

The Convergence Rate and Necessary-and-Sufficient Condition for the Consistency of Isogeometric Collocation Method

Hongwei Lin^{a,*}, Yunyang Xiong^a, Qianqian Hu^b

^aDepartment of Mathematics, State Key Lab. of CAD&CG, Zhejiang University, Hangzhou, 310027, China

^bDepartment of Mathematics, Zhejiang Gongshang University, Hangzhou, 310018, China

Abstract

Although the isogeometric collocation (IGA-C) method has been successfully utilized in practical applications due to its simplicity and efficiency, only a little theoretical results have been established on the numerical analysis of the IGA-C method. In this paper, we deduce the convergence rate of the consistency of the IGA-C method. Moreover, based on the formula of the convergence rate, the necessary and sufficient condition for the consistency of the IGA-C method is developed. These results advance the numerical analysis of the IGA-C method.

Keywords: Isogeometric collocation, consistency, necessary and sufficient condition, convergence rate

1. Introduction

In order for the integration of CAD and CAE, Hughes et. al. [1] developed the isogeometric analysis (IGA) method. Since it is based on non-linear NURBS basis functions, the IGA method can directly process the CAD models represented by NURBS, and avoid the tedious mesh transformation procedure.

Because the degree of the non-linear NURBS basis function is relatively high, it is possible to seek a numerical solution, i.e., a NURBS function, by applying the collocation method on the strong form of a differential equation. In this way, the isogeometric collocation (IGA-C) method was proposed [2]. Then unknown coefficients of the NURBS function can be determined by solving a linear system of equations, which is constructed by holding the strong form of the differential equation at some discrete points, called *collocation points*.

The IGA-C method is a simple and efficient method for solving the unknown coefficients of the NURBS function. A comprehensive study [3] revealed its superior behavior over the Galerkin method in terms of accuracy-to-computational-time ratio. Due to these merits, the IGA-C method has been successfully applied in some practical applications. However, the thorough numerical analysis for the IGA-C method is far from being established. Auricchio et. al. developed numerical analysis of the IGA-C method in one-dimensional case [2]. In the generic case, only some sufficient conditions were presented for the consistency and convergence of the IGA-C method [4].

*Corresponding author: phone number: 86-571-87951860-8304, fax number: 86-571-88206681, email: hwlin@zju.edu.cn

In this paper, we first develop the convergence rate of the consistency of the IGA-C method, and then present the necessary and sufficient condition for the consistency of the IGA-C method. Specifically, for a given boundary (or initial) problem with $\mathcal{D}T = f$ (refer to Eq. (1)), where \mathcal{D} is its differential operator. Suppose T_r is its numerical solution, represented by a NURBS function, and \mathcal{I} is an interpolation operator such that $\mathcal{I}f = \mathcal{D}T_r$. The IGA-C method is consistency, if and only if \mathcal{D} and \mathcal{I} are both uniformly bounded when the knot grid size tends to 0.

The structure of this paper is as follows. In Section 1.1, some related work is briefly reviewed. After introducing some preliminaries in Section 2, an introductory example is presented in Section 3. Moreover, the convergence rate of the consistency of the IGA-C method is deduced in Section 4, and the necessary and sufficient condition is developed in Section 5. In addition, some numerical examples are presented in Section 6. Finally, Section 7 concludes the paper.

1.1. Related work

As stated above, the IGA method [1] was proposed to advance the seamless integration of CAD and CAE, by avoiding mesh transformation. Moreover, since it has much less freedom than the traditional finite element method, the IGA method can not only save lots of computation, but also greatly improve the computational precision. Additionally, owing to the knot insertion property of the NURBS function, the shape of the computational domain represented by NURBS can be exactly kept in the mesh refinement. Due to these merits, the IGA method draws great interests in both practical applications and theoretical studies. On one hand, the IGA method has been successfully applied in lots of simulation problems, such as elasticity [5, 6], structure [7–9], and fluid [10–12], etc. On the other hand, some research on the computational aspect of the IGA method has been developed to improve the accuracy and efficiency by using reparameterization and refinement, etc. [13–18]. Recently, an optimal and totally robust multi-iterative method was developed for solving IgA Galerkin linear system [19]. For more details on the IGA method, please refer to Ref. [20] and the references therein.

Since a NURBS function has a relatively high degree, its unknown coefficients can be determined by making the strong form of the PDE hold at some collocation points, that leads to the IGA-C method [2]. Schillinger et. al. presented a comprehensive comparison between the IGA-C method and the Galerkin method, revealing that the IGA-C method is superior to the Galerkin method in terms of accuracy-to-computational-time ratio [3]. Lin et. al. developed some sufficient conditions for the consistency and convergence of the IGA-C method [4]. Moreover, Lorenzis et. al. proposed the IGA-C method for solving the boundary problem with Neumann boundary condition [21].

The IGA-C method has been successfully applied in some practical applications. For instance, the IGA-C method was employed in solving Timoshenko beam problem [22] and spatial Timoshenko rod problem [23], showing that mixed collocation schemes are locking-free independently of the choice of the polynomial degrees for unknown fields. Moreover, the IGA-C method was extended to multi-patch NURBS configurations, various boundary and patch interface conditions, and explicit dynamic analysis [24]. Recently, the IGA-C method was exploited to settle the Bernoulli-Euler beam problem [25] and the Reissner-Mindlin plate problem [26]. However, only very limited theoretical results for the IGA-C method were developed [2, 4] currently, and the numerical analysis for the IGA-C method is still far from being established.

2. Preliminaries

Suppose the IGA-C method is employed to solve the following boundary problem,

$$\begin{cases} \mathcal{D}T = f, & \text{in } \Omega \subset \mathbb{R}^d, \\ \mathcal{G}T = g, & \text{on } \partial\Omega, \end{cases} \quad (1)$$

where $\Omega \subset \mathbb{R}^d$ is a physical domain of d dimension, $\mathcal{D} : \mathbb{V} \rightarrow \mathbb{W}$ is a bounded differential operator, where \mathbb{V} and \mathbb{W} are two Hilbert spaces, $\mathcal{G}T$ is a boundary condition, and $f : \Omega \rightarrow \mathbb{R}$, $g : \partial\Omega \rightarrow \mathbb{R}$ are two given continuous functions defined on their domains. Suppose the analytical solution $T \in C^m(\Omega)$, where m is larger than or equal to the maximum order of derivatives appearing in the operator \mathcal{D} .

In the IGA method, the physical domain Ω is represented by a NURBS mapping,

$$F : \Omega_p \rightarrow \Omega, \quad (2)$$

where Ω_p is a parameter domain. Replacing the control points of F by unknown control coefficients, we get the representation of the numerical solution to the boundary problem (1), denoted as $T_r(\boldsymbol{\eta})$, $\boldsymbol{\eta} \in \Omega_p$. Meanwhile, by the inverse mapping F^{-1} , the physical domain Ω can be mapped into the parameter domain Ω_p , and then, the numerical solution T_r is still defined on the physical domain Ω through the mapping F^{-1} . Additionally, by the mapping F , the function f can be defined on Ω_p , and G on $\partial\Omega_p$.

In isogeometric analysis, while the physical domain of the boundary problem (1) is Ω , the computational domain is the parameter domain Ω_p (2). Although the operators \mathcal{D} and \mathcal{G} in Eq. (1) are performed on the variables in the physical domain, the generated formulae will be transformed into the parameter domain Ω_p for computation. Therefore, the functions in the function approximation problem in the IGA-C method should be considered to be defined on the parameter domain Ω_p .

Definition 1 (Stable operator [27]). *Let \mathbb{V} , \mathbb{W} be Hilbert spaces and $\mathcal{D} : \mathbb{V} \rightarrow \mathbb{W}$ be a differential operator. If there exists a constant $C_S > 0$ such that*

$$\|\mathcal{D}v\|_{\mathbb{W}} \geq C_S \|v\|_{\mathbb{V}}, \text{ for all } v \in D(\mathcal{D}),$$

where $D(\mathcal{D})$ represents the domain of \mathcal{D} , then the differential operator \mathcal{D} is called a stable operator.

[Remark 1:] In this paper, we suppose that the L^∞ norm $\|\cdot\|_{L^\infty}$ is equivalent to the norm $\|\cdot\|_{\mathbb{W}}$ in \mathbb{W} and the norm $\|\cdot\|_{\mathbb{V}}$ in \mathbb{V} . In other words, there exists nonnegative constants c_v , C_v , and c_w , C_w satisfying,

$$\begin{aligned} c_v \|\cdot\|_{\mathbb{V}} &\leq \|\cdot\|_{L^\infty} \leq C_v \|\cdot\|_{\mathbb{V}} \\ c_w \|\cdot\|_{\mathbb{W}} &\leq \|\cdot\|_{L^\infty} \leq C_w \|\cdot\|_{\mathbb{W}} \end{aligned}$$

Suppose $T_r(\boldsymbol{\eta})$ is an unknown NURBS function defined on the knot grid $\mathcal{T}^\rho \in \mathbb{R}^d$, $d = 1, 2, 3$. Specifically, \mathcal{T}^ρ is a knot sequence in 1D case, a rectangular grid in 2D case, and a hexahedral grid in 3D case, where ρ is the **knot grid size** defined as the following definition.

Definition 2. Given a set $\Phi \subset \mathbb{R}^d$, its **diameter** $\text{diam}(\Phi)$ is defined by

$$\text{diam}(\Phi) = \sup\{d(\mathbf{x}, \mathbf{y}), \mathbf{x}, \mathbf{y} \in \Phi\},$$

where $d(\mathbf{x}, \mathbf{y})$ denotes the Euclidean distance between \mathbf{x} and \mathbf{y} . And we call ρ as the **knot grid size** of \mathcal{T}^p , which is defined as the maximum of the diameters of the **knot intervals** of \mathcal{T}^p . That is, $\rho = \max_i\{\text{diam}([u_i, u_{i+1}])\}$ in 1D case, $\rho = \max_{ij}\{\text{diam}([u_i, u_{i+1}] \times [v_j, v_{j+1}])\}$ in 2D case, and $\rho = \max_{ijk}\{\text{diam}([u_i, u_{i+1}] \times [v_j, v_{j+1}] \times [w_k, w_{k+1}])\}$ in 3D case.

Definition 3. Let $T : \Omega_p \rightarrow \mathbb{R}$, $T \in C^0(\Omega_p)$ be a continuous function on the parameter domain Ω_p , where $C^0(\Omega_p)$ is the space of continuous functions on Ω_p . The **modulus of continuity** [28] of the function T , denoted as $\omega(T, h)$, is defined by

$$\omega(T, h) = \max\{|T(\mathbf{x}) - T(\mathbf{y})|, d(\mathbf{x}, \mathbf{y}) < h\}, h \in \mathbb{R}. \quad (3)$$

The modulus of continuity $\omega(T, h)$ satisfies the property [28],

$$\omega(T, h + k) \leq \omega(T, h) + \omega(T, k), h, k \in \mathbb{R},$$

and then

$$\omega(T, K\rho) \leq K\omega(T, \rho), K \in \mathbb{Z}. \quad (4)$$

Definition 4. Let I^p be an interpolation operator, and $I^p g$ be a spline interpolant of a function g defined on the knot grid \mathcal{T}^p . Suppose \mathbb{P} is a spline space composed of the splines with the same knot grid and degree as those of $I^p g$. **The distance of the function g to \mathbb{P}** , i.e., $\text{dist}(g, \mathbb{P})$, is defined by

$$\text{dist}(g, \mathbb{P}) = \min\{\|g - p\|_{L^\infty}, p \in \mathbb{P}\}. \quad (5)$$

3. An introductory example

Consider the following one-dimensional boundary problem:

$$\begin{cases} T'(x) = f(x), & x \in [a, b], \\ T(a) = g_1, T(b) = g_2, \end{cases} \quad (6)$$

where $f(x) \in C[a, b]$ is a continuous function, $T(x) \in C^1[a, b]$ is an analytical solution, and $g_1, g_2 \in \mathbb{R}$.

The physical domain $[a, b]$ in Eq. (6) is modeled as,

$$x(t) = \sum_{i=0}^N (a + \frac{i}{N}(b-a))B_{i,k}(t), t \in [0, 1], \quad (7)$$

where $B_{i,k}(t)$ is a B-spline basis function of order k , defined on the knot sequence,

$$G : \underbrace{0, 0, \dots, 0}_k, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}, \underbrace{1, 1, \dots, 1}_k. \quad (8)$$

Eq. (7) maps $[0, 1]$ to $[a, b]$, i.e.,

$$F_1 : [0, 1] \rightarrow [a, b]. \quad (9)$$

Then, the numerical solution $T_r(t)$ to the boundary problem (6) can be generated by replacing the coefficients $a + \frac{i}{N}(b-a)$ in $x(t)$ (7) by the unknowns coefficients p_i $i = 0, 1, \dots, N$, i.e.,

$$T_r(t) = \sum_{i=0}^N p_i B_{i,k}(t). \quad (10)$$

Note that, by the inverse mapping F_1^{-1} (9), $T_r(t)$ is defined on the physical domain $[a, b]$ (6), i.e., $T_r(t(x))$, $x \in [a, b]$.

Because,

$$\frac{dT_r(t)}{dt} = (k-1) \sum_{i=0}^{N-1} \frac{p_{i+1} - p_i}{\frac{i+k-1}{N} - \frac{i}{N}} B_{i,k-1}(t) = N \sum_{i=0}^{N-1} (p_{i+1} - p_i) B_{i,k-1}(t),$$

and,

$$\frac{dx(t)}{dt} = b - a,$$

substituting Eq. (10) into Eq. (6) yields,

$$\begin{cases} \frac{dT_r}{dx} = \frac{dT_r}{dt} \frac{dt}{dx} = \frac{dT_r}{dt} \frac{1}{\frac{dx}{dt}} = \sum_{i=0}^{N-1} \frac{N}{b-a} (p_{i+1} - p_i) B_{i,k-1}(t) = f(x(t)), \\ T_r(0) = p_0 = g_1, \quad T_r(1) = p_N = g_2. \end{cases}$$

In order for solving the unknown coefficients in Eq. (10) using the IGA-C method, a linear system is generated by sampling $N-1$ points $\tau_1, \tau_2, \dots, \tau_{N-1}$ in the interval $(0, 1)$, i.e.,

$$\begin{cases} \frac{dT_r(\tau_j)}{dx} = \sum_{i=0}^{N-1} \frac{N}{b-a} (p_{i+1} - p_i) B_{i,k-1}(\tau_j) = f(x(\tau_j)), \quad \tau_j \in (0, 1), \quad j = 1, 2, \dots, N-1, \\ T_r(0) = p_0 = g_1, \\ T_r(1) = p_n = g_2. \end{cases} \quad (11)$$

When the knot grid size $\rho = \frac{1}{N}$ of the knot sequence G (8) tends to 0, it follows $N \rightarrow +\infty$. If the control points $\frac{N}{b-a}(p_{i+1} - p_i) \rightarrow \infty$, ($\rho \rightarrow 0$), too, we have $\frac{dT_r(\tau_j)}{dx} \rightarrow \infty$, ($\rho \rightarrow 0$), $j = 1, 2, \dots, N-1$. However, because $f(x(t))$ is continuous on the close interval $[0, 1]$, $f(x(\tau_j))$ is bounded. Therefore, if the linear system (11) has a solution, there should exist $T_r(t)$ so that the control points $\frac{N}{b-a}(p_{i+1} - p_i)$ of $\frac{dT_r}{dx}$ are bounded when $\rho = \frac{1}{N} \rightarrow 0$. It results in that $\frac{dT_r}{dx}$ is also bounded when $\rho \rightarrow 0$. All of such B-spline functions $T_r(t)$ constitute a B-spline subspace, and the first order derivative operator in Eq. (6) should be bounded on the B-spline subspace when $\rho \rightarrow 0$.

4. The convergence rate

Suppose the NURBS function $T_r(\boldsymbol{\eta})$, $\boldsymbol{\eta} \in \Omega_p \subset \mathbb{R}^d$ defined on the knot grid \mathcal{T}^ρ has n unknown control coefficients p_i , i.e.,

$$T_r(\boldsymbol{\eta}) = \sum_i p_i \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} = \frac{P(\boldsymbol{\eta})}{W(\boldsymbol{\eta})}, \quad \boldsymbol{\eta} \in \Omega_p \subset \mathbb{R}^d, \quad (12)$$

where $w_i > 0$ are known weights, $B_i(\boldsymbol{\eta})$ are the B-spline basis functions, the weight function $W(\boldsymbol{\eta})$ is a known polynomial spline function, and $P(\boldsymbol{\eta})$ is a polynomial spline function with n unknown

control coefficients p_i . Moreover, the subscript \mathbf{i} in Eq. (12) is an index vector, $\mathbf{i} = (i_1, i_2, \dots, i_d)$. According to the IGA-C method, these unknown coefficients p_i can be determined by solving the following linear system of equations,

$$\begin{cases} \mathcal{D}T_r(\boldsymbol{\eta}_k) = f(\boldsymbol{\eta}_k), & k = 1, 2, \dots, n_1, \\ \mathcal{G}T_r(\boldsymbol{\eta}_l) = g(\boldsymbol{\eta}_l), & l = n_1 + 1, \dots, n, \end{cases} \quad (13)$$

where $\boldsymbol{\eta}_k (k = 1, 2, \dots, n_1)$ are collocation points inside Ω_p , and $\boldsymbol{\eta}_l (l = n_1 + 1, \dots, n)$ are collocation points on $\partial\Omega_p$. Note that, throughout this paper, the operators \mathcal{D} and \mathcal{G} are performed on the variable in the physical domain Ω (Eq. (1)).

[Remark 2:] In this paper, we assume that the coefficient matrix of the above linear system (13) is of full rank and then it has a unique solution. Otherwise, the IGA-C method is invalid.

According to the result developed in Ref. [4], $\mathcal{D}T_r$ can be represented by

$$\mathcal{D}T_r(\boldsymbol{\eta}) = \sum_i p_i \mathcal{D} \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} = \sum_i p_i \frac{\bar{B}_i(\boldsymbol{\eta})}{\bar{W}(\boldsymbol{\eta})} = \frac{\bar{P}(\boldsymbol{\eta})}{\bar{W}(\boldsymbol{\eta})}, \quad (14)$$

where $\bar{B}_i(\boldsymbol{\eta})$ is the result by applying the differential operator \mathcal{D} to $\frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})}$, $\bar{W}(\boldsymbol{\eta})$ is the power of $W(\boldsymbol{\eta})$, and $\bar{P}(\boldsymbol{\eta})$ is a polynomial B-spline function with n unknowns p_i . By Ref. [4], $\bar{P}(\boldsymbol{\eta})$ and $\bar{W}(\boldsymbol{\eta})$ both have the same break point sequence and the same knot intervals as $T_r(\boldsymbol{\eta})$. To determine these unknowns p_i in $\bar{P}(\boldsymbol{\eta})$, let $\mathcal{D}T_r(\boldsymbol{\eta})$ interpolate $\mathcal{D}T(\boldsymbol{\eta}) = f(\boldsymbol{\eta})$ at n_1 collocation points inside the domain Ω_p (refer to Eq. (13)), i.e.,

$$\mathcal{D}T_r(\boldsymbol{\eta}_k) - f(\boldsymbol{\eta}_k) = \frac{\bar{P}(\boldsymbol{\eta}_k)}{\bar{W}(\boldsymbol{\eta}_k)} - f(\boldsymbol{\eta}_k) = \frac{\bar{P}(\boldsymbol{\eta}_k) - \bar{W}(\boldsymbol{\eta}_k)f(\boldsymbol{\eta}_k)}{\bar{W}(\boldsymbol{\eta}_k)} = 0, \quad k = 1, 2, \dots, n_1. \quad (15)$$

Note that $\bar{W}(\boldsymbol{\eta}) \neq 0$ is a known function, Eq. (15) is equivalent to

$$\bar{P}(\boldsymbol{\eta}_k) = \sum_i p_i \bar{B}_i(\boldsymbol{\eta}_k) = \bar{W}(\boldsymbol{\eta}_k)f(\boldsymbol{\eta}_k), \quad k = 0, 1, \dots, n_1. \quad (16)$$

Similarly, $\mathcal{G}T_r(\boldsymbol{\eta})$ in Eq. (13) can be written as

$$\mathcal{G}T_r(\boldsymbol{\eta}) = \sum_i p_i \mathcal{G} \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} = \sum_i p_i \frac{\tilde{B}_i(\boldsymbol{\eta})}{\tilde{W}(\boldsymbol{\eta})} = \frac{\tilde{P}(\boldsymbol{\eta})}{\tilde{W}(\boldsymbol{\eta})}, \quad (17)$$

where $\tilde{B}_i(\boldsymbol{\eta})$ are the result generated by applying the operator \mathcal{G} to $\frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})}$, $\tilde{W}(\boldsymbol{\eta}) \neq 0$ is a known B-spline function, and $\tilde{P}(\boldsymbol{\eta})$ is an unknown B-spline function with n unknowns p_i . Then the linear equations $\mathcal{G}T_r(\boldsymbol{\eta}_l) = g(\boldsymbol{\eta}_l)$ in Eq. (13) are equivalent to

$$\tilde{P}(\boldsymbol{\eta}_l) = \sum_i p_i \tilde{B}_i(\boldsymbol{\eta}_l) = \tilde{W}(\boldsymbol{\eta}_l)g(\boldsymbol{\eta}_l), \quad l = n_1 + 1, n_1 + 2, \dots, n, \quad (18)$$

where $\boldsymbol{\eta}_l \in \partial\Omega_p$, $l = n_1 + 1, \dots, n$.

Therefore, combining Eqs. (16) and (18), the linear system (13) becomes

$$\begin{cases} \bar{P}(\boldsymbol{\eta}_k) = \sum_i p_i \bar{B}_i(\boldsymbol{\eta}_k) = \bar{W}(\boldsymbol{\eta}_k)f(\boldsymbol{\eta}_k), & k = 0, 1, \dots, n_1, \\ \tilde{P}(\boldsymbol{\eta}_l) = \sum_i p_i \tilde{B}_i(\boldsymbol{\eta}_l) = \tilde{W}(\boldsymbol{\eta}_l)g(\boldsymbol{\eta}_l), & l = n_1 + 1, n_1 + 2, \dots, n. \end{cases} \quad (19)$$

Since the linear system of equations (19) is equivalent to (13), then the coefficient matrix of (19) is of full rank, and it also has a unique solution.

[Remark 3:] In Eq. (13), let the functions f and g vary in $C^m(\Omega_p)$ and $C^m(\partial\Omega_p)$, respectively, and the differential operator \mathcal{D} be fixed. In addition, let the weight function $W(\boldsymbol{\eta})$ in T_r (12) be fixed as well. Then, all the numerical solutions $T_r(\boldsymbol{\eta})$ (12) generated by the IGA-C method (13) constitute a **linear spline space** $\mathbb{S}_\rho(\Omega_p)$, where ρ is the knot grid size of T_r . It should be pointed out that, all the NURBS functions T_r in the linear space $\mathbb{S}_\rho(\Omega_p)$ have the same weight function $W(\boldsymbol{\eta})$, the same knot grid with knot grid size ρ and the same degree. In order for $\rho \rightarrow 0$, the knot grid of the spline functions in $\mathbb{S}_\rho(\Omega_p)$ is refined by knot insertion, thus resulting in a series of spline spaces. Moreover, because all the numerical solutions $T_r(\boldsymbol{\eta})$ constitute the linear space $\mathbb{S}_\rho(\Omega_p)$, all of $\mathcal{D}T_r(\boldsymbol{\eta})$ compose a **linear spline space** $\mathbb{S}_{\rho,e}^d(\Omega_p)$, where e is the continuity order of the splines in $\mathbb{S}_{\rho,e}^d(\Omega_p)$. As aforementioned, $\mathcal{D}T_r$ has the same break point sequence with that of T_r (12), so they have the same knot grid size ρ .

The following Lemma 1 estimates the distance from a continuous function $f \in C^0(\Omega_p)$ to the linear space $\mathbb{S}_{\rho,e}^d(\Omega_p)$, i.e., $\text{dist}(f, \mathbb{S}_{\rho,e}^d)$. In Ref. [28, pp.146], an inequality to estimate the distance is proposed for univariate functions, and the inequality can be extended to our case.

Lemma 1. *If $\mathcal{D}T = f \in C^0(\Omega_p)$, and $T \in C^m(\Omega_p)$ (Eq. (1)), then we have*

$$\text{dist}(f, \mathbb{S}_{\rho,e}^d) = \text{dist}(\mathcal{D}T, \mathbb{S}_{\rho,e}^d) \leq \|\mathcal{D}\| K\omega(T, \rho),$$

where K is an integer related to the degree of the NURBS functions in the spline space $\mathbb{S}_\rho(\Omega_p)$.

Proof: As stated above, the NURBS functions approximating the analytical solution T constitute the linear space $\mathbb{S}_\rho(\Omega_p)$ defined on the knot grid \mathcal{T}^ρ . We select a special function from the space $\mathbb{S}_\rho(\Omega_p)$, i.e.,

$$T_r(\boldsymbol{\eta}) = \sum_i T(\tau_i) \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})},$$

and construct a spline function $(Af)(\boldsymbol{\eta})$ to approximate the function $f \in C^0(\Omega_p)$, i.e.,

$$(Af)(\boldsymbol{\eta}) = \mathcal{D}T_r(\boldsymbol{\eta}) = \mathcal{D} \sum_i T(\tau_i) \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} = \sum_i T(\tau_i) \mathcal{D} \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})},$$

where T is the analytical solution of Eq. (1), and $Af = \mathcal{D}T_r \in \mathbb{S}_{\rho,e}^d(\Omega_p)$ (defined in Remark 3). The point sequence $\{\tau_i \in \Omega_p\}$ is sampled in such a way that each knot interval of the knot grid \mathcal{T}^ρ contains at least one point, and τ_i is in the non-zero region of $B_i(\boldsymbol{\eta})$.

Suppose $u(\boldsymbol{\eta}), v(\boldsymbol{\eta}) \in C^m(\Omega_p)$. The function $u(\boldsymbol{\eta})$ is an arbitrary function in $C^m(\Omega_p)$, and,

$$v(\boldsymbol{\eta}) = T(\boldsymbol{\eta}) - \sum_i T(\tau_i) \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})}, \quad \boldsymbol{\eta} \in \Omega_p. \quad (20)$$

Note that $|v(\boldsymbol{\eta})|$ is continuous in the close set Ω_p , so $|v(\boldsymbol{\eta})|$ can take its maximum value in Ω_p . Namely, there exists $\boldsymbol{\eta}^* \in \Omega_p$ such that

$$|v(\boldsymbol{\eta}^*)| = \max_{\boldsymbol{\eta}} |v(\boldsymbol{\eta})| = \|v\|_{L^\infty}, \quad \boldsymbol{\eta} \in \Omega_p.$$

For an arbitrary value $\hat{\boldsymbol{\eta}} \in \Omega_p$, it holds,

$$\|\mathcal{D}\| = \max_{u \in C^m(\Omega_p)} \frac{\|\mathcal{D}u\|_{L^\infty}}{\|u\|_{L^\infty}} \geq \frac{\|\mathcal{D}v\|_{L^\infty}}{\|v\|_{L^\infty}} \geq \frac{|\mathcal{D}v(\hat{\boldsymbol{\eta}})|}{|v(\boldsymbol{\eta}^*)|},$$

which is equivalent to,

$$|\mathcal{D}v(\hat{\boldsymbol{\eta}})| \leq \|\mathcal{D}\| |v(\boldsymbol{\eta}^*)|. \quad (21)$$

Suppose J is the index vector set satisfying $B_i(\boldsymbol{\eta}^*) \neq 0$, $i \in J$. Because

$$\sum_i \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} = \sum_{i \in J} \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} = 1, \text{ and then } T(\boldsymbol{\eta}^*) = \sum_i T(\boldsymbol{\eta}^*) \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} = \sum_{i \in J} T(\boldsymbol{\eta}^*) \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)},$$

together with Eq. (21), we have,

$$\begin{aligned} |f(\hat{\boldsymbol{\eta}}) - (Af)(\hat{\boldsymbol{\eta}})| &= \left| \mathcal{D}T(\hat{\boldsymbol{\eta}}) - \mathcal{D} \sum_i T(\tau_i) \frac{w_i B_i(\hat{\boldsymbol{\eta}})}{W(\hat{\boldsymbol{\eta}})} \right| \\ &= |\mathcal{D}v(\hat{\boldsymbol{\eta}})| \leq \|\mathcal{D}\| |v(\boldsymbol{\eta}^*)| \quad (\text{Eq. (21)}) \\ &= \|\mathcal{D}\| \left| T(\boldsymbol{\eta}^*) - \sum_i T(\tau_i) \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} \right| \quad (\text{Eq. (20)}) \\ &= \|\mathcal{D}\| \left| \sum_i T(\boldsymbol{\eta}^*) \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} - \sum_i T(\tau_i) \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} \right| \\ &= \|\mathcal{D}\| \left| \sum_{i \in J} (T(\boldsymbol{\eta}^*) - T(\tau_i)) \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} \right| \\ &\leq \|\mathcal{D}\| \sum_{i \in J} |T(\boldsymbol{\eta}^*) - T(\tau_i)| \frac{w_i B_i(\boldsymbol{\eta}^*)}{W(\boldsymbol{\eta}^*)} \\ &\leq \|\mathcal{D}\| \max_{i \in J} |T(\boldsymbol{\eta}^*) - T(\tau_i)| \quad (\text{Definition 3}) \\ &\leq \|\mathcal{D}\| \omega(T, K\rho), \end{aligned}$$

where K is an integer related to the degree of $T_r(\boldsymbol{\eta})$. It is because that the non-zero region of $B_i(\boldsymbol{\eta})$ is determined by the degree of $T_r(\boldsymbol{\eta})$. By Eq. (4), we get

$$|f(\hat{\boldsymbol{\eta}}) - (Af)(\hat{\boldsymbol{\eta}})| \leq \|\mathcal{D}\| K\omega(T, \rho).$$

Because $Af = \mathcal{D}T_r \in \mathbb{S}_{\rho, e}^d(\Omega_p)$, and $\hat{\boldsymbol{\eta}} \in \Omega_p$ is an arbitrary value, it can be so chosen that

$$\text{dist}(f, \mathbb{S}_{\rho, e}^d) = \min\{\|f - s\|_{L^\infty}, s \in \mathbb{S}_{\rho, e}^d\} \leq |f(\hat{\boldsymbol{\eta}}) - (Af)(\hat{\boldsymbol{\eta}})| \leq \|\mathcal{D}\| K\omega(T, \rho).$$

Then the Lemma is proved. \square

Furthermore, we have

Lemma 2. *If $T(\boldsymbol{\eta}) \in C^1(\Omega_p)$, then it holds*

$$\omega(T, \rho) \leq \rho \max_{\boldsymbol{\eta} \in \Omega_p} \|\nabla T\|_E,$$

where ∇T is the gradient of T , and the norm $\|\cdot\|_E$ is defined as $\|\boldsymbol{\eta}\|_E = \|(\eta_1, \eta_2, \dots, \eta_d)\|_E = \sqrt{\eta_1^2 + \eta_2^2 + \dots + \eta_d^2}$.

Proof: Let $\mathbf{x}, \mathbf{y} \in \Omega_p$, $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_E \leq \rho$, and $c \in (0, 1)$. According to the mean value theorem, it follows that

$$\begin{aligned} |T(\mathbf{x}) - T(\mathbf{y})| &= \left| \nabla T|_{(1-c)\mathbf{x}+c\mathbf{y}} \cdot (\mathbf{x} - \mathbf{y}) \right| \leq \left\| \nabla T|_{(1-c)\mathbf{x}+c\mathbf{y}} \right\|_E \cdot \|\mathbf{x} - \mathbf{y}\|_E \\ &\leq \rho \left\| \nabla T|_{(1-c)\mathbf{x}+c\mathbf{y}} \right\|_E \leq \rho \max_{\boldsymbol{\eta} \in \Omega_p} \|\nabla T\|_E. \end{aligned}$$

Then, by the definition of $\omega(T, \rho)$ (Eq. (3)), we have

$$\omega(T, \rho) = \max\{|T(\mathbf{x}) - T(\mathbf{y})|, d(\mathbf{x}, \mathbf{y}) < \rho\} \leq \rho \max_{\boldsymbol{\eta} \in \Omega_p} \|\nabla T\|_E.$$

□

Moreover, we denote by \mathcal{I}^ρ an interpolation operator, which maps a continuous function to a spline function defined on the knot grid \mathcal{T}^ρ with knot grid size ρ . Specifically, for the continuous function $f = \mathcal{D}T \in C^0(\Omega_p)$ (refer to Eqs. (1) and (13)), we have

$$\mathcal{I}^\rho f = \mathcal{D}T_r \in \mathbb{S}_{\rho,e}^d(\Omega_p), \quad (22)$$

and the following Lemma.

Lemma 3. Suppose $\mathcal{D}T = f \in C^0(\Omega_p)$ (Eq. (1)), and $T_r \in \mathbb{S}_\rho(\Omega_p)$ (Refer to Eq. (12) and Remark 3) is the NURBS function approximating the analytical solution T . Then,

$$\|\mathcal{D}T_r - \mathcal{D}T\|_{\mathbb{W}} \leq (1 + \|\mathcal{I}^\rho\|) \text{dist}(f, \mathbb{S}_{\rho,e}^d), \quad (23)$$

where \mathcal{I}^ρ is the interpolation operator defined by Eq. (22), and $\mathbb{S}_{\rho,e}^d$ is defined as in Remark 3.

Proof: On one hand, given an arbitrary known NURBS function $T_q(\boldsymbol{\eta}) \in \mathbb{S}_\rho(\Omega_p)$ expressed as

$$T_q(\boldsymbol{\eta}) = \frac{Q(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} = \sum_i q_i \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})}, \quad \boldsymbol{\eta} \in \Omega_p, \quad (24)$$

where the weight function $W(\boldsymbol{\eta})$ and the weight w_i are the same as those in (12), two functions $h(\boldsymbol{\eta})$ and $h_b(\boldsymbol{\eta})$ can be generated by performing the operators \mathcal{D} and \mathcal{G} on T_q (see Eq. (1)), respectively, i.e.,

$$h(\boldsymbol{\eta}) = \mathcal{D}T_q(\boldsymbol{\eta}) = \sum_i q_i \mathcal{D} \left(\frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} \right), \quad h_b(\boldsymbol{\eta}) = \mathcal{G}T_q(\boldsymbol{\eta}) = \sum_i q_i \mathcal{G} \left(\frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} \right). \quad (25)$$

We construct an unknown NURBS function $T_x(\boldsymbol{\eta}) \in \mathbb{S}_\rho(\Omega_p)$ with n unknown control coefficients x_i , the same knot grid and degree with T_q ,

$$T_x(\boldsymbol{\eta}) = \frac{X(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} = \sum_i x_i \frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})}, \quad (26)$$

where the weight function $W(\boldsymbol{\eta})$ and the weight w_i are the same as those in Eqs. (12) and (24). The n unknown coefficients x_i in $T_x(\boldsymbol{\eta})$ can be obtained by making \mathcal{DT}_x and \mathcal{GT}_x interpolate $h(\boldsymbol{\eta})$ and $h_b(\boldsymbol{\eta})$ at some sampling points, respectively, similar as (13), i.e.,

$$\begin{cases} \mathcal{DT}_x(\boldsymbol{\eta}_k) = h(\boldsymbol{\eta}_k), & k = 1, 2, \dots, n_1, \\ \mathcal{GT}_x(\boldsymbol{\eta}_l) = h_b(\boldsymbol{\eta}_l), & l = n_1 + 1, \dots, n. \end{cases} \quad (27)$$

Therefore, $\mathcal{I}^p h = \mathcal{DT}_x$.

The aforementioned linear system of equations (27) can be rewritten as

$$\begin{cases} \sum_i (x_i - q_i) \mathcal{D} \left(\frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} \right) \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}_k} = 0, & k = 1, 2, \dots, n_1, \\ \sum_i (x_i - q_i) \mathcal{G} \left(\frac{w_i B_i(\boldsymbol{\eta})}{W(\boldsymbol{\eta})} \right) \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}_l} = 0, & l = n_1 + 1, \dots, n. \end{cases} \quad (28)$$

Obviously, the coefficient matrix of (28) is the same as that of the linear system (13), and is of full rank, too. Then the linear system of equations (28) has only zero solution, i.e., $x_i = q_i$, meaning that

$$\mathcal{I}^p h = \mathcal{DT}_x = \mathcal{DT}_q = h. \quad (29)$$

Therefore, we have

$$\begin{aligned} \|\mathcal{DT}_r - \mathcal{DT}\|_{\mathbb{W}} &= \|\mathcal{I}^p f - f\|_{\mathbb{W}} = \|\mathcal{I}^p f - \mathcal{I}^p h + h - f\|_{\mathbb{W}} = \|\mathcal{I}^p(f - h) - (f - h)\|_{\mathbb{W}} \\ &\leq \|f - h\|_{\mathbb{W}} + \|\mathcal{I}^p\| \|f - h\|_{\mathbb{W}} = (1 + \|\mathcal{I}^p\|) \|f - h\|_{\mathbb{W}} \\ &= (1 + \|\mathcal{I}^p\|) \text{dist}(f, \mathbb{S}_{\rho, e}^d). \quad (\text{explained in the following paragraph}) \end{aligned} \quad (30)$$

Because $T_q(\boldsymbol{\eta})$ (24) is an arbitrary NURBS function in the spline space $\mathbb{S}_\rho(\Omega_p)$, the function $h(\boldsymbol{\eta}) = \mathcal{DT}_q$ (25) is also an arbitrary NURBS function in the linear space $\mathbb{S}_{\rho, e}^d(\Omega_p)$. So in Eq.(30), the function h can be chosen from $\mathbb{S}_{\rho, e}^d(\Omega_p)$ to make $\|f - h\|_{\mathbb{W}}$ as small as possible, that is, $\text{dist}(f, \mathbb{S}_{\rho, e}^d)$. \square

Based on Lemma 1 and 3, it follows:

Lemma 4. Suppose $\mathcal{DT} = f \in C^0(\Omega_p)$ (Eq. (1)), and $T_r \in \mathbb{S}_\rho(\Omega_p)$ (Refer to Eq. (12) and Remark 3) is the NURBS function approximating the analytical solution T . Then,

$$\|\mathcal{DT}_r - \mathcal{DT}\|_{\mathbb{W}} \leq K \|\mathcal{D}\| (1 + \|\mathcal{I}^p\|) \omega(T, \rho),$$

where \mathcal{I}^p is the interpolation operator defined by Eq. (22), and K is an integer related to the degree of the splines in the spline space $\mathbb{S}_\rho(\Omega_p)$.

Moreover, due to Lemma 2 and 4, the convergence rate of \mathcal{DT}_r to \mathcal{DT} when $\rho \rightarrow 0$ is obtained as follows.

Theorem 1. Suppose the analytical solution $T \in C^1(\Omega_p)$ (Eq. (1)). We have,

$$\|\mathcal{DT}_r - \mathcal{DT}\|_{\mathbb{W}} \leq \rho K \|\mathcal{D}\| (1 + \|\mathcal{I}^p\|) \max_{\boldsymbol{\eta} \in \Omega_p} \|\nabla T\|_E.$$

Here, \mathcal{DT} , T_r , \mathcal{I}^p , and K are delineated as in Lemma 4.

In addition, if \mathcal{D} is a stable operator (Definition 1), we can get the convergence rate of T_r to T when $\rho \rightarrow 0$.

Corollary 1. *Suppose the operator \mathcal{D} in Eq. (1) is a stable differential operator, and $T \in C^1(\Omega_p)$. We have,*

$$\|T_r - T\|_V \leq \frac{K}{C_S} \|\mathcal{D}\| (1 + \|I^\rho\|) \omega(T, \rho) \leq \frac{\rho K}{C_S} \|\mathcal{D}\| (1 + \|I^\rho\|) \max_{\eta \in \Omega_p} \|\nabla T\|_E,$$

where C_S is a positive constant, T_r , I^ρ , and K are delineated as in Lemma 4.

4.1. One dimensional case

In the one dimensional case, the convergence rate can be improved. In this section, suppose the operator \mathcal{D} is a linear differential operator with constant coefficients.

Lemma 5. [28, pp. 148] *Let $g \in C^m(\Omega_p)$ be a univariate function, and $\mathbb{S}_{\rho,e}^d(\Omega_p)$ be defined as in Remark 3. It holds,*

$$\text{dist}(g, \mathbb{S}_{\rho,e}^d) \leq \gamma \rho \text{dist}(g', \mathbb{S}_{\rho,e-1}^d), \quad (31)$$

where γ is a number related to the degree of the splines in $\mathbb{S}_{\rho,e}^d(\Omega_p)$, and g' is the first order derivative of g .

Repeatedly using Lemma 5 leads to:

Lemma 6. *Suppose $f = \mathcal{D}T \in C^m(\Omega_p)$ (Eq. (1)) is a univariate function, the linear spline space $\mathbb{S}_{\rho,e}^d(\Omega_p)$ is defined as in Remark 3, and the operator \mathcal{D} is a linear differential operator with constant coefficients. We have,*

$$\text{dist}(f, \mathbb{S}_{\rho,e}^d) = \text{dist}(\mathcal{D}T, \mathbb{S}_{\rho,e}^d) \leq \Gamma \|\mathcal{D}\| \rho^\nu \|T^{(\nu)}\|_{L^\infty},$$

where $\nu = \min(m, e)$, Γ is a number related to ν and the degree of the splines in $\mathbb{S}_{\rho,e}^d(\Omega_p)$, and $T^{(\nu)}$ is the ν^{th} order derivative of T .

Proof: Because $f = \mathcal{D}T \in C^m(\Omega_p)$ is a univariate function, and \mathcal{D} is a linear differential operator with constant coefficients, we have $(\mathcal{D}T)^{(k)} = \mathcal{D}T^{(k)}$, $k = 1, 2, \dots, m$. By using Lemma 5 repeatedly, and denoting $\nu = \min(m, e)$, it follows,

$$\begin{aligned} \text{dist}(f, \mathbb{S}_{\rho,e}^d) &= \text{dist}(\mathcal{D}T, \mathbb{S}_{\rho,e}^d) \leq \gamma_1 \rho \text{dist}((\mathcal{D}T)', \mathbb{S}_{\rho,e-1}^d) \leq \gamma_1 \gamma_2 \rho^2 \text{dist}((\mathcal{D}T)'', \mathbb{S}_{\rho,e-2}^d) \\ &\leq \dots \leq \gamma_1 \gamma_2 \dots \gamma_{\nu-1} \rho^{\nu-1} \text{dist}((\mathcal{D}T)^{(\nu-1)}, \mathbb{S}_{\rho,e-\nu+1}^d) \\ &= \gamma_1 \gamma_2 \dots \gamma_{\nu-1} \rho^{\nu-1} \text{dist}(\mathcal{D}T^{(\nu-1)}, \mathbb{S}_{\rho,e-\nu+1}^d) \\ &\leq \gamma_1 \dots \gamma_{\nu-1} \rho^{\nu-1} K_\nu \|\mathcal{D}\| \omega(T^{(\nu-1)}, \rho), \quad (\text{Lemma 1}) \end{aligned}$$

where, γ_1 is a number related to the degree of the splines in $\mathbb{S}_{\rho,e}^d$ (denoted as deg), γ_2 is a number related to the degree of the splines in $\mathbb{S}_{\rho,e-1}^d$, i.e., $\text{deg} - 1, \dots$, and so on; K_ν is a number related to the degree of the splines in $\mathbb{S}_{\rho,e-\nu+1}^d$, i.e., $\text{deg} - \nu + 1$. In conclusion, $\gamma_i, i = 1, 2, \dots, \nu - 1$, and K_ν are all related to deg and $\nu = \min(m, e)$, and then we denote $\Gamma = \gamma_1 \gamma_2 \dots \gamma_{\nu-1} K_\nu$. Moreover, by Lemma 2, we have,

$$\text{dist}(f, \mathbb{S}_{\rho,e}^d) \leq \Gamma \rho^{\nu-1} \|\mathcal{D}\| \omega(T^{(\nu-1)}, \rho) \leq \Gamma \|\mathcal{D}\| \rho^\nu \|T^{(\nu)}\|_{L^\infty},$$

where, $\nu = \min(m, e)$, and Γ is a number related to ν and the degree of the splines in $\mathbb{S}_{\rho, e}^d(\Omega_p)$. \square

Based on Lemma 3 and 6, the convergence rate for the consistency of the IGA-C method in the one-dimensional case is deduced.

Theorem 2. *Suppose $f = \mathcal{D}T \in C^m(\Omega_p)$ (Eq. (1)) is a univariate function, the spline space $\mathbb{S}_{\rho, e}^d(\Omega_p)$ is defined as in Remark 3, and the operator \mathcal{D} is a linear differential operator with constant coefficients. We have,*

$$\|\mathcal{D}T_r - \mathcal{D}T\|_{\mathbb{W}} \leq \Gamma(1 + \|I^\rho\|) \|\mathcal{D}\| \rho^\nu \|T^{(\nu)}\|_{L^\infty},$$

where $\nu = \min(m, e)$, and Γ is a number related to ν and the degree of the splines in $\mathbb{S}_{\rho, e}^d(\Omega_p)$.

Moreover, if the operator \mathcal{D} is also a stable operator (Definition 1), it holds:

Corollary 2. *Suppose $f = \mathcal{D}T \in C^m(\Omega_p)$ (Eq. (1)) is a univariate function, the spline space $\mathbb{S}_{\rho, e}^d(\Omega)$ is defined as in Remark 3, and the linear differential operator with constant coefficients \mathcal{D} is stable (refer to Definition 1). We have,*

$$\|T_r - T\|_{\mathbb{V}} \leq \frac{\Gamma}{C_S} (1 + \|I^\rho\|) \|\mathcal{D}\| \rho^\nu \|T^{(\nu)}\|_{L^\infty},$$

where C_S is a positive constant, $\nu = \min(m, e)$, and Γ is a number related to ν and the degree of the splines in $\mathbb{S}_{\rho, e}^d(\Omega_p)$.

5. The necessary and sufficient condition

In this section, we will present the necessary and sufficient condition of the consistency of the IGA-C method. Because $\mathcal{D}T = f \in C^0(\Omega_p)$ and T is continuous (Eq. (1)), we have $\omega(T, \rho) \rightarrow 0$, when $\rho \rightarrow 0$. Based on Lemma 4, if $\|I^\rho\|$ and $\|\mathcal{D}\|$ are bounded, it follows $\|\mathcal{D}T_r - \mathcal{D}T\|_{\mathbb{W}} \rightarrow 0$ when $\rho \rightarrow 0$. That is, the IGA-C method is consistency. However, since $I^\rho f = \mathcal{D}T_r \in \mathbb{S}_{\rho, e}^d$ (22), and $T_r \in \mathbb{S}_\rho$ is defined on the knot grid \mathcal{T}^ρ with knot grid size ρ , the norms $\|I\|$ and $\|\mathcal{D}\|$ are both related to the knot grid size ρ . Therefore, the sufficient condition for the consistency of the IGA-C method is followed.

Lemma 7 (Sufficiency). *If the interpolation operator I^ρ (22) and differential operator \mathcal{D} (1) are both uniformly bounded when $\rho \rightarrow 0$, then the IGA-C method applied on the boundary problem (1) is consistency.*

Furthermore, the following lemma presents the necessary condition for the consistency of the IGA-C method.

Lemma 8 (Necessity). *If the IGA-C method applied on the boundary problem (1) is consistency, then the interpolation operator I^ρ (22) and the differential operator \mathcal{D} (1) are both uniformly bounded when $\rho \rightarrow 0$.*

Proof: We employ the method of proof by contradiction to show that $\mathcal{D}T_r$ is bounded when $\rho \rightarrow 0$.

The consistency of the IGA-C method means that

$$\mathcal{D}T_r \rightarrow \mathcal{D}T = f, \text{ when } \rho \rightarrow 0. \quad (32)$$

By contradiction, suppose $\mathcal{D}T_r$ is not uniformly bounded when $\rho \rightarrow 0$, i.e., $\|\mathcal{D}T_r\|_{\mathbb{W}} \rightarrow \infty$, when $\rho \rightarrow 0$. Because f is continuous, it is bounded on its domain $\Omega_p \cup \partial\Omega_p$. However, $\mathcal{D}T_r$ is unbounded when $\rho \rightarrow 0$. This violates the consistency condition (32). So the hypothesis is not true, $\mathcal{D}T_r$ is uniformly bounded when $\rho \rightarrow 0$. That is, there exists a positive constant C_r such that

$$\|\mathcal{D}T_r\|_{\mathbb{W}} \leq C_r, \text{ when } \rho \rightarrow 0.$$

Therefore, we have

$$\|\mathcal{D}\| = \sup_{\|T_r\|_{\mathbb{V}}=1} \{\|\mathcal{D}T_r\|_{\mathbb{W}}\} \leq C_r, \text{ when } \rho \rightarrow 0,$$

and (refer to Eq. (22))

$$\|I^\rho\| = \sup_{\|f\|_{L^\infty}=1} \{\|I^\rho f\|_{\mathbb{W}}\} = \sup_{\|f\|_{L^\infty}=1} \{\|\mathcal{D}T_r\|_{\mathbb{W}}\} \leq C_r, \text{ when } \rho \rightarrow 0.$$

It means that the interpolation operator I^ρ (22) and the differential operator \mathcal{D} (1) are both uniformly bounded when $\rho \rightarrow 0$. \square

Based on Lemmas 7 and 8, the necessary and sufficient condition for the consistency of the IGA-C method is followed.

Theorem 3 (Necessity and Sufficiency). *The IGA-C method applied on the boundary problem (1) is consistency, if and only if the interpolation operator I^ρ (22) and differential operator \mathcal{D} (1) are both uniformly bounded when $\rho \rightarrow 0$.*

6. Numerical examples

In this section, some numerical examples are presented to illustrate the necessary and sufficient condition of the consistency of the IGA-C method.

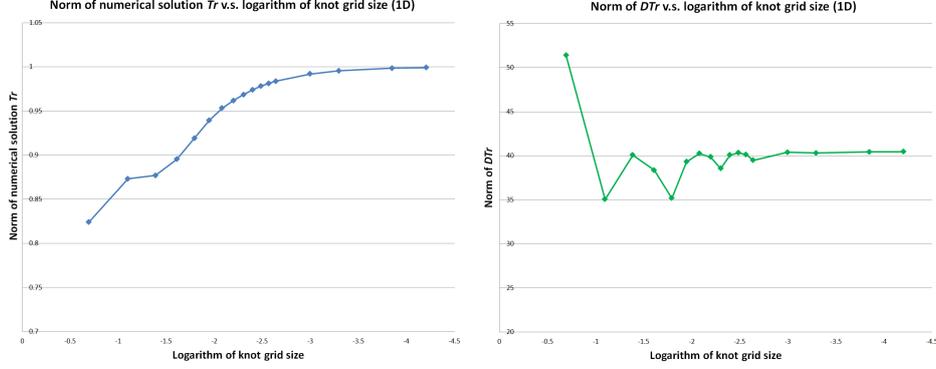
Example 1: Consider the following one-dimensional source problem:

$$\begin{cases} -T'' + T = (1 + 4\pi^2)\sin(2\pi x), & x \in \Omega = [0, 1], \\ T(0) = 0, T(1) = 0. \end{cases} \quad (33)$$

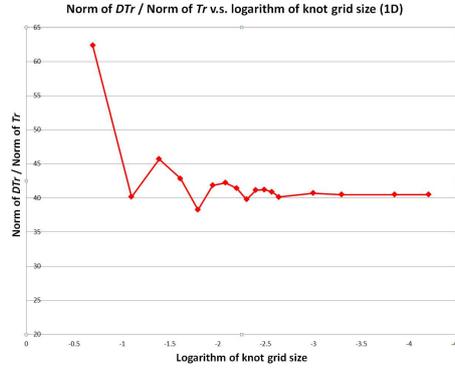
The analytical solution to the source problem is $T(x) = \sin(2\pi x)$. The physical domain is modeled by a cubic B-spline curve with control points $\{0, \frac{1}{3}, \frac{2}{3}, 1\}$ and knot vector $\{0, 0, 0, 0, 1, 1, 1, 1\}$. So the initial knot grid size is $\rho_0 = 1$. To reduce the knot grid size, we uniformly insert $k, (k = 1, 2, \dots)$ knots in $(0, 1)$. And then, the knot grid size sequence is $\rho_k = \frac{1}{k+1}, k = 0, 1, 2, \dots$.

In Fig. 1, three diagrams are demonstrated, that is,

- the norm of numerical solution T_r , i.e., $\|T_r\|_{L^\infty}$ v.s. the logarithm of knot grid size, i.e., $\ln \rho_k$ (Fig. 1(a)),
- $\|\mathcal{D}T_r\|_{L^\infty}$ v.s. $\ln \rho_k$ (Fig. 1(b)), and,
- $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ v.s. $\ln \rho_k$ (Fig. 1(c)).



(a) Diagram of $\|T_r\|_{L^\infty}$ v.s. $\ln \rho_i$ for the 1D problem. (b) Diagram of $\|\mathcal{D}T_r\|_{L^\infty}$ v.s. $\ln \rho_i$ for the 1D problem.



(c) Diagram of $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ v.s. $\ln \rho_i$ for the 1D problem.

Figure 1: In the case of one-dimensional source problem (33), $\|T_r\|_{L^\infty}$, $\|\mathcal{D}T_r\|_{L^\infty}$, and the ratio $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ are all uniformly bounded when the knot grid size sequence $\rho_k \rightarrow 0$, ($k \rightarrow \infty$).

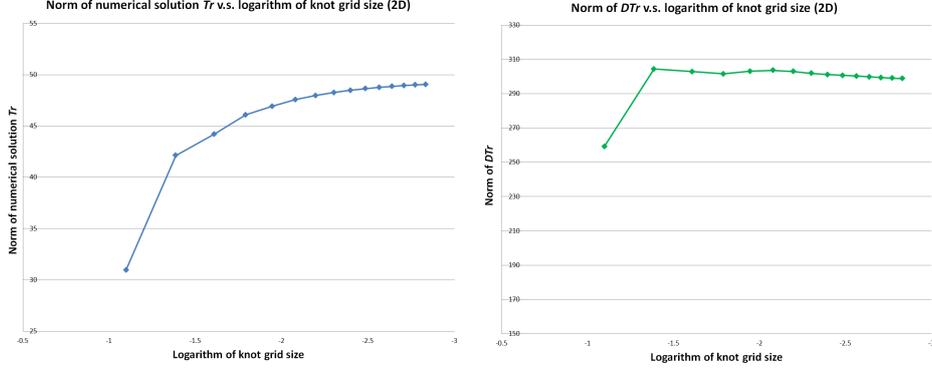
It can be seen from the diagrams in Fig. 1 that, when $k \rightarrow \infty$ and $\rho_k \rightarrow 0$, the norm of the numerical solution $\|T_r\|_{L^\infty}$ tends to the norm of the analytical solution, i.e., $\|T(x)\|_{L^\infty} = \|\sin(2\pi x)\|_{L^\infty} = 1$ (Fig. 1(a)), and $\|\mathcal{D}T_r\|_{L^\infty}$ tends to the norm of $f(x) = (1 + 4\pi^2)\sin(2\pi x)$ (33), i.e., $\|(1 + 4\pi^2)\sin(2\pi x)\|_{L^\infty} = 1 + 4\pi^2$ (Fig. 1(b)). Moreover, refer to Fig. 1(c), as an indicator of $\|\mathcal{D}\|_{L^\infty}$, the ratio $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ tends to $\frac{\|(1+4\pi^2)\sin(2\pi x)\|_{L^\infty}}{\|\sin(2\pi x)\|_{L^\infty}} = 1 + 4\pi^2$, when $\rho_k \rightarrow 0$, ($k \rightarrow \infty$). Therefore, it is uniformly bounded as $\rho_k \rightarrow 0$, ($k \rightarrow \infty$), which validates Theorem 3.

Example 2: The next example is a two-dimensional source problem:

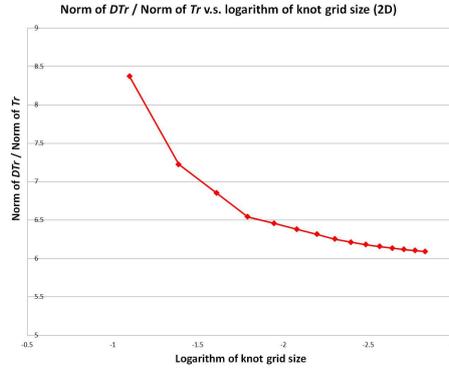
$$\begin{cases} -\Delta T + T = f, & (x, y) \in \Omega \\ T|_{\partial\Omega} = 0, \end{cases} \quad (34)$$

where,

$$\begin{aligned} f = & (3x^4 - 67x^2 - 67y^2 + 3y^4 + 6x^2y^2 + 116) \sin(x) \sin(y) \\ & + (68x - 8x^3 - 8xy^2) \cos(x) \sin(y) \\ & + (68y - 8y^3 - 8yx^2) \cos(y) \sin(x). \end{aligned}$$



(a) Diagram of $\|T_r\|_{L^\infty}$ v.s. $\ln \rho_i$ for the 2D problem. (b) Diagram of $\|\mathcal{D}T_r\|_{L^\infty}$ v.s. $\ln \rho_i$ for the 2D problem.



(c) Diagram of $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ v.s. $\ln \rho_i$ for the 2D problem.

Figure 2: In the case of two-dimensional source problem (34), $\|T_r\|_{L^\infty}$, $\|\mathcal{D}T_r\|_{L^\infty}$, and the ratio $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ are all uniformly bounded when the knot grid size sequence $\rho_k \rightarrow 0$, ($k \rightarrow \infty$).

And the analytical solution of the source problem (34) is

$$T = (x^2 + y^2 - 1)(x^2 + y^2 - 16) \sin(x) \sin(y).$$

The physical domain Ω in Eq. (34) is a quarter of an annulus, which is represented by a cubic NURBS patch with 4×4 control points. The control points and weights of the cubic NURBS patch are listed in Tables 1 and 2, respectively. The knot vectors of the cubic NURBS patch along u - and v -direction are, respectively,

$$\begin{aligned} &0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1, \\ &0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1. \end{aligned}$$

To make the knot grid size tend to 0, we uniformly insert knots in the interval $(0, 1)$ along u - and v -directions, respectively. So, the knot grid sizes are $\rho_k = \frac{1}{k+1}$, $k = 0, 1, 2, \dots$.

Fig. 2 shows the diagrams $\|T_r\|_{L^\infty}$ v.s. $\ln \rho_k$ (Fig. 2(a)), $\|\mathcal{D}T_r\|_{L^\infty}$ v.s. $\ln \rho_k$ (Fig. 2(b)), and $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ v.s. $\ln \rho_k$ (Fig. 2(c)) for the case of two-dimensional source problem (34). Similar as

the case of one-dimensional problem, $\|T_r\|_{L^\infty}$, $\|\mathcal{D}T_r\|_{L^\infty}$, and $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ are all have limit when $\rho_k \rightarrow 0, (k \rightarrow \infty)$. So they are all uniformly bounded when $\rho_k \rightarrow 0, (k \rightarrow \infty)$.

Table 1: Control points of the quarter of annulus

i	$P_{i,1}$	$P_{i,2}$	$P_{i,3}$	$P_{i,4}$
1	(1,0)	(2,0)	(3,0)	(4,0)
2	(1,2- $\sqrt{2}$)	(2, 4-2 $\sqrt{2}$)	(3,6-3 $\sqrt{2}$)	(4,8-4 $\sqrt{2}$)
3	(2- $\sqrt{2}$,1)	(4-2 $\sqrt{2}$,2)	(6-3 $\sqrt{2}$,3)	(8-4 $\sqrt{2}$, 4)
4	(0,1)	(0,2)	(0,3)	(0,4)

Table 2: Weights for the quarter of annulus

i	$w_{i,1}$	$w_{i,2}$	$w_{i,3}$	$w_{i,4}$
1	1	1	1	1
2	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$
3	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$
4	1	1	1	1

Example 3: The final example is a three-dimensional source problem:

$$\begin{cases} -\Delta T + T = f, & (x, y, z) \in \Omega, \\ T|_{\partial\Omega} = 0, \end{cases} \quad (35)$$

where

$$f = (1 + 12\pi^2) \sin(2\pi x) \sin(2\pi y) \sin(2\pi z),$$

and the analytical solution is,

$$T = \sin(2\pi x) \sin(2\pi y) \sin(2\pi z).$$

The physical domain Ω is modeled as a cubic trivariate B-spline solid with control points $P_{ijk} = (\frac{i}{3}, \frac{j}{3}, \frac{k}{3})$, $i, j, k = 0, 1, 2, 3$, and knot vectors along u -, v -, and w -directions, respectively,

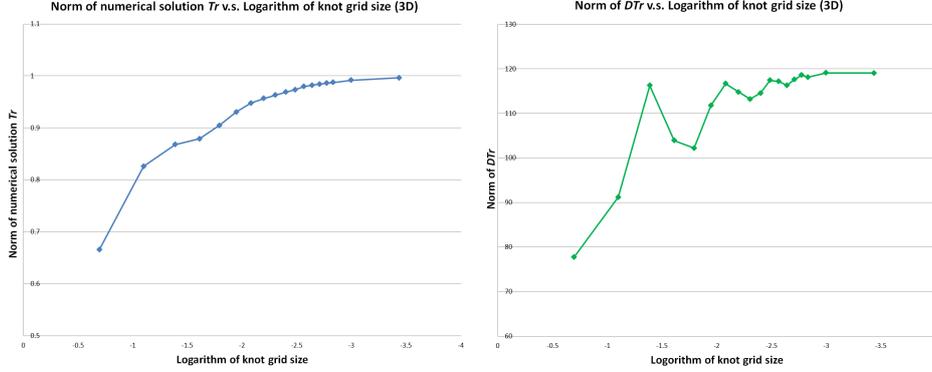
$$\begin{aligned} &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1, \\ &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1, \\ &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1. \end{aligned}$$

Similar as the one and two dimensional cases, the intervals $(0, 1)$ along u -, v -, and w -directions are uniformly inserted knots, respectively. Therefore, the knot grid size $\rho_k = \frac{1}{k+1} \rightarrow 0, k = 0, 1, 2, \dots$.

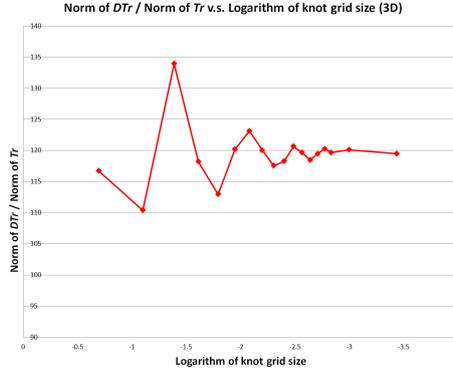
The three diagrams, i.e., $\|T_r\|_{L^\infty}$ v.s. $\ln \rho_k$ (Fig. 3(a)), $\|\mathcal{D}T_r\|_{L^\infty}$ v.s. $\ln \rho_k$ (Fig. 3(b)), and $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ v.s. $\ln \rho_k$ (Fig. 3(c)) for the case of three-dimensional source problem (35) are illustrated in Fig. 3. From these diagrams, we can see that, when $\rho_k \rightarrow 0 (k \rightarrow +\infty)$,

$$\|T_r\|_{L^\infty} \rightarrow \|T\|_{L^\infty} = 1, \quad \|\mathcal{D}T_r\|_{L^\infty} \rightarrow \|f\|_{L^\infty} = 1 + 12\pi^2, \quad \text{and,} \quad \frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}} \rightarrow 1 + 12\pi^2.$$

So they are all uniformly bounded when $\rho_k \rightarrow 0 (k \rightarrow \infty)$, too.



(a) Diagram of $\|T_r\|_{L^\infty}$ v.s. $\ln \rho_i$ for the 3D problem. (b) Diagram of $\|\mathcal{D}T_r\|_{L^\infty}$ v.s. $\ln \rho_i$ for the 3D problem.



(c) Diagram of $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ v.s. $\ln \rho_i$ for the 3D problem.

Figure 3: In the case of three-dimensional source problem (35), $\|T_r\|_{L^\infty}$, $\|\mathcal{D}T_r\|_{L^\infty}$, and the ratio $\frac{\|\mathcal{D}T_r\|_{L^\infty}}{\|T_r\|_{L^\infty}}$ are all uniformly bounded when the knot grid size sequence $\rho_k \rightarrow 0$, ($k \rightarrow \infty$).

7. Conclusions

In this paper, we developed the convergence order for the consistency and convergence of the IGA-C method, and then, deduced the necessary-and-sufficient condition for the consistency of the IGA-C method. Specifically, suppose \mathcal{D} is the differential operator of a boundary value problem with $\mathcal{D}T = f(1)$, a NURBS function T_r is the numerical solution, and I^ρ is an interpolation operator such that $I^\rho f = \mathcal{D}T_r$. First, the formula of the convergence order for the consistency of the IGA-C method is developed, which includes the norms of the operator \mathcal{D} and I^ρ . Then, the necessary-and-sufficient condition for the consistency of the IGA-C method is deduced. That is, the IGA-C method is consistency if and only if \mathcal{D} and I^ρ are both uniformly bounded when $\rho \rightarrow 0$. These results will advance the numerical analysis of the IGA-C method.

Acknowledgement

This work is supported by the Natural Science Foundation of China (Nos. 61379072, 61202201). Dr. Qianqian Hu is also supported by the Open Project Program (No. A1305) of the State Key Lab of CAD&CG, Zhejiang University.

References

- [1] T. Hughes, J. Cottrell, Y. Bazilevs, Isogeometric analysis: Cad, finite elements, nurbs, exact geometry and mesh refinement, *Computer methods in applied mechanics and engineering* 194 (39) (2005) 4135–4195.
- [2] F. Auricchio, L. Beirão da Veiga, T. Hughes, A. Reali, G. Sangalli, Isogeometric collocation methods, *Mathematical Models and Methods in Applied Sciences* 20 (11) (2010) 2075–2107.
- [3] D. Schillinger, J. A. Evans, A. Reali, M. A. Scott, T. J. Hughes, Isogeometric collocation: Cost comparison with galerkin methods and extension to adaptive hierarchical nurbs discretizations, *Computer Methods in Applied Mechanics and Engineering*, in press.
- [4] H. Lin, Q. Hu, Y. Xiong, Consistency and convergence properties of the isogeometric collocation method, *Computer Methods in Applied Mechanics and Engineering* 267 (2013) 471–486.
- [5] F. Auricchio, L. Beirão da Veiga, A. Buffa, C. Lovadina, A. Reali, G. Sangalli, A fully locking-free isogeometric approach for plane linear elasticity problems: a stream function formulation, *Computer methods in applied mechanics and engineering* 197 (1) (2007) 160–172.
- [6] T. Elguedj, Y. Bazilevs, V. Calo, T. Hughes, \bar{B} and \bar{F} projection methods for nearly incompressible linear and non-linear elasticity and plasticity using higher-order nurbs elements, *Comput. Methods Appl. Mech. Engrg* 197 (2008) 2732–2762.
- [7] J. Cottrell, A. Reali, Y. Bazilevs, T. Hughes, Isogeometric analysis of structural vibrations, *Computer methods in applied mechanics and engineering* 195 (41) (2006) 5257–5296.
- [8] T. Hughes, A. Reali, G. Sangalli, Duality and unified analysis of discrete approximations in structural dynamics and wave propagation: Comparison of p -method finite elements with k -method nurbs, *Computer methods in applied mechanics and engineering* 197 (49) (2008) 4104–4124.
- [9] W. Wall, M. Frenzel, C. Cyron, Isogeometric structural shape optimization, *Computer Methods in Applied Mechanics and Engineering* 197 (33) (2008) 2976–2988.
- [10] Y. Bazilevs, V. Calo, T. Hughes, Y. Zhang, Isogeometric fluid-structure interaction: theory, algorithms, and computations, *Computational Mechanics* 43 (1) (2008) 3–37.
- [11] Y. Bazilevs, V. Calo, Y. Zhang, T. Hughes, Isogeometric fluid-structure interaction analysis with applications to arterial blood flow, *Computational Mechanics* 38 (4) (2006) 310–322.
- [12] Y. Bazilevs, J. Gohean, T. Hughes, R. Moser, Y. Zhang, Patient-specific isogeometric fluid-structure interaction analysis of thoracic aortic blood flow due to implantation of the jarvik 2000 left ventricular assist device, *Computer Methods in Applied Mechanics and Engineering* 198 (45) (2009) 3534–3550.
- [13] Y. Bazilevs, L. Beirão da Veiga, J. Cottrell, T. Hughes, G. Sangalli, Isogeometric analysis: approximation, stability and error estimates for h-refined meshes, *Mathematical Models and Methods in Applied Sciences* 16 (07) (2006) 1031–1090.
- [14] J. Cottrell, T. Hughes, A. Reali, Studies of refinement and continuity in isogeometric structural analysis, *Computer methods in applied mechanics and engineering* 196 (41) (2007) 4160–4183.
- [15] T. Hughes, A. Reali, G. Sangalli, Efficient quadrature for nurbs-based isogeometric analysis, *Computer methods in applied mechanics and engineering* 199 (5) (2010) 301–313.
- [16] M. Aigner, C. Heinrich, B. Jüttler, E. Pilgerstorfer, B. Simeon, A. Vuong, Swept volume parameterization for isogeometric analysis, *Mathematics of Surfaces XIII* (2009) 19–44.
- [17] G. Xu, B. Mourrain, R. Duvigneau, A. Galligo, Optimal analysis-aware parameterization of computational domain in 3d isogeometric analysis, *Computer-Aided Design* 45 (4) (2013) 812–821.
- [18] G. Xu, B. Mourrain, R. Duvigneau, A. Galligo, Parameterization of computational domain in isogeometric analysis: methods and comparison, *Computer Methods in Applied Mechanics and Engineering* 200 (23) (2011) 2021–2031.
- [19] M. Donatelli, C. Garoni, C. Manni, S. Serra-Capizzano, H. Speleers, Robust and optimal multi-iterative techniques for iga galerkin linear systems, *Computer Methods in Applied Mechanics and Engineering* 284 (2015) 230–264.
- [20] T. Elguedj, T. Hughes, Isogeometric analysis of nearly incompressible large strain plasticity, *Computer Methods in Applied Mechanics and Engineering* 268 (2014) 388 – 416.
- [21] L. De Lorenzis, J. Evans, T. Hughes, A. Reali, Isogeometric collocation: Neumann boundary conditions and contact, *Computer Methods in Applied Mechanics and Engineering* 284 (2015) 21–54.

- [22] L. Beirão da Veiga, C. Lovadina, A. Reali, Avoiding shear locking for the timoshenko beam problem via isogeometric collocation methods, *Computer methods in applied mechanics and engineering* 241 (2012) 38–51.
- [23] F. Auricchio, L. Beirão da Veiga, J. Kiendl, C. Lovadina, A. Reali, Locking-free isogeometric collocation methods for spatial timoshenko rods, *Computer Methods in Applied Mechanics and Engineering* 263 (15) (2013) 113–126.
- [24] F. Auricchio, L. Beirão da Veiga, T. Hughes, A. Reali, G. Sangalli, Isogeometric collocation for elastostatics and explicit dynamics, *Computer methods in applied mechanics and engineering* 249 (2012) 2–14.
- [25] A. Reali, H. Gomez, An isogeometric collocation approach for bernoulli–euler beams and kirchhoff plates, *Computer Methods in Applied Mechanics and Engineering* 284 (2015) 623–636.
- [26] J. Kiendl, F. Auricchio, L. B. da Veiga, C. Lovadina, A. Reali, Isogeometric collocation methods for the reissner–mindlin plate problem, *Computer Methods in Applied Mechanics and Engineering* 284 (2015) 489–507.
- [27] P. Solin, *Partial differential equations and the finite element method*, Wiley-Interscience, 2006.
- [28] C. De Boor, *A practical guide to splines*, Vol. 27, Springer Verlag, 2001.