between the resulting displacement and a prescribed one for both steady-state and unsteady heat conduction (in the latter case, the prescribed displacement is specified at a given time). The method of Lohan et al. (2017) designs 2D structures for heat conduction in what amounts to an MMOS method. The structure, which has higher conductivity than its surrounding medium, is defined as the union of circles, and an element pseudo-density of 0 or 1 is assigned depending on whether the centroid is outside or inside of the structure, respectively. A gradient-free space colonization algorithm is used for the optimization. The work in Yu et al. (2019) addresses thermal-fluid problems with the MMC method to design the pipes in a cooling device, with the objective function being a weighted sum of the thermal compliance and the power dissipation of the system.

Finally, in the area of electromagnetics, the work of Li et al. (2019) designs a tree-like power distribution network on an integrated circuit by minimizing the power mean voltage subject to a volume constraint. Liu and Du (2019) use MMC to solve an inverse problem, namely to reconstruct the shapes of objects inside a body from electrical impedance measurements made on the surface of the body.

9 Discussion

Feature-mapping is a powerful and promising approach in structural optimization. As detailed in this review, there is a wide range of choices to accomplish the two main ingredients of these methods, namely the mapping of high-level parameters onto a non-body-fitted mesh for analysis and the combination of features.

We believe there is still ample room for further exploring the aspects of feature-mapping and to advance novel techniques, which will likely happen based on applications of feature-mapping to solve problems not (or not as efficiently) solvable by established methods, notably density-based and level-set topology optimization.

One area that we identified for potential further work is to consider more problems where high-level geometric constraints are essential for the application. For example, this could include manufacturing constraints and problems where connectivity between two or more points is essential.

An often discussed benefit of feature-mapping approaches is the potentially straightforward transfer of optimized designs to the CAD environment (due to use of a high-level geometric parameterization). However, there is little evidence to demonstrate this benefit in the literature.

Topology optimization using solid components allows the possibility of having components with different functionalities and/or different physical models that are simultaneously considered (Guo et al. 2014). However, to our knowledge, this ability has not yet been demonstrated or exploited—e.g., to solve problems driven by complex multi-disciplinary interactions.

The issue of local minima and initial design-dependent solutions is currently one of the key challenges of feature-mapping methods, particularly for hybrid methods and topology optimization. Some remedies to this problem are suggested in Section 8.7. However, we consider there is scope for new methods and ideas to solve (or at least mitigate) this issue.

Feature-mapping methods for topology optimization have mostly considered Boolean unions of components or holes. While the use of the Boolean union in conjunction with other operations (e.g., intersection and subtraction) was already conceived in the pioneering work of Mei et al. (2008), this has not been fully exploited in methods for topology optimization (with the exceptions of the methods that employ hollow components and coatings noted in Section 8.9). Hybrid methods, on the other hand, have demonstrated this possibility. This is an important capability because it allows to obtain more complex shapes (albeit still with high-level parameters), and because it mirrors the way in which CAD systems construct complex geometries. We thus believe there is room for topology optimization methods to further incorporate and demonstrate this capability.

Finally, in preparing this review we have noticed that there is a level of reinvention of methods and techniques. We hope this review helps researchers understand the different aspects of feature-mapping and the various techniques proposed to date in order to avoid duplication and to duly credit existing work.

Acknowledgments We thank Dr. Lukas Pflug from the Department of Mathematics at the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany, for fruitful discussion and support.

The initiative for this review goes back to critical yet constructive comments by Prof. Kurt Maute, from the University of Colorado Boulder, USA.

We also thank Prof. Horea Ilies from the University of Connecticut, USA, for guidance and insight into some of the geometric aspects of this work.

The first author acknowledges support by Deutsche Forschungsgemeinschaft (DFG) in the framework of the collaborative research center CRC 814 (subproject C2). The third author thanks the support of the US National Science Foundation, award CMMI-1634563.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Replication of results This review paper does not introduce any new methodology, and to replicate results of the different feature-mapping techniques discussed here, we refer the reader to the original works. The code to solve the simple test problems of Sections 3 and 4.4 is available by request from the authors.

References

Adalsteinsson D, Sethian JA (1999) The fast construction of extension velocities in ss methods. J Comput Phys 148(1):2–22

Bai J, Zuo W (2020) Hollow structural design in topology optimization via moving morphable component method. Struct Multidiscip Optim 61(1):187–205

Bakhtiarinejad M, Lee S, Joo J (2017) Component allocation and supporting frame topology optimization using global search algorithm and morphing mesh. Struct Multidiscip Optim 55(1):297–315

Bell B, Norato J, Tortorelli D (2012) A geometry projection method for continuum-based topology optimization of structures. In: 12th AIAA Aviation Technology, Integration, and Operations (ATIO) conference and 14th AIAA/ISSMO multidisciplinary analysis and optimization conference. https://doi.org/10.2514/6.2012-5485

Belytschko T, Black T (1999) Elastic crack growth in finite elements with minimal remeshing. International Journal for Numerical Methods in Engineering 45(5):601–620

Belytschko T, Gracie R, Ventura G (2009) A review of extended/generalized finite element methods for material modeling. Modelling and Simulation in Materials Science and Engineering 17(4):043,001

Bendsøe MP (1989) Optimal shape design as a material distribution problem. Struct Multidiscip Optim 1:193–202

Bendsøe MP, Kikuchi N (1988) Generating optimal topologies in structural design using a homogenization method. Comput Methods Appl Mech Eng 71(2):197–224

Bendsøe MP, Sigmund O (1999) Material interpolation schemes in topology optimization. Archive of Applievd Mechanics 69(9):635–654

Bendsøe MP, Sigmund O (2003) Topology optimization: theory, method and applications, 2nd edn. Springer

Bloomenthal J, Wyvill B (1990) Interactive techniques for implicit modeling. In: ACM SIGGRAPH Computer graphics, vol 24. ACM, pp 109–116

Braibant V, Fleury C (1984) Shape optimal design using B-splines. Comput Methods Appl Mech Eng 44(3):247–267

Bujny M, Aulig N, Olhofer M, Duddeck F (2018) Identification of optimal topologies for crashworthiness with the evolutionary level set method. International Journal of Crashworthiness 23(4):395–416

Cai S, Zhang W (2015) Stress constrained topology optimization with free-form design domains. Comput Methods Appl Mech Eng 289:267–290

Chen J, Shapiro V, Suresh K, Tsukanov I (2007) Shape optimization with topological changes and parametric control. Int J Numer Methods Eng 71(3):313–346

Cheng G, Mei Y, Wang X (2006) A feature-based structural topology optimization method. In: IUTAM Symposium on Topological Design Optimization of Structures, Machines and Materials. Springer, pp 505–514

Cheng L, Liu J, Liang X, To AC (2018) Coupling lattice structure topology optimization with design-dependent feature evolution for additive manufactured heat conduction design. Comput Methods Appl Mech Eng 332:408–439

Chu S, Gao L, Xiao M, Li H (2019) Design of sandwich panels with truss cores using explicit topology optimization. Compos Struct 210:892–905

Coniglio S (2019) Optimisation topologique à formalisme Eulérien et Lagrangien appliquée à la conception d'un ensemble propulsif. PhD thesis, Université de Toulouse

Coniglio S, Morlier J, Gogu C, Amargier R (2019) Generalized geometry projection: a unified approach for geometric feature based topology optimization. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-019-09362-8

Cui T, Sun Z, Liu C, Li L, Cui R, Guo X (2020) Topology optimization of plate structures using plate element-based moving morphable component (MMC) approach. Acta Mechanica Sinica 36(2):412–421

Deng H, To AC (2020) Linear and nonlinear topology optimization design with projection-based ground structure method (P-GSM). Int J Numer Methods Eng 121(11):2437–2461

Deng J, Chen W (2016) Design for structural flexibility using connected morphable components based topology optimization. Sci China Technol Sci 59(6):839–851

Deng J, Pedersen CB, Chen W (2019) Connected morphable components-based multiscale topology optimization. Front Mech Eng 14(2):129–140

van Dijk NP, Maute K, Langelaar M, Van Keulen F (2013) Level-set methods for structural topology optimization: a review. Struct Multidiscip Optim 48(3):437–472

Du B, Yao W, Zhao Y, Chen X (2019) A moving morphable voids approach for topology optimization with closed B-splines. J Mech Design 141(8):081,401

Du B, Zhao Y, Yao W, Wang X, Huo S (2020) Multiresolution isogeometric topology optimisation using moving morphable voids. Comput Model Eng Sci 122(3):1119–1140

Dunning PD (2018) Minimum length-scale constraints for parameterized implicit function based topology optimization. Structural and Multidisciplinary Optimization 58(1):155–169. https://doi.org/10.1007/s00158-017-1883-1

Eschenauer HA, Kobelev VV, Schumacher A (1994) Bubble method for topology and shape optimization of structures. Struct Optim 8(1):42–51

Gai Y, Zhu X, Zhang YJ, Hou W, Hu P (2020) Explicit isogeometric topology optimization based on moving morphable voids with closed b-spline boundary curves. Struct Multidiscip Optim 61(3):963–982

Gao HH, Zhu JH, Zhang WH, Zhou Y (2015) An improved adaptive constraint aggregation for integrated layout and topology optimization. Comput Methods Appl Mech Eng 289:387–408

Garcia MJ, Gonzalez CA (2004) Shape optimisation of continuum structures via evolution strategies and fixed grid finite element analysis. Struct Multidiscip Optim 26(1-2):92–98

García-Ruíz M, Steven G (1999) Fixed grid finite elements in elasticity problems. Eng Comput 16(2):145–164. https://doi.org/10.1108/02644409910257430

Guest JK, Zhu M (2012) Casting and milling restrictions in topology optimization via projection-based algorithms. In: ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp 913–920

Guo X, Ni C, Cheng G, Du Z (2012) Some symmetry results for optimal solutions in structural optimization. Struct Multidiscip Optim 46(5):631–645

Guo X, Du Z, Cheng G, Ni C (2013) Symmetry properties in structural optimization: some extensions. Struct Multidiscip Optim 47(6):783–794

Guo X, Zhang W, Zhong W (2014) Doing topology optimization explicitly and geometrically - a new moving morphable components based framework. J Appl Mech 81(8):081,009

Guo X, Zhang W, Zhang J, Yuan J (2016) Explicit structural topology optimization based on moving morphable components (MMC) with curved skeletons. Comput Methods Appl Mech Eng 310:711–748

Guo X, Zhou J, Zhang W, Du Z, Liu C, Liu Y (2017) Self-supporting structure design in additive manufacturing through explicit topology optimization. Comput Methods Appl Mech Eng 323:27–63

Ha SH, Guest JK (2014) Optimizing inclusion shapes and patterns in periodic materials using discrete object projection. Struct Multidiscip Optim 50(1):65–80

Haftka RT, Grandhi RV (1986) Structural shape optimization: a survey. Comput Methods Appl Mech Eng 57(1):91–106

Hansbo A, Hansbo P (2004) A finite element method for the simulation of strong and weak discontinuities in solid mechanics. Comput Methods Appl Mech Eng 193(33-35):3523–3540

Haslinger J, Mäkinen R (2003) Introduction to shape optimization: theory, vol 7. SIAM

Hoang VN, Jang GW (2017) Topology optimization using moving morphable bars for versatile thickness control. Comput Methods Appl Mech Eng 317:153–173

Hoang VN, Nguyen NL, Nguyen-Xuan H (2020a) Topology optimization of coated structure using moving morphable sandwich bars. Struct Multidiscip Optim 61(2):491–506

Hoang VN, Nguyen NL, Tran P, Qian M, Nguyen-Xuan H (2020b) Adaptive concurrent topology optimization of cellular composites for additive manufacturing. JOM 72(6):2378–2390

Hou W, Gai Y, Zhu X, Wang X, Zhao C, Xu L, Jiang K, Hu P (2017) Explicit isogeometric topology optimization using moving morphable components. Comput Methods Appl Mech Eng 326:694–712

Imam MH (1982) Three-dimensional shape optimization. Int J Numer Methods Eng 18(5):661–673

Kang Z, Wang Y (2013) Integrated topology optimization with embedded movable holes based on combined description by material density and level sets. Comput Methods Appl Mech Eng 255:1–13

Kasolis F, Wadbro E, Berggren M (2012) Fixed-mesh curvature-parameterized shape optimization of an acoustic horn. Struct Multidiscip Optim 46(5):727–738

Kazemi H, Vaziri A, Norato JA (2018) Topology optimization of structures made of discrete geometric components with different materials. J Mech Design 140(11):111,401

Kazemi H, Vaziri A, Norato J (2019) Topology optimization of multi-material lattices for maximal bulk modulus. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol 59186. American Society of Mechanical Engineers, pp V02AT03A052

Kazemi H, Vaziri A, Norato JA (2020) Multi-material topology optimization of lattice structures using geometry projection. Comput Methods Appl Mech Eng 363:112,895

Kim DH, Lee SB, Kwank BM, Kim HG, Lowther DA (2008) Smooth boundary topology optimization for electrostatic problems through the combination of shape and topological design sensitivities. IEEE Trans Magn 44(6):1002–1005

Lang C, Makhija D, Doostan A, Maute K (2014) A simple and efficient preconditioning scheme for Heaviside enriched XFEM. Comput Mech 54(5):1357–1374

Le C, Bruns T, Tortorelli D (2011) A gradient-based, parameter-free approach to shape optimization. Comput Methods Appl Mech Eng 200(9):985–996

Lee S, Kwak BM (2008) Smooth boundary topology optimization for eigenvalue performance and its application to the design of a flexural stage. Eng Optim 40(3):271–285

Lee SB, Kwak BM, Kim IY (2007) Smooth boundary topology optimization using B-spline and hole generation. International Journal of CAD/CAM 7(1):11–20

Lei X, Liu C, Du Z, Zhang W, Guo X (2019) Machine learning-driven real-time topology optimization under moving morphable component-based framework. J Appl Mech 86(1):011,004

Li B, Liu H, Zheng S (2018) Multidisciplinary topology optimization for reduction of sloshing in aircraft fuel tanks based on SPH simulation. Struct Multidiscip Optim 58(4):1719–1736

Li B, Xuan C, Liu G, Hong J (2019) Generating constructal networks for area-to-point conduction problems via moving morphable components approach. J Mech Design 141(5):051,401

Li L, Wang MY, Wei P (2012) XFEM Schemes for level set based structural optimization. Frontiers of Mechanical Engineering 7(4):335–356

Li Y, Wei P, Ma H (2017) Integrated optimization of heat-transfer systems consisting of discrete thermal conductors and solid material. Int J Heat Mass Transfer 113:1059–1069

Lian R, Jing S, He Z, Shi Z (2020) Geometric boundary feature extraction method based on moving morphable components (MMC) for topology optomization results. In: 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, vol 1, pp 2299–2303

Lin HY, Rayasam M, Subbarayan G (2015) ISOCOMP: Unified geometric and material composition for optimal topology design. Struct Multidiscip Optim 51(3):687–703

Liu C, Zhu Y, Sun Z, Li D, Du Z, Zhang W, Guo X (2018) An efficient moving morphable component (MMC)-based approach for multi-resolution topology optimization. Struct Multidiscip Optim 58(6):2455–2479

Liu D, Du J (2019) A moving morphable components based shape reconstruction framework for electrical impedance tomography. IEEE Trans Med Imag 38(12):2937–2948

Liu J, Ma Y (2016) A survey of manufacturing oriented topology optimization methods. Adv Eng Softw 100:161–175

Liu J, Ma YS (2015) 3d level-set topology optimization: a machining feature-based approach. Structural and Multidisciplinary Optimization 52(3):563–582

Liu J, Gaynor AT, Chen S, Kang Z, Suresh K, Takezawa A, Li L, Kato J, Tang J, Wang CC et al (2018b) Current and future trends in topology optimization for additive manufacturing. Struct Multidiscip Optim 57(6):2457–2483

Liu P, Kang Z (2018) Integrated topology optimization of multi-component structures considering connecting interface behavior. Comput Methods Appl Mech Eng 341:851–887

Liu T, Wang S, Li B, Gao L (2014) A level-set-based topology and shape optimization method for continuum structure under geometric constraints. Struct Multidiscip Optim 50(2):253–273

Lohan DJ, Dede EM, Allison JT (2017) Topology optimization for heat conduction using generative design algorithms. Struct Multidiscip Optim 55(3):1063–1077

Makhija D, Maute K (2014) Numerical instabilities in level set topology optimization with the extended finite element method. Struct Multidiscip Optim 49(2):185–197

Mei Y, Wang X, Cheng G (2008) A feature-based topological optimization for structure design. Adv Eng Softw 39(2):71–87

Moës N, Dolbow J, Belytschko T (1999) A finite element method for crack growth without remeshing. International Journal for Numerical Methods in Engineering 46(1):131–150

Moës N, Cloirec M, Cartraud P, Remacle JF (2003) A computational approach to handle complex microstructure geometries. Comput Methods Appl Mech Eng 192(28-30):3163–3177

Najafi AR, Safdari M, Tortorelli DA, Geubelle PH (2015) A gradient-based shape optimization scheme using an interface-enriched generalized fem. Comput Methods Appl Mech Eng 296:1–17

Nguyen TH, Paulino GH, Song J, Le CH (2010) A computational paradigm for multiresolution topology optimization (MTOP). Struct Multidiscip Optim 41(4):525–539

Niu B, Wadbro E (2019) On equal-width length-scale control in topology optimization. Struct Multidiscip Optim 59(4):1321–1334

Noël L, Duysinx P (2017) Shape optimization of microstructural designs subject to local stress constraints within an XFEM-level set framework. Struct Multidiscip Optim 55(6):2323–2338

Noël L, Miegroet LV, Duysinx P (2016) Analytical sensitivity analysis using the extended finite element method in shape optimization of bimaterial structures. Int J Numer Methods Eng 107(8):669–695

Norato J, Haber R, Tortorelli D, Bendsøe MP (2004) A geometry projection method for shape optimization. International Journal for Numerical Methods in Engineering 60(14):2289–2312

Norato J, Bell B, Tortorelli D (2015) A geometry projection method for continuum-based topology optimization with discrete elements. Comput Methods Appl Mech Eng 293:306–327

Norato JA (2018) Topology optimization with supershapes. Struct Multidiscip Optim 58(2):415–434

Overvelde JT (2012) The moving node approach in topology optimization. Master's thesis, Delft University of Technology

Pollini N, Amir O (2020) Mixed projection-and density-based topology optimization with applications to structural assemblies. Struct Multidiscip Optim 61(2):687–710

Qian X (2013) Topology optimization in B-spline space. Comput Methods Appl Mech Eng 265:15–35

Qian Z, Ananthasuresh G (2004) Optimal embedding of rigid objects in the topology design of structures. Mech Based Design Struct Mach 32(2):165–193

Raponi E, Bujny M, Olhofer M, Aulig N, Boria S, Duddeck F (2019) Kriging-assisted topology optimization of crash structures. Comput Methods Appl Mech Eng 348:730–752

Rozvany GI (2011) On symmetry and non-uniqueness in exact topology optimization. Struct Multidiscip Optim 43(3):297–317

Rozvany GI, Zhou M, Birker T (1992) Generalized shape optimization without homogenization. Struct Optim 4(3-4):250–252

Saxena A (2011) Are circular shaped masks adequate in adaptive mask overlay topology synthesis method? J Mech Design 133(1):011,001

Seo YD, Kim HJ, Youn SK (2010) Isogeometric topology optimization using trimmed spline surfaces. Comput Methods Appl Mech Eng 199(49-52):3270–3296

Shan P (2008) Optimal embedding objects in the topology design of structure. Master thesis, Dalian University of Technology, (in Chinese)

Shapiro V (2002) Solid modeling. Handbook of computer aided geometric design 20:473–518

Shapiro V (2007) Semi-analytic geometry with R-functions. ACTA numerica 16:239–303

Sharma A (2017) Advances in design and optimization using immersed boundary methods. Phd Thesis, University of Colorado Boulder

Sharma A, Villanueva H, Maute K (2017) On shape sensitivities with Heaviside-enriched XFEM. Struct Multidiscip Optim 55(2):385–408

Sharpe C, Seepersad CC, Watts S, Tortorelli D (2018) Design of mechanical metamaterials via constrained Bayesian optimization. In: ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp V02AT03A029–V02AT03A029

Sigmund O, Maute K (2013) Topology optimization approaches. Struct Multidiscip Optim 48(6):1031–1055

Sigmund O, Torquato S (1997) Design of materials with extreme thermal expansion using a three-phase topology optimization method. J Mech Phys Solids 45(6):1037–1067

Smith H, Norato JA (2020) A MATLAB code for topology optimization using the geometry projection method. Structural and Multidisciplinary Optimization. https://doi.org/10.1007/s00158-020-02552-0

Smith HA, Norato J (2019a) A geometry projection method for the design exploration of wing-box structures. In: AIAA Scitech 2019 forum, p 2353

Smith HA, Norato JA (2019b) Geometric constraints for the topology optimization of structures made of primitives. In: SAMPE Conference proceedings. Charlotte. https://doi.org/10.33599/nasampe/s.19.1518

Sokolowski J, Zolesio JP (1992) Introduction to shape optimization. In: Introduction to shape optimization. Springer, pp 5–12

Stegmann J, Lund E (2005) Discrete material optimization of general composite shell structures. Int J Numer Methods Eng 62(14):2009–2027

Stolpe M (2010) On some fundamental properties of structural topology optimization problems. Struct Multidiscip Optim 41(5):661–670

Stolpe M (2016) Truss optimization with discrete design variables: a critical review. Struct Multidiscip Optim 53(2):349–374

Stolpe M, Svanberg K (2001) An alternative interpolation scheme for minimum compliance topology optimization. Struct Multidiscip Optim 22(2):116–124

Sun J (2020) Topology optimization for removing internal resonances of a rotating thin plate. J Sound Vibr 480:115420. https://doi.org/10.1016/j.jsv.2020.115420

Sun J, Tian Q, Hu H (2018a) Topology optimization of a three-dimensional flexible multibody system via moving morphable components. Journal of Computational and Nonlinear Dynamics 13(2):021,010

Sun J, Tian Q, Hu H, Pedersen NL (2018b) Simultaneous topology and size optimization of a 3d variable-length structure described by the ale–ancf. Mech Mach Theory 129:80–105

Sun J, Tian Q, Hu H, Pedersen NL (2018c) Topology optimization of a flexible multibody system with variable-length bodies described by ALE–ANCF. Nonlinear Dynamics 93(2):413–441

Sun J, Tian Q, Hu H, Pedersen NL (2019) Topology optimization for eigenfrequencies of a rotating thin plate via moving morphable components. J Sound Vib 448:83–107

Svanberg K (1987) The method of moving asymptotes-a new method for structural optimization. Int J Numer Methods Eng 24(2):359–373

Tahhan M (2019) Topology optimization of space frames via geometry projection. Master's thesis, University of Connecticut

Takalloozadeh M, Yoon GH (2017) Implementation of topological derivative in the moving morphable components approach. Finite Elem Anal Des 134:16–26

Terada K, Asai M, Yamagishi M (2003) Finite cover method for linear and non-linear analyses of heterogeneous solids. International Journal for Numerical Methods in Engineering 58(9):1321–1346

Tröltzsch F (2010) Optimal control of partial differential equations: theory, methods, and applications, vol 112. American Mathematical Soc

Van Miegroet L, Duysinx P (2007) Stress concentration minimization of 2d filets using x-FEM and level set description. Struct Multidiscip Optim 33(4-5):425–438

Villanueva CH, Maute K (2014) Density and level set-XFEM schemes for topology optimization of 3-d structures. Comput Mech 54(1):133–150

Wall WA, Frenzel MA, Cyron C (2008) Isogeometric structural shape optimization. Comput Methods Appl Mech Eng 197(33):2976–2988

Wang F, Jensen JS, Sigmund O (2012) High-performance slow light photonic crystal waveguides with topology optimized or circular-hole based material layouts. Photonics and Nanostructures-Fundamentals and Applications 10(4):378–388

Wang F, Lazarov BS, Sigmund O, Jensen JS (2014a) Interpolation scheme for fictitious domain techniques and topology optimization of finite strain elastic problems. Comput Methods Appl Mech Eng 276:453–472

Wang MY, Zong H, Ma Q, Tian Y, Zhou M (2019a) Cellular level set in B-splines (CLIBS): a method for modeling and topology optimization of cellular structures. Comput Methods Appl Mech Eng 349:378–404

Wang N, Yang Y (2009) Structural design optimization subjected to uncertainty using fat Bézier curve. Comput Methods Appl Mech Eng 199(1-4):210–219

Wang R, Zhang X, Zhu B (2019b) Imposing minimum length scale in moving morphable component MMC-based topology optimization using an effective connection status (ECS) control method. Comput Methods Appl Mech Eng 351:667–693

Wang Y, Luo Z, Zhang X, Kang Z (2014b) Topological design of compliant smart structures with embedded movable actuators. Smart Materials and Structures 23(4):045,024

Watts S, Tortorelli DA (2017) A geometric projection method for designing three-dimensional open lattices with inverse homogenization. Int J Numer Methods Eng 112(11):1564–1588

Wei P, Wang MY, Xing X (2010) A study on x-FEM in continuum structural optimization using a level set model. Comput Aided Des 42(8):708–719

Wein F, Stingl M (2018) A combined parametric shape optimization and ersatz material approach. Struct Multidiscip Optim 57(3):1297–1315

Weiss BM, Hamel JM, Ganter MA, Storti DW (2018) Data-driven additive manufacturing constraints for topology optimization. In: ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol 2A. American Society of Mechanical Engineers, p V02AT03A031

Wormser M, Wein F, Stingl M, Körner C (2017) Design and additive manufacturing of 3d phononic band gap structures based on gradient based optimization. Materials 10(10):1125

Xia L, Zhu J, Zhang W (2012a) Sensitivity analysis with the modified Heaviside function for the optimal layout design of multi-component systems. Comput Methods Appl Mech Eng 241:142–154

Xia L, Zhu J, Zhang W (2012b) A superelement formulation for the efficient layout design of complex multi-component system. Struct Multidiscip Optim 45(5):643–655

Xia L, Zhu J, Zhang W, Breitkopf P (2013) An implicit model for the integrated optimization of component layout and structure topology. Comput Methods Appl Mech Eng 257:87–102

Xian Y, Rosen DW (2020) Morphable components topology optimization for additive manufacturing. Struct Multidiscip Optim 62(1):19–39. https://doi.org/10.1007/s00158-019-02466-6

Xie X, Wang S, Xu M, Wang Y (2018) A new isogeometric topology optimization using moving morphable components based on R-functions and collocation schemes. Comput Methods Appl Mech Eng 339:61–90

Xie X, Zheng H, Jonckheere S, Desmet W (2019) Explicit and efficient topology optimization of frequency-dependent damping patches using moving morphable components and reduced-order models. Comput Methods Appl Mech Eng 355:591–613

Xie X, Wang S, Xu M, Jiang N, Wang Y (2020a) A hierarchical spline based isogeometric topology optimization using moving morphable components. Comput Methods Appl Mech Eng 360:112,696

Xie X, Wang S, Ye M, Xia Z, Zhao W, Jiang N, Xu M (2020b) Isogeometric topology optimization based on energy penalization for symmetric structure. Front Mech Engi 15(1):100–122

Xue R, Liu C, Zhang W, Zhu Y, Tang S, Du Z, Guo X (2019) Explicit structural topology optimization under finite deformation via moving morphable void (MMV) approach. Comput Methods Appl Mech Eng 344:798–818

Yan S, Wang F, Sigmund O (2018) On the non-optimality of tree structures for heat conduction. Int J Heat Mass Transfer 122:660–680

Yazid A, Abdelkader N, Abdelmadjid H (2009) A state-of-the-art review of the x-FEM for computational fracture mechanics. Appl Math Model 33(12):4269–4282

Yu M, Ruan S, Wang X, Li Z, Shen C (2019) Topology optimization of thermal–fluid problem using the MMC-based approach. Struct Multidiscip Optim 60(1):151–165

Zhang J, Zhang W, Zhu J, Xia L (2012) Integrated layout design of multi-component systems using XFEM and analytical sensitivity analysis. Comput Methods Appl Mech Eng 245:75–89

Zhang S, Norato JA (2017) Optimal design of panel reinforcements with ribs made of plates. J Mech Design 139(8):081,403

Zhang S, Norato JA (2018) Finding better local optima in topology optimization via tunneling. In: ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp V02BT03A014–V02BT03A014

Zhang S, Norato JA, Gain AL, Lyu N (2016a) A geometry projection method for the topology optimization of plate structures. Struct Multidiscip Optim 54(5):1173–1190

Zhang S, Gain AL, Norato JA (2017a) Stress-based topology optimization with discrete geometric components. Comput Methods Appl Mech Eng 325:1–21

Zhang S, Gain AL, Norato JA (2018a) A geometry projection method for the topology optimization of curved plate structures with placement bounds. Int J Numer Methods Eng 114(2):128–146

Zhang S, Gain AL, Norato JA (2020a) Adaptive mesh refinement for topology optimization with discrete geometric components. Comput Methods Appl Mech Eng 364:112,930

Zhang W, Zhu J (2006) A new finite-circle family method for optimal multi-component packing design. WCCM VII, Los Angeles

Zhang W, Xia L, Zhu J, Zhang Q (2011) Some recent advances in the integrated layout design of multicomponent systems. J Mech Design 133(10):104,503

Zhang W, Zhong W, Guo X (2015) Explicit layout control in optimal design of structural systems with multiple embedding components. Comput Methods Appl Mech Eng 290:290–313

Zhang W, Li D, Zhang J, Guo X (2016b) Minimum length scale control in structural topology optimization based on the moving morphable components (MMC) approach. Comput Methods Appl Mech Eng 311:327–355

Zhang W, Yuan J, Zhang J, Guo X (2016c) A new topology optimization approach based on moving morphable components (MMC) and the ersatz material model. Struct Multidiscip Optim 53(6):1243–1260

Zhang W, Chen J, Zhu X, Zhou J, Xue D, Lei X, Guo X (2017b) Explicit three dimensional topology optimization via moving morphable void (MMV) approach. Comput Methods Appl Mech Eng 322:590–614

Zhang W, Li D, Yuan J, Song J, Guo X (2017c) A new three-dimensional topology optimization method based on moving morphable components (MMCs). Comput Mech 59(4):647–665

Zhang W, Yang W, Zhou J, Li D, Guo X (2017d) Structural topology optimization through explicit boundary evolution. J Appl Mech 84(1):011,011

Zhang W, Zhao L, Gao T, Cai S (2017e) Topology optimization with closed B-splines and Boolean operations. Comput Methods Appl Mech Eng 315:652–670

Zhang W, Zhou J, Zhu Y, Guo X (2017f) Structural complexity control in topology optimization via moving morphable component (MMC) approach. Struct Multidiscip Optim 56(3):535–552

Zhang W, Zhou Y, Zhu J (2017g) A comprehensive study of feature definitions with solids and voids for topology optimization. Comput Methods Appl Mech Eng 325:289–313

Zhang W, Li D, Zhou J, Du Z, Li B, Guo X (2018b) A moving morphable void (MMV)-based explicit approach for topology optimization considering stress constraints. Comput Methods Appl Mech Eng 334:381–413

Zhang W, Liu Y, Du Z, Zhu Y, Guo X (2018c) A moving morphable component based topology optimization approach for rib-stiffened structures considering buckling constraints. J Mech Design 140(11):111,404

Zhang W, Song J, Zhou J, Du Z, Zhu Y, Sun Z, Guo X (2018d) Topology optimization with multiple materials via moving morphable component (MMC) method. Int J Numer Methods Eng 113(11):1653–1675

Zhang W, Jiang S, Liu C, Li D, Kang P, Youn SK, Guo X (2020b) Stress-related topology optimization of shell structures using IGA/TSA-based moving morphable void (MMV) approach. Comput Methods Appl Mech Eng 366:113,036

Zhang W, Li D, Kang P, Guo X, Youn SK (2020c) Explicit topology optimization using IGA-based moving morphable void (MMV) approach. Comput Methods Appl Mech Eng 360:112,685

Zhou M, Wang MY (2013) Engineering feature design for level set based structural optimization. Comput Aided Des 45(12):1524–1537

Zhou Y, Zhang W, Zhu J, Xu Z (2016) Feature-driven topology optimization method with signed distance function. Comput Methods Appl Mech Eng 310:1–32

Zhou Y, Zhang W, Zhu J (2019) Concurrent shape and topology optimization involving design-dependent pressure loads using implicit B-spline curves. Int J Numer Methods Eng 118(9):495–518

Zhu B, Chen Q, Wang R, Zhang X (2018) Structural topology optimization using a moving morphable component-based method considering geometrical nonlinearity. J Mech Des 140(8):081,403

Zhu J, Zhang W, Beckers P, Chen Y, Guo Z (2008) Simultaneous design of components layout and supporting structures using coupled shape and topology optimization technique. Struct Multidiscip Optim 36(1):29–41

Zhu JH, Gao HH, Zhang WH, Zhou Y (2015) A multi-point constraints based integrated layout and topology optimization design of multi-component systems. Struct Multidiscip Optim 51(2):397–407

Zhu JH, Guo WJ, Zhang WH, Liu T (2017) Integrated layout and topology optimization design of multi-frame and multi-component fuselage structure systems. Struct Multidiscip Optim 56(1):21–45