Original articles

A representation and comparison of three cubic macro-elements

Ema Ĉešek [a], Jan Grošelj [a,b,*], Andrej Kolar-Požun [c,a], Maruša Lekše [a,b],
Gašper Domen Romih [a], Ada Šadl Praprotnik [a,b], Matija Šteblaj [a]

[a] *Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia*
[b] *Institute of Mathematics, Physics and Mechanics, Jadranska 19, 1000 Ljubljana, Slovenia*
[c] *Parallel and Distributed Systems Laboratory, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia*

ARTICLE INFO

ABSTRACT

The paper is concerned with three types of cubic splines over a triangulation that are characterized by three degrees of freedom associated with each vertex of the triangulation. The splines differ in computational complexity, polynomial reproduction properties, and smoothness. With the aim to make them a versatile tool for numerical analysis, a unified representation in terms of locally supported basis functions is established. The construction of these functions is based on geometric concepts and is expressed in the Bernstein–Bézier form. They are readily applicable in a range of standard approximation methods, which is demonstrated by a number of numerical experiments.

1. Introduction

Splines over triangulations are a widely used representation framework in numerical approximation methods. Thanks to the flexibility of triangulations as a meshing technique, they can be used to approximate data on complicated domains and, within a domain, by different levels of coarseness. In the classical sense, a spline over a triangulation is a function whose restriction to each triangle of the triangulation is a bivariate polynomial of a fixed total degree. At vertices and across edges of the triangles, these polynomials are glued together by a certain order of smoothness. High order of smoothness is desirable but not easy to achieve as smoothness constraints depend on geometry of the triangulation. For low polynomial degrees (relative to smoothness order) this prevents local and stable construction of general splines and affects their approximation power (see, e.g., [8]).

In this paper we consider three macro-elements, i.e., splines over a triangulation that are on each triangle uniquely determined by a limited number of degrees of freedom related to the region in the vicinity of the triangle. Traditionally, such splines are characterized by interpolation functionals associated with vertices of the triangulation and often (although undesirably) with additional points on the edges or inside the triangles (see, e.g., [2,8]). We focus on splines that are locally polynomials of total degree at most three, i.e., cubic splines, which are practically attractive on account of low evaluation cost. However, constructing a fast convergent and smooth approximation of this type is not trivial.

In order to construct a spline that is C^1-smooth at the vertices, it is natural to fix three degrees of freedom for each vertex which intrinsically represent the value and the gradient of the spline at the vertex. On a triangle of the triangulation this amounts to nine degrees of freedom if the three vertices of the triangle are considered. Since a bivariate cubic polynomial is determined by ten degrees of freedom, there is only one free parameter left for each triangle. This is not sufficient to impose global C^1-smoothness. Keeping the construction limited to the nine degrees of freedom corresponding to a triangle, it is not even possible to achieve

---

cubic precision. On a more positive note, it is sufficient to ensure reproduction of quadratic polynomials, e.g., by using the classical Zienkiewicz element [1], which is the first macro-element that we take into consideration.

If the spline construction does not ensure the reproduction of polynomials to the highest possible degree, we cannot account for optimal convergence rates in approximation methods. To overcome this, we can sacrifice the complete locality of the construction on a triangle and use also the degrees of freedom corresponding to the neighboring triangles. Using the technique for imposing C^1-smoothness across an edge proposed in [5], we develop a C^0-smooth macro-element that is, for a triangle with no edges on the boundary of the domain, determined by eighteen degrees of freedom.

Global C^1-smoothness of a macro-element is in approximation methods not always necessary and is indeed often renounced to keep the construction computationally simple and effective. However, it often has a positive effect on the accuracy of solutions and also provides visually more pleasant results. A common approach for achieving C^1-smoothness without increasing the spline degree is to refine the triangulation in such a way that each triangle is split into three smaller triangles. A classical Clough–Tocher spline [3] over such a refinement has three degrees of freedom per vertex and additional degrees of freedom related to directional derivatives along the edges. It is possible to reduce these extra degrees of freedom by restricting the directional derivatives to be linear but such a constraint only allows reproduction of quadratic polynomials. We instead follow the construction introduced in [11] that ensures full reproduction. This provides a C^1-smooth macro-element with cubic precision that is based on the same amount of data as the second macro-element.

With the aim to use the described constructions in approximation methods, we establish a unified representation of the three macro-elements in terms of locally supported basis functions. They are defined and analyzed using the Bernstein–Bézier techniques. For each vertex, there are three basis functions that are uniquely determined by a choice of a triangle associated with the vertex. This geometric approach reflects recent ideas in constructing B-spline-like bases for other common macro-elements (see, e.g., [4,6,7,9,14]).

Finally, we provide a numerical comparison of the macro-elements. The proposed representation is utilized in function approximation and solving boundary value problems. We perform tests of the best L^2 function approximation and scattered data approximation. In the context of boundary value problems, we consider the model Poisson problem and conduct examples showing the behavior of the macro-elements in the isoparametric finite element method and immersed penalized boundary method.

The remainder of the paper is organized as follows. In Section 2 we review relevant Bernstein–Bézier techniques for representing cubic splines and imposing C^1-smoothness. In Section 3 we provide the construction of the three macro-elements and define basis functions associated with the vertices of the triangulation. Section 4 contains numerical examples. We conclude the paper with some remarks in Section 5.

2. Bernstein–Bézier spline techniques

The aim of this section is to present some preliminaries related to the construction of cubic splines over a triangulation in terms of the Bernstein–Bézier form. We give a particular emphasis on the constraints for (local) C^1-smoothness that is tackled in Section 3.

2.1. Cubic splines over a triangulation

Let $\Theta \subset \mathbb{R}^2$ be a polygonal domain. A triangulation $\triangle = (V, E, T)$ of Θ is a discretization of the domain determined by a finite set of vertices V, a finite set of edges E, and a finite set of triangles T.

More precisely, for $n_v \in \mathbb{N}$, the set $V = \{v_1, v_2, \ldots, v_{n_v}\}$ is a set of points in Θ. The set $E = \{e_1, e_2, \ldots, e_{n_e}\}$ consists of $n_e \in \mathbb{N}$ line segments, each of the form $e_l = [v_i, v_j]$ for a pair of two distinct vertices $v_i, v_j \in V$. Similarly, the set $T = \{t_1, t_2, \ldots, t_{n_t}\}$ consists of $n_t \in \mathbb{N}$ triangles, each of the form $t_m = [v_i, v_j, v_k]$ for a triplet of three affinely independent vertices $v_i, v_j, v_k \in V$. Vice versa, V and E are the set of vertices and the set of edges of all triangles in T, respectively. We impose that the intersection of two distinct triangles from T is empty, a point in V, or a line segment in E, and the union of triangles in T is Θ. Additionally, we denote by $E^b \subseteq E$ the set of cardinality $n_e^b \in \mathbb{N}$ comprising all edges that lie on the boundary of Θ.

The vector space of cubic splines over a triangulation $\triangle$ of smoothness order $r \in \{0, 1\}$ is defined by

$$\mathbb{S}_3^r(\triangle) := \left\{ S \in C^r(\Theta) : \ S|_{t_m} \in \mathbb{P}_3, \ t_m \in T \right\},$$

where, in general, $\mathbb{P}_d$ denotes the space of bivariate polynomials of total degree at most $d \in \mathbb{N}_0$. The dimension of $\mathbb{S}_3^0(\triangle)$ is $n_v + 2n_e + n_t$. Expressing the dimension of $\mathbb{S}_3^1(\triangle)$ is much more complicated and is theoretically not fully resolved. It is known that

$$\dim(\mathbb{S}_3^1(\triangle)) \geq 3n_v^b + 2(n_v - n_v^b) + 1, \tag{1}$$

where $n_v^b \in \mathbb{N}$ is the number of points in V that lie on the boundary of Θ. Moreover, the dimension depends on the geometry of the triangulation and is larger than the lower bound in (1) if the triangulation has vertices in the interior of Θ in which the edges meet with only two different slopes (see, e.g., [8], for more details). Therefore, it is unclear how to construct and use such splines in approximation methods. In Section 3 we consider related spline spaces that can be more easily characterized.

2.2. Bernstein–Bézier form

As the restriction $S|_{t_m}$ of a spline $S \in \mathbb{S}_3^0(\triangle)$ to a triangle $t_m = [v_i, v_j, v_k] \in T$ is an element of $\mathbb{P}_3$, it can be uniquely expressed in the cubic Bernstein basis. Let

$$\mathbb{D}_3 := \left\{ (\delta_0, \delta_1, \delta_2) \in \mathbb{N}_0^3 : \ \delta_0 + \delta_1 + \delta_2 = 3 \right\}$$

be a set consisting of ten multi-indices. For $\boldsymbol{d} = (d_0, d_1, d_2) \in \mathbb{D}_3$ we define the Bernstein basis polynomial $B_{m,\boldsymbol{d}} \in \mathbb{P}_3$ associated with the triangle t_m as

$$B_{m,\boldsymbol{d}}(\boldsymbol{p}) := \frac{3!}{d_0! \, d_1! \, d_2!} \left(\tau_{m,i}(\boldsymbol{p}) \right)^{d_0} \left(\tau_{m,j}(\boldsymbol{p}) \right)^{d_1} \left(\tau_{m,k}(\boldsymbol{p}) \right)^{d_2},$$

where $(\tau_{m,i}, \tau_{m,j}(\boldsymbol{p}), \tau_{m,k}(\boldsymbol{p})) \in \mathbb{R}^3$ denotes the barycentric coordinates of a point $\boldsymbol{p} \in \mathbb{R}^2$ with respect to t_m, i.e., the unique triplet satisfying

$$1 = \tau_{m,i}(\boldsymbol{p}) + \tau_{m,j}(\boldsymbol{p}) + \tau_{m,k}(\boldsymbol{p}), \quad \boldsymbol{p} = \tau_{m,i}(\boldsymbol{p})\boldsymbol{v}_i + \tau_{m,j}(\boldsymbol{p})\boldsymbol{v}_j + \tau_{m,k}(\boldsymbol{p})\boldsymbol{v}_k.$$

It is well known that the polynomial $S|_{t_m}$ can be uniquely represented in the Bernstein–Bézier form

$$S|_{t_m} = \sum_{\boldsymbol{d} \in \mathbb{D}_3} \beta_{m,\boldsymbol{d}}(S) \, B_{m,\boldsymbol{d}},$$

where $\beta_{m,\boldsymbol{d}} : \mathbb{S}_3^0(\triangle) \to \mathbb{R}$ applied to S corresponds to the dual functional of $B_{m,\boldsymbol{d}}$ applied to $S|_{t_m}$. We refer the reader to, e.g., [8] for more details on this topic.

2.3. Techniques for imposing C^1-smoothness

A spline $S \in \mathbb{S}_3^0(\triangle)$ is C^1-smooth at a vertex $\boldsymbol{v}_i \in V$ if and only if there exists a polynomial $P \in \mathbb{P}_3$ such that $S(\boldsymbol{v}_i) = P(\boldsymbol{v}_i)$ and $\nabla S|_{t_m}(\boldsymbol{v}_i) = \nabla P(\boldsymbol{v}_i)$ for every triangle $t_m \in T$ with a vertex at $\boldsymbol{v}_i$. Here and hereafter, ∇ denotes the gradient operator. Let us consider the first order Taylor polynomial

$$Q_i(\boldsymbol{p}) = P(\boldsymbol{v}_i) + \nabla P(\boldsymbol{v}_i) \cdot (\boldsymbol{p} - \boldsymbol{v}_i),$$

where $\cdot$ denotes the inner product of two vectors. As $Q_i(\boldsymbol{v}_i) = P(\boldsymbol{v}_i)$ and $\nabla Q_i(\boldsymbol{v}_i) = \nabla P(\boldsymbol{v}_i)$, the standard vertex interpolation formulas for the Bernstein–Bézier form (see, e.g., [8]) imply

$$\beta_{m,(3,0,0)}(S) = Q_i(\boldsymbol{v}_i), \quad \beta_{m,(2,1,0)}(S) = Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_j), \quad \beta_{m,(2,0,1)}(S) = Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_k), \qquad (2)$$

where $t_m = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k]$. This provides a convenient alternative characterization of C^1-smoothness at a vertex, i.e., S is C^1-smooth at $\boldsymbol{v}_i$ if and only if there exists a linear polynomial Q_i such that (2) holds for every triangle $t_m \in T$ with a vertex at $\boldsymbol{v}_i$. Hence, by fixing a linear polynomial for each vertex, all but one coefficient in the Bernstein–Bézier form of $S|_{t_m}$ are uniquely determined.

Let us consider triangles $t_m = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] \in T$ and $t_{m'} = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_{k'}] \in T$ that have the edge $e_l = [\boldsymbol{v}_i, \boldsymbol{v}_j] \in E \setminus E^b$ in common. Provided that S is C^1-smooth at the vertices $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$, the C^1-smoothness across $[\boldsymbol{v}_i, \boldsymbol{v}_j]$ can be achieved by selecting $\beta_{m,(1,1,1)}(S)$ and $\beta_{m',(1,1,1)}(S)$ such that the value of

$$\xi_l(S) := \tau_{m,i}(\boldsymbol{v}_{k'})\beta_{m,(2,1,0)}(S) + \tau_{m,j}(\boldsymbol{v}_{k'})\beta_{m,(1,2,0)}(S) + \tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(1,1,1)}(S) - \beta_{m',(1,1,1)}(S) \qquad (3)$$

is equal to zero. However, imposing C^1-smoothness across all interior edges of $\triangle$ is challenging as it results in a global system of equations depending on the geometry of the triangulation. In fact, since for certain triangulations the lower bound (1) is attained, fixing three degrees of freedom per vertex implies that such a global system in general does not have a solution.

A convenient technique for imposing C^1-smoothness across a single edge e_l without selecting an additional parameter was proposed in [5]. If $S|_{t_m \cup t_{m'}}$ is C^2-smooth, then

$$
\begin{aligned}
\beta_{m',(1,0,2)}(S) = {} & \tau_{m,i}(\boldsymbol{v}_{k'})^2 \beta_{m,(3,0,0)}(S) + 2\tau_{m,i}(\boldsymbol{v}_{k'})\tau_{m,j}(\boldsymbol{v}_{k'})\beta_{m,(2,1,0)}(S) \\
& + \tau_{m,j}(\boldsymbol{v}_{k'})^2 \beta_{m,(1,2,0)}(S) + 2\tau_{m,i}(\boldsymbol{v}_{k'})\tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(2,0,1)}(S) \\
& + 2\tau_{m,j}(\boldsymbol{v}_{k'})\tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(1,1,1)}(S) + \tau_{m,k}(\boldsymbol{v}_{k'})^2 \beta_{m,(1,0,2)}(S)
\end{aligned}
\qquad (4a)
$$

and

$$
\begin{aligned}
\beta_{m',(0,1,2)}(S) = {} & \tau_{m,i}(\boldsymbol{v}_{k'})^2 \beta_{m,(2,1,0)}(S) + 2\tau_{m,i}(\boldsymbol{v}_{k'})\tau_{m,j}(\boldsymbol{v}_{k'})\beta_{m,(1,2,0)}(S) \\
& + \tau_{m,j}(\boldsymbol{v}_{k'})^2 \beta_{m,(0,3,0)}(S) + 2\tau_{m,i}(\boldsymbol{v}_{k'})\tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(1,1,1)}(S) \\
& + 2\tau_{m,j}(\boldsymbol{v}_{k'})\tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(0,2,1)}(S) + \tau_{m,k}(\boldsymbol{v}_{k'})^2 \beta_{m,(0,1,2)}(S).
\end{aligned}
\qquad (4b)
$$

All the values appearing in (4) are fixed by vertex interpolation, i.e., by expressions analogous to (2), except for $\beta_{m,(1,1,1)}(S)$. Motivated by this, we define the functional

$$
\begin{aligned}
\gamma_{m,k}(S) := {} & \frac{1}{2\tau_{m,k}(\boldsymbol{v}_{k'})(1 - \tau_{m,k}(\boldsymbol{v}_{k'}))} \Big(\beta_{m',(1,0,2)}(S) + \beta_{m',(0,1,2)}(S) \\
& - \tau_{m,i}(\boldsymbol{v}_{k'})^2 (\beta_{m,(3,0,0)}(S) + \beta_{m,(2,1,0)}(S)) \\
& - 2\tau_{m,i}(\boldsymbol{v}_{k'})\tau_{m,j}(\boldsymbol{v}_{k'})(\beta_{m,(2,1,0)}(S) + \beta_{m,(1,2,0)}(S)) \\
& - \tau_{m,j}(\boldsymbol{v}_{k'})^2 (\beta_{m,(1,2,0)}(S) + \beta_{m,(0,3,0)}(S)) \\
& - 2\tau_{m,i}(\boldsymbol{v}_{k'})\tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(2,0,1)}(S) \\
& - 2\tau_{m,j}(\boldsymbol{v}_{k'})\tau_{m,k}(\boldsymbol{v}_{k'})\beta_{m,(0,2,1)}(S) \\
& - \tau_{m,k}(\boldsymbol{v}_{k'})^2 (\beta_{m,(1,0,2)}(S) + \beta_{m,(0,1,2)}(S)) \Big),
\end{aligned}
\qquad (5)
$$

which is obtained by combining the equations in (4) and expressing $\beta_{m,(1,1,1)}(S)$.

Setting $\beta_{m,(1,1,1)}(S) = \gamma_{m,k}(S)$ and $\beta_{m',(1,1,1)}(S) = \gamma_{m',k'}(S)$ has some immediate benefits. Even if the assumptions leading to (4) do not hold, such choice ensures that $\xi_l(S) = 0$, i.e., S is C^1-smooth across e_l. Moreover, if the interpolation values at the vertices $\boldsymbol{v}_i$, $\boldsymbol{v}_j$, $\boldsymbol{v}_k$, $\boldsymbol{v}_{k'}$ come from the same polynomial $P \in \mathbb{P}_3$, then $S|_{t_m \cup t_{m'}} = P$. For the proofs of these statements we refer the reader to [5].

3. Cubic macro-elements

In this section we present three constructions of cubic splines over a triangulation $\triangle$, each of them determined by three degrees of freedom associated with every vertex of $\triangle$. The splines are specified by providing the Bernstein–Bézier form on each triangle of $\triangle$ in terms of linear polynomials Q_i associated with the vertices $\boldsymbol{v}_i \in V$. For the first and third construction this is a reinterpretation of the results derived in [1,11], respectively, whereas the second construction appears to be new. In the second part of the section we propose locally supported basis functions for a numerically convenient representation of the considered splines.

3.1. C^0-smooth macro-element with quadratic precision

First, we recast the Zienkiewicz element [1] that provides a C^0-smooth spline construction over $\triangle$ with quadratic precision. Let $t_m = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] \in T$ be an arbitrary triangle of $\triangle$. We prescribe a spline $S \in \mathbb{S}_3^0(\triangle)$ by setting

$$\beta_{m,(3,0,0)}(S) = Q_i(\boldsymbol{v}_i), \qquad \beta_{m,(2,1,0)}(S) = Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_j), \qquad \beta_{m,(2,0,1)}(S) = Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_k), \tag{6a}$$

$$\beta_{m,(0,3,0)}(S) = Q_j(\boldsymbol{v}_j), \qquad \beta_{m,(0,2,1)}(S) = Q_j(\tfrac{2}{3}\boldsymbol{v}_j + \tfrac{1}{3}\boldsymbol{v}_k), \qquad \beta_{m,(1,2,0)}(S) = Q_j(\tfrac{2}{3}\boldsymbol{v}_j + \tfrac{1}{3}\boldsymbol{v}_i), \tag{6b}$$

$$\beta_{m,(0,0,3)}(S) = Q_k(\boldsymbol{v}_k), \qquad \beta_{m,(1,0,2)}(S) = Q_k(\tfrac{2}{3}\boldsymbol{v}_k + \tfrac{1}{3}\boldsymbol{v}_i), \qquad \beta_{m,(0,1,2)}(S) = Q_k(\tfrac{2}{3}\boldsymbol{v}_k + \tfrac{1}{3}\boldsymbol{v}_j), \tag{6c}$$

and

$$\beta_{m,(1,1,1)}(S) = \tfrac{1}{3}Q_i(\tfrac{1}{2}\boldsymbol{v}_i + \tfrac{1}{4}\boldsymbol{v}_j + \tfrac{1}{4}\boldsymbol{v}_k) + \tfrac{1}{3}Q_j(\tfrac{1}{2}\boldsymbol{v}_j + \tfrac{1}{4}\boldsymbol{v}_k + \tfrac{1}{4}\boldsymbol{v}_i) + \tfrac{1}{3}Q_k(\tfrac{1}{2}\boldsymbol{v}_k + \tfrac{1}{4}\boldsymbol{v}_i + \tfrac{1}{4}\boldsymbol{v}_j). \tag{7}$$

Following the discussion in Section 2.3, such a spline is C^1-smooth at each vertex of $\triangle$ but only C^0-smooth across the edges of $\triangle$.

We denote by $\mathbb{S}_1 \subset \mathbb{S}_3^0(\triangle)$ the vector space consisting of all splines that can be obtained by the described construction. It holds that $\dim(\mathbb{S}_1) = 3n_v$ and $\mathbb{P}_2 \subset \mathbb{S}_1$.

3.2. C^0-smooth macro-element with cubic precision

Next, we describe the Bernstein–Bézier form on a triangle $t_m = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] \in T$ that provides a C^0-smooth spline construction over $\triangle$ with cubic precision. For this we assume that t_m has at most two edges in E^b. Additionally, we choose the ordering of the vertices of t_m in such a way that $[\boldsymbol{v}_i, \boldsymbol{v}_j] \in E \setminus E^b$ and, if t_m has at most one boundary edge, $[\boldsymbol{v}_j, \boldsymbol{v}_k] \in E \setminus E^b$.

We prescribe a spline $S \in \mathbb{S}_3^0(\triangle)$ by setting $\beta_{m,\boldsymbol{d}}(S)$, $\boldsymbol{d} \in \mathbb{D}_3 \setminus \{(1,1,1)\}$, as in (6) and

$$\beta_{m,(1,1,1)}(S) = \begin{cases} \tfrac{1}{3}\gamma_{m,k}(S) + \tfrac{1}{3}\gamma_{m,i}(S) + \tfrac{1}{3}\gamma_{m,j}(S) & \text{if all edges of } t_m \text{ are in } E \setminus E^b, \\ \tfrac{1}{2}\gamma_{m,k}(S) + \tfrac{1}{2}\gamma_{m,i}(S) & \text{if two edges of } t_m \text{ are in } E \setminus E^b, \\ \gamma_{m,k}(S) & \text{if one edge of } t_m \text{ is in } E \setminus E^b, \end{cases}$$

where the functionals $\gamma_{m,i}$, $\gamma_{m,j}$, $\gamma_{m,k}$ are defined as in (5). Following the discussion in Section 2.3, the resulting spline is again C^1-smooth at each vertex of $\triangle$ and C^0-smooth across the edges of $\triangle$. The main difference in comparison with the spline given in Section 3.1 is that the value $\beta_{m,(1,1,1)}(S)$ depends not only on the polynomials Q_i, Q_j, Q_k but also on the polynomials associated with the vertices of the triangles in T that share an edge with t_m. Nonetheless, the construction remains local.

We denote by $\mathbb{S}_2 \subset \mathbb{S}_3^0(\triangle)$ the vector space consisting of all splines that can be obtained by the described construction. It holds that $\dim(\mathbb{S}_2) = 3n_v$ and $\mathbb{P}_3 \subset \mathbb{S}_2$.

3.3. C^1-smooth macro-element with cubic precision

In order to obtain a C^1-smooth spline, a standard macro-element approach is to split each triangle of the triangulation into a number of smaller ones. Here, we use the Clough–Tocher splitting [3] which refines a triangle $t_m = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] \in T$ into three triangles

$$t_{m_k} = [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_{ijk}], \quad t_{m_i} = [\boldsymbol{v}_j, \boldsymbol{v}_k, \boldsymbol{v}_{ijk}], \quad t_{m_j} = [\boldsymbol{v}_k, \boldsymbol{v}_i, \boldsymbol{v}_{ijk}]$$

determined by the split point $\boldsymbol{v}_{ijk} = \tfrac{1}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_j + \tfrac{1}{3}\boldsymbol{v}_k$. These triangles define a finer triangulation denoted by $\triangle_{CT}$. For an illustration, see Fig. 1 (left).

In what follows we provide a C^1-smooth spline construction over $\triangle_{CT}$ inspired by [11]. It differs from the classical one in that it does not introduce extra degrees of freedom related to directional derivatives along the edges. As it was already proposed in [3],
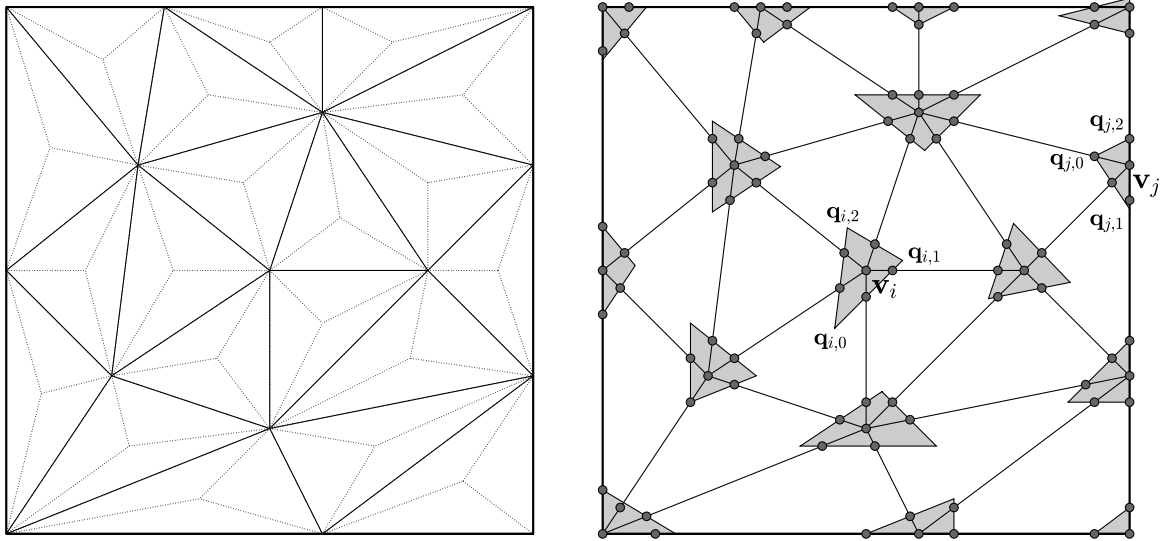
Fig. 1. A triangulation $\triangle$ of the unit square depicted by the uninterrupted lines. The figure on the left also shows the Clough–Tocher refinement $\triangle_{CT}$ of $\triangle$ determined by the dotted lines. The figure on the right depicts a configuration of triangles associated with the vertices of $\triangle$ suitable for construction of basis functions. Each triangle is obtained as the smallest triangle that contains the gray colored points in the vicinity of the vertex.

these degrees of freedom can be simply avoided by imposing that the directional derivatives are linear but such a constraint reduces precision of splines to quadratic polynomials. On the other hand, the concepts introduced in [11] ensure cubic precision.

We prescribe a spline $S \in \mathbb{S}_3^1(\triangle_{CT})$ by first setting

$$\beta_{m_k,(3,0,0)}(S) = Q_i(\boldsymbol{v}_i), \qquad \beta_{m_k,(2,1,0)}(S) = Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_j), \qquad \beta_{m_k,(2,0,1)}(S) = Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_{ijk}),$$

$$\beta_{m_k,(0,3,0)}(S) = Q_j(\boldsymbol{v}_j), \qquad \beta_{m_k,(1,2,0)}(S) = Q_j(\tfrac{2}{3}\boldsymbol{v}_j + \tfrac{1}{3}\boldsymbol{v}_i), \qquad \beta_{m_k,(0,2,1)}(S) = Q_j(\tfrac{2}{3}\boldsymbol{v}_j + \tfrac{1}{3}\boldsymbol{v}_{ijk}),$$

and, assuming t_m has at most two edges in E^b,

$$\beta_{m_k,(1,1,1)}(S) = \frac{1}{3}Q_i(\tfrac{2}{3}\boldsymbol{v}_i + \tfrac{1}{3}\boldsymbol{v}_j) + \frac{1}{3}Q_j(\tfrac{2}{3}\boldsymbol{v}_j + \tfrac{1}{3}\boldsymbol{v}_i)$$

$$+ \frac{1}{3}\begin{cases} \gamma_{m,k}(S) & \text{if } [\boldsymbol{v}_i, \boldsymbol{v}_j] \in E \setminus E^b, \\ \frac{1}{2}\gamma_{m,i}(S) + \frac{1}{2}\gamma_{m,j}(S) & \text{if } [\boldsymbol{v}_i, \boldsymbol{v}_j] \in E^b \text{ and } [\boldsymbol{v}_j, \boldsymbol{v}_k], [\boldsymbol{v}_k, \boldsymbol{v}_i] \in E \setminus E^b, \\ \gamma_{m,j}(S) & \text{if } [\boldsymbol{v}_i, \boldsymbol{v}_j], [\boldsymbol{v}_j, \boldsymbol{v}_k] \in E^b \text{ and } [\boldsymbol{v}_k, \boldsymbol{v}_i] \in E \setminus E^b, \\ \gamma_{m,i}(S) & \text{if } [\boldsymbol{v}_i, \boldsymbol{v}_j], [\boldsymbol{v}_k, \boldsymbol{v}_i] \in E^b \text{ and } [\boldsymbol{v}_j, \boldsymbol{v}_k] \in E \setminus E^b. \end{cases}$$

We analogously define the values $\beta_{m_j,(d_0,d_1,d_2)}(S)$ and $\beta_{m_i,(d_0,d_1,d_2)}(S)$ for $d_2 \leq 1$. The remaining values are given by

$$\beta_{m_k,(1,0,2)}(S) = \beta_{m_j,(0,1,2)}(S) = \tfrac{1}{3}\beta_{m_k,(2,0,1)}(S) + \tfrac{1}{3}\beta_{m_k,(1,1,1)}(S) + \tfrac{1}{3}\beta_{m_j,(1,1,1)}(S),$$

$$\beta_{m_i,(1,0,2)}(S) = \beta_{m_k,(0,1,2)}(S) = \tfrac{1}{3}\beta_{m_i,(2,0,1)}(S) + \tfrac{1}{3}\beta_{m_i,(1,1,1)}(S) + \tfrac{1}{3}\beta_{m_k,(1,1,1)}(S),$$

$$\beta_{m_j,(1,0,2)}(S) = \beta_{m_i,(0,1,2)}(S) = \tfrac{1}{3}\beta_{m_j,(2,0,1)}(S) + \tfrac{1}{3}\beta_{m_j,(1,1,1)}(S) + \tfrac{1}{3}\beta_{m_i,(1,1,1)}(S),$$

and

$$\beta_{m_k,(0,0,3)}(S) = \beta_{m_i,(0,0,3)}(S) = \beta_{m_j,(0,0,3)}(S) = \tfrac{1}{3}\beta_{m_k,(1,0,2)}(S) + \tfrac{1}{3}\beta_{m_i,(1,0,2)}(S) + \tfrac{1}{3}\beta_{m_j,(1,0,2)}(S).$$

Following the discussion in Section 2.3, the resulting spline is C^1-smooth at each vertex of $\triangle$ and also C^1-smooth across each edge of $\triangle$. Moreover, it is by construction C^1-smooth inside each triangle of $\triangle$ and consequently C^1-smooth everywhere on the domain Θ.

We denote by $\mathbb{S}_3 \subset \mathbb{S}_3^1(\triangle_{CT})$ the vector space consisting of all splines that can be obtained by the described construction. It holds that $\dim(\mathbb{S}_3) = 3n_v$ and $\mathbb{P}_3 \subset \mathbb{S}_3$.

3.4. Basis functions

In the following we define basis functions for the spaces $\mathbb{S}_\ell$, $\ell = 1, 2, 3$. For a fixed ℓ, we associate with every vertex $\boldsymbol{v}_i \in V$ three functions $B_{i,r}^\ell : \Theta \to \mathbb{R}$, $r = 0, 1, 2$, that form a basis of $\mathbb{S}_\ell$.
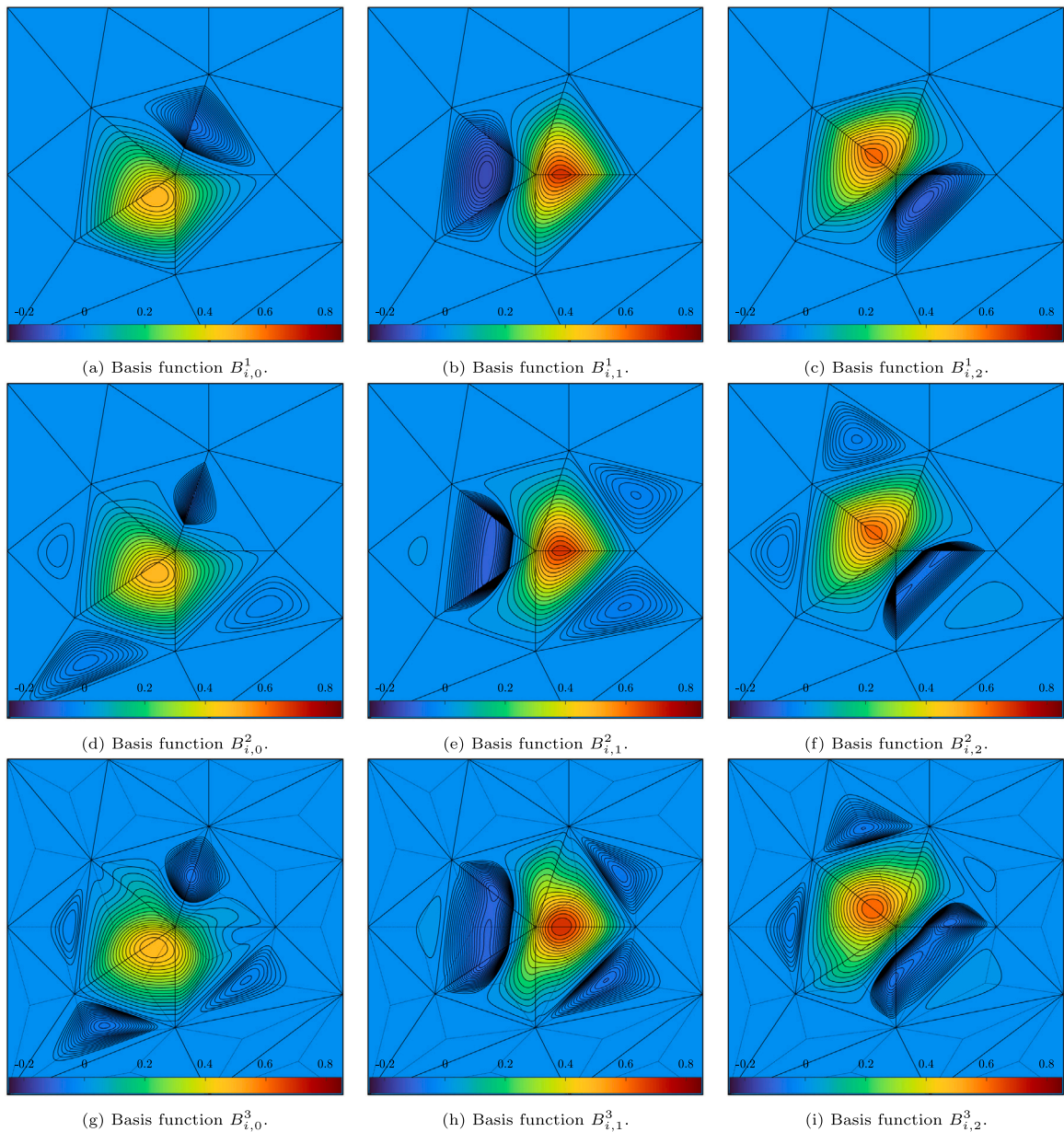
Fig. 2. Contour plots of the basis functions $B_{i,r}^{\ell}$ for $r = 0, 1, 2$ and $\ell = 1, 2, 3$ that are associated with the vertex $\boldsymbol{v}_i$ in the interior of the domain and defined based on the triangle $[\boldsymbol{q}_{i,0}, \boldsymbol{q}_{i,1}, \boldsymbol{q}_{i,2}]$ shown in Fig. 1.

The function $B_{i,r}^{\ell} \in \mathbb{S}_{\ell}$ is uniquely specified by choosing a linear polynomial $Q_{i,r}$ and setting $Q_i = Q_{i,r}$ and $Q_j = 0$, $j \neq i$, in the construction of splines described in Sections 3.1, 3.2, and 3.3. Independently of the choice of $Q_{i,r}$, this definition ensures that the support of $B_{i,r}^{\ell}$ is local. The construction in Section 3.1 implies that the support of $B_{i,r}^1$ is contained in the union of triangles of $\triangle$ that have a vertex at $\boldsymbol{v}_i$. The supports of $B_{i,r}^2$ and $B_{i,r}^3$ are, according to the constructions in Sections 3.2 and 3.3, supplemented by the triangles of $\triangle$ that share an edge with a triangle that is in the support of $B_{i,r}^1$.

Any choice of linearly independent linear polynomials $Q_{i,r}$, $r = 0, 1, 2$, yields a set of linearly independent functions $B_{i,r}^{\ell}$. Endorsing the ideas introduced in [4], we opt for polynomials determined by a barycentric coordinate system. Let $[\boldsymbol{q}_{i,0}, \boldsymbol{q}_{i,1}, \boldsymbol{q}_{i,2}] \subset \mathbb{R}^2$ be a triangle of a small area that contains $\boldsymbol{v}_i$ and the point $\frac{5}{6}\boldsymbol{v}_i + \frac{1}{6}\boldsymbol{v}_j$ for every edge $[\boldsymbol{v}_i, \boldsymbol{v}_j] \in E$. Such choice is motivated by the stability analysis in [10] and numerical tests made in Example 3 (presented in Section 4). Fig. 1 (right) shows an example of such triangles that have minimal area.
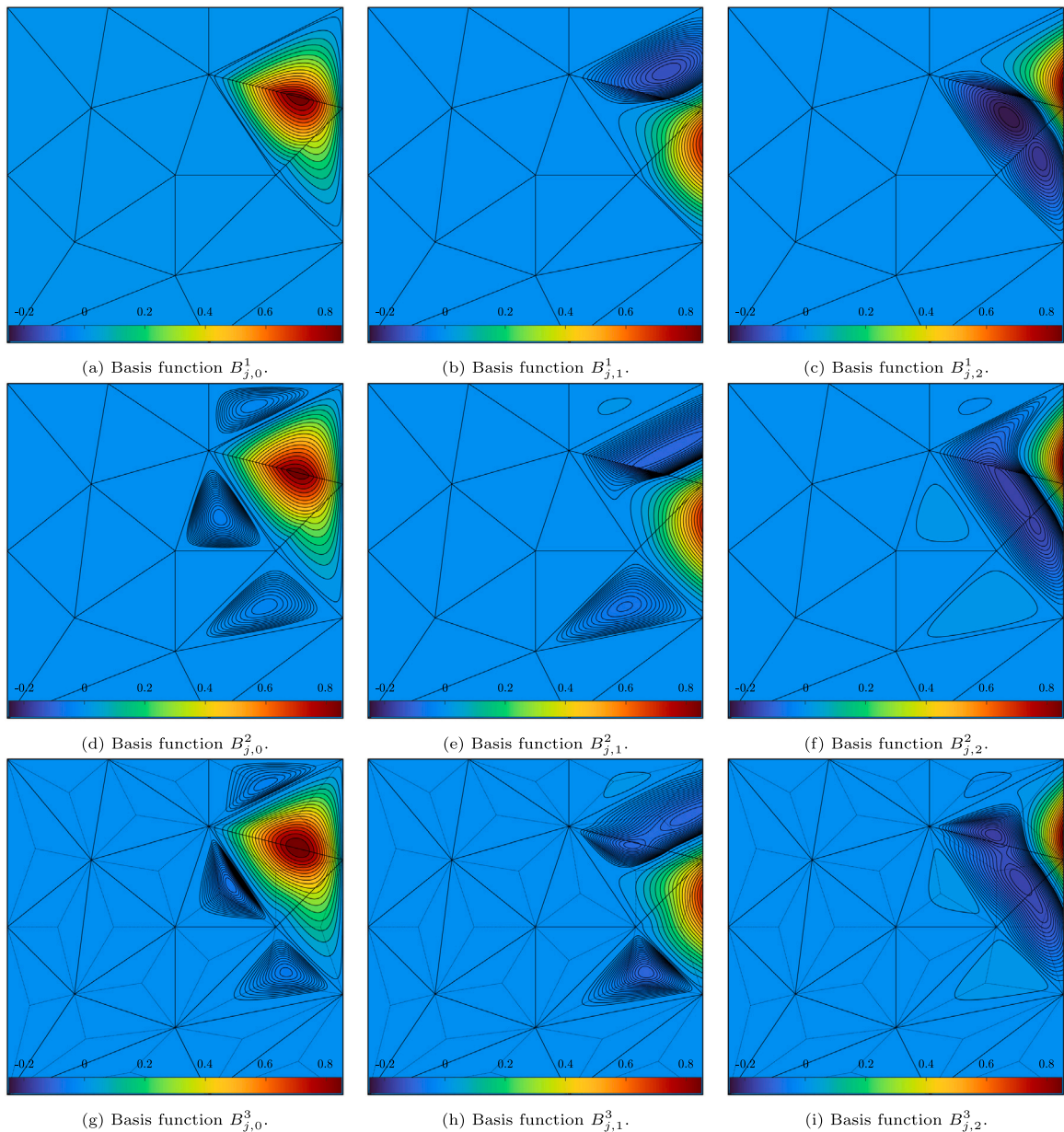
Fig. 3. Contour plots of the basis functions $B_{j,r}^{\ell}$ for $r = 0, 1, 2$ and $\ell = 1, 2, 3$ that are associated with the vertex $\boldsymbol{v}_j$ on the boundary of the domain and defined based on the triangle $[\boldsymbol{q}_{j,0}, \boldsymbol{q}_{j,1}, \boldsymbol{q}_{j,2}]$ shown in Fig. 1. Notice that the functions $B_{j,0}^{\ell}$ are zero on the boundary since the points $\boldsymbol{q}_{j,1}$ and $\boldsymbol{q}_{j,2}$ (marked in Fig. 1) lie on the boundary of the domain.

We define $(Q_{i,0}, Q_{i,1}, Q_{i,2})$ to be the barycentric coordinates with respect to $[\boldsymbol{q}_{i,0}, \boldsymbol{q}_{i,1}, \boldsymbol{q}_{i,2}]$. One obvious property of these linear polynomials is that $Q_{i,r}(\boldsymbol{q}_{i,r}) = 1$ and $Q_{i,r}(\boldsymbol{q}_{i,\rho}) = 0$ for every $\rho \in \{0, 1, 2\} \setminus \{r\}$. Examples of the resulting functions $B_{i,r}^{\ell}$ are shown in Figs. 2 and 3.

Remark 1. In the construction of basis functions for the Powell–Sabin macro-element provided in [4], the triangle $[\boldsymbol{q}_{i,0}, \boldsymbol{q}_{i,1}, \boldsymbol{q}_{i,2}]$ associated with a vertex $\boldsymbol{v}_i \in V$ is constrained to contain a certain set of points in order to ensure the nonnegativity of basis functions. This is not the case here as it is unclear how to achieve nonnegativity of the basis functions for $\mathbb{S}_2$ and $\mathbb{S}_3$. On the other hand, for the considered basis of $\mathbb{S}_1$, the nonnegativity can be easily enforced by requiring that $[\boldsymbol{q}_{i,0}, \boldsymbol{q}_{i,1}, \boldsymbol{q}_{i,2}]$ contains the point $\frac{2}{3}\boldsymbol{v}_i + \frac{1}{3}\boldsymbol{v}_j$ for every edge $[\boldsymbol{v}_i, \boldsymbol{v}_j] \in E$ with a vertex at $\boldsymbol{v}_i$ and the point $\frac{1}{2}\boldsymbol{v}_i + \frac{1}{4}\boldsymbol{v}_j + \frac{1}{4}\boldsymbol{v}_k$ for every triangle $[\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] \in T$ with a vertex at $\boldsymbol{v}_i$. This can be directly observed from (6) and (7) by definition of the basis functions $B_{i,r}^1$, $r = 0, 1, 2$.

3.5. Spline representation

A spline $S \in \mathbb{S}_\ell$, $\ell \in \{1, 2, 3\}$, can be uniquely represented as a linear combination of the basis functions $B_{i,r}^\ell$ for $\boldsymbol{v}_i \in V$ and $r \in \{0, 1, 2\}$. To express the coefficients of this representation, we introduce the linear functionals

$$\beta_{i,r}^\ell(S) := S(\boldsymbol{v}_i) + \nabla S(\boldsymbol{v}_i) \cdot (\boldsymbol{q}_{i,r} - \boldsymbol{v}_i). \tag{8}$$

Let Q_i be the linear polynomial that by constructions in Sections 3.1, 3.2, 3.3 specifies the value and the gradient of S at $\boldsymbol{v}_i$. From the discussion in Section 2.3 we see that $\beta_{i,r}^\ell(S) = Q_i(\boldsymbol{q}_{i,r})$. In particular, by definitions of the basis functions in Section 3.4, we have

$$\beta_{i,r}^\ell(B_{j,\rho}^\ell) = \begin{cases} 1 & \text{if } i = j \text{ and } r = \rho, \\ 0 & \text{otherwise}, \end{cases}$$

and hence

$$S = \sum_{i=1}^{n_v} \sum_{r=0}^{2} \beta_{i,r}^\ell(S) B_{i,r}^\ell. \tag{9}$$

This result has some interesting consequences.

As the constant $(x, y) \mapsto 1$ is an element of $\mathbb{S}_\ell$, (8) and (9) imply

$$1 = \sum_{i=1}^{n_v} \sum_{r=0}^{2} B_{i,r}^\ell(x, y)$$

for every $(x, y) \in \Theta$, i.e., the basis functions form a partition of unity. Moreover, considering the linear polynomials $(x, y) \mapsto x$ and $(x, y) \mapsto y$, it follows by the same line of arguments that

$$(x, y) = \sum_{i=1}^{n_v} \sum_{r=0}^{2} \boldsymbol{q}_{i,r} B_{i,r}^\ell(x, y)$$

for every $(x, y) \in \Theta$, i.e., the points $\boldsymbol{q}_{i,r}$, $r = 0, 1, 2$, that define the triangle associated with $\boldsymbol{v}_i \in V$ are the Greville points of the basis functions $B_{i,r}^\ell$.

4. Numerical comparison

This section provides a number of numerical examples with the aim to demonstrate the behavior of the cubic splines presented in Section 3. We consider the problems of finding the best (discrete) L^2 approximation to a function and solving the Poisson boundary value problem via the isoparametric finite element method and the immersed penalized boundary method.

4.1. Preliminaries

In what follows we deal with methods that require the representation of a spline $S \in \mathbb{S}_\ell$, $\ell \in \{1, 2, 3\}$, in terms of basis functions. For this purpose we use the bases introduced in Section 3.4. To make the notation simpler, we denote the functions $B_{i,r}^\ell$, $i = 1, 2, \ldots, n_v$, $r = 0, 1, 2$, for a fixed ℓ by S_k with the indices k ranging from 1 to $N = 3n_v$. The spline S is expressed in the form

$$S = \sum_{k=1}^{N} c_k S_k \tag{10}$$

for coefficients $c_k \in \mathbb{R}$. Additionally, we assume that the basis functions are enumerated in such a way that $S_1, S_2, \ldots, S_n$ for $n < N$ are zero on the boundary of Θ, and $S_{n+1}, S_{n+2}, \ldots, S_N$ are nonzero on the boundary of Θ. This arrangement is important for the finite element method described in Section 4.3.

Example 1. Let Θ be the unit square partitioned by the triangulation $\triangle$ shown in Fig. 1. The triangulation has 16 vertices, hence $N = 48$. The basis functions are determined by the shaded triangles, one for each vertex. For each vertex on the boundary of Θ that is shared by two collinear edges, i.e., a boundary vertex that is not a corner of the square, the triangle is positioned in a way that two of its vertices lie on the boundary. This means that of the three basis functions associated with the vertex only two are nonzero on the boundary (see Fig. 3), which implies that $n = 24$. An important aspect of such choice is that it in general ensures that the restrictions of $S_{n+1}, S_{n+2}, \ldots, S_N$ to $\partial\Theta$ are linearly independent (see [6] for a more detailed discussion).

Example 2. Fig. 4 shows three subsequent refinements of the triangulation presented in Example 1 and together with it forms a sequence of four triangulations. In each step we refine the triangulation by subdividing each triangle into four smaller triangles determined by the vertices and midpoints of the edges. We also adjust and supplement the set of triangles associated with the vertices so that the basis functions on the refined triangulations satisfy the properties discussed in Example 1.
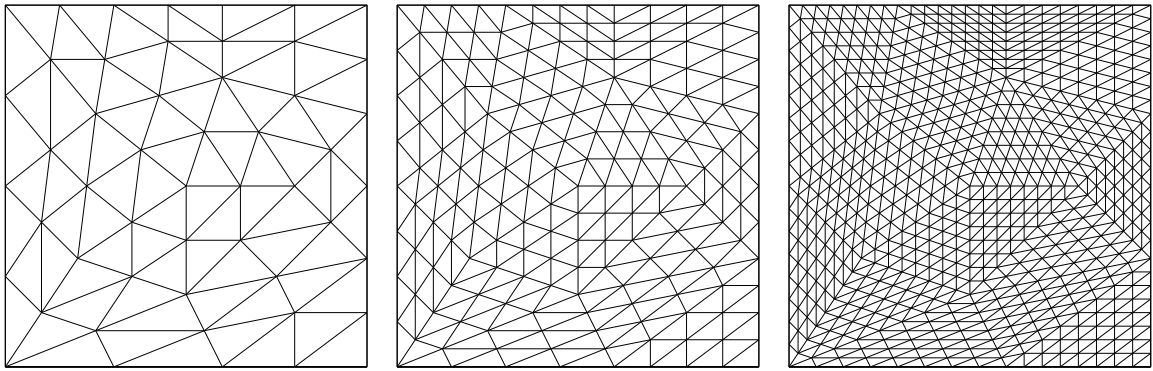
Fig. 4. A sequence of triangulations obtained by subsequently refining the triangulation of the unit square shown in Fig. 1.
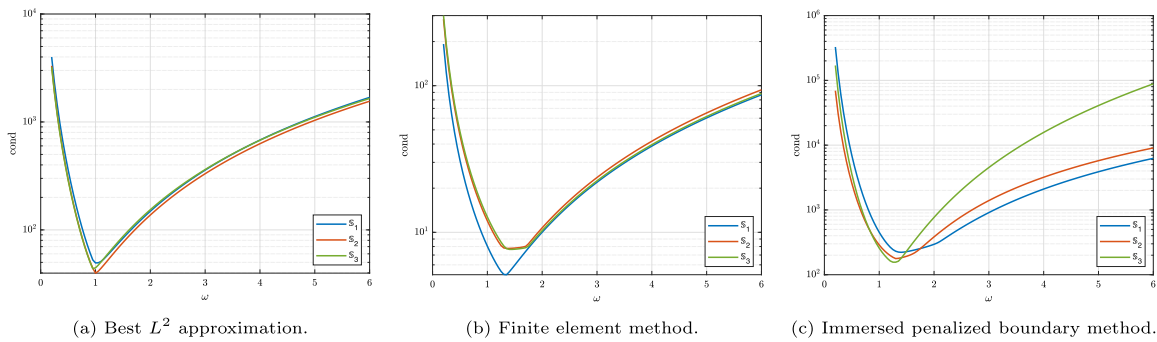

(a) Best L^2 approximation. (b) Finite element method. (c) Immersed penalized boundary method.

Fig. 5. The condition numbers of matrices arising in different approximation methods with respect to the scaling factor ω that determines the size of triangles used in the construction of the basis functions for $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, as described in Example 3.

In the process of refinement described in Example 2, the length h of the longest edge in the triangulation is halved, which allows us to numerically inspect the convergence rates of the methods. With respect to the polynomial precision properties, it is expected that the methods based on $\mathbb{S}_1$ converge with rate at most 3, and the methods based on $\mathbb{S}_2$ and $\mathbb{S}_3$ converge with rate at most 4, i.e., the approximation error in the best case behaves as $\mathcal{O}(h^3)$ and $\mathcal{O}(h^4)$, respectively. In the numerical tests conducted over the triangulations from Examples 1 and 2, the approximation error is measured as the maximal absolute difference between exact solution and approximation computed on the 401×401 uniform grid of points in the unit square.

The methods described in Sections 4.2 and 4.3 reduce to solving a system of linear equations $\boldsymbol{A}\,\boldsymbol{c} = \boldsymbol{b}$. The matrix $\boldsymbol{A}$ is of size either $N \times N$ or $n \times n$, and the vectors $\boldsymbol{c}$ and $\boldsymbol{b}$ are of length either N or n. In this respect the elements of $\boldsymbol{c}$ are either all or the first n coefficients c_k in (10) collected increasingly by k. The elements of $\boldsymbol{A}$ are denoted by either $a_{j,k}^N$ or $a_{j,k}^n$, and the elements of $\boldsymbol{b}$ are denoted by either b_j^N or b_j^n with the indices j and k ranging from 1 to N or n, respectively. We measure stability of the spline representation by the 2-norm condition number (cond) of the matrix $\boldsymbol{A}$.

Example 3. For the triangulation discussed in Example 1 we inspect the condition numbers of the matrices $\boldsymbol{A}$ arising in the best L^2 approximation (Example 4), finite element method (Example 8), and immersed penalized boundary method (Example 10) with respect to the size of the triangles that define the basis functions. Suppose $[\boldsymbol{q}_{i,0}, \boldsymbol{q}_{i,1}, \boldsymbol{q}_{i,2}]$ are the triangles associated with the vertices $\boldsymbol{v}_i \in V$ depicted in Fig. 1 (right), and let

$$\boldsymbol{q}_{i,r}^\omega = \boldsymbol{v}_i + \omega(\boldsymbol{q}_{i,r} - \boldsymbol{v}_i), \quad r = 0, 1, 2,$$

for a factor $\omega > 0$. Fig. 5 depicts the condition numbers in dependence of $\omega \in [0.2, 6]$ for the considered basis of $\mathbb{S}_\ell$, $\ell \in \{1, 2, 3\}$, determined by the triangles $[\boldsymbol{q}_{i,0}^\omega, \boldsymbol{q}_{i,1}^\omega, \boldsymbol{q}_{i,2}^\omega]$. The results indicate that $\omega = 1$ ensures near-optimal stability.

The expressions for the elements of $\boldsymbol{A}$ and $\boldsymbol{b}$ typically include integrals over a domain. If S is a spline over a triangulation $\triangle$ of the domain Θ that is defined piecewise on each triangle t_m, $m \in \{1, 2, \ldots, n_t\}$, of $\triangle$ by a polynomial P_m, the integral of S over the domain is always considered as the sum of integrals of P_m over the triangles of $\triangle$. More precisely,

$$\int_\Theta S \, \mathrm{d}\Theta = \sum_{m=1}^{n_t} \left(\int_{t_m} P_m \, \mathrm{d}t_m \right).$$

This way the integral exists even if S is not well defined on the edges of $\triangle$.

Table 1
Approximation errors and condition numbers for the problem described in Example 4.

N	$\mathbb{S}_1$		$\mathbb{S}_2$		$\mathbb{S}_3$	
	Error	cond	Error	cond	Error	cond
48	1.88e−01	4.94e+01	1.97e−01	3.97e+01	2.07e−01	4.45e+01
153	1.86e−02	4.58e+01	9.62e−03	4.12e+01	9.20e−03	4.25e+01
543	2.33e−03	4.69e+01	8.49e−04	4.66e+01	8.15e−04	5.08e+01
2043	3.89e−04	4.95e+01	8.99e−05	5.04e+01	8.39e−05	5.71e+01
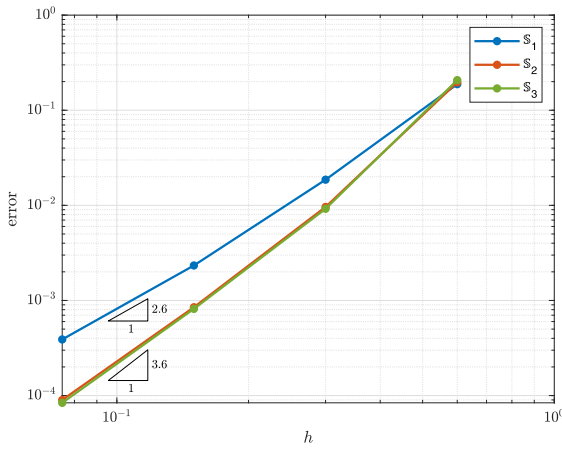
(a) Errors for Example 4.
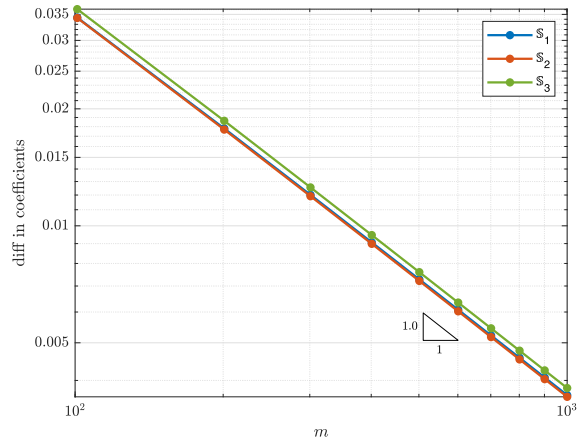
(b) Convergence discussed in Example 5.

Fig. 6. Errors of the approximations from $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, for the problem discussed in Example 4 (left) and the convergence of the coefficient vectors corresponding to the discrete best L^2 approximation when the number of points is increased as described in Example 5 (right).

In the following we briefly discuss some common methods for function approximation and solving boundary value problems and provide several numerical examples. We comment the results in Section 5.

4.2. Function approximations

Let Θ be a polygonal domain and $\triangle$ its triangulation. The best L^2 approximation $S \in \mathbb{S}_\ell$, $\ell \in \{1, 2, 3\}$, of a given function $F \in L^2(\Theta)$, i.e., the solution to the minimization problem

$$\arg\min_{S \in \mathbb{S}_\ell} \int_\Theta (F - S)^2 \, d\Theta,$$

is determined by the conditions

$$\int_\Theta (F - S) \, S_j \, d\Theta = 0, \quad j = 1, 2, \dots, N. \tag{11}$$

If we represent S in the form (10), the computation of the unknown coefficients c_k amounts to solving the system $\boldsymbol{A} \boldsymbol{c} = \boldsymbol{b}$ determined by

$$a_{j,k}^N = \int_\Theta S_k \, S_j \, d\Theta, \quad b_j^N = \int_\Theta F \, S_j \, d\Theta.$$

The matrix $\boldsymbol{A}$ is the Gram matrix associated with the functions $S_1, S_2, \dots, S_N$.

Example 4. We compute the best L^2 approximation in $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, to the Franke's test function

$$F(x, y) = 0.75 \exp\left(-\tfrac{1}{4}(9x - 2)^2 - \tfrac{1}{4}(9y - 2)^2\right) + 0.75 \exp\left(-\tfrac{1}{49}(9x + 1)^2 - \tfrac{1}{10}(9y + 1)\right)$$
$$+ 0.5 \exp\left(-\tfrac{1}{4}(9x - 7)^2 - \tfrac{1}{4}(9y - 3)^2\right) - 0.2 \exp\left(-(9x - 4)^2 - (9y - 7)^2\right) \tag{12}$$

on the unit square partitioned by the sequence of triangulations discussed in Example 2. The results are provided in Table 1. Fig. 6 (left) depicts the errors with respect to the length of the longest edge of the underlying triangulation. Fig. 7 shows the contour plots of the approximations over the initial triangulation.
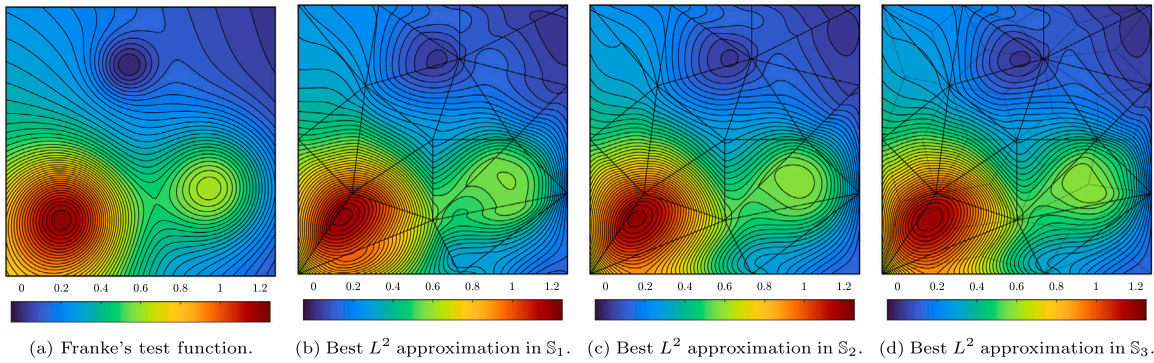
(a) Franke's test function. (b) Best L^2 approximation in $\mathbb{S}_1$. (c) Best L^2 approximation in $\mathbb{S}_2$. (d) Best L^2 approximation in $\mathbb{S}_3$.

Fig. 7. Contour plots of the Franke's test function and its best L^2 approximations in $\mathbb{S}_\ell$, $\ell = 1, 2, 3$.

Table 2
Approximation errors and condition numbers for the problem described in Example 5.

N	$\mathbb{S}_1$		$\mathbb{S}_2$		$\mathbb{S}_3$	
	Error	cond	Error	cond	Error	cond
48	1.76e−01	4.63e+01	1.85e−01	3.69e+01	1.95e−01	4.15e+01
153	1.87e−02	3.91e+01	9.59e−03	3.41e+01	9.20e−03	3.63e+01
543	2.33e−03	3.87e+01	8.48e−04	3.45e+01	8.18e−04	3.81e+01
2043	3.90e−04	3.90e+01	9.00e−05	3.55e+01	8.38e−05	4.21e+01

In the discrete version of the best L^2 approximation we use given values $F(\boldsymbol{p}_d)$, $d = 1, 2, \ldots, M$, at $M \geq N$ points $\boldsymbol{p}_d \in \Theta$ and search for

$$\arg\min_{S \in \mathbb{S}_\ell} \sum_{d=1}^{M} \left(F(\boldsymbol{p}_d) - S(\boldsymbol{p}_d) \right)^2.$$

To this end, we replace the conditions in (11) by

$$\sum_{d=1}^{M} \left(F(\boldsymbol{p}_d) - S(\boldsymbol{p}_d) \right) S_j(\boldsymbol{p}_d) = 0, \quad j = 1, 2, \ldots, N.$$

This yields the system $\boldsymbol{A}\,\boldsymbol{c} = \boldsymbol{b}$ determined by

$$a_{j,k}^N = \sum_{d=1}^{M} S_k(\boldsymbol{p}_d)\, S_j(\boldsymbol{p}_d), \quad b_j^N = \sum_{d=1}^{M} F(\boldsymbol{p}_d)\, S_j(\boldsymbol{p}_d).$$

Alternatively, the vector of coefficients $\boldsymbol{c}$ can be expressed as the best least squares solution to the overdetermined system $\boldsymbol{S}\,\boldsymbol{c} = \boldsymbol{f}$, where $\boldsymbol{S} \in \mathbb{R}^{M \times N}$ is the collocation matrix determined by element $S_k(\boldsymbol{p}_d)$ at position (d, k), and $\boldsymbol{f} \in \mathbb{R}^M$ is the vector with element $F(\boldsymbol{p}_d)$ at position d.

Example 5. We repeat Example 4 for the discrete best L^2 approximation based on the 201×201 uniform grid of points in the unit square. Table 2 reveals that results are similar to the ones observed in Example 4. For the triangulation discussed in Example 1, Fig. 6 (right) shows linear decay in 2-norm of the difference between the coefficient vector corresponding to the best L^2 approximation and the coefficient vector corresponding to the discrete best L^2 approximation on the $m \times m$ uniform grid of points ($M = m^2$) for $m \in \{101, 201, \ldots, 1001\}$.

Since scattered function data to be approximated is often subjected to noise, it is common to use a penalty term in the minimization of the discrete L^2 norm to prevent high oscillation of the approximant. Following [12], we can search for

$$\arg\min_{S \in \mathbb{S}_\ell} \left(\sum_{d=1}^{M} \left(F(\boldsymbol{p}_d) - S(\boldsymbol{p}_d) \right)^2 + \lambda \int_{\Theta} \boldsymbol{e} \cdot (\boldsymbol{H}_S \odot \boldsymbol{H}_S) \cdot \boldsymbol{e} \, d\Theta \right),$$

where $\lambda \geq 0$ is the penalty parameter, $\boldsymbol{e} = (1, 1)$, $\boldsymbol{H}_S$ is the Hessian matrix of S, and $\odot$ denotes the Hadamard product. The solution to the penalized discrete best L^2 approximation problem can be computed by solving the system $\boldsymbol{A}\,\boldsymbol{c} = \boldsymbol{b}$ for

$$\boldsymbol{A} = \boldsymbol{S}^T \boldsymbol{S} + \lambda \boldsymbol{H}, \quad \boldsymbol{b} = \boldsymbol{S}^T \boldsymbol{f},$$

where $\boldsymbol{H} \in \mathbb{R}^{N \times N}$ is determined by the element $\int_{\Theta} \boldsymbol{e} \cdot (\boldsymbol{H}_{S_k} \odot \boldsymbol{H}_{S_j}) \cdot \boldsymbol{e} \, d\Theta$ at position (j, k).
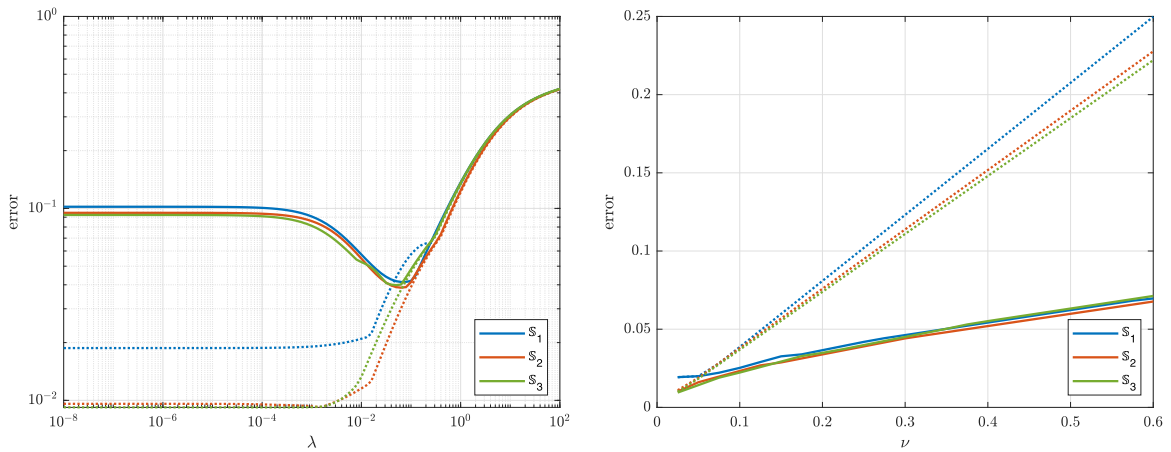
Fig. 8. The approximation error with respect to the penalty parameter λ and noise magnitude ν for the problem described in Example 6.

Example 6. As in Example 5 we approximate the function (12), but this time we corrupt the function values computed on the 201×201 uniform grid of points by adding random noise drawn from the standard uniform distribution on the interval $(-\nu, \nu)$ for $\nu > 0$. We consider approximations obtained by the penalized method over the triangulation shown in Fig. 4 (left). Fig. 8 (left) shows the approximation errors in dependence of $\lambda \in [10^{-8}, 10^2]$ with ν fixed to 0.25. The dotted lines correspond to the approximations obtained for uncorrupted data ($\nu = 0$). On the other hand, Fig. 8 (right) shows the approximation errors in dependence of $\nu \in [0.025, 0.6]$ where for each ν and each spline space the value of λ is computed in advance to be near-optimal. The dotted lines correspond to the approximations obtained without penalization ($\lambda = 0$). Note that on the unit square the function (12) attains values between 0 and 1.22 and thus the values of ν in this test represent magnitudes of noise approximately from 2 to 50 percent of the function range.

4.3. Approximations to boundary value problems

Let $\Omega \subset \mathbb{R}^2$ be a connected bounded domain that can be parametrized by a C^1-smooth mapping $P \in \mathbb{S}_\ell \times \mathbb{S}_\ell$ defined on a polygonal domain $\Theta \subset \mathbb{R}^2$ triangulated by $\triangle$. We assume P is a bijection with a nonvanishing Jacobian on Θ that maps the boundary $\partial\Theta$ of Θ to the boundary $\partial\Omega$ of Ω.

Example 7. Fig. 9 (left) shows a domain Ω that can be exactly parametrized on the unit square Θ by a geometry mapping $P \in \mathbb{P}_2 \times \mathbb{P}_2$. Both components of this mapping (shown in the middle and on the right of Fig. 9) are contained in $\mathbb{S}_\ell$, $\ell \in \{1, 2, 3\}$, and can be represented by the basis functions $S_1, S_2, \ldots, S_N$ with the help of (8) and (9).

For given functions $F \in L^2(\Omega)$ and $G \in L^2(\partial\Omega)$ we consider the boundary value problem governed by the Poisson equation and the Dirichlet boundary conditions. The goal is to find the function U satisfying

$$\begin{cases} -\nabla^2 U(\boldsymbol{p}) = F(\boldsymbol{p}) & \text{if } \boldsymbol{p} \in \Omega \setminus \partial\Omega, \\ U(\boldsymbol{p}) = G(\boldsymbol{p}) & \text{if } \boldsymbol{p} \in \partial\Omega. \end{cases} \tag{13}$$

In the isoparametric finite element method based on $\mathbb{S}_\ell$ we search for an approximation to U of the form $S \circ P^{-1}$, $S \in \mathbb{S}_\ell$. After the Galerkin discretization in $\mathrm{span}(S_1 \circ P^{-1}, S_2 \circ P^{-1}, \ldots, S_n \circ P^{-1})$ this amounts to finding S that admits

$$\int_\Omega \nabla(S \circ P^{-1}) \cdot \nabla(S_j \circ P^{-1}) \, \mathrm{d}\Omega = \int_\Omega F(S_j \circ P^{-1}) \, \mathrm{d}\Omega, \quad j = 1, 2, \ldots, n,$$

or, equivalently,

$$\int_\Theta \nabla S \cdot \boldsymbol{J}_P^{-1} \boldsymbol{J}_P^{-T} \cdot \nabla S_j \left| \det(\boldsymbol{J}_P) \right| \, \mathrm{d}\Theta = \int_\Theta (F \circ P) \, S_j \left| \det(\boldsymbol{J}_P) \right| \, \mathrm{d}\Theta, \quad j = 1, 2, \ldots, n, \tag{14}$$

where $\boldsymbol{J}_P$ is the Jacobian matrix of P.

We modify the expression (10) of S to

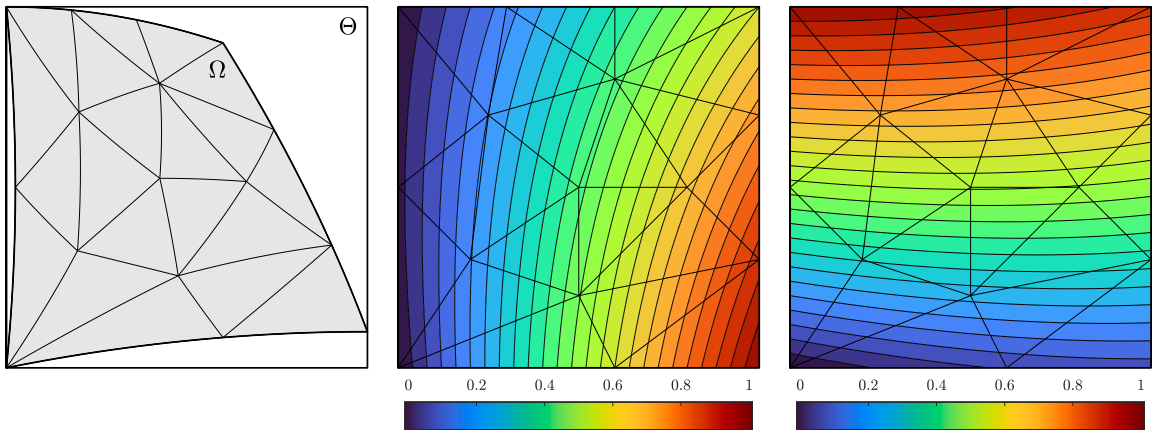$$S = S_0 + \sum_{k=1}^n c_k S_k, \quad S_0 = \sum_{k=n+1}^N c_k S_k.$$

Fig. 9. A domain with curved boundary Ω (left) discussed in Example 7, which is contained in the unit square Θ. The domain Ω can be parametrized by a mapping whose components are quadratic polynomials (the contour plots of the first and the second component are shown in the middle and on the right, respectively). The left figure also depicts the partition of the domain that is obtained by using this mapping to map the triangulation from Example 1.

Table 3
Approximation errors and condition numbers for the problem described in Example 8.

n	$\mathbb{S}_1$		$\mathbb{S}_2$		$\mathbb{S}_3$	
	Error	cond	Error	cond	Error	cond
24	1.25e−01	8.04e+00	1.20e−01	1.20e+01	1.12e−01	1.27e+01
109	3.82e−02	3.82e+01	2.00e−02	4.35e+01	1.89e−02	4.55e+01
459	5.69e−03	1.55e+02	1.28e−03	1.71e+02	1.35e−03	1.79e+02
1879	7.85e−04	6.20e+02	1.06e−04	6.80e+02	1.09e−04	7.11e+02

Table 4
Approximation errors and condition numbers for the problem described in Example 9.

n	$\mathbb{S}_1$		$\mathbb{S}_2$		$\mathbb{S}_3$	
	Error	cond	Error	cond	Error	cond
24	1.91e−01	8.63e+00	9.36e−02	9.85e+00	9.45e−02	1.05e+01
109	3.49e−02	3.36e+01	3.29e−02	3.54e+01	3.07e−02	3.74e+01
459	5.92e−03	1.44e+02	2.24e−03	1.40e+02	2.22e−03	1.48e+02
1879	9.83e−04	5.98e+02	1.99e−04	5.86e+02	2.07e−04	6.10e+02

First, we compute the coefficients c_k, $k = n+1, n+2, \ldots, N$, so that S_0 is the discrete best L^2 approximation to $G \circ P$ for a dense set of points on the boundary of Θ. Then we transform (14) into the system $\boldsymbol{A}c = \boldsymbol{b}$ determined by

$$a_{j,k}^n = \int_\Theta \nabla S_k \cdot \boldsymbol{J}_P^{-1} \boldsymbol{J}_P^{-T} \cdot \nabla S_j \left| \det(\boldsymbol{J}_P) \right| \, \mathrm{d}\Theta, \quad b_j^n = \int_\Theta \left((F \circ P) \, S_j - \nabla S_0 \cdot \boldsymbol{J}_P^{-1} \boldsymbol{J}_P^{-T} \cdot \nabla S_j \right) \left| \det(\boldsymbol{J}_P) \right| \, \mathrm{d}\Theta.$$

The matrix $\boldsymbol{A}$ is the isoparametric stiffness matrix associated with the functions $S_1, S_2, \ldots, S_n$.

Example 8. Let $P(x, y) = (x, y)$ and $\Omega = \Theta$ the unit square triangulated by the sequence of triangulations discussed in Example 2. Applying the finite element method, we approximate the solution to the problem (13) with the manufactured solution

$$U(x, y) = \sin\left(2\pi(1 - x)(1 - y)\right). \tag{15}$$

The results are provided in Table 3 and in Fig. 10 (left).

Example 9. We repeat Example 8 for the mapping P and domains Ω and Θ as defined in Example 7. The results are provided in Table 4 and Fig. 10 (right). Here, the approximation errors are computed based on the 401×401 uniform grid of points mapped by P. Fig. 11 shows the contour plots of the approximations obtained for the initial triangulation.

In [13] an interesting alternative to the finite element method is proposed. As in the finite element method, the problem (13) is observed in the weak formulation but by considering all basis functions as test functions. The boundary conditions are imposed approximately with a penalty parameter, hence the name immersed penalty boundary method.

Instead of a parametric mapping between Θ and Ω, we choose a polygonal domain Θ in such a way that it contains Ω. Also, we choose $M \in \mathbb{N}$ points $\boldsymbol{p}_d$, $d = 1, 2, \ldots, M$, on the boundary of Ω. Moreover, we opt for a penalty parameter $\lambda > 0$ to ensure that S
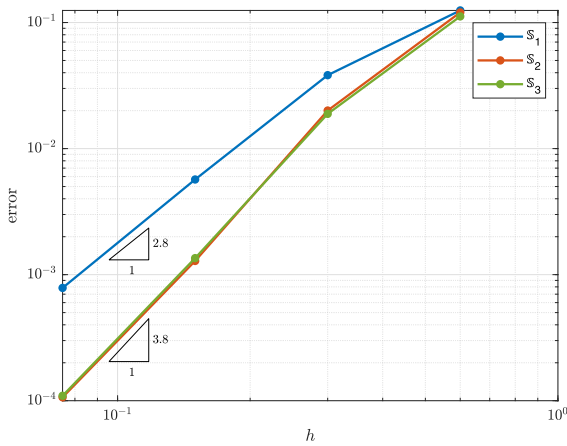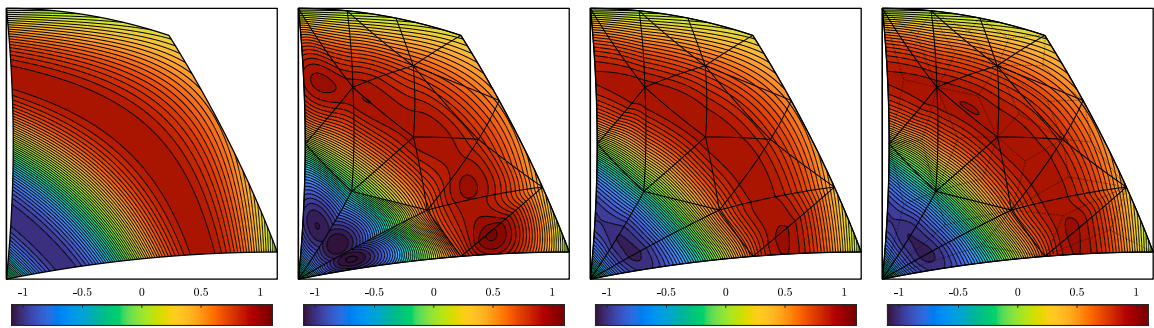
(a) Errors for Example 8.
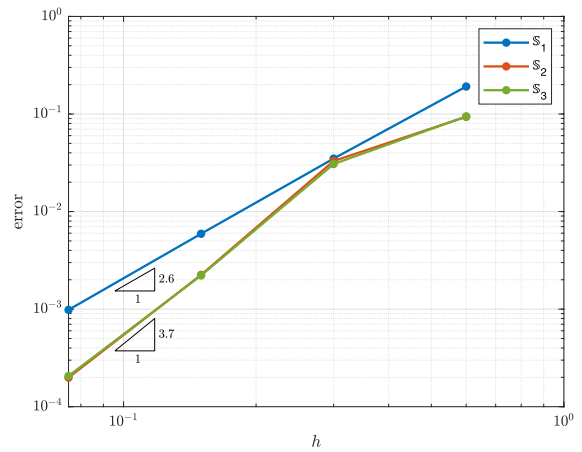
(b) Errors for Example 9.

Fig. 10. Errors of approximations from $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, for the problems discussed in Examples 8 and 9 with respect to the length h of the longest edge in the triangulation.

(a) Function to approximate. (b) Approximation from $\mathbb{S}_1$. (c) Approximation from $\mathbb{S}_2$. (d) Approximation from $\mathbb{S}_3$.

Fig. 11. Contour plots of the function (15) and its approximations from $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, which are obtained by the isoparametric finite element method as discussed in Example 9.

approximately satisfies the boundary conditions and a penalty parameter $\mu > 0$ to ensure that S is approximately C^1-smooth. Then we search for an approximation $S \in \mathbb{S}_\ell$ that solves

$$\underset{S \in \mathbb{S}_\ell}{\arg\min} \left(\sum_{j=1}^{N} \left(\int_\Theta \left(F + \nabla^2 S \right) S_j \, d\Theta \right)^2 + \lambda \sum_{d=1}^{M} \left(G(\boldsymbol{p}_d) - S(\boldsymbol{p}_d) \right)^2 + \mu \sum_{l=1}^{n_e - n_e^b} \xi_l(S)^2 \right),$$

where ξ_l is as defined in (3) and we impose (for notational convenience) that the edges in $E \setminus E^b$ are enumerated by the indices from 1 to $n_e - n_e^b$. As $S \in \mathbb{S}_3$ is C^1-smooth, the third sum in the objective function should be ignored for $\ell = 3$. The solution to the minimization problem expressed in the form (10) can be obtained by solving the system $\boldsymbol{A}\,\boldsymbol{c} = \boldsymbol{b}$ for

$$\boldsymbol{A} = \boldsymbol{L}^T \boldsymbol{L} + \lambda\, \boldsymbol{S}^T \boldsymbol{S} + \mu \boldsymbol{E}^T \boldsymbol{E}, \quad \boldsymbol{b} = -\boldsymbol{L}^T \boldsymbol{f} + \lambda\, \boldsymbol{S}^T \boldsymbol{g},$$

where

- $\boldsymbol{L} \in \mathbb{R}^{N \times N}$ is determined by the element $\int_\Theta (\nabla^2 S_k)\, S_j \, d\Theta$ at position (j, k),
- $\boldsymbol{S} \in \mathbb{R}^{M \times N}$ is determined by the element $S_k(\boldsymbol{p}_d)$ at position (d, k),
- $\boldsymbol{E} \in \mathbb{R}^{(n_e - n_e^b) \times N}$ is determined by the element $\xi_l(S_k)$ at position (l, k),
- $\boldsymbol{f} \in \mathbb{R}^N$ is determined by the element $\int_\Theta F\, S_j \, d\Theta$ at position j, and
- $\boldsymbol{g} \in \mathbb{R}^M$ is determined by the element $G(\boldsymbol{p}_l)$ at position l.

Example 10. We consider the problem from Example 8 and solve it by using the immersed penalized boundary method on the unit square. We opt for $\lambda = 1$, $\mu = 1$, and $M = 800$ with the points $\boldsymbol{p}_d$ spaced uniformly along the boundary of the domain. The results are provided in Table 5 and Fig. 12 (left).

Table 5
Approximation errors and condition numbers for the problem described in Example 10.

N	$\mathbb{S}_1$		$\mathbb{S}_2$		$\mathbb{S}_3$	
	Error	cond	Error	cond	Error	cond
48	3.00e−01	4.52e+02	1.56e−01	2.88e+02	1.29e−01	2.57e+02
153	8.09e−02	4.44e+03	2.41e−02	4.33e+03	1.92e−02	3.51e+03
543	1.21e−02	5.67e+04	3.75e−03	6.36e+04	1.52e−03	5.26e+04
2043	2.80e−03	8.15e+05	8.53e−04	9.57e+05	1.12e−04	8.23e+05
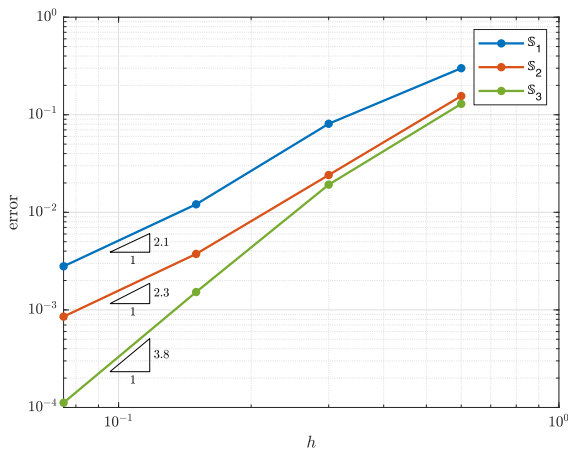
(a) Errors for Example 10.
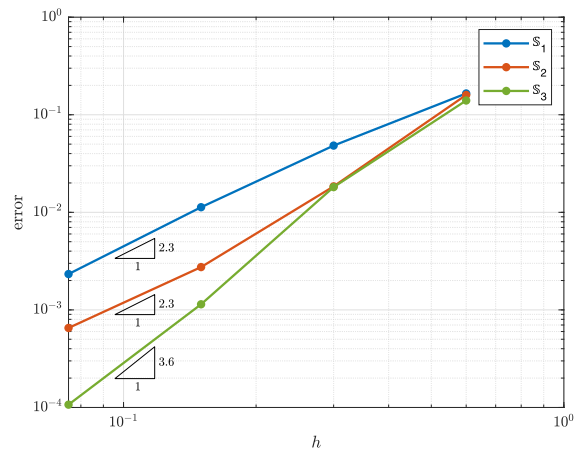
(b) Errors for Example 11.

Fig. 12. Errors of approximations from $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, for the problems discussed in Examples 10 and 11 with respect to the length h of the longest edge in the triangulation.

Table 6
Approximation errors and condition numbers for the problem described in Example 11.

N	$\mathbb{S}_1$		$\mathbb{S}_2$		$\mathbb{S}_3$	
	Error	cond	Error	cond	Error	cond
48	3.00e−01	4.52e+02	1.56e−01	2.88e+02	1.29e−01	2.57e+02
153	8.09e−02	4.44e+03	2.41e−02	4.33e+03	1.92e−02	3.51e+03
543	1.21e−02	5.67e+04	3.75e−03	6.36e+04	1.52e−03	5.26e+04
2043	2.80e−03	8.15e+05	8.53e−04	9.57e+05	1.12e−04	8.23e+05

Example 11. This is a counterpart to Example 9. The domain Ω presented in Example 7 is contained in the unit square Θ. To apply the immersed penalized boundary method we use the same parameters λ, μ, and M as in Example 10 but map the uniformly spaced points from the boundary of the unit square to the boundary of Ω by P to obtain $p_d \in \partial \Omega$. The results are provided in Table 6 and Fig. 12 (right). The approximation errors are computed in the same way as in Example 9. Fig. 13 shows contour plots of the approximations over the initial triangulation.

Example 12. Finally, we report some more details on the choice of the penalization parameters λ and μ in Example 11. We focus on the approximations over the triangulation shown in Fig. 4 (left). Fig. 14 (left) justifies the choice $\lambda = 1$. For $\mu = 1$ it shows the relation between the approximation error on the entire domain and on the boundary when the parameter λ, which controls the importance of the boundary condition, takes the values between 10^{-4} and 10^4. Fig. 14 (right) supports the choice $\mu = 1$. For $\lambda = 1$ it shows the relation between the approximation error and the C^1-smoothness error when the parameter μ, which controls the importance of smoothness, takes the values between 10^{-4} and 10^4. The C^1-smoothness error of an approximation S is measured as the maximal absolute value of $\xi_l(S)$, $l = 1, 2, \ldots, n_e - n_e^b$.

5. Conclusion

In the paper, three cubic macro-elements were considered that have the same characterization, i.e., they are determined by assigning three degrees of freedom to each vertex of the triangulation. However, the macro-elements differ in several aspects.

The first one is the classical Zienkiewicz element, which is on a triangle of the triangulation constructed by considering only the data corresponding to the three vertices of the triangle. It does not reproduce cubic polynomials and is only C^0-smooth, but
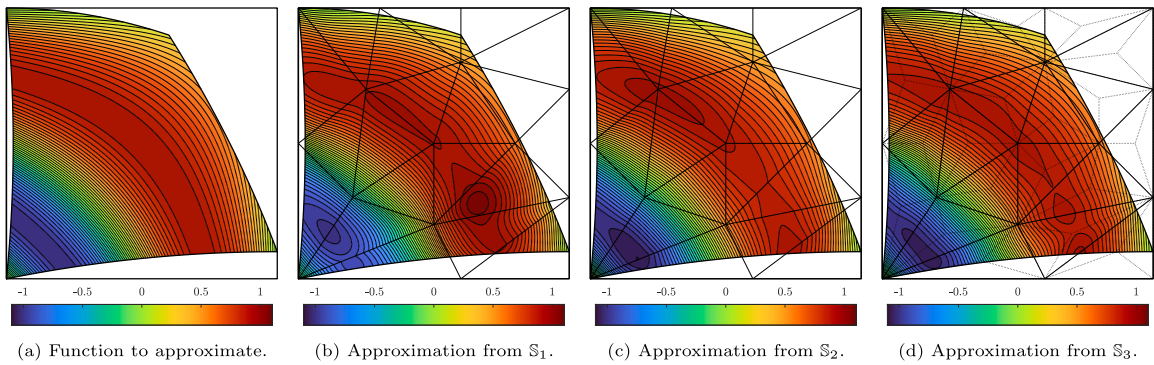
(a) Function to approximate.   (b) Approximation from $\mathbb{S}_1$.   (c) Approximation from $\mathbb{S}_2$.   (d) Approximation from $\mathbb{S}_3$.

Fig. 13. Contour plots of the function (15) and its approximations from $\mathbb{S}_\ell$, $\ell = 1, 2, 3$, which are obtained by the immersed penalized boundary method as discussed in Example 11.
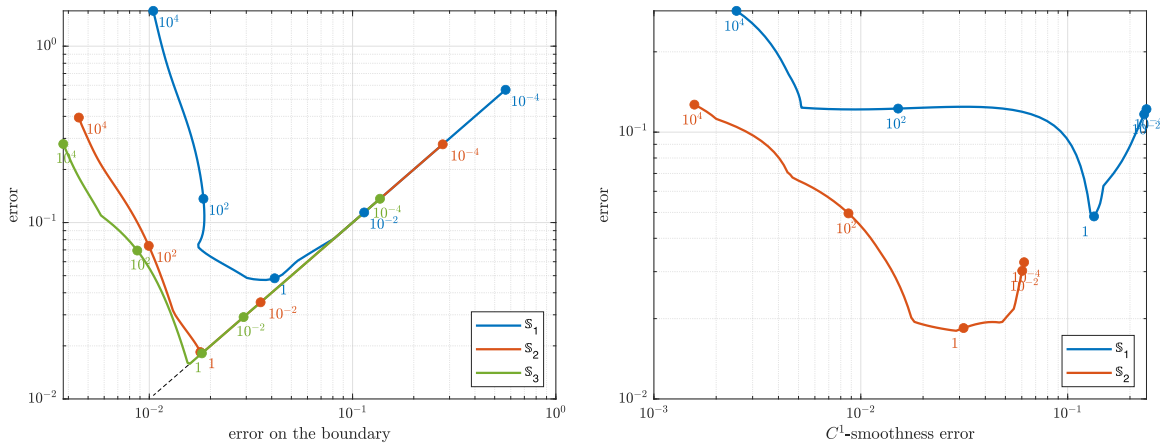

Fig. 14. Impact of the penalization parameters λ and μ on the accuracy of the approximations discussed in Example 12.

it is simple to compute. The second one has cubic precision, which is achieved by taking advantage of the data corresponding to neighboring triangles. This considerably complicates the implementation but does not significantly increase the computational cost. However, the macro-element is still only C^0-smooth. The third approach guarantees both cubic precision and C^1-smoothness, thanks to using the underlying Clough–Tocher refinement of the triangulation. The disadvantage is that the refinement triples the number of triangles to work with and correspondingly slows down the computations.

It was shown in the paper that despite different constructions, the considered macro-elements can be represented in a similar way by a basis consisting of three basis functions per vertex. The basis was utilized to numerically inspect the macro-elements in a variety of approximation methods. Tests indicate that in terms of stability the proposed representations behave very much alike. However, there is, as expected, no definite answer which of the three macro-elements performs best in terms of accuracy and computational cost, and the choice must be made with respect to the problem at hand. We finish with some remarks drawing the conclusions from the numerical experiments presented in Section 4.

Remark 2. The lack of cubic precision implies nonoptimal convergence speed, which is a serious shortcoming when high-accuracy approximation is needed. This is clearly indicated in Examples 4, 5, 8, and 9.

Remark 3. In the approximation of noisy data reported in Example 6 all three macro-elements perform comparably. Rather surprisingly, neither the cubic precision nor the C^1-smoothness provide noteworthy advantage.

Remark 4. For the considered problems of the best L^2 approximation (Examples 4 and 5) and solving the Poisson boundary problem via the finite element method (Examples 8 and 9), the two macro-elements with cubic precision provide practically identical accuracy. Nonetheless, the appearance of the approximations is considerably different as the C^0-smooth macro-element is prone to sharp transitions across the edges of the triangulation. This suggests the following. When only accuracy is important, it might be more efficient to use the C^0-smooth macro-element. If visual appearance matters, the results indicate that the C^1-smooth macro-element does not impair the quality of approximation.

Remark 5. The lack of C^1-smoothness in the immersed penalized boundary method (Examples 10 and 11) clearly degrades the convergence rate of the approximations. The C^0-smooth approximations are less accurate than those obtained by the finite element method (Examples 8 and 9). The use of a penalty parameter to sanction the C^1-smoothness violation improves the visual appearance of the approximations (see Fig. 13) but does not match the accuracy of the C^1-smooth approximations. In fact, the C^1-smooth macro-element performs remarkably well and provides as accurate results as in the finite element method.

Acknowledgments

References

[1] G.P. Bazaley, Y.K. Cheung, B.M. Irons, O.C. Zienkiewicz, Triangular elements in plate bending - conforming and non-conforming solutions, in: Proceedings of the Conference on Matrix Methods in Structural Mechanics, Wright Patterson, A.F. Base, Ohio, 1965, pp. 547–576.

[2] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, Society for Industrial and Applied Mathematics, 2002.

[3] R.W. Clough, J.L. Tocher, Finite element stiffness matrices for analysis of plates in bending, in: Conf. on Matrix Methods in Structural Mechanics, Wright–Patterson Air Force Base, Ohio, 1965, pp. 515–545.

[4] P. Dierckx, On calculating normalized Powell–Sabin B-splines, Comput. Aided Geom. Design 15 (1997) 61–78.

[5] T.A. Foley, K. Opitz, Hybrid cubic Bézier triangle patches, in: T. Lyche, L.L. Schumaker (Eds.), Mathematical Methods in Computer Aided Geometric Design II, Academic Press, 1992, pp. 275–286.

[6] J. Grošelj, M. Knez, Generalized C^1 Clough–Tocher splines for CAGD and FEM, Comput. Methods Appl. Mech. Engrg. 395 (2022) 114983.

[7] J. Grošelj, H. Speleers, Construction and analysis of cubic Powell–Sabin B-splines, Comput. Aided Geom. Design 57 (2017) 1–22.

[8] M.J. Lai, L.L. Schumaker, Spline Functions on Triangulations, Cambridge University Press, 2007.

[9] A. Lamnii, M. Lamnii, H. Mraoui, A normalized basis for condensed C^1 Powell–Sabin-12 splines, Comput. Aided Geom. Design 34 (2015) 5–20.

[10] J. Maes, E. Vanraes, P. Dierckx, A. Bultheel, On the stability of normalized Powell–Sabin B-splines, J. Comput. Appl. Math. 170 (2004) 181–196.

[11] S. Mann, Cubic precision Clough–Tocher interpolation, Comput. Aided Geom. Design 16 (1999) 85–88.

[12] L.L. Schumaker, Spline Functions: Computational Methods, Society for Industrial and Applied Mathematics, Philadelphia, 2015.

[13] L.L. Schumaker, Solving elliptic PDE's on domains with curved boundaries with an immersed penalized boundary method, J. Sci. Comput. 80 (2019) 1369–1394.

[14] H. Speleers, A normalized basis for reduced Clough–Tocher splines, Comput. Aided Geom. Design 27 (2010) 700–712.