



Research
Intelligent Manufacturing—Article

From Computer-Aided Design (CAD) Toward Human-Aided Design (HAD): An Isogeometric Topology Optimization Approach



Yingjun Wang^a, Mi Xiao^b, Zhaohui Xia^b, Peigen Li^b, Liang Gao^{b,*}

^aNational Engineering Research Center of Novel Equipment for Polymer Processing, The Key Laboratory of Polymer Processing Engineering of the Ministry of Education, Guangdong Provincial Key Laboratory of Technique and Equipment for Macromolecular Advanced Manufacturing, South China University of Technology, Guangzhou 510641, China
^bState Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China

ARTICLE INFO

Article history:

Received 16 March 2022

Revised 25 May 2022

Accepted 4 July 2022

Keywords:

Human-aided design

CAD/CAE integration

Isogeometric analysis

Structural optimization

Isogeometric topology optimization

ABSTRACT

In this paper, the novel design mode of human-aided design (HAD) is proposed to replace conventional computer-aided design (CAD). In HAD, computers can automatically complete the whole product design via a new isogeometric topology optimization (ITO), while humans just assist to slightly modify the design to meet requirements. An embedded domain ITO is presented to design complex models with irregular design domains, and editable geometric models of optimized results can be automatically generated based on layered ITO results. Three examples are tested to verify the proposed HAD mode, including a 3D cantilever beam with a regular design domain, an automotive part with an irregular design domain, and a Messerschmitt-Bölkow-Blohm (MBB) beam with a multiscale structure. The results demonstrate that the proposed HAD mode can automatically deliver high-quality optimized models; thus, it has great potential as a revolutionary technology to change the current design mode from CAD to HAD.

© 2022 © 2022 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The finite-element method (FEM), which is the computational core of computer-aided engineering (CAE), originated in the 1950s and 1960s, yet—notably—the computer-aided design (CAD) used to provide geometric models for the FEM appeared ten years later than the FEM. After nearly half a century of development, the techniques in CAD and CAE continue to mature. Nowadays, all kinds of commercial CAD and CAE software have emerged, such as Catia, Unigraphics NX (UG), Creo, Nastran, Abaqus, and ANSYS.

The advent of this software has given engineers numerous powerful new tools. As the Internet evolved, CAD/CAE migrated into the cloud; as a result, powerful design tools are now globally available and have become an integral part of the engineering landscape. However, humans still play the leading role during the whole product design, and computers are just auxiliary tools that help people design. In this design mode, humans use various tools to design and analyze a product to achieve the design requirements,

including geometric shapes and structural performance. To design a geometric shape, a CAD modeling tool (SolidWorks, Catia, Creo, etc.) is used, which has its own geometric representation such as boundary representation (B-rep) and construct solid geometry (CSG) [1]. To analyze structural performance, a CAE analysis tool (ANSYS, Abaqus, Nastran, etc.) is used, which is presented by discrete elements that are quite different from those in CAD modeling systems. Therefore, humans must manually convert the models between CAD modeling and CAE analysis, which may take more than 80% of the overall design time for complex models [2]. Moreover, current product design heavily depends on human experience; as a result, the product has to be iteratively designed until the final design is obtained (Fig. 1(a)). This iterative design process is time-consuming, and it is difficult for a design quality that is determined by the experience of the designer to reach an optimal performance. To reduce the human burden in product design, two issues need to be addressed: ① how to reduce the need for human interactions with CAD and CAE models, and ② how to automatically generate high-quality product design by computers.

Two alternative ways have been employed to reduce the need for human interactions with CAD and CAE models. One is software integration; for example, SolidWorks integrates the analysis

* Corresponding author.

E-mail address: gaoliang@mail.hust.edu.cn (L. Gao).

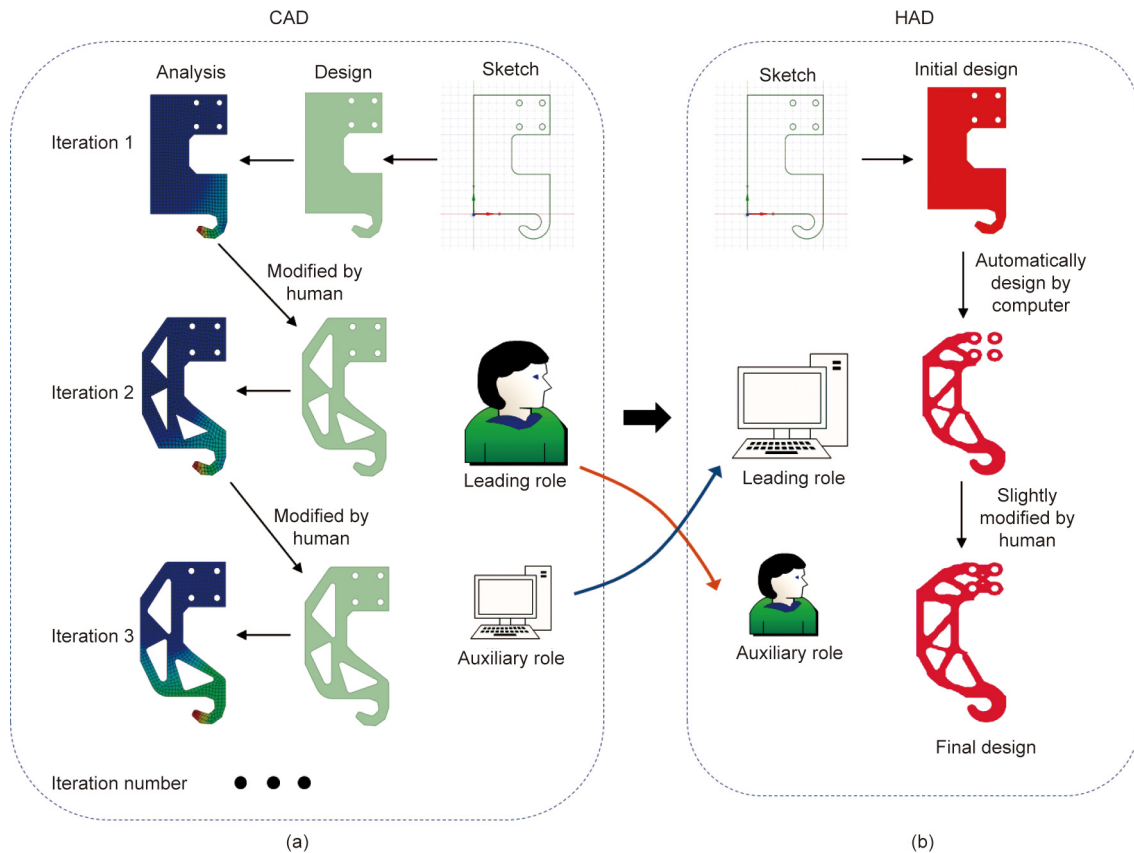


Fig. 1. A workflow illustration of CAD and HAD. (a) CAD; (b) HAD. HAD: human-aided design.

modules in COMSOL [3]. The other is to employ a data exchange interface, such as the data interface between Catia and Abaqus. Unfortunately, CAD and CAE models are still different in the aforementioned ways. To unify CAD and CAE models with the same geometric representation, Hughes et al. [4] proposed an isogeometric analysis (IGA) method in 2005, in which geometric and analytical models are described by the same mathematical representation. Compared with FEM, IGA has higher accuracy and efficiency and is now widely used in the fields of vibration, acoustics, fatigue, fluid-structure interaction, and shape and size optimization [5–8]. Therefore, the IGA method is used in this work to integrate CAD and CAE models.

To automatically generate high-quality product design by means of computers, intelligent structural optimization methods must be used. In the field of structural optimization, topology optimization (TO) aims to computationally find the optimal material distribution in a design domain under given constraints [9], such as stress and displacement [10,11]. With the help of TO, designers can obtain a high-performance design scheme automatically without a wealth of knowledge and experience. Since CAD and CAE models have different geometric representations in TO, the optimized results must be reconstructed to generate CAD models, which is complicated and time-consuming. To address this issue, the IGA is employed in TO to replace the FEM, in what is termed as isogeometric TO (ITO).

The concept of ITO was first proposed by Seo et al. [12], who used the spline surface and trimmed curves to construct a complex design domain without the patch coupling technique. Since then, certain branches of ITO have been rapidly developed, including density-based and level set-based ITO. Solid isotropic material with penalization (SIMP) is a popular and typical density-based

method that is easy to implement, and SIMP-based ITO has gained considerable research attention. Kumar and Parthasarathy [13] introduced B-spline finite elements to construct the density function and perform structural analysis. Hassani et al. [14] employed IGA instead of FEM in TO, where the material density is interpolated by the non-uniform rational B-spline (NURBS) basis function and can be viewed as continuous within the whole design domain. Qian [15] restricted the design space to the B-spline space and embedded an arbitrarily shaped design domain into a rectangular domain, where the B-splines are used to construct the density field. Xie et al. [16] used a truncate hierarchical B-spline to locally refine the mesh and reduce the computational cost significantly in ITO. Multi-material and multiscale problems can also be solved within the framework of ITO. To gain multi-material structures, Taheri and Suresh [17] proposed a new density-based ITO, in which the exact geometrical model is constructed by NURBS. Gao et al. [18] introduced a NURBS-based multi-material interpolation (N-MMI) model and developed a multi-material ITO (M-ITO) method. Wang et al. [19] investigated a multiscale ITO to construct a porous structure with a lattice infill. Lieu and Lee [20,21] proposed the addition of a novel multiresolution scheme to ITO and successfully applied it to multi-material problems; this approach yielded high-resolution designs with a lower computational cost.

Unlike SIMP-based ITO, level-set-based ITO employs a higher dimensional level set function (LSF) to achieve the implicit description of structural geometries. Shojaee et al. [22] combined the level set method (LSM) based on the radial basis function with IGA for TO. Then, Wang and Benson [23] proposed an innovative ITO by interpolating the LSF with NURBS shape functions, where the same NURBS basis functions are used to integrate CAD, CAE, and TO models. In addition, Wang and Benson [24] introduced the fast

point-in-polygon algorithm and trimmed elements in the framework of level-set-based ITO to provide higher accuracy and efficiency. Ghasemi et al. [25] implemented a novel design method for flexoelectric materials, in which IGA is combined with level set and pointwise density mapping techniques. Subsequently, Ghasemi et al. [26] further extended the method to the multi-material level-set-based TO of flexoelectric composites, successfully dealing with coupled problems and microstructures. Lee et al. [27] developed an isogeometric topological shape optimization method with the dual evolution of NURBS curves and level sets, where the topological variation of NURBS curves is ensured by the implicit LSF.

Although ITO has great potential for integrating CAD modeling, CAE analysis, and structural optimization, it is still restricted by the bottlenecks of IGA and TO—that is, the issues of how to deal with complex models with irregular geometries and how to automatically generate editable CAD models of TO results. To break through the limitations of conventional ITO for irregular geometric shapes, we herein propose a new ITO for a complex design domain based on the embedded domain method that can implement the IGA of an irregular model in a regular embedded domain. To obtain editable CAD models for ITO results, an automatic generation method is developed by using standard CAD operations to construct NURBS surfaces based on the contours layer by layer, where the contour of each layer is directly represented by the same control points used in modeling, analysis, and TO.

Based on the new ITO proposed in this paper, we propose a new design mode, human-aided design (HAD), based on the new ITO. In HAD, computers play the leading role and can automatically complete the whole product design with high quality; humans only need to slightly modify the design to meet requirements (Fig. 1(b)). Similar to revolutionary intelligent technologies such as smart homes [28], smart factories [29], and automatic driving [30], which greatly reduce human operations and are gradually changing the role played by humans from a leading to an auxiliary role, HAD is an intelligent design mode that holds great potential for replacing CAD in the next industrial revolution.

To further describe HAD and the theory behind it, the remainder of this paper is organized as follows: In Section 2, the NURBS-based ITO is summarized. Section 3 presents the HAD and its implementation based on ITO. Three numerical examples are discussed in Section 4 to demonstrate HAD for both macroscale and multiscale structures. Finally, conclusions and future work are discussed in Section 5.

2. Summary of NURBS-based ITO

In IGA, the same NURBS functions are used to represent both CAD and CAE models. A standard NURBS curve consists of a knot vector in the ξ -axis Ξ , basis functions R_i^p (where p denotes the degree of the basis functions; i is the index of the basis functions), and the corresponding control points \mathbf{P}_i , which can be written as follows:

$$\text{Curve}(\xi) = \sum_{i=1}^n R_i^p(\xi) \mathbf{P}_i \quad (1)$$

where n is the number of the basis functions.

It should be emphasized that the basis functions R_i^p are defined by a series of positive weight factors ω_i and the recursive B-spline basis functions N_i^p , as below:

$$R_i^p(\xi) = \frac{N_i^p(\xi) \omega_i}{\sum_{i=1}^n N_i^p(\xi) \omega_i} \quad (2)$$

The high-dimensional cases can be obtained from Eq. (1) by adding the knot vectors in the η - and ζ -axis, which can be expressed as follows:

$$\text{Surface}(\xi, \eta) = \sum_{i=1}^n \sum_{j=1}^m R_{ij}^{p,q}(\xi, \eta) \mathbf{P}_{ij} \quad (3)$$

$$\text{Solid}(\xi, \eta, \zeta) = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l R_{ijk}^{p,q,r}(\xi, \eta, \zeta) \mathbf{P}_{ijk}$$

where m , j , and q are the number, the index, and the degree of the basis functions in η direction, respectively; l , k , and r are the number, the index, and the degree of the basis functions in ζ direction.

For the linear isotropic elastic problem, the equilibrium equation in the form of the Galerkin method can be written as follows:

$$\int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{v}^h)^T \mathbf{E} \boldsymbol{\varepsilon}(\mathbf{u}^h) d\Omega = \int_{\Omega} \mathbf{f} \mathbf{v}^h d\Omega + \int_{\Gamma_N} \mathbf{h} \mathbf{v}^h d\Gamma_N \quad (4)$$

where Ω and Γ_N denote the design domain and its Neumann boundaries, respectively; $\boldsymbol{\varepsilon}$ is the strain vector and \mathbf{E} is the elastic tensor matrix; \mathbf{f} and \mathbf{h} represent the body force and the prescribed boundary traction vectors, respectively; \mathbf{u}^h is the trial solution displacement and \mathbf{v}^h is the weight, and superscript h denotes \mathbf{u}^h and \mathbf{v}^h belong to H^1 space. \mathbf{g} denotes the prescribed boundary displacement vector. Considering a set \mathcal{A} that contains all NURBS basis functions and a subset \mathcal{B} containing the basis functions that are equal to 0 on the Dirichlet boundary, to solve Eq. (4), the i th element of trial solution displacement \mathbf{u}^h and weight \mathbf{v}^h should be interpolated by the basis functions with the control variables

$$\begin{cases} \mathbf{u}_i^h = \sum_{j \in \mathcal{A} \setminus \mathcal{B}} R_j \mathbf{g}_i + \sum_{i \in \mathcal{B}} R_i \mathbf{d}_i \\ \mathbf{v}_i^h = \sum_{j \in \mathcal{A} \setminus \mathcal{B}} R_j \mathbf{e}_i \end{cases} \quad (5)$$

where \mathbf{e} is the arbitrary value to hold for all $\mathbf{v}^h \in H^1$, and \mathbf{g} and \mathbf{d} denote the displacement on the Dirichlet boundary and the displacement, respectively. Substituting Eq. (4) into Eq. (3), the Galerkin form can be written into a more compact matrix form as follows:

$$\mathbf{K} \mathbf{d} = \mathbf{F} \quad (6)$$

where the elements of the stiffness matrix \mathbf{K} and the load vector \mathbf{F} can be expressed as

$$\begin{cases} K_{ij} = \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{R}_i)^T \mathbf{E} \boldsymbol{\varepsilon}(\mathbf{R}_j) d\Omega \\ F_j = \int_{\Omega} \mathbf{f} \mathbf{R}_j d\Omega + \int_{\Gamma_N} \mathbf{h} \mathbf{R}_j d\Gamma_N \end{cases} \quad (7)$$

As an example, the classical minimum compliance problem is given here to illustrate the ITO mathematical model. Considering a NURBS-based density field, the formulation of the minimum compliance problem can be expressed as follows:

$$\begin{aligned} & \text{find } \rho_{ijk} \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, m; k = 1, 2, \dots, l) \\ & \min J(\mathbf{u}, \rho) = \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u})^T \mathbf{E}(\rho) \boldsymbol{\varepsilon}(\mathbf{u}) d\Omega \\ & \text{s.t. } \begin{cases} G(\rho) = \frac{1}{|\Omega|} \int_{\Omega} \rho(\xi, \eta, \zeta) v_0 d\Omega - V_0 \leq 0 \\ \mathbf{a}(\mathbf{u}, \delta \mathbf{u}) = L(\delta \mathbf{u}), \mathbf{u}|_{\Gamma_N} = \mathbf{g}, \forall \delta \mathbf{u} \in H(\Omega) \\ 0 \leq \rho_{ijk} \leq 1 \end{cases} \end{aligned} \quad (8)$$

where ρ denotes the densities of control points; $J(\cdot)$ is the objective function; $G(\rho)$ is the volume constraint; v_0 is the volume fraction of an element, and V_0 denotes the maximum permissible material; $\mathbf{a}(\mathbf{u}, \delta \mathbf{u}) = L(\delta \mathbf{u})$ denotes the bilinear form of the equilibrium equation, and $\delta \mathbf{u}$ is the virtual displacement field in the kinematically admissible space $H(\Omega)$, in which \mathbf{a} is the strain energy, L is the work by external load. Compared with the classical SIMP-based TO, the density function ρ in ITO is interpolated by the NURBS basis functions of the control points as Eq. (9), which is similar to Eqs. (1) and (3).

$$\rho(\xi, \eta, \zeta) = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l R_{ijk}^{p,q,r}(\xi, \eta, \zeta) \Phi_{ijk} \quad (9)$$

In Eq. (9), $\Phi_{i,j,k}$ are the densities of the control points. Then, the sensitivities of the objective and constraint functions in Eq. (8) can be obtained by calculating the derivatives of objective and constraint functions with respect to the densities of control points, and design variables are updated by optimization methods, such as the optimality criterion (OC) [31] and the method of moving asymptotes (MMA) [32]. If the convergence condition is achieved, the optimized layout will be obtained.

3. HAD based on ITO

3.1. ITO for a complex design domain

To be compatible with CAD modeling systems, NURBS should be used in ITO; however, the tensor-product representation of NURBS impedes ITO when dealing with engineering models with complex initial design domains or geometric constraints [24]. The root cause of this problem is that representing complex geometric models using a complete NURBS patch is quite difficult, and 3D splines do not exist in CAD modeling systems due to the B-rep of CAD models, which makes an IGA-suitable model unavailable. Remarkable analysis mesh-independent density interpolation techniques have been used to construct continuous density fields for irregular design domain [33,34] and have the potential to implement ITO for complex design domains; however, they separate the density field and the physical field. To better solve this problem, a new IGA we recently proposed [35] is used to implement an analysis of complex models. This IGA is based on an embedded domain and is capable of analyzing complex models with arbitrary geometries. In this section, we introduce two key issues of IGA based on an embedded domain: ① how to identify element types in the embedded domain, and ② how to complete the numerical integration of trimmed elements.

In IGA based on an embedded domain, an irregular model can be embedded into a regular domain, so that a simple unfitted structural mesh can be used to approximate the solution field to avoid generating boundary-conforming meshes. As shown in Fig. 2, the elements of the embedded domain are divided into three types: ① fictitious elements, ② solid elements, and ③ trimmed elements.

By computing the sign minimum distance (SMD) from the vertices of the embedded domain mesh to the CAD model boundary (e.g., the SMDs of points I and J represented by d_I and d_J in Fig. 3(a)), all SMDs of the vertices will be obtained and can be used as LSF values to construct an implicit representation of the CAD model. In the implicit representation, the CAD model boundary is represented by the zero level set isosurface. In this work, the SMD (i.e., the LSF value) of a vertex in the CAD model is defined

as a positive, and that of a vertex outside of the CAD model is defined as a negative. Therefore, the elements of the embedded domain mesh can be divided into solid elements, trimmed elements, and fictitious elements, as shown in Fig. 3(b), according to the following identification rule: ① If all SMDs of the vertices are positive, then the element is a solid element; ② if the SMDs of the vertices have both positive and negative values, then the element is a trimmed element; and ③ if all SMDs of the vertices are negative, then the element is a fictitious element. A highly efficient element classification method is proposed in Ref. [35], which only needs to compute the LSF values of the vertices of trimmed elements.

Based on the abovementioned implicit representation, the numerical integration of trimmed elements can be implemented by integrating the Gauss points in the solid domain of trimmed elements. Using the LSF values of the vertices, the LSF values of the Gauss points can be calculated by interpolation, as follows:

$$\phi = \sum_A N_A \phi_A \tag{10}$$

where ϕ is the LSF value of a Gauss point, and N_A and ϕ_A are the shape function and LSF value of the A th vertex of the corresponding embedded domain element (note that linear interpolation is used in this work, since only the SMDs at the vertices of the embedded domain mesh are computed; a more accurate interpolation rule can be used when more points are used to compute the SMDs). If the LSF value is not negative, the Gauss point belonging to the solid domain should be retained. Otherwise, the Gauss point is in the fictitious domain and should be removed. To increase integral accuracy, an adaptive subdivision method is used, which subdivides a trimmed element into four sub-elements (2D problem) or eight sub-elements (3D problem) to increase the number of Gauss points. This subdivision, which can be hierarchically implemented until the integral accuracy meets the requirement, has been successfully used to deal with the integration of trimmed elements in IGA [36,37]. An illustration of Gauss point generation for an embedded domain is shown in Fig. 4.

Based on the above element classification and trimmed element integration, a complex model with irregular geometry can be analyzed in an embedded domain by IGA. Using IGA based on an embedded domain, the ITO for complex design domains can be implemented in the same way as for simple regular domains.

3.2. Automatic generation of editable models of the ITO results

After the ITO for a complex design domain is implemented, CAD models of the design results need to be generated. Because the result models of TO often need to be further edited to adapt to

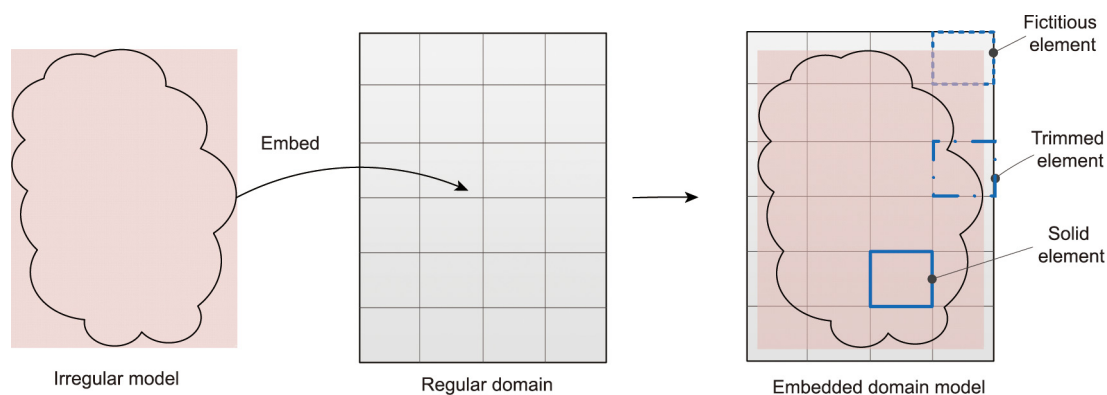


Fig. 2. Illustration of an embedded domain model.

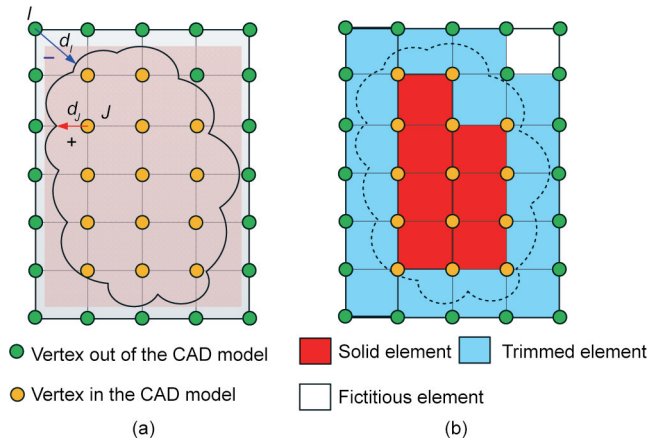


Fig. 3. SMDs of vertices and element type classification. (a) SMDs of vertices; (b) element types of the embedded domain mesh.

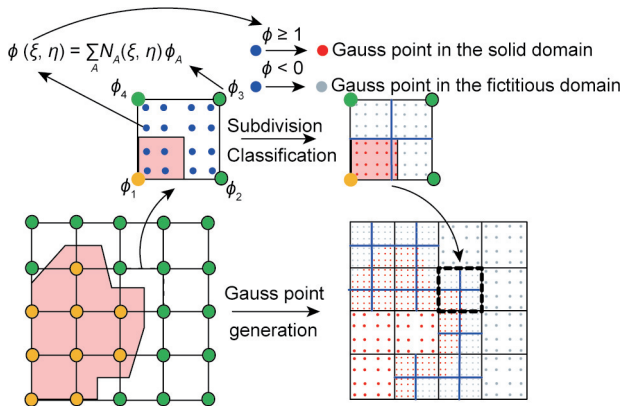


Fig. 4. Gauss point generation for an embedded domain.

downstream applications (e.g., manufacturing, prototyping, validation, and (design) exploration [38]), editable CAD models of TO results are necessary in industry. An automatic generation method is proposed for editable models based on ITO, which innovatively combines a B-rep-based CAD construction method with ITO and generates selectable and editable B-rep CAD models of ITO results.

In order to make full use of the advantages of unified geometric representation in ITO, the design variables are used as the high-dimensional coordinates of the control points. Thus, ITO results in 2D problems are represented by 3D NURBS surfaces (e.g., $S(x, y, \rho)$), where the first two dimensions correspond to the 2D coordinates of the design domain, and the third dimension represents the values of the design variables. In 3D ITO problems, the results are represented by 4D hypersurfaces (e.g., $HS(x, y, z, \rho)$). Next, a series of NURBS surfaces are extracted from the hypersurfaces layer by layer—for example, a series of $S(x, y, \rho) = HS(x, y, z = z_\beta, \rho)$ is extracted, where z_β is the z coordinate of the control points at the β th layer (the extraction can also be implemented in the x or y direction). In this way, the 3D ITO results can be represented by several layers of NURBS surfaces.

Inspired by several mature geometric algorithms (e.g., the surface/plane intersection (SPI) algorithm, surface skinning algorithm, and plane trimming algorithm) and the post-processing methods of Costa et al. [39,40], three geometric algorithms and post-processing processes are integrated into the same system to process the layered ITO results. For 2D problems, the traditional method of extracting iso-curves is replaced by the SPI algorithm,

and the results are represented in a trimmed plane form. For 3D problems, a new procedure from the definition of ITO to the construction of 3D B-rep CAD geometry is shown in Fig. 5. The overall construction procedure can be divided into five steps:

Step 1: ITO problem solving. Here, it is assumed that 3D ITO problems use hexahedral grids as their control mesh. After the standard ITO procedure, the resulting 4D hypersurface is replaced by several NURBS surfaces by regarding the design variables as high-dimensional coordinates, as follows:

$$HS(x_1, x_2, x_3, \rho) = \cup S_\beta(x_1, x_2, \rho), \quad \beta = 1, 2, \dots, N \quad (11)$$

where x_1, x_2 , and x_3 represent the coordinates of the control points; N is the number of layers in the x_3 direction, and S_β is the surface of the β th layer along the x_3 direction, which can be stored in standard CAD data (e.g., IGS file).

Step 2: construction of contour curves. In this step, closed contour curves are extracted from each surface layer. Similar to the method of intercepting iso-curves, the NURBS surfaces of each layer are intersected with a threshold plane at a certain altitude between 0 and 1. The altitude value of the threshold plane is chosen in such a way that the optimization constraints (e.g., the volume constraint) are met, where dichotomy can be used to determine the altitude. The closed contour curves can be represented as follows:

$$C_\beta(x_1, x_2) = S_\beta(x_1, x_2, \rho) \cap P_\beta(x_3 = \rho_t) \quad (12)$$

where C_β is the closed contour curve of the β th layer along the x_3 direction, P_β is the threshold plane whose altitude is $x_3 = \rho_t$, and ρ_t is the threshold density.

The SPI algorithm is a mature and accurate geometric algorithm [41,42]. A marching method, which has proved to be a widespread and effective method [43], is used to solve the SPI problem. The output of the contour curves extraction using the SPI algorithm is several sequences of closed NURBS curves at different altitudes.

Step 3: construction of the lateral surfaces. In this step, the lateral surfaces $LS(\cdot)$ of the B-rep optimized model are generated from several sequences of NURBS contour curves called section curves, as follows:

$$LS(x_1, x_2, x_3) = \int C_\beta(x_1, x_2) dx_3 \quad (13)$$

where the lateral surface can be regarded as the integration of section curves along the x_3 direction.

A surface skinning algorithm is used in this step to generate the NURBS information of the lateral surfaces. Skinning is a process by which a surface is formed by a set of curves [44,45]. This is actually the lofting operation in CAD systems, which was widely used in the automotive and aerospace industries decades ago [46]. The input curves in this step should meet certain conditions; for example, curves corresponding to each other should be in the same direction, and the topology of the section curves should be consistent at different heights. The gained section curves are used as input data for the surface skinning procedure. Finally, the external and internal lateral surfaces in the NURBS representation are obtained as the results of this step.

In addition, one contour may correspond to two or more contours in the adjacent layer, which will lead to a topology change. However, as long as the distance between adjacent layers is small enough, the lateral surfaces between the layers can be constructed by means of a surface extruding procedure using one layer's contour curves. Moreover, the accuracy of the entity model can be improved using a surface blending algorithm [47].

Step 4: construction of the top/bottom patch. To generate a B-rep solid model, each solid boundary is constructed. In this step, the top and bottom patches are generated using a plane trimming algorithm [48]. The input data of this step are the contour curves in

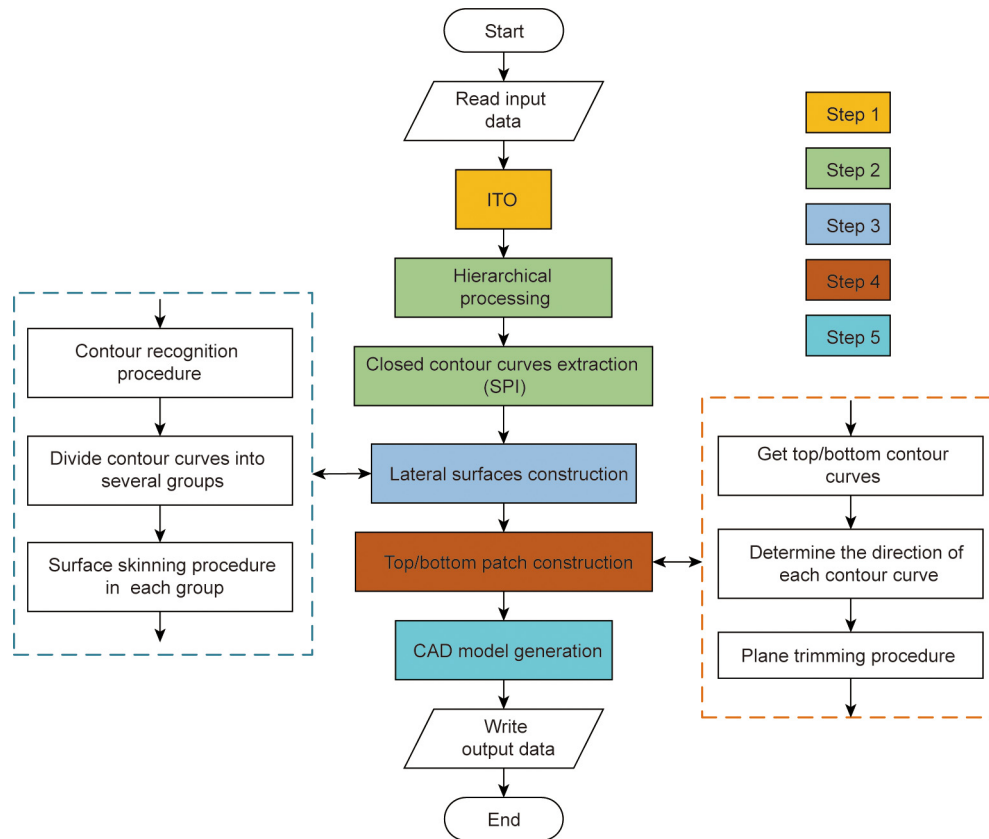


Fig. 5. The overall construction procedure of ITO result models.

the top and bottom layer and the layer’s altitude. The output data are the top and bottom patches, which are represented in trimmed plane form as follows:

$$P_{\text{output}} = P_{\text{input}}(A < C) \cup C \tag{14}$$

where P_{output} and P_{input} are the output and input patches, and $(A < C)$ denotes the area inside contour curve C .

Step 5: CAD model generation. The final CAD model is represented in B-rep form; it consists of the entire boundary vertices, edges, and faces. After the steps above, both lateral surfaces and top/bottom trimmed patches are generated. Thus, the final step is to combine the IGS files of every surface and patch, which can be done by modifying the IGS files.

In summary, the proposed method generates a selectable and editable CAD model based on the ITO results in five steps with the help of several geometry algorithms. It should be noted that the CAD model is an exact B-rep CAD model whose geometric information comes directly from the coordinates and densities of the control points.

3.3. Procedure of ITO-based HAD

In HAD, humans only need to define a design domain and requirements (i.e., the constraints of ITO); then, computers automatically generate an optimized CAD model with a high performance that meets all the requirements. In this study, the CAD model resulting from ITO is represented by NURBS, so it is editable and can be further modified by humans to meet the requirements.

The procedure for ITO-based HAD is shown in Fig. 6. Computers play the leading role in HAD and can complete most of the design

work without human interaction. In addition, computers iteratively optimize the structure by means of ITO to obtain the structure with the best performance. Unlike conventional CAD, which depends on human experience, the ITO is automatic and based on optimization theory. Since the same mathematical representation is used in the CAD modeling, CAE analysis, and ITO, the optimized result includes the geometric information so that an editable CAD model can be automatically generated using standard geometric algorithms. In HAD, human interaction is required in two main aspects: ① defining an optimization problem and constraints to meet the design requirements, and ② modifying optimized CAD models to further improve the product design. Therefore, humans play an auxiliary role, which is why we call this new design mode “HAD.”

4. Implementation and application of HAD

4.1. HAD with a regular design domain

In HAD, when the computer receives the boundary condition input of an optimization problem, it will automatically initiate the ITO program. A CAD model of the ITO results will be generated through the automatic generation method. A simple cantilever beam example is presented here to illustrate the automatic generation method and verify the effect of the method. The design domain and the boundary conditions of the cantilever beam are shown in Fig. 7(a). The design domain is divided into $40 \times 20 \times 20$ and $60 \times 30 \times 30$ quadratic 3D NURBS elements, respectively, the volume fraction (VF) of the structure is set at $\text{VF} = 0.3$, and the Young’s modulus and the Poisson’s ratio are 1 and 0.3, respectively. The input load $F = 1$

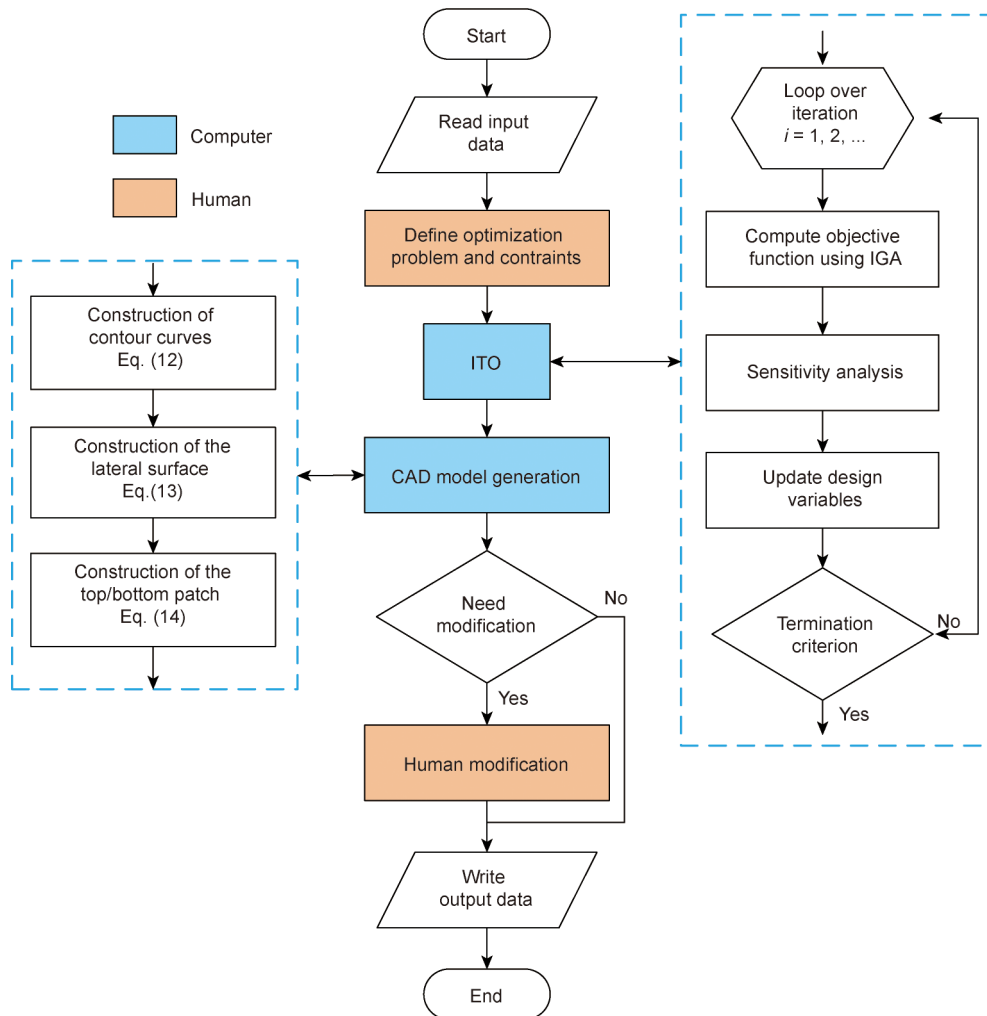


Fig. 6. The procedure flowchart of ITO-based HAD.

is applied at the center of the right face and downward vertically. A fixed constraint is applied to the left face, and the design objective is to achieve a structure with minimum compliance. A hexahedral 3D ITO control mesh is shown in Fig. 7(b); its scale is reduced to $14 \times 7 \times 7$ to avoid too many control points being shown in the figure.

Fig. 8 shows the optimized result structure of the 3D cantilever beam represented by elements. It can be seen that the optimized geometries are consistent for the two mesh scales, which demonstrates the mesh independency of the proposed ITO method. The figure only shows the elements whose control point density is above the density $\rho = 0.5$. Fig. 9 shows the iterative history of the 3D cantilever beam with $40 \times 20 \times 20$ elements. After 53 iterations in the ITO program, the minimum compliance of the cantilever beam is optimized to 29.86. Next, the NURBS information of the 3D cantilever beam is input into the automatic generation method to generate the CAD model.

A detailed procedure chart of the automatic generation method is shown in Fig. 10. The input of the method is the NURBS information (including the knot vector, control points' coordinates, and density), which is obtained from the ITO program. First, the density of the control points is extracted layer by layer to replace the 4D hypersurface result with multiple 3D NURBS surfaces. The selection of the control point layers

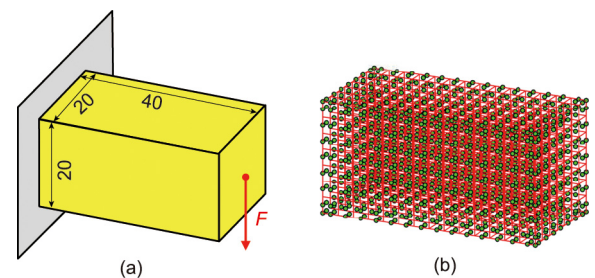


Fig. 7. The design domain of the 3D cantilever beam. (a) Boundary conditions; (b) 3D representation of the IGA control mesh. F : input load.

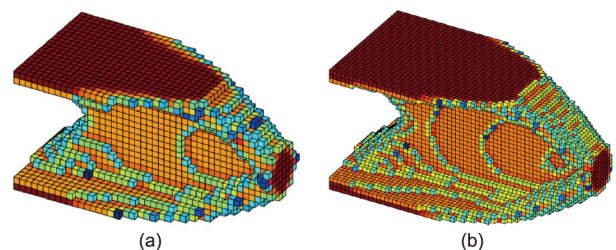


Fig. 8. The optimized structure of the 3D cantilever beam. (a) $40 \times 20 \times 20$ elements; (b) $60 \times 30 \times 30$ elements.

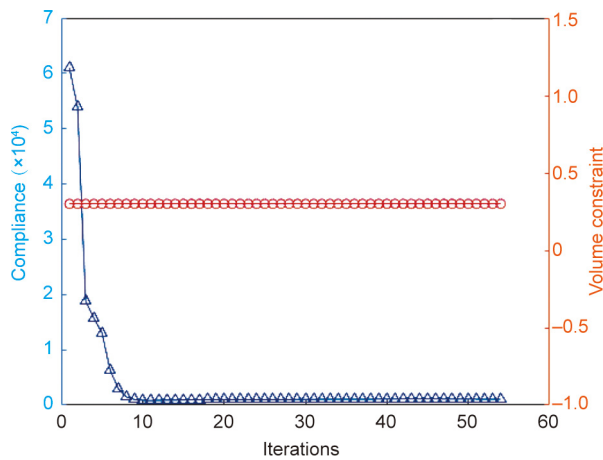


Fig. 9. The ITO convergence history of the cantilever beam with $40 \times 20 \times 20$ elements.

can be along the x , y , or z directions, and the distance between each layer can be nonuniform. In this example, 20 layers are selected along the y direction (i.e., the vertical direction), and the distances between them are uniform, as shown in Fig. 10(a). Second, these NURBS surfaces are input to the SPI procedure one by one, as shown in Fig. 10(b). Three representative SPI procedures are shown in Figs. 10(c)–(e), where the intersection curves are highlighted in yellow and the threshold planes are shown in gray. It should be noted that the altitude value of the threshold plane is chosen in such a way that the optimization constraints are met, and each layer will intersect with a threshold plane at the same relative altitude to generate contour curves. The output of the SPI procedure of 20 NURBS surfaces is 20 closed contour curves. To reduce the number of control points, a data reduction method can be adopted [49,50]. Third, these contour curves are input to the surface skinning procedure to generate a lateral surface, as shown in Fig. 10(h). In this step, a NURBS curve reconstruction procedure is used to unify the number of control points for each curve to 50. To construct a smooth surface, the order of the skinning procedure should be chosen wisely, and is chosen as 10 in this example. Fig. 10(j) shows the output of the surface skinning procedure. Next, the top and bottom boundaries should be constructed using the plane trimming algorithm shown in Fig. 10(f). The input of this procedure is the contour curves from the top/bottom layer and their altitude. The output patches are shown in Figs. 10(i) and (k). Finally, all the model boundaries gained above, including the lateral surface and the top/bottom patches, are combined to generate the final CAD model, which is shown in Fig. 10(l).

Theoretically, the accuracy of the final CAD model depends on the accuracy selection of each geometric algorithm. In this cantilever beam example, the top contour curve has 286 control points, and the error precision of the SPI algorithm is chosen to be 0.001. The order of the surface skinning algorithm and the curve reconstruction control points number can also influence the accuracy of the CAD model. The final ITO CAD model constructed by the automatic generation method is selectable and editable in a CAD system. Fig. 11(a) shows the control points of the ITO CAD model in a CAD system; designers can edit the model by changing the coordinates of the control points. Fig. 11(b) is an edited example in which the designer improved the manufacturability of the model by changing the control points of the left side. In this way, the ITO model can be used for any further process, including CAE analysis and shape optimization.

4.2. HAD with a complex design domain

To demonstrate that the proposed method can be used for irregular design domains, an ITO example for half of an automotive part is used here. As shown in Fig. 12(a), half of an automotive part embedded in a regular domain with $36 \times 36 \times 18 = 23\,328$ quadratic 3D NURBS elements is subjected to a downward force $F = 10^4$, and the right side face of the design domain is fixed. In this example, the Young's modulus and the Poisson's ratio are 10^9 and 0.3, respectively. The design objective is to achieve a structure with maximum stiffness (i.e., minimum compliant) under a volume fraction constraint VF that is set to 0.4. The final optimized structure represented by elements whose relative densities are larger than 0.5 is shown in Fig. 12(b), where the optimized result is completely in the irregular design domain. After 197 iterations, the ITO converges, and the minimum compliance is convergence to 5.47×10^4 . The iterative curve of the automotive part is shown in Fig. 13.

An editable CAD model of the optimized result can be automatically generated by our automatic generation method, as described in Section 3.2. This example contains a tree-like bifurcation structure. In general, a skeletonization technique is believed to be applicable in this case [51], but such a technique may lose the geometric information of the sections. In our method, as shown in Fig. 14, the final boundary-represented CAD model is successfully constructed. However, there are several topological changes in the section curves sequence, which means that the lateral surface of the model cannot be obtained using only a surface skinning procedure. Thus, the dual-track sweeping and surfaces blending algorithms [47] are introduced here to deal with the topology changing between section layers. The branch like structure is shown in Fig. 14(b). The continuity between the blending surface and the skinning surface at the joint can reach G3, which means that the control points of the surface change uniformly without a mutation or jump point. The final CAD model of the automotive part is compatible with CAD systems and can also be edited by designers. This example demonstrates that the automatic generation of ITO results is also effective for irregular CAD models.

4.3. HAD for multiscale ITO

In this example, the design of a Messerschmitt-Bölkow-Blohm (MBB) beam [52] is investigated to test the effectiveness of the multiscale ITO method. Graded body-centered cubic (BCC) lattices are filled into the MBB beam using an isogeometric mapping strategy [53] and shape interpolation technology. Fig. 15 shows the design domain of the MBB beam with length $L = 250$, width $W = 41.67$, and height $H = 50$, which is discretized by $60 \times 10 \times 12$ quadratic elements. The distributed force $F = 1000$ is applied vertically down the middle of the top face, and the displacements of the support edges located at the left and right of the bottom face are partly fixed. The volume fraction is set to 0.5. The Young's modulus of the base material is $E_s = 2750$, and the Poisson's ratio is $\mu = 0.38$. When the relative difference of the compliance between two consecutive iterations is less than 0.01 or the maximum number of iterations reaches 200, the optimization will be terminated.

The optimized graded lattice sandwich structures (GLSSs) of the MBB beam under a global volume fraction constraint of 0.5 are presented in Fig. 16. It can be observed that high-density BCCs with high stiffness are distributed along the loading domain on the upper face sheet to the support domain on the lower face sheet, constituting the main force transfer path. A considerable number of low-density BCCs located at the secondary force transfer paths contribute to resisting shearing deformation and stably supporting the upper and lower solid face sheets.

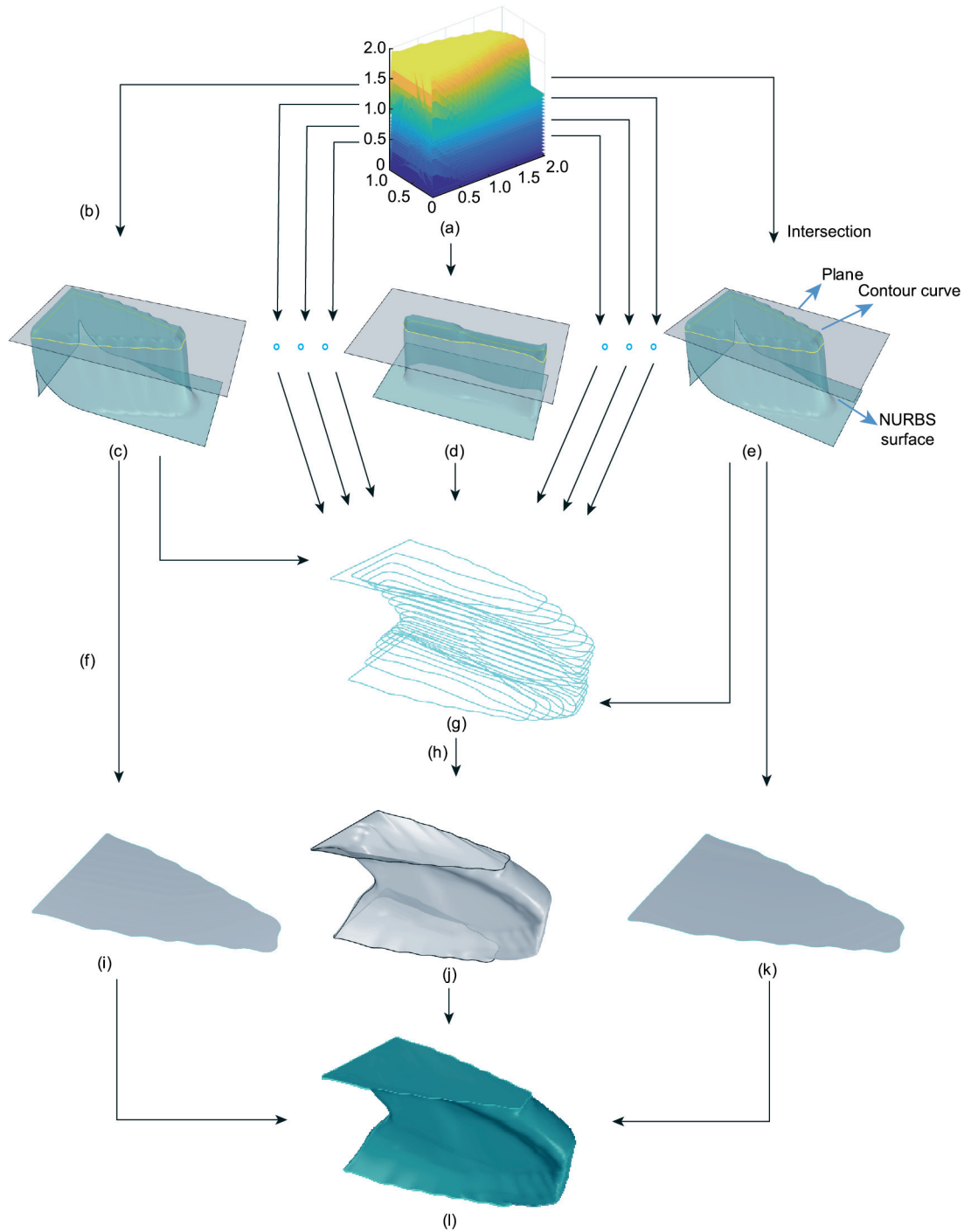


Fig. 10. The procedure for constructing an editable CAD model of the 3D cantilever beam based on the ITO result. (a) ITO result surfaces; (b) SPI procedure; (c) top surface; (d) middle surface; (e) bottom surface; (f) plane trimming procedure; (g) contour curves of each layer; (h) surface skinning procedure; (i) top patch; (j) lateral surface; (k) bottom patch; (l) ITO result model.

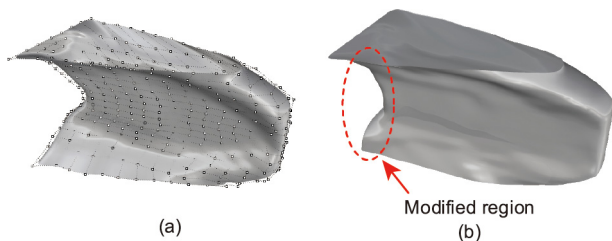


Fig. 11. Edited model in a CAD system. (a) Control points of the surface; (b) the modified model.

As shown in the convergence history in Fig. 17, 180 iterations are implemented. The compliance of the optimized MBB beam by multiscale ITO is 596.3. Compared with a traditional uniform lattice sandwich structure under a global volume constraint of 0.5 [52], the compliance performance is improved by 24.66% via the multiscale ITO. It should be noted that the microstructural unit cells in Fig. 16 have sharp changes at their interfaces, and severe stress concentration may be caused. Effective methods exist that can be used to deal with this issue, such as that in Ref. [54].

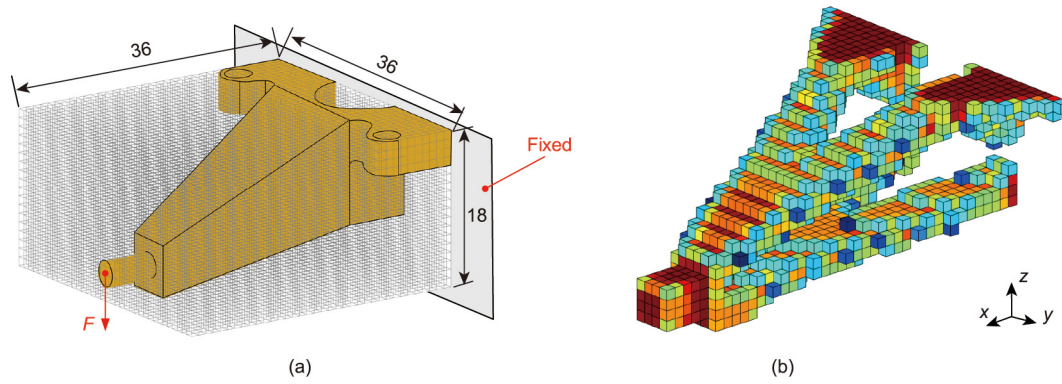


Fig. 12. Half of the automotive part model with (a) its boundary conditions and (b) its optimized structure.

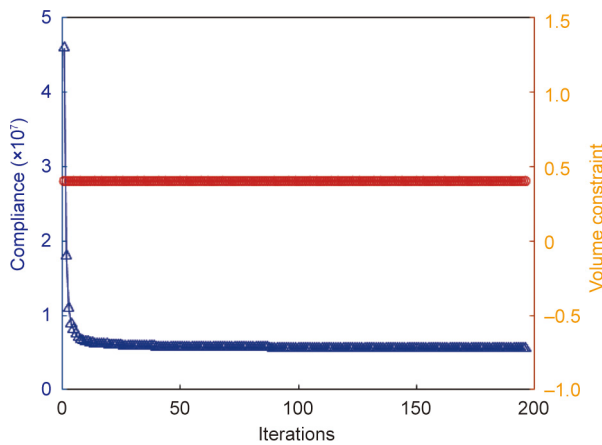


Fig. 13. The convergence history of the automotive part created using ITO.

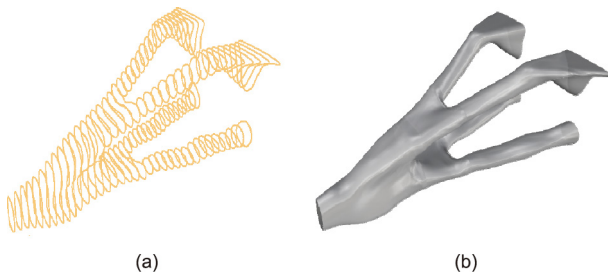


Fig. 14. Automatic construction process of the automotive part. (a) All contour curves along the x direction; (b) the boundary-represented CAD model.

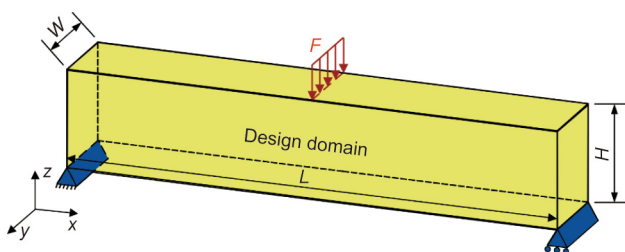


Fig. 15. Design domain of the MBB beam. W : width; H : height; L : length; F : distributed force.

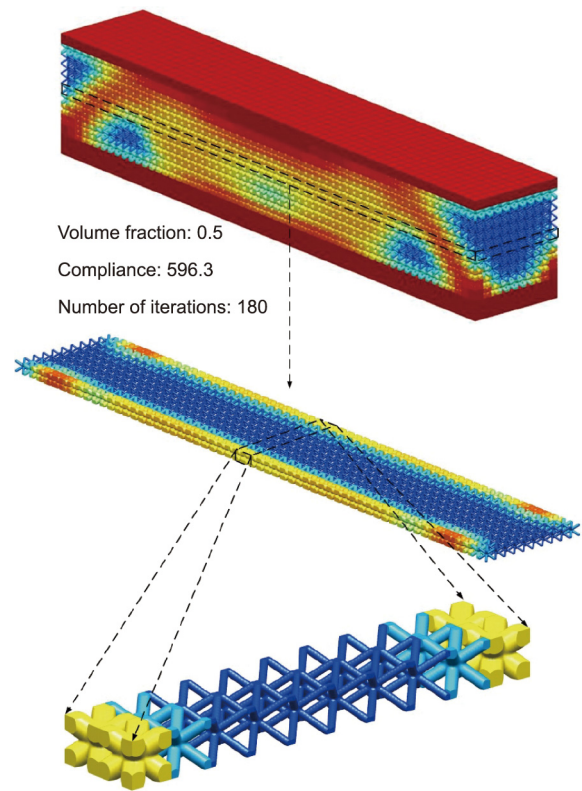


Fig. 16. Optimized design of the MBB beam under a global volume constraint of 0.5.

5. Conclusions and future work

In this paper, we proposed a novel HAD mode via ITO. In HAD, computers take on the leading role during the design procedure, in which editable high-performance product structures can be automatically generated via ITO, and humans (i.e., designers) just slightly modify the optimized structures to meet design requirements if necessary. In this paper, the same NURBS representation was used for modeling, analysis, and TO, and an embedded domain method was proposed to extend ITO from regular design domains to irregular design domains. An automatic generation method based on CAD operations was introduced to generate editable models of ITO results. Three examples were tested to verify the proposed HAD. The first example demonstrated that an editable optimization result can be automatically generated. The second example illustrated that HAD can successfully deal with a design with irregular design domains. The third example showed the

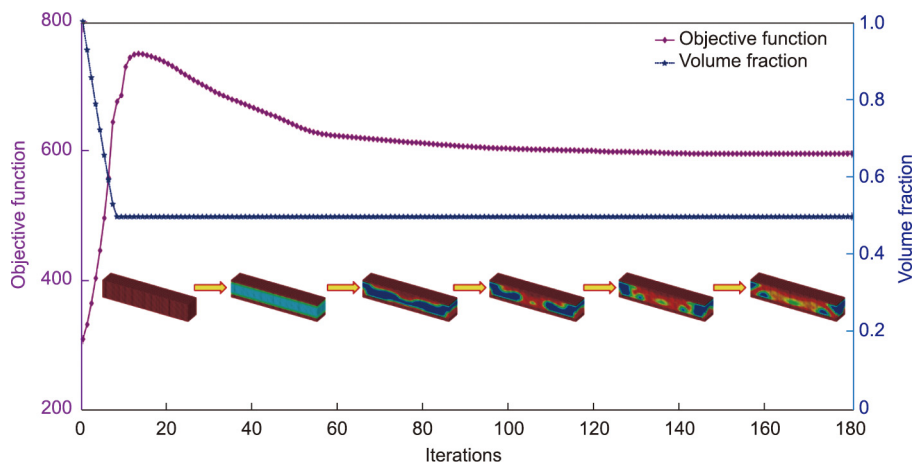


Fig. 17. Iterative curves of the multiscale ITO for the MBB beam.

advantage of HAD in multiscale design for structures with complex geometry.

In this work, the hierarchical subdivision of trimmed elements may greatly increase the number of integration points to ensure the accuracy of the integration. Moreover, since the 3D-optimized CAD models are constructed based on the contour curves of section layers, it is difficult to automatically generate a junction surface if the topological structure dramatically changes between adjacent layers. In the future, we will focus on addressing these two issues and extending the proposed ITO to other problems such as heat conduction [55], fluid-structure interaction [56], and electromagnetism [57].

Acknowledgments

This research was supported by the National Key Research and Development Program of China (2020YFB1708300), National Natural Science Foundation of China (52075184), Natural Science Foundation of Hubei Province (2019CFA059), and Tencent XPLOER PRIZE.

Compliance with ethics guidelines

Yingjun Wang, Mi Xiao, Zhaohui Xia, Peigen Li, and Liang Gao declare that they have no conflict of interest or financial conflicts to disclose.

References

- [1] Ault HK. 3D geometric modeling for the 21st century. *Eng Des Graph J* 1999; 63(2):33–42.
- [2] Bazilevs Y, Calo VM, Cottrell JA, Evans JA, Hughes TJR, Lipton S, et al. Isogeometric analysis using T-splines. *Comput Methods Appl Mech Eng* 2010;199(5–8):229–63.
- [3] Deng YM, Lam YC, Tor SB, Britton GA. A CAD–CAE integrated injection molding design system. *Eng Comput* 2002;18(1):80–92.
- [4] Hughes TJR, Cottrell JA, Bazilevs Y. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput Methods Appl Mech Eng* 2005;194(39–41):4135–95.
- [5] Nguyen VP, Anitescu C, Bordas SPA, Rabczuk T. Isogeometric analysis: an overview and computer implementation aspects. *Math Comput Simul* 2015;117:89–116.
- [6] Wang Y, Wang Z, Xia Z, Poh LH. Structural design optimization using isogeometric analysis: a comprehensive review. *Comput Model Eng Sci* 2018;117(3):455–507.
- [7] Gao J, Xiao M, Zhang Y, Gao L. A comprehensive review of isogeometric topology optimization: methods, applications and prospects. *Chin J Mech Eng* 2020;33:87.
- [8] Lieu QX, Lee J, Lee D, Lee S, Kim D, Lee J. Shape and size optimization of functionally graded sandwich plates using isogeometric analysis and adaptive hybrid evolutionary firefly algorithm. *Thin-Walled Struct* 2018;124:588–604.
- [9] Bendsøe MP, Sigmund O. *Topology optimization: theory, methods, and applications*. Berlin: Springer; 2004.
- [10] Le C, Norato J, Bruns T, Ha C, Tortorelli D. Stress-based topology optimization for continua. *Struct Multidiscipl Optim* 2010;41(4):605–20.
- [11] Huang X, Xie YM. Evolutionary topology optimization of continuum structures with an additional displacement constraint. *Struct Multidiscipl Optim* 2010;40(1–6):409–16.
- [12] Seo YD, Kim HJ, Youn SK. Isogeometric topology optimization using trimmed spline surfaces. *Comput Methods Appl Mech Eng* 2010;199(49–52):3270–96.
- [13] Kumar AV, Parthasarathy A. Topology optimization using B-spline finite elements. *Struct Multidiscipl Optim* 2011;44(4):471–81.
- [14] Hassani B, Khanzadi M, Tavakkoli SM. An isogeometrical approach to structural topology optimization by optimality criteria. *Struct Multidiscipl Optim* 2012;45(2):223–33.
- [15] Qian X. Topology optimization in B-spline space. *Comput Methods Appl Mech Eng* 2013;265:15–35.
- [16] Xie X, Wang S, Xu M, Jiang N, Wang Y. A hierarchical spline based isogeometric topology optimization using moving morphable components. *Comput Methods Appl Mech Eng* 2020;360:112696.
- [17] Taheri AH, Suresh K. An isogeometric approach to topology optimization of multi-material and functionally graded structures. *Int J Numer Methods Eng* 2017;109(5):668–96.
- [18] Gao J, Luo Z, Xiao M, Gao L, Li P. A NURBS-based multi-material interpolation (N-MMI) for isogeometric topology optimization of structures. *Appl Math Model* 2020;81:818–43.
- [19] Wang Y, Xu H, Pasini D. Multiscale isogeometric topology optimization for lattice materials. *Comput Methods Appl Mech Eng* 2017;316:568–85.
- [20] Lieu QX, Lee J. A multi-resolution approach for multi-material topology optimization based on isogeometric analysis. *Comput Methods Appl Mech Eng* 2017;323:272–302.
- [21] Lieu QX, Lee J. Multiresolution topology optimization using isogeometric analysis. *Int J Numer Methods Eng* 2017;112(13):2025–47.
- [22] Shojaei S, Mohamadianb M, Valizadeh N. Composition of isogeometric analysis with level set method for structural topology optimization. *Int J Optim Civ Eng* 2012;2(1):47–70.
- [23] Wang Y, Benson DJ. Isogeometric analysis for parameterized LSM-based structural topology optimization. *Comput Mech* 2016;57(1):19–35.
- [24] Wang Y, Benson DJ. Geometrically constrained isogeometric parameterized level-set based topology optimization via trimmed elements. *Front Mech Eng* 2016;11(4):328–43.
- [25] Ghasemi H, Park HS, Rabczuk T. A level-set based IGA formulation for topology optimization of flexoelectric materials. *Comput Methods Appl Mech Eng* 2017;313:239–58.
- [26] Ghasemi H, Park HS, Rabczuk T. A multi-material level set-based topology optimization of flexoelectric composites. *Comput Methods Appl Mech Eng* 2018;332:47–62.
- [27] Lee SW, Yoon M, Cho S. Isogeometric topological shape optimization using dual evolution with boundary integral equation and level sets. *Comput Aided Des* 2017;82:88–99.
- [28] Sovacool BK, Furszyfer Del Rio DD. Smart home technologies in Europe: a critical review of concepts, benefits, risks and policies. *Renew Sustain Energy Rev* 2020;120:109663.
- [29] Chen B, Wan J, Shu L, Li P, Mukherjee M, Yin B. Smart factory of Industry 4.0: key technologies, application case, and challenges. *IEEE Access* 2017;6:6505–19.
- [30] Meiring GAM, Myburgh HC. A review of intelligent driving style analysis systems and related artificial intelligence algorithms. *Sensors* 2015;15(12):30653–82.
- [31] Rozvany GIN, Bendsøe MP, Kirsch U. Layout optimization of structures. *Appl Mech Rev* 1995;48(2):41–119.

- [32] Svanberg K. The method of moving asymptotes—a new method for structural optimization. *Int J Numer Methods Eng* 1987;24(2):359–73.
- [33] Kang Z, Wang Y. Structural topology optimization based on non-local Shepard interpolation of density field. *Comput Methods Appl Mech Eng* 2011;200(49–52):3515–25.
- [34] Wang Y, Kang Z, He Q. An adaptive refinement approach for topology optimization based on separated density field description. *Comput Struct* 2013;117:10–22.
- [35] Wang Y, Gao L, Qu J, Xia Z, Deng X. Isogeometric analysis based on geometric reconstruction models. *Front Mech Eng* 2021;16(4):782–97.
- [36] Schillinger D, Dedè L, Scott MA, Evans JA, Borden MJ, Rank E, et al. An isogeometric design-through-analysis methodology based on adaptive hierarchical refinement of NURBS, immersed boundary methods, and T-spline CAD surfaces. *Comput Methods Appl Mech Eng* 2012;2012(249–252):116–50.
- [37] Guo Y, Heller J, Hughes TJR, Ruess M, Schillinger D. Variationally consistent isogeometric analysis of trimmed thin shells at finite deformations, based on the STEP exchange format. *Comput Methods Appl Mech Eng* 2018;336:39–79.
- [38] Subedi SC, Verma CS, Suresh K. A review of methods for the geometric post-processing of topology optimized models. *J Comput Inf Sci Eng* 2020;20(6):060801.
- [39] Costa G, Montemurro M, Pailhès J. NURBS hyper-surfaces for 3D topology optimization problems. *Mech Adv Mater Struct* 2021;28(7):665–84.
- [40] Costa G, Montemurro M. Eigen-frequencies and harmonic responses in topology optimisation: a CAD-compatible algorithm. *Eng Struct* 2020;214:110602.
- [41] Barnhill RE, Farin G, Jordan M, Piper BR. Surface/surface intersection. *Comput Aided Geom Des* 1987;4(1–2):3–16.
- [42] Grandine TA, Klein IV FW. A new approach to the surface intersection problem. *Comput Aided Geom Des* 1997;14(2):111–34.
- [43] Barnhill RE, Kersey SN. A marching method for parametric surface/surface intersection. *Comput Aided Geom Des* 1990;7(1–4):257–80.
- [44] Woodward CD. Skinning techniques for interactive B-spline surface interpolation. *Comput Aided Des* 1988;20(8):441–51.
- [45] Piegl L, Tiller W. Algorithm for approximate NURBS skinning. *Comput Aided Des* 1996;28(9):699–706.
- [46] Lin CY, Liou CS, Lai JY. A surface-lofting approach for smooth-surface reconstruction from 3D measurement data. *Comput Ind* 1997;34(1):73–85.
- [47] Hartmann E. Blending an implicit with a parametric surface. *Comput Aided Geom Des* 1995;12(8):825–35.
- [48] Farouki RT. Trimmed-surface algorithms for the evaluation and interrogation of solid boundary representations. *IBM J Res Develop* 1987;31(3):314–34.
- [49] Hennig P, Müller S, Kästner M. Bézier extraction and adaptive refinement of truncated hierarchical NURBS. *Comput Methods Appl Mech Eng* 2016;305:316–39.
- [50] Hiemstra RR, Calabrò F, Schillinger D, Hughes TJR. Optimal and reduced quadrature rules for tensor product and hierarchically refined splines in isogeometric analysis. *Comput Methods Appl Mech Eng* 2017;316:966–1004.
- [51] Yin G, Xiao X, Cirak F. Topologically robust CAD model generation for structural optimisation. *Comput Methods Appl Mech Eng* 2020;369:113102.
- [52] Xiao M, Liu X, Zhang Y, Gao L, Gao J, Chu S. Design of graded lattice sandwich structures by multiscale topology optimization. *Comput Methods Appl Mech Eng* 2021;384:113949.
- [53] Xu M, Xia L, Wang S, Liu L, Xie X. An isogeometric approach to topology optimization of spatially graded hierarchical structures. *Compos Struct* 2019;225:111171.
- [54] Liu X, Gao L, Xiao M, Zhang Y. Kriging-assisted design of functionally graded cellular structures with smoothly-varying lattice unit cells. *Comput Methods Appl Mech Eng* 2022;390:114466.
- [55] Dbouk T. A review about the engineering design of optimal heat transfer systems using topology optimization. *Appl Therm Eng* 2017;112:841–54.
- [56] Lundgaard C, Alexandersen J, Zhou M, Andreasen CS, Sigmund O. Revisiting density-based topology optimization for fluid-structure-interaction problems. *Struct Multidiscipl Optim* 2018;58(3):969–95.
- [57] Nishi S, Yamada T, Izui K, Nishiwaki S, Terada K. Isogeometric topology optimization of anisotropic metamaterials for controlling high-frequency electromagnetic wave. *Int J Numer Methods Eng* 2020;121(6):1218–47.